Eléments de théorie de l'information quantique, décohérence et codes correcteurs quantiques.

Harold Ollivier

Elements of quantum information theory, decoherence and error correcting codes

ÉCOLE POLYTECHNIQUE

Harold Ollivier
INRIA - Codes
BP 105, Domaine de Voluceau
F-78153, Le Chesnay,
France

Presented at École Polytechnique
in partial fulfillment of the
requirements for the degree of
"Docteur en mathématiques".

Committee:
President : J.-M. Steyaert
Thesis director : P. Charpin
Referees : R. Jozsa & R. Laflamme
Jury : P. Grangier & J.-P. Tillich (invited)

Contents

List of Figures

Preface

THE THEORY OF decoherence was my initial interest for modern quantum mechanics. This first encounter was promoted by W. H. Zurek at Los Alamos National Laboratory with whom I learned also quantum information. The first part of this thesis is the result of this fruitful and enjoyable collaboration. In essence, it introduces information theoretical quantities for the study of the quantum-classical transition. The first two chapters of this first part aim at giving basic notions of quantum mechanics and of the theory of decoherence. They are followed by two research articles in their almost original version ([OZ02a] in collaboration with W. H. Zurek, and [OPZ03a] in collaboration with D. Poulin and W. H. Zurek). The first one shows that an absence of quantum correlation is not only a prerequisite for decoherence, but also a possible signature. The second investigates the emergence of objective information for quantum systems. In particular, it explains that an environment monitoring a quantum system broadcasts with high sensitivity a single type of information about the system; which later acquires the status of objective information.

The second topic of this work concerns error correction for quantum information devices. This part starts with two brief introductions to the field of quantum information and of quantum error correction. About this second point, I would like to encourage the interested reader to focus first on Gottesman's PhD thesis [Got97a] for the exceptional clarity of this work concerning quantum block codes. Those two chapters are followed by an attempt toward a theory of quantum convolutional codes. Some of these results were presented in [OT03a, OT04a] with J.-P. Tillich. I would also like to mention that part of this work was accomplished while visiting R. Laflamme at Perimeter Institute thanks to D. Poulin's initiative.

Finally, my last interest concerns some implementations of quantum information processing on physical devices. Being in Paris, I had the great opportunity to work on the wonderful experiment of Haroche's group. My contributions were quite modest since those are only propositions of experiments. However, I particularly appreciated to work with P. Milman on two different topics, cloning in cavity QED [MOR03a, MOY+03a] and an implementation of the Toffoli gate [OM03a].

H. Ollivier

Waterloo, Canada
October, 2004

Acknowledgments

I would like to express my deepest gratitude to Wojciech Zurek. He triggered my interest for foundations of quantum mechanics, but also for quantum information theory at large. Thanks to his numerous invitations to Los Alamos National Laboratory, he made this initial interest become my subject of research, and gave me an exceptional entry-point into the community.

But this thesis could not have been completed without Pascale Charpin. Her curiosity for quantum computers has been the decisive element for joining INRIA at "Projet CODES". She accepted to be my advisor while letting me a complete freedom in the choice of my research projects and collaborators. She took this risk and she shall be warmly thanked for this decision.

I am also greatly indebt to David Poulin, a fantastic co-author, very keen on seeing profound connections between subjects that seem first unrelated. Our passionate discussions always brought in the end a better view of what I had to do next. His steady motivation and enthousiasm for research and for the promotion of information theory in quantum mechanics has been a wonderful encouragement.

The cavity QED proposals presented here would not have been completed without Perola Milman. I am grateful for her endless enthousiasm and patience while explaining me all the details of the ENS setting. Her quest for the "perfect" pulse sequence always improved our first trials.

I would like to thank Jean-Pierre Tillich for his persistence in thinking that generalizing convolutional codes to the quantum setting was an interesting problem. One year of intermittent discussions finally convinced me. His knowledge of classical codes and his support has since then always been key elements in achieving this generalization.

Julia Kempe and Raymond Laflamme have played an essential role through their wise guidance and their vast knowledge of quantum mechanics and quantum computing. Their encouragement and advice were crucial in the completion of many aspects of this thesis. Raymond Laflamme should also be thanked for his invitations to Perimeter Institute. I could work with David Poulin, meet some of the best researchers in quantum computing and discover Canada in winter.

To Camille Negrevergne, I would like to express my gratitude for his patience in listening and for making wise comments on very early and then hardly understandable developments of this work. He put me back on the right track more often than he will admit.

During this thesis I enjoyed many discussions with several colleagues, Robin Blume-Kohout, Diego Dalvit, Jacek Dziarmaga, Joseph Emerson, Ernesto Galvaõ, Philippe Grangier, Leonid Gurvits, Frédédric Magniez, Dominic Mayers, Mike Mosca, Jean-Michel Raimond, Miklos Santha, Rolando Somma, Lorenza Viola. Their contributions range from technical points concerning this work, to sharing their thoughts on quantum mechanics. Sometime we started a project and finished it, sometimes we simply abandoned it, but some of them are still promising and ongoing.

I would like to thank Richard Jozsa for being a member of my committee in spite of our chaotic organization.

A great deal of other persons contributed to my work in a less direct, but nonetheless crucial way. First among them, Nicolas Sendrier has kept a constant watch on where I was and what I was doing. He always carefully paid attention to my integration in the group even though my scientific interests were only partially overlapping with the traditional activity of the group. His curiosity for quantum computing has been often very comforting. I would also like to thank all my colleagues at INRIA for making the group such a pleasant place to be. Their openness to quantum computing has been truly appreciated. The administrative staff, that is to say Christelle Guiziou-Cloitre, was great. She caught many mistakes in my flight reservations, avoiding me to stay only 3 days when I was invited 2 months, and always knows who to call to solve the unsolvable.

Many friends should be thanked for dragging me out of my work or for bringing me back in : Anders, Camille, Chris, David, Eléonore, Eric, Erika, Ernesto, Gian-Luca, Guido, Isabelle, Julia, Laure, Marion, Michael, Michela, Nicolas, Patrick, Perola, Robin, Rolando, Stéphane.

Yeo-Jin "'Ingie" Chung has been my constant, trustful and wise confident during all this work. Her attention brought me the confidence that I needed to go to the end. She cannot be thanked enough by just a few words.

Last but not least, my grand-parents, parents and my brother were always associated to this adventure that started long ago. They went through hundreds of "excitation-deception" cycles. I would like to express my gratitude for their commitment in this thesis, their constant love, as well as for everything else.

This work was mainly funded by INRIA but also benefited from the support of other agencies and institutions such as: ACI Sécurité informatique, Fond National pour la Science, GdR Information quantique, DARPA, Los Alamos National Laboratory, National Science Foundation, National Security Agency, Perimeter Institute, NATO, RESQ.

H. O.

I The quantum-classical transition: an information theoretical point of view

1 Introduction to quantum mechanics

1.1 Prologue

A LOT OF attention has been drawn to quantum mechanics in the past decade. Part of this renewal is due to the discovery of unexpected computational capabilities of quantum information processors [Pre98a, NC00a, KSV02a] — devices that use quantum systems to process information. Quantum computing, as a new interdisciplinary field, motivates most of the current experimental as well as theoretical researches on quantum mechanics (for an overview of the goals and achievements of quantum information processing see [HH02a]). It also promotes the analysis of the links between quantum theory and computer science. Therefore, unlike classical[1] computer science, which does not necessitate a deep understanding of the nature of the physical principles at work in a computer, quantum information processing requires to spend some time learning quantum mechanics. Time is needed to develop an intuition about quantum mechanics.

This chapter introduces some important concepts partly in the perspective of the quantum processing of information. However, it would be misleading to think about quantum mechanics solely through this specific application. This part, devoted to the emergence of a classical world out of the quantum substrate, is an example of a more complicated picture in which quantum information processing and other areas of quantum mechanics are deeply inter-related: on the one hand, various interpretations of quantum mechanics appeal to a process called decoherence (for an introduction to decoherence, see [Zur91a, Zur02a] and references therein) to insure the transition from a quantum to a classical world, while on the other, the very same process is responsible for the presence of errors in quantum computers (for an introduction to quantum error correction, see [Pre98a] Chap. 7).

More precisely, the following sections will focus both on the interpretation problems of the theory and on their relationship with the concept of information in quantum mechanics. The presentation will deliberately emphasize a relatively new trend among physicists: information is the only relevant quantity in our incessant attempt to understand the world we live in [Har00a, Har01a, Fuc02b]. In this perspective, quantum mechanics will be reduced to a small set of axioms. These axioms are a reasonable choice because they lead to a theory whose predictions are confirmed by our everyday experience: computer hardware — based on quantum effects in solids — and laser pointers are some examples of applications directly derived from quantum mechanics.

[1] Here classical must be understood as being the opposite of quantum.

1.2 Foundations

1.2.1 General considerations about physical theories

Has science finally triumphed over obscurantism? Darwinism has undoubtedly been one of the most significant advances in the comprehension of the organization of life on earth — if not about the origin of life. However, it is being increasingly opposed to other sorts of arguments. Even if the answer to this question is not yet definitive, science should finally win the battle. Its advantage is that it does not require faith to establish results. On the contrary, it only proceeds from the elementary observation of events repeating themselves over time. This repeatability is the most fundamental statement one could formulate about the universe: it allows the learning process to take place, and simultaneously gives the intuition of a causal structure. The goal of any scientific theory is to take into account this causality to extract laws and principles that will, in some sense, crystallize our empirical understanding: I. Newton, by observing apples falling, had the intuition of the existence of a universal law governing the attraction of masses.

Such simple remark constraints what a reasonable theory can be. In short, it must summarize our experimental knowledge and propose an interpretation for the observed phenomena. More technically, a theory that describes physics must predict the outcomes of experiments carried out by a fair observer on a given particular system. Of course, when predictions and experiments disagree, it is time to modify the theory. Because, logic and mathematics are also concerned with statements that are repeatable — theorems hold as long as their associated assumptions are fulfilled — it seems natural to express the laws of physics in mathematical terms. In turn, the mathematical structure of the theory will have to:

- include a description of the state of the system, all the necessary — and only the necessary — information about it to predict outcomes of experiments;
- give a way to compute the evolution in time of the latent properties of the system, its state — it might necessitate to input some additional objective parameters such as coupling constants, etc. —;
- describe how information can be extracted from the state of the system, define measurement processes.

While the first two requirements are rather natural and expected, the last one seems so trivial in classical physics that it is almost never mentioned. For quantum mechanics, it has however a prominent place in the theory. It is also at the source of the main difference between quantum and classical information which W. H. Zurek summarizes as: "in classical physics what is known about the state can be dissociated from that state, while in quantum physics what is known about the state cannot be treated separately from the state" [Zur03b]. Indeed, the measurement of quantum systems has always been a subject of puzzlement. It was partly responsible of A. Einstein's discontent about the

new theory which he expressed as "God doesn't play dice". Researches on the measurement process together with its associated interpretation problems did undergo a revival under the impulse of quantum information processing. Some thought experiments have been implemented on physical devices confirming the validity of quantum mechanics, but more surprisingly the constraints imposed by the measurement process appeared to be fruitful for cryptographic purposes: E. Knill said "Quantum information is about what you can't do" and we shall see that because of quantum measurements, we can't do much!

The other parts of this manuscript will be devoted to applications of quantum information, but here we shall pursue on the theoretical implications raised by the quantum measurement problem and more generally the quantum-classical transition — the emergence of an objective reality out of the quantum substrate.

1.2.2 Axioms

Before introducing the fundamental axioms of quantum mechanics, it is important to mention a few facts. First of all, these axioms, since they define a physical theory, must satisfy the constraints evoked earlier. They must define the state of a system mathematically, give its possible evolutions and explain what are the measurements. This program will guide us throughout the rest of this section. Second, the reader should keep in mind that different variants of this axiomatic description exist, but all lead to the same physical theory. Of course, a particular choice of axioms is, in some sense already an interpretation of the theory. The motivation of the particular choice made here was to have a compact and comprehensive description of the theory.

Definition 1.2.1 (State). *The state of a physical system is a compact description of the properties of the system.*

Axiom 1 (Pure state). *The state of a quantum system is represented by a ray in a Hilbert space.*

For sake of simplicity, in the rest of this manuscript we will only consider the case where the Hilbert space is finite dimensional. Infinite or continuous Hilbert spaces are of course important. Their definition, uses, and interest can be found in standard textbooks [CTDL77a], but for our purpose they will be an unnecessary complication. Mathematically, a ray is the equivalence class of a non-zero vector with respect to the multiplication by a complex number — a one-dimensional subspace. However, the convention in quantum mechanics is not to manipulate the whole equivalence class directly, but to access it through one of its elements. Such element, which symbolizes the state of the system, is called the state vector or the ket vector $|\psi\rangle$ in Dirac's notation. Its complex conjugate, also corresponding to linear forms over the state space, is the bra $\langle\psi|$. The inner product, the braket, of $|\psi\rangle$ by $|\phi\rangle$ written $\langle\psi|\phi\rangle$, is the canonical inner product of the Hilbert space. The convention in quantum mechanics is to impose a normalization condition for state vectors, i.e. $\langle\psi|\psi\rangle = 1$. Because this condition does not

fully characterize a single vector in a ray, the normalized state vectors $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ correspond to the same physical state. This is the so called invariance by a global phase. Finally, the projection onto the vector $|\psi\rangle$ is noted $|\psi\rangle\langle\psi|$.

Definition 1.2.2 (Evolution). *The evolution of a quantum system is responsible for the change of its presumed properties, the change of its state.*

Axiom 2 (Evolution of pure states). *The time evolution of a quantum system is given by the Schrödinger equation,*

$$\frac{d}{dt}|\psi(t)\rangle = -iH(t)|\psi(t)\rangle\,, \tag{1.1}$$

where $H(t)$ is a hermitian matrix called the Hamiltonian of the system.

Because the Hamiltonian is hermitian, the evolution operator that allows to compute the evolved state vector at time t from the initial condition at $t = 0$ is unitary. For instance, with an Hamiltonian H independent of t, one of its eigenstate, $|\psi_i\rangle$, associated to the eigenvalue E_i evolves according to,

$$|\psi_i\rangle \to e^{-iE_it}|\psi_i\rangle\,, \tag{1.2}$$

while a generic state is transformed according to

$$\sum_i \alpha_i|\psi_i\rangle \to \sum_i \alpha_i e^{-iE_it}|\psi_i\rangle\,. \tag{1.3}$$

Here, it is important to note that quantum mechanical evolutions are linear and invertible — properties that are usually not encountered in classical physics. The linearity is thought to be responsible for the quantum parallelism and the power of quantum computers. It is also involved in one of the most famous no-go theorem of quantum information, the no-cloning theorem [WZ82a, Die82a]. Finally, remark that evolutions in quantum mechanics are deterministic: there is no randomness involved neither in the description of the state nor in its evolution. Randomness will come only through measurements.

Definition 1.2.3 (Measurement). *A measurement is an action that leads to extraction of classical information about the state of a physical system.*

Axiom 3 (Born's rule). *Each measurement on a quantum system is associated to a hermitian matrix O called observable. The possible outcomes o_i of the measurement are the eigenvalues of O. When the state of the system is $|\psi\rangle$, the probability of getting a particular outcome o_i, is given by*

$$p_i = \langle\psi|\,\Pi_i\,|\psi\rangle\,, \tag{1.4}$$

where Π_i is the orthogonal projector onto the subspace associated to the eigenvalue o_i.

Quantum measurements are therefore inherently random. Even though the state of a quantum system is perfectly known, some measurement results are known only probabilistically.

Such situation is never encountered in classical physics where a perfectly known state implies deterministic results — at least with perfect measuring devices. For instance, suppose that the state of a classical object, e.g. a ball, is perfectly known. Then, measuring its position and its momentum gives deterministic values. On the other hand, suppose that a two-dimensional quantum system with basis $\{|0\rangle, |1\rangle\}$ is prepared in the state $(|0\rangle + |1\rangle)/\sqrt{2}$. This corresponds to a perfectly known state since it gives a complete description of its state vector $|\psi\rangle$. Nonetheless, the measurement of the observable $|0\rangle\langle 0| - |1\rangle\langle 1|$ gives the value $+1$ with probability $\frac{1}{2}$ and -1 with probability $\frac{1}{2}$. In quantum theory, there are measurement results that are intrinsically uncertain.

This is not the only surprising consequence of Born's rule. The example above shows that quantum measurements are an extremely poor tool — but nevertheless the only one available — for gaining information about the state of a system: only one bit of information about the state of the system is extracted by such measurement whereas, on the other hand, the state itself is parametrized by two complex numbers. In fact, when o_i is observed, this bit of information only tells that before the measurement the state $|\psi\rangle$ was not orthogonal to Π_i. In quantum mechanics, there exists a gap between the state vector description and the actual information one can obtain about the system through one-shot measurements [Fuc96a, Fuc98a]. Once again such situation is not encountered in classical physics: the state of a system is equally parametrized by its properties, e.g. position and momentum, but all these quantities can in principle be determined simultaneously and with arbitrary accuracy in a single operation. The realization of this fundamental difference should already point toward the potential difficulties of controlling quantum systems in the perspective of quantum information processing. Surprisingly such control is possible and quantum error correcting codes are an example of a feedback mechanism involving measurements of quantum systems (see Part II).

This concludes this rapid presentation of the axioms of quantum mechanics. They answer to three fundamental questions: how is a state represented, how does it evolve, and what does the theory predict about the reality?

1.2.3 Collapse of the wave-function

The reader familiar with quantum mechanics is probably accustomed to find a fourth axiom in the description of the theory. It was a choice to let it aside and not to consider it as fundamental. In fact, called equivalently the collapse of the wave-packet or the collapse of the state vector, it can be derived from the previous construction. The approach proposed here will try to give an appropriate explanation of its relevance rather than stating it as a postulate.

To arrive at the necessity of the collapse of the wave-function, we will first examine quantum measurements in the light of classical physics. Namely, quantum mechanics has been introduced to explain physics at the atomic or subatomic level — to explain the

behavior of the fundamental constituents of classical objects. Under certain conditions, it is should thus be possible to recover classical mechanics from quantum mechanical laws. In particular, classical measurements should be quantum measurements restricted to macroscopic objects.

Let us now analyze briefly what happens in the course of classical measurements. There are two main motivations to conduct a measurement on a physical system: either its state is unknown and the measurement is meant to give information about it, or we know the state of the system at a given time and we want to check whether its dynamics follows a given set of laws. The latter case can also serve to infer information about another system (the one which is measured acts as a probe). Our everyday experience shows us that, when a system is being measured, in most of the cases, this measurement can be conducted in a way that does not disturbs the system: by looking at a ball we learn its position without ever modifying it. However, because of the newly acquired information our description of what we are thinking about the properties of the system changes. We update the state of the system because we have reduced our ignorance about it. This procedure is well known in probability theory and is called the Bayes rule [CT91a]. Namely if two random variables X and Y have a joint probability distribution $p_{XY}(i,j)$ the observation of a particular realization, e.g. j_0, of the random process for Y leads us to update our description of X. Conditionally to the information about Y, the probability distribution of X becomes $p_{XY}(i,j_0)/\sum_i p_{XY}(i,j_0)$. We shall see that the collapse of the wave-function is the exact analog of the Bayes rule.

We now return to the fully quantum case. As a probabilistic theory, quantum mechanics is also subject to the Bayes rule. When an observer gets aware of a measurement result, he should update the probability distribution of the results so that it corresponds to his current knowledge of the system. The particularity of quantum mechanics is that all probabilities derive from a more fundamental description, the state vector. Hence, he should translate the Bayes rule to state vectors in order to properly account for a gain of information. The collapse of the wave function corresponds precisely to this updating procedure: the state of the system, after the outcome o_i of observable O has been obtained, must be such that, in absence of dynamical evolution, if O is measured again the same result o_i occurs deterministically.

Proposition 1.2.1. *For a non-degenerate observable O — whose eigenspaces are one-dimensional — the update of the state vector when outcome o_i associated with the projector Π_i is obtained must follow the simple rule:*

$$|\psi\rangle \to \frac{\Pi_i |\psi\rangle}{\sqrt{\langle\psi|\,\Pi_i\,|\psi\rangle}}. \tag{1.5}$$

Proof 1.2.1. To arrive at this conclusion it is sufficient to realize that for having a deterministic result when the same subsequent measurement is performed the state vector must have support only in the subspace given by the projection operator Π_i. Since this subspace is one-dimensional it defines a ray in the Hilbert space, and therefore the state of the system. The square root factor is added to respect the quantum mechanical conventions that represent rays by vectors normalized to 1. ∎

I hope that this simple derivation,

- justifies the choice of removing the collapse of the wave function from the axiomatic description of quantum mechanics;
- tackles the almost mystical origins of this collapse;
- shows that information is required to claim for a collapse.

This third remark points out that even though the collapse is instantaneous, it only concerns the knowledge we have about the state of a system and, most importantly, is conditional to a measurement result. Hence, there is no instantaneous action at a distance.[2]

1.3 Open quantum systems

After this short review of the foundations of quantum mechanics, we will start to elaborate on them. In this perspective, open quantum systems are the first obvious generalization once the quantum formalism is introduced. They are particularly important for understanding the quantum-classical transition as well as the origins of the errors in quantum computers. Contrarily to closed quantum systems, open systems are not explicitly described by state vectors. Their openness, synonymous of interaction and exchange of energy with other quantum systems such as an environment, requires a probabilistic treatment at the level of the state. These systems tend to get correlated to other systems that are inaccessible to measurements and which perturb them at random. While this case is not directly taken into account by the axioms that we introduced earlier, it derives from them: open quantum systems can always be viewed as subsystems of a bigger closed quantum system.

The mathematical treatment of open quantum systems must take into account these new possibilities. It appeals to a slight modification of their definition: the state of a system is a linear form f over the algebra of its observables. The value of f for a given observable O is the average of O for a collection of many quantum systems all prepared in the state corresponding to f. Since any linear form f over the observables can be written

$$f : O \mapsto \operatorname{Tr} \rho O, \tag{1.6}$$

where ρ is another hermitian matrix. The state of a system is thus completely described by ρ itself, and the state of the system is consequently identified with ρ, which is called the density matrix.

Of course this formalism also takes into account closed quantum systems. If we note that the average value of O for a state $|\psi\rangle$ is $\langle\psi| O |\psi\rangle = \operatorname{Tr} |\psi\rangle\langle\psi| O$ we conclude that

[2]Hence, using the collapse of the wave-function in bipartite quantum correlated systems for the purpose of superluminal communication is bound to fail. One party cannot collapse the other party's description of the overall state without sending information — which takes place at a speed lower than the speed of light.

the system is equally described by the density matrix $\rho = |\psi\rangle\langle\psi|$. By definition, the observable $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ should average to 1 for any state, thus valid density matrices must satisfy $\mathrm{Tr}\rho = 1$. Similarly, for any physical state the average value of a positive operator must be positive as well, which imposes density matrices to be positive semi-definite.

Note also that pure states — the states whose density matrix is a one-dimensional projector — are extremal points of the convex set of admissible density matrices. All others — the mixed states — can be viewed as the result of a lack of information about the preparation of the state of the system. Any density matrix can be written as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| , \tag{1.7}$$

where $|\psi_i\rangle$ are pure states and $\{p_i\}_i$ is a probability distribution over these states. Hence, ρ can be obtained by preparing the state $|\psi_i\rangle$ with probability p_i. However, even though the above decomposition is not unique, different preparations of the same ρ cannot be distinguished by a measurement. This is a direct consequence of the definition of ρ: it is the mapping between observables and their average values. Thus, by construction, ρ contains all the information about the physical system. Note that this indistinguishably principle has had in the recent years a quite astonishing application as it ensures the unconditional security of some quantum cryptographic schemes (for the most simple examples, see [BB84a, SP00a]).

The introduction of this new formalism was motivated by the necessity to describe the state of a system interacting with its environment. We shall now pursue this goal and explain how these mixed states arise from pure states of a larger bipartite system. As we already mentioned, quantum mechanics was originally constructed for isolated systems. It is nevertheless possible to consider just a single subsystem of a bigger quantum system containing it. The whole is supposed to be in a pure state, $|\psi\rangle$, and to follow the usual axioms of quantum mechanics. Affecting a state to the subsystem $\mathcal{A}$ — we suppose $\mathcal{A}$–$\mathcal{B}$ constitutes a closed system — consistently to the density matrix formalism implies to find a matrix $\rho_{\mathcal{A}}$ such that for any observable O of $\mathcal{A}$ we get the correct average value by computing $\mathrm{Tr}\,\rho_{\mathcal{A}} O$. Note that measuring O on the system $\mathcal{A}$ without doing anything to $\mathcal{B}$ is equivalent to measure $O \otimes I$ on the whole — I does not bring information about B. Hence, we must have

$$\mathrm{Tr}\,\rho_{\mathcal{A}} O = \mathrm{Tr}\,(|\psi\rangle\langle\psi|\, O \otimes I). \tag{1.8}$$

This defines the following application

$$\mathrm{Tr}_{\mathcal{B}} : |\psi\rangle\langle\psi| \mapsto \rho_{\mathcal{A}} \tag{1.9}$$

which is called the partial trace over $\mathcal{B}$. For example, suppose that $\mathcal{A}$ and $\mathcal{B}$ are two-dimensional systems, and that their combined state is $(|0\rangle\,|0\rangle + |1\rangle\,|1\rangle)/\sqrt{2}$. The partial trace over $\mathcal{B}$ gives the result

$$\mathrm{Tr}_{\mathcal{B}} \left(\frac{1}{2}(|0\rangle\,|0\rangle + |1\rangle\,|1\rangle)(\langle 0|\,\langle 0| + \langle 1|\,\langle 1|)\right) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.10}$$

which is not a pure state anymore: this matrix does not correspond to a one dimensional projector. In short, mixed states of Eq. (1.7) can also arise as the result of a lack of information about the correlations of a system: here, the correlations of $\mathcal{A}$ and $\mathcal{B}$ are simply ignored and lead to uncertainty about $\mathcal{A}$. Any non-degenerate observable will have a $\frac{1}{2}$ probability for each of its outcomes.

The possible evolutions of open quantum systems are obtained in a similar way. The system of interest is considered as part of a larger and closed quantum system which follows Schrödinger's equation. The effect of this evolution for the system of interest is obtained by taking the appropriate partial trace. The complete mathematical treatment will not be reproduced here, but instead will give the result directly [Kra83a, Pre98a, NC00a] : allowed evolutions for open quantum systems are through completely positive trace preserving maps. Note that while the positivity of an operator L is defined as the positivity of the operators $L(\rho)$ for ρ positive, the complete positivity is the positivity of $L \otimes I$ for all finite size identity matrices I. This result is fundamental both for its theoretical and practical implications. However, in this manuscript, we will not use it directly but rather specify our environment and its interaction with the quantum system under scrutiny. We will compute the evolution of the combined quantum state and take the partial trace in order to find the exact transformation of the open system of interest.

Pursuing the idea of considering larger systems, it is also possible to define generalized measurements. They are historically related to open quantum systems but also apply to closed quantum systems [neu43a, Neu54a, Kra83a]. They are constructed by letting the system interact with an ancillary quantum system and by measuring the combined state. In this case, we are not any more confronted to observables but to Positive Operator Value Measures (POVM). It is given mathematically as a decomposition of the identity into trace decreasing positive semi-definite operators, $\{E_i\}_i$ such that $\sum_i E_i = I$. The probability of getting the outcome E_i is given by

$$p_i = \mathrm{Tr}\,(\rho E_i). \tag{1.11}$$

Note that this equation is the analog of Eq. (1.4) when E_i is replaced by the projector Π_i. However, contrarily to standard measurements, there is no direct way to deduce into what state the wave-function collapses when a measurement result has been obtained. To arrive at a result, one must come back to the description of the POVM as a projective measurement onto the larger Hilbert space of the system and the ancillas. There, the result about the collapse of the wave-function can be applied. When this procedure is followed, each operator E_i is decomposed as $E_i = A_i A_i^\dagger$ for some A_i's. When the particular outcome i is obtained, the state of the measured system must be updated according to:

$$\rho \mapsto \frac{A_i \rho A_i^\dagger}{\mathrm{Tr}\, A_i \rho A_i^\dagger}. \tag{1.12}$$

Unfortunately, the operators A_i generally depends on the specific implementation of the POVM (see the discussion in [Pre98a]). Thus, to correctly update the state of a system after a measurement, one must have a precise description of the measuring device itself.

This concludes our short presentation of quantum mechanics for open quantum systems. Of course a lot more could be said but this would be unnecessary: while open quantum systems are of fundamental interest, the concepts used in this manuscript are simple enough to be understood with the help of these elementary notions.

1.4 Superposition and entanglement

Quantum theory differs from classical mechanics partly because of the linearity of the Schrödinger equation: given the evolution of two state vectors, the evolution of any linear combination of the evolutions is the linear combination of the evolutions, Eq. (1.3). For example, suppose the position of a particle is governed by quantum mechanics and two trajectories are known. Then, the trajectory composed of the superposition — in the sense of the superposition of the state vectors describing these trajectories — also satisfies the Schrödinger equation. Nothing prevents these states to exist. Actually, interference experiments with one photon, and even quite large molecules [NAZ02a, HHB+04a], prove that they describe the physical reality. This superposition principle of quantum mechanics is analogous to the one of elementary electricity with Kirchhof's laws.

However, note that this concept is totally different from the one introduced earlier about the density matrices: a mixed state is a convex combination of pure states, but in this case the addition is carried out for density matrices representing the state and not for state vectors (i.e. vectors in a Hilbert space).

As a preview of the potential problems of the quantum-classical transition, one can ask the following question: why are superpositions of trajectories for classical objects never observed even though they are ultimately made of quantum particles? In other words, what prevents the observation of superpositions of macroscopic states, while nothing in the axioms of quantum mechanics suggests the presence of a selection principle that would eliminate them?

A possible rephrasing of this question would lead to consider the origin of the violation of the basis invariance of quantum mechanics at a macroscopic level. For instance, the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ of a two-dimensional system is a superposition of the vectors $|0\rangle$ and $|1\rangle$. But if one defines $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, we have the identity $|0\rangle = (|+\rangle + |-\rangle)/\sqrt{2}$ which shows that $|0\rangle$ can also be considered as a superposition of $|+\rangle$ and $|-\rangle$. In fact, this kind of relation holds in more general situations. Then, one usually concludes that there is no preferred basis in quantum mechanics that would lead to a sensible interpretation of superpositions. However, for large systems a preferred basis emerges and rules out superpositions of its states. Those are the classical states of macroscopic objects: for a ball there is no superposition in the position basis.

The superposition principle could be just a curiosity of quantum mechanics allowing strange manifestations, but this is not the end of the story. R. P. Feynman realized its before-then unexpected implications [Fey82a, Fey84a, Fey86a]. A generic state of n two-dimensional systems is a superposition of the 2^n basis states. Hence, to represent the state of such system, it generally requires to store about 2^n complex numbers. This

rapidly exceeds the capacities of the biggest computers to simulate interacting quantum systems. However, if the information about the combined state is stored directly into quantum mechanical systems it requires only n two level systems. Therefore, such representation could enhance our abilities to simulate large quantum systems. This simple remark is certainly at the source of quantum computation, but it took some time before D. Deutsch proved that this assertion was indeed correct [Deu85a], and that the quantum computing idea could work, at least in principle.

Finally, and before concluding this chapter, we should mention an additional manifestation of the superposition principle for composite quantum systems, called entanglement. This notion is best explained by considering the following example: the state $(|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$ of two systems $\mathcal{A}$ and $\mathcal{B}$ each having two dimensions cannot be written as a product state — a state $|\psi_A\rangle|\psi_B\rangle$. Proving such statement is not a difficult task, and is a consequence of the fact that product states can only be transformed into other product states by changes of basis on each composite system separately. In addition, it is possible to prove that this kind of correlations cannot be created by classical means [Bel64a]. It thus becomes a physical resource when $\mathcal{A}$ and $\mathcal{B}$ cannot interact with each other in a quantum way. The immense interest for entanglement is due partly to its very fundamental consequences for quantum theory (e.g. it constraints through the Bell inequalities [Bel64a] to abandon A. Einstein's local realism point of view); its applications in quantum communication (e.g. [BW92a] or [BBC+93a]); its implication in the speed-up of pure state quantum computers [JL02a].

1.5 Conclusion

After this brief introduction to quantum mechanics, the manuscript will deal with to different subjects. The first one will explore the interpretations of quantum mechanics and how the classical world emerges from the quantum realm. The other subject concerns quantum information and some of its implementations. It will focus on error correcting codes and on the design of experiments aimed at realizing elementary quantum operations.

If one thing should be remembered from this introduction it would be undoubtedly the role of information. Information constraints the mathematical representation of the state of quantum system. It also explains the collapse of the wave-function. It gives the intuition that quantum systems might have a greater ability to store information than classical bits. Finally, the next chapters will exemplify its use as a powerful tool to investigate the emergence of a classical world out of the quantum realm.

2 Introduction to decoherence and einselection

2.1 The need for an interpretation

QUANTUM THEORY HAS been verified by an impressive amount of experiments such as tests of Bell's inequalities [AGR81a]. Physicists have no doubt about of its predictions. However, there is still a domain in which quantum theory fails to provide an unquestionable framework of understanding: the existence of a macroscopic and classical world made of microscopic quantum particles. In the same vein, the universe as a closed quantum system should allow superpositions as a consequence of the linearity of Schrödinger's equation. Such superpositions are nonetheless never observed. This simple example seems to violate the basis invariance of quantum mechanics: the universe exhibits classical features. Less singular examples can be taken in our everyday life but show always the same behavior: only one alternative emerges from the possible superpositions. Even though quantum theory does not indicate a limit to the validity of Schrödinger's equation, a single classical world exists.

The easiest way to solve this problem raised by the quantum theory is to impose a limit to the validity of the quantum description. That is to prescribe the impossibility to apply quantum mechanics to macroscopic objects. Mesoscopic physics therefore constitutes a fuzzy border line between the quantum and the classical regimes. Nonetheless, this solution does not bring any answer to the initial question: why would quantum theory not be applicable to macroscopic systems? This question became even more acute since the discovery that macroscopic systems must sometimes be considered as quantum objects: Weber bars—gravity wave detectors weighting about a ton—must be described in terms of quantum mechanical oscillators.

2.2 The measurement problem

In the perspective of a rigorous analysis, the quantum-classical transition is often reduced to a more restrictive framework. This framework is the one of quantum measurements. The translation of the emergence of a classical world—of the choice of classical outcomes in a measurement—is then called the measurement problem.

More precisely, the typical quantum measurement context consists in two quantum systems—the system $\mathcal{S}$ and the apparatus $\mathcal{A}$—, which interact in a carefully controlled way. Because of this interaction, $\mathcal{S}$ and $\mathcal{A}$ get correlated. When $\mathcal{A}$ is initially in a perfectly known state—a pure state—, this process corresponds to a transfer of information from the system to the apparatus. For instance, a two dimensional system $\mathcal{S}$

in the state $(|0\rangle + |1\rangle)/\sqrt{2}$ could interact with an apparatus $\mathcal{A}$ in the initial state $|0\rangle$. For some Hamiltonian and for some interaction time, the combined state of $\mathcal{S}$–$\mathcal{A}$ can be transformed in the following way:

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,|0\rangle \rightarrow (|0\rangle\,|0\rangle + |1\rangle\,|1\rangle)/\sqrt{2}. \tag{2.1}$$

Here, an obvious one-to-one correlation between the $\{|0\rangle, |1\rangle\}$ bases of the system and of the apparatus is created. Then, it might seem legitimate to claim that the measurement of the system with respect to the $\{|0\rangle, |1\rangle\}$ basis has been performed. This is not the case. If we rewrite the above correlated state with the help of $|0\rangle = (|+\rangle + |-\rangle)/\sqrt{2}$ and $|1\rangle = (|+\rangle - |-\rangle)/\sqrt{2}$, we get the following identity,

$$(|0\rangle\,|0\rangle + |1\rangle\,|1\rangle)/\sqrt{2} = (|+\rangle\,|+\rangle + |-\rangle\,|-\rangle)/\sqrt{2}. \tag{2.2}$$

At this stage, it is impossible to maintain that the measurement process is complete. With respect to the state of the pair, each basis of $\mathcal{S}$, $\{|0\rangle, |1\rangle\}$ as well as $\{|+\rangle, |-\rangle\}$, is equally correlated with a set of perfectly distinguishable states of the apparatus.

Indeed, this matter of fact is the consequence of the presence of quantum correlations between the system and the apparatus. Before considering this idea in detail (see Chap. 3), we shall continue with our initial example and describe quickly how such problem can be effectively solved. Suppose that there is a transformation applied to the state of the apparatus which maps any superposition of states, $\alpha\,|0\rangle + \beta\,|1\rangle$ into the density matrix $|\alpha|^2\,|0\rangle\langle 0| + |\beta|^2\,|1\rangle\langle 1|$. When applied to the entangled state of $\mathcal{S} - -\mathcal{A}$, Eq. (2.2), this transformation leads to $\frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)$. In such case, the basis invariance exemplified earlier does not hold anymore: there is only one basis in which the system and the apparatus keep a one-to-one correlation in spite of this process. Then, and only then, would it be possible to conclude that the system has been measured in the $\{|0\rangle, |1\rangle\}$ basis.

The aim of the various interpretations of quantum mechanics is to provide an explanation for this transformation which conveniently removes the basis invariance of the apparatus. As a consequence, it will also give a way to understand the quantum-classical transition as an effect of the emergence of a preferred basis for quantum systems interacting with an environment.

In the following sections, we briefly describe how the measurement problem is taken into account in the most common interpretations of quantum mechanics.

2.3 Copenhagen interpretation

The core of Copenhagen interpretation (see [Boh63a] and references therein) is simply to appeal for a limit to the validity of quantum mechanics: the question of deciding what is measured is answered by requiring the presence of classical measuring devices to obtain information about quantum systems. In this case, there is no need to search for pointer states nor to eliminate entanglement between the system and the apparatus.

Because of its classical description, the apparatus cannot explore the basis invariance of quantum mechanics nor can it be entangled with the to-be-measured quantum system.

In fact, this point of view turns out to push the difficulty a step further: why are there classical objects, and where do they come from? Indeed, N. Bohr suggested to allow the border between the quantum and the classical to be movable. Ultimately, it should be possible to regard any system as quantum mechanical as long as the proper classical apparatus can be found. Human observers are no exception to the rule.

To rephrase the Copenhagen interpretation, it postulates the existence of a classical world outside the quantum one, and from which we can observe quantum effects. However, the fact that apparatuses themselves must sometimes be described in quantum mechanical terms also suggests that a quantum description underlies the classical existence of measuring devices. Unfortunately, this interpretation does not provide a satisfying a satisfying description of such more fundamental structure.

2.4 Many world interpretation

The many world interpretation was introduced by H. Everett in 1957 [Eve57a] and takes apparently a radically different approach. It attempts at describing quantum theory from inside rather than from outside: when one considers the universe as a whole there cannot be appropriate classical measuring devices nor external observers to record measurement results. Rather, the whole measurement process has to be described within the deterministic framework of the theory (i.e. without appealing to the existence of classical objects).

This interpretation postulates the existence of a preferred basis for writing superpositions of the state vector of the universe. This basis is chosen in a way such that if a measurement is made in this basis, the resulting states of subsystems continue to be pure states. Each element of the preferred basis is called a branch of the universe. By definition, the branches correspond to mutually exclusive events, which appears deterministic to an observer who has access only to a single branch. The many world interpretation actually states that this is the way of looking at the measurement problem: each time a system and an apparatus interact, the state vector of the universe branches and leads to mutually exclusive events. The observers can only see a single definite outcome since they are described by the branching process of the state vector of the universe.

This interpretation seems to remove cleverly the possibility of seeing consequences of the superpositions principle. However, it does not really define the preferred basis in an unambiguous way: Eq. (2.2) defines two legitimate decompositions in branches for the same superposition. Why only some branches are selected and when does this selection happen are questions that arise naturally and that are not answered by the many world's view.

2.5 Existential interpretation

2.5.1 A transition to the classical reality

The two interpretations that we reviewed above fail in properly solving the measurement problem. Both of them in fact push this problem a little further, into regions where it is more difficult to argue rigorously. On the other hand, facts are still striking: classical states exist and quantum mechanics is verified for systems of increasing size. Once again, it leads to the intuition that the classical world can certainly be explained consistently within the quantum mechanical framework.

Progress on this path started when H. D. Zeh, among others (see [GJK+96a] and references therein), realized that macroscopic systems are usually not closed systems. Thus, they are not bound to follow Schrödinger's equation and could probably escape the constraint set by its linearity. This idea was developed by W. H. Zurek [Zur81a, Zur82a, Zur91a, Zur93b, Zur98a, Zur03a], and constitutes the heart of the theory of decoherence and einselection—environment induced superselection—as a solution to the measurement problem: under certain conditions, the interaction of an uncontrolled environment can actually select a preferred basis and impose all density matrices to be diagonal in this basis. For a quantum measuring device, this means that its pointer cannot be anymore in a superposition of states. It destroys the basis invariance of entangled quantum states. What is measured and when it is measured are the two fundamental questions that can be answered by the theory of decoherence.

The rest of this section will define this process. It will provide the fundamental material required by more elaborate constructions. One such construction is used to analyze the emergence of objective properties for classical systems, Chap. 4.

2.5.2 Decoherence and einselection

Definition 2.5.1 (Decoherence). *Decoherence is the process by which a quantum system loses its possibility of being in a superposition of states.*

Such definition is voluntarily vague because decoherence can take many aspects. However, one possible source of misconceptions should be tackled readily: it is always possible to express the density matrix of a quantum system into its diagonal basis. Then, it will appear as a discrete probability distribution over the set of its eigenstates, and will therefore not appeal to the superposition principle. Did decoherence take place? Of course not. Contrarily to this trivial case, the decoherence basis—or the set of pointer states—for a system should not depend on its current state. It is fixed in advance. The decoherence process reduces any density matrix of the system to a classical mixture in this basis.

Einselection is the general possibility of quantum systems to violate the superposition principle. Thus, it must have its origin in the interaction with an unobserved quantum system. This is usually embraced by the term environment. It can be of course constituted by any quantum system or even by internal and unaccessible degrees of freedom of

the quantum system of interest. Finally, note that the study of einselection is a search for criteria and conditions that lead to the appropriate violation of the superposition principle. Decoherence is not about proving that any interaction with an environment will proceed in the emergence of classicality. Other phenomena can happen to open quantum systems besides decoherence. Noise or dissipation are some examples of such other possible processes.

The next chapters are based on the following toy model: an environment $\mathcal{E}$ interacts with the quantum system $\mathcal{S}$, and continuously performs a pre-measurement on it. This pre-measurement step is the one that lead to an entangled state between the system and the apparatus when we explained the measurement problem, Eq. (2.1). Most of the time, it is assumed that the environment is in fact made of N subsystems which independently interact with $\mathcal{S}$. This interaction is the same for all subsystems, except maybe for their coupling strength. Such decomposition in subsystems of the environment can be, for instance, the phonons for an electron in solid, or thermal photons for an atom in a cavity.

A generic Hamiltonian that realizes such pre-measurement has the form

$$H_I = \sum_{i=1}^{N} g_i A \otimes B_i, \tag{2.3}$$

where each A is an observable of the system and B_i are non-degenerate observables acting on each subsystem of the environment independently. In addition, and to simplify the analysis, the coupling constants are chosen randomly over some real interval. The effect of each of these terms in the Hamiltonian is to impose a conditional evolution to the state $|\psi_i\rangle$ of the i-th environment according to the eigenstates of A, $|a_k\rangle$:

$$|a_k\rangle \bigotimes_{i=1}^{N} |\psi_i\rangle \rightarrow |a_k\rangle \bigotimes_{i=1}^{N} \exp\left(-ig_i\lambda_k B_i t\right) |\psi_i\rangle \tag{2.4}$$

The full dynamics for the system $\mathcal{S}$—after the partial trace over the environment—leads to a density matrix at time t of the form

$$\rho(t) = \sum_{k=}^{N} |\alpha_k|^2 |a_k\rangle\langle a_k| + \sum_{k\neq l} \alpha_k \alpha_l^* z_{kl}(t) |a_k\rangle\langle a_l|, \tag{2.5}$$

where $z_{kl}(t)$ is a complex function which satisfies for t large enough $\langle z_{kl}(t)\rangle = 0$ and $\langle |z_{kl}(t)|^2\rangle = O(\exp(-N))$, with $\langle . \rangle$ denoting the average over the distribution of coupling constants.

It is now obvious that a set of states has emerged: the eigenstates of A are left untouched by the open dynamics, while those involving superpositions of the latter become the corresponding mixtures. Einselection is the result of an environment induced superselection rule—a rule that forbids superposition of quantum states because it reduces the algebra of observables to a direct sum of such algebras. Decoherence has taken place within this simple model. This gives an explicit example of the emergence of classical reality within the standard quantum mechanical framework (i.e. in absence of additional postulates).

