Technical Report

UCAM-CL-TR-595
ISSN 1476-2986

Number 595

UNIVERSITY OF CAMBRIDGE

Computer Laboratory

Representations of quantum operations, with applications to quantum cryptography

Pablo J. Arrighi

July 2004

15 JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500

http://www.cl.cam.ac.uk/

This technical report is based on a dissertation submitted 23 September 2003 by the author for the degree of Doctor of Philosophy to the University of Cambridge, Emmanuel College.

Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet:

http://www.cl.cam.ac.uk/TechReports/

ISSN 1476-2986

Abstract

Representations of quantum operations

We start by introducing a geometrical representation (real vector space) of quantum states and quantum operations. To do so we exploit an isomorphism from positive matrices to a subcone of the Minkowski future light-cone. Pure states map onto certain light-like vectors, whilst the axis of revolution encodes the overall probability of occurrence for the state. This extension of the Generalized Bloch Sphere enables us to cater for non-trace-preserving quantum operations, and in particular to view the per-outcome effects of generalized measurements. We show that these consist of the product of an orthogonal transform about the axis of the cone of revolution and a positive real symmetric linear transform. In the case of a qubit the representation becomes all the more interesting since it elegantly associates, to each measurement element of a generalized measurement, a Lorentz transformation in Minkowski space. We formalize explicitly this correspondence between 'observation of a quantum system' and 'special relativistic change of inertial frame'. To end this part we review the state-operator correspondence, which was successfully exploited by Choi to derive the operator-sum representation of quantum operations. We go further and show that all of the important theorems concerning quantum operations can in fact be derived as simple corollaries of those concerning quantum states. Using this methodology we derive novel composition laws upon quantum states and quantum operations, Schmidt-type decompositions for bipartite pure states and some powerful formulae relating to the correspondence.

Quantum cryptography

The key principle of quantum cryptography could be summarized as follows. *Honest parties communicate using quantum states. To the eavesdropper these states are random and non-orthogonal. In order to gather information she must measure them, but this may cause irreversible damage. Honest parties seek to detect her mischief by checking whether certain quantum states are left intact.* Thus tradeoff between the eavesdropper's information gain, and the disturbance she necessarily induces, can be viewed as the power engine behind quantum cryptographic protocols. We begin by quantifying this tradeoff in the case of a measure distinguishing two non-orthogonal equiprobable pure states. A formula for this tradeoff was first obtained by Fuchs and Peres, but we provide a shorter, geometrical derivation (within the framework of the above mentioned conal representation). Next we proceed to analyze the Information gain versus disturbance tradeoff in a scenario where Alice and Bob interleave, at random, pairwise superpositions of two message words within their otherwise classical communications. This work constitutes one of the few results currently available regarding d-level systems quantum cryptography, and seems to provide a good general primitive for building such protocols. The proof crucially relies on the state-operator correspondence formulae derived in the first part, together with some methods by Banaszek. Finally we make use of this analysis to prove the security of a 'blind quantum computation' protocol, whereby Alice gets Bob to perform some quantum algorithm for her, but prevents him from learning her input to this quantum algorithm.

Acknowledgements

I am very grateful to my superviser, Dr. Anuj Dawar, who has never failed his role. For some good advice and within the Computer Laboratory still, I would like to thank Prof. Glynn Winskel and Dr. Alan Mycroft. I am also very appreciative of my collaborators: Dr. Louis Salvail, Prof. Frank Kelly on one occasion, and above all Christophe Patricot − with whom thoughts met friendship in an exceptional manner.

A small number of great friends made my stay in Cambridge into extremely happy days: special thanks to Sébastien Loisel, Christophe Patricot, Sandra Lucas, Emilie Stephenson and Tom Kelly. I am also indebted to my uncles: Dr. Arturo Lezama made me discover quantum theory when I was little, Dr. Jean-Michel Arrighi provided encouragements and financial support throughout my studies, together with Dr. Paul Arrighi and my two grandfathers. Parents are never thanked enough. I thank my father, Pierre Arrighi, for the diagrams in this thesis and so many other things. I thank my mother, Claudia Lezama, for the title of chapter 2 and so many other things. I am also grateful to Jacques Coral and Ilse Arrighi.

I thank my sisters, Gaïa Lezama, Barbara Arrighi, Malena Arrighi who still love me in spite of the well-known series of events who took me to England, and away from their childhoods.

Thank you Elvire, for this last year and those ahead of us.

Collaborations and Publications

- Chapter 2 has been published under the title

  Quantum Computation explained to my Mother

  in the Bulletin of the EATCS, **80**, 134-142, (2003), and in Current Trends in Theoretical Computer Science, Ed. G. Paun, G. Rozenberg, World Scientific, (2004). It is also available at arXiv:quant-ph/0212062.

- The results in Chapter 3 and Chapter 6 have been accepted for publication as one single article under the title

  The Conal representation of Quantum States and Non Trace-Preserving Quantum Operations

  in Phys. Rev. A, **68**, 042310, (2003), and in Virtual Journal on Quantum Information, **3**, 10. They are also available at arXiv:quant-ph/0212062.

  This work is the outcome of a collaboration with Christophe Patricot from the DAMTP, Faculty of Mathematics, University of Cambridge, UK.

- The results in Chapter 4 have been published under the title

  A Note on the Correspondence between Qubit Quantum Operations and Special Relativity

  in J. Phys. A, **36**, L287-L296, (2003), and are available at arXiv:quant-ph/0212135.

  This work is the outcome of a collaboration with Christophe Patricot.

- The results in Chapter 5 have been published under the title

  On Quantum Operations as Quantum States

  in Annals of Phys. **311**, 26-52, (2004), and are available at arXiv:quant-ph/0307024.

  This work is the outcome of a collaboration with Christophe Patricot.

- The results in Chapter 7 have been published under the title

  Quantum Decoys

  in Int. J. Quantum Information, and are available at arXiv:quant-ph/0308050.

- The results in Chapter 8 are available under the title

  Blind Quantum Computation

  at arXiv:quant-ph/0309152.

  This work is the outcome of a collaboration with Dr. Louis Salvail from BRICS, Department of Computer Science, University of Aarhus, Denmark. Moreover

  Lemma 8.2 is due to Prof. Frank Kelly from the DPMMS, Faculty of Mathematics, University of Cambridge, UK.

Contents

To my son,

Introduction

Chapter 1

Overture

Nous dirons les choses au fur et à mesure que nous les verrons et que nous saurons.
Et ce qui doit rester obscur le sera malgré nous.

—*Jules Supervielle*

1.1 Scope and aims

This thesis presents one long line of thoughts in the field of quantum information theory. The argument begins within the mathematical foundations of quantum theory, and ends with a precise cryptographic protocol. But because the articulation must be placed somewhere let us say that the thesis pursues two aims.

First we aim to provide a better understanding of quantum states and quantum operations. To make our intentions more specific: we attempt to improve the common intuition and prove novel properties about the density matrix formalism of quantum theory. To make our methods more specific: we thoroughly exploit several formal correspondences between the mathematical structures of quantum theory and some other well-known mathematical structures. In this manner one point of view comes to complement the other, one property translates into another. Thus our work joins and adds to those of many who have studied representations of quantum states of quantum operations. In particular it builds upon the Generalized Bloch Sphere representation and upon the state-operator correspondence to achieve its ends. In the end we achieve a representation of non trace-preserving quantum operations having numerous desirable properties and a connection with special relativity, as well as a number of other significant mathematical results. It is this armoury which enables us to tackle the quantum information theoretical problems next described.

Second we aim to contribute to quantum cryptography at the fundamental level. To make our intentions more specific: we attempt to tame the tradeoff between information gain about a quantum state and the disturbance this necessarily causes to the quantum state. Moreover we attempt to use this analysis so as to open the door to a novel type of quantum cryptographic

applications. A commonplace about quantum theory is to say that *measurements modify the*

object measured. Quantum cryptographic protocols rely upon this fact to detect the presence of a malevolent eavesdropper. Thus to quantify these tradeoffs between information gain and disturbance is a crucial problem, yet it is also an extremely difficult one. To make our method more specific: we exploit the representation of non trace-preserving quantum operations we developed earlier, together with some novel formulae arising from the state-operator correspondence. In this manner we obtain, respectively, a geometrical rederivation of the well-known formula for the simplest possible such tradeoff scenario, and a previously unknown formula for an important more elaborate scenario. Armed with the latter result we can describe a protocol related to, but different from, secure two-party computation: Alice gets Bob to perform some quantum algorithm for her, but prevents Bob from learning her input to this quantum algorithm.

1.2 Generalities

Quantum cryptography

For eighty years scientists have suffered the oddity of Quantum Theory, cursing its counter-intuitiveness as they were working through the implications. Quantum measurements in particular could be resented as a limitation. Why can we not just learn all there is to know about a quantum state? Why can we not observe a quantum state without irreversibly modifying this quantum state?

The attitude shift began in 1969, when Wiesner suggested one could actually exploit the strange properties of quantum measurements as a way of achieving unforgeable banknotes [64]. The idea might have seemed too original at the time, and one must await the year 1984 for Bennett and Serge Brassard to (re)invent quantum cryptography [7]. Their protocol, well-established nowadays and commonly referred to as BB84, causes much immediate interest. The two authors claim to achieve secure key distribution with unconditional security, something that was judged impossible back then. In other words they describe a manner in which two remote parties could generate a long secret string, and this in spite of an eavesdropper standing in between. Moreover the security is information theoretical and does not rely upon any other assumption (e.g. presence of noise, computer power limitations) than that of the laws of quantum physics being true.

One fundamental fact makes the BB84 protocol (and Wiesner's unforgeable banknotes) possible. The malevolent eavesdropper, as she seeks to gather information about the unknown quantum states included in the transmissions will, in effect, cause damage to these states, in a way that she cannot repair. Honest parties use this to detect her tampering and modify their behaviour as a consequence. This mechanism is the cornerstone upon which quantum cryptography is based.

To provide Bennett and Brassard's proposal with a full-blown proof of security turned out

to be an extremely difficult mathematical problem, however. The many attempts made to quantify directly the tradeoff between how much information the eavesdropper can eavesdrop, and what then are her chances of getting caught, have all turned out notoriously complicated [22][47]. In order to circumvent these hardships Ekert proposed another secure quantum key distribution protocol [18], where the notion of entanglement between quantum states plays a central role. The ideas and notions thereby developed have had an important impact on the field [33][3], to the point even where they ended up inspiring some of the key steps in Shor and Preskill's latest and nicest-known BB84 proof of security [56].

This long detour is somewhat symbolic of the current state of affairs in quantum cryptography. In spite of being so central in quantum information theory, the tradeoff between how much information one may gain about a quantum system *versus* how much disturbance the observation must necessarily cause to the system, tends to be avoided. Those quantum cryptographers who propose novel protocols find their way round the problem; most of their proofs are a witty blend of quantum error correction, Bell inequalities, together with the particular symmetries of the protocol in question. Even the closely related question of optimal quantum cloning (copy) of quantum states has received more attention [24][66]. But the two issues cannot be equated trivially; the former seems to lie at the limit, as N tends to infinity, of a non-universal asymmetric 1 to N quantum cloning machine.

Information gain versus disturbance tradeoff

Those quantum cryptographers who, against the trend, tackle information gain versus disturbance tradeoff scenarios must therefore be driven by the search of novel fundamental properties of quantum information. At least this was the case with Fuchs as he tackled, together with Peres [21], the seemingly simple case of an eavesdropper trying to distinguish two non-orthogonal equiprobable pure quantum states ensemble. For this purpose he does acknowledge some amount of disappointment:

> There are at least two (disappointing) things to notice about [our] equations (...). The first is the energy that had to be expended in order to work out a tradeoff relation for one of - surely - the simplest possible cases. Given the hoped-for foundational importance of the principle, this is rather curious. (...) The second thing to note about [our] equations is the relative complexity of the curve. (...)

This article was the first to tackle a tradeoff scenario for its own sake. For discrete distributions this was also the last one. Of lesser interest for cryptography, but very important in terms of its methods was the work by Banaszek [5], who quantified the tradeoff for the continuous uniform n-dimensional ensemble. Barnum [6] made several accurate remarks upon the same ensemble, suggesting the tradeoff remains unchanged for a uniform distribution over mutually unbiased states.

The approach to quantum information theory which we develop in this thesis owes much to the problem of quantifying the information gain versus disturbance tradeoff. Our mathematical findings are benchmarked against, and applied to resolving these tradeoffs - upon which we later base the security of a novel protocol.

Quantum computation

In parallel with the blossoming of quantum cryptography one should mention the discovery and the impressive developments of quantum computation. That story originated with Feynman [20] as he suggested that no machine would simulate quantum physical phenomena better than... a quantum system. Soon afterwards Deutsch proposed a Quantum Turing Machine model of computation [16], and found, together with Jozsa, one particular problem which could be solved exponentially faster with such a quantum computer [17]. Two major quantum algorithms have been discovered since: Shor's integer factoring algorithm [55] (which incidentally renders quantum cryptography all the more necessary as the algorithm threatens to break all current public key encryption systems) and Grover's unordered search algorithm [26] (which brings only a square-root speedup but has an enormous range of potential applications).

Secure quantum computation

The history of quantum information theory has taught at least two lessons to its people. One is to pick up some classical problem in computer science, and prove it can be done much better using quantum theory. The other is to do cross-disciplinary work. The field of secure quantum computation was an opportunity to do both quantum cryptography and quantum computation, and to solve some rather involved classical scenarios.

In the traditional secure two-party computation scenario Alice has secret input x, Bob has secret input y, and both of them wish to compute $f(x, y)$. The function f is of course well-known to the two parties; the usual example is that of two millionaires who wish to compare their riches without revealing how much they have. There is no emphasis on computational load in secure two-party computation, in the sense that if Alice knew Bob's input y she might as well compute $f(x, y)$ on her own. Entering a secure two-party computation together with Bob will not help alleviate Alice's cost of computing f. In fact quite the opposite is true of current protocols. The notion of 'blind computation', on the contrary, is asymmetric. Alice is the only one to have a secret input x, Bob is the only one able to compute f. Alice wants Bob to compute $f(x)$ without him learning x. Thus an obvious motivation for Alice to enter a blind computation together with Bob is to unload the computational task of computing f, but without having to trust Bob.

Whilst on the one hand issues of multi-party secure computation related scenarios enjoy a vast amount of publication in the non-quantum computer science literature (as for instance [1][9][65] to cite only a few), this is not, on the other hand, the case of blind computation related scenarios which suffer as their most influential result a no-go theorem by Abadi, Feigenbaum and Kilian [2] together with a general disbelief that this can be achieved. There was a short regain of interest when Sander and Tschudin [52] gave a blind computation protocol for evaluating polynomials relying on homomorphic encryption schemes, themselves based on computational assumptions.

The uneven distribution of articles in quantum information is merely a reflection of the situation in the non-quantum computer science literature. Much effort has gone, on the one hand, into devising secure multi-party quantum computation protocols. In order to do so scientists have long sought to provide an unconditional oblivious transfer, the primitive upon which multi-party computation is founded in the classical realm [35]. But oblivious transfer can be shown to imply bit commitment (the latter cryptographic primitive is readily constructed from the former), whose security is in turn impossible to obtain from quantum theory alone. This no-go theorem was derived independently by Lo and Chau [44] for perfectly secure bit commitment protocols, and by Mayers [45][46] for the more general unconditionally secure bit commitment protocols - i.e. those where the probability of a breach is vanishingly small as the security parameter (the number of rounds) is increased. It follows that oblivious transfer is also impossible to obtain unconditionally, a fact that is all the more discouraging if

we consider that the primitive is itself a subcase of secure computation. Some other negative consequences of the theorem are discussed in [43]. Yet there has been number of results since, for some specific cryptographic tasks [36], with more than two players [14], using computational assumptions which are thought to resist quantum computational power [15], or even relying upon relativistic effects [34]. Blind quantum computation protocols, on the other hand, have not been studied so far. The closest specimen is provided in [12], but there are fundamental differences, and some problems with this protocol.

The present thesis approaches the issue of blind computation, which has not been studied in a quantum information theoretical setting so far.

Representations of quantum operations and quantum states

Quantum theory is firmly founded upon a set of simple axioms, as defined for open systems by Von Neumann [60]. However these axioms produce rather intricate optimization problems as soon as they are applied to even the simplest quantum information theoretical scenario. Still several researchers wish to believe that within quantum theory lies some intimate relation between 'information' and 'matter'. Landauer's little sentence [40] 'Information is physical' has become somewhat of a motto in the field of quantum information theory, and prominent researchers such as Zeilinger and Brukner seem to have vowed to recast quantum theory altogether along these lines [67]:

> In contrast to the theories of relativity, quantum mechanics is not yet based on a generally accepted conceptual foundation. It is proposed here that the missing principle may be identified through the observation that all knowledge in physics has to be expressed in propositions and that therefore the most elementary system represents the truth value of the proposition, i.e., it carries just one bit of information. (...)

Without always pursuing such grand aims the field certainly makes an abundant use of alternative representations of quantum states and quantum operations - so as to simplify, or at least to bring useful insight to the calculations. There are two well-established families of such representations, whose origins clearly predate quantum information theory.
The first one is the Bloch sphere picture of a qubit [8], related to the standard notion of three-dimensional spin (or polarization), and generalized to n-dimensional quantum systems it seems by Hioe and Eberly [29].
The second one is the operator sum representation of quantum operations by Kraus [37]. This result was rediscovered independently by Choi [13] but relying upon an intriguing correspondence between quantum states and quantum operations, through a map which can be traced back to the work of Jamiolkowski [32].

Since then a number of researchers have sought to modernize these representations in the light of quantum information theory. Noticeable examples include Zanardi [66], Sudarshan [58], Ruskai [51], Verstraete [61]. A considerable amount of efforts were devoted, in particular, to taming the geometry of trace-preserving quantum operations.
Trace-preserving quantum operations are indeed important, for they precisely correspond to the set of all physically allowed evolutions having probability of occurrence one on every input state. On the contrary this type of operations excludes all those evolutions which may be successfully undertaken with some finite probability, or whose probability depends upon the input state. The effects of a quantum measurement *knowing* that some outcome occurs, for instance, do not belong to this class, even though they will be the first thing one may wish to visualize in a quantum information theoretical scenario.

The present thesis approaches the theme of representations of quantum states and quantum operations with a view towards applications in quantum information theory. It focuses in particular upon representations of non trace-preserving quantum operations, and upon the intriguing role of the state-operator correspondence.

1.3 Summary of contributions

As we exploit an isomorphism from $Herm_d(\mathbb{C})$ *into* $\mathbb{R}^{d^2}$ *we contribute in the following manner:*

- By constructing a convenient geometrical (real vector space) representation of quantum states. Because these do not need to be normalized we find that they map into a subcone of a Minkowski cone in $\mathbb{E}^{1,d^2-1}$, whose vertical cross-sections are nothing but generalized Bloch spheres. We show that the conal representation has numerous elegant properties: pure states map into light-like vectors, unitary operations correspond to orthogonal transforms about the axis, and positive operations are represented by a subset of the real symmetric positive matrices. The latter can also be drawn in the cone - this allows us to visualize non trace-preserving quantum operations and their effects.

- By providing, for the conal representation of a qubit, explicit formulae for the coordinates of states after non trace-preserving quantum operations, or for the scalar product of two post-measurement states:

$$\frac{1}{4}[2(\underline{E_m}.\underline{\rho^0})(\underline{E_m}.\underline{\rho^1}) - (\eta_{\mu\mu'}\underline{E_m}_{\mu}\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho^0}_{\nu}\underline{\rho^1}_{\nu'})].$$

  These constitute a sufficient armoury to deal, using only four-vectors, with the most general evolutions to happen on a qubit.

- By showing, in the case of a qubit still, that the conal representation elegantly associates, to each measurement element of a generalized measurement, a special relativistic Lorentz transformation in Minkowski space. More precisely each measurement element acts proportionally to an element of the restricted Lorentz group together with future-directed null boosts, giving rise to the following formulae:

$$\begin{aligned}\eta_{\mu\nu}\underline{\rho_m}_{\mu}\underline{\rho_m}_{\nu} &= \eta_{\mu\nu}\underline{V_m}_{\mu}\underline{V_m}_{\nu}\eta_{\mu'\nu'}\underline{\rho}_{\mu'}\underline{\rho}_{\nu'}\\ \underline{\rho_m}_{0} &= \eta_{\mu\nu}\underline{V_m}_{\mu}\underline{\rho}_{\nu}\\ \eta_{\mu\nu}\underline{\rho}_{\mu}\underline{\rho}_{\nu} &= 2([Tr(\rho)]^2 - Tr(\rho^2)).\end{aligned}$$

  The rescaling introduced turns out to bring null boosts to finite linear maps in a natural and unifying manner. We formalize explicitly this correspondence between 'observation of a quantum system' and 'special relativistic change of inertial frame', thereby providing an original outlook upon, and a generalization of, the well-known group isomorphism of 2×2 unimodular complex matrices onto the restricted Lorentz group.

- By recovering geometrically the formula given by Fuchs and Peres for the information gain versus disturbance tradeoff, as it arises when attempting to distinguish two non-orthogonal equiprobable quantum states.

As we exploit an isomorphism from elements of $M_{mn}(\mathbb{C})$ to linear maps from $M_n(\mathbb{C})$ to $M_m(\mathbb{C})$ we contribute in the following manner:

- By providing two triangular decompositions for pure states of a bipartite system, i.e. local changes of basis so that vectors in $\mathbb{C}^m \otimes \mathbb{C}^n$ may be written with triangular coefficients only.

- By providing two original algebraic tests on Completely Positive-preserving maps: one testing extremality in the set of Trace-preserving operations, the other regarding the factorizability or single operator decomposition, i.e. $\widehat{\$}(\rho)$ is of the form $\hat{V}\rho\hat{V}^\dagger$ for all ρ, if and only if

$$\Big(\mathrm{Tr}(\widehat{\$}(\mathbb{I}_n))\Big)^2 - \sum_{jl}\mathrm{Tr}\big(\widehat{\$}(E_{jl})^\dagger\widehat{\$}(E_{jl})\big) = 0 \tag{1.1}$$

  These are particularly interesting in the sense that they do not depend on the operator sum decompositions of these maps.

- By endowing $\mathrm{Herm}^{+}_{n^2}(\mathbb{C})$ with a semi-group structure stemming from the composition law on quantum operations. The composition law defines a group when restricted to the set of totally entangled (pure) states, and yields a group isomorphism between maximally entangled (pure) states and $SU(n)$. In addition we show that the set of quantum operations is stable under component-wise product.

- By providing a number of useful formulae which arise from the state-operator correspondence such as (with \$ the state corresponding to the quantum operation $\widehat{\$}$):

$$\widehat{\$}\big((\rho^\dagger\rho)^t\big) = \mathrm{Tr}_2\big((\mathbb{I}_m \otimes \rho)\$(\mathbb{I}_m \otimes \rho^\dagger)\big)$$
$$\mathrm{Tr}\big(\sigma\widehat{\$}(\rho)\big) = \mathrm{Tr}\big((\sigma \otimes \rho^t)\$\big)$$

  The first formula suggests potential physical interpretations of the state-operator correspondence. The latter formula will simplify those many mathematical problems in quantum cryptography which require a careful optimization of the fidelities induced by a linear operator $\widehat{\$}$.

- By discovering, using the latter formula, the information gain versus disturbance tradeoff in an elaborate scenario. Suppose the eavesdropper performs an individual attack such that, whenever a canonical basis state $\{|j\rangle\}_{j=1\ldots n}$ is sent, she is able to identify which with probability G. Then, whenever a pairwise superposition $\{(1/n^2, \rho_{jk}\}_{j,k=1\ldots n}$, with $\rho_{jk} = \frac{(|j\rangle+i|k\rangle)(\langle j|-i\langle k|)}{2}$, is sent, her disturbance D is bounded below under the following tight inequality:

$$D \geq 1 - F(G)$$
$$F(G) = \frac{1}{2} + \frac{1}{2n}\Big(\sqrt{G} + \sqrt{(n-1)(1-G)}\Big)^2.$$

  where, for optimal attacks, G varies from $\frac{1}{n}$ to 1 as D varies from 0 to $\frac{1}{2} - \frac{1}{2n}$.

- By highlighting the central, transversal role of the state-operator isomorphism in various issues of quantum information theory.

As we build upon these last mathematical results we contribute

- By introducing a blind quantum computation protocol for the class of functions which admit an efficient procedure to generate random input-output pairs. In this scenario

Alice wants Bob to compute some well-known function f upon her input x, but wants to prevent Bob from learning anything about x. The protocol relies upon the newly found information gain versus disturbance tradeoffs to achieve unconditional security against the most general attack: whenever Bob gathers $\log(n)+\log(G)$ bits of Shannon mutual information about the input, he must get caught with probability at least $1-F(G)^N$, where n denotes the size of the input and N is a security parameter, whilst G and $F(G)$ remain as in the previous paragraph.

1.4 Outline

Chapter 2 is addressed to the widest audience possible, and should not be thought to set the tone of the thesis. There we introduce complex matrices, the postulates of quantum theory and discuss the simplest known quantum algorithm.

Part I looks at representations of quantum operations, with a view towards quantum information theory.

Chapter 3 introduces a useful extension the Generalized Bloch Sphere which caters for non-trace-preserving quantum operations. The properties of this conal representation of quantum states and quantum operations are systematically explored.

Chapter 4 highlights the mathematical correspondence between generalized measurement elements and changes of observers in special relativistic space-time, as it arises from the conal representation. Possible interpretations are discussed.

Chapter 5 reviews and formalizes another correspondence, from quantum operations on the one hand, to quantum states on the other hand. The systematic approach taken yields an original presentation of the fundamental theorems about quantum operations, together with some novel results.

Part II exploits the general methods developed in the first part so as to tackle a small set of important problems in quantum information theory.

Chapter 6 recovers geometrically a well-known tradeoff between the information gain when attempting to distinguish two non-orthogonal equiprobable pure states, and the disturbance this causes.

Chapter 7 investigates a similar tradeoff in a scenario where the eavesdropper seeks to distinguish canonical basis states of a n-dimensional quantum system, but sometimes a pairwise superposition of two canonical basis states is interleaved so as to detect his malevolent measurements.

Chapter 8 provides a novel security protocol, whereby Alice gets Bob to perform some quantum algorithm for her, but prevents him from learning her input to this quantum algorithm.

Finally Chapter 9 presents the conclusions of this thesis and how these may serve as departure points for further promising research.

1.5 How to read the thesis

Non-specialist readers have their chapter (starting page 27) but should probably not risk themselves further. *Specialist readers* are advised to browse through Appendix A, which presents the notation used throughout the thesis.

Further down the thesis readers will face a thematic ordering. *Someone who focuses on the foundations of quantum theory* will be content reading only Part I. Real vector space representations of quantum states and quantum operations bring an invaluable amount of intuition to quantum physics, and the connection with special relativity is somewhat intriguing. Moreover the state-operator correspondence is demonstrated to be a powerful mathematical tool, so powerful in fact one may wonder about its physical meaning.
Someone with an exclusive interest in cryptography will find satisfaction reading only Part II. The information gain versus disturbance tradeoff lies at the heart of cryptography. To understand it geometrically should prove useful, and so should the general n-dimensional quantum cryptographic primitive which we have built upon this tradeoff. The application to blind quantum computation is in fact a somewhat striking use of this primitive.

There are numerous mathematical dependencies, however, which may prevent the cryptographer from having an all too peaceful reading. The geometrical rederivation of Fuchs and Peres' information gain versus disturbance tradeoff is framed in the conal representation introduced on Chapters 3 and 4, whilst the derivation of the information gain versus disturbance tradeoff for the n-dimensional quantum cryptographic primitive relies upon the state-operator formulae given in Chapter 5. Every chapter of every part is in reality just one element of a whole long story.

Yet we have tried, out of consideration for the *opportunistic reader*, to make it possible for one to pick up a precise topic of interest. For example if looking for special relativity related material the reader should go straight to Chapter 4; if looking for secure two-party computation related material the reader should go straight to Chapter 8, etc... This was done at the cost of a few repetitions, which the *thorough mind* must forgive.

Chapter 2

Quantum computation explained to my mother

Technology, sufficiently advanced, is indistinguishable from magic.
—Arthur C. Clarke

There are many falsely intuitive introductions to quantum theory and quantum computation in a handwave. There are also numerous documents which teach those subjects in a mathematically sound manner. To our knowledge the presentation in this chapter is of the shortest in the latter category. The aim is to deliver a short yet rigorous and self-contained introduction to Quantum Computation, whilst assuming the reader has no prior knowledge of anything but the fundamental operations on real numbers. Successively we introduce complex matrices; the postulates of quantum theory and the simplest quantum algorithm.

In the mind of every scientist exists the intimate belief that one's subject is not as difficult as it may seem and, certainly, everyone should be able to understand it given a decent opportunity and a finite amount of effort. The present chapter is addressed to the widest audience possible, as an invitation to enter the subject of quantum information theory. In particular, and since this thesis is written in a computer science department, we find it useful to dissipate the popular fear of quantum mechanics − over just a few efficient pages.
This mathematically minimalist presentation should not be thought to set the tone of the thesis. Yet it constitutes an interesting exercise *per se*: experienced physicists have expressed a sense of curiosity when faced with the smallest set of mathematical definitions upon which quantum theory is anchored; as well as an interest to use this path of presentation for their courses.

2.1 Some mathematics

We will begin this introduction with about four pages of mathematics, mainly definitions. These notions constitute the vocabulary, the very language of quantum theory, and every single one of them will find its use in the second part, when we introduce the postulates of quantum theory.

2.1.1 Complex numbers

A *real* number is a number just like you are used to. E.g. -4.3, 0, $\sqrt{2}$ are all real numbers. A *complex* number, on the other hand, is just a pair of real numbers. I.e. suppose z is a complex number, then z must be of the form (a, b) where a and b are real numbers.
Now we must teach the reader how to add or multiply complex numbers. Suppose we have two complex numbers $z_1 = (a_1, b_1)$ and $z_2 = (a_2, b_2)$. Addition first: $z_1 + z_2$ is defined to be the pair of real numbers $(a_1 + a_2, b_1 + b_2)$. And now multiplication (when we put two numbers next to one another, with no sign in between that means they are multiplied): $z_1 z_2$ is defined to be the pair of real numbers given by $(a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1)$.
Sometimes we want to change the sign of the second (real) component of the complex number z. This operation is called *conjugation*, and is denoted by an upper index '*', i.e. z^* is defined to be the pair of real numbers $(a, -b)$.
Another useful operation we do on a complex number is to take its *norm*. The norm of $z = (a, b)$ is defined to be the real number $\sqrt{a^2 + b^2}$. This operation is denoted by two vertical bars surrounding the complex number, in other words $|z|$ is simply a notation for $\sqrt{a^2 + b^2}$.

2.1.2 Matrices

A *matrix* of *things* is a table containing those things, for instance: $\begin{pmatrix} \heartsuit & \spadesuit \\ \diamondsuit & \clubsuit \end{pmatrix}$ is a matrix of card suits. We shall call this matrix M for use in later examples.

A matrix does not have to be square. We say that a matrix is $m \times n$ if it has m horizontal lines and n vertical lines. For instance a *column* is a $1 \times n$ matrix e.g: $\begin{pmatrix} \heartsuit \\ \diamondsuit \end{pmatrix}$. Similarly a *row* is a $m \times 1$ matrix, e.g. $\begin{pmatrix} \heartsuit & \spadesuit \end{pmatrix}$ is a row.

The *ij-component* of a matrix designates the 'thing' which is sitting at vertical position i and horizontal position j in the table, starting from the upper left corner. For instance the 2 1-component of M is $\diamondsuit$. If A is a matrix then the ij-component of A is denoted A_{ij}, e.g. here you have that $M_{11} = \heartsuit,\ M_{21} = \diamondsuit$ etc.

Given a matrix we often need to make its vertical lines into its horizontal lines and vice-versa. This operation is called *transposition* and is written '^{t}'. In other words if A is the $m \times n$ matrix with ij-component A_{ij}, then A^t is defined to be the $n \times m$ matrix which has ij-component A_{ji}. Thus we have $A^t_{ij} = A_{ji}$. Here are two examples:

$$M^t = \begin{pmatrix} \heartsuit & \diamondsuit \\ \spadesuit & \clubsuit \end{pmatrix} \quad ; \quad \begin{pmatrix} \heartsuit \\ \diamondsuit \end{pmatrix}^t = \begin{pmatrix} \heartsuit & \diamondsuit \end{pmatrix}$$

2.1.3 Matrices of numbers

Let us now consider matrices of numbers. The good thing about numbers (real or complex, it does not matter at this point) is that you know how to add and multiply them. This particularity will now enable us to define addition and multiplication *of matrices of these numbers*.

In order to add two matrices A and B they must both be $m \times n$ matrices (they have the same size). Suppose A has ij-components. Then $A + B$ is defined to be the $m \times n$ matrix with ij-components $A_{ij} + B_{ij}$.

If we now want to multiply the matrix A by the matrix B it has to be the case that the number of vertical lines of A equals that of the number of horizontal lines of B. Now suppose A is an $m \times n$ matrix with ij-components A_{ij}, whilst B is $n \times r$ and has pq-components B_{pq}. Then AB is defined to be the $m \times r$ matrix with iq-components $A_{i1}B_{1q} + A_{i2}B_{2q} + .. + A_{in}B_{nq}$. To make things clear let us work this out explicitly for general 2×2 matrices of numbers:

$$\text{Let} \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

$$\text{Then} \quad A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix} \quad \text{and} \quad AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

2.1.4 Matrices of complex numbers

Matrix addition and multiplication work on numbers, whether they are real or complex. But from now we look at matrices of complex numbers only, upon which we define one last operation called *dagger*. To do a dagger operation upon a matrix is to transpose the matrix and then to conjugate all the complex numbers it contains. This operation is denoted '†'. We thus have $A^\dagger_{ij} = A^*_{ji}$, in other words if A is the $m \times n$ matrix with ij-component A_{ij}, then $A^\dagger$ is defined to be the $n \times m$ matrix which has ij-component A^*_{ji}.
Quite a remarkable $n \times n$ matrix of complex numbers is the one we call 'the *identity* matrix'. It is defined such that its ij-component is the complex number $(0,0)$ when $i \neq j$, and the complex number $(1,0)$ when $i = j$. The $n \times n$ identity matrix is denoted $\mathbb{I}_n$, as in:

$$\mathbb{I}_1 = \begin{pmatrix} (1,0) \end{pmatrix} \quad \text{and} \quad \mathbb{I}_2 = \begin{pmatrix} (1,0) & (0,0) \\ (0,0) & (1,0) \end{pmatrix}$$

Having defined the identity matrices we are now able to explain what it means to be a *unitary* matrix of complex numbers. Consider M an $n \times n$ matrix of complex numbers. M is said to be a unitary matrix if (and only if) it is true that $M^\dagger M = \mathbb{I}_n$. Similarly consider V a $1 \times n$ matrix of complex numbers (a column). V is said to be a unit column if (and only if) it is true that $V^\dagger V = \mathbb{I}_1$.

2.1.5 Some properties

The reader may choose to skip the following three properties, but they will be needed in order to fully understand the comments which follow postulates 2.2 and 2.3.

Property 2.1 *Let A be an $n \times m$ matrix of complex numbers and $\mathbb{I}_m$ the $m \times m$ identity matrix. We then have that $A\mathbb{I}_m = A$. In other words multiplying a matrix by the identity matrix leaves the matrix unchanged.*

Proof: First note that a complex number (a,b) multiplied by the complex number $(1,0)$ is, by definition of complex number multiplication, given by $(1a - 0b, 0a + 1b)$, which is just (a,b) again. Likewise note that a complex number (a,b) multiplied by the complex number $(0,0)$ is given by $(0a - 0b, 0a + 0b)$, which is just $(0,0)$. Now by definition of matrix multiplication the iq-component of $A\mathbb{I}_m$ is given by: (we denote $\mathbb{I}_m$ by just $\mathbb{I}$ until the end of the proof)

$$\begin{aligned}(A\mathbb{I})_{iq} &= A_{i1}\mathbb{I}_{1q} + A_{i2}I_{2q} + .. + A_{in}\mathbb{I}_{nq} \\ &= A_{i1}(0,0) + A_{i2}(0,0) + .. + A_{iq}(1,0) + .. + A_{in}(0,0)\end{aligned}$$

The second line was obtained by replacing the $\mathbb{I}_{pq}$ with their value, which we know from the definition of the identity matrix. Now using the two remarks at the beginning of the proof

we can further simplify this equation:

$$\begin{aligned}(A\mathbb{I})_{iq} &= (0,0) + (0,0) + .. + A_{iq} + .. + (0,0) \\ &= A_{iq} \quad \text{by complex number addition.}\end{aligned}$$

Thus the components of $A\mathbb{I}$ are precisely those of A. □

Property 2.2 *Let A be an $m \times n$ matrix of complex numbers and B be an $n \times r$ matrix of complex numbers. Then the following equality is true:*

$$(AB)^\dagger = B^\dagger A^\dagger$$

Proof: First note that

$$((a_1, b_1) + (a_2, b_2))^* = (a_1, b_1)^* + (a_2, b_2)^* \tag{2.1}$$

This is obvious since

$$\begin{aligned}((a_1, b_1) + (a_2, b_2))^* &= (a_1 + a_2, b_1 + b_2)^* \\ &= (a_1 + a_2, -b_1 - b_2) \quad \text{and} \\ (a_1, b_1)^* + (a_2, b_2)^* &= (a_1, -b_1) + (a_2, -b_2) \\ &= (a_1 + a_2, -b_1 - b_2) \quad \text{as well.}\end{aligned}$$

Likewise note that

$$((a_1, b_1)(a_2, b_2))^* = (a_1, b_1)^*(a_2, b_2)^* \tag{2.2}$$

and also

$$(a_1, b_1)(a_2, b_2) = (a_2, b_2)(a_1, b_1) \tag{2.3}$$

again this is easily verified by computing the left-hand-side and the right-hand-side of those equalities.
Now by definition of matrix multiplication we have that

$$(AB)_{iq} = A_{i1}B_{1q} + A_{i2}B_{2q} + .. + A_{in}B_{nq}$$

Thus the components of $(AB)^\dagger$ are given by

$$\begin{aligned}(AB)^\dagger_{iq} &= (AB)^*_{qi} \\ &= A^*_{q1}B^*_{1i} + A^*_{q2}B^*_{2i} + .. + A^*_{qn}B^*_{ni} \\ &= B^*_{1i}A^*_{q1} + B^*_{2i}A^*_{q2} + .. + B^*_{ni}A^*_{qn}\end{aligned}$$

where we used equations (2.1) and (2.2) to obtain the second line, and equation (2.3) to obtain the third line. Now consider the components of $B^\dagger A^\dagger$. By definition of matrix multiplication we have that

$$\begin{aligned}(B^\dagger A^\dagger)_{iq} &= B^\dagger_{i1}A^\dagger_{1q} + B^\dagger_{i2}A^\dagger_{2q} + .. + B^\dagger_{in}A^\dagger_{nq} \\ &= B^*_{1i}A^*_{q1} + B^*_{2i}A^*_{q2} + .. + B^*_{ni}A^*_{qn}\end{aligned}$$

where the last line was obtained using the fact that $A^\dagger_{ij} = A^*_{ji}$. Thus the components of $(AB)^\dagger$ are precisely those of $B^\dagger A^\dagger$. □

Property 2.3 *Let V be a $n \times 1$ unit matrix of complex numbers (a column). Then it is the case that:*

$$|V_{11}|^2 + |V_{21}|^2 + .. + |V_{n1}|^2 = 1$$

Proof: First let $z = (a, b)$ be a complex number, and note that

$$\begin{aligned}z^*z &= (a^2 + b^2, 0) \\ &= (|z|^2, 0)\end{aligned}$$

Now consider the 11-component of $V^\dagger V$. We have:

$$\begin{aligned}(V^\dagger V)_{11} &= V^\dagger_{11}V_{11} + V^\dagger_{12}V_{21} + .. + V^\dagger_{1n}V_{n1} \\ &= V^*_{11}V_{11} + V^*_{21}V_{21} + .. + V^*_{n1}V_{n1}\end{aligned}$$

where we used successively: the definition of matrix multiplication, and $A^\dagger_{ij} = A^*_{ji}$. The last line can be further simplified using our first remark, namely:

$$V^*_{i1}V_{i1} = (|V_{i1}|^2, 0)$$

Thus

$$\begin{aligned}(V^\dagger V)_{11} &= (|V_{11}|^2, 0) + (|V_{21}|^2, 0) + .. + (|V_{n1}|^2, 0) \\ &= (|V_{11}|^2 + |V_{21}|^2 + .. + |V_{n1}|^2, 0)\end{aligned}$$

Because V is unit the last line must be equal to $(1, 0)$, and so we have proved the property. □

2.2 Quantum theory

Quantum theory is one of the pillars of modern physics. The theory is 100 years old and thoroughly checked by experiments; it enables physicists to understand and predict the behavior of any closed (perfectly isolated from the rest of the world) physical system. Usually these are small systems such as atoms, electrons, photons etc. (only because they are generally less subject to outside interactions).

2.2.1 States

Postulate 2.1 *The* state *of a* closed physical system *is wholly described by a unit* $n \times 1$ *matrix of complex numbers.*

Comments. In other words a state is given by a column of n complex numbers

$$V = \begin{pmatrix} V_{11} \\ \vdots \\ V_{n1} \end{pmatrix} \quad \text{such that} \quad V^\dagger V = \mathbb{I}_1.$$

What we mean by closed physical system is just about anything which is totally isolated from the rest of the world. The number of components n varies depending on how complicated the system is; it is called the *degrees of freedom* or the *dimension* of the system. The postulate itself is extremely short and simple. It is nonetheless puzzling as soon as you attempt to apprehend it with your classical intuition.

Example. Consider a coin, which insofar as we have always observed, can either be 'heads ⊙' or 'tails ⊛'. Thus we will suppose it has $n = 2$ degrees of freedom, and we will further assume that the state:

$$\text{'heads } \odot\text{' corresponds to quantum state } \begin{pmatrix} (1,0) \\ (0,0) \end{pmatrix}$$

$$\text{whilst 'tails } \circledast\text{' corresponds to quantum state } \begin{pmatrix} (0,0) \\ (1,0) \end{pmatrix}$$

Now if the coin was to be perfectly isolated from any outside interaction, it would start behaving like a quantum coin. When exactly does this start to happen, in real physical conditions, is a complicated matter and lies beyond the scope of this short presentation. To this day the theory of decoherence, which seeks to describe the transition between classical

and quantum behaviors, is still an open field of research. For such a quantum coin, however, we do know that the state:

$$\text{`} \circledcirc + \circledast \text{'} = \begin{pmatrix} (\frac{1}{\sqrt{2}}, 0) \\ (\frac{1}{\sqrt{2}}, 0) \end{pmatrix}$$

would become perfectly allowable. Thus a quantum coin can be in a *superposition* of heads and tails, i.e. it can be both heads and tails at the same time, in some proportion. Quantum theory is more general than our classical intuition: it allows for more possible states. It is as if 'heads' and 'tails' were two axes, and the quantum coin was allowed to live in the plane described by those axes. This is called the 'superposition principle', and it is essential.

2.2.2 Evolution

Postulate 2.2 *A closed physical system in state V will evolve, after a certain period of time, into a new state W according to*

$$W = UV$$

where U is a $n \times n$ unitary matrix of complex numbers.

Comments. In other words, in order to see how the quantum state of a closed physical system evolves, you have to multiply it by the matrix which describes its evolution (which we call U). U could be any matrix of complex numbers so long as it is $n \times n$ (remember V is an $n \times 1$ matrix) and verifies the condition $U^\dagger U = \mathbb{I}_n$.
Note that this postulate is coherent with the first one, because evolution under U takes an allowed quantum state into an allowed quantum state. Indeed suppose V is a valid state, i.e. an $n \times 1$ matrix verifying $V^\dagger V = \mathbb{I}_1$. By definition of the matrix multiplication an $n \times 1$ matrix multiplied by an $n \times n$ matrix is also an $n \times 1$ matrix, and thus W has the right sizes. Is it a unit column? Yes:

$$\begin{aligned} W^\dagger W &= (UV)^\dagger (UV) \quad \text{by definition of W} \\ &= V^\dagger U^\dagger UV \quad \text{by Property 2.2} \\ &= V^\dagger \mathbb{I}_n V \quad \text{since } U \text{ is unitary} \\ &= V^\dagger V \quad \text{by Property 2.1} \\ &= \mathbb{I}_1 \quad \text{since } V \text{ is unit} \end{aligned}$$

Thus W is a valid quantum state.

2.2.3 Measurement

Postulate 2.3 *When a physical system in state*

$$V = \begin{pmatrix} V_{11} \\ \vdots \\ V_{n1} \end{pmatrix}$$

is measured, *it yields outcome i with probability* $p_i = |V_{i1}|^2$. *Whenever outcome i occurs, the system is left in the state:*

$$W = \begin{pmatrix} (0,0) \\ \vdots \\ (1,0) \\ \vdots \\ (0,0) \end{pmatrix} \leftarrow i^{th} \text{ position}$$

Example. Suppose you have a quantum coin in state: ' $\circledcirc + \circledast$ ' $= \begin{pmatrix} (\frac{1}{\sqrt{2}},0) \\ (\frac{1}{\sqrt{2}},0) \end{pmatrix}$ which you decide to measure. With a probability $p_1 = |\frac{1}{\sqrt{2}}|^2 = \frac{1}{2}$ you will know that outcome '1' has occurred, in which case your quantum system will be left in state ' $\circledcirc$ ' $= \begin{pmatrix} (1,0) \\ (0,0) \end{pmatrix}$. But with probability $p_2 = \frac{1}{2}$ outcome '2' may occur instead, in which case your quantum system will be left in state '$\circledast$'.

Comments. Thus a measurement in quantum theory is fundamentally a probabilistic process. For this postulate to work well we need to be sure that the probabilities all sum up to 1 (so that something happens 100% of the time). But you can check that this is the case:

$$\begin{aligned} p_1 + ... + p_n &= |V_{11}|^2 + .. + |V_{n1}|^2 \quad \text{by postulate 2.3} \\ &= 1 \quad \text{by Property 2.3} \end{aligned}$$

The other striking feature of this postulate is that the state of the system gets *changed* under the measurement. In our example everything happens as though (this is an analogy!) the quantum coin in state '$\circledcirc+\circledast$' is asked to make up its mind between '$\circledcirc$' and '$\circledast$'. The quantum coin "decides" at random, but once it does it remains coherent with its decision: its new state is either '$\circledcirc$' or '$\circledast$'.

This feature provides the basis for quantum cryptography, which is one of the main topics of this thesis. Indeed suppose Alice and Bob want to communicate secretly over the phone, but Eve, the Eavesdropper, might be spying upon their conversation. What Alice and Bob

can do is to send quantum coins to each other across the (upgraded) phone network. As Eve attempts to measure what the honest parties are saying, she is bound to *change* the state of the coin. This will enable [7] Alice and Bob to detect her malevolent presence.

2.3 Deutsch-Jozsa algorithm

The measurement postulate will (probably) make you think that quantum theory is just a convoluted machinery whose only purpose is to describe objects which might be in 'state 1' with probability p_1, in 'state 2' with probability p_2 etc. until n. After all why bother thinking of the state '$\circledcirc + \circledast$' as a coin which is both heads '$\circledcirc$' and tails '$\circledast$' at the same time - when after it gets observed it collapses to either heads '$\circledcirc$' or tails '$\circledast$' anyway?
No. You *have* to consider that the coin is both '$\circledcirc$' and '$\circledast$' *until you measure it*, because this *is* how it behaves *experimentally* (until you measure it). In other words the only way to account for what happens between the moment you prepare your initial system and the moment you measure it is to think of the complex components of the state V as *amplitudes, proportions* and *not* as probabilities. This has much to do with what Postulate 2.2 enables us to do.
In this last part we shall illustrate this point by considering the simplest of all known quantum algorithms [17]. An *algorithm* is just a recipe that is used to systematically solve a mathematical problem. But the mathematical problem we will now introduce cannot be solved by classical means: it can only be solved using quantum theory, that is with a quantum algorithm. The fact that this algorithm *does work in practice* ought to demonstrate that the amplitudes of quantum theory permit us to do things which mere probabilities would not allow, and cannot explain.

2.3.1 The problem

A *boolean value* is something which can either be **True** or **False** (but not both). For instance the statement 'the sky is blue' has the boolean value **True** almost anywhere in the world with the exception of England, where it takes the value **False**.
A *boolean operator* is just a 'box' which takes one or several boolean values and returns one or several boolean values. In order to define our problem we need to become familiar with two boolean operators, which we now describe.
The boolean operator **Not** takes the boolean value **True** into **False** and the boolean value **False** into **True**. We denote this as follows:

$$\mathbf{Not}(\mathbf{True}) = \mathbf{False}$$
$$\mathbf{Not}(\mathbf{False}) = \mathbf{True}$$

The boolean operator **Xor** (exclusive or) takes two boolean values and returns one boolean value. It returns **True** either if the first boolean value it takes is **True** and the second one is **False** or if the second boolean value it takes is **True** and the first one is **False**. Otherwise it returns **False**. We denote this as follows:

$$\begin{aligned}\mathbf{Xor}(\mathbf{True},\mathbf{False}) &= \mathbf{True}\\ \mathbf{Xor}(\mathbf{False},\mathbf{True}) &= \mathbf{True}\\ \mathbf{Xor}(\mathbf{False},\mathbf{False}) &= \mathbf{False}\\ \mathbf{Xor}(\mathbf{True},\mathbf{True}) &= \mathbf{False}\end{aligned}$$

In other words **Xor** compares its two input boolean values: it returns **True** if they are different and **False** if they are the same. We are now ready to state the problem.

Problem 2.1 *Suppose we are given a mysterious boolean operator* $\mathbf{F}$ *(a black box) which takes one boolean value and returns another boolean value. We want to calculate* $\mathbf{Xor}(\mathbf{F}(\mathbf{False}),\mathbf{F}(\mathbf{True}))$*, i.e. the boolean value returned by* **Xor** *when applied to the two possible results of* $\mathbf{F}$*. But we are allowed to use the mysterious boolean operator* $\mathbf{F}$ *only once.*

It is clear that this problem cannot be solved classically. This is because in order to learn anything about $\mathbf{F}$ you will have to use $\mathbf{F}$. But we are allowed to do this only once. Suppose we use $\mathbf{F}$ on input boolean value **False**. This gives us $\mathbf{F}(\mathbf{False})$, but tells us nothing about $\mathbf{F}(\mathbf{True})$ which may still be either **True** or **False**. Thus we cannot compute $\mathbf{Xor}(\mathbf{F}(\mathbf{False}),\mathbf{F}(\mathbf{True}))$ and we fail to solve the problem. The same reasoning applies if we begin by using $\mathbf{F}$ to obtain $\mathbf{F}(\mathbf{True})$.
But what would happen if we had the possibility to use $\mathbf{F}$ upon an input boolean value which is both **True** and **False**, in some proportions (a superposition)?

2.3.2 The quantum setup

Now suppose that the mysterious boolean operator $\mathbf{F}$ is given in the form of a 'quantum black box' instead. To make this more precise we need to call

'$\mathbf{False},\mathbf{False}$' the quantum state $\begin{pmatrix}(1,0)\\(0,0)\\(0,0)\\(0,0)\end{pmatrix}$; '$\mathbf{False},\mathbf{True}$' the quantum state $\begin{pmatrix}(0,0)\\(1,0)\\(0,0)\\(0,0)\end{pmatrix}$;

'$\mathbf{True},\mathbf{False}$' the quantum state $\begin{pmatrix}(0,0)\\(0,0)\\(1,0)\\(0,0)\end{pmatrix}$; '$\mathbf{True},\mathbf{True}$' the quantum state $\begin{pmatrix}(0,0)\\(0,0)\\(0,0)\\(1,0)\end{pmatrix}$

We assume we have access, for one use only, to a physical device which implements **F** as a quantum evolution. This quantum evolution U must take

$$\text{`}\mathbf{True}, \mathbf{False}\text{' into `}\mathbf{True}, \mathbf{F}(\mathbf{True})\text{'}$$
$$\text{`}\mathbf{False}, \mathbf{False}\text{' into `}\mathbf{False}, \mathbf{F}(\mathbf{False})\text{'}$$

Notice that if for instance '$\mathbf{F}(\mathbf{True}) = \mathbf{True}$'then '$\mathbf{True}, \mathbf{F}(\mathbf{True})$' simply denotes the quantum state '$\mathbf{True}, \mathbf{True}$'. Furthermore we assume U takes

$$\text{`}\mathbf{True}, \mathbf{True}\text{' into `}\mathbf{True}, \mathbf{Not}(\mathbf{F}(\mathbf{True}))\text{'}$$
$$\text{`}\mathbf{False}, \mathbf{True}\text{' into `}\mathbf{False}, \mathbf{Not}(\mathbf{F}(\mathbf{False}))\text{'}$$

The quantum evolution U is fully specified in this manner. In matrix form it is given as follows:

$$\begin{pmatrix} (1-F_{\text{False}},0) & (F_{\text{False}},0) & (0,0) & (0,0) \\ (F_{\text{False}},0) & (1-F_{\text{False}},0) & (0,0) & (0,0) \\ (0,0) & (0,0) & (1-F_{\text{True}},0) & (F_{\text{True}},0) \\ (0,0) & (0,0) & (F_{\text{True}},0) & (1-F_{\text{True}},0) \end{pmatrix}$$

with:
F_{False} equal to 1 if $\mathbf{F}(\mathbf{False})$ is **True**, and 0 otherwise.
F_{True} equal to 1 if $\mathbf{F}(\mathbf{True})$ is **True** , and 0 otherwise.

Whatever the values of F_{False} and F_{True}, the matrix of complex numbers defined above is unitary, i.e. $U^\dagger U = \mathbb{I}_4$. Thus according to postulate 2.2 this mysterious quantum black box is perfectly allowable physically.
One may want to check that the matrix U does take '$\mathbf{True}, \mathbf{False}$' into '$\mathbf{True}, \mathbf{F}(\mathbf{True})$' etc., and that it is indeed unitary.
For our quantum algorithm we will need another quantum evolution:

$$H = \begin{pmatrix} (1/2,0) & (1/2,0) & (1/2,0) & (1/2,0) \\ (1/2,0) & (-1/2,0) & (1/2,0) & (-1/2,0) \\ (1/2,0) & (1/2,0) & (-1/2,0) & (-1/2,0) \\ (1/2,0) & (-1/2,0) & (-1/2,0) & (1/2,0) \end{pmatrix}$$

This H is also a unitary matrix of complex numbers.

2.3.3 The solution

Algorithm 2.1 *In order to solve problem 2.1 one may use the following algorithm:*

1. *Start with a closed physical system in quantum state '***False**, **True***'.*
2. *Evolve the system under the quantum evolution H.*
3. *Evolve the system under the quantum evolution U.*
4. *Evolve the system under the quantum evolution H.*
5. *Measure the system.*

If $\mathbf{Xor}(\mathbf{F}(\mathbf{False}), \mathbf{F}(\mathbf{True}))$ *is* **False** *the quantum measurement always yields outcome '2'. On the other hand if* $\mathbf{Xor}(\mathbf{F}(\mathbf{False}), \mathbf{F}(\mathbf{True}))$ *is* **True** *the quantum measurement always yields outcome '4'.*
Thus the algorithm always manages to determine $\mathbf{Xor}(\mathbf{F}(\mathbf{False}), \mathbf{F}(\mathbf{True}))$*, and does so with only one use of the quantum evolution U.*

Proof: In Step 1 we start with a closed physical system whose quantum state is $V = \begin{pmatrix} (0,0) \\ (1,0) \\ (0,0) \\ (0,0) \end{pmatrix}$.
After Step 2 the quantum state of the system has become HV. By working out this matrix multiplication we have $HV = \begin{pmatrix} (1/2,0) \\ (-1/2,0) \\ (1/2,0) \\ (-1/2,0) \end{pmatrix}$.
After Step 3 the quantum state of the system has become UHV. We can still work out the matrix multiplication but obviously the result now depends upon our mysterious boolean operator **F**. Indeed we have $UHV = \begin{pmatrix} (1/2 - F_{\text{False}}, 0) \\ (-1/2 + F_{\text{False}}, 0) \\ (1/2 - F_{\text{True}}, 0) \\ (-1/2 + F_{\text{True}}, 0) \end{pmatrix}$.
Notice that UHV depends both upon $\mathbf{F}(\mathbf{False})$ and $\mathbf{F}(\mathbf{True})$, in some proportions.
After Step 4 the quantum state of the system has become $HUHV$ and we have, by working out the multiplication: $HUHV = \begin{pmatrix} (0,0) \\ (1 - F_{\text{False}} - F_{\text{True}}, 0) \\ (0,0) \\ (F_{\text{True}} - F_{\text{False}}, 0) \end{pmatrix}$.
Finally in Step 5 we measure the state $HUHV$. According to Postulate 2.3 this yields:
- outcome '1' with probability 0 (never).
- outcome '2' with probability $p_2 = (1 - (F_{\text{False}} + F_{\text{True}}))^2$.

- outcome '3' with probability 0 (never).
- outcome '4' with probability $p_4 = (F_{\text{True}} - F_{\text{False}})^2$.

Now if $\mathbf{Xor}(\mathbf{F}(\mathbf{False}), \mathbf{F}(\mathbf{True}))$ is **False** then F_{False} and F_{True} have to be the same. Thus $F_{\text{False}} + F_{\text{True}}$ equals either 0 or 2, whereas $F_{\text{True}} - F_{\text{False}}$ is necessarily worth 0. As a consequence p_2 must equal 1 whereas p_4 is worth 0.
Similarly, if $\mathbf{Xor}(\mathbf{F}(\mathbf{False}), \mathbf{F}(\mathbf{True}))$ is **True** then F_{False} and F_{True} have to be different values. Thus $F_{\text{False}} + F_{\text{True}}$ is necessarily worth 1, whereas $F_{\text{True}} - F_{\text{False}}$ equals either -1 or 1. As a consequence p_2 is worth 0 whereas p_4 must equal 1. □

2.3.4 Comments

It is quite a remarkable fact that with only one use of the 'quantum black box' we succeed to determine a quantity which intrinsically depends 'on both possible values which the box may return'. Although this algorithm does not seem extremely useful in every day life, it teaches us an important lesson: the components of a quantum state must be viewed as proportions (amplitudes), not as probabilities. The quantum coin can be both heads and tails in some proportions, simultaneously, until you measure it.
Until recently this feature of quantum theory was essentially regarded as an unfortunate oddity which made the theory difficult to grasp. But we are now learning to turn this feature to our own advantage, as a means of 'exploring several possibilities simultaneously' (so to speak).
This is recent research however, and to this day not so many quantum algorithms are known. Yet we do know that Quantum Computers can factorize large integer numbers efficiently, or even find a name within an unordered list of 100 people in only 5 tries. These are quite useful things to be able to do. The best place to learn about them is [48].

With great pleasure we would have liked to make the thesis self-contained, pursuing this exposition to introduce axiomatically the necessary linear algebra, the density matrix formalism of quantum theory, special relativity, results in multi-party computation, etc. Unfortunately such a task would be somewhat unrealistic in terms of size, and would prevent an efficient presentation of our research results. Most of the required notions, however, are explained throughout the text. Chapter 5 in particular contains a review of the most important results about quantum operations and quantum states − from an original and instructive perspective.
The reader is advised to consult Appendix A for a quick account of the notation used in the thesis. The next part will define several of these again in a careful and precise fashion.

Part I

Representations of quantum operations

Chapter 3

Conal representation of quantum states and non trace-preserving quantum operations

Eve and the apple was the first great step in experimental science.

—James Bridie

We represent generalized density matrices of a d-complex dimensional quantum system as a subcone of a real pointed cone of revolution in $\mathbb{R}^{d^2}$, or indeed a Minkowski cone in $\mathbb{E}^{1,d^2-1}$. Generalized pure states correspond to certain future-directed light-like vectors of $\mathbb{E}^{1,d^2-1}$. This extension of the Generalized Bloch Sphere enables us to cater for non Trace-preserving quantum operations, and in particular to view the per-outcome effects of generalized measurements. We show that these consist of the product of an orthogonal transform about the axis of the cone of revolution and a positive real linear transform. We give detailed formulae for the one qubit case and express the post-measurement states in terms of the initial state vectors and measurement vectors.

The space of pure states of finite d-dimensional quantum mechanics, i.e. the (normalized) vectors of $\mathbb{C}^d$ up to a complex phase, is, like all complex spaces, not easy to visualize. Physical motions, let alone unitary time evolutions, have no clear geometrical interpretation. However, the set of hermitian operators on $\mathbb{C}^d$, $\mathrm{Herm}_d(\mathbb{C})$, is a d^2-dimensional *real* vector space, and as such is certainly easier to represent geometrically. States, or more generally density matrices, form of course a subset of $\mathrm{Herm}_d(\mathbb{C})$. The generalized Bloch sphere representation ([29][53][66]) is a famous application of this fact which has proved to be popular and elucidating: a given density matrix can be represented as a real vector inside a (hyper) sphere.
It turns out that this representation defined for density matrices, or unit trace positive elements of $\mathrm{Herm}_d(\mathbb{C})$, is only good at handling unitary or Trace-preserving quantum operations on density matrices: the former induce rotations of the Bloch vector, the latter affine transformations [66]. Individual outcomes of generalized measurements, for example, are not directly representable. Considering the insight the Bloch sphere representation gave to unitary and Trace-preserving operations, it seems interesting, for the mere sake of geometry at first, but mainly to give a useful picture to tackle quantum information problems, to extend it to cater for non Trace-preserving quantum operations. This is further motivated by the fact that the space of (semi-definite or definite) positive hermitian operators, hereafter denoted $\mathrm{Herm}_d^+(\mathbb{C})$, is a closed convex cone, and that all admissible quantum operations, whether Trace-preserving of not, should be a subset of the transformations of this cone. We hope to convince the reader that relaxing the unit trace condition and exploiting the conal geometry of $\mathrm{Herm}_d^+(\mathbb{C})$ is often illuminating.

In Section 3.1 we consider general quantum systems of d complex dimensions. We give a representation of the set of positive hermitian matrices $\mathrm{Herm}_d^+(\mathbb{C})$ as a subcone of a real Minkowski cone in $\mathbb{R}^{d^2}$, and analyze geometrical properties of generalized measurements in this setting. We find that our approach is particularly useful to represent per-outcome post-measurement states, and that pure states correspond to certain light-like vectors of the cone. Unitary operators on the complex system become real orthogonal transforms, while positive operators become real positive transforms. Section 3.2 should be of special interest for quantum information theorists: we treat the $d = 2$ one qubit case in full detail. We find further geometrical relations between measurement vectors, state vectors and post-measurement state vectors and give explicit formulae.

3.1 Conal representation of d-dimensional quantum systems

The state of such a system is described by a $d \times d$ density matrix. We shall express hermitian matrices as real linear combinations of Hilbert-Schmidt-orthogonal hermitian matrices, and then restrict this representation to elements of $\mathrm{Herm}_d^+(\mathbb{C})$. $\mathrm{Herm}_d^+(\mathbb{C})$ turns out to be "isomorphic" to a convex subcone of a cone of revolution in $\mathbb{R}^{d^2}$, or indeed a Minkowski future

cone in $\mathbb{E}^{1,d^2-1}$. We then analyze the effects of quantum operations on density matrices in this representation.

3.1.1 Hermitian matrices

Let $\{\tau_i\}$, $i \in \{1, \ldots, d^2-1\}$, be a Hilbert-Schmidt orthogonal basis (as in (3.1)) of $d \times d$ traceless hermitian matrices, and let τ_0 be the identity matrix $\mathbb{I}$. Throughout this chapter Latin indices will run from 1 to d^2-1, Greek indices from 0 to d^2-1, and repeated indices are summed unless specified. We take the τ_μ's to satisfy by definition:

$$\forall\ \mu, \nu \quad \mathrm{Tr}(\tau_\mu \tau_\nu) = d\delta_{\mu\nu} \tag{3.1}$$

with δ the Kronecker delta. $\{\tau_\mu\}_\mu$ is a basis of $\mathrm{Herm}_d(\mathbb{C})$, and any hermitian matrix $A \in \mathrm{Herm}_d(\mathbb{C})$ decomposes on this basis as

$$\begin{aligned} A &= \frac{1}{d}\big(\mathrm{Tr}(A)\mathbb{I} + \mathrm{Tr}(A\tau_i)\tau_i\big) \\ &= \frac{1}{d}\mathrm{Tr}(A\tau_\mu)\tau_\mu \end{aligned} \tag{3.2}$$

Letting $\underline{A} = (\underline{A}_\mu) \in \mathbb{R}^{d^2}$ with $\underline{A}_\mu = \mathrm{Tr}(A\tau_\mu)$ be the component vector of A in this particular basis, we have

$$\begin{aligned} &\forall\ A, B \in \mathrm{Herm}_d(\mathbb{C}), \quad AB = \frac{1}{d^2}\underline{A}_\mu \underline{B}_\nu \tau_\mu \tau_\nu \\ &\text{hence} \quad \mathrm{Tr}AB = \frac{1}{d}\underline{A}.\underline{B} \equiv \frac{1}{d}\underline{A}_\mu \underline{B}_\mu \end{aligned} \tag{3.3}$$

We shall call $\underline{A}$ the vector in $\mathbb{R}^{d^2}$, $\overrightarrow{A} = (\underline{A}_i)$ the restricted vector in $\mathbb{R}^{d^2-1}$, and ϕ the coordinate map:

$$\begin{aligned} \phi : \mathrm{Herm}_d(\mathbb{C}) &\to \mathbb{R}^{d^2} \\ A &\mapsto \underline{A} \end{aligned}$$

Equation (3.3) says that ϕ is an isometric isomorphism of $(\mathrm{Herm}_d(\mathbb{C}), \mathrm{Tr}(\ ))$ onto $(\mathbb{R}^{d^2}, (1/d)(\,.\,))$. Therefore any linear operator L on $\mathrm{Herm}_d(\mathbb{C})$ defines via ϕ and ϕ^{-1} an operator on $\mathbb{R}^{d^2}$, $M(L) = \phi \circ L \circ \phi^{-1}$. This definition yields the following 'composition' property :

Lemma 3.1 *If L_1, L_2 are linear operators on $Herm_d(\mathbb{C})$, then $M(L_i) = \phi \circ L_i \circ \phi^{-1}$ for $i = 1, 2$ are endomorphisms of $\mathbb{R}^{d^2}$ and satisfy*

$$M(L_1 \circ L_2) = M(L_1)M(L_2) \tag{3.4}$$

In particular, any complex $d \times d$ matrix A defines via $Ad_A : \rho \mapsto A\rho A^\dagger$ a linear operator on $\mathrm{Herm}_d(\mathbb{C})$ which corresponds to a real endomorphism $M(Ad_A) : \underline{\rho} \mapsto M(Ad_A)\underline{\rho}$; and $Ad_{AB} = Ad_A \circ Ad_B$ implies $M(Ad_{AB}) = M(Ad_A)M(Ad_B)$. As a direct consequence of this and the previous definitions , calling $GL_n(\mathbb{K})$ the group of invertible $n \times n$ matrices on the field $\mathbb{K}$, we get :

Lemma 3.2 *For any subgroup G of $GL_d(\mathbb{C})$, the following mapping*

$$\begin{aligned} \psi : G &\to \psi(G) \subset GL_{d^2}(\mathbb{R}) \\ A &\mapsto M(Ad_A) = \phi \circ Ad_A \circ \phi^{-1} \end{aligned} \tag{3.5}$$

is a group homomorphism. $\psi(G)$ is a subgroup of $GL_{d^2}(\mathbb{R})$.

Note that since $\psi(\mathbb{I}) = \psi(-\mathbb{I}) = \mathbb{I}$, ψ is not necessarily injective. Moreover ψ is certainly not linear. An interesting subgroup is the Special Unitary group $SU(d) = \{U \in GL_d(\mathbb{C}) \,/\, UU^\dagger = \mathbb{I},\ \det U = 1\}$. We call $SO(n) = \{O \in GL_n(\mathbb{R}) \,/\, OO^t = \mathbb{I},\ \det O = 1\}$ the special orthogonal group in n-dimensions.

Lemma 3.3 *Special Unitary transformations on $Herm_d(\mathbb{C})$, $Ad_U : \rho \mapsto U\rho U^\dagger$ with $U \in SU(d)$, induce rotations of $\mathbb{R}^{d^2}$ about the $\mathbb{I}$-axis. In fact the linear transforms $\psi(U) : \underline{\rho} \mapsto \psi(U)\underline{\rho}$ are special orthogonal and $\psi(SU(d))$ is a subgroup of $SO(d^2-1)$. It is a proper subgroup when $d \geq 3$. Moreover, $\psi(U(d)) = \psi(SU(d))$.*

Proof: Let $\rho = (1/d)(\mathrm{Tr}(\rho)\mathbb{I} + \underline{\rho}_i\tau_i)$ a hermitian matrix. Using (3.3) and the fact that Ad_U is Trace-preserving for U unitary:

$$\begin{aligned} \underline{U\rho U^\dagger}.\underline{U\rho U^\dagger} &= (\mathrm{Tr}\rho)^2 + \big(\psi(U)\underline{\rho}\big)_i\big(\psi(U)\underline{\rho}\big)_i \\ &= d\mathrm{Tr}(U\rho U^\dagger U\rho U^\dagger) \\ &= d\mathrm{Tr}\rho^2 = \underline{\rho}.\underline{\rho} \\ &= (\mathrm{Tr}\rho)^2 + \underline{\rho}_i\,\underline{\rho}_i \end{aligned}$$

In addition to preserving the first component $\underline{\rho}_0 = \mathrm{Tr}\rho$, $\psi(U)$ preserves the $\mathbb{R}^{d^2-1}$ scalar product $\underline{\rho}_i.\underline{\rho}_i$. For all $U \in SU(d)$, there exist $t \in \mathbb{R}$ and $B \in su(d)$ such that $U = U(t) = \exp(tB)$. Since $\det\psi(U(0)) = 1$ and $t \mapsto \det\psi(U(t))$ is continuous and has values in $\{\pm 1\}$, $\det\psi(U) = 1$. Thus $\psi(U)$ is a special rotation about the $\mathbb{I}$ -axis of $\mathbb{R}^{d^2}$. By Lemma 3.2, $\psi(SU(d))$ is a subgroup of the special orthogonal group $SO(d^2-1) \subset SO(d^2)$. As for all $\theta \in \mathbb{R}$, $Ad_U = Ad_{e^{i\theta}U}$, $\psi(U(d)) = \psi(SU(d))$.
Since the $\{\sqrt{-1}\tau_i\}$ span the Lie algebra $su(d)$, $\psi(SU(d))$ is the Adjoint group of $SU(d)$. For $d = 2$, we get the whole of $SO(3)$, but this is not the case for $d > 2$, as is easily seen looking at the dimensions:

$$\dim SU(d) = d^2 - 1, \qquad \dim SO(d^2-1) = \frac{1}{2}(d^2-1)(d^2-2)$$

and $\dim SU(d) < \dim SO(d^2-1)$ for $d > 2$. □

All the results of this section remain true of course when we just consider $\mathrm{Herm}_d^+(\mathbb{C})$. From now on, for any A complex $d \times d$ matrix we shall denote $\psi(A) = M(Ad_A)$ the real endomorphism of $\mathbb{R}^{d^2}$.

3.1.2 Generalized density matrices

In [66], Zanardi showed using a restriction of a mapping analogous to $\phi : A \mapsto \underline{A}$ that $d \times d$ density matrices lie in a convex subset of a ball $S \subset \mathbb{R}^{d^2-1}$. We shall extend this to a convex cone by considering generalized density matrices, by which we mean elements of $\mathrm{Herm}_d^+(\mathbb{C})$ having trace equal or inferior to one. As will become clear later this bigger space allows an elegant per-outcome representation of generalized measurements - because the trace of the positive matrix conveniently encodes the overall probability of occurrence for the state.

We define generalized pure states to be generalized density matrices which yield pure states after rescaling them to unit trace. Note that these are not the "states of partial purity" of the complex d-dimensional system, which are singular density matrices. In other words, generalized pure states are not the elements of the boundary of $\mathrm{Herm}_d^+(\mathbb{C})$ in the sense of characteristic functions of cones (see [23] for example).

Proposition 3.1 *The cone of positive hermitian matrices $Herm_d^+(\mathbb{C})$ is isomorphic to a convex subcone C of the following cone of revolution in $\mathbb{R}^{d^2}$:*

$$\Gamma = \{(\lambda_\mu) \in \mathbb{R}^{d^2} / \sum_{i=1}^{d^2-1} \lambda_i^2 \leq (d-1)\lambda_0^2, \lambda_0 \geq 0\} \tag{3.6}$$

The set of generalized density matrices verifies $\lambda_0 \leq 1$.
The set of generalized pure states verifies $\mathcal{C} = C \cap \partial\Gamma$, where $\partial\Gamma$ stands for the boundary of Γ.

Proof: We begin as in [66]. Let $\mathcal{P}$ denote the space of (not generalized) pure states in $\mathrm{Herm}_d(\mathbb{C})$. In addition to being positive, $\rho \in \mathcal{P}$ satisfies $\mathrm{Tr}(\rho^2) = \mathrm{Tr}(\rho) = 1$, so we have

$$\begin{aligned} \mathrm{Tr}(\rho^2) = \frac{1}{d}\underline{\rho}.\underline{\rho} = \frac{1}{d}\big((\mathrm{Tr}\rho)^2 + \underline{\rho}_i\underline{\rho}_i\big) = \frac{1}{d}\big(1 + \underline{\rho}_i\underline{\rho}_i\big) = 1, \quad \text{hence} \\ \underline{\rho}_i\underline{\rho}_i = d-1 \end{aligned} \tag{3.7}$$

The restricted vector $(\underline{\rho}_i)$ is on a (d^2-2)-sphere of radius $\sqrt{d-1}$, ∂S^{d^2-2}, where S is the corresponding ball. In $\mathbb{R}^{d^2}$, $\underline{\rho}$ pure sits in the intersection of the cylinder (3.7) and the $\underline{\rho}_0 = \mathrm{Tr}(\rho) = 1$ hyperplane, in other words on ∂S^{d^2-2} "centered" at $(1, 0, \ldots, 0)$.
Any density matrix can be expressed as a positive (convex) linear combination of pure states, and any positive (convex) linear combination of pure states defines a density matrix. Calling D the set of (not generalized) density matrices, $D \subset \overline{\mathrm{Hull}(\mathcal{P})}$ and $\mathrm{Hull}(\mathcal{P}) \subset D$. Here $\mathrm{Hull}(\mathcal{P})$ denotes the set of convex linear combinations of elements of $\mathcal{P}$. Since D is closed, $D = \overline{\mathrm{Hull}(\mathcal{P})}$, a well-known result. As $\phi : \mathrm{Herm}_d(\mathbb{C}) \to \mathbb{R}^{d^2}$ is linear and bi-continuous, $\phi(D) = \phi\big(\overline{\mathrm{Hull}(\mathcal{P})}\big) = \overline{\phi(\mathrm{Hull}(\mathcal{P}))} = \overline{\mathrm{Hull}(\phi(\mathcal{P}))}$. This set is a closed convex subset of S "centered" at $(1, 0, \ldots, 0)$:

$$\phi(\mathcal{P}) \subset \partial S^{d^2-2} \Rightarrow \overline{\mathrm{Hull}(\phi(\mathcal{P}))} \subset S$$

Calling $S^+ \equiv \overline{\mathrm{Hull}(\phi(\mathcal{P}))}$ the image set of density matrices as a subset of $\mathbb{R}^{d^2-1}$, we get:

$$\rho \in D \Leftrightarrow \mathrm{Tr}(\rho) = \underline{\rho}_0 = 1 \text{ and } (\underline{\rho}_i) \in S^+$$

Now a non-zero $\underline{\rho} \in \mathrm{Herm}_d(\mathbb{C})$ is positive if and only if $(1/\mathrm{Tr}\rho)\rho$ is positive, that is if and only if $\big((1/\mathrm{Tr}(\rho))\underline{\rho}_i\big) \in S^+$. In $\mathbb{R}^{d^2}$, recalling that $\mathrm{Tr}(\rho) \equiv \underline{\rho}_0$, this reads

$$\rho \in \mathrm{Herm}_d^+(\mathbb{C}) \Leftrightarrow \underline{\rho} \in \{(\lambda_0, (\lambda_i)) \in \mathbb{R}^{d^2} / (\lambda_i) \in \lambda_0\, S^+\} \tag{3.8}$$

This clearly defines a cone C in $\mathbb{R}^{d^2}$. As $S^+ \subset S$, C is a subcone of the cone of revolution Γ given by (3.6). ϕ being an isomorphism, C is convex and isomorphic to $\mathrm{Herm}_d^+(\mathbb{C})$. As pure states correspond to some points on the sphere ∂S^{d^2-2}, generalized pure states lie in the boundary of Γ. Calling $\mathcal{C}$ the set of vectors of C corresponding to generalized pure states,

we have $\mathcal{C} \subset C \cap \partial\Gamma$. Moreover $\mathcal{C} \supset C \cap \partial\Gamma$ follows from the fact that any rescaled positive matrix ρ such that $\mathrm{Tr}(\rho^2) = \mathrm{Tr}\rho = 1$ is a pure state. Remember that $\mathcal{C}$ is not the boundary of C, but a cone over $\phi(\mathcal{P})$, the image set of pure states. □

As we shall see in detail in Section 3.2 in $d = 2$ dimensions, generalized pure states correspond to future-directed light-like vectors in Minkowski space. We have shown that this remains true to a certain extent in d-complex dimensions, Γ being the future light-cone of Minkowski space $\mathbb{E}^{1,d^2-1}$ with metric $\eta_{\mu\nu} = \mathrm{Diag}(d-1, -1, \ldots, -1)$. Thus the appearance of a Minkowski product is to be expected.

As a consequence of Lemma 3.3, unitary transforms, since they leave $\mathrm{Herm}_d^+(\mathbb{C})$ invariant, yield rotations which leave C (globally) invariant. This fact deserves to be analyzed in detail to understand the geometry of C. For the moment however, we shall consider the geometric representation of general quantum operations in C.

3.1.3 Generalized measurements

We call a generalized measurement [48] a finite set $\{M_m\}_m$ of complex $d \times d$ matrices which satisfy $\sum_m M_m^\dagger M_m = \mathbb{I}$, i.e. when averaging over all outcomes the process is Trace-preserving. Note that we never use the repeated indices summation convention for m. The set of $\{E_m\}_m = \{M_m^\dagger M_m\}_m$ defines a Positive Operator Valued Measure (POVM), as $E_m \in \mathrm{Herm}_d^+(\mathbb{C})$ and $\sum_m E_m = \mathbb{I}$. Given a quantum state or density matrix $\rho \in D$, the generalized measurement $\{M_m\}_m$ on ρ yields outcome m with probability $p(m) = \mathrm{Tr}(E_m\rho)$, and if outcome m occurs, the post-measurement state is $\rho'_m = (1/\mathrm{Tr}(E_m\rho))(M_m\rho M_m^\dagger)$. We shall call $\rho_m = M_m\rho M_m^\dagger \in \mathrm{Herm}_d^+(\mathbb{C})$ the *unrescaled* post-measurement state.
Recall that any complex matrix can be polar-decomposed into a product of a unitary matrix and a positive matrix. For all m, there exists $U_m \in U(d)$ and $A_m \in \mathrm{Herm}_d^+(\mathbb{C})$ such that $M_m = U_m A_m$. As $E_m = M_m^\dagger M_m = A_m A_m$, $A_m = \sqrt{E_m}$, the positive square root of E_m. Using this polar decomposition, ρ'_m is represented in the cone C by

$$\begin{aligned}\underline{\rho'_m} \equiv \phi(\rho'_m) &= \frac{1}{\mathrm{Tr}(E_m\rho)}\phi(U_m\sqrt{E_m}\rho\sqrt{E_m}U_m^\dagger) \\ &= \frac{1}{\mathrm{Tr}(\sqrt{E_m}\rho\sqrt{E_m})}\psi(U_m)(\underline{\sqrt{E_m}\rho\sqrt{E_m}})\end{aligned}$$

Thus when outcome m occurs, the post-measurement state $\underline{\rho'_m}$ of $\{M_m\}_m$ is the same as that of $\{\sqrt{E_m}\}_m$ up to a rotation $\psi(U_m)$, and similarly for the unrescaled states. As a consequence we shall consider the geometrical effects of generalized measurements $\{\sqrt{E_m}\}_m$ where E_m and $\sqrt{E_m}$ are in $\mathrm{Herm}_d^+(\mathbb{C})$ and verify $\sum_m E_m = \mathbb{I}$, bearing in mind that the most general measurements just involve rotations on the post-measurement state vectors. For example, in Chapter 6, Eve is free to perform unitary transforms on her post-measurement states, and

can decide this according to the outcome m. The procedure we use to find the disturbance is to first measure with $\{\sqrt{E_m}\}_m$ and then maximize on unitary transforms acting upon post-measurement states. Using the conal representation, both sets of vectors $\{\underline{E_m}\}_m$ and $\{\underline{\sqrt{E_m}}\}_m$ are in C, and $\sum_m \underline{E_m} = (d, 0, \ldots, 0)$. This enables us to represent elements of a measurement inside C, and visualize the action of a particular non Trace-preserving operation $\sqrt{E_m}$ on a given density matrix ρ, in other words find $\underline{\rho_m}$ in terms of $\underline{E_m}$ or $\underline{\sqrt{E_m}}$.