2.5.3 Several criteria, pros and cons

Our next goal is to provide criteria to analyze the evolution of open quantum systems in the perspective of finding pointer states.

The most obvious one is probably the commutation of an observable O with the interaction Hamiltonian used to induce the conditional dynamics of the environment,

$$[O, H_I] = OH_I - H_I O = 0. \tag{2.6}$$

In the Heisenberg picture of quantum mechanics—obtained by evolving the observables rather than the states—the commutation of H_I and O is equivalent to require that O does not change under the combined dynamics of the system and its environment [CTDL77a]. The eigenstates of those observables will remain pure, while others will be transformed into mixtures. The algebra of such observables is precisely forming a direct sum of smaller algebras, signature of a superselection rule.

However, even though this criterion has proved its utility in simple situations, its ability to analyze more complex situations is very limited. In essence, quantum systems usually have their own dynamics acting together with the Hamiltonian of interaction. In such cases, dissipation generally occurs and the only observable satisfying the commutation relation is the identity: all states eventually lead to the maximally mixed state. It is nonetheless possible to motivate the use of the commutation relation with the interaction Hamiltonian for certain well identified cases. In general, it appeals for yet another criterion.

Until now, we have defined classical states as the states that do not get affected at all by the openness of the dynamics. This is in some sense too restrictive to account for the complexity of the decoherence process. At least intuitively, even when dissipation occurs, there must exist states that are almost not affected by the openness of the dynamics—the pointer ones—while the others are very sensitive to the interaction with the environment. We thus arrive quite naturally at a second criterion for decoherence: the predictability sieve. Pointer states are the ones that produce the smallest amount of von Neumann's entropy[1] over the characteristic time of the interaction [Zur93b, Zur98a].

This criterion applied to the case detailed previously gives the same results than the commutation rule. An advantage of the predictability sieve is that it works in more complicated situations. For instance, it played a crucial role in understanding why Gaussian states for the quantum Brownian motion can be considered as classical states [PHZ93a]. This single discovery has actually given rise to an impressive literature on the subject (see [PZ01a] and references therein).

[1] Von Neumann's entropy, $-\mathrm{Tr}\,\rho \log \rho$, is an indicator of the disorder of the system or equivalently of its unpredictability. Thus, the production of von Neumann's entropy characterizes the perturbation induced by the environment on the system.

2.5.4 Summary

To come back to our original questions, we have found that under certain circumstances, a quantum system can exhibit a preferred classical basis. It emerges from the quantum description via the interaction of the system with an unaccessible environment. While einselection explains what are the alternatives for the collapse of the wave-function, it does not tell why a particular outcome is actually chosen. The consensus among physicists is that the latter is a truly random process. This matter of fact is actually confirmed by all the experiments until now. We have also seen that a way to recover classical states out of quantum theory is to search for the most stable states, those that are least perturbed by the interaction with the environment. W. H. Zurek's aphorism holds [Zur98a]: "states that exist are states that persist".

In the next chapters we will continue to investigate decoherence and einselection. First we will describe a tool to analyze the quantumness of the correlations between a system and an apparatus. As a second step, we will see that the einselected states can also acquire other features of classical states such as an objective description.

3 Quantum discord

THE ORIGINAL MOTIVATION for the pointer states—states that are monitored by the environment but do not entangle with it, and are therefore stable in spite of the openness of the system—comes from the study of quantum measurements [vN32a, WZ82a, Die82a, Per93a]. When the quantum apparatus $\mathcal{A}$ interacts with the system $\mathcal{S}$ (*pre-measurement*), the $\mathcal{S}$–$\mathcal{A}$ pair becomes entangled. The nature of the resulting quantum correlations makes it impossible to ascribe any independent *reality* to, say, the state of the apparatus [Zur81a]. One dramatic manifestation of this "unreality" of the state of the apparatus is its malleability: a measurement of different observables on the state of the system will force the apparatus into mutually incompatible pure quantum states. This is a consequence of the *basis ambiguity*. It is best exhibited by noting that the $\mathcal{S}$–$\mathcal{A}$ state after the pre-measurement

$$|\psi_{\mathcal{SA}}\rangle^P = \sum_i \alpha_i |s_i\rangle |a_i\rangle \tag{3.1}$$

is typically entangled. One can rewrite it in a different basis of, e.g. the system, and one-to-one correlation with a corresponding set of pure, but not necessarily orthogonal, states of the apparatus will remain. Thus, it is obviously impossible to maintain that before the measurements the apparatus had an unknown but real (i.e. existing independently of the system) quantum state.

Decoherence leads to environment-induced superselection (or *einselection*) which singles out the pointer states and thus removes quantum excess of correlation responsible for the basis ambiguity. The density matrix of the decohering quantum apparatus loses its off-diagonal terms as a result of the interaction with the environment [Zur93b, GJK+96a, PZ01a, Zur03a]:

$$\begin{aligned} \rho^P_{\mathcal{SA}} &= |\psi_{\mathcal{SA}}\rangle^P \langle\psi_{\mathcal{SA}}|^P \\ &\rightarrow \sum_i |\alpha_i|^2 |s_i\rangle \langle s_i| \otimes |a_i\rangle \langle a_i| = \rho^D_{\mathcal{SA}}. \end{aligned} \tag{3.2}$$

Above $\langle a_i | a_j \rangle = \delta_{i,j}$ following the ideal einselection process, which transforms a pure $\rho^P_{\mathcal{SA}}$ into a decohered $\rho^D_{\mathcal{SA}}$ satisfying the superselection identity [BLOT90a, Giu00a]:

$$\rho^D_{\mathcal{SA}} = \sum_i P^{\mathcal{A}}_i \rho^D_{\mathcal{SA}} P^{\mathcal{A}}_i. \tag{3.3}$$

Above $P^{\mathcal{A}}_i$ correspond to the superselection sectors of the apparatus, e.g. the record states of its pointer (in our example $P^{\mathcal{A}}_i = |a_i\rangle \langle a_i|$). One implication of this equation is

that—once einselection has forced the apparatus to abide by Eq. (3.3)—its state can be consulted (measured) in the basis corresponding to the superselection sectors $P_i^{\mathcal{A}}$ leaving $\rho_{\mathcal{SA}}^{\mathcal{D}}$ unchanged [Zur93b, Zur03a].

Einselection, Eq. (3.2), obviously decreases correlations between $\mathcal{S}$ and $\mathcal{A}$. Yet, in a good measurement, one-to-one correlation between the pointer states of the apparatus and a corresponding set of system states must survive. We shall use two classically equivalent formulas for the mutual information to quantify the quantum and the classical strength of the correlations present in a joint density matrix $\rho_{\mathcal{SA}}$, and study the difference between these two as a measure of the quantum excess of correlations—the quantum discord—in $\rho_{\mathcal{SA}}$.

3.1 Mutual Information

In classical information theory [CT91a] the Shannon entropy, $H(X)$, describes the ignorance about a random variable X, $H(X) = -\sum_i p_X(i) \log p_X(i)$. The correlation between two random variables X and Y is measured by the mutual information:

$$\mathcal{J}(X:Y) = H(X) - H(X|Y), \tag{3.4}$$

where $H(X|Y) = \sum_j p_Y(j) H(X|Y=j)$ is the conditional entropy of X given Y. All the probability distributions are derived from the joint one, $p_{X,Y}(i,j)$:

$$p_X(i) = \sum_j p_{X,Y}(i,j), \; p_Y = \sum_i p_{X,Y}(i,j) \tag{3.5}$$

$$p_{X|Y}(i,j) = p_{X,Y}(i,j)/p_Y(j) \quad \text{(Bayes rule)}. \tag{3.6}$$

Hence, the mutual information measures the average decrease of entropy on X when Y is found out. Using the Bayes rule, Eq. (3.6), one can show that $H(X|Y) = H(X,Y) - H(Y)$. This leads to another classically equivalent expression for the mutual information:

$$\mathcal{I}(X:Y) = H(X) + H(Y) - H(X,Y). \tag{3.7}$$

One would like to generalize the concept of mutual information to quantum systems. One route to this goal, motivated by discussions of quantum information processing, has been put forward [SN96a, CA97a]. We shall pursue a different strategy, using Eqs. (3.4) and (3.7). We start by defining $\mathcal{I}$ and $\mathcal{J}$ for a pair of quantum systems.

All the ingredients involved in the definition of $\mathcal{I}$ are easily generalized to deal with arbitrary quantum systems by replacing the classical probability distributions by the appropriate density matrices $\rho_{\mathcal{S}}$, $\rho_{\mathcal{A}}$ and $\rho_{\mathcal{SA}}$ and the Shannon entropy by the von Neumann entropy, e.g. $S(\mathcal{S}) = S(\rho_{\mathcal{S}}) = -\mathrm{Tr}_{\mathcal{S}}\, \rho_{\mathcal{S}} \log \rho_{\mathcal{S}}$:

$$\mathcal{I}(\mathcal{S}:\mathcal{A}) = S(\mathcal{S}) + S(\mathcal{A}) - S(\mathcal{S},\mathcal{A}). \tag{3.8}$$

In this formula, $S(\mathcal{S}) + S(\mathcal{A})$ represents the uncertainty of $\mathcal{S}$ and $\mathcal{A}$ treated separately, and $S(\mathcal{S},\mathcal{A})$ is the uncertainty about the combined system described by $\rho_{\mathcal{SA}}$. However,

in contrast with the classical case, extracting all information potentially present in a combined quantum system described by $\rho_{\mathcal{SA}}$ will, in general, require a measurement on the combined Hilbert space $\mathcal{H}_{\mathcal{SA}} = \mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{A}}$. The quantum version of $\mathcal{I}$ has been used some years ago to study entanglement [Zur83a], and subsequently rediscovered [BP89a].

On the other hand, the generalization of $\mathcal{J}$ is not as automatic as for $\mathcal{I}$, since defining a quantum conditional entropy $S(\mathcal{S}|\mathcal{A})$ requires us to specify the state of $\mathcal{S}$ *given* the state of $\mathcal{A}$. Such statement in quantum theory is ambiguous until the to-be-measured set of states $\mathcal{A}$ is selected. We focus on perfect measurements of $\mathcal{A}$ defined by a set of one-dimensional projectors $\{\Pi_j^{\mathcal{A}}\}$. The label j distinguishes different outcomes of this measurement.

The state of $\mathcal{S}$, after the outcome corresponding to $\Pi_j^{\mathcal{A}}$ has been detected, is

$$\rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}} = \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} \Pi_j^{\mathcal{A}} / \mathrm{Tr}_{\mathcal{SA}} \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}}, \tag{3.9}$$

with probability $p_j = \mathrm{Tr}_{\mathcal{SA}} \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}}$. $S(\rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}})$ is the missing information about $\mathcal{S}$. The entropies $S(\rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}})$, weighted by probabilities, p_j, yield to the conditional entropy of $\mathcal{S}$ given the complete measurement $\{\Pi_j^{\mathcal{A}}\}$ on $\mathcal{A}$,

$$S(\mathcal{S}|\{\Pi_j^{\mathcal{A}}\}) = \sum_j p_j S(\rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}}). \tag{3.10}$$

This leads to the following quantum generalization of $\mathcal{J}$:

$$\mathcal{J}(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} = S(\mathcal{S}) - S(\mathcal{S}|\{\Pi_j^{\mathcal{A}}\}). \tag{3.11}$$

This quantity represents the information gained about the system $\mathcal{S}$ as a result of the measurement $\{\Pi_j^{\mathcal{A}}\}$.

3.2 Quantum discord

The two classically identical expressions for the mutual information, Eqs. (3.4) and (3.7), differ in a quantum case [Zur00a].[1] The quantum discord is this difference,

$$\begin{aligned} \delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} &= \mathcal{I}(\mathcal{S}:\mathcal{A}) - \mathcal{J}(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} & (3.12)\\ &= S(\mathcal{A}) - S(\mathcal{S},\mathcal{A}) + S(\mathcal{S}|\{\Pi_j^{\mathcal{A}}\}). & (3.13) \end{aligned}$$

It depends both on $\rho_{\mathcal{SA}}$ and on the projectors $\{\Pi_j^{\mathcal{A}}\}$.

[1] The equality for classically correlated systems can be seen directly on the density matrix representation. $\rho_{\mathcal{SA}}$ is diagonal in a basis made of product states on $\mathcal{S}$ and $\mathcal{A}$. Therefore, the diagonal entries of this matrix describe the joint probability distribution of two random variables, say X of the system and Y of the apparatus. As a consequence, we have $\mathcal{I}(\mathcal{S}:\mathcal{A}) = \mathcal{I}(X:Y) = \mathcal{J}(X:Y) = \mathcal{J}(\mathcal{S}:\mathcal{A})$ when $\mathcal{A}$ is measured in the basis which allows diagonalization of $\rho_{\mathcal{SA}}$.

The quantum discord is asymmetric under the change $\mathcal{S} \leftrightarrow \mathcal{A}$ since the definition of the conditional entropy $S(\mathcal{S}|\{\Pi_j^{\mathcal{A}}\})$ involves a measurement on one end (in our case the apparatus $\mathcal{A}$), that allows the observer to infer the state of $\mathcal{S}$. This typically involves an increase of entropy. Hence $S(\mathcal{S}|\{\Pi_j^{\mathcal{A}}\}) \geq S(\mathcal{S},\mathcal{A}) - S(\mathcal{A})$, which implies that for any measurement $\{\Pi_j^{\mathcal{A}}\}$,

$$\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} \geq 0. \tag{3.14}$$

The proofs are postponed to the end of this chapter.

We shall usually be concerned with the set $\{\Pi_j^{\mathcal{A}}\}$ that minimizes the discord given a certain $\rho_{\mathcal{SA}}$. Minimizing the discord over the possible measurements on $\mathcal{A}$ corresponds to finding the measurement that disturbs least the overall quantum state and that, at the same time, allows one to extract the most information about $\mathcal{S}$. Decoherence picks out a set of stable states and converts their possible superpositions into mixtures, Eq. (3.2). Moreover, an unread measurement $\{\Pi_j^{\mathcal{A}}\}$ on the apparatus has an effect analogous to einselection in the corresponding basis through the reduction postulate [vN32a]. Hence it is rather natural to expect that when the set $\{\Pi_j^{\mathcal{A}}\}$ corresponds to the superselection sectors $\{P_i^{\mathcal{A}}\}$ of Eq. (3.3), there would be no extra increase of entropy:

$$\rho_{\mathcal{SA}} = \sum_j \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} \Pi_j^{\mathcal{A}} \Rightarrow \delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} = 0. \tag{3.15}$$

Thus, following einselection, the information can be extracted from $\mathcal{S}$–$\mathcal{A}$ with a local measurement on $\mathcal{A}$ without disturbing the overall state. The state of $\mathcal{S}$ can be inferred from the outcome of the measurement on $\mathcal{A}$ only. The converse of Eq. (3.15) is also true:

$$\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} = 0 \Rightarrow \rho_{\mathcal{SA}} = \sum_j \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} \Pi_j^{\mathcal{A}}. \tag{3.16}$$

Hence, a vanishing discord can be considered as an indicator of the superselection rule, or—in the case of interest—its value is a measure of the efficiency of einselection. When δ is large for any set of projectors $\{\Pi_j^{\mathcal{A}}\}$, a lot of information is missed and destroyed by any measurement on the apparatus alone, but when δ is small almost all the information about $\mathcal{S}$ that exists in the $\mathcal{S}$–$\mathcal{A}$ correlations is locally recoverable from the state of the apparatus.

The quantum discord can be illustrated in a simple model of measurement. Let us assume the initial state of $\mathcal{S}$ is $(|0\rangle + |1\rangle)/\sqrt{2}$. The pre-measurement is a c-not gate yielding $|\psi_{\mathcal{SA}}\rangle^P = (|00\rangle + |11\rangle)/\sqrt{2}$. If $|0\rangle$ and $|1\rangle$ of $\mathcal{A}$ are pointer states, partial decoherence will suppress off-diagonal terms of the density matrix:

$$\rho_{\mathcal{SA}} = \frac{1}{2}\left(|00\rangle\langle 00| + |11\rangle\langle 11|\right) + \frac{z}{2}\left(|00\rangle\langle 11| + |11\rangle\langle 00|\right), \tag{3.17}$$

with $0 \leq z < 1$. Figure 3.1 shows δ for various values of z and various bases of measurement parametrized by θ,

$$\{\cos(\theta)\,|0\rangle + e^{i\phi}\sin\theta\,|1\rangle\,, e^{-i\phi}\sin\theta\,|0\rangle - \cos\theta\,|1\rangle\}, \tag{3.18}$$

with $\phi = 1\text{rad}$. Only in the case of complete einselection ($z = 0$) there exist a basis for which discord disappears. The corresponding basis of measurement is $\{|0\rangle, |1\rangle\}$ ($\theta = 0$), i.e. it must be carried out in the pointer basis.

3.3 Classical aspect of quantum correlations

Separability has been often regarded as synonymous of classicality. The temptation that leads one to this conclusion starts with an observation that—by definition—a separable density matrix is a mixture of density matrices

$$\rho_{\mathcal{SA}} = \sum_i p_i \rho^i_{\mathcal{SA}} \tag{3.19}$$

that have explicit product eigenstates,

$$\rho^i_{\mathcal{SA}} = \sum_j p^{(i)}_j \left|s^{(i)}_j\right\rangle \left|a^{(i)}_j\right\rangle \left\langle a^{(i)}_j\right| \left\langle s^{(i)}_j\right| \tag{3.20}$$

and hence classical correlations. One might have thought that mixing such obviously classical density matrices cannot bring in anything quantum: after all, it involves only loss of information—forgetting of the label i in $\rho^i_{\mathcal{SA}}$. Yet this is not the case. One symptom of the quantumness of a separable $\rho_{\mathcal{SA}}$ with non-zero discord is immediately apparent: unless there exists a complete set of projectors $\{\Pi^{\mathcal{A}}_j\}$ for which $\delta(\mathcal{S} : \mathcal{A})_{\{\Pi^{\mathcal{A}}_j\}} = 0$, $\rho_{\mathcal{SA}}$ is perturbed by *all* local measurements. By contrast, when $\delta(\mathcal{S} : \mathcal{A})_{\{\Pi^{\mathcal{A}}_j\}} = 0$, then the measurement $\{\Pi^{\mathcal{A}}_j\}$ on $\mathcal{A}$, and an appropriate conditional measurement (i.e. conditioned by the outcome of the measurement on $\mathcal{A}$) will reveal all of the information in $\mathcal{S}$–$\mathcal{A}$, i.e. the resulting state of the pair will be pure. Moreover this procedure can be accomplished without perturbing the $\rho_{\mathcal{SA}}$ for another observer, a bystander not aware of the outcomes.

Thus, for each outcome j there exist a set $\{\pi^{\mathcal{S}}_{j,k}\}$ of conditional one dimensional projectors such that

$$\rho_{\mathcal{SA}} = \sum_j \sum_k \pi^{\mathcal{S}}_{j,k} \Pi^{\mathcal{A}}_j \rho_{\mathcal{SA}} \Pi^{\mathcal{A}}_j \pi^{\mathcal{S}}_{j,k}, \tag{3.21}$$

and $\pi^{\mathcal{S}}_{j,k} \Pi^{\mathcal{A}}_j \rho_{\mathcal{SA}} \Pi^{\mathcal{A}}_j \pi^{\mathcal{S}}_{j,k}$ is pure for any j and k. Above, the sets $\{\pi^{\mathcal{S}}_{j,k}\}$ for different j will not coincide in general ($\{\pi^{\mathcal{S}}_{j,k}\}$ is a function of j) and do not need to commute. *Classical information is locally accessible, and can be obtained without perturbing the state of the system*: one can interrogate just one part of a composite system and discover its state while leaving the overall density matrix (as perceived by observers that do not have access to the measurement outcome) unaltered. A general separable $\rho_{\mathcal{SA}}$ does not allow for such insensitivity to measurements: information can be extracted from the apparatus but only at a price of perturbing $\rho_{\mathcal{SA}}$, even when this density matrix is separable. However, when discord disappears, such insensitivity (which may be the defining feature of "classical reality", as it allows acquisition of information without perturbation of the underlying

state) becomes possible for correlated quantum systems. This quantum character of separable density matrices with non zero discord is a consequence of the superposition principle for $\mathcal{A}$, since more than one basis $\left\{\left|a_j^{(i)}\right\rangle\right\}_j$ for the apparatus is needed in Eq. (3.20) in order to warrant a non vanishing discord.

The difference between separability and vanishing discord can be illustrated by a specific example. Fig. 3.2 shows discord for a Werner state $\rho_{\mathcal{SA}} = \frac{1-z}{4} I + z\,|\psi\rangle\langle\psi|$ with $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. It can be seen that discord is greater than 0 in any basis when $z > 0$, which contrasts with the well-known separability of such states when $z < 1/3$.

3.4 Conclusion

The quantum discord is a measure of the information that cannot be extracted by the reading of the state of the apparatus (i.e. without joint measurements). Hence the quantum discord is a good indicator of the quantum nature of the correlations. The pointer states obtained by minimizing the quantum discord over the possible measurements should coincide with the ones obtained with the predictability sieve criterion [Zur93b, PZ01a], hence showing that the accessible information remains in the most stable *pointer* states.

3.5 Proofs

Proposition 3.5.1. *The following equality holds*

$$S(\mathcal{S}|\{\Pi_j^{\mathcal{A}}\}) = S(\rho_{\mathcal{SA}}^D) - S(\rho_{\mathcal{A}}^D), \tag{3.22}$$

with $\rho_{\mathcal{SA}}^D = \sum_j \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} \Pi_j^{\mathcal{A}}$.

Proof 3.5.1. $\rho_{\mathcal{SA}}^D$ is block-diagonal. The j-th block equals $p_j \rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}}$. By doing calculations block by block one has:

$$\begin{aligned} S(\rho_{\mathcal{SA}}^D) &= \sum_j S(p_j \rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}}) && (3.23)\\ &= \sum_j p_j S(\rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}}) - \sum_j p_j \log p_j && (3.24)\\ &= S(\mathcal{S}|\{\Pi_j^{\mathcal{A}}\}) + H(\rho_{\mathcal{A}}^D), && (3.25) \end{aligned}$$

which completes the proof. ∎

Proposition 3.5.2. $\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} \geq 0$.

Proof 3.5.2. This is a direct consequence of the previous proposition and of the concavity of $S(\rho_{\mathcal{SA}}) - S(\rho_{\mathcal{A}})$ with respect to $\rho_{\mathcal{SA}}$[Weh78a]. ∎

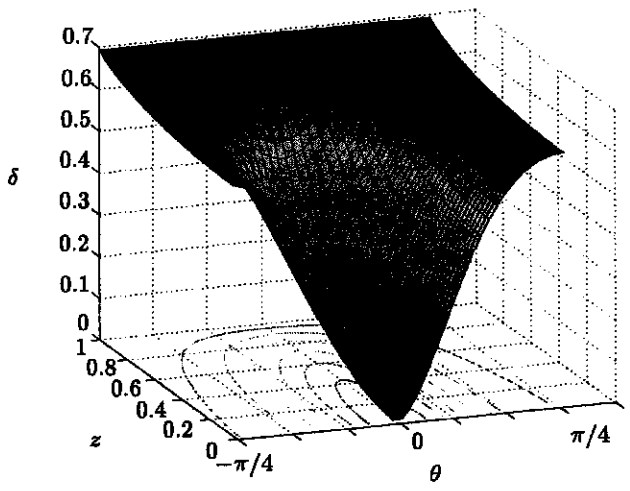

Figure 3.1: Value of the discord δ as a function of z and θ, for the states given by $(|00\rangle\langle 00| + |11\rangle\langle 11| + z\,|00\rangle\langle 11| + z\,|11\rangle\langle 00|)/2$, and a measurement basis $\{\cos(\theta)\,|0\rangle + e^{i}\sin\theta\,|1\rangle\,, e^{-i}\sin\theta\,|0\rangle - \cos\theta\,|1\rangle\}$.

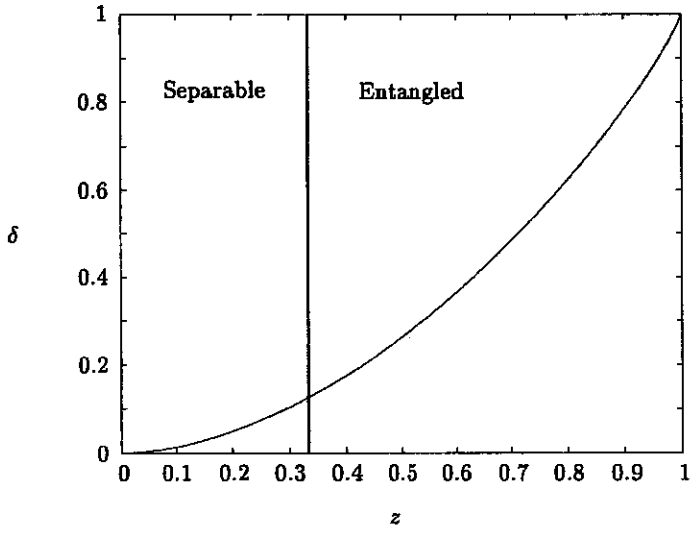

Figure 3.2: Value of the discord δ as a function of z for Werner states $\frac{1-z}{4}I + z\,|\psi\rangle\langle\psi|$, with $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. Discord does not depend on the basis of measurement in this case because both I and $|\psi\rangle$ are invariant under local rotations.

Proposition 3.5.3. $\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} = 0 \Leftrightarrow \rho_{\mathcal{SA}} = \sum_j \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} \Pi_j^{\mathcal{A}}$.

Proof 3.5.3. Proposition 3.5.1 already shows the converse. To prove the direct implication we will start with $\rho_{\mathcal{SA}}$ and $\{\Pi_j^{\mathcal{A}}\}$, a complete set of orthogonal projectors, such that $\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} = 0$. Without loss of generality, we can write the density matrix of $\mathcal{S}$–$\mathcal{A}$ as

$$\rho_{\mathcal{SA}} = \sum_j \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} \Pi_j^{\mathcal{A}} + \text{additional terms.} \tag{3.26}$$

If we choose $\{|s_i\rangle\}$ a basis of $\mathcal{S}$, and $\{|a_{k|j}\rangle\}_k$ a basis of the subspace of $\mathcal{A}$ defined by $\Pi_j^{\mathcal{A}}$, the general form of the additional terms in the above formula will be $c(s_i, s_{i'}, a_{k|j}, a_{k'|j'}) \times |s_i\rangle\langle s_{i'}| \otimes |a_{k|j}\rangle\langle a_{k'|j'}|$ with $j \neq j'$. Suppose that one of those coefficients is non-zero. By changing the basis $\{|s_i\rangle\}_i$, we can suppose $i \neq i'$. We introduce now a new density matrix $\hat{\rho}_{\mathcal{SA}}$ obtained from $\rho_{\mathcal{SA}}$ by removing the preceding matrix element and its complex conjugate. This ensures that $\hat{\rho}_{\mathcal{SA}}$ is associated with a physical state. This state thus satisfies

$$S(\hat{\rho}_{\mathcal{SA}}) > S(\rho_{\mathcal{SA}}) \text{ and } S(\hat{\rho}_{\mathcal{A}}) = S(\rho_{\mathcal{A}}). \tag{3.27}$$

The concavity of $S(\mathcal{S},\mathcal{A}) - S(\mathcal{A})$ implies inequalities:

$$\begin{aligned} S(\rho_{\mathcal{SA}}) - S(\rho_{\mathcal{A}}) &< S(\hat{\rho}_{\mathcal{SA}}) - S(\hat{\rho}_{\mathcal{A}}), \\ S(\hat{\rho}_{\mathcal{SA}}) - S(\hat{\rho}_{\mathcal{A}}) &\leq S(\rho_{\mathcal{SA}}^{D}) - S(\rho_{\mathcal{A}}^{D}). \end{aligned}$$

Then $\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} < 0$, which contradicts our primary assumption and proves our last result. ∎

Remark 3.5.1. We defined $\mathcal{J}$ with the help of a measurement associated to one-dimensional projectors. One can be interested in looking at multi-dimensional projective measurements. Depending on the context, two different generalization can be used.

For measurement purposes, one may adopt

$$\rho_{\mathcal{S}|\Pi_j^{\mathcal{A}}} = \mathrm{Tr}_{\mathcal{A}} \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} / \mathrm{Tr}_{\mathcal{SA}} \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}}, \tag{3.28}$$

since all the correlations (quantum as well as classical) between $\mathcal{S}$ and the subspace of the apparatus defined by $\Pi_j^{\mathcal{A}}$ are not observed. Proposition 1 no longer holds, but using the same techniques we still have $\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} \geq 0$ and if $\delta(\mathcal{S}:\mathcal{A})_{\{\Pi_j^{\mathcal{A}}\}} = 0$, then $\rho_{\mathcal{SA}} = \sum_j \Pi_j^{\mathcal{A}} \rho_{\mathcal{SA}} \Pi_j^{\mathcal{A}}$.

For decoherence purposes, one may prefer to define $\mathcal{J}$ as $S(\mathcal{S}) + S(\mathcal{A})^{D} - S(\mathcal{S},\mathcal{A})^{D}$. With this definition, Proposition 3 is valid. $\mathcal{J}$ now represents the average information, quantum and classical that remains in the pair $\mathcal{S}$–$\mathcal{A}$ after a decoherence process leading to einselection of the superselection sectors $\{\Pi_j^{\mathcal{A}}\}$. ∎

4 Objective properties from subjective quantum states: environment as a witness

THE KEY FEATURE distinguishing the classical realm from the quantum substrate is its objective existence. Classical states can be found out through measurements by an initially ignorant observer without getting disrupted in the process. By contrast, an attempt to discover the state of a quantum system through a direct measurement generally leads to a collapse [Boh28a, vN32a, Dir47a]: after a measurement, the state will be what the observer finds out it is, but not—in general—what it was before. Thus, it is difficult to claim that quantum states exist objectively in the same sense as their classical counterparts [Boh27a, EPR35a, FP00a].

It is by now widely appreciated that decoherence, caused by persistent monitoring of a system by the environment, can single out a preferred set of states. In simplest models, such pointer states [Zur93b, Zur98a, Zur00a, PZ01a, Zur03a, GJK+96a] are (often degenerate) eigenstates of the pointer observable which commutes with the system-environment interaction [Zur82a]. This concept can be generalized using the predictability sieve: only pointer states evolve predictably in spite of the openness of the system of interest [Zur93b, PZ01a, Zur03a]. They exist in the sense that in absence of any perturbations—save for the monitoring by the environment—they or their dynamically evolved descendants will continue to faithfully represent the state of the system. Thus, when an observer knows what is the set of pointer states, he can learn which of them represents the system without perturbing it. However, when an initially ignorant observer attempts to find out the state of the system directly, he still faces, even in the presence of decoherence, the danger of collapsing its wave packet.

Here, we build on the idea that a *direct* measurement of the system is not how observers gather data about the Universe: rather, a vast majority (if not all) of our information is obtained *indirectly* by probing a small fraction of the environment [Zur93b, Zur98a, Zur00a, Zur03a]. One may think that this twist in the story can be accounted for by adding a few links to the von Neumann chain [vN32a], but this is not the case: we shall show that the monitoring environment acquires information about the system *selectively*. More importantly, this selective spreading of information through the environment—in essence "quantum Darwinism" [Zur03a]—accounts for the objective existence of some preferred quantum states: by probing the system indirectly, hence without perturbing it, many independent observers can obtain reliable information, but only about the pointer states.

This chapter is organized as follows: first, we introduce our operational definition of objectivity. We then derive two requirements for the objective existence of an observable in the context of *einselection* (environment-induced superselection). Next, these are translated into an information theoretical framework. We prove that these requirements imply a *unique* observable: the usual pointer observable. Finally, we show that, because of quantum Darwinism, information about the pointer states is robust and, hence, objective: it can be extracted from fragments of the environment in realistic settings, where measurements are restricted and imperfect.

4.1 Requirements for objectivity

Objectivity of a property of a quantum system should not rely on pre-existence of an underlying reality as it is presumed in the classical setting. Rather, we demand that an objective property of the system of interest should be (*i*) simultaneously accessible to many observers, (*ii*) who should be able to find out what it is without prior knowledge and (*iii*) who should arrive at a consensus about it without prior agreement. As we already mentioned, the collapse of the wave packet following a direct measurement generally precludes this. However, when the system of interest $\mathcal{S}$ interacts with an environment $\mathcal{E}$ composed of many subsystems, $\mathcal{E} = \bigotimes_{k=1}^{N} \mathcal{E}_k$, the situation changes dramatically. When a property leaves a *complete* and *redundant* imprint on the environment, all three criteria are satisfied: because many copies are available (*i*) is straightforward. Moreover fractions of the environment can be measured without perturbing either the system or the rest of $\mathcal{E}$. Therefore, ignorant observers can vary their measurement strategies independently, corroborate their own results and arrive at a common description of the properties of the system. Hence, (*ii*) and (*iii*), also follow from redundancy. We shall thus identify the existence of an objective property with the existence of its complete and redundant imprint in the environment. As a consequence, our approach will focus on the study of the correlations between parts of the environment and the system of interest.

4.2 Information theoretical framework

A natural way to characterize such correlations is to use the mutual information $I(\sigma : \mathfrak{e})$ between an observable σ of $\mathcal{S}$ and a measurement $\mathfrak{e}$ on $\mathcal{E}$. In short, $I(\sigma : \mathfrak{e})$ measures one's ability to predict the outcome of measurement σ on $\mathcal{S}$ after having "peeked at the environment" through $\mathfrak{e}$. Formally, for a given density matrix $\rho^{\mathcal{SE}}$ of $\mathcal{S} \otimes \mathcal{E}$, the measurement results are random variables characterized by a joint probability distribution $p(\sigma_i, \mathfrak{e}_j) = \mathrm{Tr}\{(\sigma_i \otimes \mathfrak{e}_j)\rho^{\mathcal{SE}}\}$, where σ_i and $\mathfrak{e}_j$ are the spectral projectors of observables σ and $\mathfrak{e}$. By definition, the mutual information is the difference between the initially missing information about σ and the remaining uncertainty about σ when $\mathfrak{e}$ is known [CT91a]. Using Shannon entropy as a measure of missing information, $H(\sigma) = -\sum_i p(\sigma_i) \log p(\sigma_i)$ and $H(\sigma, \mathfrak{e}) = -\sum_{i,j} p(\sigma_i, \mathfrak{e}_j) \log p(\sigma_i, \mathfrak{e}_j)$, the mutual information is naturally defined

as $I(\sigma : \mathfrak{e}) = H(\sigma) + H(\mathfrak{e}) - H(\sigma, \mathfrak{e})$.

The first requirement for objectivity states that only when the direct measurement σ on $\mathcal{S}$ can be replaced by a measurement on $\mathcal{E}$ without noticeable loss of information, can we hope to find out about σ from the environment. This is characterized by the information about σ which can be optimally extracted from m typical environmental subsystems $\hat{I}_m(\sigma) = \max_{\{\mathfrak{e}\in\mathfrak{M}_m\}} I(\sigma : \mathfrak{e})$, where $\mathfrak{M}_m$ is the set of all measurements on those m subsystems.[1] Then, with $\mathcal{E}$ acting as a medium, the ability to eventually gain nearly all—all but a fraction $\gamma \ll 1$—of the information from the environment,

$$\hat{I}_N(\sigma) \geq (1-\gamma)H(\sigma) \tag{4.1}$$

corresponds to the complete imprinting of σ on $\mathcal{E}$. Above, N is the number of subsystems in the whole environment.

However, as a consequence of basis ambiguity [EPR35a, Eve57a, Zeh73a, Ken90a, Zur82a], information about many observables σ can be deduced by an appropriate measurement on the entire environment. Therefore, the total transfer of information, while a prerequisite for objectivity, is not a very selective criterion (see Fig. 4.1a). Clearly, to claim objectivity, it is not sufficient to have a complete imprint of the candidate property of $\mathcal{S}$ in the environment. There must be many copies of this imprint that can be accessed independently by many observers: information must be redundant.

4.3 Redundancy and its consequences

A rough estimate of the redundancy of the information about σ present in the environment can be derived from the average information about σ that can be obtained from a single subsystem of $\mathcal{E}$, $\hat{I}_1(\sigma)$. We can expect that the corresponding redundancy will be of order $N \times \hat{I}_1(\sigma)/\hat{I}_N(\sigma)$, which is indeed a useful estimate [Zur03a], but an overestimate. This is because successive measurements are likely to confirm the information about σ the observer already has. Only some of the newly acquired data will be really new: redundant information is not extensive [CT91a], $\hat{I}_m(\sigma) \neq m \times \hat{I}_1(\sigma)$.

To obtain a measure of redundancy which takes this into account, one must count the number of copies of the information $\hat{I}_N(\sigma)$ embedded in $\mathcal{E}$. To arrive at a formal definition, however, we must agree to acquire almost all—all but a fraction δ—of the information about σ present in the entire environment $\mathcal{E}$. Redundancy is thus quantified by the number of disjoint subsets of $\mathcal{E}$ containing almost all the information about σ:

$$R_\delta(\sigma) = N/m_\delta(\sigma). \tag{4.2}$$

Above $m_\delta(\sigma)$ is precisely the smallest number m of environmental subsystems that contain almost all the information about σ, i.e. $\hat{I}_m(\sigma) \geq (1-\delta)\hat{I}_N(\sigma))$.[1]

[1] For sake of simplicity we assume that $I(\sigma : \mathfrak{e})$ does not depend crucially on a particular choice of subsystems of the environment on which $\mathfrak{e}$ is applied. However, the generalization of our results to such cases is possible and straightforward.

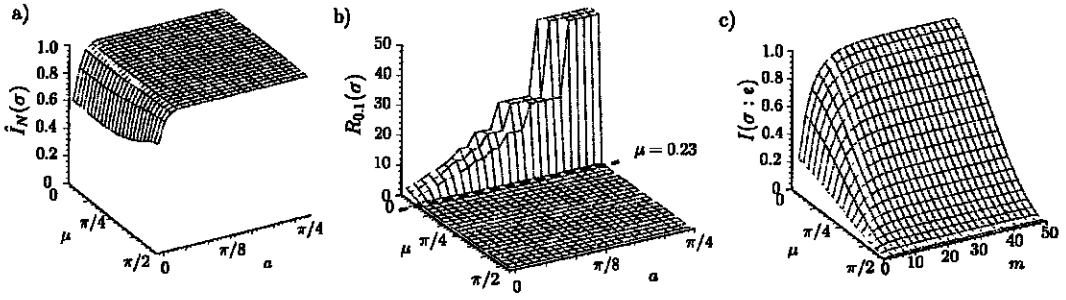

Figure 4.1: Quantum Darwinism can be illustrated using a model introduced in [Zur82a]. The system $\mathcal{S}$, a spin-$\frac{1}{2}$ particle, interacts with $N = 50$ two-dimensional subsystems of the environment through $\hat{H}^{\mathcal{SE}} = \sum_{k=1}^{N} g_k \sigma_z^{\mathcal{S}} \otimes \sigma_y^{\mathcal{E}_k}$ for a time t. The initial state of $\mathcal{S} \otimes \mathcal{E}$ is $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle^{\mathcal{E}_1} \otimes \ldots \otimes |0\rangle^{\mathcal{E}_N}$. All the plotted quantities are function of the system's observable $\sigma(\mu) = \cos(\mu)\sigma_z + \sin(\mu)\sigma_x$, where μ is the angle between its eigenstates and the pointer states of $\mathcal{S}$—here the eigenstates of $\sigma_z^{\mathcal{S}}$. **a)** Information acquired by the optimal measurement $\mathfrak{e}$ on the whole environment, $\hat{I}_N(\sigma)$, as a function of the inferred observable $\sigma(\mu)$ and the action $a_k = g_k t = a$ for all k. A large amount of information is accessible in the *whole* environment for any observables $\sigma(\mu)$ except when the interaction action a_k is very small. Thus, complete imprinting of an observable of $\mathcal{S}$ in $\mathcal{E}$ is not sufficient to claim objectivity. **b)** Redundancy of the information about the system as a function of the inferred observable $\sigma(\mu)$ and the action $a_k = g_k t = a$. It is measured by $R_{\delta=0.1}(\sigma)$, which counts the number of times 90% of the total information can be "read off" independently by measuring distinct fragments of the environment. For all values of the action $a_k = g_k t = a$, redundant imprinting is sharply peaked around the pointer observable. Redundancy is a very selective criterion. The number of copies of relevant information is high only for the observables $\sigma(\mu)$ falling inside the theoretical bound (see text) indicated by the dashed line. **c)** Information about $\sigma(\mu)$ extracted by an observer restricted to local random measurements on m environmental subsystems (e.g. $\mathfrak{e} = \mathfrak{e}^{\mathcal{E}_1} \otimes \ldots \otimes \mathfrak{e}^{\mathcal{E}_m}$ where each $\mathfrak{e}^{\mathcal{E}_k}$ is chosen at random). The interaction action $a_k = g_k t$ is randomly chosen in $[0, \pi/4]$ for each k. Because of redundancy, pointer states—and only pointer states—can be found out through this far-from-optimal measurement strategy. As our theorem establishes, the information about any other observable $\sigma(\mu)$ is equal to the information that can be obtained through its correlations with the pointer observable $\sigma_z^{\mathcal{S}}$. Information about any other observable $\sigma(\mu)$ is restricted by our theorem to be equal to the information brought about it by the pointer observable $\sigma_z^{\mathcal{S}}$, Eq. (4.3).

We now analyze what information about $\mathcal{S}$ can be shared among many fragments of $\mathcal{E}$. The answer is supplied by our main result:

Proposition 4.3.1. *When an observable ζ is (i) completely, $\hat{I}_N(\zeta) = H(\zeta)$, and (ii) redundantly, $R_{\delta=0}(\zeta) > 1$, imprinted on the environment, there exists an observable π such that,*

- *π satisfies (i) and (ii) ;*
- *the information about any other observable σ, completely and redundantly imprinted on the environment, attainable by measuring $m = m_{\delta=0}(\sigma)$ is equivalent to the information about σ that can be obtained through its correlations with π:*

$$\hat{I}_m(\sigma) = I(\sigma : \pi). \tag{4.3}$$

The observable π is then called the maximally refined observable.

The proof of this proposition is given at the end of this chapter.

Two important consequences of this proposition readily follow. (*i*) An observer who probes only a fraction of the environment is able to find out the state of the system as if he had carried out the direct measurement of π on $\mathcal{S}$. (*ii*) On the other hand, information about any other observable σ of $\mathcal{S}$ will inevitably be limited by the available correlations existing between σ and π. In essence, our proposition proves the uniqueness of redundant information, and therefore the selectivity of its proliferation.

Quantum Darwinism—the idea that the perceived *classical reality* is a consequence of the selective proliferation of information about the system [Zur03a]—is consistent with previous approaches to einselection, such as the predictability sieve, but goes beyond them. The existence of redundant information about the system, induced by specific interactions with the environment, completely defines what kind and how information can be retrieved from $\mathcal{E}$: Eq. (4.3) shows that an efficient strategy for inferring σ consists in estimating π first, and deducing from it information about σ. It also explains the emergence of a consensus about the properties of a system. Observers attempting to gain information about π—the only kind of information available in fragments of $\mathcal{E}$—will agree about their conclusions: their measurement results are directly correlated with π, and are therefore correlated with each other. Hence, observers probing fractions of the environment can act *as if* the system had a state of its own—an *objective* state (one of the eigenstates of π). By contrast, such consensus cannot arise for superpositions of pointer states, e.g. Schrödinger cats, since information about them can only be extracted by probing the whole environment, and thus cannot be obtained independently by several observers. Objectivity comes at the price of singling out a preferred observable of $\mathcal{S}$ whose eigenstates are redundantly recorded in $\mathcal{E}$. Cloning of quanta is not possible [WZ82a], but amplification of a preferred observable happens almost as inevitably as decoherence and leads to objective classical reality. The impossibility of cloning and the capacity for amplification imply selection—Darwinian "survival of the fittest".

4.4 Emergence of objectivity exemplified

In Figure 4.1b, we show for a specific model the redundancy as a function of the inferred observable $\sigma(\mu)$ (whose eigenstates are tilted by an angle μ from the pointer ones) and of the interaction action, $a_k = g_k t$ (a_k characterizes the strength of the correlations between $\mathcal{S}$ and $\mathcal{E}_k$). By carefully tracking all orders of γ and δ in Eqs. (4.1-4.3), one can show that the existence of a complete and redundant imprint of observable $\sigma(\mu)$ in the environment requires $H_2(\cos^2 \frac{\mu}{2}) \leq \delta$, where $H_2(p) = -[p \log p + (1-p)\log(1-p)]$. Inserting the actual values of the parameters chosen for our simulation, the above equation indicates that only observables with $|\mu| < 0.23$ leave a redundant imprint on the environment. This bound is in excellent agreement with our numerical results. Objective properties of the system are indeed defined uniquely by the value of the usual pointer observable $\sigma_z^{\mathcal{S}}$. Surprisingly, and as confirmed by our simulation, the interaction action a_k only plays a role in setting the value of the redundancy at its maximum, but does not affect the selectivity of our criterion. Which observable becomes objective is largely decided by the structure of the interaction Hamiltonian, (i.e. the set of pointer states), but not by its details such as strength and duration of the interaction. This ensures the stability of the pointer observable deduced from redundancy.