3.1.4 Quantum operations represented in C

One might wonder here why not just rescale all the post-measurement states and only consider the density matrices ρ'_m. The reason for *not* doing so is that the generalized density matrices encode extra information: their "height" in the cone, the first component $\underline{\rho_m}_0 = \mathrm{Tr}(E_m\rho)$, is simply the probability of their outcomes. Under a given generalized measurement, post-measurement vectors with identical first components are equiprobable. Thus the sections of C of constant λ_0 have a clear physical interpretation. Moreover note that for E_m positive and such that $\sum_m E_m = \mathbb{I}$, and for ρ a generalized density matrix, $\underline{\rho_m}_0$ is indeed always comprised between 0 and 1. We shall now need the following simple properties:

Lemma 3.4 *For $A \in Herm_d(\mathbb{C})$ and $B, C \in Herm_d^+(\mathbb{C})$,*

$$Tr(BC) \geq 0 \qquad Tr(BABA) \geq 0$$

Proof: Let $B = \sqrt{B}\sqrt{B}$, then $\mathrm{Tr}(BC) = \mathrm{Tr}(\sqrt{B}C\sqrt{B}) \geq 0$ since $\sqrt{B}C\sqrt{B} \in \mathrm{Herm}_d^+(\mathbb{C})$. Then polar decompose A into $A = U|A|$, with U unitary and $|A| \in \mathrm{Herm}_d^+(\mathbb{C})$. As $A \in \mathrm{Herm}_d(\mathbb{C})$, $A = |A|\,U^\dagger = A^\dagger$, and

$$\mathrm{Tr}(BABA) = \mathrm{Tr}(BU\,|A|\,B\,|A|\,U^\dagger) = \mathrm{Tr}(U^\dagger BU\,|A|\,B\,|A|)$$

This is non-negative by the previous result since $U^\dagger BU, |A|\,B\,|A| \in \mathrm{Herm}_d^+(\mathbb{C})$. □

Unitary transforms induce rotations in C, and generalized measurements have the following geometric properties:

Proposition 3.2 *The linear transforms $\psi(\underline{\sqrt{E_m}}) : \underline{\rho} \mapsto \underline{\rho_m}$ associated to a generalized measurement $\{\sqrt{E_m}\}_m$, $\sum_m E_m = \mathbb{I}$ correspond to real symmetric matrices which are positive. They individually map $\mathcal{C}$ into itself. In addition, for any generalized pure state θ, $\psi(\theta)$ maps C into $\mathcal{C}$.*
The probability of outcome m for a quantum system in state ρ is given by

$$p(m) = \frac{1}{d}\underline{E_m}.\underline{\rho} \tag{3.9}$$

Proof: By using (3.2) successively, we have

$$\begin{aligned}\underline{\rho_m}_\mu &= \mathrm{Tr}(\sqrt{E_m}\rho\sqrt{E_m}\tau_\mu)\\ &= \frac{1}{d}\mathrm{Tr}(\sqrt{E_m}\tau_\nu\sqrt{E_m}\tau_\mu)\underline{\rho}_\nu \\ &\equiv M^m_{\mu\nu}\,\underline{\rho}_\nu\end{aligned} \tag{3.10}$$

Clearly $M^m_{\mu\nu}$ is real symmetric by cyclicity of the trace and the fact that $\sqrt{E_m}\tau_\nu\sqrt{E_m}$ and τ_μ are hermitian. (Actually $\psi(A)$ is real for any $d\times d$ complex matrix A, and real symmetric for any A hermitian). Let $\underline{v}=(\underline{v}_\mu)\in\mathbb{R}^{d^2}$. Using (3.10) we get

$$\begin{aligned}\underline{v}^t\psi(\sqrt{E_m})\underline{v} = \underline{v}_\mu M^m_{\mu\nu}\underline{v}_\nu &= \frac{1}{d}\underline{v}_\mu\mathrm{Tr}(\sqrt{E_m}\tau_\nu\sqrt{E_m}\tau_\mu)\underline{v}_\nu\\ &= \frac{1}{d}\mathrm{Tr}(\sqrt{E_m}(\underline{v}_\nu\tau_\nu)\sqrt{E_m}(\underline{v}_\mu\tau_\mu))\geq 0\end{aligned}$$

This follows from Lemma 3.4 since $\underline{v}_\mu\tau_\mu\in\mathrm{Herm}_d(\mathbb{C})$. Hence $M^m_{\mu\nu}$ is a positive real (symmetric) matrix.
The properties on purity simply follow from general facts on quantum operations on density matrices which remain true for generalized density matrices:
For $|u\rangle\langle u|$ and $|v\rangle\langle v|$ generalized pure states, for any A complex $d\times d$ matrix and any generalized density matrix ρ,

$$\begin{aligned}A|u\rangle\langle u|A^\dagger = |A\,u\rangle\langle A\,u| \quad &\text{and}\\ |v\rangle\langle v|\rho|v\rangle\langle v| = \langle v|\rho|v\rangle|v\rangle\langle v| &\end{aligned} \tag{3.11}$$

are generalized pure states. Relation (3.9) follows from (3.2) and $\mathrm{Tr}(\sqrt{E_m}\rho\sqrt{E_m})=\mathrm{Tr}(E_m\rho)$. □

The following properties will help to give a geometrical intuition of the action of the $\psi(\sqrt{E_m})$'s. For $\sqrt{E_m}=|v\rangle\langle v|$ pure, $\langle v|\rho|v\rangle=\mathrm{Tr}(\sqrt{E_m}\rho)=(1/d)\underline{\sqrt{E_m}}.\underline{\rho}$. Thus using (3.11):

$$\psi(\sqrt{E_m})\underline{\rho}=\frac{1}{d}(\underline{\sqrt{E_m}}.\underline{\rho})\underline{\sqrt{E_m}}$$

So $\psi(\sqrt{E_m})$ is as was expected a non-normalized projection. For any $\sqrt{E_m}\in\mathrm{Herm}_d^+(\mathbb{C})$, the d^2 eigenvectors $\underline{v^\sigma}$ of $\psi(\sqrt{E_m})$ with eigenvalues λ^σ correspond to d^2 hermitian matrices $M^\sigma\equiv\phi^{-1}(\underline{v^\sigma})$ which satisfy $\sqrt{E_m}M^\sigma\sqrt{E_m}=\lambda^\sigma M^\sigma$ (no summation). As a consequence, if $\underline{\rho}\in C$ is such an eigenvector, then the rescaled density matrix ρ is such that $\rho=\rho'_m$, i.e. ρ is unchanged if outcome m occurs.

We now give the general expressions for ρ_m in terms of $\rho \equiv (1/d)\rho_\mu\tau_\mu$ and $\sqrt{E_m} \equiv (1/d)\sqrt{\varepsilon}_\nu\tau_\nu$, where we drop the index m and do not underline the components of the vectors $\underline{\rho}$ and $\underline{\sqrt{E_m}}$ for convenience. By definition:

$$\rho_m = \frac{1}{d^3}\sqrt{\varepsilon}_\mu\rho_\nu\sqrt{\varepsilon}_\sigma\tau_\mu\tau_\nu\tau_\sigma$$

Expanding this using $\tau_0 = \mathbb{I}$ and grouping the products of the τ_i's in hermitian terms, we easily derive:

$$\begin{aligned}\rho_m = \frac{1}{d^3}\Big\{ & \sqrt{\varepsilon}_0\rho_0\sqrt{\varepsilon}_0\mathbb{I} + (2\sqrt{\varepsilon}_0\rho_0\sqrt{\varepsilon}_i + \sqrt{\varepsilon}_0\rho_i\sqrt{\varepsilon}_0)\tau_i \\ & + (\frac{1}{2}\sqrt{\varepsilon}_i\rho_0\sqrt{\varepsilon}_j + \sqrt{\varepsilon}_0\rho_i\sqrt{\varepsilon}_j)(\tau_i\tau_j + \tau_j\tau_i) \\ & + \frac{1}{2}\sqrt{\varepsilon}_i\rho_j\sqrt{\varepsilon}_k(\tau_i\tau_j\tau_k + \tau_k\tau_j\tau_i)\Big\}\end{aligned} \tag{3.12}$$

To push the general d-dimensional analysis further, we need a particular choice of τ_i's whose anti-commutation relations are convenient. We now treat in full detail the $d = 2$ (one qubit) case and apply our geometric approach to a challenging quantum information theoretical problem.

3.2 The Qubit case pushed further

Applied to qubit states the representation yields two of the most familiar objects in fundamental physics: the 2×2 positive matrices yield a Minkowski future-light-cone in $\mathbb{E}^{1,3}$ whose vertical sections are nothing but Bloch spheres. In this simple case we are able to give explicit coordinates for states posterior to non Trace-preserving quantum operations. These formulae remain simple provided Minkowski products are introduced alongside the Euclidians. They constitute a sufficient armoury to deal, using only four-vectors, with the most general evolutions to happen on a qubit, and thereby complement the analysis by Ruskai *et Al.* [51] on the geometry of Trace-Preserving quantum operations.

3.2.1 The cone and the Bloch sphere

A suitable Hilbert-Schmidt orthogonal basis for 2×2 traceless hermitian matrices is given by the set of Pauli matrices:

$$\tau_1 = \mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \mathbf{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \tau_3 = \mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

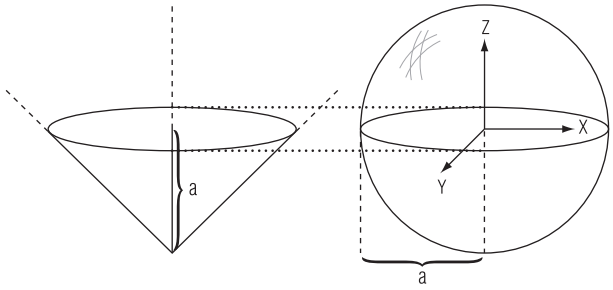

Figure 3.1: The conal representation of a qubit

Together with the identity

$$\tau_0 = \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

one may express any 2×2 hermitian matrix as a sum $A = \frac{1}{2}\underline{A}_\mu \tau_\mu$ with the $\underline{A}_\mu$'s real. The positivity conditions for those matrices turns out simple, and is a well-known result.

Lemma 3.5 *The cone of positive hermitian matrices* $Herm_2^+(\mathbb{C})$ *is isomorphic to the following cone of revolution in* $\mathbb{R}^4$*:*

$$\Gamma = \{(\lambda_\mu) \in \mathbb{R}^4 \,/\, \lambda_0^2 - \sum_{i=1}^{3} \lambda_i^2 \geq 0, \lambda_0 \geq 0\}$$

Generalized density matrices verify $\lambda_0 \leq 1$*.*
Generalized pure states lie on the boundary of Γ*.*

Proof: The eigenvalues of A are given by $\lambda_\pm = \frac{1}{2}(\underline{A}_0 \pm \sqrt{\underline{A}_i \underline{A}_i})$. A is positive if and only if $\lambda_+\lambda_- \geq 0$ and $\lambda_+ + \lambda_- \geq 0$. This is equivalent to (with $\eta_{\mu\nu} = \mathrm{Diag}(1, -1, -1, -1)$):

$$\eta_{\mu\nu}\underline{A}_\mu \underline{A}_\nu \geq 0 \wedge \underline{A}_0 \geq 0 \tag{3.13}$$

The purity condition is an obvious consequence of Proposition 3.1. □

Thus the generalized (not necessarily normalized) density matrices of a qubit cover the whole Minkowski future-light-cone in $\mathbb{E}^{1,3}$ with height less or equal to one. Taking a vertical cross-section of the cone is equivalent to fixing the trace $\underline{A}_0$ of the density matrix, which might be thought of physically as the *overall probability of occurrence* for the state. By doing so we are left with only the spin degrees of freedom along **X**, **Y**, **Z**, therefore each vertical cross-section is a Bloch sphere with radius $a = \underline{A}_0$ (see FIG. 3.1 page 53).

The ability to represent states with different traces is convenient when dealing with quantum ensembles $\{(p_x, \rho_x)\}_x$. When we seek to represent non Trace-preserving quantum operations the feature becomes absolutely crucial.

3.2.2 Post-measurement states

As we have seen in Subsection 3.1.3, the most general quantum operation can be described as $\{M_m\}_m = \{U_m\sqrt{E_m}\}_m$ with U_m unitary and $\sqrt{E_m}$ positive (the only extra feature Kraus operators allow is the possibility to ignore one's knowledge of some measurement outcomes, but in our setting this is easily dealt with by adding up the undistinguished non-normalized post-measurement states). While the action of U_m is well understood in terms of four-vectors (as a mere rotation in the Bloch Sphere, see Lemma 3.3), we are not aware of a solid real vector space geometrical framework for representing the effects of $\sqrt{E_m}$ - other than the one presented here. In Lemma 3.6, if $A \equiv \sqrt{E_m}$ while ρ is the initial state, then $A\rho A$ stands for the (not renormalized) 'post-measurement' state when outcome m has occurred (up to a unitary evolution U_m).

Lemma 3.6 *Let A and ρ be two matrices in $Herm_2^+(\mathbb{C})$. Then:*

$$\begin{aligned} A\rho A =& \frac{1}{8}[-\underline{\rho}_0(\eta_{\mu\mu'}\underline{A}_\mu\underline{A}_{\mu'}) + 2\underline{A}_0(\underline{A}.\underline{\rho})]\tau_0 \\ &+ \frac{1}{8}[\underline{\rho}_1(\eta_{\mu\mu'}\underline{A}_\mu\underline{A}_{\mu'}) + 2\underline{A}_1(\underline{A}.\underline{\rho})]\tau_1 \\ &+ \frac{1}{8}[\underline{\rho}_2(\eta_{\mu\mu'}\underline{A}_\mu\underline{A}_{\mu'}) + 2\underline{A}_2(\underline{A}.\underline{\rho})]\tau_2 \\ &+ \frac{1}{8}[\underline{\rho}_3(\eta_{\mu\mu'}\underline{A}_\mu\underline{A}_{\mu'}) + 2\underline{A}_3(\underline{A}.\underline{\rho})]\tau_3 \\ =& \frac{1}{8}[\eta_{\nu\nu'}\underline{\rho}_\nu(\eta_{\mu\mu'}\underline{A}_\mu\underline{A}_{\mu'}) + 2\underline{A}_{\nu'}(\underline{A}.\underline{\rho})]\tau_{\nu'} \end{aligned} \tag{3.14}$$

Proof: Consider

$$\underline{A} = [\ \alpha \quad \beta \quad \gamma \quad \delta\ ] \qquad \underline{\rho} = [\ a \quad x \quad y \quad z\ ]$$

We have:

$$\begin{aligned} A\rho A =& \frac{1}{8}[a(\alpha^2+\beta^2+\gamma^2+\delta^2) + 2\alpha(\beta x+\gamma y+\delta z)]\tau_0 \\ &+ \frac{1}{8}[x(\alpha^2+\beta^2-\gamma^2-\delta^2) + 2\beta(\alpha a+\gamma y+\delta z)]\tau_1 \\ &+ \frac{1}{8}[y(\alpha^2-\beta^2+\gamma^2-\delta^2) + 2\gamma(\alpha a+\beta x+\delta z)]\tau_2 \\ &+ \frac{1}{8}[z(\alpha^2-\beta^2-\gamma^2+\delta^2) + 2\delta(\alpha a+\beta x+\gamma y)]\tau_3 \end{aligned} \tag{3.15}$$

This formula can be obtained either by brute force calculation using the Pauli multiplication relations, or by exploiting the fact that Pauli matrices form a Clifford algebra I.e. $\{\tau_i, \tau_j\} = 2\delta_{ij}\tau_0$ together with equation (3.12). Regrouping the terms gives formula (3.14). □

Corollary 3.1 *Let A and ρ be two matrices in $Herm_2^+(\mathbb{C})$. $A\rho A$ can be expressed as a linear*

combination of ρ, A and the Identity:

$$A\rho A = \frac{1}{2}(\underline{A}.\underline{\rho})\, A + \frac{1}{4}(\eta_{\mu\mu'}\underline{A}_\mu\underline{A}_{\mu'})(\rho - \rho_0\tau_0)$$

This last corollary provides much geometrical insight on non Trace-preserving quantum operations. We find that the effect of $\sqrt{E_m}$ is not that difficult to visualize: the resulting state is a weighted sum of $\sqrt{E_m}$, the initial state and the identity, with real coefficients.

It is a somewhat strange fact that the structure equation (3.15) does not become apparent until one brings the Minkowski product to the rescue. The spurious appearance of special relativistic products in quantum mechanics bears some explanation in this setting however, since the Minkowski metric is intrinsically related to the characteristic function of pointed cones of revolution. We shall develop this aspect of the conal representation in Chapter 4.

Finally it is important to notice that the results expressed in these two last subsections are invariant under any orthogonal change of basis $\{\tau_i\}_i$. This is because rotations about the vertical axis leave the Minkowski product invariant. The Pauli matrices have been helpful in computing those results, but from now and in the rest of the chapter we may consider ourselves in the more general setting of Section 3.1.

3.2.3 Square and square root

In our quest towards representing non Trace-preserving quantum operations in the cone we have managed to obtain the probability of occurrence $p(m)$ in terms of $\underline{E_m}$ (Proposition 3.2). In the previous subsection we have also worked out the evolved state $\underline{\rho_m}$, but unfortunately this was done in terms of $\underline{\sqrt{E_m}}$. In order to deal fully with these operations in the cone formalism it seems, at first, that we need to understand ways of switching back and forth from $\underline{E_m}$ to $\underline{\sqrt{E_m}}$. The next Lemma is a direct consequence of equation (3.14), when $\rho = \mathbb{I}$.

Lemma 3.7 *The square of a matrix A in $Herm_2^+(\mathbb{C})$ is given by:*

$$A^2 = \underline{A}_0\, A - \frac{1}{4}(\eta_{\mu\nu}\underline{A}_\mu\underline{A}_\nu)\tau_0$$

Inversely the square root operation obeys:

$$\sqrt{A} = \frac{1}{r}(A + \frac{1}{2}\sqrt{\eta_{\mu\nu}\underline{A}_\mu\underline{A}_\nu}\,\tau_0) \qquad \textit{with} \quad r = \sqrt{\underline{A}_0 + \sqrt{\eta_{\mu\nu}\underline{A}_\mu\underline{A}_\nu}}$$

Note that A is proportional to $\sqrt{A}$ if and only if A is generalized pure or $A \propto \mathbb{I}$.
But when we seek to express a function of $\underline{E_m}$ in terms of $\underline{\sqrt{E_m}}$ (or the reverse) the next formulae become convenient.

Lemma 3.8 *Let A and ρ be two matrices in $Herm_2^+(\mathbb{C})$. The following relations hold:*

$$\begin{aligned}
\eta_{\mu\nu}\underline{\sqrt{A}}_\mu\underline{\sqrt{A}}_\nu &= 2\sqrt{\eta_{\mu\nu}\underline{A}_\mu\underline{A}_\nu} \\
\underline{A}^2.\underline{\rho} &= \underline{A}_0(\underline{A}.\underline{\rho}) - \frac{1}{2}\underline{\rho}_0(\eta_{\mu\nu}\underline{A}_\mu\underline{A}_\nu) \\
\underline{\sqrt{A}}.\underline{\rho} &= \frac{1}{r}(A.\underline{\rho} + \underline{\rho}_0\sqrt{\eta_{\mu\nu}\underline{A}_\mu\underline{A}_\nu}) \quad with \quad r = \sqrt{\underline{A}_0 + \sqrt{\eta_{\mu\nu}\underline{A}_\mu\underline{A}_\nu}}
\end{aligned}$$

On the whole taking the square root of $\underline{E}_m$ is not so easy. It would be much more convenient if we could make all calculations in terms of $\underline{E}_m$, with the added advantage that the condition

$$\sum_m E_m = 2\tau_0 \tag{3.16}$$

is easily visualized. Results in the following subsection are most useful for this purpose.

3.2.4 Inner products through quantum operations

Consider two states ρ^0, ρ^1. Suppose they undergo a quantum operation $\{M_m\}_m = \{U_m\sqrt{E_m}\}_m$ and outcome m occurs. Rather than seeking the coordinates of the rescaled post-measurement states ${\rho_m^0}'$ and ${\rho_m^1}'$, we are often interested in their positions *relative to one another*. Note this subsection reuses a number of notational conveniences introduced in Section 3.1.

Lemma 3.9 *Let ρ^0,ρ^1 be two matrices in $Herm_2^+(\mathbb{C})$ and $\sqrt{E_m}$ a measurement element in $Herm_2^+(\mathbb{C})$. The inner products of the post-measurement states satisfy:*

$$\underline{\rho_m^0}.\underline{\rho_m^1} = \frac{1}{4}[2(\underline{E_m}.\underline{\rho^0})(\underline{E_m}.\underline{\rho^1}) - (\eta_{\mu\mu'}\underline{E_m}_\mu\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho^0}_\nu\underline{\rho^1}_{\nu'})] \tag{3.17}$$

$$\underline{{\rho_m^0}'}.\underline{{\rho_m^1}'} = 2 - \frac{(\eta_{\mu\mu'}\underline{E_m}_\mu\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho^0}_\nu\underline{\rho^1}_{\nu'})}{(\underline{E_m}.\underline{\rho^0})(\underline{E_m}.\underline{\rho^1})}$$

$$\overrightarrow{\rho_m^0}.\overrightarrow{\rho_m^1} = \frac{1}{4}[(\underline{E_m}.\underline{\rho^0})(\underline{E_m}.\underline{\rho^1}) - (\eta_{\mu\mu'}\underline{E_m}_\mu\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho^0}_\nu\underline{\rho^1}_{\nu'})]$$

$$\overrightarrow{{\rho_m^0}'}.\overrightarrow{{\rho_m^1}'} = 1 - \frac{(\eta_{\mu\mu'}\underline{E_m}_\mu\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho^0}_\nu\underline{\rho^1}_{\nu'})}{(\underline{E_m}.\underline{\rho^0})(\underline{E_m}.\underline{\rho^1})} \tag{3.18}$$

Proof: By using (3.3) we have:

$$\begin{aligned}
\underline{\rho_m^0}.\underline{\rho_m^1} &= \underline{\sqrt{E_m}\rho^0\sqrt{E_m}}.\underline{\sqrt{E_m}\rho^1\sqrt{E_m}} \\
&= 2Tr(\sqrt{E_m}\rho^0\sqrt{E_m}\sqrt{E_m}\rho^1\sqrt{E_m}) \\
&= 2Tr(E_m\rho^0E_m\rho^1) \\
&= \underline{E_m\rho^0E_m}.\underline{\rho^1}
\end{aligned}$$

From there we readily obtain equation (3.17) by applying equation (3.14) once. □

By letting $\rho^0 = \rho^1 = \rho$ in the above Lemma we get:

$$
\begin{aligned}
||\underline{\rho_m}||^2 &= \frac{1}{4}[2(\underline{E_m}.\underline{\rho})^2 - (\eta_{\mu\mu'}\underline{E_m}_{\mu}\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho}_{\nu}\underline{\rho}_{\nu'})] \\
||\underline{\rho_m}'||^2 &= 2 - \frac{(\eta_{\mu\mu'}\underline{E_m}_{\mu}\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho}_{\nu}\underline{\rho}_{\nu'})}{(\underline{E_m}.\underline{\rho})^2} \\
||\overrightarrow{\rho_m}||^2 &= \frac{1}{4}[(\underline{E_m}.\underline{\rho})^2 - (\eta_{\mu\mu'}\underline{E_m}_{\mu}\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho}_{\nu}\underline{\rho}_{\nu'})] \\
||\overrightarrow{\rho_m}'||^2 &= 1 - \frac{(\eta_{\mu\mu'}\underline{E_m}_{\mu}\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho}_{\nu}\underline{\rho}_{\nu'})}{(\underline{E_m}.\underline{\rho})^2}
\end{aligned}
\tag{3.19}
$$

Equation (3.19) clearly exhibits the general property we stated in Proposition 3.2: that is if the initial state is generalized pure ($\eta_{\mu\mu'}\underline{\rho}_{\mu}\underline{\rho}_{\mu'} = 0$) or the measurement is generalized pure ($\eta_{\mu\mu'}\underline{E_m}_{\mu}\underline{E_m}_{\mu'} = 0$) then we have $||\overrightarrow{\rho_m}'|| = 1$ (pure), which implies that ρ_m is generalized pure.

The above lemma enables us to determine all the *relative positions* (angles and norms) of quantum states using relatively compact formulae which do not involve $\sqrt{E_m}$. It is only when the coordinates of each post-measurement state are required that one needs to take the impractical square root of E_m. But remember we are allowed an arbitrary rotation U_m in order to complete the quantum operation. This means we have full freedom to fix the absolute coordinates at will (so long as the relative positions are respected).

Most quantum information theoretical problems seek to evaluate the limits of quantum operations, e.g. quantum cloning [10], distinguishability [42], information gain versus disturbance tradeoff [5]. In these situations the precise individual coordinates of the states after $Ad_{\sqrt{E_m}}$ tend not to matter; usually they will need to be rotated anyhow into a position which optimizes the fidelity measure in question. What counts is the relative position of the post-measurement states. Therefore these problems can be treated comfortably in our framework. Chapter 6 provides a good example of such an application.

There are, however, some rare situations where we would like to see quantum operations act step by step, yielding precise coordinates - instead of just fixing the coordinates of the final state as we would do in order to avoid taking the square root of E_m. This is the case for instance in quantum complexity, where one needs an appreciation of how many basic computational operations it takes to accomplish some calculation. Yet in this type of problems it turns out that the basic operations can be taken to be unitary operators, with measurements only performed at the end (principle of delayed measurement [48]). Therefore these scenarios may still be analyzed comfortably within our conal representation: the basic unitary operators will just be a set of chosen real orthogonal rotations, and the final measurement statistics will be evaluated straight from E_m.

3.3 Concluding remarks

In this chapter we considered a linear embedding taking $d \times d$ positive hermitian matrices into vectors of d^2 real entries, $\phi : \rho \mapsto \underline{\rho} = (\mathrm{Tr}(\rho\tau_\mu))_\mu$. It is a well-known fact that the most general evolution a density matrix ρ may undergo is a generalized measurement $\{M_m\}_m = \{U_m\sqrt{E_m}\}_m$, where the polar decomposition was applied. In order to represent the per-outcome effect of M_m upon the real vectors we defined $\psi : A \mapsto \phi \circ Ad_A \circ \phi^{-1}$ and showed that $\psi(U_m)$ is a real orthogonal transform while $\psi(\sqrt{E_m})$ turns out to be a real positive matrix. Thus the geometrical effect of a generalized measurement can be viewed in terms of real transformations only.

Such a nice correspondence suggests quantum mechanics could be expressed elegantly over the real numbers in this manner, quite differently from its formulation in terms of real Jordan algebras [59]. However we first need to find an elegant characterization for the set of real vectors $\phi(\mathrm{Herm}_d^+(\mathbb{C}))$, and the sets of allowed orthogonal and positive transforms. For now we know that $\phi(\mathrm{Herm}_d^+(\mathbb{C}))$ is a subcone of the future-light-cone $\Gamma = \{(\lambda_\mu) \in \mathbb{R}^{d^2} / \sum_{i=1}^{d^2-1} \lambda_i^2 \leq (d-1)\lambda_0^2, \lambda_0 \geq 0\}$. Unfortunately the corresponding problem in the generalized Bloch sphere is, as soon as $d \geq 3$, regarded to be quite difficult by the researchers in the community.

One of the advantages of defining ϕ upon $\mathrm{Herm}_d^+(\mathbb{C})$ instead of the restricted set of density matrices is that $E_m = M_m^\dagger M_m$ can be visualized. In order to characterize its effects we derived rather compact and powerful formulae for the qubit case, such as the one giving the scalar product of the post-measurement states:

$$\frac{1}{4}[2(\underline{E_m}.\underline{\rho^0})(\underline{E_m}.\underline{\rho^1}) - (\eta_{\mu\mu'}\underline{E_m}_\mu\underline{E_m}_{\mu'})(\eta_{\nu\nu'}\underline{\rho^0}_\nu\underline{\rho^1}_{\nu'})]$$

By looking at such expressions it became apparent that Minkowski products have a crucial role to play in our framework, and even more so as we showed that pure quantum states correspond to light-like vectors (i.e. they sit on the boundary of Γ), even in dimensions greater than 2. Such a link with special relativity deserved to be investigated further, this is the object of the next chapter. Moreover in Chapter 6 we shall make use of the above-developed conal representation of a qubit to retrieve a classical result in quantum cryptography: Fuchs and Peres' information gain versus disturbance tradeoff.

Chapter 4

Qubit quantum operations and special relativity

Que me quiten lo bailado.
—Anónimo

We further investigate the isomorphism between non-normalized qubit states and the future cone of Minkowski space, by showing that positive operations on a qubit are proportional to pure restricted Lorentz boosts. Thus we formalize a correspondence between generalized measurements on qubit states and the Lorentz transformations of special relativity − or more precisely elements of the restricted Lorentz group together with future-directed null boosts.

This chapter develops a connection between the conal representation and special relativity, which was already hinted at in the previous chapter. The point we make is conceptual. We mathematically relate two completely different theories, quantum mechanics of the qubit on the one hand, special relativity on the other, and in the last section we daringly analyze (making our assumptions clear) whether the correspondence can be given a physical meaning. Moreover the formalism also suggests a Lorentz-invariant definition of mixedness and an interesting information conservation law. Thus the skeptical reader may simply view these results as a convenient formulation of qubit theory, in terms of real numbers only. Because this general correspondence is somewhat independent of the intricacies of the d-dimensional conal representation, we regard it as specially important to make its presentation self-contained. As a consequence this introduction must quickly overview some background material and some previous results.

As in Chapter 3 we consider generalized density matrices of a qubit, i.e. elements set of 2×2 positive complex matrices having trace inferior or equal to one. Traditionally one tends to consider normalized states only, i.e. unit trace $\mathrm{Herm}_2^+(\mathbb{C})$ matrices (density matrices). Yet relaxing this condition has a clear physical meaning and we often do so in this thesis. The most general evolution a qubit state may undergo is a generalized measurement (the only extra feature Kraus operators allow is the possibility to ignore one's knowledge of some measurement outcomes) [48]. These are described by a finite set $\{M_m\}$ of 2×2 complex matrices satisfying $\sum_m M_m^\dagger M_m = \mathbb{I}$. This last condition is crucial to ensure that averaged over all outcomes the process is Trace-preserving (conservation of probabilities), and is duly reflected in Proposition 4.3 by an appropriate rescaling. If we let $E_m = M_m^\dagger M_m$ we have that $\sum_m E_m = \mathbb{I}$, $E_m \in \mathrm{Herm}_2^+(\mathbb{C})$ and $M_m = U_m\sqrt{E_m}$ using the polar decomposition. Applied upon a density matrix ρ, the generalized measurement $\{M_m\}$ yields outcome m with probability $p(m) = \mathrm{Tr}(E_m\rho)$, in which case the post-measurement state is given by $\rho'_m = (1/\mathrm{Tr}(E_m\rho))(M_m\rho M_m^\dagger)$. As usual we call $\rho_m = M_m\rho M_m^\dagger \in \mathrm{Herm}_2^+(\mathbb{C})$ the *unrescaled* post-measurement state. Note that the generalized measurement formalism can be viewed as arising when the system is first coupled to an ancilla (through a unitary operation), which then gets measured projectively and discarded. This work takes the more axiomatic view on generalized quantum measurements.

We now restate the main ideas upon which the conal representation of the qubit is founded. Let $\{\tau_\mu\}_{\mu=0\ldots3}$ designate the set of the Pauli matrices $\mathbb{I}, \mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$. These form a Hilbert-Schmidt orthogonal basis of 2×2 hermitian matrices, that is $\forall\ \mu, \nu \quad \mathrm{Tr}(\tau_\mu\tau_\nu) = 2\delta_{\mu\nu}$ with δ the Kronecker delta. Thus any matrix $A \in \mathrm{Herm}_2(\mathbb{C})$ decomposes on this basis as

$$A = (1/2)\big(\mathrm{Tr}(A)\mathbb{I} + \mathrm{Tr}(A\tau_i)\tau_i\big) = (1/2)\mathrm{Tr}(A\tau_\mu)\tau_\mu.$$

Throughout this chapter, again, Latin indices run from 1 to 3, Greek indices from 0 to 3, and repeated indices are summed unless specified. Letting $\underline{A}_\mu = \mathrm{Tr}(A\tau_\mu)$, we shall call $\underline{A}$ the

vector $(\underline{A}_\mu) \in \mathbb{R}^4$ while $\overrightarrow{A} = (\underline{A}_i)$ will designate the restricted vector in $\mathbb{R}^3$. Note that the coordinate map

$$\begin{aligned} \phi : \mathrm{Herm}_2(\mathbb{C}) &\to \mathbb{R}^4 \\ A &\mapsto \underline{A} \end{aligned}$$

is an isometric isomorphism, in the sense that

$$\mathrm{Tr}(AB) = \frac{1}{2}\underline{A}.\underline{B} \equiv \frac{1}{2}\underline{A}_\mu \underline{B}_\mu \tag{4.1}$$

Lemma 4.1 *The cone of positive hermitian matrices $Herm_2^+(\mathbb{C})$ is isomorphic to the following cone of revolution in $\mathbb{R}^4$:*

$$\Gamma = \{(\lambda_\mu) \in \mathbb{R}^4 \,/\, \lambda_0^2 - \sum_{i=1}^{3} \lambda_i^2 \geq 0, \lambda_0 \geq 0\}$$

Generalized density matrices verifies $\lambda_0 \leq 1$.
Generalized pure states lie on the boundary of Γ.

Thus the generalized (not necessarily normalized) density matrices of a qubit cover the whole Minkowski future-light-cone in $\mathbb{E}^{1,3}$ with height less or equal to one. Taking a vertical cross-section of the cone is equivalent to fixing the trace $\underline{A}_0$ of the density matrix, which might be thought of physically as the *overall probability of occurrence* for the state. By doing so we are left with only the spin degrees of freedom along **X**, **Y** and **Z**, and therefore each vertical cross-section is a Bloch sphere with radius $a = \underline{A}_0$ (see FIG. 3.1 page 53).

Where the use of Clifford algebras is encountered such a representation of hermitian and positive 2×2 matrices is not totally uncommon: ϕ^{-1} is precisely the isomorphism used to define Dirac spinors [54] in Quantum Field Theories. Moreover Havel and Doran have already looked at quantum information theoretical processing upon four-vector representations of a qubit [27]: on page 8 they even apply a Lorentz transform to the qubit, but fall short of remarking the one-to-one relation between generalized measurements and the Lorentz transformations of special relativity. The vocabulary of geometric algebras would not have served, in any case, the exposition of the simple and general correspondence highlighted in this chapter.

We now consider the map ψ from 2×2 complex matrices to endomorphisms of $\mathbb{R}^4$ given by:

$$\psi : A \mapsto \phi \circ Ad_A \circ \phi^{-1}$$

i.e. $\psi(A)$ is the 4×4 real matrix taking a vector $\underline{\rho}$ into $\underline{A\rho A^\dagger}$. We have $\psi(AB) = \psi(A)\psi(B)$. Amongst the standard results we also have that $\psi(U)$, with U unitary, is a special orthogonal transform about the axis of revolution of the cone Γ (the proof can be found in Chapter 3). Indeed without loss of generality one can assume $\det(U) = 1$, and so the special unitary matrix can be written as:

$$U = \cos(\frac{\theta}{2})\mathbb{I} - i\sin(\frac{\theta}{2})(\overrightarrow{n_k}\tau_k) = e^{-i\frac{\theta}{2}\overrightarrow{n_k}\tau_k} \tag{4.2}$$

$$\text{and has image:} \quad \psi(U) = \begin{pmatrix} 1 & 0 \\ 0 & R_\theta(\overrightarrow{n}) \end{pmatrix}$$

Here $R_\theta(\overrightarrow{n})$ denotes the real rotation by an angle θ around the normalized axis $\overrightarrow{n}$ (to happen in the Bloch sphere). Alternatively one may use the expression $\psi(U)_{\mu\nu} = (1/2)\mathrm{Tr}(U\tau_\nu U^\dagger \tau_\mu)$. The next results are not well-known.

Lemma 4.2 *Let $\sqrt{E_m}$ be a matrix in $Herm_2^+(\mathbb{C})$, with $\underline{\sqrt{E_m}} = [\alpha\ \beta\ \gamma\ \delta]$, and E_m its square, with $\underline{E_m} = [a\ x\ y\ z]$. Then*

$$\begin{aligned}\psi(\sqrt{E_m}) &= \frac{1}{4}\begin{pmatrix} -X+2\alpha^2 & 2\alpha\beta & 2\alpha\gamma & 2\alpha\delta \\ 2\alpha\beta & X+2\beta^2 & 2\beta\gamma & 2\beta\delta \\ 2\alpha\gamma & 2\beta\gamma & X+2\gamma^2 & 2\gamma\delta \\ 2\alpha\delta & 2\beta\delta & 2\gamma\delta & X+2\delta^2 \end{pmatrix} \\ &= \frac{1}{4}\begin{pmatrix} 2a & 2x & 2y & 2z \\ 2x & X+\frac{4x^2}{2a+X} & \frac{4xy}{2a+X} & \frac{4xz}{2a+X} \\ 2y & \frac{4xy}{2a+X} & X+\frac{4y^2}{2a+X} & \frac{4yz}{2a+X} \\ 2z & \frac{4xz}{2a+X} & \frac{4yz}{2a+X} & X+\frac{4z^2}{2a+X} \end{pmatrix} \\ \text{with} \quad X &= \alpha^2-\beta^2-\gamma^2-\delta^2 = 2\sqrt{a^2-x^2-y^2-z^2}.\end{aligned} \tag{4.3}$$

Proof: $\psi(\sqrt{E_m})$ can be computed in terms of $\underline{\sqrt{E_m}}$ using the following simple formula:

$$\psi(\sqrt{E_m})_{\mu\mu'} = (1/2)\underline{\sqrt{E_m}}_{\nu}\underline{\sqrt{E_m}}_{\nu'}\mathrm{Tr}(\tau_\nu\tau_{\mu'}\tau_{\nu'}\tau_\mu)$$

This method requires lengthy calculations, subtler approaches were discussed in Chapter 3. Now let $\underline{\iota} = [1\ 0\ 0\ 0] = (1/2)\phi(\mathbb{I})$ and observe that

$$\begin{aligned}\psi(\sqrt{E_m})\underline{\iota} &\equiv \phi \circ Ad_{\sqrt{E_m}} \circ \phi^{-1}\underline{\iota} \\ &= (1/2)\phi(\sqrt{E_m}\mathbb{I}\sqrt{E_m}) \equiv (1/2)\underline{E_m}\end{aligned} \tag{4.4}$$

In other words, $(1/2)\underline{E_m}$ has as components the first column of $\psi(\sqrt{E_m})$. Thus we can now proceed to the substitutions which yield the second form of $\psi(\sqrt{E_m})$. Finally the X relation stems from:

$$\begin{aligned}\eta_{\mu\nu}\underline{\sqrt{E_m}}_{\mu}\underline{\sqrt{E_m}}_{\nu} &= 4\det(\sqrt{E_m}) \\ &= 4\sqrt{\det(E_m)} = 2\sqrt{\eta_{\mu\nu}\underline{E_m}_{\mu}\underline{E_m}_{\nu}}\end{aligned} \tag{4.5}$$

□

4.1 Quantum operations as Lorentz transforms and vice-versa

We begin by showing that elements of a generalized measurement act on a qubit either as rescaled restricted Lorentz transformations or as rescaled future-directed null boosts. Then we show that the reverse is also true. Remember that a Lorentz transform $L \equiv L^\mu_\nu$ is called restricted if it is proper ($\det L = 1$) and orthochronous ($L^0_0 > 0$). We will show that such

an L decomposes uniquely into the product of a proper spatial rotation and a pure (timelike future-directed velocity) boost. We like to think of null velocity boosts as limiting cases of restricted boosts, or effectively as elements of the topological boundary of the restricted Lorentz group, but they need to be *rescaled* to yield a finite linear transform. We shall call these (rescaled) future-directed null boosts. They are singular transforms. It turns out the rescaling introduced defines a natural unifying way of thinking about Lorentz transforms and null boosts.

If $\underline{E_m} = [a,\ x,\ y,\ z]$ corresponds to one particular measurement element $E_m = M_m^\dagger M_m$, we shall call $\underline{V_m}$ the vector of coordinates $(\underline{V_m}_\mu) = (\frac{1}{2}\eta_{\mu\nu}\underline{E_m}_\nu)$, i.e. $\underline{V_m} = [a/2,\ -x/2,\ -y/2,\ -z/2]$. Then $\underline{v_m} = 2\underline{V_m}/a$ is the corresponding normalized vector and $\overrightarrow{v_m} = [-x/a,\ -y/a,\ -z/a]$ can be thought of as a three vector velocity, whose norm is defined as usual: $v_m = (\overrightarrow{v_m}.\overrightarrow{v_m})^{1/2}$.

Proposition 4.1 *Let $\{M_m\} = \{U_m\sqrt{E_m}\}$ be a generalized measurement on a qubit, with U_m unitary and $\sqrt{E_m}$ positive. Then for all m such that E_m is not projective, we have:*

$$\psi(M_m) = \sqrt{\eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu}R_m L(\underline{v_m}) \tag{4.6}$$

where $R_m = \psi(U_m)$ is a proper rotation about the axis of the cone and $L(\underline{v_m})$ is a pure restricted Lorentz boost of normalized velocity $\underline{v_m}$. Thus $\psi(M_m)$ is a restricted Lorentz transform up to a (strictly positive) scalar. Similarly, if E_m is projective, $\psi(M_m) = (a/2)R_m L(\underline{v_m})$, where $L(\underline{v_m})$ is a rescaled pure future-directed null boost of null velocity $\underline{v_m}$.

Proof: First recall that $\psi(M_m) = \psi(U_m)\psi(\sqrt{E_m})$, and by (4.2), $\psi(U_m)$ is a special orthogonal transformation about the axis of the cone, so a restricted Lorentz transform. Suppose $\underline{E_m}$ (hence $\underline{v_m}$) timelike future-directed. Letting $\gamma \equiv 2a/X$ in (4.3) and using the definition of $\overrightarrow{v_m}$, we get:

$$\frac{4}{X}\psi(\sqrt{E_m}) = \begin{pmatrix} \gamma & -\gamma\overrightarrow{v_m}^T \\ -\gamma\overrightarrow{v_m} & \mathbb{I} + \frac{\gamma^2}{1+\gamma}\overrightarrow{v_m}\overrightarrow{v_m}^T \end{pmatrix} \equiv L(\underline{v_m}) \tag{4.7}$$

As $\gamma = 1/\sqrt{1-v_m^2}$, $L(\underline{v_m})$ is precisely a pure Lorentz boost of velocity $\underline{v_m}$ (see [54] for example). Since $\underline{v_m}$ is timelike future-directed, $\psi(M_m)$ is a restricted Lorentz transform up to the factor $X/4 = (1/2)\sqrt{\eta_{\mu\nu}\underline{E_m}_\mu\underline{E_m}_\nu} = \sqrt{\eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu}$.
Now, when $\underline{E_m}$ is null (E_m projective) this factor vanishes and γ becomes infinite. Nevertheless one can write (4.3) for $X = 0$ as

$$\frac{2}{a}\psi(\sqrt{E_m}) = \begin{pmatrix} 1 & -\overrightarrow{v_m}^T \\ -\overrightarrow{v_m} & \overrightarrow{v_m}\overrightarrow{v_m}^T \end{pmatrix} \tag{4.8}$$

We can see that this is in fact a pure null boost *rescaled* by a factor γ^{-1}. Indeed, when

$v_m \to 1$ the right-hand-side of (4.7) becomes

$$L^{null}(\underline{v_m}) \sim \gamma \begin{pmatrix} 1 & -\overrightarrow{v_m}^T \\ -\overrightarrow{v_m} & \overrightarrow{v_m}\overrightarrow{v_m}^T \end{pmatrix}$$

and since $\frac{a}{2} = \gamma\sqrt{\eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu}$, we precisely get

$$\psi(\sqrt{E_m}) \sim \sqrt{\eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu}L^{null}(\underline{v_m})$$

Here the Minkowski product vanishes and the *unrescaled* pure null velocity boost is infinite. Nevertheless rescaling $L^{null}(\underline{v_m})$ by the factor γ^{-1} yields the right-hand-side of (4.8); thus $\psi(\sqrt{E_m})$ indeed corresponds to a rescaled pure null boost, which of course is not an element of the Lorentz group. □

As we said previously the natural rescaling by the Minkowski product precisely corresponds to an appropriate rescaling of generalized Lorentz transforms bringing null boosts to finite linear maps. Formally the essence of this Proposition can be thought of as a consequence of the Alexandrov-Zeeman theorems relating the causality group (Lorentz group and dilatations) to the Minkowski causal structure, though this approach would not cover null velocity boosts. Note that the rescaled pure null velocity boosts (right-hand-side of (4.8)) are in fact proportional to projections on the null four vectors $\underline{E_m}$.
Maybe the reader wonders here why the Lorentz pure boosts corresponding to positive measurement elements E_m are parameterized by $\underline{v_m}$ and not $\underline{E_m}$. However since E_m is an operator acting on states and not a state, $\underline{E_m}$ is better thought of as a *co-vector*, or element of the *dual* space, in the same way as momenta are dual to positions in usual Special Relativity. The (contravariant) vector corresponding to $\underline{E_m}$ is precisely $2\underline{V_m}$, thus in the space of states, and not operators, E_m is represented by $2\underline{V_m}$. The factor of two was introduced merely for convenience.

The following relations suggest the Minkowski product of the state vector of a qubit is an important quantum information theoretical quantity:

Proposition 4.2 *Let $\{M_m\}$ be a generalized measurement, $\underline{\rho}$ a state vector and $\psi(M_m)\underline{\rho} \equiv \underline{\rho_m}$ the unrescaled post-measurement state vector if outcome m occurs. We have:*

$$\eta_{\mu\nu}\underline{\rho_m}_\mu\underline{\rho_m}_\nu = \eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu\eta_{\mu'\nu'}\underline{\rho}_{\mu'}\underline{\rho}_{\nu'} \tag{4.9}$$

$$\underline{\rho_m}_0 = \eta_{\mu\nu}\underline{V_m}_\mu\underline{\rho}_\nu \tag{4.10}$$

$$\eta_{\mu\nu}\underline{\rho}_\mu\underline{\rho}_\nu = 2([\mathrm{Tr}(\rho)]^2 - \mathrm{Tr}(\rho^2)) \tag{4.11}$$

Proof: We make use of the previous proposition. Equation (4.6) implies

$$\eta_{\mu\nu}\underline{\rho_m}_\mu\underline{\rho_m}_\nu = \eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu\eta_{\mu'\nu'}(R_mL(\underline{v_m})\underline{\rho})_{\mu'}(R_mL(\underline{v_m})\underline{\rho})_{\nu'}$$

and (4.9) follows since $R_mL(\underline{v_m})$ is a Lorentz transform. This relation remains true of course when $\underline{V_m}$ is light-like (E_m projective), since so is $\underline{\rho_m}$. (Purity relations of Chapter 3).
For the second equation note that $\underline{\rho_m}_0 = \mathrm{Tr}(E_m\rho) = (1/2)\underline{E_m}.\underline{\rho}$, where the isometry (4.1) was applied. Introducing the definition of $\underline{V_m}$ in this last equation yields the required result. Equation (4.11) can be shown explicitly using the components of ρ and ρ^2, but it seems more interesting to use our isomorphism $\phi : \rho \to \underline{\rho}$. Consider the linear map on $\mathbb{E}^{1,3}$, $\Lambda : (\underline{\rho}_\mu) \to (\eta_{\nu\mu}\underline{\rho}_\nu)$ (musical isomorphism). Then $\widetilde{\Lambda} : \rho \to \phi^{-1} \circ \Lambda \circ \phi(\rho)$ is a linear map on $\mathrm{Herm}_2(\mathbb{C})$. One finds easily $\widetilde{\Lambda}(\rho) = (\mathrm{Tr}\rho)\mathbb{I} - \rho$. Using the fact that ϕ is an isometry (4.1), we get

$$\eta_{\mu\nu}\underline{\rho}_\mu\underline{\rho}_\nu \equiv (\Lambda\underline{\rho}).\underline{\rho} = 2\mathrm{Tr}(\widetilde{\Lambda}(\rho)\rho) = 2([\mathrm{Tr}\rho]^2 - \mathrm{Tr}(\rho^2))$$

□

It seems interesting that this quantity, invariant under Lorentz transforms on the state vector $\underline{\rho}$, in fact measures the mixedness of qubit states: recall that a density matrix ρ is pure if and only if $\mathrm{Tr}(\rho^2) = (\mathrm{Tr}\rho)^2$. Not only is purity preserved under a formal Lorentz boost, so is this notion of mixedness. Moreover this quantity maps according to the simple relation (4.9) under a generalized measurement. Note that since $\eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu \leq 1$, the mixedness always decreases given a measurement outcome. But (4.9) and (4.10) suggest much more: the mixedness of post-measurement states and their probabilities are invariant if both the initial vector $\underline{\rho}$ and the measurement vectors $\underline{V_m}$ are Lorentz transformed. However, the set of transformed measurement vectors does not sum to the identity, and it is unclear how to interpret it as a quantum measurement. In Section 4.2 we will discuss the way a boosted observer perceives measurement probabilities, but without using the approach equation (4.10) might suggest. We now show that any Lorentz transformation can be thought of as an element of a generalized measurement up to scale.

Proposition 4.3 *Let L a restricted Lorentz transform or a rescaled future-directed null boost of $\mathbb{E}^{1,3}$. L decomposes as $L = RL(\underline{v})$ where R is a proper Lorentz rotation and $L(\underline{v})$ a pure velocity boost, rescaled when $\underline{v}$ is null. Then there exists a particular element M_1 of a measurement scheme $\{M_m\}$, $\sum_m M_m^\dagger M_m = \mathbb{I}$, such that for any qubit ρ,*

$$L\underline{\rho} \propto \psi(M_1)\underline{\rho} \tag{4.12}$$

Thus the effect of a Lorentz boost on a qubit can essentially be viewed as applying a particular measurement element whose outcome occurs. More precisely there exits a family of such

possible measurement elements $M(\lambda) = U\sqrt{E(\lambda)}$ *defined by* $U = U(R)$ *as in (4.2) and* $\sqrt{E(\lambda)}$ *satisfying the following:*
If $L = RL(\underline{v})$ *is a restricted Lorentz transform:*

$$\underline{\sqrt{E(\lambda)}} = (1+\sqrt{1-v^2})^{-1/2}[\lambda(1+\sqrt{1-v^2}), -\lambda\overrightarrow{v}] \qquad \textit{with} \quad 0 < \lambda \leq \sqrt{\frac{2}{1+v}}$$

while if $L = RL(\underline{v})$ *is a rescaled future-directed null boost:*

$$\underline{\sqrt{E(\lambda)}} = [\lambda,\ -\lambda\overrightarrow{v}] \quad \textit{with} \quad 0 < \lambda \leq 1.$$

Proof: For completeness we first show the decomposition of restricted Lorentz transforms L into $L = RL(\underline{v})$ as above. This relies on the well-known spinor representation of the restricted Lorentz group, or the two-to-one group homomorphism between unimodular 2×2 complex matrices and restricted Lorentz transforms (see [54] for example):

$$\begin{aligned} \psi : SL(2,\mathbb{C}) &\to SO(1,3)^+ \\ A &\mapsto \psi(A) \equiv \phi \circ Ad_A \circ \phi^{-1} \end{aligned}$$

Indeed as Ad_A preserves the determinant and ϕ is such that for all $\rho \in \mathrm{Herm}_2(\mathbb{C})$, $\det\rho = (1/4)\eta_{\mu\nu}\underline{\rho}_\mu\underline{\rho}_\nu$, $\psi(A)$ preserves the Minkowski product. The fact that $\psi(A) \in SO(1,3)^+$ and that ψ is two-to-one and onto can be checked explicitly. Let L any restricted Lorentz transform. There exits a unique $A \in SL(2,\mathbb{C})$ such that $\psi(\pm A) = L$. Polar decompose A into $A = U|A|$ with U unitary and $|A|$ positive. (U is in fact special unitary and $|A|$ positive definite since $\det A = 1$, and by unicity of the polar decomposition for A non-singular, $-A = (-U)|A|$). Applying Proposition 4.1 to $|A|$ with $\det|A|^2 = 1$, $\psi(|A|)$ is a *pure* restricted Lorentz boost, thus $L = \psi(U)\psi(|A|)$ provides a decomposition. Since $\psi(U) = \psi(-U)$, this decomposition is unique.
Thus given $L = RL(\underline{v})$, with R a proper rotation and $L(\underline{v})$ a pure boost of future-directed timelike velocity $\underline{v} = [1,\ \overrightarrow{v}]$, we use Proposition 4.1 to find $M = U\sqrt{E}$ such $\psi(M) \propto L$. $U = U(R)$ is given by (4.2) and we choose $\underline{E} = [1,\ -\overrightarrow{v}]$.
We then have to find $\lambda > 0$ such that λM can be part of a measurement scheme. This is equivalent to $\lambda^2 M^\dagger M$ positive (satisfied) and $\mathbb{I} - \lambda^2 M^\dagger M$ positive too. (λM and $-\lambda M$ are equivalent in terms of measurement elements). With $\lambda M = U\sqrt{E(\lambda)}$, we have

$$\underline{E(\lambda)} = [\lambda^2,\ -\lambda^2\overrightarrow{v}],$$

from which we find $\underline{\sqrt{E(\lambda)}}$ using (4.4): $\underline{\sqrt{E(\lambda)}} = (1+\sqrt{1-v^2})^{-1/2}[\lambda(1+\sqrt{1-v^2}),\ -\lambda\overrightarrow{v}]$.
Then requiring $\mathbb{I} - E(\lambda)$ positive is equivalent to $(\lambda > 0)$: $\lambda \leq \sqrt{\frac{2}{1+v}}$

Applying Proposition 4.1 we get:

$$\psi(M(\lambda)) = \frac{\lambda^2}{2}\sqrt{1-v^2}RL(\underline{v}).$$

Thus for such λ the measurement elements $M(\lambda) = U\sqrt{E(\lambda)}$ are all possible measurements whose occurrence is equivalent up to factor to the restricted Lorentz boost $L = RL(\underline{v})$.

Now let L a rescaled future-directed null boost. As we have shown, any restricted Lorentz transform can be decomposed into a product of a proper rotation and a boost of timelike future-directed velocity. Future-directed null boosts are no exception, and thus the rescaled null boosts L may be assumed to be the product of a rotation R and a rescaled null pure boost $L(\underline{v})$ of type (4.8). The rotation can be dealt with as in the previous case. Defining $\underline{E} = [1,\ -\overrightarrow{v}]$ null future-directed, we have $L(\underline{v}) \propto \psi(\phi^{-1}(\sqrt{\underline{E}}))$. Then again we consider $\underline{E(\lambda)} = \lambda^2\underline{E}$ $(\lambda > 0)$ such that $\mathbb{I} - E(\lambda)$ is positive. This is equivalent to $0 < \lambda \leq 1$, and using (4.4) we have $\sqrt{\underline{E(\lambda)}} = [\lambda,\ -\lambda\overrightarrow{v}]$, which gives $\psi(\sqrt{M(\lambda)}) = (\lambda^2/2)RL(\underline{v})$. Note that the scaling factor is always less than 1, indeed less than $\sqrt{(1-v)/(1+v)}$ in the restricted case, and 1/2 in the null case. □

Overall we have shown that elements of generalized measurements on a qubit are *equivalent* to rescaled restricted or null Lorentz transforms. Projective measurement elements are future-directed null boosts, while mixed ones correspond to restricted Lorentz boosts. One can of course think of these linear transforms as elements or limits of elements of the causality group of $\mathbb{E}^{1,3}$.

4.2 Discussion

The following is a somewhat original discussion of Propositions 4.1 to 4.3. Our formalism and its consequences suggest that qubit states may be viewed as spatio-temporal objects, or indeed as four-vectors of a Minkowski spacetime. This differs only slightly from the notion of spin as a spatial polarization direction, and thus may apply to 2 dimensional quantum systems whose degrees of freedom can be thought of as spacelike. We shall adopt this point of view from now, i.e consider naively qubits as four-vectors, and analyze the physical implications. Let us begin by merely rephrasing the content of the correspondence that was established in Section 4.1. Suppose Alice proceeds to a generalized measurement $\{M_m\} = \{U_m\sqrt{E_m}\}$, $\sum_m M_m^\dagger M_m = \mathbb{I}$ on a qubit density matrix ρ (ρ is unit trace). With probability $p(m) = \mathrm{Tr}(E_m\rho)$ this will yield her a (non-normalized) post-measurement state $\rho_m = M_m\rho M_m^\dagger$. This rather common situation turns out to be equivalent, according to Proposition 4.1, to the following less usual scenario:

Scenario 1: Suppose Alice is standing at the origin of an inertial frame of Minkowski spacetime, contemplating the four-vector $\underline{\rho}$. Say she gives herself a set of rotations $\{R_m\}$ and four-vectors $\{\underline{V_m}\}$ such that $\sum_m \underline{V_m} = [1\ 0\ 0\ 0]$. Now, with probability $p(m) = \eta_{\mu\nu}\underline{V_m}_\mu\underline{\rho}_\nu$ she chooses to Lorentz boost herself up to velocity vector $\underline{v_m} = \underline{V_m}/\underline{V_m}_0$, to rotate the resulting space-frame by R_m and to rescale her coordinates by a factor of $\sqrt{\eta_{\mu\nu}\underline{V_m}_\mu\underline{V_m}_\nu}$ (we are assuming E_m is not projective). She then looks back upon her object of contemplation and sees $\underline{\rho_m}$, the unrescaled post-measurement state. The case with E_m projective is the limit of the previous one when the boost vector $\underline{v_m}$ becomes null, and the rescaling yields finiteness of the corresponding linear transform.
Therefore a quantum measurement can be thought of, up to scale, as the observer taking a Lorentz boost relative to his or her qubit. Notice that applying a second quantum measurement $\{N_n\}$ similarly corresponds to the observer taking a second (successive) Lorentz transformation at random amongst $\{L_n\}$, say. Thus qubit quantum mechanics can easily be axiomatized within the mathematics of special relativity, and pure measurement elements go hand-in-hand with future-directed null boosts.

Difficulties are prompt to arise when one seeks to equate a measurement interaction, in which the qubit is physically acted upon, with a (somewhat passive) coordinate transformation in Minkowski spacetime: indeed the latter is purely kinematical, thus reversible, whereas the former usually implies a collapse of the state. In the following scenario we dissociate one from the other. In other words we consider special relativity and qubit quantum theory in their most usual fashion, save for the fact that we continue to interpret the spin as a four-vector.

Scenario 2: Suppose Alice is at the origin of an inertial frame of Minkowski space, together with a qubit density matrix ρ (unit trace) which we think of as a (normalized) spacetime vector $\underline{\rho}$. If we consider the point of view of Bob as he passes by in an inertial frame, this suggests that Bob sees a boosted version of ρ, i.e. a state $\Lambda\underline{\rho}$. This seemingly innocuous point raises an important issue however: Λ is not restricted to Bloch sphere rotations, and thus may indeed not correspond to a unitary operation. To understand its effect upon ρ we must refer to Proposition 4.3: Λ acts, up to a factor, as a measurement element M_1 whose outcome always happens, even though $\mathrm{Tr}(M_1\rho M_1^\dagger) \neq 1$. Thus $\{M_1\}$ can be thought of as a non trace-preserving quantum operation ($M_1M_1^\dagger \neq \mathbb{I}$) which systematically occurs. We shall let $\underline{\rho^{Bob}} \equiv \Lambda\underline{\rho} \propto \underline{M_1\rho M_1^\dagger}$ and proceed to reassure the reader: such a phenomenon would not violate the principle of relativity. Bob does not *make happen* a non trace-preserving quantum operation on the qubit. The laws of quantum mechanics remain exactly the same in every inertial frame: only the *change of observers*, or more precisely the way a boosted observer perceives a non-boosted state, is a non-orthodox quantum operation. If Bob were then to decelerate down to the speed of Alice, his mathematical description of the qubit would return to be ρ again.
Now suppose Alice measures ρ under a generalized measurement $\{N_n\}$. The probability

associated with the transition from ρ to ρ_n is given by $p(n) \equiv \mathrm{Tr}(N_n^\dagger N_n \rho)/\mathrm{Tr}(\rho) = \underline{\rho_n}_0$, as usual when $\underline{\rho}$ is normalized. As Bob passes, he sees the initial state $\underline{\rho^{Bob}} = \Lambda\underline{\rho}$, and the post-measurement states $\underline{\rho_n^{Bob}} = \Lambda\underline{\rho_n}$. Remember that the probability associated to a state is simply given by the first component of its vector representation. Assuming Λ is a pure boost of non-null normalized velocity $\underline{v(\Lambda)}$, we get:

$$p^{Bob}(n) \equiv \frac{\mathrm{Tr}(\rho_n^{Bob})}{\mathrm{Tr}(\rho^{Bob})} = \frac{\underline{\rho_n^{Bob}}_0}{\underline{\rho^{Bob}}_0} = \frac{p(n) - \overrightarrow{v(\Lambda)}.\overrightarrow{\rho_n}}{1 - \overrightarrow{v(\Lambda)}.\overrightarrow{\rho}} \geq 0$$

In other words the probabilities associated with the transitions from $\underline{\rho}$ to $\underline{\rho_n}$, in the same way as lengths of objects, are not invariant under a change of observer. Thus if one believes probabilities are absolute quantities independent of notions of space and time, one must abandon trying to interpret the qubit as a four-vector.
Otherwise, the notion of probability as a physical quantity needs to be redefined ($\sum_n p(n)$ is not conserved, as the probability of a state transforms just like the time-component of a four vector). The idea is disturbing, and certainly worth comparing with the contraction of any spatial object (a ruler, say) under a Lorentz boost. As he passes by Bob will see Alice's $20cm$ ruler shrunk down to $15cm$. But what we now have is that if Alice's quantum ruler has half a chance of being $22cm$ long, and another half chance of measuring $18cm$, it may well turn out that Bob instead perceives a quantum ruler of length $17cm$ with probability a third, and $14cm$ two third of the times.

Allowing the Lorentz boosts Λ to act on $\underline{\rho}$ as on spacetime vectors thus seems a radical departure from Quantum Field Theories in Minkowski space, where the approach is to seek *unitary* representations of the Poincaré group, i.e. the full Lorentz group together with translations. However, Poincaré invariance (see [63] for example) does not require any given state of a theory to transform unitarily under a change of *observer*: for any two inertial observers Alice and Bob, it requires the existence, given any state of the theory possibly measured by Alice in her frame, of another state of the theory measured by Bob in his frame, such that the statistics of their measurement outcomes on their respective states are the same. In this sense, the action of a particular Poincaré transform on a state in Quantum Field Theory corresponds to a change of *inertial frame*: it maps a given solution for an inertial family of observers to another *equivalent* solution for another family of observers, hence it simply cannot change the measurement statistics. Our second scenario does not involve a change of inertial frame, but just a change of observer. It is true that nonetheless, Alice's non-boosted qubit viewed by a boosted observer Bob, though not necessarily unitarily equivalent to the same non-boosted state viewed by Alice, should be an admissible state of the theory which could be measured by Bob to yield measurement statistics with the usual properties. We are not in this case, since in scenario 2, Bob is not performing a quantum operation on Alice's qubit. Note also that in the formalism developed above, pure states, whether viewed in their

inertial frame or not, remain pure.

But if we begin to think of quantum measurement outcome probabilities as not invariant under Lorentz transformations, then the Von Neumann entropy should not be either. On the other hand the invariant quantity $\eta_{\mu\nu}\underline{\rho}_\mu\underline{\rho}_\nu$ seems a good measure of the mixedness of ρ, an idea which is strongly supported by its equivalent form (4.11). With $I(\underline{\rho})$ proportional to the logarithm of $\eta_{\mu\nu}\underline{\rho}_\mu\underline{\rho}_\nu$ equation (4.9) becomes:

$$I(\underline{\rho_m}) = I(\underline{V_m}) + I(\underline{\rho})$$

This result is rather interesting as an information conservation law.

More generally we feel that the correspondence between qubit quantum operations and special relativity transforms deserves further attention. The lines of thought suggested in this discussion section need to be anchored in firmer ground and generalized to higher dimensional quantum systems. Although most of the mathematical results of this chapter stem from the exceptional isomorphism between $\mathrm{Herm}_2^+(\mathbb{C})$ and the future cone of Minkowski space, the results in Chapter 3 give us hope to find a special relativistic interpretation to d-dimensional systems also. Moreover the thorough study of the properties of quantum states and quantum operations, which is pursued the next chapter, is likely to turn out helpful for this purpose.

Chapter 5

On quantum operations as quantum states

Marins qui rêvez en haute mer, les coudes appuyés sur la lisse, craignez de penser longtemps dans le noir de la nuit à un visage aimé.

—*Jules Supervielle*

We formalize Jamiolkowski's correspondence between quantum states and quantum operations isometrically, and harness its consequences. This correspondence was already implicit in Choi's proof of the operator sum representation of Completely Positive-preserving linear maps; we go further and show that all of the important theorems concerning quantum operations can be derived directly from those concerning quantum states. As we do so the discussion first provides an elegant and original review of the main features of quantum operations. Next (in the second half of the chapter) we find more results stemming from our formulation of the correspondence. Thus we provide a factorizability condition for quantum operations, and give two novel Schmidt-type decompositions of bipartite pure states. By translating the composition law of quantum operations, we define a group structure upon the set of totally entangled states. The question whether the correspondence is merely mathematical or can be given a physical interpretation is addressed throughout the text: we provide formulae which suggest quantum states inherently define a quantum operation between two of their subsystems, and which turn out to have applications in quantum cryptography.

This chapter is concerned with the properties of positive matrices (quantum states) and the linear maps between these, i.e. Positive-preserving linear maps and Completely Positive-preserving linear maps (quantum operations), as provided by the density matrix formalism of finite dimensional quantum theory. The analysis we carry out is formal and mathematical, and although it focuses on quantum information theoretical issues, it should have applications in other domains. So far the thesis has concentrated upon finding real vector space representations of quantum states and quantum operations, but in this chapter we remain upon the complex field. We follow, instead, another driving line: formalizing and exploiting systematically an isomorphism from hermitian matrices to Hermitian-preserving linear maps and quantum states to quantum operations. To our knowledge, this isomorphism was first suggested by Jamiolkowski [32], and later exploited by Choi [13] to obtain the operator sum representation theorem. However the latter had already been independently derived by Kraus [37] (see also [38]) in infinite dimensions. Our investigation shows that the isomorphism between states and operations has a much wider range of implications, whether to simplify the proofs of well-known results or to point out novel properties, both technical and geometrical. The presentation is rigorous and self-contained, it contains some introductory material presented from an original perspective.

In Section 5.1, after setting our conventions, we relate vectors to matrices, and matrices to superoperators, the idea being to map an $mn \times mn$ matrix to a linear operator from $n \times n$ matrices to $m \times m$ matrices. These isomorphisms are often viewed pragmatically as rearrangements of the coordinates of vectors or matrices, but we formalize them more abstractly as norm-preserving bijections between tensor product spaces. We derive original formulae relating to these isomorphisms which we use throughout the chapter. One of them will simplify those numerous mathematical problems in quantum cryptography which require a careful optimization of the fidelities induced by a quantum operation. This formal setting leads in Subsection 5.1.3 to the state-operator equivalence, inherently present in the works many, but rarely exploited as such: non-normalized quantum states of an mn-dimensional system are equivalent to quantum operations from an n-dimensional system to an m-dimensional one. We use this correspondence in Subsection 5.2.1 to rederive all the main properties of quantum operations from those of quantum states: the operator sum decomposition and its unitary degree of freedom stem from the spectral decomposition and Hughston-Josza-Wooters theorems; the factorizability of quantum operations up to a trace-out corresponds to the purification of quantum states; and the polar decomposition of matrices is equivalent to the Schmidt decomposition of pure states. Next, in Subsection 5.2.2, we consider properties of states (or operations) whose translation in terms of operators (or states) was unknown to us previously. Mainly we give a factorizability condition for quantum operations, i.e. a criteria for an operator to be single operator in the operator sum representation; and we find two original triangular decompositions of pure states of a bipartite system. Throughout the section the

normalization of density matrices is unimportant. Yet for completeness the reader is reminded of the well known Trace-preserving conditions in Subsection 5.2.3 (both in terms of states and operators). Moreover we highlight the fact that maximally entangled pure states of a bipartite system go hand in hand with isometric maps from one subsystem to the other (unitary maps in case both systems have the same dimension). Choi's extremal Trace-preserving condition is also presented and recasted in terms of the rank of an easily constructed matrix.
Section 5.3 is devoted to geometrical structures of quantum states. We exploit the composition law on Completely Positive-preserving maps to define a semi-group structure on the states of n^2-dimensional quantum systems, and show that the subset of totally entangled pure states is isomorphic to the group of invertible $n \times n$ matrices defined up to phase (with maximally entangled pure states corresponding to unitary transforms as in [39]). These group isomorphisms have profound structural meaning, and are useful in finding nice coordinate charts on such spaces. We also give an exotic composition law on operators stemming from the Schur product on states. In Subsection 5.3.2 we make use of the dual mapping between states and positive functionals, and readily show that the space of Positive-preserving maps is dual to that of separable states of a bipartite system. This yields a simple result which is in fact equivalent to Peres' separability criterion. More generally the notion of duality seems to help provide possible physical interpretations of the state-operator correspondence formulae, notably as we show that the effect of any quantum operation can be viewed as the trace out of a particular local single operation on its corresponding state.
We conclude in Section 5.4 and give a table summarizing the main results of this chapter.

5.1 The setting

Notation. The present chapter makes use of the same notational conventions that were introduced in Chapter 3 (see also Appendix A), together with some minor additions. For convenience we now recall the lot: we will denote by $M_d(\mathbb{C})$ the set of $d \times d$ matrices of complex numbers, and by $\mathrm{Herm}_d(\mathbb{C})$ its hermitian subset. Amongst the latter we will denote by $\mathrm{Herm}_d^+(\mathbb{C})$ the set of positive matrices, and also refer to it as the set of (non-normalized) states of a d-dimensional quantum system. An important subset of $\mathrm{Herm}_{mn}^+(\mathbb{C})$ is the set of *separable states*, i.e. those which can be written in the form

$$\rho = \sum_x \lambda_x \rho_1^x \otimes \rho_2^x$$

where $\lambda_x \geq 0$ and the ρ_1^x and ρ_2^x belong to $\mathrm{Herm}_m^+(\mathbb{C})$ and $\mathrm{Herm}_n^+(\mathbb{C}^)$ respectively. Later we shall denote this set by $\mathrm{Herm}_{mn}^S(\mathbb{C})$.
Throughout the dagger operation † will be somewhat overloaded, in a manner which has now become quite standard: as usual a ket $A = \sum A_i|i\rangle$ will be taken into a bra $A^\dagger = \sum A_i^* \langle i|$,

while a matrix $\hat{A} = \sum A_{ij}|i\rangle\langle j|$ will be mapped into its conjugate transpose $\hat{A}^\dagger = \sum A^*_{ij}|j\rangle\langle i|$. In other words, † takes kets into bras using the canonical complex scalar product for vectors, i.e. $B^\dagger \equiv [A \mapsto (B, A) = \sum B^*_i A_i \equiv B^\dagger A]$, but for linear maps on vectors it denotes the usual adjoint operation defined with respect to the same scalar product. We also make frequent use of the conjugation operation * which is defined in the canonical basis to take kets $A = \sum A_i|i\rangle$ into $A^* = \sum A^*_i|i\rangle$, and similarly on bras. Linearity will refer to complex linearity.

Definition 5.1 *A linear map $\Omega : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ is Hermitian-preserving if and only if for all ρ in $Herm_m(\mathbb{C})$, $\Omega(\rho)$ belongs to $Herm_n(\mathbb{C})$.*

The following is a well-known fact:

Remark 5.1 *If $\Omega : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ is a Hermitian-preserving linear map, then so is $\Omega \otimes \mathbb{I}_r$.*

Proof: Let us denote by $\{\tau_i\}$ and $\{\tau_j\}$ two sets of hermitian matrices forming a basis of $\mathrm{Herm}_m(\mathbb{C})$ and $\mathrm{Herm}_r(\mathbb{C})$ respectively, considered as a real vector spaces. $\{\tau_i \otimes \tau_j\}$ forms a basis for $\mathrm{Herm}_{mr}(\mathbb{C})$. Now consider $Z \in \mathrm{Herm}_{mr}(\mathbb{C})$, so that $Z = \sum_{ij} z_{ij}\tau_i \otimes \tau_j$ with $z_{ij} \in \mathbb{R}$. We then have

$$\begin{aligned}(\Omega \otimes \mathbb{I}_r)Z &= \sum_{ij} z_{ij}\Omega(\tau_i) \otimes \tau_j = \sum_{ij} z_{ij}\Omega(\tau_i)^\dagger \otimes \tau_j^\dagger = \sum_{ij} \big(z_{ij}\Omega(\tau_i) \otimes \tau_j\big)^\dagger \\ &= ((\Omega \otimes \mathbb{I}_r)Z)^\dagger\end{aligned}$$

□

Definition 5.2 *A linear map $\Omega : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ is Positive-preserving if and only if for all ρ in $Herm^+_m(\mathbb{C})$, $\Omega(\rho)$ belongs to $Herm^+_n(\mathbb{C})$.*

A Positive-preserving map is necessarily Hermitian-preserving since any hermitian matrix can be expressed as the difference of two positive matrices. Note also that having $\Omega : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ a Positive-preserving linear map does not imply that $\Omega \otimes \mathbb{I}_r$ is also Positive-preserving. *Example.* The map

$$\begin{aligned}{}^t : \mathrm{Herm}_2^+(\mathbb{C}) &\to \mathrm{Herm}_2^+(\mathbb{C}) \\ \rho &\mapsto \rho^t\end{aligned}$$

is clearly Positive-preserving, but $({}^t \otimes \mathbb{I}_2)$ is not: indeed let $|\beta\rangle = |00\rangle + |11\rangle$, $|\gamma\rangle = |01\rangle + |10\rangle$ and $|\delta\rangle = |01\rangle - |10\rangle$. Note that $|00\rangle$, $|11\rangle$, $|\gamma\rangle$ and $|\delta\rangle$ form an orthogonal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$.

$$\begin{aligned}({}^t \otimes \mathbb{I}_2)(|\beta\rangle\langle\beta|) &= |00\rangle\langle 00| + |10\rangle\langle 01| + |01\rangle\langle 10| + |11\rangle\langle 11| \\ &= |00\rangle\langle 00| + |11\rangle\langle 11| + |\gamma\rangle\langle\gamma| - |\delta\rangle\langle\delta|\end{aligned}$$

which is not positive since $\langle\delta|({}^t\otimes\mathbb{I}_2)(|\beta\rangle\langle\beta|)|\delta\rangle < 0$.

Definition 5.3 *A linear map $\Omega : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ is Completely Positive-preserving if and only if for all r and for all ρ in $Herm^+_{mr}(\mathbb{C})$, $(\Omega \otimes \mathbb{I}_r)(\rho)$ belongs to $Herm^+_{nr}(\mathbb{C})$.*

5.1.1 Isomorphisms

Next we relate vectors of $\mathbb{C}^m \otimes \mathbb{C}^n$ to endomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$. The tensor split of $\mathbb{C}^{mn}$ into $\mathbb{C}^m \otimes \mathbb{C}^n$ is considered fixed, as will be all tensor splits throughout the chapter unless specified otherwise. (Notions of entanglement will refer to a particular tensor product of spaces, given a priori.) Let $\{|i\rangle\}$ and $\{|j\rangle\}$ be orthonormal basis of $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively, which we will refer to as canonical.

Isomorphism 5.1 *The following linear map*

$$\begin{aligned} \hat{} : \mathbb{C}^m \otimes \mathbb{C}^n &\to End(\mathbb{C}^n \to \mathbb{C}^m) \\ A &\mapsto \hat{A} \\ \sum_{ij} A_{ij}|i\rangle|j\rangle &\mapsto \sum_{ij} A_{ij}|i\rangle\langle j| \end{aligned}$$

where $i = 1, \ldots, m$ and $j = 1, \ldots, n$, is an isomorphism taking vectors A into $m \times n$ matrices $\hat{A}$. It is isometric in the sense that:

$$\forall A, B \in \mathbb{C}^m \otimes \mathbb{C}^n, \quad B^\dagger A = Tr(\hat{B}^\dagger \hat{A}) \tag{5.1}$$

Proof: This is trivial, but note that the definition of this isomorphism is basis dependent. □

Following a very convenient notation introduced by Sudarshan [57][58] we will often use a semicolon ';' to separate output indices (on the left) from input indices (on the right), together with the repeated indices summation convention. For instance the matrix $\hat{A} : \mathbb{C}^n \to \mathbb{C}^m$ will be denoted $A_{i;j}$, so that $w = \hat{A}v$ is simply written as $w_i = \hat{A}_{i;j}v_j$. Thus the 'hat' operation acts as follows:

$$\text{if } A \equiv A_{ij} \text{ then } \hat{A} \equiv \hat{A}_{i;j} \text{ with } \hat{A}_{i;j} = A_{ij} \tag{5.2}$$

Another useful interpretation of this operation is provided in [61], by considering the canonical maximally entangled state of $\mathbb{C}^n \otimes \mathbb{C}^n$, $|\beta\rangle = \sum |j\rangle|j\rangle$. Indeed we have:

$$\begin{aligned} A &= (\hat{A} \otimes \mathbb{I}_n)|\beta\rangle \\ \hat{A} &= (\mathbb{I}_m \otimes \langle\beta|)(A \otimes \mathbb{I}_n) \end{aligned} \tag{5.3}$$

We now use the previous isomorphism to relate elements of $M_{mn}(\mathbb{C})$ to linear maps from $M_n(\mathbb{C})$ to $M_m(\mathbb{C})$. This formalizes some of the key steps by Choi [13] and finds its origins in

the work of Jamiolkowski [32]. We highlight the isometric property of this bijection.

Isomorphism 5.2 *The following linear map:*

$$\begin{aligned} \widehat{\ } : \mathbb{C}^{mn} \otimes (\mathbb{C}^{mn})^\dagger &\longrightarrow End(M_n(\mathbb{C}) \to M_m(\mathbb{C})) \\ \$ &\longmapsto [\widehat{\$} : \rho \mapsto \widehat{\$}(\rho)] \\ \textit{such that } \ AB^\dagger &\longmapsto [\rho \mapsto \hat{A}\rho\hat{B}^\dagger] \quad \textit{i.e.} \\ \sum_{ijkl} A_{ij}B^*_{kl}|i\rangle|j\rangle\langle k|\langle l| &\longmapsto [\rho \mapsto \sum_{ijkl} A_{ij}B^*_{kl}|i\rangle\langle j|\rho|l\rangle\langle k|\ ] \end{aligned}$$

where $i, k = 1, \ldots, m$ and $j, l = 1, \ldots, n$, is an isomorphism. It is isometric in the sense that:

$$\forall\ \$, \mathscr{E} \in M_{mn}(\mathbb{C}), \quad Tr(\mathscr{E}^\dagger \$) = \sum_{jl} Tr\big(\widehat{\mathscr{E}}(E_{jl})^\dagger\, \widehat{\$}(E_{jl})\big), \tag{5.4}$$

where $\{E_{jl} = |j\rangle\langle l|\}$ is the canonical basis of $M_n(\mathbb{C})$.

Before we give a proof we shall reassert Sudarshan's notation in this case. Suppose $\$ = \sum \$_{ijkl}|i\rangle|j\rangle\langle k|\langle l|$ so that we can write $\$ \equiv \$_{ij;kl}$. We then have:

$$\begin{aligned} &\widehat{\$} \equiv \widehat{\$}_{ik;jl} \ \text{ with } \ \widehat{\$}_{ik;jl} = \$_{ij;kl} \\ &\text{so that } \ \widehat{\$} : \rho_{j;l} \mapsto \widehat{\$}(\rho)_{i;k} = \widehat{\$}_{ik;jl}\rho_{j;l} \end{aligned} \tag{5.5}$$

This notation views $End(M_n(\mathbb{C}) \to M_m(\mathbb{C}))$ as $m^2 \times n^2$ matrices, or as *superoperators*, thus admitting the usual Hilbert-Schmidt inner-product:

$$\mathrm{Tr}((\widehat{\mathscr{E}}^\dagger_{jl;ik})(\widehat{\$}_{i'k';j'l'})) \tag{5.6}$$

where $\widehat{\mathscr{E}}^\dagger_{jl;ik}$ is an $n^2 \times m^2$ matrix. The superoperator formalism simply consists of labelling a linear operator on matrices by a super-matrix, or more generally a linear map on tensors by a bigger tensor, and hence helps define operator norms. In fact it will turn out to be a cornerstone of the state-operator correspondence. It has had many applications in physics, amongst them the super-scattering or 'dollar' operator formalism introduced in Quantum Field Theory by Hawking [28], which, in contrast with the S-matrix formalism, allows non-unitary evolutions (hence our notation).

Proof *of Isomorphism 5.2.* Elements of $\mathbb{C}^{mn} \otimes (\mathbb{C}^{mn})^\dagger$ are all of the form $\sum_x A_x B_x^\dagger$, and thus by linearity the map $\widehat{\ }$ is fully determined by the above. The fact that it is an isomorphism is made obvious by Equation (5.5).

Now let $\mathscr{E} \equiv \mathscr{E}_{ij;kl}$ and $\$ = \$_{ij;kl}$. We now show that the notion of inner product given by (5.6) is precisely that of the RHS of Equation (5.4). Since $\widehat{\mathscr{E}}_{ik;jl} = \widehat{\mathscr{E}}(E_{jl})_{i;k}$ and $\widehat{\mathscr{E}}^\dagger_{jl;ik} = \widehat{\mathscr{E}}^*_{ik;jl}$, we

have

$$\begin{aligned}\mathrm{Tr}((\widehat{\not{\mathbb{C}}}^{\dagger}_{jl;ik})(\widehat{\$}_{i'k';j';l'})) &= \widehat{\not{\mathbb{C}}}^{\dagger}_{jl;ik}\widehat{\$}_{ik;jl} \\ &= \sum_{ikjl} \widehat{\not{\mathbb{C}}}(E_{jl})^{\dagger}_{k;i}\widehat{\$}(E_{jl})_{i;k} \\ &= \sum_{jl} \mathrm{Tr}(\widehat{\not{\mathbb{C}}}(E_{jl})^{\dagger}\,\widehat{\$}(E_{jl}))\end{aligned}$$

Finally notice that $\widehat{\not{\mathbb{C}}}^{\dagger}_{jl;ik} = \widehat{\not{\mathbb{C}}}^{*}_{ik;jl} = \not{\mathbb{C}}^{*}_{ij;kl}$, using (5.5). Thus (5.6) is also equal to the LHS of (5.4):

$$\begin{aligned}\mathrm{Tr}((\widehat{\not{\mathbb{C}}}^{\dagger}_{jl;ik})(\widehat{\$}_{i'k';j';l'})) &= \widehat{\not{\mathbb{C}}}^{\dagger}_{jl;ik}\widehat{\$}_{ik;jl} \\ &= \not{\mathbb{C}}^{*}_{ij;kl}\$_{ij;kl} = \not{\mathbb{C}}^{\dagger}_{kl;ij}\$_{ij;kl} \\ &= \mathrm{Tr}(\not{\mathbb{C}}^{\dagger}\$)\end{aligned}$$

□

In terms of the canonical maximally entangled state $|\beta\rangle$ of $\mathbb{C}^n \otimes \mathbb{C}^n$, using (5.3), we have that

$$\$ = (\widehat{\$} \otimes \mathbb{I}_n)(|\beta\rangle\langle\beta|) \tag{5.7}$$

Note that $|\beta\rangle\langle\beta| = \sum E_{jl} \otimes E_{jl}$, so we get

$$\$ = \sum_{jl} \widehat{\$}(E_{jl}) \otimes E_{jl} \tag{5.8}$$

This relation is quite handy when one seeks to visualize the isomorphism in terms of matrix manipulation. It is clear that the isomorphisms $\widehat{\ }$ and $\widehat{\ }$ are biased towards interpreting states in $\mathbb{C}^{mn} = \mathbb{C}^m \otimes \mathbb{C}^n$ as linear operations from states in the second subspace $\mathbb{C}^n$ into states in the first subspace $\mathbb{C}^m$. This will be made explicit in the forthcoming theorems. Without difficulty we could do the contrary and view states in $\mathbb{C}^{mn}$ as operations from states in $\mathbb{C}^m$ to states in $\mathbb{C}^n$:
For $A = \sum_{ij} A_{ij}|i\rangle|j\rangle \in \mathbb{C}^{mn}$, let $\check{A} = \sum_{ij} A_{ij}|j\rangle\langle i|$, i.e. $\check{A} \equiv \check{A}_{j;i} = A_{ij}$, so that $\check{A} = \hat{A}^t$. For $\$ = AB^{\dagger} \in M_{mn}(\mathbb{C})$ let $\breve{\$} : M_m(\mathbb{C}) \to M_n(\mathbb{C})$, $\rho \mapsto \sum_{ijkl} A_{ij}B^{*}_{kl}|j\rangle\langle i|\rho|k\rangle\langle l|$, which implies:

$$\breve{\$} \equiv \breve{\$}_{jl;ik} = \$_{ij;kl} = \widehat{\$}_{ik;jl}, \;\; \text{i.e.} \;\; \breve{\$}_{jl;ik} = \widehat{\$}^{t}_{jl;ik}. \tag{5.9}$$

In this case Equation (5.8) becomes:

$$\$ = \sum_{ik} E_{ik} \otimes \breve{\$}(E_{ik}) \tag{5.10}$$

Note that with the usual tensor product convention of taking the right-hand-side matrix as the one to be plugged into each component of the left-hand-side matrix, Equation (5.10) is simply written $\$ = (\breve{\$}(E_{ik}))_{ik}$, which is precisely Choi's formalism. Thus many view these two Isomorphisms as rearrangements of the coordinates of vectors or matrices. Although all would work equally well with $\breve{}$, from now on we shall keep to our initial version of the isomorphisms, taking the second subspace into the first.