4.5 Robustness of information

One can gain further insight into the role of redundancy by considering an analogy between the environment and a noisy communication channel. As in classical coding theory, a high redundancy protects the information against a wide range of errors. In the context of environment-induced superselection, it also holds the key to the objective existence of pointer states: the most common "error"—which as a rule happens always in the course of our everyday measurements, and which can be only rarely avoided in carefully controlled laboratory experiments—is the loss of most environmental subsystems. However, thanks to redundancy, information about the einselected states is still available from small fragments of $\mathcal{E}$.

Non-optimal measurements on the environment is another form of error which should be also overcome by redundancy. Objective information must be extractable through "realizable"—hence, not necessarily optimal—measurements for many observers to arrive at an operational consensus about the state of a system. For instance, human eyes can only measure photons *separately*, yet we can still learn about the position of objects. This issue is considered for our model in Fig. 4.1c. Here, even local (i.e. spin by spin) *random* measurements eventually acquire the entire information available in $\mathcal{E}$ about the pointer states. Though surprising, this result naturally follows from quantum Darwinism. Almost any observable of $\mathcal{S}$ is completely imprinted on the environment (see Fig. 4.1a). However, as our proposition establishes and Fig. 4.1b illustrates (for non γ and δ small), only the observables which can be completely deduced from the maximally refined one, $\sigma_z^{\mathcal{S}}$, are imprinted redundantly in the environment. Therefore only information about pointer states can tolerate errors, i.e. can be extracted by non-optimal

measurements. In short, *not only is the information about the pointer observable easy to extract from fragments of the environment, it is impossible to ignore!*

4.6 Conclusion

Quantum Darwinism capitalizes on some ideas that arose in the context of decoherence and einselection, but goes beyond them in an essential fashion. Existence of records in the environment has been noted before [Zur93b, GJK+96a, Zur98a, Zur00a, Zur82a, Hal99b, BHJ03a], and the fact that it is easiest to find out about the pointer observable has been also appreciated [DDZ01a]. Here, however, we have described an even more dramatic turn of events—environment as a broadcast medium—, which may seem fanciful until we realize that it describes rather accurately what happens in the real world. For instance, human observers acquire all of their visual data by intercepting a small fraction of their photon environment. An operational notion of objectivity emerges from redundant information as it enables many independent observers to find out the state of the system without disturbing it. Furthermore, *objective* observables are robust—insensitive to changes in the strategy through which the environment is interrogated, as well as to variations of the strength and duration of the interaction between $\mathcal{S}$ and $\mathcal{E}$, etc.

4.7 Proof

First, we briefly introduce the notation that used in the proof.

- The system $\mathcal{S}$ has interacted with an environment $\mathcal{E} = \bigotimes_{k=1}^{N} \mathcal{E}_k$, and their common state is represented by the combined density matrix $\rho^{\mathcal{SE}}$. From this common state we derive reduced density matrices for the system, or the system and part of the environment, $\rho^{\mathcal{S}} = \mathrm{Tr}_{\mathcal{E}}\, \rho^{\mathcal{SE}}$, $\rho^{\mathcal{SE}_k} = \mathrm{Tr}_{\mathcal{E}-\mathcal{E}_k}\, \rho^{\mathcal{SE}}$, etc.
- π, σ, and ζ are (possibly degenerate) observables on the system $\mathcal{S}$, and $\{\pi_i\}_i$, $\{\sigma_i\}_i$, and $\{\zeta_i\}_i$ their spectral decompositions.
- $\mathfrak{p}^{\mathcal{E}_k}$, $\mathfrak{s}^{\mathcal{E}_k}$ and $\mathfrak{z}^{\mathcal{E}_k}$ are measurements on the environment $\mathcal{E}_k$, and $\{\mathfrak{p}^{\mathcal{E}_k}\}_j$, $\{\mathfrak{s}^{\mathcal{E}_k}\}_j$, $\{\mathfrak{z}^{\mathcal{E}_k}\}_j$ their respective spectral decompositions.

Theorem 4.7.1. *Suppose that $I(\sigma : \mathfrak{s}^{\mathcal{E}_k}) = H(\sigma)$, and $I(\zeta : \mathfrak{s}'^{\mathcal{E}_k}) = H(\zeta)$ for all k. Then, there exist a measurement $\mathfrak{p}$ on the environment $\mathcal{E}_k$ such that it gives complete information about an observable π on $\mathcal{S}$ which commutes with both σ and ζ. This observable is more refined than σ and ζ in the sense that $I(\sigma : \pi) = H(\sigma)$ and $I(\zeta : \pi) = H(\zeta)$.*

Note that for simplicity we assumed that the partition of the environment is such that each subsystem $\mathcal{E}_k$ has one copy of the information about σ or ζ. This can be

obtained from the general case by grouping subsystems of the environment to make larger subsystems that satisfy this property.

Before going to the proof we give an intermediate proposition.

Proposition 4.7.1. *Under the asumptions of the theorem, the observables σ and ζ commute on the support of ρ^S.*

Proof 4.7.1. We start the proof by noting the following fact. The hypotheses impose that for any k:

$$H(\sigma|\mathfrak{s}^{\mathcal{E}_k}) = 0 = \sum_j p^k(j) H(\sigma|\mathfrak{s}_j^{\mathcal{E}_k}), \tag{4.4}$$

where the different probabilities involved are defined by, $p^k(i,j) = \mathrm{Tr}\,(\rho^{S\mathcal{E}}\sigma_i \otimes \mathfrak{s}_j^{\mathcal{E}_k})$, $p^k(j) = \sum_i p^k(i,j)$, and $p^k(i|j) = p^k(i,j)/p^k(j)$. The equality of Eq. (4.4) can be satisfied if and only if there exist a function $i^k(j)$ such that $p(i^k(j)|j) = 1$ for all j's.

We now construct a new measurement $\tilde{\mathfrak{s}}^{\mathcal{E}_k}$ whose projectors are defined by:

$$\tilde{\mathfrak{s}}_i^{\mathcal{E}_k} = \sum_{j,\, i^k(j)=i} \mathfrak{s}_j^{\mathcal{E}_k}. \tag{4.5}$$

This very same construction can be performed to obtain $\tilde{\mathfrak{z}}^{\mathcal{E}_k}$. Now, these two measurements retain the original property,

$$H(\sigma|\tilde{\mathfrak{s}}^{\mathcal{E}_k}) = H(\zeta|\tilde{\mathfrak{z}}^{\mathcal{E}_k}) = 0, \tag{4.6}$$

but now also satisfy

$$H(\tilde{\mathfrak{s}}^{\mathcal{E}_k}|\sigma) = H(\tilde{\mathfrak{z}}^{\mathcal{E}_k}|\zeta) = 0. \tag{4.7}$$

Hence, $\mathrm{Tr}\,\rho^{S\mathcal{E}}\sigma_i = \mathrm{Tr}\,\rho^{S\mathcal{E}}\sigma_i \otimes \tilde{\mathfrak{s}}_i^{\mathcal{E}_k} = \mathrm{Tr}\,\rho^{S\mathcal{E}}\tilde{\mathfrak{s}}_i^{\mathcal{E}_k}$, from which we deduce that

$$\sigma_i\rho^{S\mathcal{E}}\sigma_i = \sigma_i \otimes \tilde{\mathfrak{s}}_i^{\mathcal{E}_k}\rho^{S\mathcal{E}}\sigma_i \otimes \tilde{\mathfrak{s}}_i^{\mathcal{E}_k} = \tilde{\mathfrak{s}}_i^{\mathcal{E}_k}\rho^{S\mathcal{E}}\tilde{\mathfrak{s}}_i^{\mathcal{E}_k}. \tag{4.8}$$

In other terms, because of the one-to-one correspondence between measurements on the system and measurement on the environment, making an indirect measurement of σ by using $\tilde{\mathfrak{s}}^{\mathcal{E}_k}$ leads to the same update rule of the reduced density matrix $\rho^{S\mathcal{E}}$ as if the direct measurement σ was performed.

The same series of equalities also hold for ζ and $\tilde{\mathfrak{z}}^{\mathcal{E}_k}$, and for all values of k.

Suppose that $\tilde{\mathfrak{s}}^{\mathcal{E}_1}$ and $\tilde{\mathfrak{z}}^{\mathcal{E}_2}$ are performed. Because these two measurements are applied to different parts of the environment, they commute and we must have (using the equalities above and taking the partial trace over $\mathcal{E}$):

$$\sigma_i\zeta_j\rho^S\zeta_j\sigma_i = \zeta_j\sigma_i\rho^S\sigma_i\zeta_j. \tag{4.9}$$

Using $(\sigma_i)^2 = \sigma_i$ and $(\zeta_j)^2 = \zeta_j$, we conclude that $\zeta_j\sigma_i\zeta_j\sigma_i = \zeta_j\sigma_i$ on the support of ρ^S ($\sigma\zeta$ is trace decreasing). Therefore $\sigma_i\zeta_j$ is a projector on the support of ρ^S, and we have $\zeta_j\sigma_i = \sigma_i\zeta_j$ on the support of ρ^S. ∎

The theorem now follows easily:

Proof 4.7.2. Let us consider the observable $\pi = \{\pi_{ij}\}_{ij} = \{\sigma_i\zeta_j\}_{ij}$. This is a legitimate observable since it is hermitian as a consequence of the previous proposition.

By the same kind of argument as the one employed for showing that σ and ζ commute, it is straightforward to show that for any $\mathcal{E}_k$, $\tilde{\mathfrak{s}}^{\mathcal{E}_k}$ and $\tilde{\mathfrak{z}}^{\mathcal{E}_k}$ commute. Therefore, $\mathfrak{p}^{\mathcal{E}_k} = \{\mathfrak{p}_{ij}^{\mathcal{E}_k}\}_{ij} = \{\tilde{\mathfrak{s}}_i^{\mathcal{E}_k}\tilde{\mathfrak{z}}_j^{\mathcal{E}_k}\}_{ij}$ defines a legitimate measurement.

Firstly, because σ and σ' commute, once π_{ij} has been obtained, the outcome of any subsequent measurement σ or ζ is deterministic:

$$\operatorname{Tr}\sigma_i\pi_{ij}\rho^S\pi_{ij}\sigma_i = \operatorname{Tr}\pi_{ij}\rho^S\pi_{ij}, \tag{4.10}$$

and similarly for ζ_j. Hence, we have $I(\sigma:\pi) = H(\sigma)$ and $I(\zeta:\pi) = H(\zeta)$.

The last part of the proof consists in showing that the measurement $\mathfrak{p}^{\mathcal{E}_k}$ allows to retrieve all the information about π. This is easy to check, using the commutation relations established earlier and the fact that each outcome of $\tilde{\mathfrak{s}}^{\mathcal{E}_k}$ (resp. $\tilde{\mathfrak{z}}^{\mathcal{E}_k}$) corresponds to a given outcome of σ (resp. ζ). For each value of (i,j) we have:

$$\begin{aligned}
\operatorname{Tr}\tilde{\mathfrak{s}}_i^{\mathcal{E}_1}\tilde{\mathfrak{z}}_j^{\mathcal{E}_1}\rho^{S\mathcal{E}}\tilde{\mathfrak{z}}_j^{\mathcal{E}_1}\tilde{\mathfrak{s}}_i^{\mathcal{E}_1} &= \operatorname{Tr}\zeta_j\otimes(\tilde{\mathfrak{s}}_i^{\mathcal{E}_1}\tilde{\mathfrak{z}}_j^{\mathcal{E}_1})\rho^{S\mathcal{E}}\zeta_j\otimes(\tilde{\mathfrak{z}}_j^{\mathcal{E}_1}\tilde{\mathfrak{s}}_i^{\mathcal{E}_1}) && (4.11)\\
&= \operatorname{Tr}\zeta_j\otimes(\tilde{\mathfrak{z}}_j^{\mathcal{E}_1}\tilde{\mathfrak{s}}_i^{\mathcal{E}_1})\rho^{S\mathcal{E}}\zeta_j\otimes(\tilde{\mathfrak{s}}_i^{\mathcal{E}_1}\tilde{\mathfrak{z}}_j^{\mathcal{E}_1}) && (4.12)\\
&= \operatorname{Tr}(\zeta_j\sigma_i)\otimes(\tilde{\mathfrak{z}}_j^{\mathcal{E}_1}\tilde{\mathfrak{s}}_i^{\mathcal{E}_1})\rho^{S\mathcal{E}}(\sigma_i\zeta_j)\otimes(\tilde{\mathfrak{s}}_i^{\mathcal{E}_1}\tilde{\mathfrak{z}}_j^{\mathcal{E}_1}) && (4.13)\\
&= \operatorname{Tr}\pi_{ij}\otimes\mathfrak{p}_{ij}^{\mathcal{E}_1}\rho^{S\mathcal{E}}\pi_{ij}\otimes\mathfrak{p}_{ij}^{\mathcal{E}_1} && (4.14)
\end{aligned}$$

which leads to $I(\pi_{ij}:\mathfrak{p}_{ij}^{\mathcal{E}_1}) = H(\pi)$, and concludes the proof. ∎

Prop. 4.3.1 is obtained by successively refining in this way any pair of completely and redundantly imprinted observables.

II Quantum convolutional error correcting codes

5 Quantum information processing: basics

THE FIRST HALF of the 20st century has ended the industrial revolution started in the middle of the 19th century. It also announced our modern world made of science, technology and information. It was a time of great scientific discoveries. General relativity is certainly the most widely "known" in the non-scientific public. Less spectacular, quantum mechanics and information theory had probably more impact on our everyday life. The ever increasing use of computers and more generally of information processing devices is the indicator of such a fruitful collaboration: quantum mechanics serves to design and understand semiconductors — the support of information —, while information theory and computer science define the potential uses and limits of automated treatment.

Both fields continue to grow very rapidly. New techniques are developed constantly, but recent years have witnessed the emergence of another kind of interaction between those theories: quantum information processing. It is a new computational paradigm resulting in the storage and manipulation of information in a quantum mechanical way. In quantum computers, quantum mechanics is not only used to design the physical devices but also to directly process information in a completely new and before unaccessible way.

This chapter will give a brief overview of the concepts and achievements in this field. It will only use some elementary notions of classical information theory, complexity and quantum mechanics. Before actually entering into the details of quantum information processing, we shall first explain rapidly the reasons that made the encounter of information with the quantum almost unavoidable.

5.1 Information and the quantum

The seminal work of A. Turing [Tur36a] had prepared the basis of modern computer science even before World War II. However, it remained quite a theoretical construction until it was used in practice to design machines aimed at decyphering texts automatically. In parallel, information theory was discovered by C. Shannon in 1948 [Sha48a]. Together with Turing's work it demonstrated the validity of the fundamental concepts related to information. For the very first time, it was realized that information was an abstract concept that could lead to a mathematical analysis: information is fungible — it is independent of its physical representation.

The discovery of this property of course promoted an abstract treatment of information. Simultaneously, it also encouraged the search for physical representations of

information in order to build processing units. Among all the candidates, it is the representation as an electrical charge which allows the greatest variety of uses. It is the only one — beside the representation of information in the brain — to allow not only storage and broadcast but automated processing. The first devices were build with vacuum tubes, but after the invention of transistors, the semiconductor technology became the only used for reliable memories and processors.

The advantages of the new technology were immediately recognized for information processing: the size of the devices was smaller than ever, thus allowing for a greater complexity in terms of the possible operations performed by the circuits. As an example of this unexpected breakthrough, Popular Mechanics wrote in 1949: "Computers in the future may weight no more than 1.5 tons". Even more unexpected was the rate at which these techniques developed. The complexity of processors started an incredible growth summarized by G. Moore [Moo65a]: "the number of transistors in a chip doubles approximately every 18 months". 40 years after this statement, this empirical law is still up to date. However, specialists of processor design seem more and more convinced that we will soon reach a fundamental limit to Moore's law. To continue building these always more complex devices, an exponential increase of the density of integrated circuits is needed to keep their size reasonable. A simple extrapolation of Moore's law shows that the size of transistors will approach the atomic scale around 2017 [Moo97a]. At this time, there will be no choice but to cooperate with quantum mechanics. Better be prepared to the new rules.

Will this be a limit to the capacity of performing more complex tasks with computers? If the quantum effects arising at this scale are seen only as limiting factors for information processing, the answer will be yes. But what if quantum mechanics could help? Part of the motivation in favor of the quantum computing idea relies on this hope. Various results confirmed the idea that, at least theoretically, quantum mechanics is improving our capacity to compute [Sho94a, Sim97a, Gro96a, KSV02a] and to communicate [Raz99a, BCT99a, Lo00a]. But large scale physical devices are still quite a way beyond our experimental achievements.

From a more physical point of view, information has gained all along the past century an increasing role. In fact, through the concepts of ignorance and disorder — two fundamentals of thermodynamics — physicists had an idea of the nature of information. This notion was first exploited in Boltzmann's formula for the entropy of a particle with fixed energy,

$$S = k_B \log W, \tag{5.1}$$

where k_B is a constant — converting bits into Joules — and W is the number of equivalent states with a fixed energy[1]. The entropy is then presented as the amount of ignorance about the state of the system, once its energy is fixed. Thermodynamics further developed the use of entropy to a high degree of refinement and naturally introduced information theoretical notions.

[1] More elaborated versions of the significance of W exists. In general, W is the number of states that satisfy a given set of constraints.

One of the best example illustrating the emergence of information in physics dates back to 1929 [Szi29a]. According to its 2nd law, thermodynamics forbids extraction of work out of a single heat bath. Nevertheless, Maxwell's demon seem to violate this principle. More precisely, suppose a single particle gas is contained in a box divided in two by a mobile piston. Initially, the position of the particle is unknown. The Maxwell's demon measures it, and transmits this information to a reversible mechanism, which will extract work from this situation by letting the piston relax in one or the other direction. This cycle can be started again by removing the piston, and by putting it reversibly back to the central position of the box. For each cycle, the whole device would then violate the 2nd law of thermodynamics — work is extracted from pressure fluctuations from a single heat bath. The solution to this paradox is to properly account for all the resources, and in particular those required by the information processing. For each pressure measurement, a properly initialized memory cell is needed. A reset operation must be performed after each movement of the piston. In turn, this reset is necessarily accompanied by entropy production which in this case just compensate for the extraction of work [Szi29a, Lan61a, Ben82b].

After all the efforts by the information theory community to promote information as an abstract concept, R. Landauer realized that [Lan91a] "information is physical". In other words, in spite of its fungibility information must be represented in a physical way to allow its automated treatment [Zur94a]: "no information without representation".

This success, proving the relevance of information for physics, was related to other questions. Among them, was the question of minimal energy and entropy requirements to make a calculation. Bennett concluded that any classical information processing could be done entirely in a reversible way [Ben73a]. Fredkin and Toffoli later proposed reversible circuit models of computation [FT82a]. This result contributed to draw attention to the ways of processing information at the physical level, i.e. according to the laws of physics.

It is thus quite naturally that quantum systems were examined on the basis of their computational capabilities. In 1982, R. P. Feynman promoted the study of the computing capabilities of quantum systems for simulating many-body quantum mechanics [Fey82a, Fey84a]. As we already mentioned in the first chapters of this manuscript, the classical simulation of n two-level quantum systems generally requires $O(2^n)$ complex numbers for classical computers whereas only $O(n)$ quantum bits[2] should be sufficient for a quantum computer. This conjecture was later proved by D. Deutsch [Deu85a, Deu89a]. Quantum computing is one alternative to continue performing more complex operations in spite of the inevitable stop in the miniaturization of classical electronic devices.

[2] A proper definition of a quantum bit is given in the next section.

5.2 Fundamentals

5.2.1 The qubit

The qubit is the elementary unit of quantum information. Like its classical analog — the bit — it is an abstract object which can be implemented in many different ways.

Definition 5.2.1 (Qubit). *A qubit is any system whose state space has a two dimensional Hilbert space available for information storage and manipulation according to the laws of quantum mechanics.*[3]

The canonical example of a qubit is a two-level quantum system — for instance NMR quantum processors use the nuclear spin of ^{1}H or ^{13}C. Here, the whole state space is used to perform information processing. It is the quantum computing convention to express the state of a qubit in a fixed but arbitrary orthonormal basis $\{|0\rangle, |1\rangle\}$, called the computational basis. In spite of this canonical example, it is important to realize that qubits are not necessary two-level systems. In fact any system with a two-dimensional subspace can serve as a qubit as well. This remark holds the key to error protection strategies such as quantum error correcting codes and noiseless subsystems [ZR97b, ZR97c, VKL00a].

The evolution of an isolated qubit is given by the axioms of quantum mechanics: it can undergo any unitary evolution of $SU(2)$. A useful set of such transformations is given by the Pauli matrices:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.2}$$

together with the Hadamard transform,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{5.3}$$

where the above transformations are defined for the computational basis.

The measurements follow Born's rule and the state of the qubit is updated according to the collapse of the wave function. Positive Operator Value Measures (POVM), which were introduced for open quantum systems, are also available for single qubits. In practice, they are often implemented as projective measurements on a larger Hilbert space.

Instead of going directly to the many-qubit case, we can detail a little the information storage capabilities of a single qubit. Since its beginning, quantum mechanics was expected to fit into the wide framework of thermodynamics. There was no apparent

[3] A slightly more general definition relies on the existence of a subalgebra of observables isomorphic to the one generated by the usual 2-dimensional Pauli matrices (see next paragraph).

reason justifying quantum systems as exceptions to the general rule of thermodynamics. Von Neumann proposed a generalization of the entropy for density matrices ρ,

$$S(\rho) = -\mathrm{Tr}\,\rho \log \rho. \tag{5.4}$$

It was later shown that this natural candidate was the only choice sharing many properties with its classical analog, Shannon's entropy. In essence, von Neumann's entropy is the entropy of the probability distribution $\{p_i\}_i$ of the eigenstates of ρ:

$$S(\rho) = S\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = H(\{p_i\}) \tag{5.5}$$

Since orthogonal states are perfectly distinguishable by a single-shot measurement process, $S(\rho)$ can be seen as the information content of a measurement of ρ in its eigenbasis.

Therefore, a single qubit in the maximally mixed state,

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5.6}$$

has an information content of one bit, thus leading to the fortunate conclusion that an isolated qubit can store exactly one classical bit [Hol73a]. In fact, it is through the correlations that can exist between many qubits that the difference between the classical and the quantum arises, as it can be seen in very simple examples of quantum communication protocols such as teleportation [BBC+93a] or dense-coding.

5.2.2 Many qubits

Processing many qubits is not conceptually different from the single qubit case. As a whole, they simply constitute a bigger quantum system which follows the axioms of quantum mechanics. The only difference lies in the greater complexity of the operations that can be performed. The analysis of this structure has been analyzed in detail to define the notion of quantum complexity.

The state space of n qubits taken together is the n-fold tensor product of the single qubit Hilbert space. The computational basis is the canonical basis of the tensor product space, $\{\bigotimes_{i=1}^{n} |\epsilon_i\rangle\}_{\epsilon_i=0,1}$. The evolutions permitted by quantum mechanics for a closed system of qubits are transformations from $SU(2^n)$. We have seen earlier that, for open systems, completely positive trace preserving maps are possible evolutions. In the context of quantum information processing, these maps are usually implemented as unitary transformations on a larger Hilbert space: the qubits of interest are complemented with ancillary qubits initially in the fiducial state $|0\rangle$; a unitary transformation is performed on the whole; finally, the ancillary qubits are traced out to recover only the qubits of interest.

Similarly, each measurement is reduced to a projective measurement in the computational basis. This reduction is possible because any other measurement can be obtained as a preliminary evolution which maps the to-be-measured observable onto the computational basis and the above projective measurement.

5.2.3 Circuit model

For classical and quantum processing, algorithms are simple recipes for solving a given problem. The complexity of an algorithm can be simply considered as an arbitrary accounting rule [KSV02a]. It counts the number of elementary or of costly steps required to arrive at the solution of a problem. Here, the definition of elementary step is the key concept which is defined by the computational model.

For instance, classical complexity in the classical framework can be determined by counting the access to memory cells or by counting the number of floating point operations required to map an instance of the problem to its solution. For quantum complexity, the situation is quite similar: the accounting rule of the standard quantum computing model — or circuit model — consists in counting the number of elementary gates in the quantum algorithm.

More precisely, each qubit start in a definite state of the computational basis. This state constitutes the — classical — input of the algorithm. The algorithm itself is a succession of elementary unitary operations — i.e. acting on a small number of qubits at a time — leading, after a measurement in the computational basis, to the result with high probability — typically greater than $\frac{2}{3}$. A general result about unitary evolutions states that they can be approximated up to arbitrary accuracy by a finite product of gates drawn from a finite universal set (This result is due to Solovay and Kitaev but has never been published by the authors. However, it can be found in [NC00a]). Thus, within this context, the complexity of a quantum algorithm is naturally defined as the number of universal gates that need to be applied before the final measurement. A possible set of universal gates is the Hadamard gate, the phase rotation of $\pi/8$ and the `c-not` gate:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \; R = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/8} \end{pmatrix}, \; \mathtt{c-not} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{5.7}$$

where the action is defined on the computational basis. Other sets exists such as the Toffoli gate and the Hadamard transform. Graphically, each qubit is represented by a horizontal line, and gates as boxes overlapping the qubits it acts on. The circuit is then run from left to right, indicating in which order to apply the different elementary operations.

The above decomposition as universal gates, although very useful for theoretical purposes, is often replaced by a decomposition into slightly more complex gates — a finite set which is not minimal but still universal — in order to simplify the circuit description or in order to take advantage of the specificities of the physical device that implements them. For example, the control-Pauli operations are quite often seen on encoding circuits for error correction (see Chap. 7).

The model that we have presented here is certainly the most widely used in quantum computation. However, other models, such as the one-way quantum computer [Bri01a],

exist, and do have a priori another appropriate definition of complexity. In the rest of this manuscript, and especially for the description of the convolutional quantum convolutional codes, we will only use the circuit model of quantum computing.

5.3 Alorithms and communcation tasks

When one is asked to tell what is the most important achievement in quantum computing the obvious answer is Shor's algorithm [Sho94a]. It shows how a quantum computer can be used to factor an n-digit number efficiently — with only a polynomial circuit complexity. The work of P. W. Shor is based on the possibility of efficient implementation of another primitive, the quantum Fourier transform (see [NC00a] for a clear presentation of the quantum circuit). Indeed, the accomplishment of P. W. Shor was the result of a long research investigating the possibilities of quantum information processing. The Deutsch-Jozsa algorithm [DJ92a] was certainly one of the important step to this result together with Simon's period finding algorithm [Sim94a, Sim97a]. The importance of Shor's algorithm itself is due to the supposed difficulty of factoring large integers on a classical computer. This difficulty is used to guarantee the security of one of the mostly used asymmetric cryptographic protocol, RSA.

The discovery of this algorithm is certainly one of the main reasons for the investment in both time and money for building a quantum computer. In parallel, it also motivated a more systematic research to know which algorithms could be implemented in a more efficient way on a quantum computer. One other result has been obtained for the unstructured database search by L. K. Grover [Gro96a]. Here, it does not result in an exponential speed-up but in a more modest quadratic speed-up. Contrarily to the previous case, it can be proved that it is better by a square-root than any classical algorithm. It thus opens a firm gap between quantum and classical information processing. This type of speed-up is nowadays thought to be possible for many other problems. In fact it seems to result from the manipulation of probability amplitudes rather than of probabilities.

Naturally, since their discovery these two specific examples have been widely studied and improved. Shor's algorithm was a particular case of a more general framework: the Abelian hidden subgroup problem. Grover's search has also been extended to different kind of databases, not only unstructured ones. But all these require to have a fully functional quantum computer able to perform long computations without errors and which are able to use many qubits.

Following the initial ideas that brought interest onto quantum information processing several results have been obtained about the simulation of quantum systems by quantum computers, e.g. [GS01a, BCMS01a]. The evolution of a large number of interacting systems — an a priori difficult task for classical computers — is tractable on a quantum computer, at least for a wide class of Hamiltonians. Even more interesting is the fact that some of these simulations are actually far less resources demanding than solving general mathematical problems. Therefore, this kind of algorithms could be run in a

quite near future on primitive quantum information processors. Experimental results tend to confirm this possibility.

Quantum communication, which also includes quantum cryptography, is another very active area where our currently limited abilities to manipulate the quantum are nonetheless very useful. The achievements in this domain also set a gap between quantum and classical information processing: some tasks are impossible to perform without a help from the quantum world.

One of the first application is quantum key distribution discovered by C. H. Bennett and G. Brassard in 1984 [BB84a]. Using this protocol, two parties could after an authentication phase generate a shared random key with unconditional security [SP00a]. A closer look at the protocols in fact indicates that the security of this scheme is a consequence of the restricted abilities offered by quantum mechanics to eavesdrop the channel — it is impossible to measure a qubit in an unknown state without perturbing it [Fuc98a].

Many other primitives useful for cryptography have been under scrutiny leading to quantum protocols for secret sharing or coin flipping. Quantum data hiding and quantum fingerprinting are other possible examples.

More astonishingly some protocols prove the superiority of quantum information over its classical counterpart (see for example [Raz99a, BCT99a, Lo00a]): the possibility to teleport an unknown quantum state to another location with the use of a pre-established entangled pair and a classical communication channel is one of them. But more generally, quantum communication complexity — equivalent to counting the number of qubits needed to be sent to perform a task between two parties — has also proved some gaps between quantum and classical versions of the same problem. For a review of these algorithms, the reader in invited to refer to "Quantum Computation and Quantum Information" by Nielsen and Chuang [NC00a].

6 Quantum error correction

6.1 Decoherence and control

THE PREVIOUS CHAPTER exemplified how quantum information processing changed computer science. Quantum computing has triggered important theoretical advances such as solving difficult problems but also helped proving results in classical complexity. However, the road leading to functioning devices is paved with many serious difficulties. Namely, quantum computers need to manipulate entangled states involving many qubits in order to outperform their classical counterparts [JL02a]. These states are particularly sensitive to the presence of noise or interaction with their environment. The challenge which consists in avoiding the decoherence process—which turns these computers into at most classical devices—seems overwhelming.

To be more precise, entanglement can be destroyed by an unwanted evolution acting on a single qubit of the entangled state. Unfortunately, these evolutions are almost inevitable when these rather small systems are manipulated by macroscopic objects—i.e. the classical device which contains the quantum registers. Hence, quantum information is very often subject to errors in the course of a computation. In fact, present day technologies allow at best about 100 elementary gates to be performed before an error occurs. This figure can be compared to the 10^{-12} error rate of a CD-writer... As a consequence, error correction strategies to protect the quantum states must be applied at each step of the algorithms.

One obvious but naive solution to the problems of errors and noise induced by the environment would be to isolate the quantum registers. The evolution of the qubits would be closer than the ideal unitary evolutions. Unwanted interactions will certainly be suppressed, but this option would also have some disadvantages: the rate at which designed operations can be performed will decrease. In fact, algorithms that run on quantum computers are made of classical data—e.g. a pulse sequence in an NMR spectrometer—later used by other classical devices to manipulate the qubits. A better insulation of the quantum registers would therefore reduce the coupling with the classical parts of the quantum computer. Hence, even though the decoherence time would increase, the number of gates that can be realized before complete loss of information is not expected to be significantly improved by this strategy. Another element against it is the necessity to perform measurements. They extract some information and amplify it to the macroscopic level. This again requires the presence of classical devices close to the quantum register and might cause unwanted decoherence.

As we can see, decoherence and control of the quantum registers are two contradicting objectives that need to be met before real quantum computers can be build. One path to this goal was the creation of a theory of quantum error correcting codes [Sho95a, Got97a]. It allowed the study of the requirements on the codes and on the physical device in order to effectively reduce the noise and decoherence processes during a computation. This specific issue is the focus of fault tolerant proposals for quantum computers [Sho96a, Got97a, Pre98a]—a combination of error correction schemes that reduce the noise even though it requires a large overhead in the number of gates and qubits to perform the algorithm.

In this chapter we will describe some of the results obtained in the past years. We will emphasize the similarities with classical error correcting codes and try to build some intuition about some of the problems encountered in the quantum domain. Finally, note that other noise protection schemes exist. The theory of noiseless subsystems [ZR97b, ZR97c, VKL00a] is one of them and has some advantages over error correcting codes for some specific applications.

6.2 Classical error correction

6.2.1 Error model

An obvious, but nevertheless essential, point to consider before studying error correction strategies is the error model. It is a probabilistic description of the errors that can affect the bits during the transmission. In the classical domain one of the most widely used is the binary symmetric memoryless channel. In this framework, errors affect each bit independently with a fixed probability. Hence, it is defined by a real number p such that with probability $1-p$ the bit is transmitted without error, while with probability p its value is flipped during the transmission.

6.2.2 Linear codes

The goal of coding theory is to analyze ways for improving the stability of information given a particular error model. Even though classical coding theory is a very developed field, we will only give a basic overview of some specific codes whose formalism is similar to the one developed in the quantum setting. These are linear codes [JZ99a].

Definition 6.2.1 ((n,k) binary code). *A (n,k) binary code is a set of 2^k n-dimensional binary vectors. Each vector in the set is called a codeword.*

Intuitively, such structure can protect information because each possible k-bit string corresponds to a unique codeword, and because this codeword contains some redundancy: it uses n bits instead of the original k bits of the to-be-encoded string. When the codeword is sent instead of the original string, noise might affect some of the bits. Fortunately, the presence of redundant information might prevent the erasure of all the copies of the information that needs to be transmitted. The error correction procedure

uses redundancy to recover the original codeword and therefore the initial bit string. The goal of coding theory is to describe ways to generate this redundancy and to exploit it to later correct the maximum number of errors for an overhead of $n-k$ bits for each k-bit string of to-be-transmitted information.

Additional structure is often added to the set of codewords to ease the encoding, decoding and error correction as well as the analysis of the code itself. A linear structure is such a possible additional constraint on the codewords. Namely, if y and y' are two codewords, it requires that $y+y'$ obtained by adding modulo two each components must also be a codeword. In other terms, there exists an $n \times k$ matrix G such that any codeword is obtained via Gx where x is a k-dimensional binary vector. Most of the time this also specifies an easy encoding procedure which consists in associating x to the codeword Gx. The matrix G is called the generating matrix of the code. The linearity of the code also allows to describe the set of codeword by a very small number of parameters—polynomial in n and k—whereas one would expect it to be exponential in k for a random code.

The error correction procedure for linear codes is based on the calculation of the syndrome. One defines a matrix H—called a parity check matrix—such that $HG=0$ where H is $(n-k)\times n$ and of rank $n-k$. This equivalently defines the code since y is a codeword if and only if $Hy=0$. In the case of a received signal z, each component of the vector Hz—the whole vector is called a syndrome—brings information about the possible error that happened during the transmission. The error correction algorithm uses this information together with the error model to infer what has been really sent.

6.2.3 Example

The simplest example of error correcting code is the 3-bit repetition code. It is obtained by introducing redundancy in the form of copying three times the to-be-protected information. When 0 has to be transmitted $(0,0,0)$ is sent over the channel, while $(1,1,1)$ is associated to 1. The generating matrix is therefore equal to

$$G=\begin{pmatrix}1\\1\\1\end{pmatrix}. \tag{6.1}$$

For the binary symmetric channel[1], when we receive $(0,1,0)$ we conclude that the most likely sent codeword is $(0,0,0)$, which corresponds to the to-be-protected bit 0. To arrive at this conclusion, two methods are available: either through a majority vote, which also decodes the protected information, or through the parity check matrix H. In this specific example a possible choice for H is:

$$H=\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix}. \tag{6.2}$$

[1]Further suppose that the error probability p is less than $1/2$.

Calculating the syndrome, here $(1,1)$, indicates that the received signal does not correspond to a valid codeword. The error correction procedure should find the binary string which is the closest to the received signal and which has an all-zero syndrome. This notion of distance between binary strings is given by the error model, and in our example corresponds to the error model is the Hamming distance—the number of bits that differ in the two strings.

Of course both result coincide. However, in the latter procedure if we suppose that instead of transmitting 0 the emitter wanted to transmit 1 and the same error occurred on the emitted codeword—a bit flip on the second bit—, then the received signal would have been $(1,0,1)$. It is important to note that this vector corresponds to the same syndrome $(1,1)$. This particular point of syndrome based error correction procedures will be used later in the quantum case: the value of the syndrome is completely independent from the information that has been sent over the channel. It only characterizes the error that happened during the transmission.

6.3 Quantum error correction

6.3.1 Error model

Like for classical codes an error model is required for a proper discussion of the error correction capabilities of a code. In this work we will only mention the depolarizing channel. It is the generalization to the quantum case of the binary symmetric memoryless channel. It acts on each qubit independently and applies the identity with probability $(1-p)$, a bit flip X with probability $p/3$, a phase flip Z with probability $p/3$ and a combination of those two, Y, with probability $p/3$. For a qubit state ρ the corresponding trace preserving completely positive map is given by:

$$\rho \mapsto (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z). \tag{6.3}$$

Hence, for a n-qubit string the possible errors are all the elements of the Pauli group $\mathcal{G}_n = \mathrm{sp}\{I,X,Y,Z\}^{\otimes n}$. The aim of the error correction procedure will be to find the most likely element of $\mathcal{G}_n$ which could have affected the qubit state during the transmission.

6.3.2 Example

Instead of directly describing the theory of quantum codes, it is suitable to have an idea of the difficulties that arise in the quantum framework. We abandon temporarily the error model just introduced and turn into an even simpler case. The channel we will use is the quantum binary symmetric channel for the computational basis: with probability $1-p$ nothing happens, while with probability p a bit flip X is applied to the qubit.

We can try to adapt the 3-bit error correcting code to the quantum domain. To protect $|0\rangle$ and $|1\rangle$ we need two three-qubit states, $|0\rangle|0\rangle|0\rangle$ and $|1\rangle|1\rangle|1\rangle$. Hence, when a generic state $\alpha|0\rangle + \beta|1\rangle$ has to be protected, $\alpha|0\rangle|0\rangle|0\rangle + \beta|1\rangle|1\rangle|1\rangle$ is sent instead. With a bit flip on the second qubit, the received state becomes $\alpha|0\rangle|1\rangle|0\rangle + \beta|1\rangle|0\rangle|1\rangle$.

The receiver must then apply an error correction step to recover the initial information. At this stage only one strategy is possible. If the receiver were to recover the information by a majority vote, it would collapse the state vector onto $|0\rangle|1\rangle|0\rangle$ or $|1\rangle|0\rangle|1\rangle$ with probability $|\alpha|^2$ and $|\beta|^2$ respectively. It will thus destroy irremediably the quantum superposition. On the other hand, if he applies a syndrome extraction step with two ancillary qubits, the overall state would be

$$(\alpha|0\rangle|1\rangle|0\rangle + \beta|1\rangle|0\rangle|1\rangle)|1\rangle|1\rangle, \tag{6.4}$$

which does not induce a collapse even if the last two qubits are measured. Then, the same error estimation procedure as in the classical case can be applied and would successfully correct the error. The initial quantum information is preserved.

Syndrome extraction procedures are mandatory in the quantum domain as it is the only way to preserve essential superpositions. This is one of the fundamental characteristics of quantum codes. One might further ask whether it is always possible to find a circuit transforming the to-be-protected quantum information into codewords, and how the syndrome can actually be measured. The answers to these questions are brought by the stabilizer formalism.

6.3.3 Stabilizer formalism

Definition 6.3.1 ((n,k) binary quantum code). *A (n,k) binary quantum code is a 2^k-dimensional subspace of the Hilbert space of n qubits.*

Defining a code is thus equivalent to define a particular subspace in a larger Hilbert space. In the stabilizer formalism (see [Got97a] and references therein), it is done by considering an Abelian subgroup S of the Pauli group $\mathcal{G}_n$ over n qubits, $\mathcal{G}_n = \mathrm{sp}\{I, X, Y, Z\}^{\otimes n}$. The code subspace is by definition the set of all vectors $|\psi\rangle$ such that $|\psi\rangle = S|\psi\rangle$. It is possible to show that to have a (n,k) code the number of independent generators of S must be equal to $n-k$. Note that this definition of the code is somewhat very similar to the one involving the parity check matrix H for a classical linear code.

In fact, the analogy between S and H can be pushed further way. A set of independent generators $\{M_i\}_i$ of S is equivalent to the rows of H in that they allow to calculate the syndrome for the corresponding quantum code. Because operators in $\mathcal{G}_n$ have only two eigenvalues $+1$ and -1, and because these operators either commute—$ABB^\dagger A^\dagger = I$—or anti-commute—$ABB^\dagger A^\dagger = -I$—, if an error operator $E \in \mathcal{G}_n$ is applied to the transmitted state vector, either at least one generator anti-commutes with E or all the elements of S commute with E.

To know whether E commutes with all the M_j is a relatively easy task, which can be conducted without knowing E itself. The only needed ingredient is the received signal $|\psi\rangle$: if M_i anti-commutes with E then we have $M_i|\psi\rangle = -|\psi\rangle$. This phase difference can be measured by a quantum circuit. In this case, it is possible to gain information

about the error without ever gaining information about the state itself, without inducing a collapse of the state vector.

For the example given earlier, the stabilizer group is generated by ZZI and IZZ which are equivalent to the syndrome $(1,1,0)$ and $(0,1,1)$ for the classical code.

As in the classical case, the error estimation procedure infers the error that affected the transmitted state according to a particular error model. Correction can take place followed by the decoding operation to recover the initial k-qubit state.

6.3.4 Specific issues

The encoding of a classical code is a very easy task obtained by multiplying the to-be-protected k-bit string by the generating matrix of the code, G. The situation is more complicated in the quantum case. The no-cloning theorem forbids to blindly copy the quantum state. This is a big restriction for introducing redundancy. In fact, the procedure for encoding is rather complicated. From the description of the stabilizer group by its generators, one can obtain a matrix which can be put in a standard form to produce an encoding network for the code. With this circuit, the decoding procedure is simply implemented by running the encoding one backward: in the stabilizer formalism, the gates used for encoding are their own inverses.

We voluntarily do not mention fault-tolerance issues nor the criteria to know whether an error is detectable or correctable. These concepts will not be used in the next chapters as we focus only on defining a particular class of codes. We do not analyze their intrinsic performance in terms of error correction as new tools need to be developed for these specific issues. However, for understanding quantum error correction as a whole, the interested reader should refer to Gottesman's PhD thesis [Got97a].

7 Quantum convolutional codes

7.1 Introduction

QUANTUM INFORMATION SCIENCE has been developed in the past two decades as a way to process information more efficiently than with classical means. It lead to great theoretical advances and to impressive experimental realizations (see [Pre98a, NC00a] for a review). The main results motivating the interest for quantum computation concern integer factorization [Sho94a] and unsorted database search [Gro97a]. Both contribute to the widely accepted idea that quantum computers are intrinsically more powerful than their classical analogs, and justify the ever increasing interest for this new model of computation.

In parallel to these developments, the difficulty of building quantum information processing devices has been throughly pointed out: the quantum world is extremely sensitive to interactions with its surrounding environment [Zur91a, Zur02a, Zur03a]. This process, called decoherence, is responsible for the instability of the fragile quantum superpositions necessary to obtain a speedup over classical computation [JL02a]. In absence of any control over the decoherence process, these quantum devices would be turned into—at best—classical computers. Fortunately, the discovery of quantum error correction schemes [Sho95a], together with their fault-tolerant implementation [Got98a] cleared the future of quantum computation: quantum codes protect from unwanted evolutions and noise, whereas their fault-tolerant implementation guarantees that, below a certain error rate, quantum information processing can be done without loss of coherence [DS96a, Zal96a, Got98a, AB97a]. However, fault-tolerant quantum computation usually requires a large overhead in costly quantum resources.

On the other hand, quantum communication protocols—e.g. quantum key distribution—achieve the production of large numbers of qubits often represented by some degrees of freedom of light modes. For most protocols, the manipulation of quantum bits is very limited and errors occur mainly during the transmission—loss of photons, noise, etc. In the perspective of quantum communication, we develop a theory of quantum convolutional error correcting codes. These codes are largely inspired by their classical analogs [Lee97a, JZ99a] and share many of their properties: efficient encoding and decoding circuits and an efficient maximum likelihood error estimation procedure for any memoryless channel.

Even more crucially, and as it is the case in the classical context, these codes can deal with infinitely long streams of "to-be-protected" information without introducing unacceptable delays in the transmission, and yet being inequivalent to block codes. This

is one of the motivations for the introduction of such error protection schemes. Indeed, it is expected that convolutional codes transform memoryless channels into channels with correlated errors once decoding has been performed. Thus, by concatenating a convolutional code with a suitably chosen block code, it would be possible to take advantage of this specific biais, and to improve the general performance of concatenated schemes. This study has not been performed yet, but is subject to future research.

This chapter is organized as follows: sec. 7.2 describes the structure of quantum convolutional codes and introduces an appropriate formalism; sec. 7.3 provides an encoding circuit for this class of codes; sec. 7.4 studies error propagation properties, and sec. 7.5 details the efficient maximum likelihood error estimation algorithm. Throughout the text, abstract concepts are readily applied to a previously introduced example of quantum convolutional code [OT03a].

7.2 Structure of quantum convolutional codes

All error protection strategies share many common ingredients. They are specifically designed to enhance communication for a given error model. Once this error model is known, they define the structure in which quantum information will be stored and, as second step, explain how information can be manipulated within this structure. Quantum error correcting codes impose the information to be stored in a subspace of the total Hilbert space of the physical qubits. This subspace, $\mathcal{C}$, is called the code subspace. $\mathcal{C}$ is usually further decomposed into—e.g. single qubit—subspaces for which elementary operations are then provided.

However, to arrive at a practical definition of a quantum error correction scheme, it is usually necessary to further restrict the possibilities offered by the above general program. One such restriction leads to stabilizer codes. Those are often compared to classical linear codes: they are defined by a set of linear equations—called syndromes—which allow an efficient description of $\mathcal{C}$ together with a great flexibility in their design. To facilitate the introduction of quantum convolutional codes, we will use abundantly the stabilizer formalism, even though convolutional codes can be generalized to a wider framework[1].

More precisely, the code subspace $\mathcal{C}$ of any stabilizer code is defined as the largest subspace stabilized by an Abelian group S acting on the N physical qubits of the code. In practice, S is a subgroup of the multiplicative Pauli group $G_N = \mathrm{sp}\{I, X, Y, Z\}^{\otimes N}$, where I, X, Y, Z are the well known Pauli matrices. The description of $\mathcal{C}$ is further simplified by the introduction of a set of independent generators $\{M_i\}$ of S. This leads to the definition of $\mathcal{C}$ in terms of syndrome equations:

$$\forall i,\ |\psi\rangle = M_i\,|\psi\rangle \ \Leftrightarrow\ |\psi\rangle \in \mathcal{C}. \tag{7.1}$$

The code vectors are common eigenstates with eigenvalue $+1$ of all the operators M_i.

[1] In particular, our main theorem concerning error propagation in sec. 7.4 does not rely on the stabilizer formalism.

7.2.1 Definition

The particularity of convolutional codes is to impose a specific form to the generators of the stabilizer group such that on-line encoding, decoding and error correction become possible even in the presence of an infinitely long to-be-protected stream of information.

However, convolutional codes do not consider groups of qubits independently of each other: the encoding operation cannot be decomposed into a tensor product of encoding operations acting on a small number of qubits. By contrast, an (N, K)-block code can protect such stream only by cutting it into successive K-qubit blocks. As a result, the code subspace defined by these independent applications can be decomposed as a tensor product of the N-qubit subspaces of each output block. Furthermore, increasing the parameter K is usually not an option as it requires, in most cases, a quadratic overhead in the complexity of the encoding circuit [Got97a] and, more dramatically, an exponentially growing complexity of the error estimation algorithm[2].

Quantum convolutional codes are especially designed to offer an alternative to small block codes in counteracting the effect of decoherence and noise over long-distance communications while using a limited overhead of costly quantum resources.

Definition 7.2.1 (**(n, k, m) convolutional code**). *The stabilizer group S of a convolutional code with parameters (n, k, m) is given by:*

$$S = \mathrm{sp}\{M_{j,i} = I^{\otimes j \times n} \otimes M_{0,i},\ 1 \leq i \leq n-k,\ 0 \leq j\}, \tag{7.2}$$

where $M_{0,i} \in G_{n+m}$. Above $M_{j,i}$'s are required to be independent and to commute with each other.

Remark 7.2.1. As expected, the length of the code (i.e. the number N of physical qubits of the code) as well as the number of logical qubits are left unspecified. In fact, the maximum value of the integer j controls this length implicitly. However, and contrarily to block codes, this maximum value does not need to be known in advance for encoding and decoding qubits. Instead, it will be fixed a posteriori when the transmission ends. This specific issue will be addressed in sec. 7.3. Hence, in most situations the length of the code can simply be set to infinity. The only associated restriction is to consider operators whose support[3] has size of order 1. This also explains why in Eq. (7.2) the $M_{j,i}$'s seem to have different length: in the rest of the article we simply assume that the operators are "padded" by identities on the right-most physical qubits to adjust them to the appropriate length. ∎

With this remark in mind, the structure of the stabilizer group generators can be

[2]This holds for random codes without particular structure—not belonging to a restricted class— and with constant rate as K increases.

[3]In this article the definition of support of an element A of the Pauli group is — rather unconventionally — the smallest block of consecutive qubits on which A acts non-trivially.

summarized easily with the help of a semi-infinite matrix M:

$$M = \left(\begin{array}{ccccc} \boxed{\begin{array}{c} M_{0,1} \\ \vdots \\ M_{0,n-k} \end{array}} & & & & \\ & \boxed{\begin{array}{c} M_{1,1} \\ \vdots \\ M_{1,n-k} \end{array}} & & & \\ & & \ddots & & \end{array} \right) \tag{7.3}$$

Each line of the matrix represents one of the $M_{j,i}$ and each column a different qubit. A given entry in M is thus the Pauli matrix for the corresponding qubit and generator. The rectangles represent graphically which qubits are potentially affected by the action of the generators. The form of Eq. (7.3) visually emphasizes the structure of convolutional codes:

- M has a block-band structure;
- the overlap of m qubits between two neighboring sets of generators forces to consider the code subspace as a whole.

By contrast, for a block code used repeatedly to protect an infinitely long stream of qubits, the above parameter m would be equal to 0.

Remark 7.2.2. In addition to the above generators, and in order to properly account for the finiteness of real-life communications, a few other generators will be added to the matrix M. This will however not interfere with the rest of this section. ∎

Example 7.2.1 (5-qubit convolutional code). The following generators satisfy the conditions for defining a quantum convolutional code with parameters $n = 5$, $k = 1$ and $m = 2$:

$$\begin{array}{rcl} M_{0,1} & = & Z\ X\ X\ Z\ I\ I\ I \ldots, \\ M_{0,2} & = & I\ Z\ X\ X\ Z\ I\ I \ldots, \\ M_{0,3} & = & I\ I\ Z\ X\ X\ Z\ I \ldots, \\ M_{0,4} & = & I\ I\ I\ Z\ X\ X\ Z \ldots, \\ M_{j,i} & = & I^{\otimes 5j} \otimes M_{0,i},\ 0 < j. \end{array} \tag{7.4}$$

∎

7.2.2 Polynomial representation

Although, it is in principle possible to carry out a complete the analysis of the code with the matrix M only, we will introduce a polynomial formalism which greatly simplifies this task. Such formalism is the exact translation of the polynomial formalism for classical

convolutional codes. Its advantage is to capture in a convenient and efficient way the fact that the generators in M are n-qubit shifted versions of the $M_{0,i}$'s.

More precisely, for a (n, k, m) convolutional code, we define the delay operator D acting on any element A of the Pauli group of the physical qubits with bounded support[3] by:

$$D[A] = I^{\otimes n} \otimes A, \tag{7.5}$$

with the same "padding rule" as before. Naturally, one can consider powers of D as repeated applications of the delay operator. For instance, the generators of the code can now be written as:

$$M_{j,i} = D^j[M_{0,i}],\ 0 \leq j,\ 1 \leq i \leq n-k. \tag{7.6}$$

Therefore, and to further continue with simplifications, it is obviously not necessary to keep more than the first $n-k$ lines of the matrix M defined in Eq. (7.3). All the omitted ones can be easily recovered by applying D the appropriate number of time.

In addition to applying a single D^j to an element of the Pauli group, it is, under certain conditions, possible to consider more complex operations—for instance, these will be necessary for deriving the encoding circuit. Namely, consider A, an element of the Pauli group with bounded support, such that A and $D^j[A]$ commute for any value of j. Then, the full polynomial ring $GF_2[D]$ can act on A. For $P(D) = \sum_j \alpha_j D^j$, the action of $P(D)$ on A is naturally defined as:

$$P(D)[A] = \prod_j \alpha_j D^j[A]. \tag{7.7}$$

Above, the commutation relation is crucial: the sum operation in $GF_2[D]$ is commutative and must therefore be translated into another commutative operation—here the product—on the multiplicative group spanned by $\{D^j[A]\}_j$.

Finally, we will sometimes use a short hand in our notation and, instead of restricting ourselves only to polynomials in D, consider formal Laurent series acting on A. In such case, we do not really need to define the action of negative powers of D, but we impose that, at the end of the calculation—possibly concerning several operators—, all the negative powers of D are removed by globally[4] applying the smallest possible positive power of D. For instance, if we end with

$$L(D)[A] = \left(\sum_{j=-p}^{q} \alpha_j D^j \right)[A], \tag{7.8}$$

it will be turned into

$$\begin{aligned} P(D)[A] &= \left(D^p \sum_{j=-p}^{q} \alpha_j D^j \right)[A] \\ &= \left(\sum_{j=-p}^{q} \alpha_j D^{j+p} \right)[A]. \end{aligned} \tag{7.9}$$

[4]This means on all the operators involved in the calculation.

In practice, the representation of the code generators as a matrix M with entries I, X, Y, Z is often replaced by the one of [CRSS97a]. In this representation, the first $n-k$ generators of an (n,k,m) convolutional code would be written as a pair of $(n+m)\times(n-k)$ binary matrices arranged side by side[5]. Each line corresponds to a generator and each column to a qubit. A 1 for the left matrix indicates the presence of an X or Y and, similarly, a 1 for the right matrix indicates the presence of a Y or Z. Within this framework, it is easy to realize that the polynomial formalism can be fruitfully extended to lead an even more compact notation for the generators of the stabilizer group.

First, recall that the addition of two pairs of binary vectors simply results in the multiplication of the corresponding generators provided that these commute. For instance, suppose A and B are two elements of G_n, and $(A_X|A_Z)$, $(B_X|B_Z)$ their respective representations as pair of binary vectors. In such case, the operator $A\otimes B$ is represented by $(A_X : B_X|A_Z : B_Z)$ where ":" indicates the concatenation of the vectors. With the polynomial formalism, we also have $A\otimes B = A\times D[B]$, which leads to[6] $(A_X : B_X|A_Z : B_Z) = (A_X|A_Z) + D[(B_X|B_Z)]$. Here, the commutation of A and $D[B]$ is trivially verified since their supports do not intersect.

This last equality suggests the following modification of the representation. A generic element P of the Pauli group of the physical qubits with bounded support is represented by a pair of length n vectors with coefficients in $GF_2[D]$ such that, by definition,

$$\begin{aligned}(P_X|P_Z) &= (P_X^{(0)} : P_X^{(1)} : P_X^{(2)} : \ldots | P_Z^{(0)} : P_Z^{(1)} : \ldots) \\ &= (P_X^{(0)} + D\times P_X^{(1)} + D^2\times P_X^{(2)} + \ldots | P_Z^{(0)} + D\times P_Z^{(1)} + \ldots), \quad (7.10)\end{aligned}$$

where the $P_X^{(j)}$'s and $P_Z^{(j)}$'s are length n binary vectors.

Example 7.2.2 (5-qubit convolutional code). These new concepts can be used to give the representation of the generators found in Ex. 7.2.1 as a pair of polynomial matrices.

$$M = \left(\begin{array}{ccccc|ccccc} 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & D & 0 & 1 & 0 & 0 \\ D & 0 & 0 & 0 & 1 & 0 & D & 0 & 1 & 0 \end{array}\right). \quad (7.11)$$

The correspondence between this notation and the usual notation as pairs of binary matrices is given by Eq. 7.10 The above matrix corresponds to the first 4 generators of the code. The others are derived by multiplying both matrices by successive powers of D, which induces shifts of 5 qubits. ∎

[5]This representation as a pair of binary vectors or matrices is not restricted to elements of the stabilizer group, and can indeed be used for any element of G_{n+m}.

[6]Here again we apply the implicit "padding rule" to adjust the length of the vectors.

7.2.3 Generalized commutator

In the context of block codes, the main reason justifying the introduction of the representation of the elements of the Pauli group as pairs of binary vectors [CRSS97a, Got97a] is the existence of an easy way to compute the group commutator. We shall see below that the same kind of advantage holds for the representation as pair of polynomial vectors.

First, consider two elements $A = (A_X|A_Z)$ and $B = (B_X|B_Z)$ of G_N. It is then easy to check on their representation as pair of binary vectors that,

$$AB = BA \Leftrightarrow A_X B_Z + A_Z B_X = 0, \tag{7.12}$$

where we use the standard inner product of two vectors of length N and addition modulo 2.

Now suppose that P and Q are two elements of the Pauli group of the physical qubits of an (n, k, m) convolutional code, and $(P_X(D)|P_Z(D))$, $(Q_X(D)|Q_Z(D))$ their representation as pair of polynomial vectors. Using the above method, one can conclude that the commutation of P and Q is simply expressed by:

$$PQ = QP \Leftrightarrow \sum_l P_X^{(l)} Q_Z^{(l)} + P_Z^{(l)} Q_X^{(l)} = 0, \tag{7.13}$$

where $P_X(D) = \sum_j P_X^{(j)} D^j$ with $P_X^{(j)}$ a binary vector of length n and similarly for $P_Z(D)$, $Q_X(D)$, $Q_Z(D)$. This also leads to,

$$D^r[P]D^s[Q] = D^s[Q]D^r[P] \Leftrightarrow \sum_l P_X^{(l+s)} Q_Z^{(l+r)} + P_Z^{(l+s)} Q_X^{(l+r)} = 0. \tag{7.14}$$

The last equation is particularly interesting since its right hand side is the coefficient of D^{s-r} in $P_X(D)Q_Z(1/D) + P_Z(D)Q(1/D)$. Therefore, one can readily conclude that the representation as pair of polynomial vectors allows an easy computation of the "generalized commutation relation"—i.e. the commutation of any n-qubit shifted version of P with any n-qubit shifted version of Q—:

$$\begin{gathered} \forall r, s,\ D^r[P]D^s[Q] = D^s[Q]D^r[Q] \\ \Leftrightarrow \\ P_X(D)Q_Z(1/D) + P_Z(D)Q_X(1/D) = 0. \end{gathered} \tag{7.15}$$

We will see below that this property of the polynomial representation is crucial as it allows the derivation of almost all the encoded Pauli operators by considering only the first $n - k$ generators $M_{0,i}$'s of the stabilizer group.

7.2.4 Encoded Pauli operators

The encoded Pauli operators for a quantum error correcting code are some operators of the Pauli group of the physical qubits which allow the manipulation of the information

without requiring any decoding. More precisely, these are operators that leave the code subspace $\mathcal{C}$ globally invariant, but which have a non-trivial action on it. Indeed, it is possible to require such operators to reproduce exactly the commutation relations of the Pauli group for the encoded qubits. This is mathematically expressed by [Got97a][7]:

$$\begin{aligned} \overline{X}_i,\ \overline{Z}_i &\in N(S)/S, && (7.16)\\ [\overline{X}_i, \overline{X}_j] &= 0, && (7.17)\\ [\overline{Z}_i, \overline{Z}_j] &= 0, && (7.18)\\ [\overline{X}_i, \overline{Z}_j] &= 0,\ i \neq j, && (7.19)\\ \{\overline{X}_i, \overline{Z}_i\} &= 0, && (7.20) \end{aligned}$$

where the index i in $\overline{X}_i$ and $\overline{Z}_i$ denotes the i-th logical qubit.

In the rest of this paragraph we exploit Eq. (7.15) to find an algorithmic procedure for deriving the $\overline{X}_i$'s and $\overline{Z}_i$'s. First we define the standard polynomial form of M and, as a second step, we translate Eqs. (7.16–7.20) into a set of equations for polynomial vectors which can be solved easily.

To obtain the standard polynomial form for the generators of the stabilizer group one can perform two Gaussian eliminations[8] on M written in its representation as pair of polynomial matrices over $GF_2[D]$. This can be done by using line additions, column swaps and multiplication of a line by a power of D:

$$M_{\text{std}} = \left(\begin{array}{ccc|ccc} \overbrace{A(D)}^{r} & \overbrace{B(D)}^{n-k-r} & \overbrace{C(D)}^{k} & \overbrace{E(D)}^{r} & \overbrace{F(D)}^{n-k-r} & \overbrace{G(D)}^{k} & \}r \\ 0 & 0 & 0 & J(D) & K(D) & L(D) & \}n-k-r \end{array}\right) \qquad (7.21)$$

where $A(D)$ and $K(D)$ are diagonal matrices with polynomial coefficients, and where r is the rank of the X-part of M.

By definition, $A(D)$ has full rank. In fact, this holds for $K(D)$ as well: if it was not the case, then there would exist a line with zeroes everywhere except for at least one position in the first r columns of the Z-part. Then, the operator corresponding to this line cannot commute in the generalized sense with all the other generators, which would contradict the assumption that the stabilizer group S generated by M_{std} is Abelian.

We now turn to the determination of the encoded Pauli operators. Here, we restrict our search to operators that preserve the convolutional nature of the code: we want to find a finite set of independent operators with bounded support which generate through n-qubit shifts—almost all—the encoded Pauli operators[9]. This can be accomplished by considering a k-line matrix,

$$\overline{X} = (U_1(D), U_2(D), U_3(D) | V_1(D), V_2(D), V_3(D)), \qquad (7.22)$$

[7] In all this article, and following the notation of [Got97a], the encoded Pauli operators are denoted by, e.g. $\overline{X}$ and $\overline{Z}$.

[8] See also [Got97a] for a similar procedure for block codes

[9] For the purpose of introducing the theory of quantum convolutional codes, it is not necessary to consider encoded Pauli operators that do not respect the convolutional structure of the code. However, in more elaborated error correction scheme, this might prove to be useful.

representing the encoded $\overline{X}$ operators—the rest of the discussion shows that such encoded Pauli operators exist. Since these operators can be multiplied by any element of the stabilizer group, $U_1(D)$ and $V_2(D)$ can be set to 0. The generalized commutation with the lines of M imposed by Eq. (7.16) can be simply written:

$$\left(\begin{array}{ccc|ccc} A(D) & B(D) & C(D) & E(D) & F(D) & G(D) \\ 0 & 0 & 0 & J(D) & K(D) & L(D) \end{array}\right) \left(\begin{array}{c} V_1^T(1/D) \\ 0 \\ \hline V_3^T(1/D) \\ 0 \\ U_2^T(1/D) \\ U_3^T(1/D) \end{array}\right) = \\ = \left(\begin{array}{c} 0 \\ 0 \end{array}\right). \tag{7.23}$$

On the other hand, Eq. (7.17) is expressed by

$$U_3(D)V_3^T(1/D) + V_3(D)U_3^T(1/D) = 0, \tag{7.24}$$

which can be trivially satisfied with $V_3(D) = 0$ and $U_3(D) = \Lambda(D) \times I$, where $\Lambda(D)$ is a non-zero polynomial of $GF_2[D]$. This choice guarantees that the operators in $\overline{X}$ together with their n-qubit shifted versions are independent of each other and from the generators of S. In this case, Eq. (7.23) becomes,

$$\left(\begin{array}{c} A(D)V_1(1/D)^T + F(D)U_2(1/D)^T + G(D)U_3(1/D)^T \\ K(D)U_2(1/D)^T + L(D)U_3(1/D)^T \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \end{array}\right). \tag{7.25}$$

Then we can write the encoded $\overline{X}$ operators:

$$\begin{aligned} U_1(D) &= 0 && (7.26)\\ U_2(D) &= L^T(1/D)K^{-1}(1/D)\Lambda(D) && (7.27)\\ U_3(D) &= \Lambda(D) \times I && (7.28)\\ V_1(D) &= \left(U_2(D)F(1/D)^T + \Lambda(D)G(1/D)^T\right) A^{-1}(1/D) && (7.29)\\ V_2(D) &= 0 && (7.30)\\ V_3(D) &= 0. && (7.31) \end{aligned}$$

One must realize that the encoded Pauli operators $\overline{X}$ are not yet properly defined as the division by polynomials is in general problematic. The reason is that generic polynomial fractions cannot be written as finite formal Laurent series. Thus, the operators that they describe have an unbounded support. In such case, and without further modifications, the formalism introduced earlier imposes transmissions of infinite length. However, when the result of the division can be written with a finite Laurent series, such operation is permitted.

Definition 7.2.2 (Conditioning polynomial). *The non-zero polynomial $\Lambda(D)$ with minimum degree such that the equations (7.26–7.31) only involve finite Laurent series is called the conditoning polynomial of the code.*

As it can be seen easily, the conditional polynomial always exists, and the $\overline{X}$'s operators are well defined. They correspond to operators with a finite support, respecting the convolutional structure of the code.

We now turn to the derivation of some $\overline{Z}$'s by applying the same tools. First note that once the $\overline{X}$'s are fixed, there is a unique set of valid $\overline{Z}$'s. Quite surprisingly, we will also see here that it is not always possible to impose to the $\overline{Z}$'s the convolutional structure—the invariance by n-qubit shifts.

For instance, first define the k-line matrix

$$\overline{Z} = (0, U_2'(D), U_3'(D) | V_1'(D), 0, V_3'(D)). \tag{7.32}$$

Above, the zeroes have been set for the same reason as in the derivation of the $\overline{X}$'s. In addition to satisfying an equation similar to Eq. (7.23), the matrix $\overline{Z}$ must anti-commute in the generalized sense with $\overline{X}$, Eq. (7.20). Equivalently, this can be expressed as $V_3'(D)U_3(1/D)^T = I$, which can be fulfilled if and only if $V_3' = I/\Lambda(D)$. As discussed above, only when $\Lambda(D)$ is a monomial in D does $V_3'(D)$ correspond to a valid polynomial vector (i.e. $1/\Lambda(D)$ is a bounded Laurent series). In this latter case, we obtain $\overline{Z}$:

$$\begin{aligned}
U_1'(D) &= 0 && (7.33)\\
U_2'(D) &= 0 && (7.34)\\
U_3'(D) &= 0 && (7.35)\\
V_1'(D) &= C^T(1/D)A(1/D)^{-1}/\Lambda(1/D) && (7.36)\\
V_2'(D) &= 0 && (7.37)\\
V_3'(D) &= I/\Lambda(1/D). && (7.38)
\end{aligned}$$

Note that for $\Lambda(D)$ to be a monomial, all the $A_{i,i}(D)$'s must be monomials as well, so that Eq. (7.36) is automatically well defined.

Remark 7.2.3. The obvious question raised by this derivation concerns the case where $\Lambda(D)$ is not a monomial. The rigorous answer will be given in sec. 7.4 where it will be shown that if such code were to be used, it would have bad error propagation properties.[10] One can also consider the following hand-waving argument: when $\Lambda(D)$ is not a monomial, and for a finite length communication, the $\overline{Z}$'s have a support with a size of the order of the length of the code. Thus, if one implements an encoded phase flip by applying individual Z's on the physical qubits with finite precision, then for long streams of to-be-protected information this will result in an error with probability close to 1. ∎

[10] Only the $\overline{X}$ operators are used to derive the encoding circuit. Then, if one renounces to manipulate information in its encoded form, the code can be, in principle, successfully used to protect quantum information.

Finally, and to conclude this section on the structure of convolutional codes, we should count how many logical qubits are described by our construction in the case of a finite transmission. To simplify this discussion, we define the integer λ as the highest degree in the polynomial matrices $\overline{X}$ and $\overline{Z}$. This sets an upper bound on the size of the support of any of the $\overline{X}$'s and $\overline{Z}$'s: they extend on at most $\lambda+1$ consecutive n-qubit blocks. Further consider the stabilizer group S generated by the $\{M_{j,i}\}$ for $0 < i \leq n-k$ and $0 \leq j < p$ with $p > \lambda$. In this case, the above derivation leads to at least[11] $k \times (p + \lceil m/n \rceil - \lambda)$ logical qubits while we used $(n-k) \times p$ generators and $(p+\lceil m/n \rceil) \times n$ physical qubits. Therefore, only $\lceil m/n \rceil \times (n-k) + \lambda k$ logical qubits—a number independent of p—do not follow the convolutional structure of the code. These will simply be discarded in the encoding process as this does not change the asymptotic rate of the code. This can be done consistently with the stabilizer formalism by adding their encoded $\overline{Z}$ operators to the generators of S.

One should further note that the presence of these sacrificed qubits is due to an uneven protection against errors at the beginning and at the end of the stream. There, each physical qubit is involved in a smaller number of generators of the stabilizer group than those in the middle of the stream. Therefore, extra logical qubits with poor error protection capabilities are created. To ensure an even protection, it is then natural to simply discard them.

Example 7.2.3 (5-qubit convolutional code). By working out the method given above with the polynomial matrix of Ex. 7.2.2, it is easy to find the standard form of M,

$$M_{\text{std}} = \left(\begin{array}{ccccc|ccccc} D & 0 & 0 & 0 & 1 & 0 & D & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1+D & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & D & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & D & 0 & 1 & 0 & 0 \end{array}\right). \tag{7.39}$$

The $\overline{X}$ operators are obtained from a single 5-dimensional vector, with the conditioning polynomial $\Lambda(D)$ equal to 1:

$$\begin{array}{rcl} \overline{X} & = & (0,0,0,0,1|0,1,1,0,0), \\ \overline{Z} & = & (0,0,0,0,0|D,1,1,1,1). \end{array} \tag{7.40}$$

■

7.3 Encoding

This section provides an operational method to arrive at an encoding circuit which respects the convolutional structure of the code: a simple unitary operation—independent of the length of the to-be-protected stream—and its n-qubit shifted versions will be

[11] Here, we consider an integer number of physical n-qubit blocks. I wrote "at least" because it is possible that the support of some of the $\overline{X}$ and $\overline{Z}$ is smaller than $n \times (\lambda+1)$.

applied successively to arrive at the protected state. Therefore, the complexity of this scheme in terms of number of gates in the encoding circuit only grows linearly with the number of encoded qubits. This is of particular relevance since dealing with convolutional codes as if they were generic block codes would lead to an encoding circuit with quadratic gate complexity. It would also require increasing precision in the applications of the encoding gates and would cause severe delays in the transmission of the information.

The derivation of the encoding circuit will nonetheless be very similar to the one for block codes [Got97a]. Here, instead of the usual standard form for the generators, we use the standard polynomial form. The circuit that will be obtained is relative to the encoding of $q \times k$ logical qubits. The encoded Pauli operators corresponding to these qubits will be denoted $\overline{X}_{j,i}$ and $\overline{Z}_{j,i}$. For instance, $\overline{X}_{0,i}$ is defined as the i-th line of the $\overline{X}$ matrix derived in the previous section, $\overline{X}_{j,i} = D^j[\overline{X}_{0,j}]$, and similarly for the $\overline{Z}$'s.

The encoding circuit maps the to-be-protected qubits $c_{j,i}$ onto the code subspace. Its action on the computational basis can be written as:

$$|c_{0,1}, \ldots, c_{q-1,k}\rangle \rightarrow \left(\prod_{i,j} \frac{1 + M_{j,i}}{\sqrt{2}} \right) \prod_{r,s} \overline{X}_{s,r}^{c_{s,r}} |0, \ldots, 0\rangle \,, \tag{7.41}$$

for $c_{s,r} \in \{0,1\}$, $0 < i \leq n-k$, $0 \leq j < q+\lambda$, $0 \leq s < q$ and $1 \leq r \leq k$.[12] This operation can be decomposed in two steps. The first one, $\prod_{r,s} \overline{X}_{s,r}^{c_{s,r}}$, applies the different flip operators depending on the value of the to-be-protected qubits in the computational basis. The second, $\prod_{i,j}(1 + M_{j,i})/\sqrt{2}$, projects this state onto the code subspace.[13]

We first focus on the conditional application of the $\overline{X}$'s:

$$|c_{0,1}, \ldots, c_{q-1,k}\rangle \rightarrow \prod_{r,s} \overline{X}_{s,r}^{c_{s,r}} |0, \ldots, 0\rangle \,. \tag{7.42}$$

The number of n-qubit blocks involved in the right hand side of Eq. (7.41) is equal to $q+\lambda+\lceil m/n \rceil$. Hence, the first requirement is to supplement the to-be-protected stream of information with some ancillary qubits prepared in the $|0\rangle$ state. Both are arranged in the following way:

$$\begin{aligned} &|c_{0,1}, \ldots, c_{q-1,k}\rangle \rightarrow \\ &|\overbrace{0 \ldots 0}^{n\times\lambda-k}, c_{0,1} \ldots c_{0,k}, \overbrace{0 \ldots 0}^{n-k}, c_{1,0} \ldots c_{1,k}, \overbrace{0 \ldots 0}^{n-k}, c_{q-1,k} \ldots c_{q-1,k}, \overbrace{0 \ldots 0}^{\lceil m/n \rceil \times n} \rangle. \end{aligned} \tag{7.43}$$

The notation $\overline{X}_{s,r}^{c_{s,r}}$ means that $\overline{X}_{s,r}$ needs to be applied on the all-zeroes state if and only if $c_{s,r} = 1$. Now, in the standard polynomial form, $\overline{X}_{s,r}$ has a factor X exactly at

[12] Here λ is defined as the in the previous section. With this definition, the operators $\overline{X}$ have support on at most $\lambda + 1$ consecutive n-qubit blocks. The choice $j < q + \lambda$ then ensures that the support of each logical qubit is covered by the same number of generators of the stabilizer group.

[13] The way of writing this projection follows from the realization that any element of the stabilizer group is a product where each generator appears at most once—any element of the Pauli group is its own inverse.

the position of $c_{s,r}$ in the state of the right hand side of Eq. (7.43). Therefore, if all the other logical qubits are set to zero, the output state of Eq. (7.42) can be obtained from the right hand side of Eq. (7.43) by applying $\overline{X}_{s,r}$—without the above mentioned X—conditioned on qubit $c_{s,r}$. Unlike for quantum bock codes, these conditional operations can confuse each other when the conditioning polynomial $\Lambda(D)$ (see sec. 7.2) is not a monomial[14]. In this situation, applying $\overline{X}_{s,r}$ might flip some control qubits $c_{s',r}$ for $s' < s$. Therefore, these modified qubits $c_{s',r}$ cannot be used anymore to condition the application of $\overline{X}_{s',r}$. This also indicates the way-out of this problem: when the $\overline{X}$'s are applied by increasing successively the index s by one, there is no risk that one application flips a qubit later used to condition another $\overline{X}$.

The rest of the encoding circuit must implement the effect of the projection onto the code subspace for this partially encoded state:[15]

$$\prod_{r,s} \overline{X}_{s,r}^{c_{s,r}} |0,\ldots,0\rangle \rightarrow \left(\prod_{i,j} \frac{1+M_{j,i}}{\sqrt{2}} \right) \prod_{r,s} \overline{X}_{s,r}^{c_{s,r}} |0,\ldots,0\rangle . \tag{7.44}$$

There are two classes of $M_{j,i}$'s. Either $M_{j,i}$ is a tensor product of I's and Z's only, or there is a polynomial $A_{i,i}$ on the i-th column of the X part when it is expressed in the standard polynomial form (see sec. 7.2). In the first case, nothing needs to be done. In the latter, consider the i-th qubit of the $(j + \deg A_{i,i}(D))$-th n-qubit block in Eq. (7.44). The resulting state is an equal weight superposition of a state with a $|0\rangle$ and a state with a $|1\rangle$ on the previously mentioned qubit. This can be created by first applying a Hadamard gate for this qubit, which later controls the application of $M_{j,i}$—ignoring the X factor for the control. If there is a Z factor for the control qubit, it does not need to be conditioned on anything and can be applied right after the Hadamard gate. Once again, since $A_{i,i}$'s are not required to be monomials, the above operations might confuse each other when a control qubit, supposedly still in its initial $|0\rangle$ state, has indeed already been modified. As before, this can be overcome by applying the conditional gates and increasing the index j one by one successively.

Remark 7.3.1. For sake of simplicity in the presentation of the whole encoding circuit, the usual simplifications corresponding to the removal of control-Z gates acting on a target in state $|0\rangle$ have not been described. Of course, these should be performed to obtain a simpler circuit. ∎

Remark 7.3.2. Note also that the circuit described in this section encodes the qubits on-line:

[14] Here, for sake of generality, we describe the encoding circuit without imposing $\Lambda(D)$ to be a monomial even though, in this case, the encoding shows bad error propagation properties.

[15] The method described here details how to obtain the encoding circuit when the generators $M_{j,i}$ have a positive sign. When this is not the case, the procedure described here must be modified so that a Z gate is applied to the qubit conditioning the application of those particular $M_{j,i}$ with a negative sign.

- the second step rotating the partially encoded state into the code subspace can start before all the $\overline{X}$'s are applied;
- sending the qubits can be done before all the stream has been encoded.

This is a simple consequence of the fact that each conditional gate in the circuit acts only on the last $\lambda + 1$ n-qubit blocks. ∎

Example 7.3.1 (5-qubit convolutional code). For the code given in Ex. 7.2.1, the circuit realizing the application of the encoded $\overline{X}$ operators, Eq. (7.42) is given in Fig. 7.1. The full encoding circuit is presented in Fig. 7.2, where all the simplifications have been implemented (it consisted in removing all Z operations applied to qubits in the $|0\rangle$ state).

Here, the existence of a sacrificed logical qubit is clearly apparent: the first qubit is never involved in any gate and does not contain any quantum information (it remains in the $|0\rangle$ state). This comes from the finiteness of the to-be-protected sequence: at the beginning and at the end of the stream, there are less commutation constraints for the encoded Pauli operators imposed by the generators in M. Thus, it is not surprising that there exist a finite number of encoded Pauli operators that do not follow the convolutional structure. The encoding procedure given in this section forces these logical qubits to be in their logical $|0\rangle$ state: it is taken care of by setting the first λ n-qubit blocks to the all-zeroes state. ∎

7.4 Error propagation and on-line decoding

The previous section was devoted to the derivation of the encoding circuit for quantum convolutional codes. It showed how the standard polynomial form for the generators of the code leads to an automated procedure for finding an on-line encoding circuit. In this section, the focus shifts to decoding quantum convolutional codes. The need for a clear discussion on this issue comes from the specificity of convolutional codes: usual decoding circuits—obtained by running the encoding one in reverse direction—require to wait for the last logical qubit before running them. This is not a practical option as it would cause long transmission delays.

Here, we show that the existence of an on-line decoding circuit is implied by a more fundamental property of the encoding operation: the absence of catastrophic errors. These errors will be defined carefully below, but we can already mention that they are not specific to quantum codes. Rather they, and more generally all the error propagation problems considered in this section, are also encountered in the theory of classical convolutional codes [Lee97a, JZ99a].

To build our intuition on the error propagation problems that might arise when using convolutional codes, consider a generic encoding circuit as derived in the previous section (see also Fig. 7.3). Because of the overlap between the generators on m qubits as defined in Eq. (7.3), quantum information is propagated from one n-qubit block to another. As a consequence, even though the to-be-protected stream of information is in a separable state, say $|0,\ldots,0\rangle$, the encoded state is not, in general, separable with respect to any

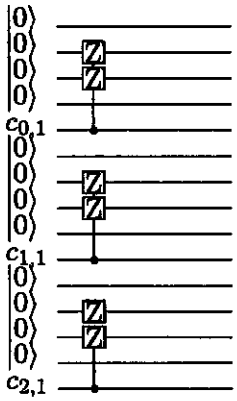

Figure 7.1: Circuit for generating the state $\prod_{r,s}\left(\overline{X_{s,r}}\right)^{c_{s,r}}|0,\ldots,0\rangle$ for the 5-qubit convolutional code. For obvious reasons, the control-Z operations have been kept even though they act on $|0\rangle$ and should be simplified: this part of the encoding circuit would be reduced to no circuit at all!

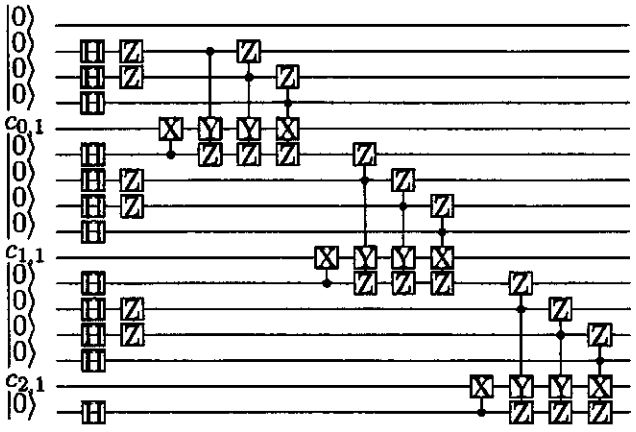

Figure 7.2: Circuit for encoding the first three qubits of a stream of quantum information with the 5-qubit convolutional code. Here, we have removed unnecessary Z gates acting on quibts in the $|0\rangle$ state. Note that the first physical qubit corresponds to a sacrificed logical qubit: in our encoding procedure, it is not used to store information, since it offers no protection to errors. For "pedagogical purposes", it has not been removed from the circuit, but in practice, it is not necessary to transmit it.

bipartite cut. In spite of their relatively simple form—invariant by shifts of n-qubit—encoding circuits apply global unitary transformations that cannot be casted in tensor products of smaller unitary operations.

The good spreading of quantum information induced by the particular structure of convolutional codes might in some cases have a bad consequences: nothing prevents an error affecting a finite number of qubits before the complete decoding of the stream to propagate infinitely through the decoding circuit. Such error is called catastrophic.

Definition 7.4.1 (Catastrophic error). *Consider the encoding scheme of a (n, k, m) quantum convolutional code protecting $q \times k$ logical qubits. A catastrophic error is an error that affects $O(1)$ qubits before the end of the decoding operation and that can only be corrected by a unitary transformation whose size of support grows with q, for large q.*

Remark 7.4.1. The theory of classical convolutional codes explicitly shows the existence of catastrophic errors for some convolutional encoders. As these are a special case of quantum codes—their generators are tensor products of I's and Z's—it proves the existence of catastrophic errors for some quantum encoding circuits. ∎

7.4.1 Catastrophicity condition

In this paragraph, we will find a catastrophicity condition for convolutional encoders without relying on the stabilizer description of the code. Instead, we simply assume a generic form for the encoding operation of $q \times k$ to-be-protected qubits:

$$C(q) = T_{\text{erm}} \times D^{q-1}[U] \times \ldots \times D[U] \times U \times I_{\text{nit}}, \tag{7.45}$$

where I_{nit} and T_{erm} are two fixed unitary transformations, respectively the initialization—acting at the beginning of the to-be-protected stream of information—, and the termination—acting on the last qubits of the stream.[16] The unitary U has a finite support independent of q. In the standard encoding presented in the previous section, U corresponds to the encoding of k consecutive qubits containing information—i.e. it corresponds to applying some $\overline{X}$'s and some $M_{j,i}$'s. The presence of I_{nit} and T_{erm} is due to the sacrificed logical qubits at the beginning and at the end of the encoded stream. The typical arrangement of the unitary operations $D^i[U]$ far from the beginning and the end of the stream of information is depicted in Fig. 7.3.

Proposition 7.4.1. *A quantum convolutional encoder is non-catastrophic if and only if the encoding operation $C(q)$ can be decomposed in the following way for large q:*

$$C(q) = \tilde{T}_{\text{erm}}(q) \times \left(\prod_{i=0}^{\lfloor q/l_t \rfloor} D^{il_t}[U_t] \right) \times \ldots \times \left(\prod_{i=0}^{\lfloor q/l_1 \rfloor} D^{il_1}[U_1] \right) \times \tilde{I}_{\text{nit}}(q), \tag{7.46}$$

[16] The delay operator, initially introduced only for elements of the Pauli group with finite support, is easily generalized to handle unitary matrices with finite support.

where $\tilde{I}_{\text{init}}(q)$ and $\tilde{T}_{\text{erm}}(q)$ are modified initialization and termination steps which can vary with q, but whose support is bounded; $\{U_j\}_j$ is a finite set of unitary operators independent of q—thus with bounded support—such that $D^i[U_j]$ and $D^{i'}[U_j]$ commute; and l_j's are integers independent of q.

Even though this condition might seem at first sight quite complicated, it corresponds to a reordering of the unitaries—or gates—in the quantum circuit which is easy to understand. The new circuit must have the following form: first an initialization step is performed;[17] then, there are t layers of unitaries (each of them made out of a single unitary, e.g. U_i, and its n-qubit shifted versions) such that the gates inside a layer commute with each other; finally it is followed by a termination step, $\tilde{T}_{\text{erm}}(q)$ with bounded support. This structure resembles a pearl-necklace as it can be seen on Fig. 7.4.

Proof 7.4.1 (Sufficiency). To simplify the discussion, we will consider the case where the error E occurs before the beginning of the decoding operation. This is not general, since the definition of non-catastrophicity also imposes to consider errors occurring on a partially decoded stream. Nonetheless, the proof presented here can be easily adapted for this other case.

Here, we have to show that for q large, whenever E has bounded support, $C(q)^\dagger EC(q)$ has a bounded support as well. Since $\tilde{T}_{\text{erm}}(q)$ has a bounded support at the end of the stream, it is always possible to increase q such that E and $\tilde{T}_{\text{erm}}(q)$ commute. Therefore, after simplifying $C(q)^\dagger EC(q)$ by $\tilde{T}_{\text{erm}}$, we have:

$$\begin{aligned} C(q)^\dagger EC(q) \;=\; & \tilde{I}_{\text{init}}(q)^\dagger \times \left(\prod_{i=0}^{\lfloor q/l_1\rfloor} D^{il_1}[U_1^\dagger]\right) \times \ldots \times \left(\prod_{i=0}^{\lfloor q/l_t\rfloor} D^{il_t}[U_t^\dagger]\right) \times E \times \\ & \times \left(\prod_{i=0}^{\lfloor q/l_t\rfloor} D^{il_t}[U_t]\right) \times \ldots \times \left(\prod_{i=0}^{\lfloor q/l_1\rfloor} D^{il_1}[U_1]\right) \times \tilde{I}_{\text{init}}(q). \end{aligned} \tag{7.47}$$

Similarly, in the above equation all the $D^{il_t}[U_t]$ whose support does not intersect the one of E commute with it and can be simplified (recall also that the $D^{il_t}[U_t]$ also commute with each other). Only a finite number of the $D^{il_t}[U_t]$'s remain, say $\{D^{il_t}[U_t]\}_{i\in I_t}$. Note that for q large, this number is independent of q. We thus have

$$\begin{aligned} C(q)^\dagger EC(q) \;=\; & \tilde{I}_{\text{init}}(q)^\dagger \times \left(\prod_{i=0}^{\lfloor q/l_1\rfloor} D^{il_1}[U_1^\dagger]\right) \times \ldots \times \left(\prod_{i=0}^{\lfloor q/l_{t-1}\rfloor} D^{il_{t-1}}[U_{t-1}^\dagger]\right) \times \\ & \times E_1 \times \\ & \times \left(\prod_{i=0}^{\lfloor q/l_{t-1}\rfloor} D^{il_{t-1}}[U_{t-1}]\right) \times \ldots \times \left(\prod_{i=0}^{\lfloor q/l_1\rfloor} D^{il_1}[U_1]\right) \times \tilde{I}_{\text{init}}(q) \end{aligned} \tag{7.48}$$

[17] Because of possible side effects $\tilde{I}_{\text{init}}(q)$ can depend on q but the size of its support must be of order 1 and it can act non-trivially only on the first few qubits

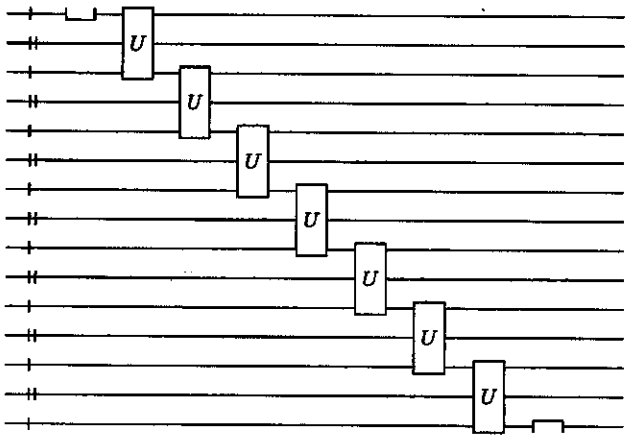

Figure 7.3: Typical encoding circuit for a convolutional code. The circuit is run from left to right. Horizontal lines of a given type (i.e. with single or double vertical bar) always represent the same number of qubits. The unitary operation U is implemented as a series of elementary gates acting only on the qubits with which it intersects.