5.1.2 Useful formulae

The following two lemmas are simple but useful results related to isomorphisms 1 and 2.

Lemma 5.1 *Let $A, B \in \mathbb{C}^m \otimes \mathbb{C}^n$, so that $AB^\dagger \in \mathbb{C}^{mn} \otimes (\mathbb{C}^{mn})^\dagger$, and let Tr_1 and Tr_2 denote the partial traces on $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively. Then we have:*

$$Tr_1(AB^\dagger) = (\hat{B}^\dagger \hat{A})^t \tag{5.11}$$

$$Tr_2(AB^\dagger) = \hat{A}\hat{B}^\dagger \tag{5.12}$$

Proof: let $A \equiv A_{ij}$ and $B \equiv B_{kl}$ with $i, k = 1, \ldots, m$ and $j, l = 1, \ldots, n$. $AB^\dagger = A_{ij}B^*_{kl}|i\rangle\langle k| \otimes |j\rangle\langle l|$. Thus taking Tr_1 sets $i = k$ and taking Tr_2 sets $j = l$:

$$\mathrm{Tr}_1(AB^\dagger)_{j;l} = A_{ij}B^*_{il} = \hat{B}^\dagger_{l;i}\hat{A}_{i;j} = (\hat{B}^\dagger\hat{A})^t{}_{j;l}$$

$$\mathrm{Tr}_2(AB^\dagger)_{i;k} = \hat{A}_{ij}B^*_{kj} = \hat{A}\hat{B}^\dagger{}_{i;k}$$

□

Lemma 5.2 *Suppose* $\widehat{\ }$ *is defined for* n *fixed and for all* d *such that it takes any element of* $\mathbb{C}^{dn} \otimes (\mathbb{C}^{dn})^\dagger$ *to a linear map from* $M_n(\mathbb{C})$ *to* $M_d(\mathbb{C})$*:*

$$\forall d, \quad \widehat{\ } : \mathbb{C}^{dn} \otimes (\mathbb{C}^{dn})^\dagger \longrightarrow End(M_n(\mathbb{C}) \to M_d(\mathbb{C})),$$

and let Tr_1 *denote the partial trace on the first* r*-dimensional subsystem of any system. We then have:*

$$\forall\ \$\ \in \mathbb{C}^{rmn} \otimes (\mathbb{C}^{rmn})^\dagger, \quad \widehat{Tr_1(\$)} = Tr_1 \circ \widehat{\$}$$

in other words Tr_1 *and* $\widehat{\ }$ *commute.*

Proof: In the following, $i, k = 1, \ldots, m$ and $j, l = 1, \ldots, n$ as usual, while $p, q = 1, \ldots, r$. Let $\$ \equiv \$_{pij;qkl} \in \mathbb{C}^{rmn} \otimes (\mathbb{C}^{rmn})^*$, and $\rho = \rho_{j;l} \in M_n(\mathbb{C})$.
Then $\widehat{\$}(\rho)_{pi;qk} = \widehat{\$}_{piqk;jl}\rho_{j;l} = \$_{pij;qkl}\rho_{j;l}$ is in $M_{rm}(\mathbb{C})$. Since Tr_1 sets $p = q$, $(\mathrm{Tr}_1 \circ \widehat{\$})(\rho)_{i;k} = \$_{pij;pkl}\rho_{j;l}$. On the other hand $\mathrm{Tr}_1(\$)_{ij;kl} = \$_{pij;pkl}$ so $\widehat{\mathrm{Tr}_1(\$)} \equiv \widehat{\mathrm{Tr}_1(\$)}_{ik;jl} = \$_{pij;pkl}$, thus $\widehat{\mathrm{Tr}_1(\$)}(\rho)_{i;k} = \$_{pij;pkl}\rho_{j;l}$. □

Next we give a novel and powerful formula relating linear operations $\widehat{\$}$ to trace outs of matrix multiplications involving $\$$.

Proposition 5.1 *Let* $\widehat{\$}$ *a linear map from* $M_n(\mathbb{C})$ *to* $M_m(\mathbb{C})$*,* σ*,* ρ *two elements of* $M_n(\mathbb{C})$*,* κ*,* τ *two elements of* $M_m(\mathbb{C})$*. Then we have:*

$$\kappa\widehat{\$}(\rho\sigma)\tau = Tr_2\big((\kappa \otimes \rho^t)\$(\tau \otimes \sigma^t)\big) \tag{5.13}$$

where Tr_2 *denotes the partial trace over the second system* $\mathbb{C}^n$ *in* $\mathbb{C}^m \otimes \mathbb{C}^n$*. This implies that for all* $\rho \in M_n(\mathbb{C})$ *and* $\kappa \in M_m(\mathbb{C})$*,*

$$Tr\big(\kappa\widehat{\$}(\rho)\big) = Tr\big((\kappa \otimes \rho^t)\$\big). \tag{5.14}$$

Proof: Since $(\kappa \otimes \rho^t)_{ij;kl} = \kappa_{ik}\rho^t_{jl}$, $(\tau \otimes \sigma^t)_{ij;kl} = \tau_{ik}\sigma^t_{jl}$, and tracing out $\mathbb{C}^n$ consists of setting $j = l$, we have

$$\begin{aligned}
(\kappa \otimes \rho^t)\$(\tau \otimes \sigma^t)_{ij;kl} &= \kappa_{ii'}\rho^t_{jj'}\$_{i'j';i''j''}\tau_{i''k}\sigma^t_{j''l} \\
\mathrm{Tr}_2\big((\kappa \otimes \rho^t)\$(\tau \otimes \sigma^t)\big)_{i;k} &= \kappa_{ii'}\rho_{j'l}\$_{i'j';i''j''}\tau_{i''k}\sigma_{lj''} \\
&= \kappa_{ii'}\$_{i'j';i''j''}\rho_{j'l}\sigma_{lj''}\tau_{i''k} \\
&= \kappa_{ii'}\widehat{\$}_{i'i'';j'j''}(\rho_{j'l}\sigma_{lj''})\tau_{i''k} \\
&= \kappa\widehat{\$}(\rho\sigma)\tau.
\end{aligned}$$

Equation (5.14) follows immediately by letting $\tau = \mathbb{I}_m$, $\sigma = \mathbb{I}_n$ and taking the total trace. □

From Equation (5.13) one can also derive the following interesting formula: $\forall \rho \in M_n(\mathbb{C})$,

$$\widehat{\$}\big((\rho^\dagger \rho)^t\big) = \mathrm{Tr}_2\big((\mathbb{I}_m \otimes \rho)\$(\mathbb{I}_m \otimes \rho^\dagger)\big) \tag{5.15}$$

We shall come back to Equation (5.15) in Subsection 5.3.2, with a more physical point of view. For now note that the equation is slightly more general than the one given in [61]p4, and that its equivalent form for $\breve{}$ is clearly seen to define a map from the first subspace into the second:

$$\breve{\$}\big((\rho^\dagger \rho)^t\big) = \mathrm{Tr}_1\big((\rho \otimes \mathbb{I}_n)\$(\rho^\dagger \otimes \mathbb{I}_n)\big).$$

Moreover the original Equation (5.14) will have a wide range of applications in the field of quantum information theory. This is because many of the mathematical problems raised by quantum cryptography require a careful optimization of the fidelities induced by a linear operator $\widehat{\$}$. By means of this formula such involved expressions can elegantly be brought to just the trace of the product of two matrices, as we shall see in Chapter 7.

5.1.3 The correspondence

We proceed to give the well-known three fundamental theorems about isomorphism 2.

Theorem 5.1 *The linear operation* $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *is* Hermitian-preserving *if and only if* $\$$ *belongs to* $Herm_{mn}(\mathbb{C})$.

Proof: [$\Rightarrow$] Suppose $\widehat{\$}$ Hermitian-preserving, then by Remark 5.1 so is $(\widehat{\$} \otimes \mathbb{I}_n)$. Now since $|\beta\rangle\langle\beta|$ is hermitian it must be the case that $(\widehat{\$} \otimes \mathbb{I}_n)(|\beta\rangle\langle\beta|) = \$$ is hermitian. We used Equation (5.7) for the last equality.
[$\Leftarrow$] Suppose $\$ Hermitian, so that $\$_{ij;kl} = \$^*_{kl;ij}$. Let $\rho_{jl} = \rho^*_{lj} \in \mathrm{Herm}_n(\mathbb{C})$. Using (5.5) we have

$$\begin{aligned}\widehat{\$}(\rho)_{i;k} &= \widehat{\$}_{ik;jl}\rho_{jl} = \$_{ij;kl}\rho_{jl} \\ &= \$^*_{kl;ij}\rho^*_{lj} = (\widehat{\$}_{ki;lj}\rho_{lj})^* \\ &= \widehat{\$}(\rho)^*_{k;i}\end{aligned}$$

so that $\widehat{\$}$ is Hermitian-preserving. □
This result first appeared in [50]. In terms of components, $\widehat{\$}$ is Hermitian-preserving if and only if $\$_{ij;kl} = \$^*_{kl;ij}$, or equivalently $\widehat{\$}_{ik;jl} = \widehat{\$}^*_{ki;lj}$.

Theorem 5.2 *The linear operation* $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *is* Positive-preserving *if and only if* $\$$ *belongs to* $Herm_{mn}(\mathbb{C})$ *and is such that for all separable state* ρ *in* $Herm^+_{mn}(\mathbb{C})$, $Tr(\$\rho) \geq 0$.

Proof: $\$$ is Hermitian by theorem 5.1 since $\widehat{\$}$ is Hermitian-preserving. Using Equation (5.14) in the following, with $\rho, \rho_1 \in \mathrm{Herm}_n^+(\mathbb{C})$ and $\sigma, \rho_2 \in \mathrm{Herm}_m^+(\mathbb{C})$, we have:

$$\begin{aligned}
&\widehat{\$} \text{ is Positive-preserving} \\
\Leftrightarrow &\forall \rho, \ \ \forall \sigma, \ \ \mathrm{Tr}\big(\sigma \widehat{\$}(\rho)\big) \geq 0 \\
\Leftrightarrow &\forall \rho, \ \ \forall \sigma, \ \ \mathrm{Tr}\big((\sigma \otimes \rho^t)\$\big) \geq 0 \\
\Leftrightarrow &\forall \rho_1, \ \ \forall \rho_2, \ \ \mathrm{Tr}\big(\$(\rho_1 \otimes \rho_2)\big) \geq 0 \\
\Leftrightarrow &\forall \rho \in \mathrm{Herm}_{mn}^+(\mathbb{C}) \ \text{ separable}, \ \ \mathrm{Tr}(\$\rho) \geq 0
\end{aligned}$$

□

This result is shown for instance in [31], in a different manner. We shall come back to its geometrical consequences in Section 5.3.

Theorem 5.3 *The linear operation* $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *is* Completely Positive-preserving *if and only if* $\$$ *belongs to* $Herm^+_{mn}(\mathbb{C})$.

Proof: [$\Rightarrow$] Suppose $\widehat{\$}$ Completely Positive-preserving. Since $|\beta\rangle\langle\beta|$ is positive it must be the case that $(\widehat{\$} \otimes \mathbb{I}_n)(|\beta\rangle\langle\beta|) = \$$ is positive. We used Equation (5.7) for the last equality.
[$\Leftarrow$] Suppose $\$$ positive. We want to show that for all r, $\widehat{\$} \otimes \mathbb{I}_r : M_{nr}(\mathbb{C}) \to M_{mr}(\mathbb{C})$ is Positive-preserving. Let $\mathbb{E} \in M_{(mr)(nr)}(\mathbb{C})$ be such that:

$$\widehat{\mathbb{E}} = \widehat{\$} \otimes \mathbb{I}_r.$$

Explicitly, with $s, t, u, v = 1, \ldots, r$, and $i, k = 1, \ldots, m$ and $j, l = 1, \ldots, n$ as usual,

$$\begin{aligned}(\widehat{\$} \otimes \mathbb{I}_r)_{(is)(kt);(ju)(lv)} &= \delta_{su}\delta_{tv}\widehat{\$}_{ik;jl} \\ &= \delta_{su}\delta_{tv}\$_{ij;kl} \\ &= \widehat{\mathbb{E}}_{(is)(kt);(ju)(lv)} \\ &= \mathbb{E}_{(is)(ju);(kt)(lv)}\end{aligned} \tag{5.16}$$

where we have used (5.5) to switch from $\widehat{\$}$ to $\$$ and $\widehat{\mathbb{E}}$ to $\mathbb{E}$. Let $V_{(kt)(lv)} \in \mathbb{C}^{(mr)(nr)}$. Using (5.16) and the fact that $\$ \in \mathrm{Herm}^+_{mn}(\mathbb{C})$, we get

$$\begin{aligned}V^\dagger \mathbb{E} V &= V^*_{(is)(ju)}\mathbb{E}_{(is)(ju);(kt)(lv)}V_{(kt)(lv)} \\ &= V^*_{isjs}\$_{ij;kl}V_{ktlt} \geq 0,\end{aligned}$$

hence $\mathbb{E} \in \mathrm{Herm}^+_{(mr)(nr)}(\mathbb{C})$. Then, by theorem 5.2, $\widehat{\$} \otimes \mathbb{I}_r$ is Positive-preserving if for all $\rho_1 \in \mathrm{Herm}^+_{rm}(\mathbb{C})$ and $\rho_2 \in \mathrm{Herm}^+_{rn}(\mathbb{C})$, $\mathrm{Tr}(\mathbb{E}(\rho_1 \otimes \rho_2)) \geq 0$. This follows directly since $\rho_1 \otimes \rho_2$ and $\mathbb{E}$ are positive. □

This result first appeared in [13] with a different proof. The (possibly non-normalized) states of a mn-dimensional quantum system, or elements of $\mathrm{Herm}^+_{mn}(\mathbb{C})$, are thus in one-to-one correspondence with the (possibly non Trace-preserving) quantum operations, or Completely Positive-preserving maps, taking an n-dimensional system into an m-dimensional system. We claim that virtually all of the important, well-established results about quantum operations are in direct correspondence with those regarding quantum states, through the use of theorem 5.3. In [13][39][58], the operator sum representation for Completely Positive-preserving maps is derived in the proof of theorem 5.3, but in our approach we will think of it as stemming directly from the properties of quantum states.

5.2 Properties of quantum states and quantum operations

5.2.1 Properties rediscovered via the correspondence

Property 5.1 (Decomposition, degree of freedom.) *A matrix ρ is in $Herm_d^+(\mathbb{C})$ if and only if it can be written as*

$$\rho = \sum_x A_x A_x^\dagger$$

where each A_x is a d-dimensional vector. Two decompositions $\{A_x\}$ and $\{B_y\}$ correspond to the same state ρ if and only if there exists an isometric matrix U (i.e. $U^\dagger U = \mathbb{I}$) such that $A_x = \sum U_{xy} B_y$. There is a decomposition $\{A_x\}$ with $rank(\rho) \leq d$ non-zero elements and such that $A_{x'}^\dagger A_x \propto \delta_{xx'}$.

Corollary 5.1 (Operator sum representation.) *A linear map $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ is Completely Positive-preserving if and only if it can be written as*

$$\widehat{\$} : \ \rho \mapsto \sum_x \hat{A}_x \rho \hat{A}_x^\dagger$$

where each $\hat{A}_x$ is an $m \times n$ matrix. Two decompositions $\{\hat{A}_x\}$ and $\{\hat{B}_y\}$ correspond to the same $\widehat{\$}$ if and only if there exists an isometric matrix U (i.e. $U^\dagger U = \mathbb{I}$) such that $\hat{A}_x = \sum U_{xy} \hat{B}_y$. There is a decomposition $\{\hat{A}_x\}$ with $r \leq mn$ elements and such that $Tr(\hat{A}_{x'}^\dagger \hat{A}_x) \propto \delta_{xx'}$. r will be referred to as the Choi rank of $\widehat{\$}$, as this is the decomposition having the least number of elements.

Proof *of Property 5.1.* This is the spectral decomposition theorem for positive matrices, together with the unitary degree of freedom theorem by Hughston, Josza and Wooters [48]p103. □

Proof: *of Corollary 5.1.* Consider $\widehat{\$}$ a Completely Positive-preserving linear operator. By theorem 5.3, $\$ is positive, and so Property 5.1 provides decompositions upon that state. One may translate back these decompositions in terms of quantum operations using Isomorphism 2: this yields nothing but Corollary 5.1. □

Notice that the Choi rank of $\widehat{\$}$ is equal to $rank(\$)$.

Property 5.2 (Purification.) *A matrix ρ is in $Herm_d^+(\mathbb{C})$ if and only if it can be written as*

$$\rho = Tr_1(\rho_{pure}) \quad with \quad \rho_{pure} = VV^\dagger$$

where V is an rd-dimensional vector and Tr_1 traces out the first r-dimensional subsystem (r can be chosen equal to $rank(\rho) \leq d$).

Corollary 5.2 (Factorizable then trace representation.) *A linear map* $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *is Completely Positive-preserving if and only if it can be written as*

$$\widehat{\$} : \rho \mapsto Tr_1(\widehat{\$}_{pure}(\rho)) \quad with \quad \widehat{\$}_{pure} : \rho \mapsto \hat{V}\rho\hat{V}^\dagger$$

where $\hat{V}$ *is an* $rm \times n$ *matrix and* Tr_1 *traces out the first* r*-dimensional subsystem (*r *may be chosen equal to* $rank(\widehat{\$}) \leq mn$*). Moreover if* $\widehat{\$}$ *decomposes as* $\{\hat{A}_x\}$ *we have:*

$$\hat{V}^\dagger\hat{V} = \sum_x \hat{A}_x^\dagger \hat{A}_x \tag{5.17}$$

Proof *of Property 5.2.* [$\Rightarrow$] Suppose ρ decomposes as $\{A_x\}$ and let $V = \sum |x\rangle A_x$, with $\{|x\rangle\}$ an orthonormal basis of an ancilla system.

$$\rho_{pure} = VV^\dagger = \sum_{xy} |x\rangle\langle y| \otimes A_x A_y^\dagger$$
$$\mathrm{Tr}_1(\rho_{pure}) = \sum_{xy} \langle y|x\rangle A_x A_y^\dagger = \sum_x A_x A_x^\dagger = \rho$$

If $\{A_x\}$ is a spectral decomposition of ρ it counts $rank(\rho)$ elements, and thus r can be chosen to equal $rank(\rho)$.

$$[\Leftarrow] \qquad \langle\psi|\rho|\psi\rangle = \sum_i \langle i|\langle\psi|VV^\dagger|i\rangle|\psi\rangle \geq 0 \quad \text{since}$$
$$\forall i \quad \langle i|\langle\psi|VV^\dagger|i\rangle|\psi\rangle \geq 0$$

□

The second corollary is not traditionally thought of as a 'quantum operation equivalent' of quantum state purification. We now explicitly show how the result is again trivially obtained from Property 5.2, by virtue of Theorem 5.3.

Proof *of Corollary 5.2.* Consider $\widehat{\$}$ a Completely Positive-preserving linear operator. By Theorem 5.3, $\$ is positive, and so Property 5.2 gives $\$ = \mathrm{Tr}_1(\$_{pure})$, $\$_{pure} = VV^\dagger$, where the ancilla system can be chosen to be of dimension $r = rank(\$)$. As a consequence we can use Lemma 5.2 to retrieve $\widehat{\$} = \mathrm{Tr}_1(\widehat{\$}_{pure})$, $\widehat{\$}_{pure} : \rho \mapsto \hat{V}\rho\hat{V}^\dagger$.

Moreover, denote by $\mathrm{Tr}_{1'}$ the partial trace over the m-dimensional system. For $V = V_{xij}|x\rangle|i\rangle|j\rangle$, let $\hat{V} \equiv V_{xij}|x\rangle|i\rangle\langle j|$ the corresponding $rm \times n$ matrix. Since $\mathrm{Tr}_1(\rho_{pure}) = \rho$ with $\rho_{pure} = VV^\dagger$, and $\rho = \sum_x A_x A_x^\dagger$, we get

$$(\mathrm{Tr}_{1'} \circ \mathrm{Tr}_1)(VV^\dagger) = \sum_x \mathrm{Tr}_{1'}(A_x A_x^\dagger) \quad \text{implying}$$
$$\hat{V}^\dagger\hat{V} = \sum_x \hat{A}_x^\dagger \hat{A}_x \quad \text{by Equation (5.11)}$$

□

Notice that whenever $\widehat{\$}$ is Trace-preserving, then Equation (5.17) reads $\hat{V}^\dagger \hat{V} = \mathbb{I}_n$, so that $\hat{V}$ is isometric. Thus we have derived as a simple consequence of properties of state purification that any Trace-preserving quantum operation can arise as the trace-out of an isometric operation.

Property 5.3 (Schmidt decomposition.) *Consider $\rho = VV^\dagger$ a non-normalized pure state in $Herm^+_{mn}(\mathbb{C})$ with $V = \sum V_{ij}|i\rangle|j\rangle$ in the canonical basis of $\mathbb{C}^m \otimes \mathbb{C}^n$. Then there exists some positive reals $\{\lambda_i\}$ and some orthogonal basis $\{|\psi_i\rangle\}$ and $\{|\phi_i\rangle\}$ of $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively, such that*

$$V = \sum_{i=1}^{r} \lambda_i |\psi_i\rangle|\phi_i\rangle,$$

with $r \leq m$ and $r \leq n$. Moreover:

$$Tr_1(\rho) = \sum_{i=1}^{r} \lambda_i^2 |\phi_i\rangle\langle\phi_i| \quad (n \times n \text{ positive})$$
$$Tr_2(\rho) = \sum_{i=1}^{r} \lambda_i^2 |\psi_i\rangle\langle\psi_i| \quad (m \times m \text{ positive})$$

Corollary 5.3 (Polar decomposition.) *Consider $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$, $\rho \mapsto \hat{V}\rho\hat{V}^\dagger$ a factorizable Completely Positive-preserving linear map, with $\hat{V} = \sum V_{ij}|i\rangle\langle i|$. Then there exists some positive reals $\{\lambda_i\}$ and some orthogonal basis of $\mathbb{C}^m$ and $(\mathbb{C}^n)^\dagger$, namely $\{|\psi_i\rangle\}$ and $\{\langle\phi_i^*|\}$, such that*

$$\hat{V} = \sum_{i=1}^{r} \lambda_i |\psi_i\rangle\langle\phi_i^*|$$

with $r \leq m$ and $r \leq n$. In other words:

$$\hat{V} = UJ = KU \quad \text{with}$$
$$J = \sqrt{\hat{V}^\dagger\hat{V}} = \sum_{i=1}^{r} \lambda_i |\phi_i^*\rangle\langle\phi_i^*| \quad (n \times n \text{ positive})$$
$$K = \sqrt{\hat{V}\hat{V}^\dagger} = \sum_{i=1}^{r} \lambda_i |\psi_i\rangle\langle\psi_i| \quad (m \times m \text{ positive})$$
$$U = \sum_{i=1}^{n} |\psi_i\rangle\langle\phi_i^*| \quad (m \times n \text{ isometric, i.e. } U^\dagger U = \mathbb{I}_n)$$

Proof *of Property 5.3.* Let $\rho = VV^\dagger$, $V = \sum V_{ij}|i\rangle|j\rangle$, and Tr_2 the partial trace on the last n-dimensional system. Since $\rho^A = \mathrm{Tr}_2(\rho)$ is in $\mathrm{Herm}^+_m(\mathbb{C})$, we can write

$$\rho^A = \sum_{i=1}^{r} \lambda_i^2 |\psi_i\rangle\langle\psi_i|$$

where $\{\lambda_i\}$ are strictly positive reals, $r \leq m$, and $\{|\psi_i\rangle\}$ is an orthonormal family of vectors which we may complete into an orthonormal basis of $\mathbb{C}^m$. By expressing the first subspace of

V in this basis we can of course write:

$$V = \sum_{i=1}^{r} |\psi_i\rangle|\tilde{\phi}_i\rangle \quad \text{with} \quad |\tilde{\phi}_i\rangle = (\langle\psi_i| \otimes \mathbb{I}_n)V$$

We have:

$$\begin{aligned}
\langle\tilde{\phi}_i|\tilde{\phi}_j\rangle &= \mathrm{Tr}(|\tilde{\phi}_j\rangle\langle\tilde{\phi}_i|) \\
&= \mathrm{Tr}((\langle\psi_i| \otimes \mathbb{I})VV^\dagger(|\psi_j\rangle \otimes \mathbb{I})) \\
&= \mathrm{Tr}((|\psi_j\rangle\langle\psi_i| \otimes \mathbb{I})VV^\dagger) \\
&= \mathrm{Tr}(|\psi_j\rangle\langle\psi_i|\rho^A) \\
&= \lambda_i^2\delta_{ij}
\end{aligned}$$

Thus $\{|\phi_i\rangle = |\tilde{\phi}_i\rangle/\lambda_i\}$ is an orthonormal family of vectors in $\mathbb{C}^n$, which we may again complete into an orthonormal basis.
We now have $V = \sum \lambda_i|\psi_i\rangle|\phi_i\rangle$, from which it is straightforward to verify that

$$\mathrm{Tr}_1(\rho) = \sum_{i=1}^{r} \lambda_i^2|\phi_i\rangle\langle\phi_i|$$

□

The well-known connection between the Schmidt decomposition and the polar decomposition (itself trivially equivalent to the singular value decomposition) is now shown to arise naturally using the state-operator correspondence.

Proof *of Corollary 5.3.* Consider $\widehat{\$} : \rho \mapsto \hat{V}\rho\hat{V}^\dagger$. Using Isomorphism 2 the corresponding state in $\mathrm{Herm}^+_{mn}(\mathbb{C})$ is $\rho = VV^\dagger$. Applying the Schmidt decomposition theorem yields

$$\begin{aligned}
V &= \sum_{i=1}^{r} \lambda_i|\psi_i\rangle|\phi_i\rangle \quad \text{and thus} \\
\hat{V} &= \sum_{i=1}^{r} \lambda_i|\psi_i\rangle\langle\phi_i^*|
\end{aligned}$$

with $\{|\psi_i\rangle\}$ and $\{\langle\phi_i^*| = \langle\phi_i|^*\}$ some orthogonal basis of $\mathbb{C}^m$ and $(\mathbb{C}^n)^\dagger$ respectively. Now if we call U the $m \times n$ isometric (i.e. $U^\dagger U = \mathbb{I}_n$) matrix $\sum_{i=1}^n |\psi_i\rangle\langle\phi_i^*|$, we have that $\hat{V} = UJ = KU$, with

$$\begin{aligned}
K &= \sum_{i=1}^{r} \lambda_i|\psi_i\rangle\langle\psi_i| = \sqrt{\mathrm{Tr}_2(VV^\dagger)} = \sqrt{\hat{V}\hat{V}^\dagger} \\
J &= \sum_{i=1}^{r} \lambda_i|\phi_i^*\rangle\langle\phi_i^*| = \left(\sqrt{\mathrm{Tr}_1(VV^\dagger)}\right)^t = \sqrt{\hat{V}^\dagger\hat{V}}.
\end{aligned}$$

In the above K is $m \times m$ whilst J is $n \times n$, and the last equality of each line was derived from Equations (5.12) and (5.11). □

Thus it seems that all the standard results about quantum operations are in correspondence with those concerning quantum states. Of course although we derived the properties of operators from those of states, we could equally have done the opposite. Next we seek to apply the same principle to derive new results, as we consider properties of states and operations which do not yet have any equivalent in terms of, respectively, operations and states.

5.2.2 Properties discovered via the correspondence

We first derive a factorizability condition on quantum operations by making use of the well-known property:

Property 5.4 (Purity condition.) *Let ρ a matrix in $Herm_d^+(\mathbb{C})$. Then ρ is non-normalized pure, i.e. of the form $\rho = VV^\dagger$, if and only if*

$$Tr(\rho)^2 - Tr(\rho^2) = 0$$

Corollary 5.4 (Factorizability condition.) *Let $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ a Completely Positive-preserving linear operator. Then $\widehat{\$}$ is of the form $\widehat{\$} : \rho \mapsto \hat{V}\rho\hat{V}^\dagger$, i.e. it is factorizable, if and only if*

$$\Big(Tr(\widehat{\$}(\mathbb{I}_n))\Big)^2 - \sum_{jl} Tr\big(\widehat{\$}(E_{jl})^\dagger \widehat{\$}(E_{jl})\big) = 0 \tag{5.18}$$

or equivalently in terms of indices

$$(\widehat{\$}_{ii;jj})^2 - \widehat{\$}^*_{ik;jl}\widehat{\$}_{ik;jl} = 0.$$

Proof *of Property 5.4.*[⇒] is obvious since ρ pure has only got one non-zero eigenvalue.
[⇐] Suppose ρ has eigenvalues $\{\lambda_i\}$. The purity condition amounts to

$$(\sum_i \lambda_i)^2 = \sum_i \lambda_i^2 \quad \text{implying} \quad \sum_{i<j} \lambda_i\lambda_j = 0.$$

For the last relation to hold, since the λ_i's are positive there can be at most one value of i such that $\lambda_i \neq 0$. □

Proof *of Corollary 5.4.* $\widehat{\$}$ is factorizable is equivalent to $\$$ being pure, thus by Property 5.4 to

$$\mathrm{Tr}(\mathbb{I}_{mn}\$)^2 - \mathrm{Tr}(\$^2) = 0.$$

Using $\$^\dagger = \$$ Equation (5.18) is a direct application of Equation (5.4) upon this last equation, as can be seen from

$$\mathbb{I}_{mn} = \sum_{kl} |kl\rangle\langle kl|$$
$$\text{so that } \widehat{\mathbb{I}_{mn}} : \rho \mapsto \sum_{kl} E_{kl}\rho E_{kl}^\dagger$$
$$\text{and } \widehat{\mathbb{I}_{mn}} : E_{jl} \mapsto \delta_{jl}\,\mathbb{I}_m$$

□

Next we give two new vector decompositions which stem from classical results on matrix decomposition.

Property 5.5 (One-sided triangular decomposition.) *Let $\rho = VV^\dagger$ a non-normalized pure state in $Herm^+_{mn}(\mathbb{C})$, with $V = \sum V_{ij}|i\rangle|j\rangle$ in the canonical basis, and suppose $m \geq n$. Then there exists some orthogonal basis of $\mathbb{C}^m$, namely $\{|\psi_i\rangle\}$, such that*

$$V = \sum_{i \leq j}^{j=n} \mu_{ij}|\psi_i\rangle|j\rangle$$

Proof: According to the QR decomposition theorem [30] the $m \times n$ matrix $\hat{V}$ can be decomposed as $\hat{V} = QR$, where Q is $m \times n$ and verifies $Q^\dagger Q = \mathbb{I}_n$ whilst R is $n \times n$ upper triangular. Thus we have:

$$\hat{V} = Q \sum_{i \leq j}^{j=n} \mu_{ij}|i\rangle\langle j|$$
$$\hat{V} = \sum_{i \leq j}^{j=n} \mu_{ij}|\psi_i\rangle\langle j|$$
$$V = \sum_{i \leq j}^{j=n} \mu_{ij}|\psi_i\rangle|j\rangle$$

Since Q is isometric, the $\{|\psi_i\rangle = Q|i\rangle\}$ are orthonormal and can be extended to form a basis of $\mathbb{C}^m$. □

On the one hand Property 5.5 is less powerful than the Schmidt decomposition, in the sense that it yields 'upper triangular' coefficients μ_{ij} instead of the neat diagonal form $V = \sum_i^r \lambda_i|\psi_i\rangle|\phi_i\rangle$. On the other hand however Property 5.5 requires a change of basis for the first subsystem only. Such a distinction is perfectly analogous to what separates the polar decomposition (or more expressively its singular value decomposition corollary) from the QR decomposition when speaking about matrices. Just like the QR decomposition the one-sided state triangularization is easily computed.
Schur's triangularization theorem can also be given a quantum state equivalent, as we now explain. This seems of lesser interest however, since the procedure involves two changes of basis, one for each subsystem - a case which seems better covered by the Schmidt decomposition (though here the two basis are simply related).

Property 5.6 (Two-sided triangular decomposition.) *Let $\rho = VV^\dagger$ a non-normalized pure state in $Herm^+_{m^2}(\mathbb{C})$, with $V = \sum V_{ij}|i\rangle|j\rangle$ in the canonical basis. Then there exists some orthogonal basis of $\mathbb{C}^m$, namely $\{|\psi_i\rangle\}$ such that*

$$V = \sum_{i \leq j}^{j=m} \mu_{ij}|\psi_i\rangle|\psi_j^*\rangle$$

*where * denotes complex conjugation of the coordinates of a vector in the canonical basis. Moreover the set $\{\mu_{ii}\}$ is the set of the Schmidt coefficients $\{\lambda_i\}$ of V (as defined in Property 5.3).*

Proof: According to Schur's decomposition theorem [30] the matrix $\hat{V}$ can be decomposed as $\hat{V} = UTU^\dagger$, where U is unitary and T is upper triangular and has the singular values of $\hat{V}$ in the diagonal (i.e. precisely the λ_i's of the polar decomposition and of the Schmidt decomposition). And so we have:

$$\hat{V} = U \sum_{i \leq j}^{j=m} \mu_{ij} |i\rangle\langle j| U^\dagger$$

$$\hat{V} = \sum_{i \leq j}^{j=m} \mu_{ij} |\psi_i\rangle\langle\psi_j|$$

$$V = \sum_{i \leq j}^{j=m} \mu_{ij} |\psi_i\rangle|\psi_j^*\rangle$$

Since U is unitary the $\{|\psi_i\rangle = U|i\rangle\}$ are orthonormal and can be extended to form a complete basis of $\mathbb{C}^m$. □

5.2.3 Trace-preserving quantum operations

The results of Subsection 5.2.1, although extremely useful in quantum theory (quantum information theory in particular), are in fact general results on positive matrices and Completely Positive-preserving linear maps. The same is true of Subsection 5.2.2, and this is the reason why we have barely mentioned the unit trace condition on density matrices so far. Yet in quantum theory the states must have trace one (unless we start to consider the trace as encoding some overall probability of occurrence), and quantum operation must be Trace-preserving (so that they may always occur). We now give an account of the main known results related to these restrictions, augmented with some results stemming from the state-operator correspondence.

Definition 5.4 *A linear map $\Omega : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ is* Trace-preserving *if and only if for all ρ in $M_n(\mathbb{C})$, $Tr(\Omega(\rho)) = Tr(\rho)$.*

Definition 5.5 *The state $(1/d)\mathbb{I}_d \in Herm_d^+(\mathbb{C})$ is called the* maximally mixed *state of $\mathbb{C}^d$. Moreover we say that $\$ \in Herm_{mn}^+(\mathbb{C})$ is a* maximally entangled *state of $\mathbb{C}^m \otimes \mathbb{C}^n$ if and only*

if $\$$ *is pure and verifies either of*

$$(n \leq m) \quad Tr_1(\$) = \mathbb{I}_n$$
$$(m \leq n) \quad Tr_2(\$) = \mathbb{I}_m$$

depending on the integers m *and* n *(if* $m = n$ *the two conditions are equivalent).*

Lemma 5.3 (Trace-preserving linear maps.) *A Completely positive-preserving linear map* $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *with decomposition* $\{\hat{A}_x\}$ *is* Trace-preserving *if and only if one of the following six equivalent conditions is satisfied:*

$$(i)\ \sum \hat{A}_x^\dagger \hat{A}_x = \mathbb{I}_n, \qquad (ii)\ \widehat{\$}_{kk;jl} = \delta_{jl},$$

In terms of $\breve{\$}$ *this is*

$$(iii)\ \sum \check{A}_x \check{A}_x^\dagger = \mathbb{I}_n, \qquad (iv)\ \breve{\$}(\mathbb{I}_m) = \mathbb{I}_n,$$

In terms of the state $\$$ *this is*

$$(v)\ Tr_1(\$) = \mathbb{I}_n, \qquad (vi)\ \$_{kj;kl} = \delta_{jl}.$$

Proof: We have $\widehat{\$}(\rho) = \sum \hat{A}_x \rho \hat{A}_x^\dagger$ or using components $\widehat{\$}(\rho)_{i;k} = \widehat{\$}_{ik;jl}\rho_{jl}$, so that $\mathrm{Tr}\big(\widehat{\$}(\rho)\big) = \mathrm{Tr}\big(\sum \hat{A}_x^\dagger \hat{A}_x \rho\big) \equiv \widehat{\$}_{kk;jl}\rho_{jl}$. Thus (i) and (ii) follow immediately. Using that $\check{A} = \hat{A}^t$ and $\breve{\$}_{jl;kk} = \breve{\$}(\mathbb{I}_m)_{j;l} = \widehat{\$}_{jl;kk}$ from (5.9), we get (iii) and (iv). (v) and (vi) follow from (i) and (ii) using (5.11) and (5.5) respectively. □

Note that these conditions imply, but are not equivalent to, $(1/n)\$$ having unit trace. This is because '$\$$ has unit trace' reads:

$$\mathrm{Tr}(\$) = \mathrm{Tr}\big(\widehat{\$}(\mathbb{I}_n)\big) = \mathrm{Tr}\big(\breve{\$}(\mathbb{I}_m)\big) = 1$$

$$\text{or } \$_{kl;kl} = \widehat{\$}_{kk;ll} = \breve{\$}_{ll;kk} = 1.$$

Thus we have shown that Trace-preserving quantum operations $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ go hand in hand with unit trace states $(1/n)\$ \in \mathrm{Herm}^+_{mn}(\mathbb{C})$ whose partial trace on the first subsystem yields the maximally mixed state: $\mathrm{Tr}_1\big((1/n)\$\big) = (1/n)\mathbb{I}_n$. We immediately obtain the following, which is a generalization of a result in [39] and [61]:

Lemma 5.4 (Unitary maps.) *Let* $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *a Completely Positive-preserving map. Then* $\widehat{\$}$ *is* isometric *(i.e it can be written as* $\widehat{\$} : \rho \mapsto \hat{U}\rho\hat{U}^\dagger$ *with* $\hat{U}^\dagger\hat{U} = \mathbb{I}_n$*) if and only if* $n \leq m$ *and the corresponding state* $\$$ *is* maximally entangled *(i.e. pure with* $Tr_1(\$) = \mathbb{I}_n$*). Equivalently, in terms of indices,* $\widehat{\$}$ *must verify* $\widehat{\$}_{kk;jl} = \delta_{jl}$ *and*

$$\sum_{jl} Tr\big(\widehat{\$}(E_{jl})^\dagger \widehat{\$}(E_{jl})\big) = n^2$$

Remark 5.2 (Bistochastic maps.) $\widehat{\$}$ *is* bistochastic*, i.e. it is Trace-preserving and satisfies* $\widehat{\$}(\mathbb{I}_n) = \mathbb{I}_m$*, if and only if the state* $\$$ *satisfies* $Tr_1(\$) = \mathbb{I}_n$ *and* $Tr_2(\$) = \mathbb{I}_m$*. Thus bistochastic maps cannot be factorizable whenever* $m \neq n$*.*

Proof: The Lemma follows immediately from Lemma 5.3 and Corollary 5.4. The remark follows from Lemma 5.3 and the fact that $\mathrm{Tr}_2(\$) = \widehat{\$}(\mathbb{I}_m)$. □

The set of states $\$ \in \mathrm{Herm}^+_{mn}(\mathbb{C})$ satisfying $\mathrm{Tr}_1(\$) = \mathbb{I}_n$ is convex, hence its extremal points correspond to extremal Trace-preserving quantum operations. Recall that the extremal elements of a convex set S are those which cannot be written as sums of two distinct elements of S. Extremal elements are important since they generate S, and so we now restate Choi's well-known theorem about extremal Trace-preserving maps (without reproducing the proof).

Theorem 5.4 (Extremal Trace-preserving.) *Let $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ a Trace-preserving Completely Positive-preserving linear map with decomposition $\{\hat{A}_x\}$ and Choi rank r (i.e. $r = rank(\$)$). Then $\widehat{\$}$ is extremal in the set of Trace-preserving Completely Positive-preserving maps if and only if one of the following three equivalent conditions is satisfied:*
(i) the span of the set $\{\hat{A}_x^\dagger \hat{A}_y\}$ in $M_n(\mathbb{C})$ is r^2-dimensional;
(ii) the span of the set $\{\check{A}_x \check{A}_y^\dagger\}$ in $M_n(\mathbb{C})$ is r^2-dimensional;
(iii) the span of the set $\{Tr_1(A_x A_y^\dagger)\}$ in $M_n(\mathbb{C})$ is r^2-dimensional.