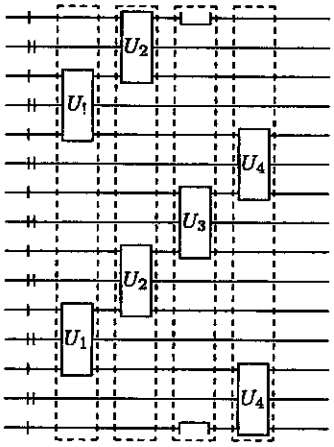

Figure 7.4: Example of pearl-necklace structure for the encoding circuit. We have depicted four layers of unitaries, U_1 through U_4. Here, the condition of commutation inside a layer is guaranteed by the disjointness of the support of the different unitaries $\{D^j[U_i]\}_j$.

where $E_1 = \left(\prod_{i\in I_t} D^{il}[U_t^\dagger]\right) \times E \times \left(\prod_{i\in I_t} D^{il}[U_t]\right)$ has a bounded support, independent of q. The rest of the proof follows immediately by applying the same technique to the remaining layers: another step generates E_2, by considering E_1 instead of E and U_{t-1} instead of U_t. Following the same arguments, E_2 has a bounded support independent of q and so will $E_3, \ldots, E_t$. Thus it proves that $C(q)^\dagger E C(q) = \tilde{I}_{\text{nit}}(q)^\dagger E_t \tilde{I}_{\text{nit}}(q)$ has bounded support. ∎

Proof 7.4.2 (Necessity). To prove that this condition is necessary, we will show that a non-catastrophic encoding operation $C(q)$ can be put in the special form of Eq. (7.46), for q large. The outline of the proof is the following: we will work on the circuit of the decoding operation $C(q)^\dagger$, obtained by running the encoding circuit in the reverse direction (see Fig. 7.5). Our goal is to convert this decoding circuit into an equivalent one which displays the pearl-necklace structure. To do so, we will consider a possible—but yet very particular—error which could occur on the physical qubits during the transmission. The chosen error indeed corresponds to a local reordering of the unitaries in $C(q)^\dagger$. Since the encoding is supposed to have no catastrophic errors, this local reordering can be compensated by applying a unitary operation with finite support after complete decoding. This will give us an identity between two decoding circuits, which we can apply as many times as required to arrive at the pearl-necklace structure.

More specifically, consider the decoding unitary operation,

$$C(q)^\dagger = I_{\text{nit}}^\dagger \times U^\dagger \times D[U^\dagger] \times \ldots \times D^{q-1}[U^\dagger] \times T_{\text{erm}}^\dagger. \tag{7.49}$$

We define the integer l such that U and $D^i[U]$ have disjoint support for $|i| > l$.[18] The circuit identity that will be derived is:

$$C(q)^\dagger = D^{q-l'}[V^\dagger] \times \tilde{C}(q)^\dagger, \tag{7.50}$$

where V has finite support extending on l' n-qubit blocks, and where $\tilde{C}(q)^\dagger$ is obtained from $C(q)^\dagger$ by locally reordering its last $2l+1$ unitaries U:

$$\begin{aligned} \tilde{C}(q)^\dagger \;=\; & I_{\text{nit}}^\dagger \times U^\dagger \times \ldots \times D^{q-2l-3}[U^\dagger] \times \\ & \times D^{q-2l-2}[D^{l+1}[U^\dagger] \times U^\dagger] \times D^{q-2l-1}[D^{l+1}[U^\dagger] \times U^\dagger] \times \ldots \\ & \ldots \times D^{q-l-2}(D^{l+1}[U^\dagger] \times U^\dagger) \times T_{\text{erm}}^\dagger. \end{aligned} \tag{7.51}$$

Consider E, a unitary operation, defined by:

$$\begin{aligned} E \;=\; & (D^{l+1}[U^\dagger] \times U^\dagger) \times D[D^{l+1}[U^\dagger] \times U^\dagger] \times \ldots \times D^l[D^{l+1}[U^\dagger] \times U^\dagger] \times \\ & \times D^{2l+1}[U] \times D^{2l}[U] \times \ldots \times D[U] \times U. \end{aligned} \tag{7.52}$$

An illustration of the arrangement of the unitaries in E is presented on Fig. 7.6 for $l = 1$. By construction, E satisfies:

$$\begin{aligned} \tilde{C}(q)^\dagger \;=\; & I_{\text{nit}}^\dagger \times U^\dagger \times \ldots \times D^{q-2l-3}[U^\dagger] \times \ldots \\ & \times D^{q-2l-2}[E] \times \\ & \times D^{q-2l-2}[U^\dagger] \times D^{q-2l-1}[U^\dagger] \times \ldots \times D^{q-1}[U^\dagger] \times T_{\text{erm}}^\dagger, \end{aligned} \tag{7.53}$$

[18] This integer exists because U has finite support.

which simply corresponds to the initial decoding operation $C(q)^\dagger$ with an error E happening between the unitaries $D^{q-2l-3}[U]$ and $D^{q-2l-2}[U]$. Since, the encoding is non-catastrophic, there exists a unitary $V^\dagger$ with finite support—also obviously independent of q—such that $C(q)^\dagger = D^{q-l'}[V^\dagger] \times \tilde{C}(q)$, where l' is the size of the support of V counted in number of n-qubit blocks, which gives the circuit identity (see Figs. 7.7 & 7.8 for the local reordering implied by Eqs. (7.49–7.53).

Moreover, this identity concerns only the unitary operations around the position where E is applied. It is then possible to apply it at repeated intervals—e.g. separated from $\max(l, l') + 1$ n-qubit blocks—in the decoding circuit. It is then straightforward to show that $\tilde{C}(q)^\dagger$—and similarly $\tilde{C}(q)$—has the form of Eq. (7.46), and to conclude the proof (see Figs. 7.9 & 7.10). ∎

Remark 7.4.2. Note also, that this demonstrates the possibility of on-line decoding for non-catastrophic quantum convolutional codes: in this form, the "directionality" of the quantum circuit which imposed to begin the decoding at the end of the received stream disappeared. ∎

Example 7.4.1 (5-qubit convolutional code). A possible arrangement of gates exhibiting the pearl-necklace structure for the encoding circuit of the convolutional code defined in Ex. 7.2.1 is given in Fig. 7.11. ∎

7.4.2 Catastrophicity condition for standard encoders

Proposition 7.4.2. *Encoders derived from the standard polynomial form are non-catastrophic if and only if $\Lambda(D)$ is a monomial.*

Proof 7.4.3. Simple commutations rules between controlled gates can be used to show that when $\Lambda(D)$ is a monomial, the quantum circuit can be put in the form of Eq. (7.46). To prove the necessity, suppose $\Lambda(D)$ is not a monomial and consider the decoding circuit for this code. More precisely, focus on the qubits that control the application of $\overline{X}_{0,1}, \ldots \overline{X}_{q-1,1}$. If the decoding circuit is restricted to those qubits only, the only two-qubit gates that are used are controlled-NOT's. Thus, this part of the quantum circuit in fact implements a rate 1 classical convolutional encoder with feedback. This encoder links its output stream $y(D)$ with its input $x(D)$ through (see [Lee97a] for a rapid introduction to classical convolutional codes and their polynomial formalism),

$$y(D) = x(D) + (\Lambda(1/D) - 1)y(D). \tag{7.54}$$

Thus, an error affecting the input stream—corresponding to a bit flip in the quantum case—propagates to an infinite number of output bits when $\Lambda(D)$ is not a monomial:

$$y(D) = \frac{x(D)}{\Lambda(1/D)}. \tag{7.55}$$

Similarly, in the quantum case, a single bit flip could propagate to an infinite number of qubits. Thus non-catastrophic standard encoders have a monomial $\Lambda(D)$. ∎

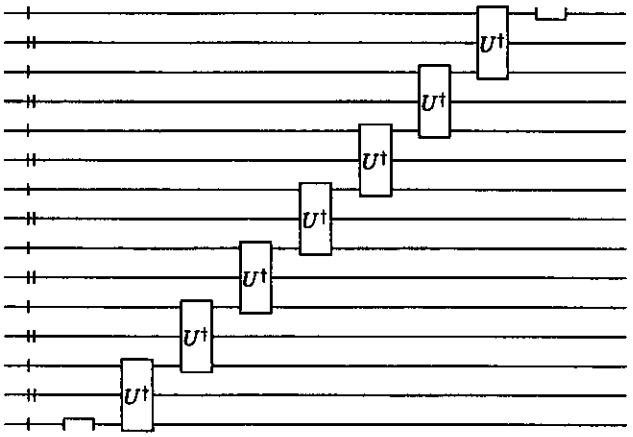

Figure 7.5: Typical decoding circuit for a convolutional code. The circuit is obtained by running the encoding circuit in reverse direction and with appropriate Hermitian conjugates.

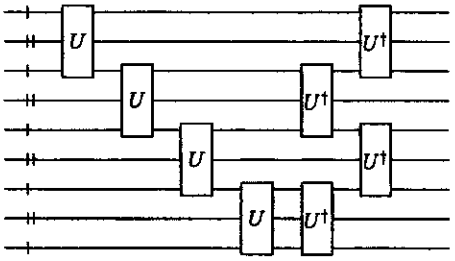

Figure 7.6: Error operation E as defined in Eq. (7.52). Here, $l = 1$ because $D^i[U]$ commutes with U for $i > 1$. When introduced in the decoding circuit, such operation induces a local reordering of the unitaries $U^\dagger$.

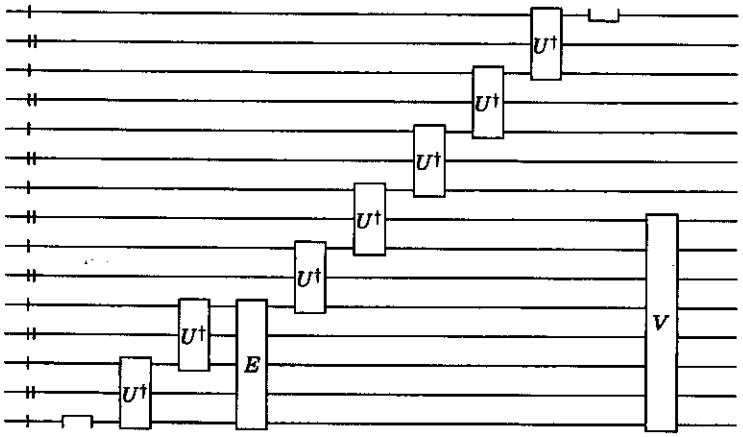

Figure 7.7: Derivation of a circuit identity for decoding. Because there is no catastrophic error, the effect of applying E as defined in Eq. (7.6) in the decoding circuit can be corrected by a unitary operation V with finite support: this circuit induces the same unitary transformation on the received stream of information.

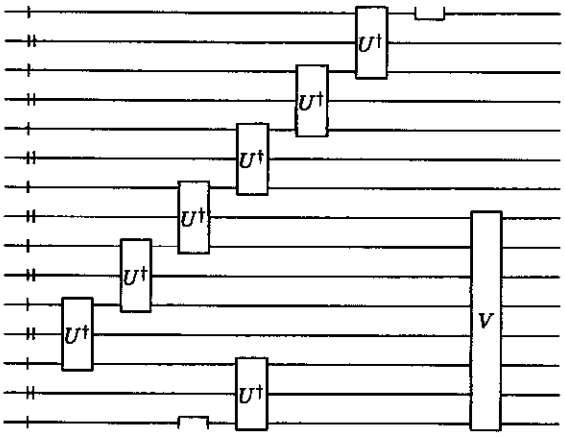

Figure 7.8: Local reordering in the decoding circuit. By using the specific form of E, this circuit is equivalent to the ones given in Figs. 7.5 & 7.8

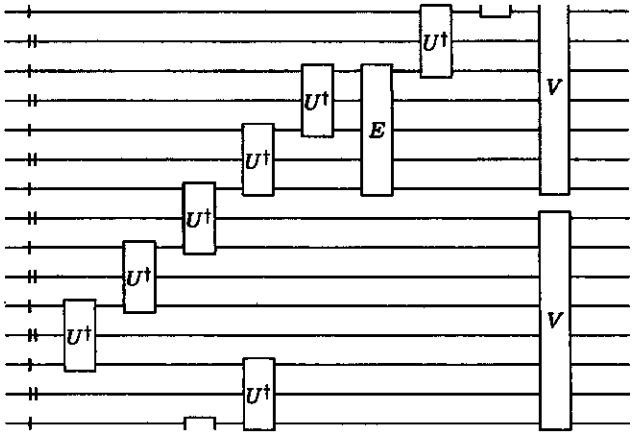

Figure 7.9: Global reordering of the decoding circuit. Exploiting the circuit identity described in Fig. 7.8, the fact that it corresponds to a local reordering only (i.e. only a finite number of unitaries with bounded support are involved in this identity), and the invariance of the initial decoding circuit by n-qubit shifts, it is possible to induce local reorderings at regular intervals in the decoding circuit.

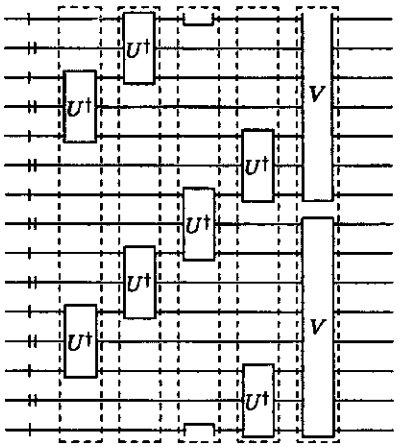

Figure 7.10: Pearl-necklace structure after global reordering of the decoding circuit. Each layer of the structure is identified by a dashed box. The necessity of introducing new definitions for the initialization and termination steps, I_{nit} and T_{erm}, is due to the impossibility of applying the local reordering when few $U^\dagger$'s remain at the beginning or at the end of the decoding circuit (less than the number of n-qubit blocks involved in the support of V).

Remark 7.4.3. Note also that the condition "$\Lambda(D)$ is a monomial" is equivalent to having the $\overline{Z}$ operators efficiently described with the polynomial formalism. These two questions are in fact intimately related. The application of a $\overline{Z}$ can be done before encoding by applying the corresponding Z to the physical unprotected qubit. It is well known that phase flips propagate through controlled-NOT gates from the target to the control. Here, this phase flip propagates in the same way the bit flip of the proof propagates in the decoding circuit. The number of qubits affected by this Z operation after running the encoding increases linearly with q, the number of k-qubit blocks to be protected. More generally, the non-catastrophicity condition shows that contrarily to classical convolutional codes, an operation with finite support acting before encoding cannot propagate to an infinite number of qubits after encoding. ∎

7.5 Error estimation algorithm

The last subject that must be addressed to arrive at a theory of quantum convolutional codes is the error estimation algorithm. A naive attempt at finding the most likely error could be to search among all the possible errors. In turn, this usually implies an exponential complexity in the number of encoded qubits, thus making this scheme impractical for large amounts of to-be-protected information. In this section, a maximum likelihood estimation algorithm with a linear complexity is provided. This algorithm is similar to its classical analog, known as the Viterbi algorithm [Vit67a, Lee97a, JZ99a].

7.5.1 Notation

To simplify the description of the algorithm, some additional notation will be useful. Recall Eq. (7.3) which defines the generators of the stabilizer group $M_{j,i}$. The expression "block j" will refer to the qubits involved in $M_{j,i}$ for $i = 1, \ldots, n-k$. The qubits are numbered in increasing number from left to right, so that the first m qubits and the last n qubits of the second block are those separated on Eq. (7.3) by a dashed line. Note also that due to the convolutional nature of the code and because of the definition of m, the last m qubits of block j are the same as the first m qubits of block $j+1$. The syndrome $s_{j,i}$ for a received stream of information is the result of the projective measurement associated to the $M_{j,i}$. It is equal to $+1$ (resp. -1) if the measured state belongs to the $+1$ (resp. -1) eigenspace of $M_{j,i}$. An element of the Pauli group of the transmitted qubits is said to be compatible with the syndrome $s_{j,i}$ if it commutes (resp. anti-commutes) with $M_{j,i}$ when $s_{j,i} = 1$ (resp. -1). An error candidate up to block j is an operator of the Pauli group defined on all the qubits up to block j and which satisfies all the syndromes up to block j. The likelihood of an error candidate is the logarithm of the probability of getting this particular error pattern according to the channel model. Since we consider memoryless channels, the likelihood is the sum of the logarithms of single-qubit-error probabilities.

7.5.2 Quantum Viterbi algorithm

The algorithm examines the syndromes block by block and updates a list of error candidates among which one of them coincides with the most likely error. All this algorithm is classical except the syndrome extraction procedure.

The value of the syndrome is obtained by the usual phase estimation circuit: an ancillary qubit is prepared in the $|0\rangle$ state; undergoes a Hadamard gate; conditionally applies $M_{j,i}$; is once again modified by a Hadamard gate; and is measured according to the Z observable. The result of this measure is the value of the syndrome $s_{j,i}$.

Algorithm 7.5.1 (Quantum Viterbi algorithm).
Inputs:

1. *The list of syndromes* $\{s_{j+1,i}\}_i$ *for* $i = 1, \ldots, n-k$;
2. *a list* $\{E_j^{(e)}\}_{e \in \{I,X,Y,Z\}^{\otimes m}}$ *of error candidates up to block* j *such that the element* $E_j^{(e)}$ *corresponding to the index* e *has a tensor product decomposition ending by* e *for its last* m *qubits and maximizes the likelihood given the previous constraint.*

The list $\{E_j^{(e)}\}_e$ *is constructed recursively.*

Step $j+1$: *For a given value of* $e' \in\in \{I,X,Y,Z\}^{\otimes m}$, *consider all the possible* n*-qubit extensions of the elements of* $E_j^{(e)}$ *such that:*

- *they satisfy the syndromes* $s_{j+1,i}$ *for* $i = 1, \ldots, n-k$;
- *they have the prescribed tensor product decomposition* e' *on their last* m *positions.*

By construction, these extensions are error candidates up to block $j+1$. *For each element* e' *of* $\{I,X,Y,Z\}^{\otimes m}$ *select one such extension with maximum likelihood—take one at random among them in case of tie. This constitutes the new list of error candidates* $\{E_{j+1}^{(e')}\}_{e'}$.

When all the syndromes have been taken care of in this way, select the most likely error candidate of the list. This error candidate is one of the most likely errors compatible with all the syndromes.

Proof 7.5.1. Consider a most likely error E_p for the whole p blocks of syndromes. The truncation of this error to the first $p-1$ blocks, E_{p-1}, is by construction an error candidate up to block $p-1$. This error candidate has maximum likelihood given its decomposition on the last m qubits. If it was not the case, another error candidate, $\tilde{E}_{p-1}$, with the same decomposition on the last m qubits could be extended up to block p by concatenation with the last n Pauli operators of E. It would therefore have a strictly greater likelihood than E. Recursively, this property holds for E_j: it has maximum likelihood given its tensor product decomposition on the last m positions. Thus, at each step j of the algorithm, one element of the list coincides with the most likely error up to block j. ∎

Remark 7.5.1. Note that in the encoding of quantum convolutional codes, we chose to set to $|0\rangle$ some logical qubits that were not described by the polynomial formalism. This was done formally by adding their $\overline{Z}$ operators to the stabilizer group of the code. Hence either the first and last steps of the algorithm should be modified to take into account these extra syndromes. ■

Remark 7.5.2. It is also important to understand that in the error estimation algorithm presented above, the most likely error is known only at the end of the algorithm. However, in practice the error candidates considered at step j all coincide except on the last few blocks. Hence, the most likely error is known except on the last few blocks. Some simulations for a depolarizing channel with error probability less than 0.05 showed that keeping two blocks in the 5-qubit convolutional code was enough to estimate the most likely error with high probability. ■

7.6 Conclusion

This chapter showed the basis of quantum convolutional coding. An appropriate polynomial formalism has been introduced to handle the codes efficiently and to make calculations consistently with their specific structure. A procedure for deriving an encoding circuit with linear gate complexity has been given together with a condition which warrants the good behavior of this circuit with respect to error propagation effects. Finally, the quantum Viterbi algorithm has been given explicitly. This algorithm finds the most likely error with a complexity growing linearly with the number of encoded qubits.

More importantly, as the reader familiar with classical convolutional codes can notice, other error estimation algorithms, such as Bahl's [BCJR74a] algorithm—a stepping stone toward turbo-decoding—, can readily be employed with the codes described here. Hence, quantum convolutional codes open a new range of efficient error correction strategies.

Further studies on this subject will include characterizations of errors that remain after decoding, as well as appropriate concatenation schemes. Theoretical quantities such as the free distance, the constraint length, etc, should also be generalized to the quantum case. Of course, finding an example of turbo code will be another subject of attention.

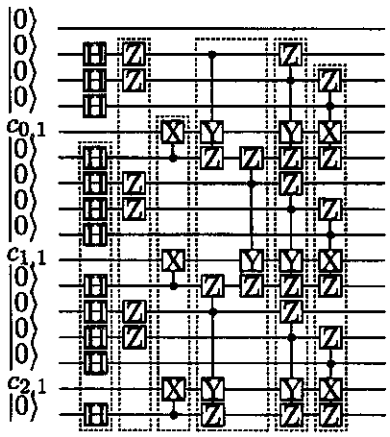

Figure 7.11: Encoding circuit for the 5-qubit convolutional code with the pearl-necklace structure. Each dashed box represents a different layer in which the unitaries commute. Note that the first three Hadamard gates cannot be put into a layer, but rather form the unitary I_{nit}.

III Cavity QED proposals

8 Building a quantum computer

IN PARALLEL TO the theoretical developments induced by the emergence of quantum computing, many efforts have been dedicated to implement these novel ideas with physical devices. The inherent instability and sensitivity of quantum information makes the task extremely challenging.[1] Currently, a large number of different experimental options are under scrutiny. Some are specifically designed for quantum information processing, while others have been built to perform tests of quantum mechanics, but have now been successfully recycled.

In this chapter, we will review five stringent requirements that ensure that a quantum system is well suited for being a building block of quantum computer. Even though they might not be necessary, they are often considered as such, and are definitely a good starting point for analyzing experimental proposals of quantum information processors. These requirements are summarized in the "DiVincenzo criteria" [Div00a]. After a brief review of these criteria, we will look at some widely used experimental setups for quantum information processing. Finally, we will detail one of them—cavity QED—as a preparation for the next two chapters dedicated to universal quantum cloning and the making of an elementary gate.

8.1 DiVincenzo Criteria

When quantum information appeared, most of the researchers thought that the speed-up of quantum computers was due to computing in a Hilbert space rather than with real numbers. DiVincenzo studied this assertion more carefully. He popularized the idea that a Hilbert space is not enough to gain over classical computers. Through simple criteria, he defines the necessary improvements for any experimental setting to become a viable quantum information processor. However, not meeting all of these requirements does not necessarily mean that nothing interesting can be done with those devices.

[1] In my opinion, this unescapable matter of fact is a consequence of the physical nature of quantum information. Classical information—mainly because of its discreteness—benefits from the existence of intrinsic error correction at the physical level. For instance, the electric charge of a capacitor in a memory cell, representing the value of a bit, might well decrease over time, but yet it will retain the original value of the bit long enough so that the content of the memory can be refreshed only once in a while. However, any continuous transformation for an unencoded qubit corresponds to a change of the quantum information it contains. As a consequence, quantum information requires encoding at the physical level, which is not necessary for classical information.

- *The implementation must be scalable with well characterized qubits.* It is probably the most important of all criteria because it states that our efforts in studying small scale quantum information processing devices should help us in building large scale ones. More precisely, building a large quantum computer should be reduced to allow many small elementary devices exchange quantum information. The history of classical computers exemplifies this concept as new generations of processors usually reuse most of the technology developed at earlier stages. On the quantum side, it seems that this requirement is hardly met by any of the proposed implementations. Optical networks are the exception among them because they offer quite easily scalability, but having a thousand-qubit optical quantum computer would probably require a whole building by itself.

  It is important to note, that this requirement has another important justification: it imposes a natural tensor product structure to each of the qubits to avoid the exponential growth of resources associated with unary representations of the Hilbert space used in the computation. It is in this sense prior to any other requirements. For instance, quantum computer that would use the Hilbert space of a single circular Rydberg atom is not scalable since doubling the size of the state space requires to access directly twice as many energy levels. In comparison, using well characterized qubits reduces the energy requirements of this simple scheme to only a logarithmic growth.

- *It must be possible to initialize the qubits to a simple fiducial state.* As users of classical computers, we all know what this requirement is: there should be a reset button to reboot the computer once in a while. However, for quantum computers rebooting is an operation that must be performed very carefully at the begining of each new calculation. Any algorithm starts with each qubit in a given initial state—often taken to be $|0\rangle$ of the computational basis. The qubits that are used in error correcting codes also need a reset as they cool the qubits used for the computation. They need to evacuate the excess of entropy from the quantum register. This operation is far from being granted as qubits in quantum computers follow reversible evolutions. In fact, this reset imposes an interaction of the register with a cold environment in order to drive them to a fixed given state.[2] Once again, this operation puts forward the inevitable trade-off between decoherence and control in quantum computers: qubits must be isolated as much as possible to follow unitary transformations, and at the same time they should also be measurable which necessitates an interaction with the external classical world.

[2] Here the environment can actually be a measuring device and a control device. The control makes a conditional evolution upon the measurement result to transform the measured state into the initial one. The classical registers recording the measurement result extract entropy from the qubit and evacuate it when another initialization is required. Note that this setting is very similar to the Maxwell's demon setting seen earlier.

- *The decoherence time must be long compared to the gate operation time.* In other words, the natural behavior of the system should allow for keeping the quantum information during a sufficiently long time so that some quantum gates can be applied to the qubit before an error occurs. In the worst case scenario, most of those gates would be used to implement error correction, while only very few of them would actually perform the useful calculation. Note that such schemes explicitly require scalability as adding layers of error correction requires adding more physical qubits.

  The study of the efficiency of error correction for quantum computation is the goal of fault-tolerance. For given codes and error models it gives a threshold value such that: *i*) below the threshold, error correction scheme actually improve the number of operations that can be performed before an error occurs, while *ii*) above the threshold, error correction scheme actually introduces more errors because quantum information is manipulated too often. In simple cases, it is possible to derive the value of the threshold, thus obtaining indications on the real value of this threshold. Currently derivations have provided numbers varying between 10^{-12} up to 10^{-5}, depending on the model used. However, the main point here is not that the value of the threshold is low, but that a threshold exists, and that it is independent of the length of the computation and the number of qubits involved in the scheme.

  On the practical side, most experimental schemes have a decoherence versus gate operation time ratio of the order 0.05 and still need a tremendous amount of work to reach the safe 10^{-4} zone.

- *It must be possible to implement a universal set of gates.* Any quantum algorithm is a succession of elementary gates which in turn have to be casted into universal gates. Failing to implement a universal set simply restricts the kind of algorithms that can be performed on the computer. It does not allow the state to explore the full Hilbert space at its disposal.

- *It must be possible to measure each qubit specifically.* In all the quantum algorithms, the qubits contain the information about the solution of the problem at the end of the unitary evolution. Therefore, it is imperative to be able to extract this information from the qubits themselves. There is no possibility of a read-out directly at the quantum level. Most of the new techniques develop for satisfying this requirement will actually be greatly appreciated by other areas of physics as they provide very sensitive detectors.

These criteria have been a subject of focus for all experimental groups around the world in the past few years. Recently the criteria were also used to grant money to different experimental propositions according to the potential satisfaction to these requirements. However, a closer examination indicates that not meeting some of these does not ruin the future of an experimental proposal for processing quantum information. For example scalability is often mentioned to be the main problem of NMR based

quantum processors. But, on the other hand, this technology is the only one at the moment for which we can hope that it will offer a few tenths of useful qubits in a quite near future. Certainly factoring large integers will still be impossible but interesting physics problems could be tackled quite easily with such computational power.

In conclusion, these criteria should be taken as guidelines, but not borderlines to differentiate good and evil.

8.2 Experimental achievements

Present day techniques just allow for building very small scale quantum information processing devices. None of them actually meet all the DiVincenzo criteria [HH02a], but they must be regarded as incredible experimental achievements. First of all they improve our knowledge about quantum mechanics itself: most of the thought experiments that were designed by the founding fathers of the quantum theory have been conducted and have confirmed the predictions of the theory. Quantum information helped rephrasing these into a new communication oriented framework and shed a new light on quantum mechanics. However, in most cases, this was not the original motivation for building these devices. They were thought to test how some algorithms could be implemented effectively by using quantum systems. The result of such research is that the field is achieving goals that everybody thought inaccessible for the next 50 years at least: factoring the number 15 using Shor's was performed at MIT in 2001.

The relevance of these experiments is also due to the importance of evaluating the capacity of specific devices to process information in a quantum way. Note also that, because all of them are going toward a unique goal, they share many common aspects and difficulties, and many solutions to their problems as well. Most techniques used to reduce noise or to improve the precision of the control on the qubits are not device-specific: the robust one qubit pulse sequence for implementing a π-rotation was developed for NMR but could be used for trapped ions, or cavity QED as well.

It is of course well possible that none of these techniques and implementations will ever succeed in building a quantum computer, but they increase our knowledge of the behavior of quantum systems in general. This knowledge will be crucial when the appropriate approach will be found.

8.2.1 Cavity QED

This will be described in greater details in the following section. In short it uses the electronic structure of neutral atoms to store quantum information. The interaction between atoms is mediated by a micro-wave or optical cavity and allows the multi-qubit gates to be performed. Detection of the state varies between the different proposals: ionization is one possible method, but will destroy the carrier of the quantum information, shining a laser pulse onto the atom is less destructive but achieves lower performances while keeping the atom intact.

quantum processors. But, on the other hand, this technology is the only one at the moment for which we can hope that it will offer a few tenths of useful qubits in a quite near future. Certainly factoring large integers will still be impossible but interesting physics problems could be tackled quite easily with such computational power.

In conclusion, these criteria should be taken as guidelines, but not borderlines to differentiate good and evil.

8.2 Experimental achievements

Present day techniques just allow for building very small scale quantum information processing devices. None of them actually meet all the DiVincenzo criteria [HH02a], but they must be regarded as incredible experimental achievements. First of all they improve our knowledge about quantum mechanics itself: most of the thought experiments that were designed by the founding fathers of the quantum theory have been conducted and have confirmed the predictions of the theory. Quantum information helped rephrasing these into a new communication oriented framework and shed a new light on quantum mechanics. However, in most cases, this was not the original motivation for building these devices. They were thought to test how some algorithms could be implemented effectively by using quantum systems. The result of such research is that the field is achieving goals that everybody thought inaccessible for the next 50 years at least: factoring the number 15 using Shor's was performed at MIT in 2001.

The relevance of these experiments is also due to the importance of evaluating the capacity of specific devices to process information in a quantum way. Note also that, because all of them are going toward a unique goal, they share many common aspects and difficulties, and many solutions to their problems as well. Most techniques used to reduce noise or to improve the precision of the control on the qubits are not device-specific: the robust one qubit pulse sequence for implementing a π-rotation was developed for NMR but could be used for trapped ions, or cavity QED as well.

It is of course well possible that none of these techniques and implementations will ever succeed in building a quantum computer, but they increase our knowledge of the behavior of quantum systems in general. This knowledge will be crucial when the appropriate approach will be found.

8.2.1 Cavity QED

This will be described in greater details in the following section. In short it uses the electronic structure of neutral atoms to store quantum information. The interaction between atoms is mediated by a micro-wave or optical cavity and allows the multi-qubit gates to be performed. Detection of the state varies between the different proposals: ionization is one possible method, but will destroy the carrier of the quantum information, shining a laser pulse onto the atom is less destructive but achieves lower performances while keeping the atom intact.

8.2.2 Optical networks

Photons are good prototypes for storing qubits. Moreover, all the techniques to manipulate them individually are mastered since quite a time. The information itself can be stored into the polarization or the mode of the field. Their evolution is simply obtained by placing onto the light path different optical elements such as plates and crystals. However, because photons do not interact effectively with each other, building two-qubit gates remains a difficult challenge. This can be done with the use of non-linear optics, but at relatively low success rate. A wonderful scheme has been proposed recently to build an all-linear optics quantum information processor[KLM01a]. Unfortunately, to implement it, an enormous overhead in terms of optical elements is needed. This scheme is scalable, but still very difficult to build. On the other hand, because photons rarely interact with other residual fields, they have very long decoherence time. Another problem inherent to this technique is the necessity of having single photon sources on demand, because synchronization of the different beams that must cross and interact in a very small volume is crucial. This is still a challenge to be overcome even if a lot of new and promising ideas did emerge in the last few years.

8.2.3 Solid state

Solid state proposal are numerous. From quantum Josephson junctions to quantum dots, many problems arise due to the very poor insulation of the quantum information carrier from its environment—the rest of the solid. Most of this propositions actually struggle for having a single qubit. Many problems also seem to come from the very low efficiency of the measurement. Two-qubit gates are the next big challenge to address and it is nowadays hard to make a guess on one or the other of the possible options. The advantage of these schemes is the very strong interactions between qubits giving gate operation times of the order of 10 ns which balances the short decoherence time. Also, they can rely on the existing technology for building integrated circuits.

8.2.4 NMR-based quantum processors

NMR stands for Nuclear Magnetic Resonance. It is a tool used by many chemists and material scientists to study spatial configurations of molecules or defects in solids. The working principle is to use the spin of some atoms inside the molecule. A typical experiment consists in exciting the spins with a strong magnetic field and to record their response. Because each spin is interacting with its neighbors one can deduce from the recorded signal some information about the environment of each spin.

For manipulating quantum information, the situation is quite modified: the sample used by the NMR spectrometer is made of a carefully designed molecule whose structure is very well characterized and which has been chosen accordingly to its capability to hold and process quantum information—they have very long decoherence times for each of the spins. There, nuclear spins of the molecule can represent qubits, which are manipulated individually by an external magnetic field tuned to the adequate frequency—each spin

in the molecule rotate constantly at a frequency depending upon its position in the molecule. The two-qubit gates rely on the existence of interactions between the different spins inside the molecule: depending on the state of the neighboring spins, the frequency of rotation is slightly shifted which allows conditional operations.

Decoherence for this kind of systems is relatively small compared to the gate operation time. The ratio is about 10^{-2} which indicates that 100 operations can be performed before the system completely relaxes to the thermal equilibrium.

This relatively favorable situation lead to several proposals and some experimental realizations [HSTC99a, CLK+00a], for which error correction schemes were demonstrated [KLMN01a]. The advantage of NMR spectrometers is that they are commercially available and are almost ready to implement simple algorithms. However, the choice of the molecule is a crucial point, which benefits from interactions with chemists. Another more serious problem is that the description of the state of the quantum registers is not exactly the one expected in a quantum computer: there are around 10^{23} molecules in the sample, but only a very small fraction of them is actually useful for the computation. Hence, there are no pure states inside the computer, but only mixed ones. More precisely, only a small deviation from the completely mixed state is induced by the magnetic field. In the preparation of the state the $|0\rangle \dots |0\rangle$ fiducial state, the deviation decreases exponentially with the number of qubits. In principle this can be overcome by algorithmic cooling techniques, but they require so many qubits that they are impractical with the available molecules. NMR quantum processors are thus currently not scalable, but they are still the most powerful ones and did achieve the most complicated algorithms to date.

8.2.5 Ion traps

This approach shares many common aspects with NMR. It can actually be thought of as an NMR spectrometer that works with a single molecule—each qubit is now carried by the electronic structure of ions in the trap, and the molecular binding is replaced by electrical repulsion of the ions. The detection and manipulation of single qubits is actually performed by lasers instead of a radio antenna for NMR. Multi-qubit gates benefit from the vibrational motion inside the trap to induce conditional evolution between neighboring ions.

The decoherence rate in these experiments is much higher than for NMR or optical networks. But on the other hand it seems quite possible to make the whole scheme scalable because there is no ensemble computation as in NMR. Making the whole scheme scalable would imply moving the ions inside a more complex trap with an interaction zone without actually destroying quantum coherences. Some experiments did achieve encouraging results, but once again the technical difficulties are still tremendous.

8.3 Cavity QED: experimental setting

In the following chapters, two experimental proposals for realizing simple algorithms in cavity QED are described. More specifically, the setting they rely on is the one at École Normale Supérieure in the S. Haroche, J.-M. Raimond and M. Brune group. This experiment was originally built to explore quantum electrodynamics for simple systems and decoherence. Nowadays, most of the propositions are expressed in the framework of quantum information (see [RBH01a, Ber02a] and references therein). Many simple algorithms have already been implemented and many more propositions benefit from the exceptional qualities of the cavities which give a decoherence versus gate operation time of the order 10.

8.3.1 Circular Rydberg atoms

In this scheme quantum information is stored in the electronic structure of circular Rydberg atoms. Rydberg states for an atom are states with a large principal number and a maximal orbital number. The valence electron is on a circular orbit which is well approximated by the classical Bohr orbit. More precisely, three levels $|i\rangle$, $|g\rangle$ and $|e\rangle$ are populated in this experiment and they have a principal number of 49, 50 and 51 respectively. The different transitions, $|e\rangle \leftrightarrow |g\rangle$ and $|g\rangle \leftrightarrow |i\rangle$, happen at a frequency of 51.1 GHz and 54.3 GHz. To avoid transitions between elliptical and circular states which would otherwise happen at the same frequencies, a small electrical field lifts the degeneracy between the levels. This field can also be used to change the detuning between the cavity and the transition using the Stark effect. These atoms persist in their state as long as they do not emit a photon spontaneously. This radiative lifetime has been estimated to be around 30 ms, which is much longer than the time each atom spends in the experiment (≈ 0.2 ms).

These atoms are obtained by heating rubidium at 190 ^{o}C in an oven. Atoms with a specific velocity are selected for preparation by Doppler selective optical pumping techniques. Theses atoms are then excited by lasers and radio-frequencies to obtain the circular states. The circular state purity reaches 98%. The speed uncertainty is about 2 m/s and the time interval in which the atoms are prepared is of order 2 μs. Since the number of atoms per packet follows a Poisson statistics, 0.1 atom per pulse is produced on average to warrant that two-atom events are rare.

The atoms then cross the superconducting cavity in which all the interactions aimed at processing quantum information will be performed. One very big constraint in this setting with a single cavity is that once an atom leaves the cavity, no more operation can be done.

8.3.2 Superconducting cavity

The superconducting cavity is made of two polished niobium mirrors capable of holding an electromagnetic field at a frequency close the $|e\rangle \leftrightarrow |g\rangle$ transition. This resonance frequency can be modified and controlled by piezo-electrical crystals. The quality factor

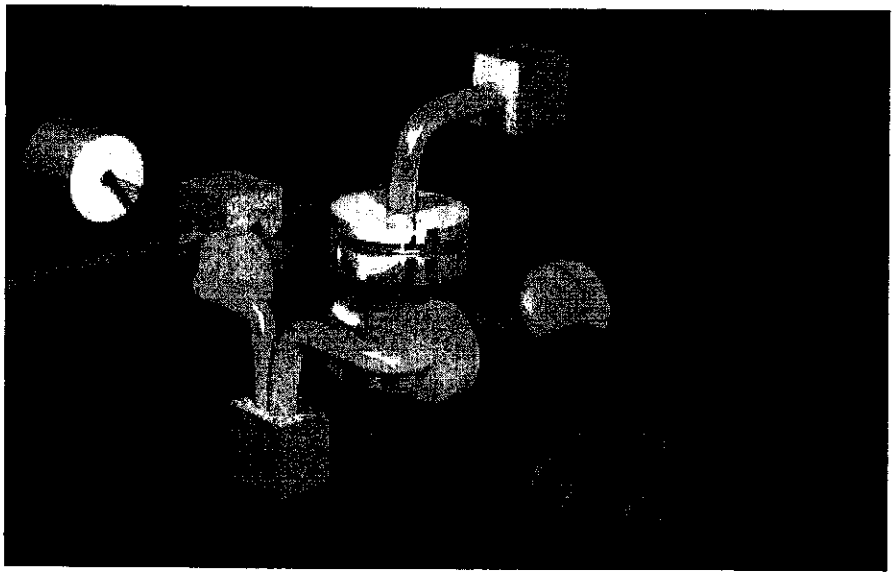

Figure 8.1: Diagram representing the cavity QED scheme. The atoms follow the path in red, from left to right. They come from the oven (white cylinder); then, they are prepared in their circular state; enter the first Ramsey zone (yellow and blue); enter the single mode high finesse cavity; go through the second Ramsey zone; and are finally detected by ionization between the red plates.

of the cavity is around 3.2×10^8 which corresponds to a photon-lifetime of about 1 ms. To obtain these values, it is actually necessary to dispose a closing ring around the gap between the mirrors. This ring has entrance and exit holes for the atomic beam. But as a result of the very inhomogeneous electrical fields around the holes, all coherence between atomic levels is lost. Hence, the coherent manipulations must be performed inside the cavity. The single atoms manipulations are actually resulting from an interaction with a classical microwave field which is injected perpendicularly to the path of the atoms by wave guides ending in the ring.

8.3.3 Detectors

The detectors measure in which state $|i\rangle$, $|g\rangle$ or $|e\rangle$ the atom exits the cavity. This is performed by a selective ionization of the atoms: depending on the state, the field strength required to ionize the atom varies. Thus, by setting a low voltage in the detector only $|e\rangle$ will loose its electron in the detector. This electron is then multiplied to obtain a macroscopic signal with a 40% efficiency. Two detectors are used and are set to observe $|e\rangle$ and $|g\rangle$.

8.3.4 Comments

This cavity QED scheme is a wonderful tool to explore very simple algorithms and elementary quantum gates. Due to the difficulty of having many cavities, and of having atoms on demand, it does not really seems reasonable to think of it as a good proposal for building quantum computers. However, many groups study neutral atoms traps which could stay longer times inside such cavities. Holding information inside the cavity would then be possible and would allow many more gates to be performed. It will then be straightforward to adapt most of the schemes that were developed in the above setting to this new configuration.

9 Universal quantum cloning in cavity QED

QUANTUM THEORY PROVIDES new and unexpected effects when compared to classical physics. Among them, the no-cloning theorem, derived in 1982 by Wooters and Zurek [WZ82a, Die82a], plays a particularly important role: While classical information can be copied perfectly and many times, quantum information cannot. This fundamental difference is a consequence of the unavoidable creation of quantum correlations. Since perfect cloning is not possible, an important question naturally arises: What is the best quantum copying operation? The answer to this question is context-dependent. On the one hand, there is a single transformation that produces the best identical copies of a qubit prepared in any input states. This universal quantum cloning machine (UQCM) has been discussed for the first time in [BH96a]. On the other hand, many other rules of the game can be considered, such as state dependent cloning [BDE+98a, BCDM00a], cloning of 3-dimensional states [CDG01a] and cloning of orthogonal qubits [FIMC01a].

The quality of a copy is usually measured by the quantum fidelity [Joz94a]. This quantity is discussed, in the context of UQCM, in [BH96a] and [GM97a]. When M copies are produced from N identical pure two-dimensional states, the fidelity of the copies is given by $F(N,M) = (NM+N+M)/(M(N+2))$. For the simplest case of two copies produced from one input state, this expression reduces to $F(1,2) = 5/6$. The complete understanding of the fidelity behavior versus N and M is still a subject of debate, with connections to the measurement and state estimation problems [Ban01a, BEM98a]. Beyond these fundamental problems, the interest of quantum cloning machines also encompasses a wide area of quantum information processing, including quantum cryptography, teleportation [GG01b], eavesdropping, state preservation and measurement-related problems, as well as quantum algorithm improvements [GH00a].

The derivation of the optimal UQCM transformation has led to several proposals [SWZ00a, KSW00a] for its experimental implementation. Most of them, based on the Buzek and Hillery quantum logics network [BBHB97a], use the quantum optics framework. Experimental quantum cloning has been realized up to now only with photons as the carriers of quantum information. This information was either encoded in different degrees of freedom of the same photon (polarization and position) [HLL+01a] or in the photon polarization only [LSHB02a, FGR+02a]. An alternative network adapted to NMR-based quantum information processors has also been proposed and experimentally implemented [CJF+02a].

In this paper, we propose an implementation of the $1 \rightarrow 2$ UQCM operating for atomic states in the Cavity QED (CQED) context [RBH01a]. The quantum information is coded on electronic levels of long-lived highly excited Rubidium (Rb) atoms. Our

protocol realizes, with four atoms, the transformation described in [BH96a], with an original quantum logics network based on the resonant interaction between the atoms and two high-Q niobium superconducting microwave cavities C_a and C_b. We discuss, at the end of this paper, an adaptation of the scheme using two different modes of a single cavity [RBO+01a, BOM+02a], making the proposal implementation more realistic with our present cavity QED set-up. This paper focuses on the quantum logics protocol. The interested reader can find more details about the experimental techniques in [RBH01a].