Notice that this is a slightly different formulation from the one given in [13], where the $\{\hat{A}_x^\dagger \hat{A}_y\}$ have to form a linearly independent set. This implies that the $\{\hat{A}_x\}$ are automatically linearly independent themselves, and hence there must be r of them. Since different decompositions can give the same operation, we thought it better to express the extremality conditions in terms of any decomposition, and not just a minimal one.
Proof: We just prove the equivalence with Choi's formulation. Let $\{\hat{V}_\alpha\}$ be a minimal decomposition of $\widehat{\$}$. Then $\mathrm{Span}(\{A_x\}) = \mathrm{Span}(\{V_\alpha\})$ since both are equal to the support (the image space) of $\$ (see Corollary 5.1); and trivially $\mathrm{Span}(\{\hat{A}_x\}) = \mathrm{Span}(\{\hat{V}_\beta\})$ implies $\mathrm{Span}(\{\hat{A}_x^\dagger \hat{A}_y\}) = \mathrm{Span}(\{\hat{V}_\alpha^\dagger \hat{V}_\beta\})$. □

Remark 5.3 *An extremal map $\widehat{\$}$ has Choi rank $r \leq n$ since it must satisfy $r^2 \leq n^2$, but this condition is not sufficient.*

Proof: Suppose $U_1 \neq U_2$ unitary and $\widehat{\$} : \rho \mapsto (1/2)U_1 \rho U_1^\dagger + (1/2)U_2 \rho U_2^\dagger$. Clearly $U_1^\dagger U_1 = U_2^\dagger U_2 = \mathbb{I}_n$, and thus this Trace-preserving Completely Positive-preserving map cannot be extremal Trace-preserving. Yet it has Choi rank 2 regardless of a choice for n. □
By pushing the consequences of Choi's theorem further we obtain the following original criteria for extremal Trace-preserving linear maps:

Proposition 5.2 (Extremal Trace-preserving.) *Let $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ a Trace-preserving Completely Positive-preserving linear map of Choi rank r (i.e. $r = rank(\$)$) with $\$ its corresponding state, and $\widehat{\$}^\dagger : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ its adjoint map (i.e. $\widehat{\$}^\dagger \equiv \widehat{\$}^\dagger_{jl;ik}$). Then $\widehat{\$}$ is extremal if and only if one of the following equivalent conditions is satisfied:*

(i) The Choi rank of $\widehat{\$}^\dagger \circ \widehat{\$}$ *is equal to* r^2*,*
(ii) $\$$ *is such that the state in* $Herm^+_{n^2}(\mathbb{C})$ *defined by*

$$\mathbb{C}_{jj';ll'} = \$^*_{ij;kl}\$_{ij';kl'}$$

has rank r^2*.*

Proof: If $\widehat{\$}$ has operator sum decomposition $\{\hat{A}_y\}$ i.e. $\widehat{\$}(\rho) = \sum_y \hat{A}_y \rho \hat{A}_y^\dagger$, then we get that $\widehat{\$}^\dagger$ has decomposition $\{\hat{A}_x^\dagger\}$ i.e. $\widehat{\$}^\dagger(\sigma) = \sum_x \hat{A}_x^\dagger \rho \hat{A}_x$. This can be seen using $\widehat{\$}^\dagger_{jl;ik} \equiv \widehat{\$}^*_{ik;jl}$ for example. Thus $\widehat{\$}^\dagger \circ \widehat{\$}$ has decomposition $\{\hat{A}_x^\dagger \hat{A}_y\}$, and (i), using Corollary 5.1, is equivalent to (i) in Theorem 5.4.
Next we restate (i) using indices and Equation (5.5).

$$\begin{aligned}(\widehat{\$}^\dagger \circ \widehat{\$})_{jl;j'l'} &= \widehat{\$}^\dagger_{jl;ik}\widehat{\$}_{ik;j'l'} \\ &= \widehat{\$}^*_{ik;jl}\widehat{\$}_{ik;j'l'} \\ &= \$^*_{ij;kl}\$_{ij';kl'} \\ &\equiv \widehat{\mathbb{C}}_{jl;j'l'} = \mathbb{C}_{jj';ll'}.\end{aligned}$$

Since $\widehat{\$}^\dagger \circ \widehat{\$}$ is a Completely Positive-preserving map from $M_n(\mathbb{C})$ to $M_n(\mathbb{C})$, $\mathbb{C}$ is in $\mathrm{Herm}^+_{n^2}(\mathbb{C})$ by Theorem 5.3. We see that (i) is equivalent to (ii). □
The relation between condition (i) and (ii) suggests that the composition law on quantum operations could yield, through Isomorphism 5.2, an interesting structure upon states. We pursue this idea in the following section.

5.3 Induced geometrical structure

The beginning of this section is maybe aimed at a mathematically-minded reader. We investigate simple algebraic and geometric properties stemming from the operator state correspondence. These yield a nice group theoretic description of totally entangled states of a bipartite system (Proposition 5.4), and a description of Positive-preserving maps as dual to separable states (Theorem 5.2 restated). Proposition 5.6 however unravels a possible physical interpretation of the correspondence.

5.3.1 Composition laws

We make use of some elementary facts about operators or positive matrices to define new composition laws on the spaces of operators or positive matrices.
First, the set of Completely Positive-preserving linear maps from $M_n(\mathbb{C})$ into itself is stable under composition. This induces the following semi-group structure for states (recall that semi-group elements do not need to have an inverse):

Proposition 5.3 *If* $ *and* €′ *are in* $Herm^+_{n^2}(\mathbb{C})$*, then so is*

$$\$ \diamond €′ \equiv (\$ \diamond €′)_{ij;kl} = \$_{ii';kk'} €′_{i'j;k'l}, \tag{5.19}$$

where all the indices run from 1 *to* n*.* $(Herm^+_{n^2}(\mathbb{C}), \diamond)$ *is a semi-group with identity element the canonical maximally entangled state* $|\beta\rangle\langle\beta| \equiv \delta_{ij}\delta_{kl}$*.*
The set of non-normalized pure states, the set of unentangled states and the set of separable states (together with $|\beta\rangle\langle\beta|$*), are sub-semi-groups of* $(Herm^+_{n^2}(\mathbb{C}), \diamond)$*. More precisely,*

$$\begin{aligned} &(AA^\dagger) \diamond (BB^\dagger) = VV^\dagger \ \ where \ \ \hat{V} = \hat{A}\hat{B} \\ &(\mu_1 \otimes \mu_2) \diamond (\sigma_1 \otimes \sigma_2) = Tr(\mu_2^t \sigma_1)\mu_1 \otimes \sigma_2 \end{aligned} \tag{5.20}$$

Proof: The composition law is just the transcription of $(\widehat{\$} \circ \widehat{€′})_{ik;jl} = \widehat{\$}_{ik;i'k'}\widehat{€′}_{i'k';jl}$ using (5.5), and the identity element is clearly $\delta_{ij}\delta_{kl}$. Next, the composition of two factorizable operations is factorizable and trivially yields (5.20). Let $€′ = \sigma_1 \otimes \sigma_2$ and $\$ = \mu_1 \otimes \mu_2$ two unentangled states. We have using Equation (5.15):

$$\begin{aligned} \widehat{€′}(\rho) &= \mathrm{Tr}_2\big((\mathbb{I} \otimes \rho^t)(\sigma_1 \otimes \sigma_2)\big) = \big(\mathrm{Tr}(\sigma_2^t \rho)\big)\sigma_1, \\ \text{hence } \widehat{\$} \circ \widehat{€′}(\rho) &= \mathrm{Tr}\big(\mu_2^t(\mathrm{Tr}(\sigma_2^t\rho)\sigma_1)\big)\mu_1 = \mathrm{Tr}(\mu_2^t\sigma_1)\mathrm{Tr}(\sigma_2^t\rho)\mu_1 \end{aligned}$$

and the last equation follows immediately.
Since the composition law is bilinear, the space of separable states of $\mathrm{Herm}^+_{n^2}$, together with the identity $|\beta\rangle\langle\beta|$, is also a sub-semi-group of $(\mathrm{Herm}^+_{n^2}(\mathbb{C}), \diamond)$. □

It seems natural at this point to look for subgroups of $(\mathrm{Herm}^+_{n^2}(\mathbb{C}), \diamond)$. Clearly the largest subgroup corresponds to the set of invertible quantum operations $\widehat{\$}$, of which it is difficult to give a physical description in terms of the states $\$$: we just require $\widehat{\$}_{ik;jl}$ to be invertible. Since unentangled states yield projections (as was illustrated in the proof above), they are not in this group; yet mixtures of them (separable states) may well yield invertible operations.

Definition 5.6 *The positive definite matrices of* $Herm^+_d(\mathbb{C})$ *are sometimes called the* totally mixed *states of* $\mathbb{C}^d$*.*
Moreover we say that $\$ \in Herm^+_{mn}(\mathbb{C})$ *is a* totally entangled *state of* $\mathbb{C}^m \otimes \mathbb{C}^n$ *if and only if* $\$$ *is pure and verifies either of*

$$\begin{aligned} &(n \leq m) \quad Tr_1(\$) \text{ is totally mixed} \\ &(m \leq n) \quad Tr_2(\$) \text{ is totally mixed} \end{aligned}$$

depending on the integers m *and* n *(if* $m = n$ *the two conditions are indifferent).*

As in Chapter 3 we let $GL_n(\mathbb{C})$ denote the group of invertible $n \times n$ complex matrices, $U(1)$ its (normal) subgroup of matrices of the type $e^{i\theta}\mathbb{I}_n$, and $SU(n)$ the group of special unitary $n \times n$ matrices, i.e. matrices U satisfying $U^\dagger U = UU^\dagger = \mathbb{I}_n$ and $\det U = 1$. We have the following:

Proposition 5.4 *The set of totally entangled pure states in $Herm^+_{n^2}(\mathbb{C})$, equipped with the composition law $\diamond$, is a group which is isomorphic to the group $GL_n(\mathbb{C})/U(1)$. Its subset of maximally entangled states is a subgroup isomorphic to $SU(n)$.*

Proof: Let us denote by T the set of totally entangled (pure) states in $\mathrm{Herm}^+_{n^2}(\mathbb{C})$. Note that for any $\hat{A} \in M_n(\mathbb{C})$, $\hat{A}$ is invertible if and only if $\hat{A}\hat{A}^\dagger$ is invertible, which by (5.12) is equivalent to $\mathrm{Tr}_2(AA^\dagger)$ invertible, in other words $AA^\dagger$ totally entangled. Thus $T = \{AA^\dagger \ / \ \hat{A} \in GL_n(\mathbb{C})\}$, and from (5.20), $(T, \diamond)$ is a group with identity element $|\beta\rangle\langle\beta|$.

$$\begin{aligned} \chi : GL_n(\mathbb{C}) &\to T \\ \hat{A} &\mapsto AA^\dagger \end{aligned}$$

is then trivially a group homomorphism, since $\chi(\hat{A}\hat{B}) = AA^\dagger \diamond BB^\dagger$ by (5.20). χ is clearly onto, and its kernel is $U(1)$. Thus $GL_n(\mathbb{C})/U(1)$ is isomorphic to T.
χ restricted to $U(n)$ maps onto the set of maximally entangled states by Lemma 5.4, so that $SU(n) = U(n)/U(1)$ is isomorphic to it. □

These results are useful when one seeks to parameterize certain pure states of an n^2-dimensional system. The description of pure states in $\mathrm{Herm}_{n^2}(\mathbb{C})$ in terms of the homogeneous space $SU(n^2)/SU(n^2-1)$ is well-known, but yields a very complicated parameterization since one must mod out the $SU(n^2-1)$. We have shown that we can in fact parameterize the set of maximally entangled (pure) states of $\mathrm{Herm}_{n^2}(\mathbb{C})$ in terms of (the Euler angles of) $SU(n)$, without having to mod out any redundancy. The parameterization could have potential applications in the study of entanglement, Bell states and EPR scenarios.

Next one can also define an original semi-group structure on the set of Completely Positive-preserving maps by using an exotic composition law (the Schur product $\triangle$) on the set of states:

Proposition 5.5 *If $\widehat{\$}$ and $\widehat{\in}$ are Completely Positive-preserving maps from $M_n(\mathbb{C})$ to $M_m(\mathbb{C})$, then so is*

$$\widehat{\$} \triangle \widehat{\in} \equiv (\widehat{\$} \triangle \widehat{\in})_{ik;jl} = \widehat{\$}_{ik;jl}\widehat{\in}_{ik;jl}$$

where the summation convention is suspended, and $i, k = 1, \ldots, m$, and $j, l = 1, \ldots, n$. This composition law is obviously commutative, and the set of factorizable operations is stable under it.

Proof: This stems, via Theorem 5.3, from the stability of the set of positive matrices under of the Schur (or Hadamard) product [30]. I.e. the fact that the component-wise product of two

positive matrices is a positive matrix, when applied to \$ times $\mathbb{E}$, induces the corresponding result for $\widehat{\$}$ times $\widehat{\mathbb{E}}$.
We use the same symbol $\triangle$ to denote all component-wise products of matrices. If $\widehat{\$}$ and $\widehat{\mathbb{E}}$ have decompositions $\{\hat{A}_x\}$ and $\{\hat{B}_y\}$ respectively, then $\widehat{\$} \triangle \widehat{\mathbb{E}}$ has decomposition $\{\hat{A}_x \triangle \hat{B}_y\}$: this implies the stability of factorizable operations. □

5.3.2 Duality: states and functionals

When relating operators and states of a physical theory notions of duality between vector spaces are often illuminating: operators sometimes induce functionals on the space of states, which can in turn be thought of as states. In finite-dimensional Quantum Mechanics, a given positive matrix can either represent a state or a positive functional, and we can switch from one to the other easily.
So far we have equipped the algebra of complex $d \times d$ matrices, $M_d(\mathbb{C})$, with the complex-bilinear form: $(\mathbb{E}, \$) = \mathrm{Tr}(\mathbb{E}^\dagger \$)$. This non-degenerate form naturally defines a canonical pairing of $M_d(\mathbb{C})$ with $\widetilde{M_d}(\mathbb{C})$, the space of linear functionals on $M_d(\mathbb{C})$:

$$\begin{aligned} \widetilde{\ } : M_d(\mathbb{C}) &\longrightarrow \widetilde{M_d}(\mathbb{C}) \\ \mathbb{E} &\longmapsto [\widetilde{\mathbb{E}} : \$ \mapsto \mathrm{Tr}(\mathbb{E}^\dagger \$)] \end{aligned}$$

Since $\widetilde{\ }$ is an (anti-linear) isomorphism, any linear functional on $M_d(\mathbb{C})$ has a unique antecedent by $\widetilde{\ }$, thus is uniquely represented by an element of $M_d(\mathbb{C})$. Let $\{E_{ij}\}_{1 \le i,j \le d}$ a canonical basis of $M_d(\mathbb{C})$ and $\{\widetilde{E}_{kl}\}_{1 \le k,l \le d}$ its corresponding peered basis, i.e. $\widetilde{E}_{kl}(E_{ij}) \equiv \mathrm{Tr}(E_{kl}^\dagger E_{ij}) = \delta_{ik}\delta_{jl}$. Then the functional of $\mathbb{E}$, namely $\widetilde{\mathbb{E}}$, is represented in the peered basis by $\mathbb{E}^*$. Indeed, $\mathbb{E}^*_{kl} \widetilde{E}_{kl}(\$_{ij} E_{ij}) = \mathbb{E}^*_{kl} \$_{kl} = \widetilde{\mathbb{E}}(\$)$.
When restricted to the real vector space of hermitian matrices $\mathrm{Herm}_d(\mathbb{C})$, $(\mathbb{E}, \$) \mapsto \mathrm{Tr}(\mathbb{E}\$)$ yields a real scalar product, and $\widetilde{\mathrm{Herm}_d}(\mathbb{C})$ is defined similarly. It then becomes possible to define the dual (sometimes called polar) of a subspace $\mathcal{S}$ of $\mathrm{Herm}_d(\mathbb{C})$ as follows:

$$\mathcal{S}^\star \equiv \{ \widetilde{\sigma} \in \widetilde{\mathrm{Herm}_{mn}}(\mathbb{C}) \ / \ \forall \rho \in \mathcal{S}, \ \widetilde{\sigma}(\rho) \ge 0 \} \tag{5.21}$$

The convex cone of hermitian positive matrices $\mathrm{Herm}_d^+(\mathbb{C})$ is clearly self-dual under this dual pairing:

$$\begin{aligned} \mathbb{E} \in \mathrm{Herm}_d^+(\mathbb{C}) &\Leftrightarrow \forall \$ \in \mathrm{Herm}_d^+(\mathbb{C}), \ \ \mathrm{Tr}(\mathbb{E}\$) \ge 0 \\ &\Leftrightarrow \forall \$ \in \mathrm{Herm}_d^+(\mathbb{C}), \ \ \widetilde{\mathbb{E}}(\$) \ge 0 \\ &\Leftrightarrow \widetilde{\mathbb{E}} \in \mathrm{Herm}_d^+(\mathbb{C})^\star \end{aligned}$$

In the last line we have used the definition (5.21). Thus $\mathrm{Herm}_d^+(\mathbb{C})^\star = \widetilde{\mathrm{Herm}_d}^+(\mathbb{C})$, hence the set of non-normalized states is isomorphic to that of non-normalized linear probability distributions on states, i.e. functionals which are positive on $\mathrm{Herm}_d^+(\mathbb{C})$. In this sense, if $\not{\mathbb{C}}$ is an element of $\mathrm{Herm}_d^+(\mathbb{C})$, then $\not{\mathbb{C}}^* \equiv \not{\mathbb{C}}^t$, represents its dual element, or associated linear probability distribution, and conversely. We shall now explain why this picture is illuminating.

Separable states and Positive-preserving maps

We now denote by $\mathrm{Herm}_{mn}^S(\mathbb{C})$ the set of separable states of $\mathbb{C}^m \otimes \mathbb{C}^n$, and define its dual space by (5.21):

$$\mathrm{Herm}_{mn}^S(\mathbb{C})^\star \equiv \{\widetilde{\sigma} \in \widetilde{\mathrm{Herm}}_{mn}(\mathbb{C}) \ / \ \forall \rho \in \mathrm{Herm}_{mn}^S(\mathbb{C}), \ \widetilde{\sigma}(\rho) \geq 0 \},$$

This is a convex cone too. The geometrical meaning of Theorem 5.2 is now clear in this formalism:

Theorem 5.2 (restatement.) *A linear operation $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ is Positive-preserving if and only if the linear functional of its associated state $\$$, namely $\widetilde{\$}$, is in $Herm_{mn}^S(\mathbb{C})^\star$. In other words, the convex cone of Positive-preserving maps is isomorphic to the dual of the convex cone of separable states.*

Remember that inclusions are reversed by duality:

$$\begin{aligned} &\mathrm{Herm}_{mn}^S(\mathbb{C}) \subsetneq \mathrm{Herm}_{mn}^+(\mathbb{C}) \\ \Leftrightarrow &\mathrm{Herm}_{mn}^+(\mathbb{C})^\star \subsetneq \mathrm{Herm}_{mn}^S(\mathbb{C})^\star. \end{aligned}$$

Since not all states are separable, this confirms the fact that Positive-preserving maps are not necessarily Complete Positive-preserving.

Remark 5.4 *The set of $\$$ in $Herm_{mn}(\mathbb{C})$ such that $\widehat{\$}$ is Positive-preserving, i.e. such that $\widetilde{\$}$ belongs to $Herm_{mn}^S(\mathbb{C})^\star$, is stable under the transposes ^{t_1} on $\mathbb{C}^m$ and ^{t_2} on $\mathbb{C}^n$.*

Proof: For $\widehat{\$}$ Positive-preserving, $\widehat{\$} \circ {}^{t_2} \equiv \widehat{\$^{t_2}}$ and ${}^{t_1} \circ \widehat{\$} \equiv \widehat{\$^{t_1}}$ are Positive-preserving too. From this simple observation we readily obtain that the set of the $\$$ is stable under partial transpositions. □

Remark 5.5 *Remark 5.4 is equivalent the Peres' criterion [49] for separability, which states that the set of the separable states $Herm_{mn}^S(\mathbb{C})$ is stable under partial transposition.*

Proof: [Peres ⇒ Remark 5.4] If $ is such that $\widetilde{\$}$ belongs to $\mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star}$, then we have that

$$\begin{aligned} & \forall \mathbb{\in} \in \mathrm{Herm}_{mn}^{S}(\mathbb{C}) \quad \mathrm{Tr}(\$\mathbb{\in}) \geq 0 \\ \Rightarrow \quad & \forall \mathbb{\in} \in \mathrm{Herm}_{mn}^{S}(\mathbb{C}) \quad \mathrm{Tr}(\$\mathbb{\in}^{t_2}) \geq 0 \quad \text{by Peres} \\ \Rightarrow \quad & \forall \mathbb{\in} \in \mathrm{Herm}_{mn}^{S}(\mathbb{C}) \quad \mathrm{Tr}(\$^{t_2}\mathbb{\in}) \geq 0 \end{aligned}$$

which means, by definition, that $\widetilde{\$^{t_2}}$ belongs to $\mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star}$. The same applies with ^{t_1}.
[Remark 5.4 ⇒ Peres] Now let $ belong to $\mathrm{Herm}_{mn}^{S}(\mathbb{C})$. Since $\mathrm{Herm}_{mn}^{S}(\mathbb{C})$ is a closed convex set containing 0 we have, by the bipolar theorem (see for instance [62]), that $\mathrm{Herm}_{mn}^{S}(\mathbb{C}) = \mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star\star}$. Thus $ belongs to $\mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star\star}$, and so

$$\begin{aligned} & \forall \widetilde{\mathbb{\in}} \in \mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star} \quad \mathrm{Tr}(\$\mathbb{\in}) \geq 0 \\ \Rightarrow \quad & \forall \widetilde{\mathbb{\in}} \in \mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star} \quad \mathrm{Tr}(\$\mathbb{\in}^{t_2}) \geq 0 \quad \text{by Remark 5.4} \\ \Rightarrow \quad & \forall \widetilde{\mathbb{\in}} \in \mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star} \quad \mathrm{Tr}(\$^{t_2}\mathbb{\in}) \geq 0 \end{aligned}$$

which means, by definition, that $\$^{t_2}$ belongs to $\mathrm{Herm}_{mn}^{S}(\mathbb{C})^{\star\star} = \mathrm{Herm}_{mn}^{S}(\mathbb{C})$. The same applies with ^{t_1}: we have recovered Peres' criterion. □
That the Peres' criterion corresponds to the simple fact that $\widehat{\$}$ Positive-preserving implies $\widehat{\$}\circ^{t}$ Positive-preserving is a somewhat striking fact. This insight may well help to build tighter criterions: recently the Horodeckis [31] have been following this line of thought.

Physical interpretation of formulae

When attempting to characterize separability the notions of duality seem to play a simplifying role, as they help to clarify the correspondence induced by Isomorphism 2. Thus one may wonder if these concepts could facilitate the interpretation of other results in this chapter. We now give a formulation of quantum operations $\widehat{\$}$ in terms of single operations on their corresponding state $.

Proposition 5.6 *Let $ represent a non-normalized quantum state of a bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^m \otimes \mathbb{C}^n$ shared by Alice and Bob. Suppose Bob performs on $ a local generalized measurement $\{\mathbb{I}_m \otimes M_n^{(x)}\}_x$. Call $\mathbb{I}_m \otimes M$ the element whose outcome occurs and let $\rho_B \equiv (M^{\dagger}M)^{t} \in Herm_n^{+}(\mathbb{C})$.*
Then the unrescaled post-measurement state as viewed by Alice is precisely $\widehat{\$}(\rho_B)$. Thus the effect of any quantum operation $\widehat{\$}$ can be viewed as the trace out of a particular local single operation on its corresponding state $.

Proof: The unrescaled post-measurement state is simply $\$^M = (\mathbb{I}_m \otimes M)\$(\mathbb{I}_m \otimes M^\dagger)$. Using (5.15) this yields for Alice the state:

$$\begin{aligned}\mathrm{Tr}_2(\$^M) &= \mathrm{Tr}_2\big((\mathbb{I}_m \otimes M)\$(\mathbb{I}_m \otimes M^\dagger)\big)\\ &= \widehat{\$}\big((M^\dagger M)^t\big)\\ &= \widehat{\$}(\rho_B)\end{aligned}$$

□

The fact that there is a transpose corroborates the idea of duality. Indeed, first $M^\dagger M$ is thought of as defining a functional $\sigma \mapsto \mathrm{Tr}(M^\dagger M\sigma)$, but then as we think of a quantum operation as acting on states we act upon its transpose. The map $M^\dagger M \mapsto \widehat{\$}\big((M^\dagger M)^t\big)$, though it is Positive-preserving, is not Completely Positive-preserving since it can be written as $\widehat{\$} \circ {}^t$. However the same map defined from states to states, i.e. $(M^\dagger M)^t \mapsto \widehat{\$}((M^\dagger M)^t)$, is Completely Positive-preserving. Proposition 5.6 suggests that quantum states in $\mathrm{Herm}^+_{mn}(\mathbb{C})$ inherently defines a quantum operation between their two subsystems.

5.4 Summary and concluding remarks

In this chapter we made several new contributions, some technical, others more geometrical. Amongst the technical results we provided two triangular decompositions for pure states of a bipartite system, i.e. local changes of basis so that vectors in $\mathbb{C}^m \otimes \mathbb{C}^n$ may be written with triangular coefficients only. We also gave two original algebraic tests on Completely Positive-preserving maps: one regarding extremality in the set of Trace-preserving operations, the other testing the factorizability or single operator decomposition. The latter is particularly interesting since it does not depend on the operator sum decompositions of these maps. The formulae in Proposition 5.1 should yield simplifications in optimization of fidelities of quantum operations as encountered for instance in quantum cryptographic problems.

On the more geometrical side we endowed $\mathrm{Herm}^+_{n^2}(\mathbb{C})$ with a semi-group structure stemming from the composition law on quantum operations. This in turn provided a group isomorphism between totally entangled (pure) states and $GL_n(\mathbb{C})/U(1)$, and maximally entangled (pure) states and $SU(n)$. This result sheds light on the geometry of entangled states as it suggests, for future work, simple parameterizations and bi-invariant metrics on the corresponding (group-)submanifolds of the set of pure states in $\mathrm{Herm}^+_{n^2}(\mathbb{C})$. In addition we showed that the set of quantum operations is stable under component-wise product.

These contributions are interesting enough by themselves, but perhaps the most significant achievement of this chapter is to demonstrate the central, transversal role of the state-operator isomorphism as formalized in Isomorphism 2 and justified by Theorem 5.3. We have shown that virtually all the main results regarding states/operators can be elegantly obtained as corollaries of their operator/state analogue, which makes this correspondence one of the most

Matrix \$	*Linear operator* $\widehat{\$}$
Hermitian	Hermitian-preserving
Dual to separable	Positive-preserving
Positive	Completely Positive-preserving
Particular state \$	*Particular quantum operation* $\widehat{\$}$
Pure	Factorizable
Unentangled $\sigma_1 \otimes \sigma_2$	Projection $\rho \mapsto \big(\mathrm{Tr}_2(\sigma_2^t \rho)\big)\sigma_1$
Separable	Sum of projections Dual to Positive-preserving
$\mathrm{Tr}_1(\$) = \mathbb{I}$	Trace-preserving
$\mathrm{Tr}_1(\$) = \mathbb{I}$ and $\mathrm{Tr}_2(\$) = \mathbb{I}$	Bistochastic
Particular ket A	*Particular evolution matrix* $\hat{A}$
Maximally entangled	Unitary
Totally entangled	Invertible
$\sum_i \lvert i\rangle\lvert i\rangle$	$\mathbb{I}$
$\sum_i \lambda_i \lvert i\rangle\lvert i\rangle$	$Diag\{\lambda_i\}$
$\sum_i \lambda_i \lvert\psi_i\rangle\lvert\psi_i^*\rangle$ with $\forall i, \lambda_i \in \mathbb{R}$ with $\forall i, \lambda_i \in \mathbb{R}^+$	 Hermitian Positive
Theorems on states	*Theorems on quantum operations*
Spectral decomposition, Unitary degree of freedom	Operator Sum decomposition, Unitary degree of freedom
Purification	$\widehat{\$}(\rho) = \mathrm{Tr}_1(U\rho U^\dagger)$
Bipartite decompositions: Schmidt One-sided triangular Two-sided triangular	Matrix decompositions: Polar QR Schur's triangularization
Purity condition	Factorizability condition
Formulae on states	*Formulae on quantum operations*
$\mathrm{Tr}_{1/2}(AB^\dagger)$	$= (\hat{B}^\dagger \hat{A})^t / \hat{A}\hat{B}^\dagger$
$\mathrm{Tr}_2\big((\kappa \otimes \rho^t)\$(\tau \otimes \sigma^t)\big)$	$= \kappa \widehat{\$}(\rho\sigma)\tau$
$\mathrm{Tr}_2(\mathbb{I} \otimes \rho)\$(\mathbb{I} \otimes \rho^\dagger))$	$= \widehat{\$}((\rho^\dagger \rho)^t)$
$\mathrm{Tr}((\sigma \otimes \rho^t)\$)$	$= \mathrm{Tr}(\sigma \widehat{\$}(\rho))$
$\mathrm{Tr}(\not{\epsilon}^\dagger \$)$	$= \sum \mathrm{Tr}(\widehat{\not{\epsilon}}(E_{jl})^\dagger \widehat{\$}(E_{jl}))$

Table 5.1: Summary of the state-operator correspondence.

fruitful linear algebraic tool in the surroundings of quantum theory (see table 5.3.2 for summary). Even for more specialist issues of quantum information theory we find that the isomorphism has a role to play, as was illustrated by the problem of characterizing separable states.

On this occasion we introduced notions of duality, which serve both to facilitate the interpretation of the state-operator correspondence and its related formulae, and to understand the underlying geometry from a slightly more abstract point of view. The formulae themselves should have numerous applications in quantum information theory (as we shall demonstrate in Chapter 7), and could also provide a novel interpretation of states versus operations in open systems (as suggested in Proposition 5.6).

Part II

Quantum cryptography

Chapter 6

The qubit information gain versus disturbance

Dans les champs de l'observation le hasard ne favorise que les esprit préparés.
—Louis Pasteur

Alice draws uniformly at random a state amongst two pure states $\{|\psi_0\rangle, |\psi_1\rangle\}$, and then sends it over a quantum channel. Eve, the malevolent eavesdropper, gains access to this $|\psi_x\rangle$ and may use this opportunity to try and learn about x. How much she learns is quantified using information theoretical notions. But at the receiving end honest Bob, whom we assume knows the value of x, gets a chance to check whether Eve interfered. Indeed, suppose Eve's measurement and further manipulations have changed $|\psi_x\rangle$ into ρ_x. If Bob measures $\{|\psi_x\rangle\langle\psi_x|, \mathbb{I} - |\psi_x\rangle\langle\psi_x|\}$ upon the received state he has a probability $1 - \langle\psi_x|\rho_x|\psi_x\rangle$ of detecting the felony. The above idealized scenario captures a key ingredient for any quantum cryptographic protocol, namely the fact that the eavesdropper cannot observe a state without causing it an irreversible, detectable damage. In spite of their central role, information gain versus disturbance tradeoffs upon discrete ensembles remain largely unknown, due to the mathematical difficulties they raise. In 1995 Fuchs and Peres managed to obtain an analytic formula for the above case of two equiprobable non-orthogonal pure states. In this chapter we rederive their result intuitively and geometrically − in the context of the conal representation.

The key principle of quantum cryptography could be summarized as follows. *Honest parties communicate using quantum states. To the eavesdropper these states are random and non-orthogonal. In order to gather information she must measure them, but this may cause irreversible damages. Honest parties seek to detect her mischief by checking whether certain quantum states are left intact.* The tradeoff between the eavesdropper's information gain (about an ensemble of quantum states), and the disturbance she necessarily induces (upon this ensemble), can thus be viewed as the power engine behind quantum cryptographic protocols.
Yet while numerous protocol-specific proofs of security have been given, information gain versus disturbance tradeoffs themselves have remained stubbornly difficult to quantify. The problem was first taken over by Fuchs and Peres [21], who tackled the seemingly simple case of the two non-orthogonal equiprobable states ensemble $\{(1/2, |\psi_0\rangle), (1/2, |\psi_1\rangle)\}$. For discrete distributions this is just about the only result available. Of lesser interest for cryptography, but very important in terms of its methods is the work by Banaszek [5], who quantified the tradeoff for the continuous uniform n-dimensional ensemble. Barnum [6] makes several accurate qualitative remarks upon the same ensemble, suggesting the tradeoff remains unchanged for uniform distributions over mutually unbiased states.

Unfortunately Fuchs and Peres' derivation involves lengthy algebra, a number of assumptions, and seems extremely difficult to apply to even slight variations of their scenario. One should be able to find a method which gives a glimpse of intuition about the geometry of optimal measurements: this is the purpose of the present chapter. The Cone, by enabling a *per-outcome* geometrical representation of generalized measurements an their effects, greatly facilitates the derivation of Fuchs and Peres' formula and enables us to visualize the family of optimal measurements. We hope this illustrates the power of the geometrical framework developed in Chapter 3. But before we begin, let us describe the exact scenario once again, formally:

Scenario 6.1 (Fuchs and Peres') *Consider a quantum channel for transmitting qubits. Suppose Alice owns a random variable* $X = \{(\frac{1}{2}, 0), (\frac{1}{2}, 1)\}$. *According to outcome* x *she prepares either* $|\psi_0\rangle$ *or* $|\psi_1\rangle$ *and sends it to Bob. Suppose that Bob, whenever the state* $|\psi_x\rangle$ *gets sent, measures*

$$\{P_{intact} = |\psi_x\rangle\langle\psi_x|\,, \quad P_{tamper} = \mathbb{I} - |\psi_x\rangle\langle\psi_x|\} \tag{6.1}$$

so as to check for tampering. Suppose Eve is eavesdropping the quantum channel, and has an interest in determining whether Alice sent $|\psi_0\rangle$ *or* $|\psi_1\rangle$.

6.1 Mathematical preliminaries

6.1.1 Information contribution

Suppose the $\{|\psi_x\rangle\}_{x=0,1}$ states Alice prepares verify the following basic relations (with ϕ defined as in Subsection 3.1.1, page 45):

$$\underline{v^x} = \phi(|\psi_x\rangle\langle\psi_x|) \qquad \sqrt{\frac{\underline{v^0}.\underline{v^1}}{2}} = d = \sqrt{1-c^2}$$

By choosing a suitable basis in the Bloch Sphere and since the $\{|\psi_x\rangle\}_x$ are pure we may fix:

$$\underline{v^0} = [\ 1 \quad c \quad d \quad 0\ ] \tag{6.2}$$

$$\underline{v^1} = [\ 1 \quad -c \quad d \quad 0\ ] \tag{6.3}$$

The most general thing Eve can ever do is to attack the states with a measurement $\{M_m\}_m$. This procedure is equivalent to first measuring $\{\sqrt{E_m}\}_m$, and then, conditional to m, applying the unitary transformation U_m, with E_m and U_m defined as in Subsection 3.1.3. It is rather interesting to observe that the second step has no other use but to 'repair' the post-measurement states as much as is possible. The first step on the other hand may partially destroy the initial states so as to collect the information Eve seeks. This is the step we now study in order to quantify her information gain.

Let Y be the random variable arising from the measurement outcomes, i.e. $Y = \{(p(m), m)\}_m$. In this chapter we quantify Eve's information gain in terms of Shannon mutual entropy (see [48]):

$$\begin{aligned} I &= H(X:Y) = H(Y) - H(Y|X) \\ &= -\sum_m p(m)\log(p(m)) + \sum_{x,m} p(x,m)\log(p(m|x)) \\ &\equiv \sum_m I_m \quad \text{with} \quad I_m = -p(m)\log(p(m)) + \sum_x p(x,m)\log(p(m|x)) \end{aligned}$$

I_m must be understood as the *information contribution* brought by the measurement element:

$$\underline{\varepsilon_m} = [\ \alpha \quad \beta \quad \gamma \quad \delta\ ] = \phi(E_m)$$

By making use of the relations (6.2),(6.3) and (3.9) one can express I_m geometrically in terms

of scalar products in the cone:

$$I_m = -(p_m + q_m)\log(p_m + q_m) + p_m \log(2p_m) + q_m \log(2q_m)$$

$$\text{with} \quad p_m = \frac{\alpha + \beta c + \gamma d}{4} = \frac{\underline{\varepsilon_m}.\underline{v^0}}{4} \equiv p(0,m) \tag{6.4}$$

$$q_m = \frac{\alpha - \beta c + \gamma d}{4} = \frac{\underline{\varepsilon_m}.\underline{v^1}}{4} \equiv p(1,m) \tag{6.5}$$

Notice that if $\underline{\varepsilon_m}$ is orthogonal to $\underline{v^1}$ (resp. $\underline{v^0}$) then $I_m = p_m$ (resp. q_m). Such a measurement element may be said to be "all or nothing": it brings a whole bit of information when it occurs, but does so only with probability p_m (resp. q_m). Taken individually these measurement elements seem ideal: they fully identify $|\psi_x\rangle$ and thus they let you reconstruct the initial state perfectly, with no disturbance at all. The downside is that failure to occur comes at a high price. In order to verify the trace-preservation condition (3.16) the other measurement elements generally become rather inefficient with respect to the tradeoff. The family of the optimizing $\{M_m\}_m$ is not constructed in such simple ways.

6.1.2 Disturbance contribution

Next we seek an expression of the *disturbance contribution* brought by each measurement element. For this purpose we must first assume outcome m has occurred. Eve knows it, and now she will try to maximize her chances of fooling Bob by applying a carefully tailored unitary evolution U_m. First we will give D_m as a function of U_m, and next proceed to the maximization which determines U_m. As was already the case in Subsection 3.2.4, page 56, upper indices x will be used to distinguish initial states, whilst lower indices m specify the measurement outcome. Moreover note that ${\rho_m^x}'$ is the *rescaled* (unit -trace) post-measurement state.

$$\begin{aligned}
p(\text{fooling Bob}|m) &= \sum_x p(x|m)\mathrm{Tr}(|\psi_x\rangle\langle\psi_x|U_m {\rho_m^x}' U_m) \\
&\equiv \frac{\sum_x p(x|m)\underline{v^x}.\underline{r_m^x}}{2} \\
&= \frac{1 + \sum_x p(x|m)\overrightarrow{v^x}.\overrightarrow{r_m^x}}{2} \quad \text{where} \\
\underline{r_m^x} &= [\ 1 \quad \overrightarrow{r_m^x}\ ] \equiv \phi(U_m {\rho_m^x}' U_m) \\
&= \frac{\phi(U_m\sqrt{E_m}|\psi_x\rangle\langle\psi_x|\sqrt{E_m}U_m)}{p(m|x)}
\end{aligned}$$

Negating back to the disturbance we obtain:

$$\begin{aligned} D &= \sum_m D_m \quad \text{with} \\ D_m &= p(\text{not fooling Bob}, m) \\ &= \frac{p(m) - \sum_x p(x,m)\overrightarrow{v^x}.\overrightarrow{r_m^x}}{2} \\ &= \frac{p(m) - \sum_x p(x,m)\|\overrightarrow{v^x}\|\,\|\overrightarrow{r_m^x}\|\cos\widehat{(\overrightarrow{v^x},\overrightarrow{r_m^x})}}{2} \end{aligned}$$

In our scenario the $\{|\psi_x\rangle\}_x$ are pure. Thus by Lemma 3.2 or equation (3.19) we have $\|\overrightarrow{v^x}\|\,\|\overrightarrow{r_m^x}\| = 1$ (i.e. the per-outcome effect of a measurement takes a pure state into a pure state). Now let us deal with $\cos\widehat{(\overrightarrow{v^x},\overrightarrow{r_m^x})}$ by making the following definitions:

$$\begin{aligned} \theta &\equiv \widehat{(\overrightarrow{v^0},\overrightarrow{v^1})} \\ \theta_m &\equiv \widehat{(\overrightarrow{r_m^0},\overrightarrow{r_m^1})} \\ \Delta_m &\equiv \theta - \theta_m \\ \omega_m &\equiv \widehat{\left(\overrightarrow{r_m^0}+\overrightarrow{r_m^1},\overrightarrow{v^0}+\overrightarrow{v^1}\right)} \end{aligned}$$

ω_m is the angle between the bisector of $(\overrightarrow{r_m^0},\overrightarrow{r_m^1})$ and that of $(\overrightarrow{v^0},\overrightarrow{v^1})$. Given that we want to minimize D_m in terms of U_m we can safely assume $\overrightarrow{r_m^0},\overrightarrow{r_m^1},\overrightarrow{v^0},\overrightarrow{v^1}$ to be coplanar. Thus D_m may now be rewritten in terms of those angles as well as p_m and q_m:

$$D_m = \frac{p_m + q_m - p_m\cos(\Delta_m - \omega_m) - q_m\cos(\Delta_m + \omega_m)}{2}$$

In this equation the values of p_m, q_m and Δ_m are fully determined by $\underline{\varepsilon_m}$, as described in (6.4),(6.5), an the 'inner product through measurement' Equation (3.18). ω_m on the other hand solely depends on U_m: it can be chosen at will by rotation in the Bloch Sphere. We now show how Eve must tune ω_m so as to minimize D_m.

$$\frac{\partial D_m}{\partial \omega_m} = 0 \Rightarrow p_m\sin(\Delta_m - \omega_m) - q_m\sin(\Delta_m + \omega_m) = 0$$

The minimum occurs at:

$$\omega_m = \arcsin\left(\frac{p_m - q_m}{\sqrt{p_m^2 + q_m^2 + 2p_mq_m\cos(2\Delta_m)}}\right)$$

which yields, after simplification:

$$D_m = \frac{p_m + q_m - \sqrt{p_m^2 + q_m^2 + 2p_mq_m\cos(2\Delta_m)}}{2}$$

6.2 Optimization and conclusion

How many elements should Eve's measurement contain? Levitin has proved that there exists a two-element measurement $\{M_m\}_{m=0,1}$ which maximizes Eve's information gain [41]. While this was never formally shown to be the case for the measurements which optimize the information gain versus disturbance tradeoff, there is strong numerical evidence in support of this assumption [21]. Suppose this is the case and let $\underline{\varepsilon_m}_\mu$ denote the μ^{th} coordinate of $\underline{\varepsilon_m}$. Using the trace-preservation constraint given by Equation (3.16) we have:

$$\delta\underline{\varepsilon_0}_\mu = -\delta\underline{\varepsilon_1}_\mu. \tag{6.6}$$

Optimizing the Tradeoff implies finding a stationary point for the disturbance while keeping the information gain fixed. We need to find $\underline{\varepsilon_0}$ such that

$$\sum_\mu \frac{\partial D}{\partial \underline{\varepsilon_0}_\mu}\delta\underline{\varepsilon_0}_\mu = 0$$

where the variations $\delta\underline{\varepsilon_0}_\mu$ are subject to the additional constraint:

$$\sum_\mu \frac{\partial I}{\partial \underline{\varepsilon_0}_\mu}\delta\underline{\varepsilon_0}_\mu = 0$$

Using equation (6.6) and $D = D_0 + D_1$ and $I = I_0 + I_1$ this gives:

$$\sum_\mu \frac{\partial D_0}{\partial \underline{\varepsilon_0}_\mu}\delta\underline{\varepsilon_0}_\mu = \sum_\mu \frac{\partial D_1}{\partial \underline{\varepsilon_1}_\mu}\delta\underline{\varepsilon_0}_\mu \tag{6.7}$$

$$\text{subject to} \quad \sum_\mu \frac{\partial I_0}{\partial \underline{\varepsilon_0}_\mu}\delta\underline{\varepsilon_0}_\mu = \sum_\mu \frac{\partial I_1}{\partial \underline{\varepsilon_1}_\mu}\delta\underline{\varepsilon_0}_\mu \tag{6.8}$$

Guided by the geometrical picture of the scenario one may consider the following attack (see FIG. 6.1):

$$\underline{\varepsilon_0} = [\ 1 \quad \beta \quad 0 \quad 0\ ]$$
$$\underline{\varepsilon_1} = [\ 1 \quad -\beta \quad 0 \quad 0\ ]$$

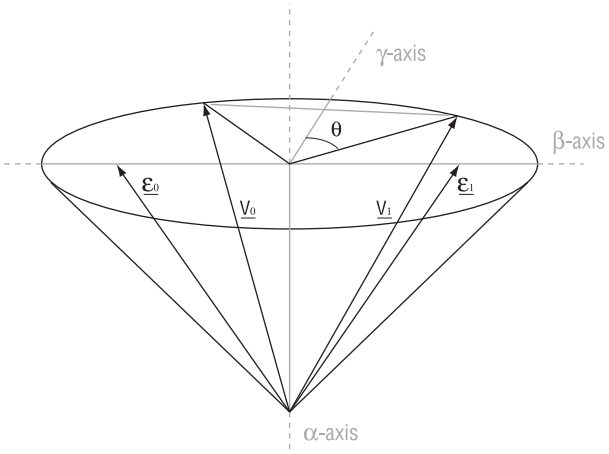

Figure 6.1: Optimal measurement family

The fact that this is indeed a solution follows from its obvious symmetries:

$$\text{For } \mu \neq 1 \quad \frac{\partial D_0}{\partial \underline{\varepsilon_0}_\mu} = \frac{\partial D_1}{\partial \underline{\varepsilon_1}_\mu} \text{ and for } \mu = 1 \quad \frac{\partial D_0}{\partial \underline{\varepsilon_0}_\mu} = -\frac{\partial D_1}{\partial \underline{\varepsilon_1}_\mu} \tag{6.9}$$

$$\text{For } \mu \neq 1 \quad \frac{\partial I_0}{\partial \underline{\varepsilon_0}_\mu} = \frac{\partial I_1}{\partial \underline{\varepsilon_1}_\mu} \text{ and for } \mu = 1 \quad \frac{\partial I_0}{\partial \underline{\varepsilon_0}_\mu} = -\frac{\partial I_1}{\partial \underline{\varepsilon_1}_\mu} \tag{6.10}$$

Substituting (6.10) in the constant information constraint (6.8), we get $\delta\underline{\varepsilon_0}_1 = 0$. Using this fact together with equation (6.9) it becomes clear that condition (6.7) is fulfilled. Thus $\underline{\varepsilon_0}$ is a stationary point. We may now proceed to compute the values of the disturbance and the information gain under this family of optimal attacks. First by making a few additional observations:

$$\begin{aligned} p_0 &= q_1 = p \\ p_1 &= q_0 = q \\ D_0 &= D_1 = D/2 \\ I_0 &= I_1 = I/2 \\ D &= \frac{1}{2} - \sqrt{p^2 + q^2 + 2pq + \cos(2\Delta_m)} \\ I &= 1 + 2p\log(2p) + 2q\log(2q) \end{aligned}$$

and second by plugging in the relations (3.18), (6.2)-(6.5), we reproduce the exact content of Fuchs and Peres' formulae:

$$\begin{aligned} D &= \frac{1}{2} - \frac{1}{2}\sqrt{1 + (c^2 - c^4)(\beta^2 - 2 + 2\sqrt{1-\beta^2})} \\ I &= \frac{1}{2}((1+\beta c)\log(1+\beta c) + (1-\beta c)\log(1-\beta c)) \end{aligned}$$

where β is a parameter ranging from 0 to 1.

Therefore we have recovered Fuchs and Peres' information gain versus disturbance formula in an elegant and geometrical manner. In the future we should be able to extend this type of analysis to the case of two non-equiprobable states - or without having to assume a two element generalised measurement. Most importantly we illustrated the uses of the conal representation developed in Chapter 3, and gathered some intuition about the family of measurements which optimizes the information gain versus disturbance tradeoffs. No doubt this understanding guided our steps as we tackled the much more complex, n-dimensional scenario we present in the next chapter.

Chapter 7

Quantum decoys

Fortune may rob our wealth, but never our courage

—Seneca

Alice communicates with words drawn uniformly amongst $\{|j\rangle\}_{j=1\ldots n}$, the canonical orthonormal basis. Sometimes however Alice interleaves quantum decoys $\frac{|j\rangle+i|k\rangle}{\sqrt{2}}$ between her messages. Such pairwise superpositions of possible words cannot be distinguished from the message words. Thus as malevolent Eve observes the quantum channel, she runs the risk of damaging the superpositions (by causing a collapse). At the receiving end honest Bob, whom we assume is warned of the quantum decoys' distribution, checks upon their integrity with a measurement. The present chapter establishes, in the case of individual attacks, the tradeoff between Eve's information gain (her chances, if a message word was sent, of guessing which) and the disturbance she induces (Bob's chances, if a quantum decoy was sent, to detect tampering). Besides secure channel protocols, quantum decoys seem a powerful primitive for constructing n-dimensional quantum cryptographic applications. Moreover the methods employed in this chapter should be of interest to anyone concerned with information gain versus disturbance tradeoffs derivations.

In this chapter we quantify the disturbance induced upon the uniform ensemble of n-dimensional states $\{(1/n^2, \rho_{jk}\}$, where j and k range from 1 to n, and ρ_{jk} stands for the density matrix of pairwise superpositions $\frac{(|j\rangle+i|k\rangle)(\langle j|-i\langle k|)}{2}$ (note that when $j = k$ this is simply the basis state $|j\rangle\langle j|$). When making use of non-orthogonal states this is no doubt a natural distribution to consider, and thus an important building block for n-dimensional cryptographic protocols. Its $\pi/2$ phase renders the 'pairing ensemble' indistinguishable from the canonical ensemble $\{(1/n, |j\rangle\langle j|)\}$, for they both have density matrix $\mathbb{I}/n$ (the maximally mixed state). This feature enables the honest parties to hide the pairwise superpositions within classical messages as a means of securing those, i.e. to use the superpositions as 'quantum decoys'. In such situations the eavesdropper seeks to gather information about the classical messages, not the decoys. Therefore we quantify her information gain with respect to the canonical ensemble $\{(1/n, |j\rangle\langle j|)\}$, as suits the following scenario best:

Scenario 7.1 (Quantum decoys) *Consider a quantum channel for transmitting n-dimensional systems having canonical orthonormal basis $\{|j\rangle\}$. Suppose Alice's message words are drawn from the canonical ensemble $\{(1/n, |j\rangle\langle j|)\}_{j=1\ldots n}$, whilst her* quantum decoys *are drawn from the pairing ensemble $\{(1/n^2, \rho_{jk}\}_{j,k=1\ldots n}$, with $\rho_{jk} = \frac{(|j\rangle+i|k\rangle)(\langle j|-i\langle k|)}{2}$. Alice sends Bob, over the quantum channel, either a message word or a decoy. Suppose that Bob, whenever a quantum decoy ρ_{jk} gets sent, measures*

$$\{P_{intact} = \big(\frac{|j\rangle + i|k\rangle}{\sqrt{2}}\big)\big(\frac{\langle j| - i\langle k|}{\sqrt{2}}\big)\,, \quad P_{tamper} = \mathbb{I} - P_{intact}\} \tag{7.1}$$

so as to check for tampering. Suppose Eve is eavesdropping the quantum channel, and has an interest in determining Alice's message words.

Little is known as we write these lines regarding cryptographic protocols involving n-dimensional quantum systems (where n is left to vary), save for two interesting articles [3][11] (these focus on mutually unbiased states). We hope our main result will prove a useful contribution to this line of research:

Claim 7.1 (Statement of security) *Referring to Scenario 1, suppose Eve performs an individual attack such that, whenever a message word gets sent, she is able to identify which with probability G (mean estimation fidelity).*
Then, whenever a quantum decoy gets sent, the probability D (induced disturbance) of Bob detecting the tampering is bounded below under the following tight inequality:

$$D \geq \frac{1}{2} - \frac{1}{2n}\Big(\sqrt{G} + \sqrt{(n-1)(1-G)}\Big)^2 \tag{7.2}$$

For optimal attacks G varies from $\frac{1}{n}$ to 1 as D varies from 0 to $\frac{1}{2} - \frac{1}{2n}$.

The remainder of this chapter is dedicated to proving the above statement. The method is highlighted, as it seems applicable to several similar problems in quantum cryptography.

In Section 7.1 we provide the necessary mathematical results required to prove Claim 7.1. We recall, in particular, a key inequality upon eigenvalues of measurement elements (first obtained in [5]), as well as a powerful formula arising from the state-operator correspondence (first obtained in Chapter 5). Section 7.2 exploits the latter formula to express the probability of Bob *not* detecting the tampering (induced fidelity) as a linear functional upon the positive matrix corresponding to Eve's attack. This brings about crucial simplifications, finally placing us in a position to apply the inequality. We do so in Section 7.3, and prove our claim.

7.1 Mathematical methods

In this section we let $\{|i\rangle\}$ and $\{|j\rangle\}$ be orthonormal basis of $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively, which we will refer to as canonical.
The following result is a minor generalization of some steps by Banaszek [5].

Proposition 7.1 (Inequality) *Consider a vector of complex numbers $v = (a_{jr})_{jr}$ together with a function $j : \mathbb{N} \longrightarrow \mathbb{N}$. We then have:*

$$f \leq \left(\sqrt{g} + \sqrt{(m-1)(n-g)}\right)^2$$

With

$$g = \sum_r |a_{j_{(r)}r}|^2$$
$$f = \sum_r |\sum_{j=0}^{m-1} a_{jr}|^2$$

And subject to $||v||^2 = n$.

Proof. Further let

$$v_j = (a_{jr})_r \quad ; \quad v_{j_{(r)}} = (a_{j_{(r)}r})_r$$
$$v'_j = (a_{jr})_r \text{ with r such that } j_{(r)} \neq j$$

and notice that $g = ||v_{j_{(r)}}||^2$, $f = \sum_{ij} v_i.v_j^*$. The Cauchy-Schwartz inequality yields:

$$v_i.v_j^* \leq ||v_i||\,||v_j||$$
$$f \leq (\sum_{j=0}^{m-1} ||v_j||)^2$$
$$f \leq (\sqrt{g} + \sum_{j=0}^{m-1} ||v'_j||)^2 \tag{7.3}$$

The quadratic/arithmetic mean inequality yields:

$$\frac{1}{m-1}\sum_{j=0}^{m-1}||v'_j|| \leq \sqrt{\frac{1}{m-1}\sum_{j=0}^{m-1}||v'_j||^2}$$
$$\leq \sqrt{\frac{n-g}{m-1}} \tag{7.4}$$

Combining Inequalities (7.3) and (7.4) yields the lemma. □

For convenience we now remind the reader of some important definitions and results from the state-operator correspondence. These were largely developed in Chapter 5, where all the proofs can be found. First let us relate vectors of $\mathbb{C}^m \otimes \mathbb{C}^n$ to endomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$.

Isomorphism 7.1 *The following linear map*

$$\hat{} : \mathbb{C}^m \otimes \mathbb{C}^n \to End(\mathbb{C}^n \to \mathbb{C}^m)$$
$$A \mapsto \hat{A}$$
$$\sum_{ij} A_{ij}|i\rangle|j\rangle \mapsto \sum_{ij} A_{ij}|i\rangle\langle j|$$

where $i = 1 \ldots m$ and $j = 1 \ldots n$, is an isomorphism taking mn vectors A into $m \times n$ matrices $\hat{A}$.

Second we relate elements of $M_{mn}(\mathbb{C})$ to linear maps from $M_n(\mathbb{C})$ to $M_m(\mathbb{C})$.

Isomorphism 7.2 *The following linear map:*

$$\widehat{\ } : \mathbb{C}^{mn} \otimes (\mathbb{C}^{mn})^\dagger \longrightarrow End(M_n(\mathbb{C}) \to M_m(\mathbb{C}))$$
$$\$ \longmapsto [\widehat{\$} : \rho \mapsto \widehat{\$}(\rho)]$$
$$\textit{such that } AB^\dagger \longmapsto [\rho \mapsto \hat{A}\rho\hat{B}^\dagger] \quad \textit{i.e.}$$
$$\sum_{ijkl} A_{ij}B^*_{kl}|i\rangle|j\rangle\langle k|\langle l| \longmapsto [\rho \mapsto \sum_{ijkl} A_{ij}B^*_{kl}|i\rangle\langle j|\rho|l\rangle\langle k|\ ]$$

where $i, k = 1 \ldots m$ and $j, l = 1 \ldots n$, is an isomorphism.

Definition 7.1 *A linear map $\Omega : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ is Completely Positive-preserving if and only if for all r and for all ρ in $Herm^+_{mr}(\mathbb{C})$, $(\Omega \otimes \mathbb{I}_r)(\rho)$ belongs to $Herm^+_{nr}(\mathbb{C})$.*

Completely Positive-preserving linear maps from quantum states in $\mathrm{Herm}^+_n(\mathbb{C})$ to quantum states in $\mathrm{Herm}^+_m(\mathbb{C})$ are exactly those which are physically allowable. They correspond, via Isomorphism 7.2, to quantum states in $\mathrm{Herm}^+_{mn}(\mathbb{C})$:

Theorem 7.1 [13] *The linear operation $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ is* Completely Positive-preserving *if and only if $\$$ belongs to $Herm^+_{mn}(\mathbb{C})$.*

Definition 7.2 *A linear map* $\widehat{\$} : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *is* Trace-preserving *if and only if for all* ρ *in* $M_n(\mathbb{C})$, $Tr(\Omega(\rho)) = Tr(\rho)$.

Completely Positive-preserving linear maps having unit probability of occurrence on every input quantum state are exactly those which are Trace-preserving. They correspond, via Isomorphism 7.2, to quantum states in $\mathrm{Herm}^{+}_{mn}(\mathbb{C})$ verifying

$$\mathrm{Tr}_1(\$) = \mathbb{I}_n. \tag{7.5}$$

Proposition 7.2 (State-operator formula) *Let* $\widehat{\$}$ *a linear map from* $M_n(\mathbb{C})$ *to* $M_m(\mathbb{C})$ *and* σ, ρ *two elements of* $M_n(\mathbb{C})$. *We have:*

$$Tr\big(\sigma\widehat{\$}(\rho)\big) = Tr\big((\sigma \otimes \rho^t)\$\big)$$

As with many quantum cryptographic problems our analysis will require a careful optimization of the fidelity induced by a quantum operation $\widehat{\$}$. By means of the above formula we shall be able to write the induced fidelity as a linear functional upon \$. This step is crucial to the next section (Lemma 7.3).

7.2 Preliminary calculations

The purpose of these calculations is to express Eve's information gain and induced disturbance in terms of the eigenvalues of her measurement elements.

7.2.1 Information gain

There exists several well-motivated manners in which to quantify Eve's information gain. The one we shall adopt focuses on her ability to make a guess after the measurement. Compared with the Shannon mutual entropy we used in the previous chapter, the mean estimation fidelity is advantageously close in nature to the notion of disturbance.

Definition 7.3 (Mean estimation fidelity) *The* mean estimation fidelity *of a generalized measurement* $\{\hat{A}_r\}$ *with guesses* $\{|\psi_r\rangle\}$ *w.r.t to an ensemble* $\{(p_i, |\phi_i\rangle)\}$ *is defined by:*

$$\begin{aligned} G &= \sum_{r,i} p(r,i) |\langle \phi_i | \psi_r \rangle|^2 \\ &= \sum_{r,i} p_i \langle \phi_i | \hat{A}_r^\dagger \hat{A}_r | \phi_i \rangle \, Tr(|\phi_i\rangle\langle\phi_i|\psi_r\rangle\langle\psi_r|) \end{aligned}$$

The mean estimation fidelity is to be understood as the average fidelity between the measurer's *guess* knowing outcome r occurred (the $|\psi_r\rangle$'s) and the i^{th} state which was indeed originally sent to him (the $|\phi_i\rangle$'s).

Notice it is justified to consider that Eve's preferred attack is a generalized measurement. In general she could perform a quantum operation, which leaves her the possibility to regroup several measurement outcomes into one likelier outcome. But there is no information to be gained by ignoring the break-up of the likelier measurement outcome. In fact this would simply force some of the $|\psi_r\rangle$'s to be equal: the induced disturbance can only be made worse. In our scenario Eve gathers information about the canonical ensemble $\{(1/n, |j\rangle)\}_{j=1\ldots n}$, for which one obtains

$$G = \frac{1}{n}\sum_r \mathrm{Tr}\big(\langle j|\hat{A}_r^\dagger \hat{A}_r|j\rangle\, |j\rangle\langle j|\psi_r\rangle\langle\psi_r|\big)$$

Clearly Eve's optimal guess knowing outcome r occurred is $|j_{(r)}\rangle$ such that $\langle j_{(r)}|\hat{A}_r^\dagger \hat{A}_r|j_{(r)}\rangle = \max_j \langle j|\hat{A}_r^\dagger \hat{A}_r|j\rangle$.
As a consequence

$$\begin{aligned} G &= \frac{1}{n}\sum_r \langle j_{(r)}|\hat{A}_r^\dagger \hat{A}_r|j_{(r)}\rangle \\ &= \frac{1}{n}\sum_r \mathrm{Tr}\big(\hat{A}_r|j_{(r)}\rangle\langle j_{(r)}|\hat{A}_r^\dagger\big) \\ &= \frac{1}{n}\sum_r \mathrm{Tr}\big(Id\otimes|j_{(r)}\rangle\langle j_{(r)}|A_r A_r^\dagger\big) \end{aligned}$$

where we applied Proposition 7.2. This yields:

Lemma 7.1 (Estimation as a linear functional) *Let $\widehat{\$} \equiv \{\hat{A}_r\}$ be a generalized measurements with best guess $|j_{(r)}\rangle$, and $\$ \equiv \{A_r\}$ its corresponding quantum state.*
Further let

$$€ = \frac{1}{n}\sum_r Id\otimes|j_{(r)}\rangle\langle j_{(r)}|\otimes|r\rangle\langle r|$$

With $j_{(r)}$ such that $\langle j_{(r)}|\hat{A}_r^\dagger \hat{A}_r|j_{(r)}\rangle = \max_j \langle j|\hat{A}_r^\dagger \hat{A}_r|j\rangle$.
Then the mean estimation fidelity of $\widehat{\$}$ with respect to the canonical ensemble is given by

$$G = \sum_r Tr\big(€\,(A_r\otimes|r\rangle)(A_r\otimes|r\rangle)^\dagger\big). \tag{7.6}$$

As we have seen the generalized measurement is equivalently described, using Isomorphism 7.1, by $\{A_r\}$, a set of non-zero non-normalized n^2-dimensional vectors. Further consider the larger vector $v = (A_{ijr})_{ijr}$, i.e. with r itself an index of the complex components. The trace-preserving condition upon the generalized measurement is easily seen to imply that $||v||^2$ should be equal to n. From Lemma 7.1 it is clear that when seeking an upper bound for G under this fixed norm constraint, we may assume v to take the form $v = (A_{jjr})_{jjr}$, because of the identity matrix on the first subsystem of €. As we shall explain in subsection 7.2.2 this can be done at no cost for the mean induced fidelity. This way we reach the following

Lemma:

Lemma 7.2 (Information) *Consider a generalized measurement* $\{\hat{A}_r\}$, $\sum_r \hat{A}_r^\dagger \hat{A}_r = Id$, $\hat{A}_r$ *diagonal for all* r, *acting upon an* n*-dimensional system. Then the mean estimation fidelity w.r.t the canonical ensemble verifies*

$$G \leq \frac{1}{n}\, g$$

with $g = \sum_r |A_{j_{(r)} j_{(r)} r}|^2$ *and* $j_{(r)}$ *such that* $|A_{j_{(r)} j_{(r)} r}|^2 = \max_j |A_{jjr}|^2$.

7.2.2 Disturbance

The notion of disturbance refers to Bob's chances of detecting Eve's alteration of the state originally sent. For this purpose Bob can, at best, project the received state upon the span of the original state. Thus the disturbance verifies $D = 1 - F$, where F is the induced fidelity.

Definition 7.4 (Induced fidelity) *The* fidelity induced *by a quantum operation* $\widehat{\$}$ *upon an ensemble* $\{(p_i, |\phi_i\rangle)\}$ *is defined by:*

$$F = \sum_i p_i\, Tr(|\phi_i\rangle\langle\phi_i| \widehat{\$}(|\phi_i\rangle\langle\phi_i|))$$

The induced fidelity is to be understood as the average fidelity between the output of the quantum operation (the $\widehat{\$}(|\phi_i\rangle\langle\phi_i|)$'s) and its input (the $|\phi_i\rangle\langle\phi_i|$'s). A straightforward application of Proposition 7.2 yields: (with * denoting componentwise complex conjugation as usual)

$$F = \sum_i p_i \mathrm{Tr}\big(\big(|\phi_i\rangle\langle\phi_i| \otimes |\phi_i^*\rangle\langle\phi_i^*|\big)\$\big) \tag{7.7}$$

In our scenario Eve is tested on the pairing ensemble $\{(1/n^2, \rho_{jk}\}_{j,k=1\ldots n}$, with $\rho_{jk} = \frac{(|j\rangle + i|k\rangle)(\langle j| - i\langle k|)}{2}$, for which one obtains:

$$\begin{aligned}
4\,\rho_{jk}\otimes\rho_{jk}^* &= |jj\rangle\langle jj| + |jj\rangle\langle kk| + i|jj\rangle\langle jk| - i|jj\rangle\langle kj| \\
&\quad + |kk\rangle\langle jj| + |kk\rangle\langle kk| + i|kk\rangle\langle jk| - i|kk\rangle\langle kj| \\
&\quad - i|jk\rangle\langle jj| - i|jk\rangle\langle kk| + |jk\rangle\langle jk| - |jk\rangle\langle kj| \\
&\quad + i|kj\rangle\langle jj| + i|kj\rangle\langle kk| - |kj\rangle\langle jk| + |kj\rangle\langle kj| \\
\rho_{jk}\otimes\rho_{jk}^* + \rho_{kj}\otimes\rho_{kj}^* &= \frac{1}{2}\big(|jj\rangle + |kk\rangle\big)\big(\langle jj| + \langle kk|\big) \\
&\quad + \frac{1}{2}\big(|jk\rangle - |kj\rangle\big)\big(\langle jk| - \langle kj|\big) \\
\sum_{jk} \rho_{jk}\otimes\rho_{jk}^* &= \frac{1}{4}\sum_{jk} \Big(\big(|jj\rangle + |kk\rangle\big)\big(\langle jj| + \langle kk|\big) + \big(|jk\rangle - |kj\rangle\big)\big(\langle jk| - \langle kj|\big)\Big)
\end{aligned} \tag{7.8}$$

We now proceed to express Equation (7.8) in terms of projectors. Regarding the subspace of repeated indices we observe that:

$$\sum_{jk}\big(|jj\rangle+|kk\rangle\big)\big(\langle jj|+\langle kk|\big) = 2\sum_{jk}\big(|jj\rangle\langle jj| + |jj\rangle\langle kk|\big)$$
$$= 2n\sum_{j}|jj\rangle\langle jj| + 2\big(\sum_{j}|jj\rangle\big)\big(\sum_{j}\langle jj|\big)$$

As regards the subspace of non-repeated indices the vectors $|jk\rangle - |kj\rangle$ are already orthogonal to each other, so long as we maintain $j < k$. Combining our newly found spectral decomposition with Equation (7.7) yields:

Lemma 7.3 (Fidelity as a linear functional) *Let $\widehat{\$}$ be a quantum operation, and $\$$ its corresponding quantum state.*
Further let

$$\pounds = \frac{1}{2n}P_{rep} + \frac{1}{2n}P_\beta P_{rep} + \frac{1}{n^2}\sum_{j<k}\big(\frac{|jk\rangle - |kj\rangle}{\sqrt{2}}\big)\big(\frac{\langle jk| - \langle kj|}{\sqrt{2}}\big)P_{nonrep}$$
$$With \quad P_{rep} = \sum_j |j\rangle\langle j|\otimes|j\rangle\langle j| \quad P_{nonrep} = \mathbb{I} - P_{rep}$$
$$|\beta\rangle = \frac{1}{\sqrt{n}}\sum_j |jj\rangle \quad and \quad P_\beta = |\beta\rangle\langle\beta|$$

Then the fidelity induced by $\widehat{\$}$ upon the pairing ensemble is given by

$$F = Tr(\pounds\ \$). \tag{7.9}$$

Using Theorem 7.1 $\$$ is positive and may be thus be written $\$ = \sum A_r A_r^\dagger$, with $\{A_r\}$ a set of non-zero non-normalized n^2 vectors. From Lemma 7.3 it is clear that, when seeking an upper bound for F, we may assume all our A_r to lie in the subspace of projector P_{rep}. In other words $A_r = \sum_j \lambda_j^r |jj\rangle$. We then have, using Lemma 7.3 still:

$$F = \frac{1}{2n}\sum_{rj}(\lambda_j^r)^2 + \frac{1}{2n^2}\sum_{rjk}\lambda_j^r\lambda_k^r$$

Corresponding to a measurement $\{\hat{A}_r\}$, $\hat{A}_r = \sum_j \lambda_j^r |j\rangle\langle j|$, $\sum_{rj}(\lambda_j^r)^2 = n$ under Equation (7.5). This way we reach the following Lemma:

Lemma 7.4 (Disturbance) *Consider a generalized measurement $\{\hat{A}_r\}$, $\sum_r \hat{A}_r^\dagger \hat{A}_r = \mathbb{I}$ acting upon an n-dimensional system. Then the disturbance induced upon the pairing ensemble verifies*

$$D \geq \frac{1}{2} - \frac{1}{2n^2}f$$

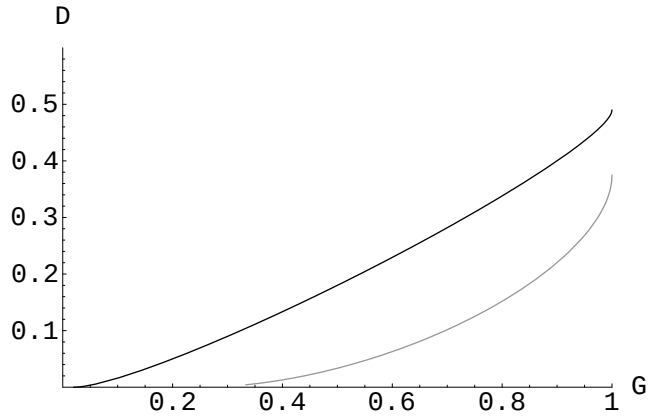

Figure 7.1: **Decoys' information gain versus disturbance.** *In grey* $n=4$*, in black* $n=50$*.*

with $f = \sum_r |\sum_{j=0}^{n-1} A_{jjr}|^2$.

7.3 Optimization and conclusion

We are now set to prove Claim 7.1. From Proposition 7.1 we immediately have

$$\frac{1}{2} - \frac{1}{2n^2} f \geq \frac{1}{2} - \frac{1}{2n^2}\big(\sqrt{g} + \sqrt{(n-1)(n-g)}\big)^2.$$

Applying Lemma 7.2 and 7.4 yields

$$D \geq \frac{1}{2} - \frac{1}{2n^2}\big(\sqrt{nG} + \sqrt{n(n-1)(1-G)}\big)^2$$

which in turn is nothing but Inequality (7.2). A plot of the curve is shown in Figure 7.3. As was the case with the continuous uniform ensemble [5] the generalized measurement family

$$\{\hat{A}_r\}, \qquad \hat{A}_r = \sqrt{G}|r\rangle\langle r| + \sqrt{\frac{1-G}{n-1}}\big(\mathbb{I} - |r\rangle\langle r|\big)$$

saturates the tradeoff for any fixed $G \in [1/n, 1]$. This may come as no surprise since the corresponding n^2 vectors A_r verify

$$A_r = \lambda|rr\rangle + \mu|\beta\rangle, \quad \lambda \equiv \sqrt{G} - \sqrt{\frac{1-G}{n-1}}, \quad \mu \equiv \sqrt{\frac{1-G}{n-1}}.$$

In other words these unit vectors $\{A_r\}$ can be thought of as superpositions of Eve's two extreme attacks: on the one hand $\lambda = 1$ yields the projective measurement $\{|r\rangle\langle r|\}$ maximizing the mean estimation fidelity, whilst on the other hand $\mu = 1$ yields the 'do nothing' measurement $\{\mathbb{I}\}$ minimizing the disturbance. Viewed from the perspective of Lemma 7.3, Eve, as she seeks to be more conservative, increases her component in the subspace of P_β.

The generalized measurement family 'measure $\{|r\rangle\langle r|\}$ with probability p else leave it alone' does not saturate the tradeoff, but linear combinations of *pure states corresponding to measurement elements* do. Stated in this simple manner, our result suggests the state-operator correspondence method developed in this paper could establish itself as a very natural procedure for deriving quantum cryptographic security bounds in general.

We believe there could be many quantum cryptographic applications of quantum decoys. The next chapter develops their use for the purpose of blind computation.

Chapter 8

Blind quantum computation

One began to hear it said that World War I was the chemists' war,
World War II was the physicists' war, and World War III (may it never come)
will be the mathematicians' war.

—Philip J. Davies

I do not know how World War III will be fought,
but I do know how World War IV will: with sticks and stones.

—Albert Einstein

We investigate the possibility of *having someone carry out the work of executing a function for you, but without letting him learn anything about your input.* Say Alice wants Bob to compute some well-known function f upon her input x, but wants to prevent Bob from learning anything about x. The situation arises for instance if client Alice has limited computational resources in comparison with mistrusted server Bob, or if x is an inherently mobile piece of data. Could there be a protocol whereby Bob is forced to compute $f(x)$ *blindly*, i.e. without observing x? We provide such a blind computation protocol for the class of functions which admit an efficient procedure to generate random input-output pairs, e.g. factorization. The setting is quantum, the security is unconditional, the eavesdropper is as malicious as can be.

In the traditional secure two-party computation scenario[65, 1] Alice has secret input x, Bob has secret input y, and both of them wish to compute $f(x, y)$. The function f is of course well-known to the two parties; the usual example is that of two millionaires who wish to compare their wealth without disclosing how much they own[65]. Most protocols for secure two-party computation are symmetric with respect to the computing power each party should carry out during the execution. In these scenarios, if Alice knew Bob's input y she could compute $f(x, y)$ on her own without having to invest more computing power. Entering a secure two-party computation together with Bob will not help in diminishing Alice's computing power needed to evaluate f. In fact, most implementations require both Alice and Bob to invest more computing power than what is needed for the mere evaluation of f.

Unlike secure two-party computation, blind computation is fundamentally asymmetric. Alice is the only party with a secret input x, Bob is the only one able to compute f. Alice wants Bob to compute $f(x)$ without him learning *too much* about x. Thus an obvious motivation for Alice to enter a blind quantum computation together with Bob is to unload the computational task of computing f without having to compromise the privacy of her input. One could easily imagine this occurring in a Grid architecture, or in any client-server relation with a mistrusted server retaining the computational power. To make things more precise, suppose there were only a handful of fully operational large-scale quantum computers in the world, and some hungry academic decided to make use of her timeshare as scientist to crack some Swiss bank's RSA private key x. The hungry academic (Alice) will surely want to keep x secret from the authorities handling the quantum computer (Bob), so that she does not get suspected when subsequent international money transfers come to top up her meager income. But there may be other reasons to enter a blind computation protocol than mere computational power asymmetry. For instance Bob may possess some trapdoor information about the otherwise well-known function f. Or perhaps x may represent some mobile agent's code which ought to be protected against the malicious host upon which it runs. Others may see blind quantum computation as a somewhat philosophical issue: Is it possible to carry out some work for someone whilst being prevented from knowing what the work consists in?

In the classical setting, blind computation has first been studied by Feigenbaum [19]. It was shown that for some functions f, an instance x can be encrypted by $z = E_k(x)$ in such a way that Alice can recover $f(x)$ efficiently from k and $f(z)$. The construction cannot be extended easily to general classes of functions. In particular, blind computation of the discrete logarithm function (DLF) was shown possible but no blind computation of the RSA factoring function (FACF) is known. Moreover Abadi, Feigenbaum, and Kilian [2] have shown that no NP-hard problem can be computed blindly unless the polynomial-time hierarchy collapses at the third level, and this seems to remain true when the privacy of Alice's input is only partial. Even when computational assumptions are invoked [52], none of the currently known

classical blind computation protocols applies to general classes of functions. Rather they take advantage of specific algebraic properties of particular functions. These constructions rely upon encryption that are, in some sense, homomorphic with respect to function f. Clearly, very natural candidates for f are not known to have this property like for FACF. It is not surprising that such stringent requirements do not necessarily hold when Bob is running a quantum computer.

In this chapter as in[19, 2], we are mainly concerned with information-theoretic security. That is, although we allow Bob to learn some Shannon information about Alice's input x, we show that if Bob gets too much information then he will be detected by Alice with high probability. Any server Bob who wants to remain in business should clearly avoid such a detection. Our goal consists in finding protocols for blind computation for which a good tradeoff between Bob's ability of being detected and the amount of Shannon information about Alice's input can be established. Almost privacy was recently studied by Klauck [36] in a two-party computation setting which differs from the asymmetric scenario imposed by blind computations. Moreover, the security was only considered with respect to passive adversaries. We want our solution to apply to a wider class of functions than the one considered in the classical setting while being resistant to all adversaries, not just passive ones. As far as we can tell, blind quantum computation has not been studied as such so far.

In Section 8.1 we present the basic ideas of our blind quantum computation protocol, as well as the reasons which limit their use to a certain class of function. In Section 8.2 we review and adapt a recent result in the information versus disturbance tradeoff literature. In Section 8.3 we formalize the protocol and give a proof of its security. We conclude in Section 8.4 and mention possible extensions.

8.1 Principles of a solution

Let us now explain the basic principles underlying our blind quantum computation protocol. Suppose Alice wants Bob to compute $f(x)$ whilst keeping x secret. Moreover suppose Bob possesses a quantum computer which implements f, i.e. he is able to implement a unitary transform U such that $U|q\rangle = |q\rangle|f(q)\rangle$ for all input q. In order to achieve her purpose Alice could hide her true input $|x\rangle$ amongst superpositions of other potential inputs $\frac{|q\rangle+i|q'\rangle}{\sqrt{2}}$ and send all this to Bob so that he executes U. Now if Bob attempts a measure so as to determine $|x\rangle$ he will run the risk of collapsing the superpositions. Alice may detect such a tampering when she retrieves her results. The above suggestion has a weakness however: Alice is not returned $\frac{|q\rangle+i|q'\rangle}{\sqrt{2}}$, but

$$U\frac{|q\rangle + i|q'\rangle}{\sqrt{2}} = \frac{|q; f(q)\rangle + i|q'; f(q')\rangle}{\sqrt{2}},$$

the result of Bob's computation upon the superposition Alice had sent. Since Alice does not want to compute f herself she is in general unable to check upon the integrity of such states.

There are many computational problems, however, where this need not be a problem. For example say f takes composite numbers into the list of their integer factors. Then Alice can easily (at the cost of a few multiplications) prepare several input-output pairs $\{(q, f(q))\}$. Thus if Alice hides her true input $|x\rangle$ amongst superpositions $\frac{|q\rangle+i|q'\rangle}{\sqrt{2}}$ generated in this manner, she will later be able to check whether $\frac{|q;f(q)\rangle+i|q';f(q')\rangle}{\sqrt{2}}$ are indeed being returned. Formally the idealized class of functions for which our protocol will work is defined as follows:

Definition 8.1 (Random verifiable functions) *Let S and S' denote two finite sets. A function $f : S \to S'$ is* random verifiable *if and only if there exists, for all N, an efficient probabilistic process which generates N input-output pairs $\{(q, f(q))\}$ and such that the inputs (the q's) are uniformly distributed in S.*

There are several promised problems for which we can define functions that are random verifiable. Consider the language RSA-composite which contains natural numbers of a fixed size that can be expressed by the product of two primes of the same size. The function f that returns the prime factors is also random verifiable. In this case, f can be computed efficiently on a quantum computer but not, as far as we know, on a classical computer. Another example can be obtained from the graph isomorphism problem. Let $L_{e,v}$ be the set of all pairs of isomorphic graphs with e edges and v vertices. We define function $f : L_{e,v} \mapsto S_e$, where S_e is the set of all permutations among v elements, as $f(G_0, G_1) = \sigma$ such that $\sigma(G_0) = G_1$. It is easy to verify that f is random verifiable. The following efficient classical computation does the job:

- Pick a random permutation $\sigma \in S_e$,
- Generate a random graph G_0 with e edges and v vertices,
- Output $((G_0, \sigma(G_0)), \sigma)$.

Although f is random verifiable by an efficient classical algorithm, it is not known whether even a quantum computer can evaluate f efficiently.

In this chapter, we provide a blind quantum computation protocol for random verifiable functions together with a thorough security analysis. The security is unconditional, it is expressed in information theoretical terms and relies upon the laws of physics only. As was hinted in this section our analysis will crucially depend upon the tradeoff between Bob's information gain about Alice's true input (a canonical basis state) and the disturbance he induces upon superpositions of potential inputs (pairwise superpositions of canonical basis states).

8.2 Information gain versus disturbance tradeoff

In order to construct a blind quantum computation protocol we needed to quantify the disturbance upon pairwise superpositions of n-dimensional canonical basis states, as induced when Bob seeks to learn information about the canonical basis. This is precisely what was achieved with Claim 7.1 of the preceding chapter. For convenience we now recall these results, rephrasing them in terms of induced fidelity and modifying the name of the parties to suit the present scenario.

Scenario 8.1 (Quantum decoys) *Consider a quantum channel for transmitting n-dimensional systems having canonical orthonormal basis $\{|j\rangle\}$.*
Suppose Alice's message words are drawn from the canonical ensemble $\{(1/n, |j\rangle)\}_{j=1..n}$, whilst her quantum decoys *are drawn from the pairing ensemble $\{(1/n^2, \rho_{jk}\}_{j,k=1...n}$, with $\rho_{jk} = \frac{(|j\rangle+i|k\rangle)(\langle j|-i\langle k|)}{2}$. Alice sends, over the quantum channel, either a message word or a decoy, which she later retrieves.*
Whenever she sends a quantum decoy $\frac{|j\rangle+i|k\rangle}{\sqrt{2}}$ she later measures the retrieved system with $\{P_{intact} = \left(\frac{|j\rangle+i|k\rangle}{\sqrt{2}}\right)\left(\frac{\langle j|-i\langle k|}{\sqrt{2}}\right), P_{tamper} = \mathbb{I} - P_{intact}\}$ so as to check for tampering.
Suppose Bob is eavesdropping the quantum channel, and has an interest in determining Alice's message words.

Proposition 8.1 (Quantum decoys' security) *Referring to Scenario 8.1, suppose Bob performs an individual attack such that, whenever a message word gets sent, he is able to identify which with probability G (mean estimation fidelity).*
Then, whenever a quantum decoy gets sent, the probability F (induced fidelity) of Bob's tampering not being detected by Alice is bounded above under the following tight inequality:

$$F \leq \frac{1}{2} + \frac{1}{2n}\Big(\sqrt{G} + \sqrt{(n-1)(1-G)}\Big)^2 \tag{8.1}$$

For optimal attacks G varies from $\frac{1}{n}$ to 1 as F varies from 1 to $\frac{1}{2}+\frac{1}{2n}$.

In Chapter 7 Proposition 8.1 was proven at the level of each individual transmission. In other words every time Alice sends one decoy, and later retrieves it, Bob's chances of not being detected are less than $F(G)$, as given by the right-hand-side of Equation (8.1). Here G stands for the mean estimation fidelity Bob would gain if a message word had been sent instead, and this was later announced to him.
However let us now imagine that scenario 8.1 gets repeated N times round, and Alice happens to send only decoys. In general Proposition 8.1 does not allow us to deduce Bob's chances of not being detected at all, because the probabilities $\{p(\text{Bob passing round } i) = F(G_i)\}_{i=1..N}$ may not be independent. Here G_i stands for the mean estimation fidelity Bob would gain if a message word had been sent at round i instead.
However if Scenario 8.1 is repeated conditionally upon Alice's measurement outcome we can no longer have such correlations:

Lemma 8.1 (Interactive decoys' security) *Referring to Scenario 8.1 suppose*

Step 0. Alice prepares a pool of $N+1$ quantum states consisting of one message word together with N quantum decoys.
Step 1. Alice sends Bob one quantum state drawn at random amongst those remaining in the pool.
Step 2. Alice awaits to retrieve the quantum state she sent.
Step 3. If Alice sent a quantum decoy she measures the retrieved system so as to check for tampering. Whenever she detects such a tampering she stops.
Step 4. If the pool is empty Alice stops the protocol, else she proceeds again with Step 1.

Say the message word was sent at position p. Then the probability of Bob reaching round m ($1 < m \leq N+1$) is bounded above under the following tight inequality:

$$p(\textit{Bob reaches } m) \leq \prod_{i=1, i\neq p}^{m-1} F(G_i) \tag{8.2}$$

where G_i stands for Bob's mean estimation fidelity, if p is later announced equal to i, about the message word sent at round i.