Let us first recall the optimal $1 \rightarrow 2$ UQCM transformation [BH96a]. To account for all the possibilities leading to a correct transformation—in particular, the many ways ancillas can be used in the process—, we introduce three particular states $|\mathcal{B}\rangle$, $|\mathcal{A}\rangle$, and $|\mathcal{A}_\perp\rangle$. $|\mathcal{B}\rangle$ represents the intial fiducial state of all the resources—but the to-be-cloned qubit—that are available *before* the cloning process. It includes ancillas together with some blank qubits that will receive the cloned information. $|\mathcal{A}\rangle$ and $|\mathcal{A}_\perp\rangle$ are two orthogonal states that correspond to the state of the cloner—the machine itself, not its output—*after* the process has been completed. When the computational basis of the qubits is defined by $\{|+\rangle, |-\rangle\}$, the UQCM must perform the transformation

$$\begin{aligned} |-\rangle|\mathcal{B}\rangle &\rightarrow \sqrt{\tfrac{2}{3}}|-\rangle|-\rangle|\mathcal{A}\rangle + \sqrt{\tfrac{1}{3}}|\Phi\rangle|\mathcal{A}_\perp\rangle \\ |+\rangle|\mathcal{B}\rangle &\rightarrow \sqrt{\tfrac{2}{3}}|+\rangle|+\rangle|\mathcal{A}_\perp\rangle + \sqrt{\tfrac{1}{3}}|\Phi\rangle|\mathcal{A}\rangle, \end{aligned} \tag{9.1}$$

where the first ket of the *l.h.s* represents the input qubit and $|\mathcal{B}\rangle$ is the initial state of the blank copies and of possible ancilla qubits involved in the process. In the *r.h.s*, the first two kets are the quantum clones, $|\Phi\rangle = (|+\rangle|-\rangle + |-\rangle|+\rangle)/\sqrt{2}$. The third ket represents two possible orthogonal final states, $|\mathcal{A}\rangle$ and $|\mathcal{A}_\perp\rangle$ for the ancilla qubits. Tracing with respect to the ancilla will leave us with a two-bit mixed state. By tracing again with respect to each one of them, we obtain the same one-qubit mixed state, which has a fidelity of 5/6 when compared to the original state.

Our scheme makes use of three atomic levels, $|e\rangle$, $|g\rangle$ and $|i\rangle$. The transition between levels $|e\rangle$ and $|g\rangle$ can be set in and out of resonance with the cavity mode using the Stark effect induced by an electric field applied between the Fabry Perot cavity mirrors [RBH01a]. The auxiliary level $|i\rangle$ is far off-resonant from the cavity fields and is not coupled to them. However, it can be accessed via classical microwave pulses either from level $|g\rangle$ (one-photon transition) or from level $|e\rangle$ (two-photon transition). The atomic qubit encoding is $|+\rangle = 1/\sqrt{2}(|i\rangle + |g\rangle)$ and $|-\rangle = 1/\sqrt{2}(|i\rangle - |g\rangle)$. The photon number states of each cavity mode are denoted as $|n\rangle_i$, where $i = (a, b)$.

9.1 Description of the protocol

The sequence of operations achieving the UCQM transformation is depicted in Fig. 9.1. It presents, in a space-time diagram, the space lines of the two cavity modes and of the four Rydberg atoms, A_{1-4}, involved in the process. The atom-cavity resonant interactions are represented by black lozenges. Classical microwave pulses mixing the atomic levels are represented as gray circles.

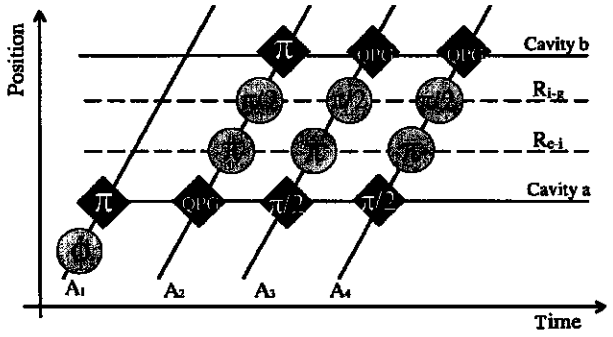

Figure 9.1: Detailed scheme of the atom field interactions in each cavity. A_1 enters the first cavity in state $\sqrt{1/3}\,|e\rangle_1 + \sqrt{2/3}\,|g\rangle_1$, prepared via a classical pulse in the Ramsey zone R_{e-i} and transfers its state to the cavity field in a π Rabi pulse. The cavity performs thus a QPG in A_2, the atom carrying the state to be cloned $\alpha\,|+\rangle_2 + \beta\,|-\rangle_2$. After being manipulated by microwave classical fields in the Ramsey zones R_{e-i} and R_{i-g}, A_2 also transfers its state to the second cavity, which is now entangled to the first one. A third atom A_3, prepared in state $|g\rangle_3$, crosses the first cavity performing a $\pi/2$ Rabi pulse. The first cavity field's state is completely recovered by the atomic states via the passage of a fourth atom, A_4, also prepared in $|g\rangle_4$, which makes a π Rabi pulse. Both of these atoms will interact resonantly with the second cavity field after classical microwave pulses manipulation in R_{e-i} and R_{i-g}. They will perform a resonant QPG, which will leave the total combined atom+second cavity field in the desired final state, corresponding to the cloning transformation.

The cavity fields are initially prepared in the vacuum state $|0\rangle_i$ [RBH01a]. The first atom, A_1, initially in state $|g\rangle_1$, is prepared in state

$$|\Psi\rangle_1 = \sqrt{\frac{2}{3}}\,|g\rangle_1 + \sqrt{\frac{1}{3}}\,|e\rangle_1\,, \tag{9.2}$$

by a classical pulse resonant with the $|g\rangle \rightarrow |e\rangle$ transition. The coefficients in the equation above are set by adjusting the duration of the classical pulse to a $\phi = \arcsin(\sqrt{1/3})$ rotation (see Fig. 9.1). The production of (9.2) can be checked in auxiliary experiments measuring the population of states $|g\rangle_1$ and $|e\rangle_1$ and the quantum coherence. The atomic state (9.2) is then transferred to C_a, through a π pulse of resonant quantum Rabi oscillation [MHN+97a, BSKM+96a]. Atom A_1 finally leaves C_a in state $|g\rangle_1$ and the cavity field is left in state $|\psi\rangle_a = \sqrt{2/3}\,|0\rangle_a + \sqrt{1/3}\,|1\rangle_a$. The first atom's final state being factored out, it will no longer be considered here.

Atom A_2, carrying the state to be cloned, crosses then C_a. It is prepared in the arbitrary state

$$|\Psi\rangle_2 = \alpha\,|+\rangle_2 + \beta\,|-\rangle_2\,, \tag{9.3}$$

where α and β are complex coefficients. Note that the preparation of this state, which is not part of the quantum cloning process, is not represented in Fig. 9.1. In an actual experiment, an additional microwave pulse acting on A_2 can be used to prepare this state. This atom interacts with the cavity field, performing a 2π quantum Rabi pulse. The purpose of this interaction is to realize a Quantum Phase Gate (QPG) as described in [RNO+99a]: it produces a π phase shift of the atom-cavity quantum state if and only if the atom is in state $|g\rangle_2$ and the cavity in state $|1\rangle_a$. As it is well know, a QPG, when expressed in the conjugate basis of the target $\{|+\rangle, |-\rangle\}$, amounts to a controlled not gate (CNOT), where the control qubit is the field state. After this interaction the total entangled atom-field state thus becomes

$$\sqrt{\frac{2}{3}}(\alpha\,|+\rangle_2 + \beta\,|-\rangle_2)\,|0\rangle_a + \sqrt{\frac{1}{3}}(\alpha\,|-\rangle_2 + \beta\,|+\rangle_2)\,|1\rangle_a\ . \tag{9.4}$$

We then send a third atom, A_3, prepared in state $|g\rangle_3$. It interacts resonantly with C_a, for a time interval corresponding to a $\pi/2$ quantum Rabi pulse, producing the state

$$\sqrt{\frac{2}{3}}(\alpha\,|+\rangle_2 + \beta\,|-\rangle_2)\,|g\rangle_3\,|0\rangle_a + \sqrt{\frac{1}{6}}(\alpha\,|-\rangle_2 + \beta\,|+\rangle_2)(|g\rangle_3\,|1\rangle_a + |e\rangle_3\,|0\rangle_a). \tag{9.5}$$

The state of C_a is finally transferred to a fourth atom, A_4, initially in $|g\rangle_4$ via a resonant π quantum Rabi pulse, creating the three-atom entangled state:

$$\sqrt{\frac{2}{3}}(\alpha\,|+\rangle_2 + \beta\,|-\rangle_2)\,|g\rangle_3\,|g\rangle_4 + \sqrt{\frac{1}{6}}(\alpha\,|-\rangle_2 + \beta\,|+\rangle_2)(|g\rangle_3\,|e\rangle_4 + |e\rangle_3\,|g\rangle_4), \tag{9.6}$$

and leaving C_a in the vacuum state, which factors out.

Classical microwave pulses then address the three atoms in the Ramsey zones, depicted in Fig. 9.1 as R_{e-i} and R_{g-i}. In R_{e-i}, $|e\rangle$ is transformed into $|i\rangle$ via a two-photon

π pulse. This does not affect state $|g\rangle$. Then, another classical $\pi/2$ pulse is applied in R_{g-i} on the three atoms, combining states $|g\rangle$ and $|i\rangle$. The sequence of transformations produced by these classical pulses can be summarized as follows:

$$\begin{aligned} |g\rangle &\longleftrightarrow |g\rangle \longleftrightarrow \sqrt{\tfrac{1}{2}}\left(|i\rangle - |g\rangle\right) = |-\rangle\,, \\ |e\rangle &\longleftrightarrow |i\rangle \longleftrightarrow \sqrt{\tfrac{1}{2}}\left(|i\rangle + |g\rangle\right) = |+\rangle\,. \end{aligned} \tag{9.7}$$

The remaining part of the protocol involves the second cavity, C_b. Atom A_2 interacts resonantly with C_b for a time interval corresponding to a π quantum Rabi pulse, transferring its state to the field mode. The final state of A_2 is $|g\rangle_2$ and also factorizes out. The total state of A_3, A_4 and C_b is then :

$$\sqrt{\frac{2}{3}}\left(\alpha\,|1\rangle_b + \beta\,|0\rangle_b\right)|-\rangle_3\,|-\rangle_4 + \sqrt{\frac{1}{6}}\left(\alpha\,|0\rangle_b + \beta\,|1\rangle_b\right)\left(|-\rangle_3\,|+\rangle_4 + |+\rangle_3\,|-\rangle_4\right). \tag{9.8}$$

Atoms A_3 and A_4 interact then independently and successively with C_b. They perform a resonant QPG, corresponding to a CNOT in the $\{|+\rangle\,,|-\rangle\}$ basis. The final state of these three systems writes, after a simple rearrangement of terms:

$$\begin{aligned} & \alpha\left[\sqrt{\tfrac{2}{3}}\,|+\rangle_3\,|+\rangle_4\,|\mathcal{A}\rangle + \sqrt{\tfrac{1}{3}}|\Phi\rangle_{3,4}\,|\mathcal{A}_\perp\rangle\right] \\ + \;& \beta\left[\sqrt{\tfrac{2}{3}}\,|-\rangle_3\,|-\rangle_4\,|\mathcal{A}_\perp\rangle + \sqrt{\tfrac{1}{3}}|\Phi\rangle_{3,4}\,|\mathcal{A}\rangle\right], \end{aligned} \tag{9.9}$$

where $|\mathcal{A}\rangle = |g\rangle_1\,|g\rangle_2\,|0\rangle_a\,|0\rangle_b$, $|\mathcal{A}_\perp\rangle = |g\rangle_1\,|g\rangle_2\,|0\rangle_a\,|1\rangle_b$ and $|\Phi\rangle_{3,4} = (|+\rangle_3\,|-\rangle_4 + |-\rangle_3\,|+\rangle_4)/\sqrt{2}$. In Eq. (9.9), the second cavity field ensures the orthogonality of $|\mathcal{A}\rangle$ and $|\mathcal{A}_\perp\rangle$ and hence is the important qubit in the ancilla's final state. This achieves the implementation of the optimal $1 \rightarrow 2$ cloning process.

Eq. (9.9) shows that this sequence actually implements the UCQM transformation given by Eq. (9.1). In this proposal, the blank state $|\mathcal{B}\rangle$ corresponds to the initial state of atoms A_1, A_3 and A_4 and of both cavity fields:

$$|\mathcal{B}\rangle = \left(\sqrt{\frac{1}{3}}\,|e\rangle_1 + \sqrt{\frac{2}{3}}\,|g\rangle_1\right)|g\rangle_3\,|g\rangle_4\,|0\rangle_a\,|0\rangle_b\,. \tag{9.10}$$

9.2 Discussion

We can now discuss the feasibility of an experimental implementation of the UQCM in a CQED system. The basic operations (quantum gates and classical field pulses) involved in the scheme have already been thoughtfully tested [RBH01a]. Their implementation thus does not present any major difficulty. The availability of an experimental configuration with two cavities can be considered as natural development of the present configurations where only one cavity is present. Note also some other interesting proposals require at least a two-cavity system [DZB+94a].

Atoms interact with C_a or C_b for a time corresponding at most to a 2π quantum Rabi pulse. The single photon Rabi frequency [MHN+97a, BSKM+96a] being

$\Omega/2\pi = 50$ kHz, the atomic velocity should be ≈ 500 m/s, in the range used in present experiments. Cavity and atomic relaxation are of course important issues. The circular Rydberg atoms lifetime is much longer than the protocol duration and is not bound to be a limiting factor. The main cause of decoherence in the present set-up is the cavity mode relaxation. The quantum information is stored in C_a only during the time interval between the passage of A_1 and A_4. Each atom may enter the cavity immediately after the preceding one has left it. The total quantum information storage time is of the order of four full atomic transit times, i.e. $\approx 2.10^{-4}$ s. This is shorter than present cavity damping times (about 1 ms). The cavity C_b stores quantum information for an even shorter time interval. Note finally that the atomic transit time between the two cavities does not matter to evaluate the influence of damping, since the quantum information is then carried by long-lived atomic systems.

An alternative implementation of our UCQM scheme uses two modes of a single cavity. In the present experimental set-up, the cavity sustains two gaussian modes, M_a and M_b, with orthogonal linear polarizations. Due to mirrors imperfections, these two modes have slightly different resonant frequencies (splitting 130 kHz). Since this splitting is much larger than the atom-field coupling Ω, the atoms resonantly interact with one mode only at same time. Stark tuning can be used to tailor atomic interactions with the two modes during the atomic transit time through the cavity.

In this scheme, A_1 leaves its state in M_a. Then, A_2 performs the CNOT operation in M_a. It is set off-resonance with both modes for a short time interval during which the microwave classical pulses are applied. Atom A_2 is then tuned to resonance with M_b for its final quantum Rabi pulse. A_3 and A_4 interact first resonantly with M_a, undergo the classical pulses while being off-resonance from the two modes and finally interact with M_b as described above. This implementation of the UCQM, requiring a single cavity, would be much simpler to realize. Each atom should interact with the cavity for a total time corresponding at most to a 3π quantum Rabi pulse (the duration of the classical pulses is negligible). The atomic velocity should be about 330 m/s, still well within the available range. The quantum information is stored in the cavity modes for a slightly longer time than in the two-cavity arrangement (four times the full transit time at the slower 330 m/s velocity). Cavity damping should thus be somewhat smaller.

The UQCM operation verification can, in principle, be performed by the usual detection techniques [RBH01a]. As mentioned above, the fidelity is, ideally, 5/6 while the trivial production of a maximally mixed state gives an average fidelity of 2/3. This means that the fidelity should be measured with a precision greater than $\approx 92\%$ $(= 1 - \frac{1}{2}(5/6 - 2/3))$. Note that, in the NMR quantum cloning experiment [CJF+02a], this degree of precision was not reached, so that the improvement due to the cloning process could not be verified. In our proposal, all the elementary operations, quantum Rabi or classical field pulses, are prone to errors. The total number of these operations is sixteen if we take into account the detection and preparation process. The necessary precision could only be reached if each pulse has a fidelity greater than $\sqrt[16]{0.92}$. This value, being about 0.995 is still out of the experimental reach (present pulse imperfections are between 3 and 10%). This figure, however, sets an interesting goal to be

reached. We can also think of measuring the quality of our copy using the Ramsey interferometer technique. It consists in applying two $\pi/2$ classical pulses in the produced state which differ by a phase ϕ, making the transformations $|+\rangle \rightarrow 1/\sqrt{2}(|+\rangle + e^{i\phi}|-\rangle)$ and $|-\rangle \rightarrow 1/\sqrt{2}(-e^{-i\phi}|+\rangle + |-\rangle)$. The measurement of the probability of finding atoms in states $|+\rangle$ or $|-\rangle$ has a sinusoidal dependence on the phase ϕ. If the atom is in maximally mixed state, the contrast of these fringes is zero, and it reaches the maximal value of one for atoms in a pure state. For an atom in a state prepared by the cloning transformation, the ideal value of this contrast is 2/3. We should naturally expect a decay of this value given all the experimental imperfections. Considering the imperfections in the interation times of the order of 3%, the expected value of the fringe contrast is of the order of 0.42. This value shows that the Ramsey fringes contrast allows one to make a better distinction between the cloning proccess and the maximally mixed state for a more realistic experimental situation.

9.3 Conclusion

We described a protocol implementing the universal optimal copying transformation in CQED. Basic quantum information operations have already been implemented in the cavity QED context [RNO+99a], and proposals that could extend these experimental realizations to more elaborated quantum information algorithms [YMB+02a, SZ02a] are naturally appealing.

The quantum logics network used in our scheme is simpler than previous ones by making use of auxiliary degrees of freedom which are discarded in the end of the process. Note also that the same protocol can be applied to the cloning of equatorial qubits [BDE+98a, BCDM00a], i.e. $|\psi\rangle = \sqrt{1/2}\left(|0\rangle + e^{i\phi}|1\rangle\right)$ by sending A_1 in state $\sqrt{1/2}(|e\rangle + |g\rangle)$.

10 Proposal for realization of a Toffoli gate via cavity-assisted atomic collision

THE RECENT DEVELOPMENT of quantum information processing has shed new light on complexity and communication theory. It is now widely accepted that some problems are solved more efficiently by quantum computers than by their classical analogs [Sho94a, Gro96a]. This matter of fact has triggered in the past years a lot of studies on theoretical and practical aspects of quantum computation. In particular, finding universal sets of gates for processing quantum information is of paramount importance: the possibility of implementing universal gates is a key requirement for building interesting quantum information processing devices. Consequently, many different physical systems have been studied for their quantum information processing capabilities (for a review see [NC00a]). Some important results concerning the implementation of gates can be found for example in [WSM01a, EJPP00a, PSD+99a], while more general questions concerning broad classes of interaction Hamiltonians (such as their universality for quantum computation) are addressed for example in [BHLS02a, BKD+01a].

One of the next challenges that need to be overcome is to design robust implementations of these sets of gates. Several constructions have already been proposed and realized for various experimental settings [THL+95a, RNO+99a, MMK+95a, GC97a, JMH98a]. These results currently serve as benchmarks for other implementations and do provide deep insights into the ability of the chosen devices to manipulate quantum systems efficiently. In this article, we describe a protocol for realizing the Toffoli gate—a three qubits "control-control-NOT"—in the Cavity QED context (CQED). This gate, together with one qubit rotations, form a universal set for quantum computing [Deu89a]. We will see that the combination of techniques used makes our protocol optimal in term of number of interactions within the constraints of our experimental setting. It is thus less resources demanding than the standard implementation as a sequence of control-NOT gates [DS94a]. Finally we estimate the performance of our protocol by taking into account imprecisions as well as decoherence effects during the protocol.

10.1 Cavity-QED toolbox

In this paragraph we review briefly the different techniques used for implementing the Toffoli gate. Further experimental details concerning our particular setting are exposed with great care in [RBH01a]. In our protocol, qubits, ie physical systems with a two-dimensional Hilbert space available for information processing, will be stored in circular

Rydberg states of highly excited Rubidium atoms and in the cavity field. The very long lifetime of these atoms together with the transition frequencies of the chosen transitions allow to deal with those complex atoms as if they had only three levels noted $|g\rangle$, $|e\rangle$ and $|i\rangle$. Transformations between those three states can be driven coherently by a classical microwave pulse adjusted nearly resonantly to the proper transition frequency. For example, if the classical field is tuned with respect to $|i\rangle \leftrightarrow |g\rangle$, a π pulse will transform any pure quantum state of the form $\alpha|e\rangle + \beta|g\rangle$ into $\alpha|e\rangle + \beta|i\rangle$. Indeed, those transformations are the analogs of the well known one qubit gates for our three level systems. In order to process quantum information in a non-trivial way, we also need an equivalent for the two-qubit gates. Within the CQED context, this involves a coupling between the atoms and a high-Q Niobium superconducting microwave cavity. The frequency of the mode inside the cavity can be adjusted in and out of resonance with the $|g\rangle \leftrightarrow |e\rangle$ transition. Thus, we can consider two distinct approaches for processing quantum information. The first one, which has been throughly used in recent experiments [NRO+99a], relies on a resonant interaction with a single mode of the cavity. As an example, for a well chosen interaction time, an initial atom-field state of the form $(\alpha|g\rangle + \beta|e\rangle)|0\rangle$ will evolve into $|g\rangle(\alpha|0\rangle + \beta|1\rangle)$. This transformation is called π-Rabi rotation. Similarly, continuing the interaction for an equal amount of time leads to the state $(\alpha|g\rangle - \beta|e\rangle)|0\rangle$. In the perspective of information processing, the first half of the transformation corresponds to transferring quantum information from the atom into the cavity, while the second half can be applied to a second atom which would thus retrieve the initial quantum information: the field inside the cavity acts as a temporary quantum memory. The second approach for building atom field interactions, referred to as cavity assisted atomic collision, has been proposed [ZG00a, Zhe01a] and experimentally tested [OBA+01a] more recently. In this setting, two atoms enter the cavity at the same time and follow an evolution conditioned upon the state of the cavity field. More precisely, the cavity is detuned from the $|e\rangle \leftrightarrow |g\rangle$ transition. In this regime, where the detuning δ is much larger than the atom-field coupling constant Ω, the effective Hamiltonian can be derived using second order perturbation theory:

$$\begin{aligned} H_e \;=\; & \lambda\left(|e_1\rangle\langle e_1| aa^+ - |g_1\rangle\langle g_1| a^+a\right. \\ & + |e_2\rangle\langle e_2| aa^+ - |g_2\rangle\langle g_2| a^+a \\ & \left. + |e_1\rangle\langle g_1| \otimes |g_2\rangle\langle e_2| + |g_1\rangle\langle e_1| \otimes |e_2\rangle\langle g_2|\right), \end{aligned} \tag{10.1}$$

where $\lambda = \Omega^2/4\delta$, and where the operators a and a^+ are the annihilation and creation operators for the cavity field. In the above formula, we see that any exchange of energy between the field and the atoms present in the cavity is forbidden, but this still allows conditional dynamics upon the photon number in the cavity mode. This interaction will be our main tool to perform the Toffoli gate in the CQED setting.

10.2 Description of the protocol

Before going into the details of the protocol, recall that the Toffoli gate is a three qubits gate. Its action on the computational basis of the three qubits is to perform a "control-control-NOT". Thus, only two basis vectors are affected by this evolution:

$$\begin{aligned} |1\rangle\,|1\rangle\,|0\rangle &\rightarrow |1\rangle\,|1\rangle\,|1\rangle \\ |1\rangle\,|1\rangle\,|1\rangle &\rightarrow |1\rangle\,|1\rangle\,|0\rangle \end{aligned} \tag{10.2}$$

where the control qubits are the first two ones, and the target is the last one. Taking into account the specific CQED setting, we chose an implementation with one cavity mode and two atoms. The photon number states of the cavity mode will be denoted $|n_c\rangle$, while the energy levels of the atoms will be written as $|i_{c,t}\rangle$, $|g_{c,t}\rangle$, $|e_{c,t}\rangle$ (the subscripts c and t are short hands for control and target). The relation between those states and the computational basis of the Toffoli gate is summarized below (obvious normalization factors have been omitted):

Computational basis	Control 1 (cavity mode)	Control 2 (Rb atom 1)	Target (Rb atom 2)
$\|0\rangle$	$\|1_c\rangle$	$\|i_c\rangle$	$\|g_t\rangle + \|e_t\rangle$
$\|1\rangle$	$\|0_c\rangle$	$\|g_c\rangle$	$\|g_t\rangle - \|e_t\rangle$

(10.3)

For sake of simplicity we will concentrate on the realization of the gate assuming that the preparation of the cavity has been already performed.

Since all our manipulations are done coherently, it is sufficient to describe the quantum evolution for basis vectors of the whole system (ie, our three qubits). The protocol can be decomposed into three distinct phases: encoding, atomic collision, and decoding. The encoding and decoding operations only involve resonant interactions between the cavity mode (control 1) and a single Rb atom (control 2). Decoding is performed by applying the encoding evolutions in reverse order. The heart of the evolution is the cavity assisted atomic collision which realizes the equivalent of the Toffoli gate on the encoded quantum information. More precisely, it realizes a control phase gate between the encoded information of the two control qubits and the target qubit. Thus, the first atom (control 2) is sent at a relatively low speed into the cavity. After the encoding, it collides with the faster moving second atom (target). When the second atom has left the cavity, the first one interacts with the cavity mode to accomplish the decoding step. Figure 10.1 summarizes the different steps required to perform the gate. Each of them will be described in detail in the following paragraphs.

The encoding starts when the atom identified as Control 2 (A_c) is sent into the cavity. First, it interacts resonantly with the cavity field and undergoes a π-Rabi rotation. In terms of the basis vectors given in Eq. (10.3), only one of them is affected by the evolution:

$$|1_c\rangle\,|g_c\rangle \rightarrow -\,|0_c\rangle\,|e_c\rangle\,. \tag{10.4}$$

At this point, for a generic input state of the Toffoli gate, all three levels of the control atom can be populated: the overall state of the cavity and Rb atom 1 cannot be

represented by a two-qubit state. However, coherent manipulation of the quantum information remains possible through the cavity assisted collision. The main purpose of the decoding operation is precisely to restore the two-qubit structure of the cavity mode and Rb atom 1. The state obtained in Eq. (10.4) almost corresponds to the needed preparation of the cavity and control atom before the atomic collision: a microwave pulse tuned to the $|g_c\rangle \leftrightarrow |i_c\rangle$ transition, and corresponding to a basis rotation exchanging $|i_c\rangle$ and $|g_c\rangle$, completes this first step of the protocol.

The main part of the protocol can now be achieved: the cavity assisted atomic collision. Thus, the cavity detuning δ is set such that the interaction Hamiltonian is given by Eq. (10.2). The target atom (A_t) then enters the cavity and interacts with the field and the control atom. The evolution of the basis states is easily computed by diagonalizing the effective Hamiltonian. After an interaction time $t_{\text{col}} = \pi/\lambda$, all states remain unaffected except the following ones:

$$\begin{aligned} |0_c\rangle |i_c\rangle (|g_t\rangle + |e_t\rangle) &\rightarrow |0_c\rangle |i_c\rangle (|g_t\rangle - |e_t\rangle) \\ |0_c\rangle |i_c\rangle (|g_t\rangle - |e_t\rangle) &\rightarrow |0_c\rangle |i_c\rangle (|g_t\rangle + |e_t\rangle). \end{aligned} \tag{10.5}$$

All operations done before the atomic collision can be considered as the preparation of the quantum systems such that they undergo the proper overall evolution described by the Hamiltonian of Eq. (10.5). Hence, to complete the whole protocol, we only need to undo all the preparatory steps: after A_t left the cavity (recall A_t is the fast moving atom), we apply the resonant pulse on A_c exchanging $|i_c\rangle$ and $|g_c\rangle$, followed by a π-Rabi rotation to extricate from A_c the information initially contained in the cavity.

This completes the overall protocol and leads to:

$$\begin{aligned} |1_c\rangle |i_c\rangle (|g_t\rangle + |e_t\rangle) &\rightarrow |1_c\rangle |i_c\rangle (|g_t\rangle + |e_t\rangle) \\ |1_c\rangle |i_c\rangle (|g_t\rangle - |e_t\rangle) &\rightarrow |1_c\rangle |i_c\rangle (|g_t\rangle - |e_t\rangle) \\ |1_c\rangle |g_c\rangle (|g_t\rangle + |e_t\rangle) &\rightarrow |1_c\rangle |g_c\rangle (|g_t\rangle + |e_t\rangle) \\ |1_c\rangle |g_c\rangle (|g_t\rangle - |e_t\rangle) &\rightarrow |1_c\rangle |g_c\rangle (|g_t\rangle - |e_t\rangle) \\ |0_c\rangle |i_c\rangle (|g_t\rangle + |e_t\rangle) &\rightarrow |0_c\rangle |i_c\rangle (|g_t\rangle + |e_t\rangle) \\ |0_c\rangle |i_c\rangle (|g_t\rangle - |e_t\rangle) &\rightarrow |0_c\rangle |i_c\rangle (|g_t\rangle - |e_t\rangle) \\ |0_c\rangle |g_c\rangle (|g_t\rangle + |e_t\rangle) &\rightarrow |0_c\rangle |g_c\rangle (|g_t\rangle - |e_t\rangle) \\ |0_c\rangle |g_c\rangle (|g_t\rangle - |e_t\rangle) &\rightarrow |0_c\rangle |g_c\rangle (|g_t\rangle + |e_t\rangle), \end{aligned} \tag{10.6}$$

which, in turn, exactly corresponds to the Toffoli gate in the computational basis of Eq. (10.3). The pulses and interactions sequence are summarized in Fig. 10.1.

10.3 Discussion

The scheme presented here proposes to use the atomic collision as the key interaction in the making of the Toffoli gate. The most fundamental reason that lead to this choice is that using only resonant interactions would not allow us to perform the gate in this very specific one mode CQED setting: it has been shown [RNO+99a], that resonant interactions can be used to design CNOT gates and hence lead to universal quantum

computation. However, implementating the Toffoli gate with CNOT gates together with one qubit gates, as presented in [DS94a], requires to address the qubits separately between each CNOT gate. Our particular experimental single mode cavity scheme does not allow such addressing, as the differents atoms are inditinguishable when they are inside the cavity. Hence, to implement the circuits of [DS94a], we would need to have more than one mode.

This fact also leads to the proof of optimality of the proposed scheme: the impossibility of manipulating the atoms individually inside the cavity imposes that atoms can only interact once with each other. This is far from enough for realizing a Toffoli gate, if these interactions are restricted to be pairwise. Therefore, at least a three quantum system interaction is required. This can be accomplished through the cavity assisted atomic collision. Then, only two options have to be considered: either a three Rb atom interaction with an empty cavity or a collision of only two Rb atoms inside a cavity containing a quantized field in a non-vacuum state. The first case is ruled out easily by noting that there do not exist a set of 8 orthogonal vectors that have a periodic evolution under the Hamiltonian of Eq. (10.2) and that would realize the phase gate associated with the Toffoli gate (ie the gate that does nothing except for the following state $|1\rangle|1\rangle|1\rangle \rightarrow -|1\rangle|1\rangle|1\rangle$). Since this gate is equivalent to the Toffoli gate up to local unitary operations, we can conclude that even with an appropriate encoding and decoding phase, this interaction will not lead to the desired evolution. On the other hand, the second option is more successful and is the one presented in this paper. Sets of 8 orthogonal vectors with the proper evolution can be found, but since none of them can be written as a tensor product of single qubit sets, encoding and decoding involving at least a two qubit interaction are required. This last remark thus proves the optimality of our protocol.

Let us now discuss the practical feasibility of this proposal. The atomic collision can only happen when the cavity is detuned from the $|e\rangle \leftrightarrow |g\rangle$ transition by an amount $\delta \gg \Omega$, where the cavity-field coupling constant $\Omega/2\pi = 50\,\mathrm{kHz}$ [RBH01a]. This can be done at any time and in the presence of Rb atoms inside the cavity by applying an external electric field to the mirrors. The time needed to change the cavity detuning is sufficiently small compared to the interaction times so that the change in Hamiltonians can be considered instantaneous. Here we choose $\delta/2\pi = 4\Omega$, which gives an interaction Hamiltonian well approximated by Eq (10.2) with $\lambda = \Omega^2/4\delta$. With these figures, the interaction time to realize a π-Rabi rotation is $2\,10^{-5}$s and for the atomic collision $1.25\,10^{-4}$s. The time required to perform the interaction with the classical microwave cavity is negligible. This imposes a velocity of the atoms of order $50\mathrm{m.s}^{-1}$. This value can be reached by means of simple atomic beam techniques with transverse laser cooling. The total interaction time with the cavity field is of order 0.18ms, still much smaller than the cavity lifetime (1ms). Decoherence effects due to the loss of a photon should then be relatively small. We present in Fig. 10.2 the result of a numerical simulation to estimate the fidelity of the gate. Dissipation effects during the realization of the gate have been accounted for by letting a photon escape from the cavity with a Poisson probability law according to various cavity lifetime. Other imperfections, like for instance

imprecision in the velocity of the atoms or misalignment of the atomic beam, can all be translated into an imprecision in the interaction time with all other parameters set at their nominal value. This defines the effective interaction time t^i_{eff} for the i-th interaction. The imperfection strength $\epsilon = \Delta t^i_{\text{eff}}/t^i_{\text{eff}}$ has been taken equal for all interactions, and the average fidelity has been computed by accumulating results for different realizations of the gate. The currently achievable precision (around 3%) and cavity lifetime (1 ms) yield an estimated fidelity above 70%. Thus, the practical realization of the Toffoli gate through this protocol is not out of reach and could set an interesting benchmark for comparing the efficiency of quantum information processing approaches using CQED. Moreover, we can see that even small improvements on either the precision of the pulses or the cavity lifetime result in achieving a fidelity of nearly 90%, thus making this realization of the gate very attractive for analyzing its experimental behavior.

We now return to the preparation of the cavity field, before the above protocol takes place. This can be done in full generality by sending a Rydberg atom (A_p) containing the desired quantum information, and by transferring its state to the cavity through a π-Rabi rotation. The protocol is then started when this ancillary atom leaves the cavity. The retrieval is accomplished by transferring the state of the cavity back into an atom (A_r) initially in the $|g\rangle$ state. Thus, the result of the Toffoli gate is contained in the state of the atoms A_c, A_t and A_r. This state, and hence the behavior of the gate, can be easily analyzed using standard quantum tomography techniques for cavity QED, ie the observation of resonant Rabi oscillations: the analysis of the gate requires only single atom rotations and accumulation of statistics.

10.4 Conclusion

We have presented a realistic scheme for implementing the Toffoli gate with currently available CQED techniques. We have shown that using a cavity assisted collision instead of only resonant interactions makes our scheme optimal given the restrictions of the experimental setting. It thus enlarges the range of possible applications of CQED to process information for realizing basic logical operations [RNO+99a, NRO+99a, OBA+01a] as well as more complicated protocols [YMB+02a, MOR03a, ZPM03a, SZ02a]. The estimated achievable fidelity (around 70% for current imprecision and decoherence levels) ensures that the behavior of the gate can be tested experimentally. The corresponding experimental results could be compared to analogous quantities for other sets of universal gates and thus provide a deeper insight into the quantum information processing capability of the CQED setting with non resonant interactions.

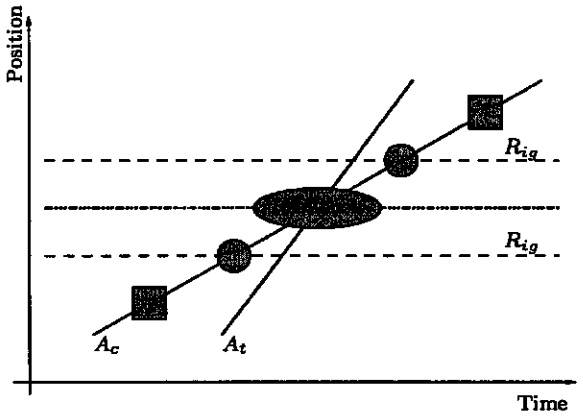

Figure 10.1: Detailed scheme of the atom-field and atom-atom interactions in the cavity. Circles are interactions with a classical microwave field tuned to the $|i\rangle \leftrightarrow |g\rangle$ transition. This field is generated in two Ramsey zones denoted R_{ig}. Squares symbolize the resonant interaction with the quantum field stored into the high-Q cavity. The duration time of each interaction is set either by controlling the length of the pulse or by applying a voltage to the mirrors of the cavity to change the atom-field coupling constant Ω.

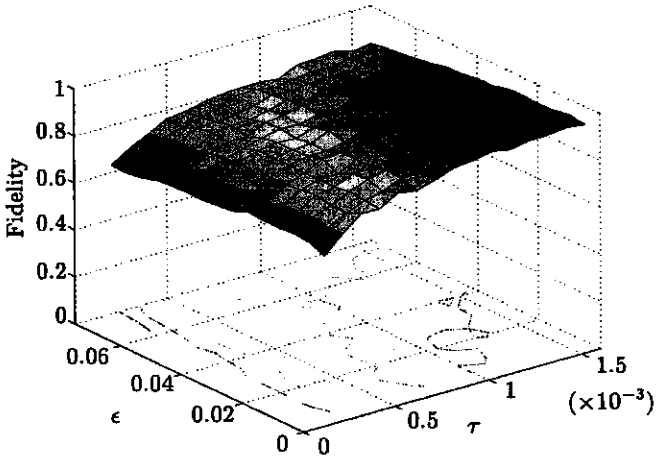

Figure 10.2: Fidelity of the Toffoli gate as a function of the photon lifetime, τ, and of the uncertainty on the effective interaction times, ϵ. For current day values, $\tau = 1$ms and $\epsilon = 3\%$ a fidelity of 0.7 is expected.

Appendix

Appendix Résumé

LA DÉCOUVERTE DE la mécanique quantique au début du XX$^{\text{ème}}$ siècle a profondément marqué le développement de la physique moderne. Moins connue du grand public que la relativité générale, c'est à elle que l'on doit la révolution du traitement de l'information grâce à l'emploi des structures de bandes du silicium, les lasers, l'imagerie médicale par résonance magnétique nucléaire, etc.

Depuis quelques années l'intérêt pour cette théorie du microscopique connaît un renouveau sans précédent. D'une part, un premier mouvement consacre la convergence de l'informatique et de la physique: l'informatique quantique. La "pose de la première pierre" de cette discipline est attribuée à R. P. Feynman [Fey82a, Fey84a] au début des années 80. On pensait alors que les machines quantiques seraient principalement dévolues à la simulation de systèmes physiques complexes. Dans les faits, les travaux accomplis jusqu'à maintenant mettent en avant les capacités inédites de ces machines à résoudre efficacement certains problèmes mathématiques réputés difficiles. On peut par exemple citer la factorisation des grands nombres entiers [Sho95a], ou encore la recherche dans des bases de données non-ordonnées. Il faut également noter que cette convergence s'est opérée au travers de la découverte d'un protocole inconditionnellement sûr pour l'échange de clefs cryptographiques [BB84a, SP00a]. Cette dernière avancée a eu pour effet de sceller définitivement le sort de la mécanique quantique à celui de l'information et de participer à la création d'une théorie de l'information quantique.

D'autre part, depuis le début des années 80, une nouvelle approche du phénomène de décohérence — ce phénomène expliquant la coexistence d'un monde microscopique quantique et d'un monde macroscopique classique — semblait vouloir se dessiner [Zur81a, Zur82a]. Elle promettait une résolution élégante de certains problèmes liés aux fondations de la mécanique quantique, problèmes qui avaient résisté à l'analyse depuis l'invention de la théorie. En particulier, elle exhibe un mécanisme physique empêchant de conduire dans la pratique la fameuse expérience du chat de Schrödinger. Mais ce ne fut pas là son seul mérite. En effet, parallèlement aux développements théoriques de l'informatique quantique, de nombreux groupes ont essayé de mettre en pratique la manipulation quantique de l'information. Bien que séduits par l'idée d'effectuer des calculs à l'échelle atomique, ils ont très rapidement constaté les difficultés de ce programme de recherche. Ils se heurtèrent à un obstacle qui paraissait insurmontable: la décohérence. Ce processus survient malheureusement de façon presque inévitable lorsque le registre de calcul n'est pas parfaitement isolé de son environnement. Heureusement, l'approche promue en ce début des années 80 a permis d'analyser rigoureusement et quantitativement ce phénomène, mais surtout a permis de montrer qu'il était concevable d'en combattre les effets au travers de stratégies de correction d'erreurs.

Finalement, on peut aussi noter que l'information quantique s'est infiltrée dans de nombreux domaines de la physique du microscopique, aidant à la formalisation de notions connues, mais également permettant la résolution élégante de problèmes expérimentaux [SV98a, BGL+03a].

Le travail présenté ici reflète dans sa structure et dans les sujets qu'il aborde une partie de ces évolutions. La première partie est consacrée aux problèmes d'interprétation de la mécanique quantique, la seconde à la correction d'erreurs quantiques et enfin la troisième à quelques propositions expérimentales visant à démontrer nos capacités à traiter l'information de façon quantique.

La transition quantique-classique : une approche via la théorie de l'information

Depuis son invention dans les années 1920, la mécanique quantique a toujours suscité de nombreuses questions quant à son interprétation. En effet, malgré des prédictions sans cesse confirmées, la mécanique quantique a une place un peu à part dans l'ensemble des grandes théories de la physique.

D'une part les axiomes qui la définissent ne peuvent être exprimés qu'en employant des notions mathématiques abstraites et pour lesquelles l'intuition physique semble absente. D'autre part, et contrairement à la relativité générale qui est une extension manifeste de la mécanique Newtonienne, il semble difficile de définir la mécanique classique comme un cas limite de la théorie quantique.

Afin de comprendre les enjeux mais aussi les solutions qui ont été apportées durant les décennies précédentes il convient de faire quelques rappels de notions élémentaires de mécanique quantique.

Axiomes de la mécanique quantique *(Chapitre 1)*

Ce qui différencie une théorie mathématique et une théorie physique c'est que la dernière rend compte d'une certaine "réalité". Ce que l'on entend par là c'est que la construction axiomatique d'une théorie se doit de répondre à trois questions incontournables :

- quels sont les systèmes physiques étudiés et comment s'écrit leur état ;
- quelle est leur évolution ;
- quelles sont les quantités relatives à l'état du système qui peuvent être mesurées expérimentalement.

Les axiomes de la mécanique quantique apportent une réponse précise à chacun de ces points (cf. [CTDL77a, Prea, NC00a]) :

- l'état d'un système est un vecteur $|\psi\rangle$ de norme unité d'un espace de Hilbert complexe ;

- l'évolution des états est régie par l'équation de Schrödinger,

$$\frac{d}{dt}|\psi(t)\rangle = -iH(t)\,|\psi(t)\rangle\,, \tag{7}$$

où $H(t)$ est une matrice hermitienne ;

- les quantités observables sont représentées par des matrices Hermitiennes et leur valeur moyenne lorsque le système est dans l'état $|\psi\rangle$ est donnée par $\langle O\rangle_\psi = \langle\psi|\,O\,|\psi\rangle$. De façon équivalente, un résultat de mesure correspond à l'obtention d'une valeur propre, e.g. θ, de O associée au projecteur P_θ. La probabilité d'un tel résultat sachant que le système est dans l'état $|\psi\rangle$ est donnée par $\langle\psi|\,P_\theta\,|\psi\rangle$. L'acquisition d'information correspondante force à mettre à jour l'état du système. Après l'obtention du résultat θ, le système est dans l'état $P_\theta\,|\psi\rangle\,/\sqrt{\langle\psi|\,P_\theta\,|\psi\rangle}$.

Décohérence et superselection induite par l'environnement *(Chapitre 2)*

Le problème de la mesure

Tout d'abord on peut se demander ce que ces axiomes ont de si étrange pour que l'on soit encore, près de 80 ans après leur formulation, en train de débattre à leur sujet.

Premièrement, comme cela a été déjà mentionné, ils sont techniques et ne font pas référence à un principe physique fort. Pour prendre l'exemple de la relativité restreinte, un des axiomes est "la vitesse de la lumière est une constante indépendante du référentiel inertiel considéré", alors que la mécanique quantique se contente de "l'état d'un système est un vecteur de norme unité dans un espace de Hilbert complexe"! Dans le premier cas, on perçoit nettement le caractère fondamental d'une telle assertion, tandis que dans le second il ne semble s'agir que d'une construction ad-hoc permettant, certes, de prédire le résultat d'expériences mais qui appelle une autre théorie plus complète et mieux motivée physiquement.

Mais ce n'est pas là le souci principal qui préoccupait et préoccupe toujours une grande partie des physiciens théoriciens. D'une part, on peut constater que la mécanique quantique telle qu'elle est décrite plus haut est une théorie linéaire : ainsi, lorsque $|\psi(t)\rangle$ et $|\phi(t)\rangle$ sont deux solutions de l'équation de Schrödinger, alors $(|\psi(t)\rangle+|\phi(t)\rangle)/\sqrt{(2)}$ est également une solution. C'est ce que l'on appelle le principe de superposition. D'autre part, les objets macroscopiques régis par la mécanique classique, et bien que constitués de systèmes microscopiques obéissant à la mécanique quantique, ne satisfont en général pas à un tel principe de superposition.

L'expérience du chat de Schrödinger Schrödinger a mis en avant l'insoutenable contradiction entre physique quantique et physique classique grâce à sa fameuse expérience de pensée.

Pour cela, il considère une boîte parfaitement isolée et contenant un mécanisme un peu particulier. Il s'agit d'un atome commandant, en fonction de son état quantique, la libération d'un gaz toxique. Lorsque l'atome est dans son état excité, rien ne se passe,

tandis que le poison est libéré dès que l'atome est dans l'état fondamental. Si maintenant, on enferme un chat dans la boîte et que l'atome est préparé dans une superposition de l'état excité et de l'état fondamental, alors l'état du chat est impliqué dans une superposition macroscopique où un des termes correspond à un chat vivant et l'autre à un chat mort. Manifestement jamais une telle situation ne se produit et tous les fondateurs de la mécanique quantique en étaient bien conscients. Ce que Schrödinger soulignait de cette manière c'était la nécessité de répondre à la question de la transition quantique-classique : pourquoi et comment se fait le passage entre la mécanique quantique et la mécanique classique alors qu'aucune d'elles ne spécifie dans ses axiomes une limite à son champ d'application.

Le modèle de von Neumann Préoccupé par cette question von Neumann [vN32a] a mis en place un cadre générique permettant d'étudier rigoureusement et quantitativement cette question. Ce cadre est appelé "modèle de mesure". Il s'agit d'un système quantique $\mathcal{S}$ qui interagit avec un appareil de mesure $\mathcal{A}$ selon les lois de la mécanique quantique. Initialement, le système et l'appareil de mesure ne sont pas corrélés, c'est à dire que leur état global s'écrit $\left|\psi^{\mathcal{S}}\right\rangle \otimes \left|\psi^{\mathcal{A}}\right\rangle$. Au cours du temps, ils interagissent selon les lois de la mécanique quantique. Ils sont donc soumis à l'équation de Schrödinger avec un terme d'interaction décrit par le Hamiltonien H_{int}. Après un temps bien défini, ils se séparent et l'interaction prend fin. Leur état générique global correspond donc à un état couplé — plus précisément à un état enchevêtré.

Pour comprendre pourquoi un tel modèle rend compte de la question posée par Schrödinger, il suffit de considérer pour $\mathcal{S}$ un atome à deux niveaux, $|0\rangle$ et $|1\rangle$, ainsi qu'un appareil de mesure possédant deux positions possibles pour son aiguille, $|d\rangle$ et $|g\rangle$. Si l'on considère que le Hamiltonien d'interaction induit l'évolution suivante,

$$|0\rangle\,|d\rangle \quad \rightarrow |0\rangle\,|d\rangle \tag{8}$$
$$|1\rangle\,|d\rangle \quad \rightarrow |1\rangle\,|g\rangle\,, \tag{9}$$

alors pour un système préparé initialement dans l'état $(|0\rangle + |1\rangle)\sqrt{2}$, l'état final doit être $(|0\rangle\,|d\rangle + |1\rangle\,|g\rangle)\sqrt{2}$. Autrement dit, l'appareil de mesure doit être dans une superposition impliquant l'aiguille dans l'état $|d\rangle$ et l'aiguille dans l'état $|g\rangle$. Comme pour le chat de Schrödinger, de telles superpositions n'ont jamais été observées. Seul l'un des deux résultats est choisi et conduit à l'exclusion de l'autre. Cette situation qui semble en contradiction avec l'expérience est appelé problème de la mesure.

Interprétations

Devant l'absence d'une réponse apportée par la théorie elle même, les physiciens ont tenté de trouver des explications appropriées à cet état de fait. C'est ce que l'on appelle les interprétations de la mécanique quantique.

L'école de Copenhague La plus fameuse d'entre elle, l'interprétation de Copenhague, a été promue par N. Bohr et suspend par décret le principe de superposition. Il motive

cette solution par le fait que contrairement au domaine quantique, le domaine classique ne souffre pas de problèmes d'interprétation. Ainsi il en fait un élément de référence pour analyser les phénomènes quantiques. Plus précisément, il postule qu'une mesure n'est réalisée que lorsque son résultat est retranscrit dans le monde classique. Ainsi, dans le modèle de von Neumann présenté plus haut, la finalisation de la mesure ne se fait que lorsque $\mathcal{A}$ est mis en contact avec un système classique. Ce n'est qu'à ce moment là qu'un résultat particulier est choisi. Cependant, comme le soulignait Schrödinger, tout système classique peut être en principe décrit par la mécanique quantique. Ainsi en repoussant peu à peu les limites du domaine quantique, on est contraint de faire se produire la transition vers le monde classique dans la conscience des observateurs.

Toutefois, il faut reconnaître que l'interprétation de Copenhague décrit bien la réalité pratique : il n'est pas envisageable de maintenir un observateur vivant suffisamment isolé du reste du monde pour qu'il soit effectivement décrit par la mécanique quantique. Pour toutes les situations rencontrées dans les expériences de laboratoire, une mesure se termine bien par un contact avec le monde classique — i.e. la mesure est effective lorsqu'un nombre est inscrit sur une feuille de papier ou dans la mémoire d'un ordinateur. Dans le cas de l'interprétation de Copenhague, le problème de la transition quantique-classique se réduit donc à comprendre pourquoi certains objets peuvent être qualifiés de "classiques" et pourquoi ils sont alors à même de finaliser les mesures sur les systèmes quantiques.

Les univers multiples L'interprétation de Copenhague n'est cependant pas la seule possible. Parmi les candidates, celle des univers multiples d'Everett [Eve57a, Whe57a] a connu un regain d'intérêt depuis le début des années 1970. Tout comme l'interprétation de Copenhague, elle ne permet pas vraiment de répondre au problème de la mesure, mais elle a le mérite de mettre l'accent sur un autre aspect de la transition quantique-classique : l'invariance de base [Zeh73a]. En effet, dans sa formulation, la mécanique quantique ne fait pas référence à des états particuliers. Ainsi, si l'on reprend l'exemple d'un système à deux niveaux $|0\rangle$ et $|1\rangle$, l'état $|+\rangle = (|0\rangle + |1\rangle)\sqrt{2}$ est tout autant une superposition que l'état $|0\rangle = (|+\rangle + |-\rangle)\sqrt{2}$, à condition de définir $|-\rangle = (|0\rangle - |1\rangle)\sqrt{2}$. Si on pense que la mécanique quantique doit permettre l'émergence de la mécanique classique, alors il convient d'expliquer pourquoi certains états de systèmes macroscopiques correspondent aux états classiques tandis que les autres — leurs superpositions — ne sont jamais observés.

Décohérence

Ainsi, depuis les débuts de la mécanique quantique, la compréhension de la transition quantique-classique restait une question ouverte. C'est dans les années 1980, que W. H. Zurek y apporta un élément de réponse concret [Zur81a, Zur82a]. En effet, il a explicité un mécanisme simple permettant de comprendre pourquoi les propriétés quantiques ne se manifestent généralement pas à notre échelle : l'interaction avec un environnement.

Ce mécanisme peut-être explicité dans le cas simple qui nous a servi à illustrer le modèle de mesure de von Neumann et à formuler le problème de la mesure. On se souvient que l'état final du système quantique et de l'appareil de mesure après leur interaction est donné par $(|0\rangle\,|d\rangle + |1\rangle\,|g\rangle)\sqrt{2}$ lorsque le système est initialement dans l'état $(|0\rangle + |1\rangle)/\sqrt{2}$. Prenons maintenant en compte un environnement quantique initialement dans l'état $|e\rangle$ évoluant soit dans l'état $|e_d\rangle$ si l'appareil de mesure est dans l'état $|d\rangle$, soit dans l'état $|e_g\rangle$ si l'appareil de mesure est dans l'état $|g\rangle$. Alors, l'état "système + appareil de mesure + environnement" après que toutes les interactions ont eu lieu est donné par:

$$\frac{1}{\sqrt{2}}(|0\rangle\,|d\rangle\,|e_d\rangle + |1\rangle\,|g\rangle\,|e_g\rangle). \tag{10}$$

En quoi cela résout-il spécifiquement le problème de la mesure? La réponse est simple. Pour cela, il faut se pencher sur l'état "système + appareil de mesure". Dans le cas où les états $|e_d\rangle$ et $|e_g\rangle$ sont orthogonaux, on obtient la matrice densité $(|0\rangle\langle 0| \otimes |d\rangle\langle d| + |1\rangle\langle 1| \otimes |g\rangle\langle g|)$. Le processus d'interaction avec l'environnement permet donc de passer de corrélations quantiques — l'état enchevêtré explicité plus haut — à des corrélations classiques — une variable aléatoire classique pour l'appareil de mesure corrélée avec certains états du système. L'avantage de ces corrélations classiques est l'absence d'ambiguïté par changement de base. En effet, une seule base de l'appareil de mesure permet d'exhiber des corrélations dont la nature peut être expliquée à l'aide d'une variable aléatoire classique pour l'appareil de mesure — i.e. en employant uniquement des états orthogonaux pour décrire son état. Bien entendu, il s'agit là d'un cas d'école. Dans la pratique, les interactions entre systèmes quantiques sont complexes et il est parfois bien difficile de distinguer dans quelles circonstances un tel mécanisme permet de résoudre le problème de la mesure.

Avant de procéder à une étude plus détaillée, il est important de noter qu'à une question fondamentale, la réponse apportée est de nature opérationnelle. Il ne s'agit pas d'ajouter un axiome à la mécanique quantique comme le requiert l'interprétation de Copenhague, mais seulement de reconnaître que les systèmes quantiques macroscopiques sont rarement des systèmes isolés. L'effet de l'environnement — un système quantique légitime — sur la capacité des systèmes quantiques à être dans des superpositions d'états "classiques" doit donc être prise en compte. De façon plus précise, de nombreux travaux théoriques étayés par des résultats expérimentaux ont montré que l'effet de l'environnement sur l'existence de superpositions d'états est en général exponentiel avec le nombre de particules quantiques élémentaires qui composent le système étudié. En résumé, le mécanisme proposé par W. H. Zurek donne une explication raisonnable à l'émergence d'un monde classique à partir de la mécanique quantique.

D'autres éléments viennent confirmer cette approche. Tout d'abord il existe des systèmes macroscopiques qui obéissent aux lois de la mécanique quantique (par exemple les barres de Weber, pesant plus d'une tonne) ce qui indique que la transition quantique-classique n'est pas fondamentalement liée à la taille des objets considérés mais bien à leur capacité à être ou ne pas être en interaction avec leur environnement. Ensuite, et ce sujet occupe la seconde partie de cette thèse, l'ensemble des travaux effectués sur les

codes correcteurs d'erreurs quantiques s'appuie directement sur le fait que la décohérence est un phénomène, certes réel, mais pas fondamental [Prea, Preb, NC00a, Got97a]. Elle peut donc être combattue. Les années qui viennent confirmeront certainement ce point de vue, qui est maintenant partagé par une très large majorité de la communauté des expérimentateurs et des théoriciens de la mécanique quantique.

Concernant le travail que j'ai effectué sur la transition quantique-classique, il consiste en deux sujets distincts:

1. La classification des différents types de corrélations présentes dans les systèmes quantiques, avec pour application l'obtention d'un critère permettant d'analyser la résolution effective du problème de la mesure ;

2. L'étude de l'émergence de propriétés objectives sous l'effet d'une interaction avec un environnement.

A propos de ce second point, il est important de noter qu'il s'agit d'une approche nouvelle tranchant radicalement avec les discussions habituelles concernant la présence ou l'absence de principe de superposition comme seule caractéristique déterminant la nature quantique ou classique des systèmes physiques. Le travail effectué s'attache à une autre caractéristique du monde classique: l'existence d'une représentation objective des états des systèmes physiques.

Discorde quantique [OZ02a] *(Chapitre 3)*

Comme cela vient d'être brièvement mentionné, parallèlement au développement de la théorie de la décohérence, la théorie de l'information quantique a commencé à progressivement réinterpréter les différentes expériences de pensées chères aux fondateurs de la mécanique quantique. En particulier, la violation des inégalités de Bell par des états enchevêtrés peut être vue comme la ressource fondamentale permettant des applications nouvelles comme la téléportation quantique [BBC+93a] ou encore le codage super-dense [Prea]. C'est dans le but de participer à une analyse des fondations de la mécanique quantique à l'aide de la théorie de l'information que j'ai travaillé sur la discorde quantique.

Pour paraphraser une fois de plus le problème de la mesure, la mécanique quantique des systèmes fermés ne privilégie aucune base de l'espace de Hilbert des états. Ainsi, il est impossible de justifier l'émergence d'états classiques pour des systèmes fermés. C'est l'invariance par changement de base : l'état d'un système ne dépend pas de la base dans laquelle son vecteur d'état est exprimé. En revanche, dans le cadre des systèmes quantiques ouverts, et sous certaines conditions, une base classique peut être dynamiquement privilégiée comme étant l'ensemble des états quantiques les plus stables.

La discorde quantique s'attaque précisément à la reconnaissance d'une telle situation à l'aide du modèle de la mesure. Elle s'attache donc à déterminer quelle est la base classique de l'appareil de mesure $\mathcal{A}$ qui émerge lorsqu'il interagit avec son environnement. Pour cela, elle propose de classer les différents types de corrélations entre le système mesuré $\mathcal{S}$ et $\mathcal{A}$ après l'interaction de ce dernier avec son environnement.

Sa définition est fondée sur le calcul de l'information mutuelle entre $\mathcal{S}$ et $\mathcal{A}$ selon deux définitions identiques pour le cas de probabilités classiques, mais différentes dans le cas de systèmes quantiques :[1]

$$\mathcal{I}(\mathcal{S}:\mathcal{A}) = S(\mathcal{S}) + S(\mathcal{A}) - S(\mathcal{S},\mathcal{A}) \quad (11)$$
$$\mathcal{J}(\mathcal{S}:\mathcal{A}) = S(\mathcal{S}) - S(\mathcal{S}|\mathcal{A}). \quad (12)$$

Alors que l'on a pour des probabilités classiques $\mathcal{I}(\mathcal{S} : \mathcal{A}) = \mathcal{J}(\mathcal{S} : \mathcal{A})$, dans le cas quantique, on obtient facilement, $\mathcal{I}(\mathcal{S} : \mathcal{A}) \geq \mathcal{J}(\mathcal{S} : \mathcal{A})$. La discorde quantique est donc définie par $\delta = \mathcal{I}(\mathcal{S} : \mathcal{A}) - \mathcal{J}(\mathcal{S} : \mathcal{A})$. Une analyse plus détaillée montre que $\mathcal{J}(\mathcal{S} : \mathcal{A})$ mesure la quantité d'information sur le système $\mathcal{S}$ qui peut être retirée de ses corrélations avec une mesure particulière sur $\mathcal{A}$, et que $\mathcal{I}(\mathcal{S} : \mathcal{A})$ est une mesure globale des corrélations entre les deux systèmes quantiques. La différence entre ces deux quantités est donc une mesure de l'information qui ne peut être extraite par la meilleure des mesures sur $\mathcal{A}$, mais qui est effectivement présente dans les corrélations entre $\mathcal{S}$ et $\mathcal{A}$.[2]

Le résultat central du chapitre 3 montre que lorsque $\delta = 0$ pour une certaine base de mesure sur $\mathcal{A}$, alors les deux systèmes quantiques sont corrélés uniquement de façon classique au travers de cette base de mesure de $\mathcal{A}$. Lorsqu'une telle situation se produit, on peut reconnaître que l'invariance par changement de base inhérente à la mécanique quantique des systèmes fermés est levée: la matrice densité représentant l'état "système + appareil de mesure" peut-être décrit en employant une variable aléatoire pour l'appareil de mesure corrélée avec certains états du système. En conclusion, c'est l'environnement qui impose à l'appareil de mesure quelles corrélations garder avec le système mesuré et donc quels sont les résultats possibles, éliminant du même coup leurs superpositions.

Une autre conséquence mise en avant par l'étude de la discorde est la conversion progressive de l'information quantique en une information classique : au cours d'une interaction avec un environnement, deux systèmes quantiques partageant initialement de l'enchevêtrement vont peu à peu voir ces corrélations diminuer tout en gardant (sous certaines conditions) leurs corrélations classiques intactes. La discorde quantique est à ce titre un indicateur de la présence du phénomène de décohérence : il s'agit d'une mesure de la diminution des ressources disponibles pour effectuer de la communication quantique.

On peut cependant noter que la définition d'information quantique donnée par l'emploi de la discorde ne correspond pas à la définition de l'enchevêtrement. En effet, le chapitre 3 fournit l'exemple des états de Werner [Wer89a] dont l'enchevêtrement est nul, mais dont la discorde est toujours positive. Ceci est dû à la différence de point de vue adoptée dans la définition de la discorde et de l'enchevêtrement : dans un cas

[1] Pour être précis, il faut noter que la définition de $\mathcal{J}$ dépend du choix d'une base de projecteurs orthogonaux agissant sur l'espace de Hilbert de $\mathcal{A}$. Afin de ne pas alourdir les notations ni la discussion qui suit, cet élément technique mais néanmoins important a été omis.

[2] Certaines corrélations quantiques ne peuvent être efficacement utilisées qu'à condition de permettre l'emploi de communication classique entre les deux systèmes quantiques.

la nature des corrélations est décrite dans la perspective d'une mesure, tandis que dans l'autre la nature quantique se manifeste au moment de la création de ces corrélations.

C'est grâce à cette spécificité que la discorde quantique a pu être utilisée pour analyser à nouveau le paradoxe du démon de Maxwell, ou encore pour quantifier les ressources nécessaires aux protocoles de synchronisation d'horloges.

Objectivité [OPZ03a] *(Chapitre 4)*

Bien que la théorie de la décohérence puisse être considérée comme un grand pas en avant sur le chemin de la compréhension de la transition quantique-classique, nous ne sommes pas encore au bout du chemin.

En effet, la décohérence met en avant un mécanisme autorisant l'apparition d'une base privilégiée pour les systèmes quantiques en interaction [Zur81a, Zur82a, Zur91a, PZ01a, Zur03a]. Il s'agit d'états classiques dans le sens où toute superposition de ces états est intrinsèquement instable et est transformée en un mélange statistique d'états stables. En revanche, une partie de la nature quantique de ces états stables subsiste : lorsqu'un système subit une mesure, il est inévitablement perturbé par cette dernière. Malgré la décohérence, une mesure aboutit toujours à une re-préparation. C'est la conséquence inévitable de l'axiome relatif à l'acquisition d'information à propos des systèmes quantiques : toute mesure induit une acquisition d'information et par conséquent doit être accompagnée d'une mise à jour de l'état du système mesuré. En d'autre termes, même pour les états stables, on observe une extrême sensibilité aux mesures.

Cette simple constatation est en contradiction presque totale avec une des caractéristiques fondamentales de la mécanique classique : l'existence d'une réalité sous-jacente indépendante de l'observateur, ou encore celle de propriétés objectives. La lune existe qu'on la regarde ou non, et son état n'est pas modifié par l'observation qu'on en fait. Une explication satisfaisante de la transition quantique-classique doit donc permettre de comprendre comment une réalité objective peut émerger d'un substrat quantique intrinsèquement dépendant de la façon dont il est observé.

Pour répondre à cette attente, nous nous sommes penchés sur la façon utilisée pour appréhender les propriétés des systèmes classiques. Nous avons mis en avant le fait qu'aucun système classique n'était en fait mesuré directement comme c'est le cas en mécanique quantique. En fait, ce n'est qu'en interceptant une fraction de l'environnement d'un système classique que nous déduisons l'ensemble de ses propriétés.

De ce point de départ, nous avons promu l'environnement d'un rôle passif, destructeur des corrélations quantiques, à un rôle actif, canal de communication entre le système et ses observateurs. En étudiant comment l'information est propagée par l'environnement à différents observateurs nous pouvons comprendre l'existence de propriétés objectives pour des systèmes macroscopiques.

Dans le chapitre 4, une mesure de redondance est introduite pour les systèmes quantiques. Elle s'appuie sur un comptage précis du nombre de fois qu'une même information peut être extraite par des mesures sur des fragments disjoints de l'environnement.

Definition .0.1 (Redondance). *La redondance de niveau δ d'une observable A du*

système est donnée par:

$$R_\delta(A) = \max_{\bigotimes_k \mathcal{E}_k = \mathcal{E}} \{\mathcal{E}_k / \hat{I}_{\mathcal{E}_k}(A) \geq (1-\delta)\hat{I}_{\mathcal{E}}(A)\}, \tag{13}$$

où $\hat{I}_{\mathcal{F}}(A) = \max_X$, *mesure sur* $_{\mathcal{F}} I(A,X)$ *et* $I(A,X)$ *est l'information mutuelle entre les variables aléatoires correspondant aux résultats des mesures* A *sur le système et* X *sur un fragment* $\mathcal{F}$ *de l'environnement* $\mathcal{E}$. ∎

Une étude approfondie de cette quantité montre qu'une seule observable A du système peut voir son information diffusée de façon complète — $\hat{I}_{\mathcal{E}}(A) \simeq H(A)$ — et redondante dans l'environnement. Ainsi, lorsqu'une propriété d'un système quantique a une grande redondance, on peut en déduire que cette quantité jouera le rôle de grandeur objective pour peu que son empreinte dans l'environnement est suffisamment fidèle. En effet, on peut montrer que la redondance, associée à l'unicité de cette information, permet à plusieurs observateurs indépendants et honnêtes d'arriver à un consensus sur certaines propriétés d'un système quantique. Dès qu'un tel consensus peut être atteint, il est possible de définir de façon opérationnelle la réalité objective comme l'ensemble des propriétés des systèmes physiques sur lesquels de tels observateurs peuvent se mettre d'accord.

Par ailleurs, il a été possible de montrer analytiquement que redondance et fidélité de l'information conduisent à la suspension de l'invariance par changement de base évoquée plus haut. De fait, l'existence d'information redondante est un critère plus restrictif que l'analyse standard de la décohérence qui ignore le rôle actif que joue l'environnement dans la transition quantique-classique.

Une fois encore, on peut remarquer qu'une question qui semblait fondamentale finit par trouver un élément de réponse de nature opérationnelle. Cependant, ce serait aller trop vite que de croire que tous les problèmes soulevés par la mécanique quantique peuvent être résolus de cette façon. En effet, cette étude préliminaire amène déjà à considérer de nouvelles questions relatives à l'interprétation de la mécanique quantique. En particulier le cadre employé pour la définition de la redondance donne une place prépondérante à la décomposition de l'environnement en sous-systèmes élémentaires. Pourquoi et comment une telle structure émerge de la théorie quantique sera certainement le prochain sujet auquel il faudra s'atteler. Par ailleurs, il est clair que la mécanique quantique ne peut être considérée comme une théorie ultime à elle seule puisqu'elle ne permet pas une description de l'espace-temps cohérente avec le principe de relativité.

Cependant, l'étude d'autres modèles menant à une réalité objective au travers de la mécanique quantique usuelle devra être poursuivie, avec l'espoir que cela permette de répondre au moins partiellement au programme formulé par C. Fuchs [Fuc02c] visant à justifier les axiomes de la mécanique quantique par la théorie de l'information.

Codes correcteurs d'erreurs quantiques

Alors que la théorie de la décohérence a été initialement introduite pour répondre à un manque d'explications satisfaisantes concernant les fondations de la mécanique quan-

tique, elle est également au cœur des difficultés rencontrées pour fabriquer les ordinateurs quantiques. Avant de résumer les résultats obtenus, il est approprié de présenter le contexte dans lequel s'est effectuée cette recherche, celui de l'informatique quantique (pour une introduction complète au calcul quantique on peut se référer à [Prea, Preb, NC00a]).

Informatique quantique *(Chapitre 5)*

Il y a vingt ans, naissait l'idée du calcul quantique [Fey82a, Fey84a]. Il s'agissait alors de simuler l'évolution de systèmes quantiques complexes avec d'autres systèmes quantiques, un peu à la manière dont les ordinateurs classiques simulent l'évolution de systèmes classiques. L'idée sous-jacente à cette proposition était la remarque simple que toutes les simulations classiques de systèmes quantiques semblent requérir une quantité de mémoire croissant de façon exponentielle avec le nombre de systèmes simulés.

Faire calculer à des systèmes quantiques l'évolution d'autres systèmes quantiques permettrait donc de s'affranchir de cette contrainte. Mais finalement, les recherches dans le domaine ont pris une tournure un peu différente à partir du moment où ont été trouvés des algorithmes résolvant des problèmes difficiles ou supposés tels, comme c'est le cas de la factorisation des nombres entiers par exemple [Sho94a]. Le principe de base de fonctionnement d'un ordinateur quantique est l'utilisation du principe de superposition. En effet, puisque les évolutions quantiques sont linéaires, il est possible, en préparant le registre de l'ordinateur dans une superposition de valeurs d'entrée, d'effectuer plusieurs calculs à la fois. C'est ce qui est communément appelé le parallélisme quantique.

Cependant, la situation décrite ici est un peu simpliste. En effet, si tous les calculs ont bien été effectués en même temps, l'obtention d'un résultat particulier reste lié à un processus probabiliste. Si aucun traitement n'était effectué sur l'état du registre quantique, rien ne distinguerait une telle machine d'un ordinateur classique probabiliste. La détermination des transformations à effectuer est la véritable difficulté de l'algorithmique quantique.

Dans certains cas, comme pour la factorisation d'entiers, ces transformations apportent un gain exponentiel par rapport au meilleur algorithme classique connu. Dans d'autres cas, comme celui de la recherche dans une base de données non-ordonnée [Gro97a], le gain n'est que quadratique, mais l'existence d'un 'gap' entre le meilleur algorithme classique possible et l'algorithme quantique peut être démontrée.

Cependant, comme nous l'avons vu plus haut, le taux auquel disparaissent les superpositions d'états est en général exponentiel dans le nombre de systèmes impliqués dans la superposition. Ici le nombre de ces systèmes est directement proportionnel au nombre de bits quantiques utilisables dans le registre de l'ordinateur. Par conséquent, même si la décohérence n'est pas un phénomène fondamental, elle semble dresser une barrière insurmontable qui remet en cause le bien-fondé de tout programme de recherche visant à construire un ordinateur quantique.

Stratégies de suppression d'erreurs *(Chapitre 6)*

A première vue, une solution simple semble pourtant s'imposer : en isolant les registres de calcul de toute interaction involontaire avec un environnement incontrôlé, on diminue d'autant l'importance de la décohérence. Malheureusement, une telle stratégie n'est pas toujours applicable aux différentes propositions expérimentales visant à la construction d'un ordinateur quantique. Plus fondamentalement, une isolation accrue des registres signifie une dégradation dans la même proportion du contrôle qui peut être effectué sur ce registre. Ainsi, une meilleure isolation prolonge le temps de cohérence mais également le temps nécessaire pour accomplir une porte élémentaire donnée (pour une introduction à la problématique contrôle-décohérence on peut lire [Bac01a]).

Il faut donc mettre en place des stratégies actives de correction d'erreurs. Plusieurs outils sont désormais à notre disposition :

- Les codes correcteurs d'erreurs quantiques, analogues aux codes classiques (pour une introduction aux codes voir [Got97a, Preb, NC00a]) ;
- Les sous-systèmes protégés de la décohérence, sous-systèmes qui, pour des raisons de symétrie de leur interaction avec l'environnement, ne subissent pas de décohérence [Zan99a, LBKW99a, BKLW00a, KBLW01a] ;
- Les contrôles actifs induisant une interaction effective connue et maîtrisée [LV00a, VKL99a, VKL00a].

Durant ce travail de thèse je me suis plus principalement penché sur les codes correcteurs d'erreurs quantiques. Je vais donc donner ici quelques bases permettant de comprendre le travail qui a été effectué sur la définition de codes quantiques convolutifs et les difficultés qui ont été résolues.

La plupart des codes correcteurs connus à l'heure actuelle dérivent d'un formalisme appelé formalisme stabilisateur [Got97a]. Ce formalisme permet une analogie simple avec les codes classiques. En effet, le sous-espace C d'un code protégeant k qubits dans n est défini comme le plus grand sous-espace stabilisé par un sous groupe abélien S du groupe de Pauli des qubits physiques, $\text{span}\{X, Y, Z, I\}^{\otimes n}$.

Sans entrer dans les détails de la définition de chacun de ces objets, il est important de noter que cette procédure identifie le sous-espace du code à l'aide d'équations linéaires sur le vecteur représentant l'état du système physique composite dans lequel est stockée l'information. En d'autre termes, ces équations peuvent êtres assimilées aux équations de syndromes des codes classiques linéaires (cf. Table 1). De façon plus précise, le groupe stabilisateur est généré par $n-k$ opérateurs indépendants, $\{M_i\}$, qui commutent deux à deux. Alors, un vecteur $|\psi\rangle$ est dans le sous-espace du code si et seulement si:

$$\forall i \in \{1, \ldots, n-k\},\ M_i |\psi\rangle = |\psi\rangle\,. \tag{14}$$

Le grand avantage de cette formulation est qu'elle contourne un résultat bien connu qui semblait compromettre à jamais la correction d'erreurs quantiques. En effet, en

1982 W. Wootters et W. H. Zurek ont montré [Die82a, WZ82a] qu'il était impossible de copier parfaitement l'état d'un système quantique inconnu. Par conséquent, introduire de la redondance dans le but de protéger de l'information contre des erreurs au moyen d'un code correcteur semblait irréalisable. En fait, la description d'un code au moyen de ses équations de syndromes ne fait aucune référence à une nécessité de copier l'information pour introduire de la redondance. Ici, la redondance n'est que la conséquence de l'espacement des différents mots de codes. Elle ne provient pas d'une répétition éventuelle de l'information. Cette capacité fut donc exploitée pleinement pour la définition des codes stabilisateurs.

Une autre question délicate que cette définition permet de résoudre concerne la procédure d'estimation d'erreurs. Comme nous l'avons vu précédemment dans un autre contexte, toute mesure change l'état du système mesuré détruisant par la même occasion une partie des superpositions quantiques dans lesquelles il pouvait être impliqué. Il est donc impossible de mesurer les qubits reçu pour effectuer l'opération d'estimation d'erreur sans détruire de cette façon toute l'information quantique qu'ils contenaient.

L'avantage de l'introduction d'un analogue quantique des équations de syndromes est de proposer une méthode permettant de préserver la nature quantique de l'information à protéger tout en acquérant le maximum d'information sur l'erreur qui a pu se produire pendant la transmission. Cette méthode consiste à réaliser la mesure quantique correspondant à chacune des observables représentées par les générateurs M_i du groupe S. Ceci s'effectue très facilement au moyen d'un qubit auxiliaire (cf. Fig. 3) Alors, les seules superpositions détruites au cours de la mesure sont des superpositions d'états correspondant à des erreurs différentes, ce qui n'affecte pas l'information protégée. Aussi surprenant que cela puisse paraître, les axiomes de la mécanique quantique nous permettent ici de limiter les erreurs à corriger à un petit nombre (appelé base d'erreurs) sans avoir à se préoccuper de leurs éventuelles superpositions [Kni96b].

Par ailleurs, le formalisme stabilisateur offre l'avantage d'une description simple du circuit quantique de codage : pour être utilisable en pratique, il faut ajouter à la description du code, celle de son circuit de codage, c'est à dire la description en terme de portes logiques élémentaires de la transformation permettant le passage de l'état des bits quantiques non-protégés à l'état des qubits physiques appartenant au sous-espace du code. La difficulté provient des exigences spécifiques au calcul quantique. Toutes les manipulations doivent être effectuées de façon réversible ce qui rend parfois ardue la détermination de la séquence de portes élémentaires permettant de réaliser une transformation donnée. Dans le cas des codes stabilisateurs, cette transformation est déduite facilement et de façon algorithmique de l'écriture des générateurs du groupe S correspondant au code étudié. Ceci accroît donc l'intérêt des codes stabilisateurs par rapport à des descriptions alternatives de codes quantiques.

Finalement, et c'est la question cruciale, il a fallu déterminer si les code quantiques permettent de combattre efficacement le phénomène de décohérence. Autrement dit, les codes quantiques sont-ils suffisamment puissants pour rendre le calcul quantique à même de franchir la barrière de la décohérence — i.e. pour balancer la tendance des systèmes quantiques de grande taille à perdre leur faculté à se superposer. Une

Classique	Quantique
n bits : $\overbrace{(0,1,\ldots,0)}^{n}$	n qubits : $\|\psi\rangle \in \overbrace{\mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_2}^{n}$
Code (n,k) : ensemble $\mathcal{C}$ de 2^k mots de code parmi les 2^n chaînes de n bits	Code (n,k) : sous espace $\mathcal{C}$ de dimension 2^k de $\mathcal{H}_2^{\otimes n}$
Code linéaire : $w \in \mathcal{C} \Leftrightarrow \forall\, i \in [[1, n-k]], h_i.w = 0$	Code stabilisateur : $\|\psi\rangle \in \mathcal{C} \Leftrightarrow \forall\, i \in [[1, n-k]], M_i \|\psi\rangle = \|\psi\rangle$
h_i chaîne de n bits	$M_i = \overbrace{X \otimes Y \otimes \ldots \otimes Z}^{n}$, ce sont les générateurs du groupe stabilisateur. Ils vérifient $[M_i, M_j] = 0$.

Table 1: Analogies entre les codes linéaires classiques et les codes stabilisateurs

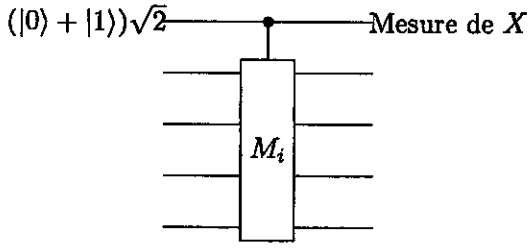

Figure 3: Mesure associée au générateur M_i du groupe stabilisateur S. Lorsque le résultat de la mesure de l'opérateur X sur le premier qubit est $+1$, l'erreur survenue dans la transmission commute avec M_i. Lorsque le résultat est -1, l'erreur anti-commute avec M_i.

réponse affirmative fut apportée dès 1996 [Sho96a, AB96a]. En utilisant des schémas de concaténation et sous des hypothèses raisonnables, il est possible d'effectuer un calcul quantique arbitrairement long avec un taux d'erreur arbitrairement bas pourvu que la décohérence soit suffisamment faible. La grande surprise ici est donc l'existence d'un seuil pour le taux d'erreur par qubit, appelé seuil pour le calcul quantique tolérant à la faute, en dessous duquel tout calcul devient théoriquement possible.

Codes convolutifs quantiques [OT03a, OT04a] *(Chapitre 7)*

Dans la lignée de ces travaux concernant l'étude des codes pour le calcul quantique, je me suis intéressé à la définition d'un analogue quantique des codes convolutifs. Comme nous l'avons vu, la plupart des codes quantiques ont pour vocation d'être utilisés pour protéger de la décohérence un registre de calcul au sein même d'un ordinateur quantique. Les codes convolutifs sont quant à eux plus orientés vers la protection de l'information dans le contexte des transmissions quantiques à longue distance, et en particulier pour la cryptographie quantique. On peut en fait penser que certains de ces codes pourraient être utilisés dans un avenir relativement proche pour améliorer les performances des systèmes de cryptographie quantique.

La principale difficulté rencontrée dans la définition des codes quantiques convolutifs est l'absence de registres à décalages. En effet, le concept même d'une mémoire dont le contenu peut être utilisé au cours d'un calcul arithmétique ne peut être employé dans le cadre quantique en raison de l'impossibilité de copier la valeur d'un bit inconnu. Ainsi, une généralisation directe des codes convolutifs classiques (qui utiliserait les registres à décalages) est vouée à l'échec. La solution à ce problème a été de ne retenir que la définition des codes convolutifs classiques impliquant leurs syndromes. Utilisant l'analogie syndromes-générateurs du groupe stabilisateur, une définition adéquate a ensuite pu être trouvée : un code convolutif de paramètres (k, n, m) est la donnée de $n-k$ opérateurs, chacun agissant sur au plus $n+m$ bits quantiques. Les générateurs du sous-groupe stabilisateur du code sont obtenus par décalage à droite de n position de ces $n-k$ opérateurs.

Definition .0.2 (Code convolutif (n, k, m)). *Le groupe stabilisateur S pour un code convolutif de paramètres (n, k, m) est défini par*

$$S = \mathrm{sp}\{M_{j,i} = I^{\otimes j \times n} \otimes M_{0,i}, 1 \leq i \leq n-k,\ 0 \leq j\}, \tag{15}$$

où $M_{0,i}$ est un élément du groupe de Pauli pour $n+m$ qubits. Les $M_{j,i}$ doivent être indépendants et commuter deux à deux.

Malgré les similitudes entre cette définition et une définition des codes convolutifs classiques en terme de leurs équations de syndromes (cf. Table 2), il reste à prouver que ces codes quantiques ont bien les propriétés désirées. Ces propriétés sont les suivantes:

1. Possibilité de protéger l'information en ligne — sans devoir attendre la fin du message pour réaliser l'encodage — tout en maintenant une taille de bloc aussi longue que le message protégé ;

2. Algorithme d'estimation d'erreur au maximum de vraisemblance de complexité linéaire vis à vis du nombre de qubits protégés lorsque le canal est sans mémoire.

Outre la définition du code, il faut donc en particulier se pencher sur le circuit de codage permettant de protéger l'information à l'aide de ce nouveau type de codes. Bien que cette question puisse paraître incongrue à qui est habitué aux codes classiques, elle revêt une importance particulière dans le cadre quantique. En effet, comme cela a déjà été mentionné plus haut, introduire la redondance nécessaire à la correction d'erreur est une question non-triviale lorsqu'il est impossible de copier la valeur des bits que l'on désire protéger.

Bien que cette tâche puisse toujours être accomplie pour les codes stabilisateurs, elle requiert la description d'une méthode générale produisant un circuit d'encodage en terme de portes quantiques élémentaires, c'est à dire réversibles et linéaires. Pour se faire, il a d'abord été nécessaire d'adapter le formalisme des codes stabilisateurs pour prendre en compte la spécificité des codes convolutifs quantiques. Pour cela un formalisme polynomial a été développé (cf. Table 3). Il permet le calcul des relations de commutation entre un opérateur quelconque et un générateur du groupe stabilisateur, ainsi qu'avec tous ses décalés de n positions. On se ramène donc à une situation proche de celle des codes en blocs : on peut trouver un circuit d'encodage assez similaire à celui des codes stabilisateurs habituels tout en maintenant la contrainte d'un délai constant entre le flot de bits à protéger et le flot de bits protégés. Ce dernier point fait de cette construction le véritable analogue des codes convolutifs classiques. En effet, cette propriété de délai fixe entre l'arrivée de bits non protégés et la sortie de bits protégés peut être considérée comme une définition alternative des codes convolutifs.

D'autre part, une étude plus poussée montre que le décodage peut s'effectuer de la même manière (avec un délai constant). Cette possibilité est cependant conditionnée par l'absence d'erreurs catastrophiques pour l'encodage envisagé. Cette absence d'erreurs se propageant à un nombre infini de bits quantiques après décodage peut, comme dans le cas classique, être exprimée sous la forme d'une condition simple, vérifiable de manière algorithmique. En revanche, la dérivation ainsi que le résultat lui même diffèrent très sensiblement de ceux donnés par Sain et Massey pour les codes classiques [MS68a] (cf. Table 4). La différence principale réside encore une fois dans les contraintes supplémentaires imposées par la mécanique quantique sur le circuit de codage : la linéarité et la réversibilité du calcul limite les possibilités de circuits effectuant l'opération de décodage.

Enfin, l'étude menée sur les codes convolutifs quantiques se termine par la description d'un algorithme d'estimation de l'erreur la plus probable pour les canaux sans mémoire analogue de l'algorithme de Viterbi [Vit67a] (cf. Table 5), mais prenant en compte les spécificités quantiques — seule une mesure du syndrome peut être effectuée.

Les résultats présentés dans ce manuscrit ne correspondent qu'aux premiers pas dans l'étude de ces codes. Ils montrent que la définition que nous avons proposée est un bon analogue de la définition des codes convolutifs classiques. Cependant, un travail complémentaire devra être mené pour accroître notre connaissance de ces codes. En particulier,

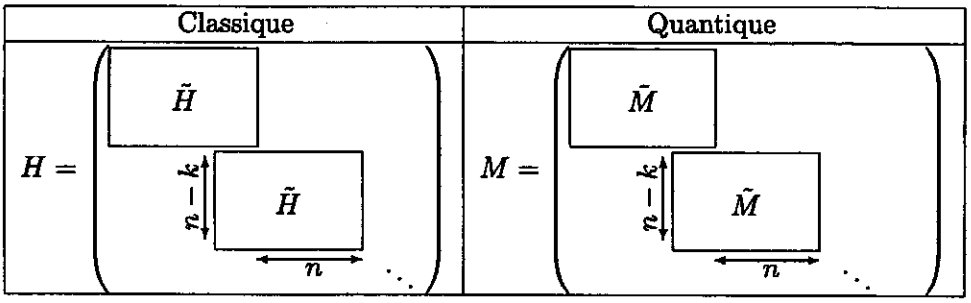

Table 2: Similitudes entre la définition d'un code convolutif classique à l'aide de sa matrice de parité et la définition des générateurs du groupe stabilisateur pour les codes convolutifs quantiques. La matrice M contient en chaque position une matrice de Pauli. Une ligne correspond à un générateur, une colonne à un qubit.

$$
\begin{array}{rcl}
M_{0,1} & = & Z\ X\ X\ Z\ I\ I\ I \ldots, \\
M_{0,2} & = & I\ Z\ X\ X\ Z\ I\ I \ldots, \\
M_{0,3} & = & I\ I\ Z\ X\ X\ Z\ I \ldots, \\
M_{0,4} & = & I\ I\ I\ Z\ X\ X\ Z \ldots, \\
M_{j,i} & = & I^{\otimes 5j} \otimes M_{0,i},\ 0 < j.
\end{array}
\qquad
M = \left(\begin{array}{ccccc|ccccc}
0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 & D & 0 & 1 & 0 & 0 \\
D & 0 & 0 & 0 & 1 & 0 & D & 0 & 1 & 0
\end{array}\right)
$$

Table 3: Le formalisme polynomial pour un exemple particulier de codes. A gauche le groupe stabilisateur, à droite la représentation polynomiale. Pour la partie de droite, la matrice se décompose en deux sous-matrices l'une représentant la partie X des générateurs, l'autre la partie Z. L'opérateur D représente un décalage de 5 qubits. Ainsi on peut se contenter de décrire le premier bloc de générateurs, le n-ième bloc étant obtenu par la multiplication de chaque entrée de la matrice par D^{n-1}.

il sera important de déterminer avec précision les conditions dans lesquelles un code convolutif doit être utilisé et comment il doit éventuellement être concaténé avec d'autres schémas de protection de l'information quantique pour pouvoir améliorer les distances sur lesquelles il est possible de transmettre efficacement l'information quantique.

Par ailleurs, une autre motivation à ce travail a été l'implication des codes convolutifs classiques dans les schémas de type turbo-codes. En effet, ces stratégies de correction d'erreurs utilisent généralement deux codes convolutifs pour atteindre des taux de transmission proches de la limite théorique fixée par le théorème de Shannon. Un équivalent quantique de ces codes permettrait sans doute d'améliorer la robustesse de la transmission et du stockage de l'information quantique, facilitant donc son traitement à grande échelle dans des structures dédiées.

Propositions expérimentales

Enfin, la dernière partie de cette thèse est dévolue à des aspects expérimentaux du traitement quantique de l'information. Ces travaux m'ont permis d'acquérir une meilleure compréhension des difficultés liées à la réalisation pratique de la manipulation quantique de l'information.

Le choix de l'électrodynamique quantique en cavité pour mener cette étude résulte de la présence à Paris d'un groupe de renommée internationale (LKB - ENS) ayant une forte activité dans le domaine de l'information quantique. Cependant, même si certains problèmes rencontrés concernent spécifiquement cette implémentation physique, la plupart des contraintes expérimentales imposées ainsi que la façon de remédier aux problèmes rencontrés sont applicables à d'autres dispositifs expérimentaux.

Dispositif expérimental *(Chapitre 8)*

Dans le cas particulier des expériences de l'ENS, l'information quantique est stockée dans le degré d'excitation électronique d'un atome de Rubidium dans un état de Rydberg [RBH01a]. De façon simple, il s'agit d'un atome très excité et donc de très grand volume, dont un des électrons peut accéder à plusieurs niveaux d'énergie permettant une représentation quantique de l'information.

Pour manipuler l'information contenue dans un seul de ces atomes, il est possible d'appliquer des champs micro-ondes dont la fréquence est ajustée avec celle de la transition entre deux niveaux d'énergie de l'atome. Dans la pratique, cette manipulation peut être effectuée de façon précise et en un temps très court, ce qui permet de s'affranchir des problèmes de décohérence.

Cependant, pour pouvoir manipuler de l'information de façon non triviale, il faut également pouvoir effectuer des portes logiques comportant plusieurs bits. D'un point de vue pratique, il s'agit de contrôler une interaction entre deux systèmes quantiques. Dans l'expérience de l'ENS, cette interaction ne se fait pas directement entre deux atomes, mais entre un atome et une cavité. Cette cavité est constituée de deux miroirs supraconducteurs refroidis aux alentours de 1 K, et ayant la capacité de conserver un photon

unique pendant un temps avoisinant 1 ms.