Proof: Suppose Bob's attack is constituted of successive measurements each yielding him G_i − where G_i stands for Bob's mean estimation fidelity, if p is afterwards announced equal to i, about the message word sent at round i. We can assume Bob's optimal strategy is of this kind since he cannot distinguish a message word from a quantum decoy until he gets caught or the pool is empty (because both canonical ensemble and the pairing ensemble have density matrix $\mathbb{I}/n$). We proceed to prove the Lemma by contradiction.
Suppose $p(\text{Bob reaches } m) > \prod_{i=1,i\neq p}^{m-1} F(G_i)$. Then there exists a k for which

$$p(\text{Bob reaches } k) \leq \prod_{i=1,i\neq p}^{k-1} F(G_i) \quad \text{and} \quad p(\text{Bob reaches } k{+}1) > \prod_{i=1,i\neq p}^{k} F(G_i).$$

For such a k we thus have

$$p(\text{Bob reaches } k{+}1|\text{Bob reaches } k) > F(G_k). \tag{8.3}$$

In other words Bob, on the k^{th} round, is capable of collecting mean estimation G_k about a message word whilst remaining undetected with probability more than $F(G_k)$ upon a quantum decoy, which is impossible. For instance say Bob now simulates the scenario in Lemma 8.1 until such a round k is reached, and then enters Scenario 8.1 with Alice. If Alice sends a message word Bob obtains mean estimation fidelity G_k. But if Alice sends a quantum decoy, Bob, according to Equation (8.3), remains undetected with probability strictly superior to $F(G_k)$. This contradicts Proposition 8.1. □

Lemma 8.2 (Concavity of circular products) *Consider $f : [0,1] \to [0,1]$ a concave, continuous function and $\{x_i\}_{i=1\ldots N+1}$ a set of real numbers in the interval $[0,1]$.*
Suppose the sum $t = \sum_{i=1}^{N+1} x_i$ is fixed. We have

$$\frac{1}{N+1}\sum_{p=1}^{N+1}\Big(\prod_{i=1,i\neq p}^{i=N+1} f(x_i)\Big) \leq f\big(\frac{t}{N+1}\big)^N.$$

Proof: By definition of concavity one has

$$\frac{1}{2}\big(f(x_1)+f(x_2)\big) \leq f\big(\frac{x_1+x_2}{2}\big) \tag{8.4}$$

$$\text{and} \quad f(x_1)f(x_2) \leq f\big(\frac{x_1+x_2}{2}\big)^2, \tag{8.5}$$

where the latter equation trivially derives from $f(x_1)f(x_2) \leq \big(\frac{f(x_1)+f(x_2)}{2}\big)^2$. Let us now show that

$$\frac{1}{N+1}\sum_{p=1}^{N+1}\prod_{i=1,i\neq p}^{i=N+1} f(x_i) \leq \frac{1}{N+1}\sum_{p=1}^{N+1}\prod_{i=1,i\neq p}^{i=N+1} f(y_i), \tag{8.6}$$

where $y_1 = y_2 = \frac{x_1+x_2}{2}$ and $y_i = x_i$ for $i = 3 \dots N+1$. This result is in fact obtained by combining (summing) two inequalities:

$$\big(f(x_1)+f(x_2)\big)\prod_{i=3}^{N+1} f(x_i) \leq \big(f(y_1)+f(y_2)\big)\prod_{i=3}^{N+1} f(y_i)$$
$$f(x_1)f(x_2)\sum_{p=3}^{N+1}\prod_{i=3,i\neq p}^{i=N+1} f(x_i) \leq f(y_1)f(y_2)\sum_{p=3}^{N+1}\prod_{i=3,i\neq p}^{i=N+1} f(y_i)$$

where former stems from Equation (8.4) and $f(x)$ positive, whilst the latter stems from Equation (8.5) and $f(x)$ positive.

Equation (8.6) expresses the fact that, whenever two elements x_i and x_j, $i \neq j$ are replaced by their mean, the value of

$$\pi(\underline{x}) \equiv \frac{1}{N+1}\sum_{p=1}^{N+1}\Big(\prod_{i=1,i\neq p}^{i=N+1} f(x_i)\Big)$$

is increased. Now let us define $\{\underline{x}^{(k)}\}$ a sequence of vectors such that $\underline{x}^{(1)} = (x_1, x_2, \dots, x_{N+1})$, and $\underline{x}^{(k)}$ is formed from $\underline{x}^{(k-1)}$ by replacing both the largest and the smallest component by their mean. As k goes to infinity this sequence of vectors tends to $\underline{x}^{(\infty)} = (\frac{t}{N+1}, \frac{t}{N+1}, \dots)$. By Equation (8.6) we have $\{\pi(\underline{x}^{(k)})\}$ an increasing sequence of real numbers. As k goes to infinity, and since $\pi(\underline{x})$ is continuous in $\underline{x}$, this sequence of real numbers tends to

$$\pi(\underline{x}^{(\infty)}) = f\big(\frac{t}{N+1}\big)^N.$$

This limit must therefore provide, for all $\underline{x}$ having components summing to t, a tight upper bound on the value of $\pi(\underline{x})$. □

8.3 Protocol and Security

We are now set to give our blind quantum computation protocol:

Protocol 8.1 (Interactive version) *Alice wants Bob to compute $f(x)$ whilst keeping her input x secret. Here f designates a random verifiable function implemented on a quantum computer by a unitary evolution U.*

Step 0. Alice efficiently computes $2N$ random input-solution pairs $(q, f(q))$ and prepares a pool of $N+1$ quantum states consisting of her true input $|x\rangle$ together with N quantum decoys $\frac{|q\rangle+i|q'\rangle}{\sqrt{2}}$ (or just $|q\rangle$ if q happens to be equal to q').

Step 1. Alice sends Bob one quantum state $|\psi\rangle$ drawn at random amongst those remaining in the pool.

Step 2. Bob supposedly computes $U|\psi\rangle$ and sends the result back to Alice.
Step 3. If $|\psi\rangle$ was a quantum decoy Alice measures the retrieved system with

$$\{P_{intact} = \frac{1}{2}\big(|q\, f(q)\rangle + i|q' f(q')\rangle\big)\big(\langle q\, f(q)| - i\langle q' f(q')|\big)\ ;\ P_{tamper} = \mathbb{I} - P_{intact}\},$$

so as to check for tampering. Whenever she detects such a tampering she stops. If on the other hand $|\psi\rangle$ was her true input Alice reads off $f(x)$.
Step 4. If the pool is empty Alice stops the protocol, else she proceeds again with Step 1.

The security of this protocol is rigorously described through the following claim.

Claim 8.1 (Statement of security) *Referring to Protocol 8.1 suppose Bob has no a priori information about Alice's true input x. Moreover suppose Bob performs an attack such that his mutual information about Alice's true input x now verifies*

$$I \leq \log(n) + \log(G)\,, \qquad G \in [\frac{1}{n}, 1]$$

Then the probability D (induced disturbance) of Alice detecting Bob's tampering is bounded below under the following inequality:

$$D \geq 1 - F(G)^N$$

Proof: We prove that the claim holds for a weakened form of Protocol 8.1, where we add (whenever Alice used to stop the protocol):
Step 5. Alice publicly announces the position in which she sent her true input $|x\rangle$, if she did.
Until this stage, however, Bob has no means of knowing at which round true input $|x\rangle$ was sent. This is because we have assumed he has no a priori knowledge about the true input. In his view the state was drawn from the canonical ensemble $\{(1/n, |j\rangle)\}_{j=1..n}$, whilst the quantum decoys were drawn from the pairing ensemble $\{(1/n^2, \rho_{jk}\}_{j,k=1...n}$ (with $\rho_{jk} = \frac{(|j\rangle+i|k\rangle)(\langle j|-i\langle k|)}{2}$), but the two are undistinguishable for they both have density matrix $\mathbb{I}/n$. We are, therefore, in the precise case of Lemma 8.1. Without loss of generality we can assume Bob's attack yields him mean estimation fidelity G_i about Alice's true input whenever the position is later announced equal to i.

First we give an upper bound upon Bob's mutual information I about true input x, expressed in terms of $G_T = \sum_p G_p$. Say the true input is at position p and suppose, on the one hand, that Bob's tampering is undetected by Alice. In this situation Bob's best chance of guessing the true input is G_p (by definition) and thus his Shannon uncertainty H_p about

Alice's true input is bounded as follows

$$\begin{aligned} H_p &\equiv \sum -p(x|\text{Bob's outcome}) \log(p(x|\text{Bob's outcome})) \\ &\geq -\lfloor \frac{1}{G_p} \rfloor G_p \log(G_p) - (1 - \lfloor \frac{1}{G_p} \rfloor G_p) \log(1 - \lfloor \frac{1}{G_p} \rfloor G_p) \\ &\geq -\log(G_p). \end{aligned}$$

The RHS of the last line is often referred to as the 'min-entropy' sometimes denoted H_∞ and is commonly used to bound uncertainties in the above manner (i.e. Shannon uncertainty is always at least H_∞). As a consequence Bob's mutual information I_p verifies

$$I_p \leq \log(n) + \log(G_p). \tag{8.7}$$

Suppose, on the other hand, that Bob's tampering is detected by Alice at round m. In this situation Bob may or may not have accessed the true input, and thus his uncertainty about Alice's true input is bounded as follows

$$(m < p) \quad H_p \geq \log(n)\ ; \quad (m \geq p) \quad H_p \geq -\log(G_p)$$

As a consequence Bob's mutual information still verifies Inequality (8.7). Averaging over all possible positions $p = 1 \ldots N+1$ Bob's mutual information verifies

$$\begin{aligned} I &= \sum_{p=1}^{N+1} \frac{1}{N+1} I_p \leq \log(n) + \sum_{p=1}^{N+1} \frac{1}{N+1} \log(G_p) \\ &\leq \log(n) + \log(\frac{G_T}{N+1}) \end{aligned}$$

where the second line was obtained using the concavity of $x \mapsto \log(x)$.

Second we give an upper bound upon the probability $p(\text{undetected})$ that Bob's tampering remains undetected by Alice, again expressed in terms of G_T. Lemma 8.1 ensures that an optimal strategy for Bob is to make a series of individual, independent measurement as in Proposition 8.1, for this would saturate Inequality 8.2 whatever the position of the true input. Suppose Bob adopts this strategy and say the true input is at position p. According to Lemma 8.1, Bob is undetected with probability

$$p(\text{undetected}|p) \leq \prod_{i=1, i\neq p}^{N+1} F(G_i).$$

Let us now average the above over all possible positions $p = 1 \dots N+1$. The probability that Bob's tampering remains undetected by Alice verifies

$$p(\text{undetected}) = \frac{1}{N+1}\sum_{p=1}^{N+1} p(\text{undetected}|p) \leq \frac{1}{N+1}\sum_{p=1}^{N+1}\big(\prod_{i=1,i\neq p}^{N+1} F(G_i)\big)$$
$$\leq F\big(\frac{G_T}{N+1}\big)^N$$

where the last line was obtained using Lemma 8.2 upon the concave, continuous function $x \mapsto F(x)$.

By letting $G = \frac{G_T}{N+1}$ we recover our claim. □

Protocol 8.1 requires $N+1$ communications between Alice and Bob. One could suggest a modification whereby Alice would send Bob her whole pool (as prepared in *Step 0*), and later proceed to check upon the integrity of each element of the pool which Bob returns, apart from her true input. Formally this yields the following protocol:

Protocol 8.2 (Non-interactive version) *Alice wants Bob to compute $f(x)$ whilst keeping her input x secret. Here f designates a random verifiable function implemented on a quantum computer by a unitary evolution U.*

Step 0. Alice efficiently computes $2N$ random input-solution pairs $(q, f(q))$ and prepares a pool of $N+1$ quantum states consisting of her true input $|x\rangle$ together with N quantum decoys $\frac{|q\rangle+i|q'\rangle}{\sqrt{2}}$ (or simply $|q\rangle$ if q happens to be equal to q').

Step 1. Alice sends Bob the large quantum state $\bigotimes_{i=1}^{N+1}|\psi_i\rangle$ constituted of a random permutation of all elements of the pool.

Step 2. Bob supposedly computes $\bigotimes_{i=1}^{N+1} U|\psi_i\rangle$ and sends the result back to Alice.

Step 3. For each location i, if $|\psi_i\rangle$ was a quantum decoy Alice measures

$$\big\{P_{intact} = \frac{1}{2}\big(|q; f(q)\rangle + i|q'; f(q')\rangle\big)\big(\langle q; f(q)| - i\langle q'; f(q')|\big)\ ;\ P_{tamper} = \mathbb{I} - P_{intact}\big\}.$$

so as to check for tampering. If on the other hand $|\psi_i\rangle$ was her true input Alice reads off $f(x)$.

When Bob is restricted to individual attacks (non-coherent attacks, i.e. Bob measures each quantum state in the pool individually) then Claim 8.1 holds also for Protocol 8.2. We omit the proof of this since it is similar, and in fact simpler than the one given for Protocol 8.1. This is largely because we need not rely upon Lemma 8.1 when assuming individual attacks: the probabilities $F(G_i)$ of Bob's tampering not being detected by Alice as she checks upon location i are independent of each other by definition in this case.

8.4 Concluding remarks

We have investigated the possibility of *having someone else carrying out the evaluation of a function for you without letting him learn anything about your input.* We gave a blind computation protocol for the class of functions which admit an efficient procedure to generate random input-output pairs. The protocol relies upon quantum physical information gain versus disturbance tradeoffs to achieve unconditional security against the most general attack: whenever the server gathers $\log(n) + \log(G)$ bits of Shannon information about the input, he must get caught with probability at least $1 - F(G)^N$ (where n denotes the size of the input and N is a security parameter). Moreover the server cannot distinguish a weary client who uses the blind computation protocol (sending one true input amongst N decoys) from a normal client who simply makes repeated use of the server (sending $N + 1$ true inputs). Thus if the server wanted to deny his services to suspected users of the protocol, he would also have to refuse the normal clients.

Our protocol could be improved in several directions.

In terms of costs one may hope to reduce the set of quantum gates needed by Alice to prepare her transmissions [12]; lower the size of the transmissions; lower the number or rounds required. We leave it an open problem to find the security properties of the non-interactive version of our protocol when Bob is allowed coherent attacks.

In terms of functionality one may wish to achieve tamper prevention rather than tamper detection (Protocol 8.1 provides this to some degree but was not analyzed as such) or to extend the class of functions admitting a blind computation protocol (identifying such a class might have consequences in complexity theory [2]).

Conclusion

Chapter 9

Achievements and Further research

I am satisfied with my life in the past years.
I have kept my good temper and do not take myself, nor others, too seriously.
—Albert Einstein

This conclusion interprets our main results in a more discursive and speculative manner than was done previously. For this purpose the once cohesive thesis is separated into four main themes of investigation, all of which may follow their own path in future work.

9.1 On quantum theory with real vector spaces

Geometrical representations constitute a privileged manner in which to provide intuition about a theory. When phrasing axioms and concepts in terms of real vectors, these are brought one step closer to our human experience of space and time. For some theories this can be done without sacrificing mathematical rigour. For quantum theory, so central to modern physics and yet so hard to comprehend in its phenomena, the motivation to do so remains enormous.

Achievements. One of the most intriguing features of quantum theory is, perhaps, the fact that measurements must modify the state of the observed system. In this thesis we have extended the most perfected available representation of quantum states and quantum operations (the generalized Bloch sphere) to encompass this 'post-selection', 'collapse of the wavefunction', phenomenon. But because the effect of a measurement *knowing that some outcome occurred* is a non trace-preserving quantum operation (i.e. it does not in general occur with probability one on every input state), the most elegant approach for our purpose was to also allow for non-normalized quantum states (i.e. which do not necessarily have probability one). Thus in our representation quantum states map into a subcone of a Minkowski cone in $\mathbb{E}^{1,d^2-1,1}$, whose vertical cross-sections are nothing but generalized Bloch spheres. This conal representation has, it turns out, several desirable properties. Pure states map into light-like vectors, unitary

operations correspond to orthogonal transforms about the axis, and positive operations are represented by a subset of the real symmetric positive matrices. The latter can also be drawn in the cone, thus enabling us to represent the measurement element themselves.
In the case of a qubit we provided explicit formulae for the coordinates of a state after a non trace-preserving quantum operation, or for the scalar product of two post-measurement states. Moreover this four-vector representation of two-dimensional quantum systems took a whole new meaning when we realized that each measurement element acts proportionally to a special relativistic Lorentz transformation in Minkowski space. The rescaling introduced turns out to bring null boosts to finite linear maps in a natural and unifying manner. Thus we have successfully provided a rigorous space-time analogue to qubit quantum mechanics.

Further research. One may argue there are not so many of us with hands-on experience of special relativity, so the question whether the correspondence between 'observation of a quantum system' and 'special relativistic change of inertial frame' is truly a simplifying one, could be labelled a matter of taste. Even so the correspondence may turn out to be a unifying one. Quantum field theories successfully unify quantum theory and special relativity, through the use of faithful unitary representations of Lorentz transformations plus translations upon the set of pure quantum states allowed by the theory. Our correspondence suggests a physically different approach instead, whereby Lorentz transformations could act non-unitarily upon the set of non-normalized mixed states. These ideas deserve further investigation. Moreover, even though our results suggest a complete real vector space formulation of quantum mechanics is possible along these lines, much remains to be learnt about the geometry of n-dimensional quantum states.

9.2 On quantum operations as quantum states

It is fair to say that the Jamiolkowski-Choi correspondence between quantum operations (from $\mathrm{Herm}_n^+(\mathbb{C})$ to $\mathrm{Herm}_m^+(\mathbb{C})$), and quantum states (elements of $\mathrm{Herm}_{mn}^+(\mathbb{C})$), had so far been reserved to only a handful of uses. The most remarkable of these was the provision of a simplified proof to the Kraus operator sum representation theorem, thereby obtained as a mere quantum operation equivalent of spectral decomposition upon states. But could it possibly be the case that all properties of quantum states have an elegant translation in terms of quantum operations? Or inversely, instead of calculating with quantum operations and investigating their properties, could it possibly be simpler to work upon their quantum state equivalents?

Achievements. In this thesis we have highlighted the central, transversal role of the state-operator isomorphism in various issues of quantum information theory, by rederiving all the main properties of quantum operations from those of quantum states. Persisting with this

translation work we then provided two triangular decompositions for pure states of a bipartite system, and two original tests on Completely Positive-preserving maps: one for extremality in the set of Trace-preserving operations, the other regarding the factorizability or single operator decomposition. These are particularly interesting in the sense that they do not depend on the operator sum decompositions of these maps.
We also endowed quantum states with a semi-group structure stemming from the composition law on quantum operations. The composition law defines a group when restricted to the set of totally entangled (pure) states, and yields a group isomorphism between maximally entangled (pure) states and $SU(n)$. Similarly we showed that the set of quantum operations is stable under component-wise product.
Finally we provided a number of useful formulae arising from the state-operator correspondence. One of these will simplify those many mathematical problems in quantum cryptography which require a careful optimization of the fidelities induced by a quantum operation, as was illustrated later when tackling information gain versus disturbance tradeoffs.

Further research. The state-operator correspondence may be just a useful one, or one that also carries a physical meaning. Consider for instance a large state \$, its corresponding quantum operation $\widehat{\$}$, and a small state $\rho = (M^\dagger M)^t$. One has $\widehat{\$}(\rho) = \mathrm{Tr}_2\big((Id_m \otimes M)\$(Id_m \otimes M^\dagger)\big)$, which seems to say a lot if one stares at it long enough.

9.3 On information gain versus disturbance tradeoffs

Quantum measurements modify the state of the observed system. This is a commonplace about quantum theory, and this is the basic principle upon which quantum cryptography relies in order to detect the malevolent eavesdropper. Yet in spite of its central role, the tradeoff between how much information can be gained about a quantum system, and how much disturbance this may cause to the quantum system, has only rarely been quantified. These problems were judged difficult, and no general method was available.

Achievements. In this thesis we began by recovering geometrically the formula given by Fuchs and Peres for the information gain versus disturbance, as it arises when attempting to distinguish two non-orthogonal equiprobable quantum states. This would not have been achieved without a representation of non trace-preserving quantum operations, such as the one earlier developed.
Having built our intuition in this manner, we proceeded to obtain the information gain versus disturbance tradeoff in a more elaborate scenario. Suppose Alice interleaves pairwise superpositions $\{\frac{|j\rangle+i|k\rangle}{\sqrt{2}}\}$ at random amongst her otherwise classical message words $\{|j\rangle\}_{j=1\ldots n}$. Moreover suppose Eve performs an individual attack such that, whenever a canonical basis state (a message word) is sent, she is able to identify which with probability G. Making use

of a general formula inspired by the state-operator correspondence we derived a tight lower bound upon $D(G)$, the disturbance she causes whenever a pairwise superposition (a quantum decoy) is sent.

Further research. The last method we developed seems applicable to a myriad of different information gain versus disturbance tradeoff scenarios, and one should of course do so whenever quantum cryptographic applications are in sight. At the more fundamental level our method suggests one can construct optimal measurement families by: first finding the two extreme attacks; second working out their corresponding state; third constructing superpositions of these states; fourth working out the attacks which correspond to these superpositions. Such a procedure would be quite elegant, and deserves further investigation. Finally let us note that our 'quantum decoys' information gain versus disturbance tradeoff could have numerous cryptographic applications other than the one next described. Since pairwise superpositions are undistinguishable from classical message words, the former may be used to secure the latter against tampering in a variety of situations.

9.4 On blind quantum computation

Following the rise of quantum cryptography many researchers have sought to design unconditionally secure two-party computation protocol. It was soon realized, however, that quantum theory alone would not provide unconditionally secure bit commitment and oblivious transfer. Faced with this elementary fact quantum cryptographers went on to investigate different assumptions and various specific multi-party computation scenarios. And, whilst all this was happening, quantum physicists continued their long complaint on how environment interaction was the cause of an unforgiving noise, a noise which would prevent humankind from playing with a quantum computer in the next thirty years, at least. But since environment interaction (such as eavesdropping), is such an obstacle to quantum computation, why then is it not just natural that the curious player could be brought to disadvantage in some secure computation-related protocols?

Achievements. In this thesis we considered an asymmetric variant of secure two-party computation, in which Alice wants Bob to compute some well-known function f upon her input x, whilst preventing him from learning too much about x. Due to their difficulty, results for such 'blind computation' problems are extremely scarce in the classical setting. There lies the first reason why, we suspect, such scenarios have so far remained untouched by quantum cryptographers. A second reason may be that solutions to these problems require cracking some complicated information gain versus disturbance tradeoff. At least this was certainly the case with the blind quantum computation protocol we constructed. Our protocol achieves unconditionally secure blind quantum computation against malicious eavesdroppers, for the

whole class of functions admitting an efficient procedure to generate random input-output pairs. It provides a strong and natural sense in which curious players can be detected in a secure computation-related protocol.

Further research. Our protocol may perhaps be improved in several directions. As regards costs one may attempt to diminish the set of quantum gates required on Alice side, or reduce the number and size of the transmissions. As regards functionalities one may seek to achieve tamper prevention rather than tamper detection. It seems an interesting problem also to try and determine the class of functions admitting a blind computation protocol. The question could serve quantum complexity theory if one was to hope for a connection between how easy a function is to secure, and how easy it is to compute. There are several such connections in the classical setting.

Although quantum theory is one hundred years old, its information theoretical perspective is much more recent. The problems raised by this young field do not have, as yet, systematic methodologies to turn towards. Thus to provide a rigorous proof of security for a slightly original quantum cryptographic protocol requires plunging into the abyss of the linear algebra which surround quantum theory. We did so and returned with general methods for quantum information theory, possible implications in theoretical physics, as well as the mathematical proof of a blind quantum computation protocol.

Appendix A

Notation

La véritable éloquence consiste à dire tout ce qu'il faut et à ne dire que ce qu'il faut.
—La Rochefoucauld

A.1 Common formalism

$a+ib$ the complex number (a, b) (i.e. with $i^2 = -1$)
^{t}, *, †, $\otimes$ transpose, conjugate, dagger, tensor product (see Chapter 2 and [48]).

$\mathbb{R}^d$ set of $1 \times d$ matrices of real numbers (also called 'real vectors').
$\mathbb{C}^d$ set of $1 \times d$ matrices of complex numbers (also called 'vectors' or 'kets').
$u, v, w, |i\rangle, |j\rangle, |\psi\rangle, |\phi\rangle$ elements of $\mathbb{C}^d$.
$|i\rangle \otimes |j\rangle = |i\rangle|j\rangle = |ij\rangle$ element of $\mathbb{C}^{dd'}$
$(\mathbb{C}^d)^\dagger$ set of $d \times 1$ matrices of complex numbers (also called 'columns' or 'bras').
$v^\dagger, \langle i| = |i\rangle^\dagger, \langle\psi| = |\psi\rangle^\dagger$ elements of $(\mathbb{C}^d)^\dagger$.

$M_d(\mathbb{C})$ set of $d \times d$ matrices of complex numbers.
H, M, U elements of $M_d(\mathbb{C})$.
$\mathrm{Herm}_d(\mathbb{C})$ set of hermitian $d \times d$ matrices of complex numbers, i.e. such that $M^\dagger = M$.
$\mathrm{Herm}_d^+(\mathbb{C})$ set of positive $d \times d$ matrices, i.e. such that for all v in $\mathbb{C}^d$, $vMv^\dagger \geq 0$. Also referred to as the set of (non-normalized) quantum states.
$\rho, E, |M| = \sqrt{M^\dagger M}$ elements of $\mathrm{Herm}_d^+(\mathbb{C})$.
$\mathrm{Herm}_{mn}^S(\mathbb{C})$ set of separable states of a mn-dimensional quantum system i.e. of the form $\rho = \sum_x \lambda_x \rho_1^x \otimes \rho_2^x$, where $\lambda_x \geq 0$ and the ρ_1^x and ρ_2^x belong to $\mathrm{Herm}_m^+(\mathbb{C})$ and $\mathrm{Herm}_n^+(\mathbb{C})$ respectively.
$\mathbb{I}_d$ the $d \times d$ identity matrix.
E_{ij} the matrix with ij-component one, and zero elsewhere.

$\mathrm{Tr}(M)$ trace of M, i.e. the sum of its diagonal elements $M_{11} + M_{22} + \ldots$.
$\mathrm{Tr}_x(M)$ partial trace of M, removing subsystem x (see [48]).
$\det(M)$ determinant of M.
Ad_M adjoint operation of M, i.e. the function which takes ρ into $M\rho M^\dagger$.

δ_{xy} Kronecker delta, equal to one if $x = y$, and zero otherwise.
$f \circ g$ composition of function f and g, i.e. the function such that for all x, $(f \circ g)(x) = f(g(x))$.
$a_x b_x$ *under the repeated indices summation convention* is the sum $\sum_x a_x b_x$, i.e. the value of the sum $\ldots + a_3 b_3 + a_4 b_4 + a_5 b_5 + \ldots$ over the entire range of x (this is also called the Einstein summation convention).

Hull(S) convex hull of the set S, i.e. the set generated by taking linear combinations of elements of S, weighted with coefficients in $[0, 1]$ and summing to one.

---

p(event|another) probability of the event knowing another.
$H(.)$ Shannon entropy (see [48]).
$H(.|.)$ Shannon conditional entropy.
$H(. : .), I(., .)$ Shannon mutual entropy.

A.2 Formalism specific to Chapters 3, 4 and 6

A, B elements of $M_d(\mathbb{C})$.
i, j (Latin) indices ranging from 1 to $d^2 - 1$.
μ, ν (Greek) indices ranging from 0 to $d^2 - 1$.
θ, Δ, ω angles ranging from 0 to 2π.
$\underline{v}, \underline{w}, \underline{r}, \underline{\varepsilon}$ elements of $\mathbb{R}^{d^2}$.
$\overrightarrow{v}$, $\overrightarrow{w}$, $\overrightarrow{n}$ elements of $\mathbb{R}^{d^2-1}$.

τ_0 the 2×2 identity matrix $\mathbb{I}_2$.
$\{\tau_\mu\}$ elements of $\mathrm{Herm}_d(\mathbb{C})$ verifying $\mathrm{Tr}(\tau_\mu \tau_\nu) = d\delta_{\mu\nu}$.

$\underline{A}_\mu$ the real number given by $\mathrm{Tr}(A\tau_\mu)$.
$\underline{A} = \phi(A)$ the vector $(\underline{A}_\mu)$, i.e. an element of $\mathbb{R}^{d^2}$.
$\overrightarrow{A}$ the vector $(\underline{A}_i)$, i.e. an element of $\mathbb{R}^{d^2-1}$.
$\underline{v}.\underline{w}$ scalar product (also called inner product), i.e. the real number given by the sum $\underline{v}_\mu \underline{w}_\mu$.
$\overrightarrow{v}.\overrightarrow{w}$ scalar product (also called inner product), i.e. the real number given by the sum $\overrightarrow{v}_i \overrightarrow{w}_i$.
$||\underline{v}||$ the norm of $\underline{v}$, i.e. the real number $\sqrt{\underline{v}.\underline{v}}$.
$||\overrightarrow{v}||$ the norm of $\overrightarrow{v}$, i.e. the real number $\sqrt{\overrightarrow{v}.\overrightarrow{v}}$.

$\psi(M)$ the action of M in the cone, i.e. $\psi(M) = \phi \circ Ad_M \circ \phi^{-1}$.
$R_\theta(\overrightarrow{n})$ rotation of angle θ about axis $\overrightarrow{n}$.
$L(\underline{v})$ pure Lorentz boost of velocity $\underline{v}$ ($\underline{v}$ must be an element of $\mathbb{R}^4$, see [54]).
$\eta = Diag(d-1, -1, -1, \ldots)$ the $d \times d$ matrix with diagonal entries $d-1$, -1, -1 etc. and zeros elsewhere.
$\mathbb{E}^{1,d^2-1}$ the set $\mathbb{R}^{d^2}$ implicitly endowed with Minkowski metric η.

A.3 Formalism specific to Chapters 5 and 7

$End(S \to S')$ set of linear functions from the set S into the set S' (these are called endomorphisms).
$\tau, \sigma, \kappa, \rho$ elements of $\mathrm{Herm}_d^+(\mathbb{C})$.

A, B, V elements of $\mathbb{C}^m \otimes \mathbb{C}^n$, i.e. elements of $\mathbb{C}^{mn}$.
$\hat{A}, \hat{B}, \hat{V}$ elements of $End(\mathbb{C}^n \to \mathbb{C}^m)$, i.e. $m \times n$ matrices.
$\check{A}, \check{B}, \check{V}$ elements of $End(\mathbb{C}^m \to \mathbb{C}^n)$, i.e. $n \times m$ matrices.

$\$, €$ elements of $M_{mn}(\mathbb{C})$.
$\Omega, \widehat{\$}, \widehat{€}$ elements of $End(M_n(\mathbb{C}) \to M_m(\mathbb{C}))$, i.e. linear functions taking $n \times n$ matrices to $m \times m$ matrices.
$\check{\$}, \check{€}$ elements of $End(M_m(\mathbb{C}) \to M_n(\mathbb{C}))$, i.e. linear functions taking $m \times m$ matrices to $n \times n$ matrices.

---

i, k are indices ranging from 1 to m.
j, l are indices ranging from 1 to n.
$\hat{A}_{i;j}$ is equal to the matrix element A_{ij} and is associated to the repeated indices summation convention, so that $w = \hat{A}v$ is simply written $w_i = \hat{A}_{i;j} v_j$.
$\widehat{\$}_{ik;jl}$ is equal to the matrix element $\$_{ij;kl}$ and is associated to the repeated indices summation convention, so that $\tau = \widehat{\$}(\rho)$ is simply written $\tau_{i;k} = \widehat{\$}_{ik;jl} \rho_{j;l}$.

Bibliography

[1] M. Abadi, J. Feigenbaum, *Secure circuit evaluation*, Journal of Cryptology, **2**(1), 1-12, (1990).

[2] M. Abadi, J. Feigenbaum, J. Kilian, *On hiding information from an oracle*, Journal of Computer and System Sciences, **39**(1), 21-50, (1989).

[3] A. Acin, N. Gisin, V. Scarani, *Security bounds in quantum cryptography using d-level systems*, arXiv:quant-ph/0303009.

[4] P. Arrighi, C. Patricot, *A Note on the Correspondence between Qubit Quantum Operations and Special Relativity*, J. Phys. A, **36**, L287-L296.

[5] K. Banaszek, *Information gain versus disturbance for a single qubit*, arXiv:quant-ph/0006062.

[6] H. Barnum, *Information-disturbance tradeoff in quantum measurement on the uniform ensemble and on the mutually unbiased bases*, arXiv:quant-ph/0205155.

[7] C.H. Bennett, G. Brassard, *Quantum cryptography: public-key distribution and coin tossing*, Proc. of IEEE International Conference on Computers, Systems and Signal Processing, 175-179, (1984).

[8] F. Bloch, *Nuclear induction*, Phys. Rev. **70**, 460, (1946).

[9] C. Cachin, J. Camenisch, J. Kilian, J. Muller, *One-round secure computation and secure autonomous mobile agents*, Proc. of the 27^{th} ICALP, LNCS, **1853**, 512-523, Springer, (2000).

[10] N. Cerf, *Asymmetric quantum cloning machines in any dimensions*, arXiv:quant-ph/9805024, J. Mod. Opt., **47**, 187, (2000).

[11] N.J. Cerf, M. Bourennane, A. Karlsson, N. Gisin, *Security of quantum key distribution using d-level systems*, Phys. Rev. Lett., **88**, 127902, (2002).

[12] A.M. Childs, *Secure assisted quantum computation*, tech. report MIT-CTP 3211, arXiv:quant-ph/0111046.

[13] M.D. Choi, *Completely Positive linear maps on complex matrices*, Lin. Alg. Appl., **10**, 285-290, (1975).

[14] C. Crepeau, D. Gottesman, A. Smith, *Multi-party quantum computation*, Proc. of the 34^{th} annual ACM symposium on Theory of computing, 643-652, (2002).

[15] C. Crepeau, F. Legare, L. Salvail, *How to convert the flavour of a quantum bit commitment*, Proc. of EUROCRYPT'01, LNCS, **2045**, 60-77, Springer, (2001).

[16] D. Deutsch, *Quantum theory. the Church-Turing principle and the universal quantum computer*, Proc. Roy. Soc. Lon. A 400, 97, (1985).

[17] D. Deutsch, R. Jozsa *Rapid solution of problems by quantum computation.* Proceedings of the Royal Society of London A, **439**, 553-558, (1992).

[18] A.K. Ekert, *Quantum cryptography based on Bell's theorem*, Phys. Rev. Lett., **67**, 661, (1991).

[19] J. Feigenbaum, *Encrypting problem instances, or,. . . , can you take advantage of someone without having to trust him?*, Proc. CRYPTO'85, 477-488, Springer, (1986).

[20] R.P. Feynman, *Simulating physics with computers*, Int. J. Th. Phys., **21**(6), 467-488, (1982).

[21] C. Fuchs, *Information gain vs. state disturbance in quantum theory*, arXiv:quant-ph/9611010 and C. Fuchs, A. Peres, arXiv:quant-ph/9512023.

[22] C.A. Fuchs, N. Gisin, R.B. Griffiths, C. Niu, A. Peres, *Optimal eavesdropping in quantum cryptography. I. Information bound and optimal strategy*, Phys. Rev. A, **56**, 1163-72, (1997).

[23] J.P. Gauntlett, G.W. Gibbons, C.M. Hull, P.K. Townsend, *BPS states of $D = 4$ $N = 1$ supersymmetry*, Commun. Math. Phys., **216**, 431-459, (2001).

[24] N. Gisin, *Quantum cloning without signaling*, arXiv:quant-ph/9801005, Phys. Lett. A, **242**, 1-3, (1998).

[25] D. Gottesman, *Stabilizer codes and quantum error correction*, Ph.D. thesis, Caltech, USA, (1997).

[26] L.K. Grover, *A fast quantum mechanical algorithm for database search*, Proc. of the 28^{th} Annual ACM Symposium on the Theory of Computing, 212-219, (1996).

[27] T.F. Havel, C.J. Doran, *Geometric algebra in quantum information processing*, arXiv:quant-ph/0004031.

[28] S.W. Hawking, *Breakdown of predictability in gravitational collapse.*, Phys. Rev. D., **14**(10), 2460-2473, (1976).

[29] F.T. Hioe and J.H. Eberly, *N-level coherence vector and higher conservation laws in quantum optics and quantum mechanics*, Phys. Rev. Lett., **47**(12), 838-841, (1981).

[30] R.A. Horn, C.R. Johnson, *Matrix analysis*, Cambridge University Press, (1985).

[31] M. Horodecki, P. Horodecki, R. Horodecki, *Separability of mixed quantum states: necessary and sufficient conditions* Phys. Lett., **223**, 1, (1996).

[32] A. Jamiolkowski, *Linear transformations which preserve trace and positive semidefinite operators*, Rep. Mod. Phys., **3**, 275-278, (1972).

[33] J. Kempe, *Multi-particle entanglement and its applications to cryptography*, Phys. Rev. A, **60**, 910-916, (1999).

[34] A.P.A. Kent, *Unconditionally secure bit commitment*, Phys. Rev. Lett, **83**, 1447, (1999).

[35] J. Kilian, *Founding cryptography on oblivious transfer*, Proc. of the 20^{th} ACM Symposium on the Theory of Computing, 20-31, ACM press, (1988).

[36] H. Klauck, *On quantum and approximate privacy*, Proc. of the 19^{th} Annual Symposium on Theoretical Aspects of Computer Science, LNCS, **2285**, 335-365, Springer, (2002).

[37] K. Kraus, *General state changes in quantum theory*, Annals of Physics, **64**, 311-315, (1971).

[38] K. Kraus, *States effects and operators: fundamental notions of quantum theory*, Springer Verlag, (1983).

[39] L.J. Landau, R.F. Streater, *On Birkhoff's theorem for doubly stochastic Completely Positive maps of matrix algebras*, Lin. Alg. Appl. **193**, 107, (1993).

[40] R. Landauer, *Irreversiblity and heat generation in the computing process*, IBM J. Res. Dev., 183-191, July 1961.

[41] L.B. Levitin, *Optimal quantum measurements for two pure and mixed states*, Quantum Communications and Measurement, V. P. Belavkin, O. Hirota and R. L. Hudson, eds., Plenum Press, New York, 439-448, (1995).

[42] L. B. Levitin, T. Toffoli, Z. D. Walton, *Information and distinguishability of ensembles of identical quantum states*, IQSA, (2001), arXiv:quant-ph/0112075.

[43] H. Lo, *Insecurity of quantum secure computation*, Phys. Rev. A, **56**, 1154-1162 (1997).

[44] H. Lo, H.F. Chau, *Why quantum bit commitment and ideal quantum coin tossing are impossible*, Physica D, **120**, 177-187, (1998).

[45] D. Mayers, *Unconditionally secure quantum bit commitment is impossible*, Phys. Rev. Lett. **78**, 3414-3417, (1997).

[46] D. Mayers, *The trouble with quantum bit commitment*, arxiv:quant-ph/9603015.

[47] D. Mayers, *Unconditional security in quantum cryptography*, JACM, **48**(3), 351-406, (2001).

[48] M.A. Nielsen, I.L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, (2000).

[49] A. Peres, *Separability criterion for density matrices*, Phys. Rev. Lett., **77**, 1413-1415, (1996).

[50] J. de Pillis, *Linear transformations which preserve hermitian and positive semidefinite operators*, Pacific J. Math. **23**, 129-137, (1967).

[51] M.B. Ruskai, S. Szarek, E. Werner, *An analysis of Completely-Positive Trace-Preserving maps on* 2×2 *matrices*, arXiv:quant-ph/0101003.

[52] T. Sander, C.F. Tschudin *Protecting mobile agents against malicious hosts*, Mobile Agents and Security, LNCS, **1419**, 44-61, Springer, (1998).

[53] J. Schlienz and G. Mahler, *Description of entanglement*, Phys. Rev. A, **52**(6), 4396-4405, (1995).

[54] R.U. Sexl, H.K. Urbantke, *Relativity, groups, particles*, Springer Physics, Springer Wien NewYork, (2001).

[55] P. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Sci. Statist. Comput., **26**, 1484, (1997).

[56] P. Shor, J. Preskill, *Simple proof of security of the BB84 quantum key distribution protocol*, Phys. Rev. Lett. **85**, 441-444, (2000).

[57] E.C.G Surdarshan, *Quantum measurements and dynamical maps*, in *From SU(3) to Gravity*, Ed. E. Gotsman, G. Tauber, Cambridge University Press, (1986).

[58] E.C.G. Sudarshan, A. Shaji, *Structure and paramatrization of generic stochastic maps of density matrices*, J. Phys. A, **36**, 5073-5081, (2003)

[59] P.K. Townsend: *The Jordan formulation of quantum mechanics: A review.*, Supersymmetry, Supergravity, and Related Topics, F. del Alguila, J.A. de Azcárraga and L.E. Ibañes, Singapore, World Scientific, (1985).

[60] J.L. Von Neumann, *Mathematische grundlagen der quantenmechanik*, Berlin, (1982).

[61] F. Verstraete, H. Verschelde, *On quantum channels.*, Internal Report 02-176, ESAT-SISTA, K.U.Leuven, (2002).

[62] C. Wagschal, *Topologie et analyse fonctionnelle* Hermann Editeur des Sciences et des Arts, Mthodes N 27669, (1995).

[63] R.M. Wald, *Quantum gravity and time reversibility*, Phys. Rev. D **51**, 2742-2755, (1980).

[64] S. Wiesner, *Conjugate coding*, Sigact News, **15**(1), 78-88, (1983), original manuscript written *circa* 1969.

[65] A.C. Yao, *How to generate and exchange secrets*, Proc. of the 27^{th} Annual Symposium on Foundations of Computer Science, 162-167, IEEE Computer Society Press, (1986).

[66] P. Zanardi, *A note on quantum cloning in d dimensions*, arXiv:quant-ph/9804011, Phys. Rev. A, **58**, 3484, (1998).

[67] A. Zeilinger, *A foundational principle for quantum mechanics*, Found. of Phys., **29**(4), (1999).