L'interaction d'un atome de Rydberg avec le champ dans la cavité est due à une modification de l'indice de réfraction induite par le passage de l'atome dans la cavité. Cette modification est elle-même liée à l'état d'excitation de l'atome. Ainsi cette interaction agit comme une porte logique conditionnant l'état du champ à l'intérieur de la cavité en fonction du niveau d'excitation de l'atome qui la traverse. Un second atome envoyé à la suite du premier peut entrer dans la cavité et interagir de nouveau avec elle, ce qui aboutit à une interaction effective entre les deux atomes concernés.

Une autre méthode, dite de collision assistée par cavité a été proposée plus récemment pour faire interagir deux atomes directement dans la cavité. Cette nouvelle possibilité s'est révélée cruciale pour l'une des propositions expérimentales décrites ici.

Les principaux obstacles que nous avons rencontrés pour le traitement de l'information quantique sont :

- La présence d'une seule cavité, ce qui empêche tout parallélisme (allongement du temps des expériences et donc plus grande sensibilité à la décohérence) ;
- L'impossibilité d'arrêter les atomes dans la cavité (deux atomes ne peuvent interagir qu'une seule fois par l'intermédiaire de la cavité) ;
- Impossibilité d'adressage individuel des atomes une fois qu'ils sont dans la cavité.

A la vue de ce bref descriptif, une question de taille demeure : quelles expériences peut-on réaliser avec ces moyens limités ?

Propositions [MOR03a, MOY+03a, OM03a] *(Chapitres 9 & 10)*

Nous nous sommes concentrés ici sur la réalisation de tâches simples : clonage optimal et porte quantique élémentaire. Elle peuvent être vues comme un banc d'essai permettant de préciser nos besoins en terme de contrôle et d'architecture des ordinateurs quantiques, tout en permettant le développement de solutions adaptées pour combattre la décohérence ou contourner les limitations intrinsèques des systèmes physiques utilisés.

Le clonage quantique optimal (cf. Fig. 4) consiste à réaliser à partir d'un état inconnu deux clones aussi proches que possible de l'état initial. Le résultat d'impossibilité évoqué plus haut interdit à cette opération d'être réalisée parfaitement. Cependant, il n'empêche pas de simplement maximiser la qualité des clones. Ce problème a été décrit en détail de façon théorique et a fait l'objet d'expériences sur d'autres systèmes physiques. Son implémentation dans le cas de l'électrodynamique quantique en cavité montre qu'avec la technologie actuelle, on ne peut espérer qu'une qualité assez médiocre des clones. En revanche, cette étude précise des objectifs qui devront être atteints dans le futur pour permettre la démonstration d'un traitement quantique complexe de l'information en utilisant l'électrodynamique quantique en cavité.

Notre seconde proposition expérimentale concerne l'implémentation d'une porte élémentaire à trois bits quantiques : la porte de Toffoli (cf. Fig. 5). Nous avons utilisé pour cela une interaction directe entre deux atomes avec une assistance par cavité ce

Proposition .0.1. *Un encodeur est non-catastrophique si et seulement si l'opération d'encodage $C(q)$ pour q blocs peut être décomposée de la façon suivante (pour q grand):*

$$C(q) = \tilde{T}_{\mathrm{erm}}(q) \times \left(\prod_{i=0}^{\lfloor q/l_t \rfloor} D^{il_t}[U_t] \right) \times \ldots \times \left(\prod_{i=0}^{\lfloor q/l_1 \rfloor} D^{il_1}[U_1] \right) \times \tilde{I}_{\mathrm{nit}}(q), \tag{16}$$

où $\tilde{I}_{\mathrm{nit}}(q)$ et $\tilde{T}_{\mathrm{erm}}(q)$ sont des étapes d'initialisation et de terminaison pour le codeur qui peuvent éventuellement dépendre de q mais dont le support est borné ; où $\{U_j\}_j$ est un ensemble fini d'opérateurs unitaires indépendants de q et tels que $D^i[U_j]$ and $D^{i'}[U_j]$ commutent ; et où les l_j sont des entiers indépendants de q.

Table 4: Condition de non-catastrophicité pour un codeur quantique.

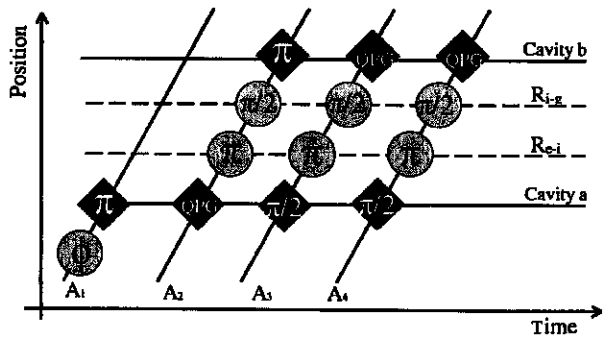

Figure 4: Résumé graphique des interactions successives permettant de réaliser le clonage optimal universel. On y voit les manipulations par les champs micro-ondes classiques effectuées dans les zones de Ramsey — i.e. R_{e-i}, R_{e-g} — ainsi que les transferts résonnants entre les atomes et les deux cavités contenant un champ quantique — avec au plus un photon.

Algorithm .0.1.
Inputs: (*i*) Liste de syndromes $\{s_{j+1,i}\}_i$ pour $i = 1, \ldots, n-k$; (*ii*) une liste $\{E_j^{(e)}\}_{e \in \{I,X,Y,Z\}^{\otimes m}}$ de candidats d'erreurs définis jusqu'au bloc j et telle que l'élément $E_j^{(e)}$ a une décomposition en produit tensoriel de matrices de Pauli comprenant l'indice e sur ses m dernières positions et qui maximise la vraisemblance étant donnée cette contrainte de terminaison. La liste $\{E_j^{(e)}\}_e$ est construite récursivement par l'algorithme.
Step $j+1$: Pour une valeur de $e' \in \{I, X, Y, Z\}^{\otimes m}$, on considère toutes les extensions de n qubits des éléments de $E_j^{(e)}$ tels que:

- ils satisfont les syndromes $s_{j+1,i}$ pour $i = 1, \ldots, n-k$;
- leur décomposition en produit tensoriel de matrices de Pauli se termine par e'.

Par construction ces extensions sont des candidats d'erreurs jusqu'au bloc $j+1$. Pour chaque élément e' de $\{I, X, Y, Z\}^{\otimes m}$ on choisit une des extensions avec vraisemblance maximale — en cas d'ex æquo on en choisit une au hasard. Ces éléments constituent la nouvelle liste $\{E_{j+1}^{(e')}\}_{e'}$.
Lorsque tous les syndromes ont été épuisés, le candidat avec la plus grande vraisemblance est une des erreurs les plus probables.

Table 5: Algorithme de Viterbi pour les codes convolutifs quantique.

qui a permis de réduire très sensiblement la complexité des manipulations requises et les sources d'erreurs. Avec des paramètres actuels, une telle porte logique pourrait être effectuée avec une précision de 70%, ce qui au regard des performances actuelles d'autres prototypes d'ordinateurs quantiques est un résultat relativement satisfaisant. Plus fondamentalement, cette étude montre que la décomposition trouvée est optimale vis à vis des contraintes auxquelles nous sommes soumis en employant le schéma expérimental du LKB. Par ailleurs, elle indique qu'un simple allongement de la durée de vie des photons dans la cavité permettrait d'atteindre une performance de 90% sans demander plus d'efforts dans le contrôle des différentes opérations effectuées au sein de la cavité. De nouveau, on voit apparaître l'importance de la décohérence (ici induite par la perte d'un photon de la cavité) dans la performance globale du traitement quantique de l'information.

Perspectives

Depuis que les premières versions de ce manuscrit ont été écrites, de nouveaux résultats ont été obtenus.

En premier lieu, la simulation d'évolutions de systèmes quantiques complexes avec des prototypes d'ordinateurs quantiques. Nous avons pu mettre en évidence une façon efficace de différencier deux régimes dynamiques en estimant de façon quantique la décroissance de la fidélité [PBKLO04a].

D'autre part, les recherches sur les codes quantiques convolutifs, les codes quantiques à matrice de parité creuse ont été poursuivies. Les grandes similitudes avec leurs analogues classiques nous font plus que jamais espérer que de telles classes de codes quantiques joueront un rôle important dans les futures structures de protection de l'information quantique.

Enfin, des résultats analytiques complémentaires sont venus étoffer notre compréhension du rôle joué par la redondance dans la transition quantique-classique. Plus précisément, il est possible de prouver dans de nombreux cas pratiques que la redondance implique l'existence d'une observable préférée dont les vecteurs propres sont les états classiques sélectionnés par l'interaction avec l'environnement.

L'avenir seul dira si un ordinateur quantique universel verra le jour, cependant, on peut dire d'ores et déjà dire que la physique et l'informatique seront à l'avenir encore profondément bouleversées par les recherches effectuées dans le cadre de l'information quantique.

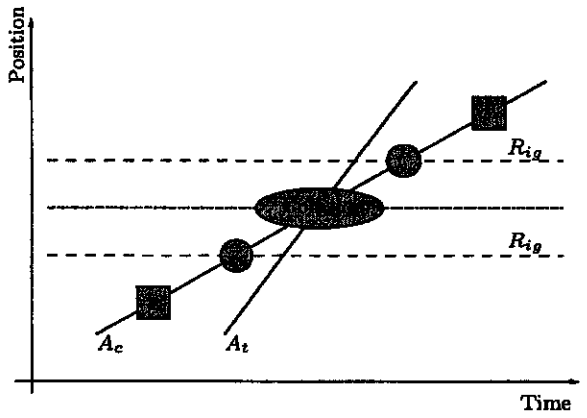

Figure 5: Représentation graphique des interactions atome-champ et atome-atome au sein de la cavité. Les cercles dénotent les interactions avec un champ micro-onde classique ajusté sur la transition $i - g$. Ce champ est généré par deux zones de Ramsey R_{ig}. Les carrés correspondent aux interactions résonantes avec le champ quantique stocké dans une cavité de grande finesse. La durée de chaque interaction est ajustée en contrôlant la longueur de chaque impulsion ou encore en appliquant une différence de potentiel sur les miroirs qui composent la cavité, ce qui change la valeur du coupage atome-champ.

Bibliography

[AB96a] AHARONOV, D., AND BEN-OR, M. Fault-tolerant quantum computation with constant error, 1996. arXiv, quant-ph/9611025.

[AB97a] AHARONOV, D., AND BEN-OR, M. Fault-tolerant quantum computation with constant error rate. In *Proc. 29th. Ann. ACM Symp. on Theory of Computing*, 1997. arXiv, quant-ph/9611025. Longer version quant-ph/9906129.

[AGR81a] ASPECT, A., GRANGIER, P., AND ROGER, G. Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.*, 47:460, 1981.

[Bac01a] BACON, D. *Decoherence, Control, and Symmetry in Quantum Computers*. PhD thesis, University of California at Berkeley, 2001.

[Ban01a] BANASZEK, K. Fidelity balance in quantum operations. *Phys. Rev. Lett.*, 86(7):1366–1369, February 2001. arXiv, quant-ph/0003123.

[BB84a] BENNETT, C. H., AND BRASSARD, G. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE international Conference on Computers, Systems and Signal Processing, Bangalore, India*, page 175, New York, 1984. IEEE Press.

[BBC+93a] BENNETT, C. H., BRASSARD, G., CRÉPEAU, C., JOZSA, R., PERES, A., AND WOOTTERS, W. K. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, 1993.

[BBHB97a] BUŽEK, V., BRAUNSTEIN, S., HILLERY, M., AND BRUSS, D. Quantum copying: a network. *Phys. Rev. A*, 56(5):3446–3452, 1997. arXiv, quant-ph/9703046.

[BCDM00a] BRUSS, D., CINCHETTI, M., D'ARIANO, G. M., AND MACCHIAVELLO, C. Phase-covariant quantum cloning. *Phys. Rev. A*, 62:012302/1–7, 2000.

[BCJR74a] BAHL, L. R., COCKE, J., JELINEK, F., AND RAVIV, J. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Inf. Theory*, 20:284–287, March 1974.

[BCMS01a] BENENTI, G., CASATI, G., MONTANGERO, S., AND SHEPELYANSKY, D. L. Efficient quantum computing of complex dynamics. *Phys. Rev. Lett.*, 87:227901, 2001. arXiv, quant-ph/0107036.

[BCT99a] BRASSARD, G., CLEVE, R., AND TAPP, A. Cost of exactly simulating quantum entanglement with classical communication. *Phys. Rev. Lett.*, 83:1874–1877, 1999. arXiv, quant-ph/9901035.

[BDE+98a] BRUSS, D., DIVINCENZO, D. P., EKERT, A., FUCHS, C. A., MACCHIAVELLO, C., AND SMOLIN, J. A. Optimal universal and state dependent quantum cloning. *Phys. Rev. A*, 57:2368, 1998. arXiv, quant-ph/9705038.

[Bel64a] BELL, J. S. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195, 1964.

[BEM98a] BRUSS, D., EKERT, A., AND MACCHIAVELLO, C. Optimal universal quantum cloning and state estimation. *Phys. Rev. Lett.*, 81:2598–2601, 1998. arXiv, quant-ph/9712019.

[Ben73a] BENNETT, C. Logical reversibility of computation. *IBM J. Res. Dev.*, 17:5225, 1973.

[Ben82b] BENIOFF, P. Quantum mechanical models of turing machines that dissipate no energy. *Phys. Rev. Lett.*, 48:1581–1585, 1982.

[Ber02a] BERTET, P. *Atomes en cavité: complémentarité et fonctions de Wigner*. PhD thesis, Université de Paris VI, 2002.

[BGL+03a] BOILEAU, J.-C., GOTTESMAN, D., LAFLAMME, R., POULIN, D., AND SPEKKENS, R. W. Robust polarization-based quantum key distribution over a collective-noise channel. *Phys. Rev. Lett.*, 92:017901, 2003. arXiv, quant-ph/0306199.

[BH96a] BUŽEK, V., AND HILLERY, M. Quantum copying: Beyond the no-cloning theorem. *Phys. Rev. A*, 54:1844, 1996. arXiv, quant-ph/9607018.

[BHJ03a] BHATTACHARYA, T., HABIB, S., AND JACOBS, K. Continuous quantum measurement and the quantum to classical transition. *Phys. Rev. A*, 67:42103, 2003.

[BHLS02a] BENNETT, C. H., HARROW, A. W., LEUNG, D. W., AND SMOLIN, J. A. On the capacities of bipartite Hamiltonians and unitary gates, 2002. arXiv, quant-ph/0205057.

[BKD+01a] BACON, D., KEMPE, J., DIVINCENZO, D., LIDAR, D. A., AND WHALEY, K. B. Encoded universality in physical implementations of a quantum computer. In *Proceedings of the International Conference on Experimental Implementation of Quantum Computation (IQC'01)*, Australia, 2001. Rinton Press. arXiv, quant-ph/0102140.

[BKLW00a] BACON, D., KEMPE, J., LIDAR, D. A., AND WHALEY, K. B. Universal fault-tolerant computation on decoherence free subspaces. *Phys. Rev. Lett.*, 85:1758–1761, 2000. arXiv, quant-ph/9909058.

[BLOT90a] BOGOLUBOV, N. N., LOGUNOV, A. A., OKSAK, A. I., AND TODOROV, I. T. *General principles of quantum field theory*. Kluver, Dordrecht, 1990.

[Boh27a] BOHR, N. In *Atti del Congresso Internazionale dei fisici*, volume 2, page 565, Como-Pavia-Roma, 1927.

[Boh28a] BOHR, N. H. D. The quantum postulate and the recent development of atomic theory. *Nature*, 121:580–590, 1928. Reprinted in [WZ83a].

[Boh63a] BOHR, N. *Essays 1958-1962 on atomic physics and human knowledge*. Interscience, New York, 1963.

[BOM+02a] BERTET, P., OSNAGHI, S., MILMAN, P., AUFFEVES, A., MAIOLI, P., BRUNE, M., RAIMOND, J. M., AND HAROCHE, S. Generating and probing a two-photon fock state with a single atom in a cavity. *Phys. Rev. Lett.*, 88(14):143601, 2002.

[BP89a] BARNETT, S. M., AND PHOENIX, S. J. D. Entropy as a measure of quantum optical correlation. *Phys. Rev. A*, 40:2404–2409, 1989.

[Bri01a] BRIEGEL, H. Universal quantum computation using only one-qubit measurements. Talk at ITP Conference on Quantum Information, 2001.

[BSKM+96a] BRUNE, M., SCHMIDT-KALER, F., MAALI, A., DREYER, J., HAGLEY, E., RAIMOND, J. M., AND HAROCHE, S. Quantum rabi oscillation: A direct test of field quantization in a cavity. *Phys. Rev. Lett.*, 76(11):1800, 1996.

[BW92a] BENNETT, C. H., AND WIESNER, S. J. Communication via one- and two-particle operators on Einstein -Podolsky- Rosen states. *Phys. Rev. Lett.*, 69:2881–2884, 1992.

[CA97a] CERF, N. J., AND ADAMI, C. Negative entropy and information in quantum mechanics. *Phys. Rev. Lett.*, 79:5194–5197, 1997.

[CDG01a] CERF, N. J., DURT, T., AND GISIN, N. Cloning a qutrit, 2001. arXiv, quant-ph/0110092.

[CJF+02a] CUMMINS, H. K., JONES, C., FURZE, A., SOFFE, N. F., MOSCA, M., PEACH, J. M., AND JONES, J. A. Approximate quantum cloning with nuclear magnetic resonance. *Phys. Rev. Lett.*, 88:187901, 2002. arXiv, quant-ph/0111098.

[CLK+00a] CORY, D. G., LAFLAMME, R., KNILL, E., VIOLA, L., HAVEL, T. F., N.BOULANT, BOUTIS, G., FORTUNATO, E., LLOYD, S., MARTINEZ, R., NEGREVERGNE, C., PRAVIA, M., SHARF, Y., TEKLEMARIAM, G., WEINSTEIN, Y. S., AND ZUREK, W. H. NMR based quantum information processing: Achievements and prospects. *Fort. der Phys. special issue, Experimental Proposals for Quantum Computation*, 48, 2000. arXiv, quant-ph/0004104.

[CRSS97a] CALDERBANK, A. R., RAINS, E. M., SHOR, P. W., AND SLOANE, N. J. A. Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.*, 78:405–408, 1997. arXiv, quant-ph/9605005.

[CT91a] COVER, T. M., AND THOMAS, J. A. *Elements of Information Theory*. Series in Telecommunication. John Wiley and Sons, New York, 1991.

[CTDL77a] COHEN-TANNOUDJI, C., DIU, B., AND LALÖE, F. *Quantum Mechanics*, volume I-II. John Wiley and Sons, 1977.

[DDZ01a] DZIARMAGA, J., DALVIT, D. A. R., AND ZUREK, W. H. Conditional dynamics of open quantum systems: The case of multiple observers. *Phys. Rev. Lett.*, 86:373–376, 2001. arXiv, quant-ph/0106036.

[Deu85a] DEUTSCH, D. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc. R. Soc. Lond. A*, 400:97–117, 1985.

[Deu89a] DEUTSCH, D. Quantum computational networks. *Proc. R. Soc. of Lond. A*, 425:73, 1989.

[Die82a] DIEKS, D. Communication by electron-paramagnetic-res devices. *Phys. Lett. A*, 92(6):271–272, 1982.

[Dir47a] DIRAC, P. A. M. *The Principles of Quantum Mechanics*. The international series of monographs on physics. Clarendon Press, Oxford, 4 edition, 1947.

[Div00a] DIVINCENZO, D. P. The physical implementation of quantum computation, 2000. arXiv, quant-ph/0002077.

[DJ92a] DEUTSCH, D., AND JOZSA, R. Rapid solution of problems by quantum computation. 439:553–558, 1992.

[DS94a] DIVINCENZO, D. P., AND SMOLIN, J. A. Results on two-bit gate design for quantum computers. In *Proceedings of the Workshop on Physics and Computation*, page 14, Los Alamitos, CA, 1994. IEEE Comput. Soc. Press.

[DS96a] DIVINCENZO, D. P., AND SHOR, P. W. Fault-tolerant error correction with efficient quantum codes. *Phys. Rev. Lett.*, 77:3260–3263, 1996. arXiv, quant-ph/9605031.

[DZB+94a] DAVIDOVICH, L., ZAGURY, N., BRUNE, M., RAIMOND, J.-M., AND HAROCHE, S. Teleportation of an atomic state between two cavities using nonlocal microwave fields. *Phys. Rev. A*, 50:895–898(R), 1994.

[Eve57a] EVERETT III, H. Relative state formulation of quantum mechanics. *Rev. Mod. Phys.*, 29(3):454–462, 1957.

[EJPP00a] EISERT, J., JACOBS, K., PAPADOPOULOS, P., AND PLENIO, M. B. Optimal local implementation of nonlocal quantum gates. *Phys. Rev. A*, 62:52317, 2000.

[EPR35a] EINSTEIN, A., PODOLSKY, B., AND ROSEN, N. Can quantum-mechanical description of reality be considered complete? *Phys. Rev.*, 47(10):777–780, May 1935.

[Fey82a] FEYNMAN, R. P. Simulating physics with computers. *Int. J. of Theor. Phys.*, 21:467, 1982.

[Fey84a] FEYNMAN, R. P. Quantum-mechanical computers. *Journal of the Optical Society of America B-Optical Physics*, 1:464, 1984.

[Fey86a] FEYNMAN, R. P. Quantum-mechanical computers. *Foundations of Physics*, 16:507–531, 1986.

[FGR+02a] FASEL, S., GISIN, N., RIBORDY, G., SCARANI, V., AND ZBINDEN, H. Quantum cloning with an optical fiber amplifier, 2002.

[FIMC01a] FIURASEK, J., IBSIDIR, S., MASSAR, S., AND CERF, N. J. Quantum cloning of orthogonal qubits, 2001. arXiv, quant-ph/0110016.

[FP00a] FUCHS, C. A., AND PERES, A. Quantum theory needs no 'interpretation'. *Phys. Today*, 53:70, 2000.

[FT82a] FREDKIN, E., AND TOFFOLI, T. Conservative logic. *Int. J. Theor. Phys.*, 21:219–253, 1982.

[Fuc96a] FUCHS, C. A. *Distinguishability and Accessible Information in Quantum Theory*. PhD thesis, The University of New Mexico, Albuquerque, NM, 1996. arXiv, quant-ph/9601020.

[Fuc98a] FUCHS, C. A. Information gain vs. state disturbance in quantum theory. *Fortsch. Phys.*, 46:535–565, 1998. arXiv, quant-ph/9611010.

[Fuc02c] FUCHS, C. Quantum mechanics as quantum information (and only a little more), 2002. arXiv, quant-ph/0205039.

[Fuc02b] FUCHS, C. A. Quantum foundations in the light of quantum information. In A. Gonis, Ed., *2001 NATO Advanced Research Workshop "Decoherence and its implications in quantum computation and information transfer"*, Mikonos, Greece, 2002. arXiv, quant-ph/0106166.

[GC97a] GERSHENFELD, N., AND CHUANG, I. L. Bulk spin-resonance quantum computation. 275:350–356, 1997.

[GG01b] GROSSHANS, F., AND GRANGIER, P. Quantum cloning and teleportation criteria for continuous quantum variables. *Phys. Rev. A*, 64:010301/1–4, 2001.

[GH00a] GALVÃO, E. F., AND HARDY, L. Building multiparticle states with teleportation. *Phys. Rev. A*, 62:12309, 2000. arXiv, quant-ph/9906080.

[Giu00a] GIULINI, D. Decoherence, a dynamical approach to superselection rules?, 2000.

[GJK+96a] GIULINI, D., JOOS, E., KIEFER, C., KUPSCH, J., STAMATESCU, I.-O., AND ZEH, H. D. *Decoherence and the Appearance of a Classical World in Quantum Theory*. Springer, Berlin, 1996.

[GM97a] GISIN, N., AND MASSAR, S. Optimal quantum cloning machines. *Phys. Rev. Lett.*, 79:2153, 1997.

[Got97a] GOTTESMAN, D. *Stabilizer codes and quantum error correction*. PhD thesis, California Institute of Technology, Pasadena, CA, 1997. arXiv, quant-ph/9705052.

[Got98a] GOTTESMAN, D. Fault tolerant quantum computation with higher dimensional systems, 1998. arXiv, quant-ph/9802007.

[Gro96a] GROVER, L. A fast quantum mechanical algorithm for database search. In *Proc. 28th Annual ACM Symposium on the Theory of Computation*, pages 212–219, New York, NY, 1996. ACM Press, New York. arXiv, quant-ph/9605043.

[Gro97a] GROVER, L. K. Quantum computers can search arbitrarily large databases by a single query. *Phys. Rev. Lett.*, 79:4709–4712, 1997.

[GS01a] GEORGEOT, B., AND SHEPELYANSKY, D. L. Exponential gain in quantum computing of quantum chaos and localization. *Phys. Rev. Lett.*, 86:2890, 2001. arXiv, quant-ph/0010005.

[Hal99b] HALLIWELL, J. J. Somewhere in the universe: Where is the information stored when histories decohere? *Phys. Rev. D*, 60:105031, 1999.

[Har00a] HARDY, L. Can we obtain quantum theory from reasonable axioms?, 2000. arXiv, quant-ph/0010083.

[Har01a] HARDY, L. Quantum theory from five reasonable axioms, 2001. arXiv, quant-ph/0101012.

[HH02a] HUGHES, R., AND HEINRICHS, T. Quantum information science and technology roadmap. Technical Report LA-UR-02-6900, Los Alamos National Laboratory, Los Alamos, NM, 2002. Collective work, web-accessible at `http://qist.lanl.gov`.

[HHB+04a] HACKERMÜLLER, L., HORNBERGER, K., BREZGER, B., ZEILINGER, A., AND ARNDT, M. Decoherence of matter waves by thermal emission of radiation. *Nature*, 427:711, 2004.

[HLL+01a] HUANG, Y.-F., LI, W.-L., LI, C.-F., ZHANG, Y.-S., JIANG, Y.-K., AND GUO, G.-C. Optical realization of quantum cloning. *Phys. Rev. A*, 64(1):12315, July 2001. arXiv, quant-ph/0006032.

[Hol73a] HOLEVO, A. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:177–183, 1973.

[HSTC99a] HAVEL, T. F., SOMAROO, S. S., TSENG, C.-H., AND CORY, D. G. Principles and demonstrations of quantum information processing by NMR spectroscopy, 1999. arXiv, quant-ph/9812086.

[JL02a] JOZSA, R., AND LINDEN, N. On the role of entanglement in quantum computational speed-up, 2002. arXiv, quant-ph/0201143.

[JMH98a] JONES, J. A., MOSCA, M., AND HANSEN, R. H. Implementation of a quantum search algorithm on a quantum computer. *Nature*, 392:344–346, 1998.

[Joz94a] JOZSA, R. Fidelity for mixed quantum states. *J. Mod. Optics*, 41:2315, 1994.

[JZ99a] JOHANNESSON, R., AND ZIGANGIROV, K. *Fundamentals of Convolutional Coding.* Digital and Mobile Communication. IEEE press, 1999.

[KBLW01a] KEMPE, J., BACON, D., LIDAR, D. A., AND WHALEY, K. B. Theory of decoherence-free fault-tolerant universal quantum compuation. *Phys. Rev. A*, 63:42307, 2001. arXiv, quant-ph/0004064.

[Ken90a] KENT, A. Against many-worlds interpretations. *Int. J. Mod. Phys.*, 5:1745–1762, 1990.

[KLM01a] KNILL, E., LAFLAMME, R., AND MILBURN, G. J. A scheme for efficient linear optics quantum computation. *Nature*, 409:46, January 2001.

[KLMN01a] KNILL, E., LAFLAMME, R., MARTINEZ, R., AND NEGREVERGNE, C. Benchmarking quantum computers: hte five-qubit error correcting code. *Phys. Rev. Lett.*, 86(25):5811–5814, 2001. arXiv, quant-ph/0101034.

[Kni96b] KNILL, E. Group representations, error bases and quantum codes, 1996. arXiv, quant-ph/9608049.

[Kra83a] KRAUS, K. *States, Effects and Operations.* Fundamental Notions of Quantum Theory. Academic Press, Berlin, 1983.

[KSV02a] KITAEV, A. Y., SHEN, A. H., AND VYALYI, M. N. *Classical and quantum computation.* Graduate studies in mathematics. American Mathematical Society, Providence, Rhodes Island, 2002.

[KSW00a] KEMPE, J., SIMON, C., AND WEIHS, C. Lambda's v's and and optimal cloning with stimulated emission, 2000. arXiv, quant-ph/0003025.

[Lan61a] LANDAUER, R. Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.*, 5:183–191, 1961.

[Lan91a] LANDAUER, R. Information is physical. *Phys. Today*, 44(5):23, 1991.

[LBKW99a] LIDAR, D. A., BACON, D., KEMPE, J., AND WHALEY, K. B. Universal fault tolerant quantum computation on a class of decoherence-free subspaces without spatial symmetry, 1999. arXiv, quant-ph/9908064.

[Lee97a] LEE, L. H. C. *Convolutional coding: fundamentals and applications*. Artech House Publishers, Boston, MA, 1997. TK5103.7.L433 1997.

[LSHB02a] LAMAS-LINARES, A., SIMON, C., HOWELL, J., AND BOUWMEESTER, D. Experimental quantum cloning of single photons. *Science*, 286:712, march 2002. 10.1126/science.1068972.

[Lo00a] LO, H.-K. Classical communication cost in distributed quantum information processing: a generalization of quantum-communication complexity. *Phys. Rev. A*, 62:12313, 2000. arXiv, quant-ph/9912009.

[LV00a] LLOYD, S., AND VIOLA, L. Control of open quantum system dynamics, 2000. arXiv, quant-ph/0008101.

[MHN+97a] MAÎTRE, X., HAGLEY, E., NOGUES, G., WUNDERLICH, C., GOY, P., BRUNE, M., RAIMOND, J. M., AND HAROCHE, S. Quantum memory with a single photon in cavity. *Phys. Rev. Lett.*, 79(4):769, 1997.

[MMK+95a] MONROE, C., MEEKHOF, D. M., KING, B. E., ITANO, W. M., AND WINELAND, D. J. Demonstration of a universal quantum logic gate. *Phys. Rev. Lett.*, 75:4714–4717, 1995.

[Moo65a] MOORE, G. Cramming more components onto integrated circuits. *Electronics*, 38(8), 1965.

[Moo97a] MOORE, G. An update on moore's law. Speech given at the INTEL Developer's Forum, 1997.

[MOR03a] MILMAN, P., OLLIVIER, H., AND RAIMOND, J. M. Universal quantum cloning in cavity QED. *Phys. Rev. A*, 67:12314, 2003. arXiv, quant-ph/0207039.

[MOY+03a] MILMAN, P., OLLIVIER, H., YAMAGUCHI, Y., BRUNE, M., RAIMOND, J.-M., AND HAROCHE, S. Simple quantum information algorithms in cavity QED. *J. Mod. Opt.*, 50(6-7):901–913, 2003.

[MS68a] MASSEY, J. L., AND SAIN, M. K. Inverses of linear sequential circuits. *IEEE Trans. Comp.*, C-17(4):330–337, 1968.

[NAZ02a] NAIRZ, O., ARNDT, M., AND ZEILINGER, A. Experimental verification of the heisenberg uncertainty principle for fullerene molecules. *Phys. Rev. A*, 65:032109, 2002.

[NC00a] NIELSEN, M. A., AND CHUANG, I. L. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, UK, 2000.

[neu43a] NEUMARK, M. A. On a representation of additive operator set functions. *Dokl. Acad. Sci. USSR*, 41:359–361, 1943.

[Neu54a] NEUMARK, M. A. Operatorenalgebren im hilbertschen raum. In *Sowjetische Arbeiten zur Funktionalanalisys*, Berlin, 1954. Verlag Kultur und Fortscritt.

[NRO+99a] NOGUES, G., RAUSCHENBEUTEL, A., OSNAGHI, S., BRUNE, M., RAIMOND, J., AND HAROCHE, S. Seeing a single photon without destroying it. *Nature*, 400:239–242, 1999.

[OBA+01a] OSNAGHI, S., BERTET, P., AUFFEVES, A., MAIOLI, P., BRUNE, M., RAIMOND, J. M., AND HAROCHE, S. Coherent control of an atomic collision in a cavity. *Phys. Rev. Lett.*, 87:37902, 2001. arXiv, quant-ph/0105063.

[OM03a] OLLIVIER, H., AND MILMAN, P. Proposal for realization of a toffoli gate via cavity-assisted collision. *Quant. Info. Comput. J.*, 6, 2003. arXiv, quant-ph/0306064.

[OPZ03a] OLLIVIER, H., POULIN, D., AND ZUREK, W. H. Objective properties from subjective quantum states: Environment as a witness, 2003. arXiv, quant-ph/0307229.

[OT03a] OLLIVIER, H., AND TILLICH, J.-P. Description of a quantum convolutional code. *Phys. Rev. Lett.*, 91(17):177902, 2003. arXiv, quant-ph/0304189.

[OT04a] OLLIVIER, H., AND TILLICH, J.-P. Quantum convolutional codes: fundamentals, 2004. arXiv, quant-ph/0401134. Submitted to IEEE Trans on Info. Theor.

[OZ02a] OLLIVIER, H., AND ZUREK, W. H. Quantum discord: A measure of the quantumness of correlations. *Phys. Rev. Lett.*, 88:17901, 2002. arXiv, quant-ph/0105072.

[PBKLO04a] POULIN, D., BLUME-KOHOUT, R., LAFLAMME, R., AND OLLIVIER, H. Exponential speed-up with a single bit of quantum information: Testing the quantum butterfly effect. *Phys. Rev. Lett.*, 2004.

[Per93a] PERES, A. *Quantum Theory: Concepts and Methods.* Kluwer Academic, Dordecht, 1993.

[PHZ93a] PAZ, J.-P., HABIB, S., AND ZUREK, W. H. Reduction of the wave packet: Preferred observable and decoherence time-scale. *Phys. Rev. D*, 47:488–501, 1993.

[Prea] PRESKILL, J. Lecture notes for physics 229: Quantum information and computation.

[Preb] PRESKILL, J. Lecture notes for physics 229: Quantum information and computation (ch.7).

[Pre98a] PRESKILL, J. Reliable quantum computers. *Proc. R. Soc. Lond. A*, 454:385, 1998.

[PSD+99a] PRICE, M. D., SOMAROO, S. S., DUNLOP, A. E., HAVEL, T. F., AND CORY, D. G. Generalized methods for the development of quantum logic gates for an NMR quantum information processor. *Phys. Rev. A*, 60:2777–2780, 1999.

[PZ01a] PAZ, J.-P., AND ZUREK, W. H. Environment-induced decoherence and the transition from quantum to classical. In R. Kaiser, C. Westbrook, and F. David, Eds., *Coherent Atomic Matter Waves*, Les Houches Lectures, pages 533–614. Springer, Berlin, 2001. arXiv, quant-ph/0010011.

[Raz99a] RAZ, R. Exponential separation of quantum and classcial communication complexity. *Proceedings of the 31st ACM Symposium on Theory of Computing*, pages 358–367, 1999.

[RBH01a] RAIMOND, J. M., BRUNE, M., AND HAROCHE, S. Colloquium: Manipulating quantum entanglement with atoms and photons in a cavity. *Rev. Mod. Phys.*, 73:565, 2001.

[RBO+01a] RAUSCHENBEUTEL, A., BERTET, P., OSNAGHI, S., NOGUES, G., BRUNE, M., RAIMOND, J. M., AND HAROCHE, S. Controlled entanglement of two field modes in a cavity quantum electrodynamics experiment. *Phys. Rev. A*, 64:050301(R), 2001.

[RNO+99a] RAUSCHENBEUTEL, A., NOGUES, G., OSNAGHI, S., BERTET, P., BRUNE, M., RAIMOND, J. M., AND HAROCHE, S. Coherent operation of a tunable quantum phase gate in cavity QED. *Phys. Rev. Lett.*, 83:5166, 1999.

[Sha48a] SHANNON, C. E. A mathematical theory of communication. *Bell System Tech. J.*, 27:379, 623, 1948.

[Sho94a] SHOR, P. W. Algorithms for quantum computation: Discrete logarithms and factoring. In S. Goldwasser, Ed., *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, pages 124–134, Los Alamitos, CA, 1994. IEEE Computer Society.

[Sho95a] SHOR, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:2493, 1995.

[Sho96a] SHOR, P. W. Fault-tolerant quantum computation. In *Proceedings of the 37th Symposium on the Foundations of Computer Science*, pages 56–65, Los Alamitos, California, 1996. IEEE press. arXiv, quant-ph/9605011.

[Sim94a] SIMON, D. On the power of quantum computation. In *Proc. 26th STOC*, pages 116–123, 1994. See also [Sim97a].

[Sim97a] SIMON, D. On the power of quantum computation. *SIAM J. Comp.*, 26(5):1474–1483, 1997.

[SN96a] SCHUMACHER, B., AND NIELSEN, M. A. Quantum data processing and error correction. *Phys. Rev. A*, 54:2629, 1996. quant-ph/9604022.

[SP00a] SHOR, P. W., AND PRESKILL, J. Simple proof of security of BB84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85:441–444, 2000. arXiv, quant-ph/0003004.

[SV98a] SCHULMAN, L. J., AND VAZIRANI, U. Scalable NMR quantum computation. In *Proceedings of the 31th Annual ACM Symposium on the Theory of Computation (STOC)*, pages 322–329, El Paso, Texas, 1998. ACM Press. arXiv, quant-ph/9804060.

[SWZ00a] SIMON, C., WEIHS, G., AND ZEILINGER, A. Optimal quantum cloning via stimulated emission. *Phys. Rev. Lett.*, 84:2993, 2000.

[SZ02a] SCULLY, M. O., AND ZUBAIRY, M. S. Cavity QED implementation of the discrete quantum Fourier transform. *Phys. Rev. A*, 65:52324, 2002.

[Szi29a] SZILARD, L. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Z. Phys.*, 53:840, 1929. Reprinted in [WZ83a].

[THL+95a] TURCHETTE, Q. A., HOOD, C. J., LANGE, W., MABUCHI, H., AND KIMBLE, H. J. Measurement of conditional phase shifts for quantum logic. *Phys. Rev. Lett.*, 75:4710–4713, 1995.

[Tur36a] TURING, A. M. On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc. 2*, 42:230–265, 1936.

[Vit67a] VITERBI, A. J. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Inf. Theor.*, 13(2):260–269, 1967.

[VKL99a] VIOLA, L., KNILL, E., AND LLOYD, S. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.*, 82:2417, 1999. arXiv, quant-ph/9809071.

[VKL00a] VIOLA, L., KNILL, E., AND LLOYD, S. Dynamical generation of noiseless quantum subsystems. *Phys. Rev. Lett.*, 85:3520, 2000. arXiv, quant-ph/0002072.

[vN32a] VON NEUMANN, J. *Mathematische Grundlagen der Quanten-mechanik.* Julius Springer-Verlag, 1932.

[Weh78a] WEHRL, A. General properties of entropy. *Rev. Mod. Phys.*, 50:221–261, 1978.

[Wer89a] WERNER, R. F. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40(8):4277–4281, October 1989.

[Whe57a] WHEELER, J. A. Assesment of everett's "relative state" formulation of quantum theory. *Rev. Mod. Phys.*, 29(3):463–465, 1957.

[WSM01a] WANG, X., SORENSEN, A., AND MOLMER, K. Multibit gates for quantum computing. *Phys. Rev. Lett.*, 86:3907, 2001.

[WZ82a] WOOTTERS, W. K., AND ZUREK, W. H. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.

[WZ83a] WHEELER, J. A., AND ZUREK, W. H., Eds. *Quantum theory and measurement.* Plenum Press, Princeton, 1983.

[YMB+02a] YAMAGUCHI, F., MILMAN, P., BRUNE, M., RAIMOND, J.-M., AND HAROCHE, S. Quantum search with two atoms collision in cavity QED. *Phys. Rev. A*, 66:010302(R), 2002. arXiv, quant-ph/0203146.

[Zal96a] ZALKA, C. Threshold estimate for fault tolerant quantum computation, 1996. arXiv, quant-ph/9612028.

[Zan99a] ZANARDI, P. Computation on a noiseless quantum code and symmetrization. *Phys. Rev. A*, 60:R729, 1999. arXiv, quant-ph/9901047.

[Zeh73a] ZEH, H. D. Toward a quantum theory of observation. *Found. Phys.*, 3:109, 1973. arXiv, quant-ph/0306151.

[ZG00a] ZHENG, S.-B., AND GUO, G.-C. Efficient scheme for two-atom entanglement and quantum information processing in cavity QED. *Phys. Rev. Lett.*, 85:2392, 2000.

[Zhe01a] ZHENG, S.-B. One-step synthesis of multiatom Greenberger-Horne-Zeilinger states. *Phys. Rev. Lett.*, 87:230404, 2001.

[ZPM03a] ZOU, X.-B., PAHLKE, K., AND MATHIS, W. Scheme for the implementation of a universal quantum cloning machine via cavity-assisted atomic collisions in cavity qed. *Phys. Rev. A*, 67:24304, 2003. arXiv, quant-ph/0303015.

[ZR97b] ZANARDI, P., AND RASETTI, M. Error avoiding quantum codes. 11:1085, 1997. arXiv, quant-ph/9710041.

[ZR97c] ZANARDI, P., AND RASETTI, M. Noiseless quantum codes. *Phys. Rev. Lett.*, 79:3306–3309, 1997. arXiv, quant-ph/9705044.

[Zur81a] ZUREK, W. H. Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? *Phys. Rev. D*, 24:1516, 1981.

[Zur82a] ZUREK, W. H. Environment-induced superselection rules. *Phys. Rev. D*, 26:1862–1880, 1982.

[Zur83a] ZUREK, W. H. Information transfer in quantum measurements: Irreversibility and amplification. In P. Meystre and M.O. Scully, Eds., *Quantum optics and experimental gravitation and measurement theory*, pages 87–116, New York, NY, 1983. Plenum. arXiv, quant-ph/0111137.

[Zur91a] ZUREK, W. H. Decoherence and the transition from quantum to classical. *Physics Today*, 44(10):36–44, 1991.

[Zur93b] ZUREK, W. H. Preferred states, predictability, classicality and the environment-induced decoherence. *Prog. of Theor. Phys.*, 89:281–312, 1993.

[Zur94a] ZUREK, W. H. *Decoherence and the existential interpretation of quantum theory, or "No information without representation"*, pages 341–350. Plenum, Dordrecht, 1994.

[Zur98a] ZUREK, W. H. Decoherence, einselection, and the existential interpretation (the rough guide). *Phil. Trans. R. Soc. Lond. A*, 356:1793, 1998. arXiv, quant-ph/9805065.

[Zur00a] ZUREK, W. H. Einselection and decoherence from an information theory perspective. *Ann. Phys.*, 9:855–864, 2000.

[Zur02a] ZUREK, W. H. Decoherence and the transition from quantum to classical – revisited. *Los Alamos Science*, 27, 2002. arXiv, quant-ph/0306072.

[Zur03a] ZUREK, W. H. Decoherence, einselection and the quantum origins of the classical. *Rev. Mod. Phys.*, 75:715–775, 2003. arXiv, quant-ph/0105127.

[Zur03b] ZUREK, W. H. Quantum darwinism and envariance, 2003. arXiv, quant-ph/0308163.

[Zur91a] ZUREK, W. H. Decoherence and the transition from quantum to classical. *Physics Today*, 44(10):36–44, 1991.

[Zur93b] ZUREK, W. H. Preferred states, predictability, classicality and the environment-induced decoherence. *Prog. of Theor. Phys.*, 89:281–312, 1993.

[Zur94a] ZUREK, W. H. *Decoherence and the existential interpretation of quantum theory, or "No information without representation"*, pages 341–350. Plenum, Dordrecht, 1994.

[Zur98a] ZUREK, W. H. Decoherence, einselection, and the existential interpretation (the rough guide). *Phil. Trans. R. Soc. Lond. A*, 356:1793, 1998. arXiv, quant-ph/9805065.

[Zur00a] ZUREK, W. H. Einselection and decoherence from an information theory perspective. *Ann. Phys.*, 9:855–864, 2000.

[Zur02a] ZUREK, W. H. Decoherence and the transition from quantum to classical – revisited. *Los Alamos Science*, 27, 2002. arXiv, quant-ph/0306072.

[Zur03a] ZUREK, W. H. Decoherence, einselection and the quantum origins of the classical. *Rev. Mod. Phys.*, 75:715–775, 2003. arXiv, quant-ph/0105127.

[Zur03b] ZUREK, W. H. Quantum darwinism and envariance, 2003. arXiv, quant-ph/0308163.