SLOVENSKÁ AKADÉMIA VIED
FYZIKÁLNY ÚSTAV
BRATISLAVA

DIZERTAČNÁ PRÁCA

2003 Mário Ziman

SLOVAK ACADEMY OF SCIENCIES
INSTITUTE OF PHYSICS
BRATISLAVA

Research Center for Quantum Information

Entanglement as a quantum structure: Applications to information processing

(PhD thesis)

Date:	March, 2003
PhD student:	Mário Ziman
Supervisor:	prof. Vladimír Bužek

Acknowledgment

I would like to express my thanks to my supervisor Vladimír Bužek, and to my colleagues, especially to Peter Štelmachovič, Marek Šašura and Martin Plesch. They created a very friendly and creative atmosphere, in which it was pleasure to work. I am glad that the years of my study I could spend among them. I wish I will never forget the things I have learnt, the things I have lived and I hope that most of the gathered experiences will be useful in my future.

I can't forget the people - unphysicists, like my Family and Friends. Thanks to them I fully enjoyed also my leisure time. Finally, I would like to thank to the only person I surely know that read the whole Thesis, to Karla. ☺

Contents

Introduction

Quantum Theory entered the world of Science at the beginning of the twentieth century. It satisfactorily explained a lot of phenomena: description of the spectral lines of the hydrogen atom, the black body radiation, the transistor effect, etc. Without any doubts we can say that Quantum Theory has been succeeding. Although there is no contradiction between its predictions and experimental reality, sometimes the vain effort to understand its rules makes the Quantum Theory "mysterious". On one hand, we wonder at its precision, but on the other hand, the comprehension of basic principles have remained misty. Simply, it is unnatural to see the world via "quantum eyes".

At first, its probabilistic essence is in deep disharmony with the classical determinism. Foundations of the Probability Theory were established by *Kolmogorov* and now the concept of probability is widely used in many different areas of Science. Quantum Theory is not an exception. However the properties of "quantum probabilities" are exceptional. To understand the phenomena such as *wave-particle duality*, or *uncertainty relations*, one needs to start thinking in the language of probabilities. Our intuition is subjected to probabilities that, in principle, cannot be eliminated.

All the differences between quantum and classical statistics originate in the *superposition principle* Combining this principle with the description of a system consisting of two distinguishable particles, we obtain states with very peculiar properties. These states can be used to violate *Bell inequalities*, to transmit two bits of information with two-level signals (*superdense coding*), to teleport a state of quantum system (*quantum teleportation*), or to distribute a cryptographic key (*quantum key distribution*) in a secure way. *Erwin Schrödinger* introduced the word *entangled particles* to name this feature of quantum states. Today the concept of entanglement is investigated in order to understand basic properties of quantum world. It was recognized that *entanglement* stands behind many purely quantum phenomena, but its role has not still been satisfactorily specified.

The development of the computers is facing with the problem of minimalization, and as we know, the adequate theory of the microworld is Quantum Theory with its own rules. Therefore, in the last decades all aspects of Quantum Theory have been studied and analyzed in order to provide us new trends in the information processing. A real boom has started after the discovery of the *Shor algorithm*, that solves an old algorithmic problem of the *prime factorization* in an efficient way. To run this algorithm one needs a device ruled by the laws of Quantum Theory, i.e. a *quantum computer*. Proposals of the experimental realization of this idea (*trapped ions, quantum dots, cavities QED, etc.*) are now studied and tested in laboratories all around the world. To transform this idea into reality we must avoid the phenomenon of *decoherence*, which is responsible for the destruction of quantum coherence (superposition) and, consequently, the quantum advantages cannot be exploited. In a sense the entanglement stands also behind this phenomenon, because the decoherence is caused by interactions of the system with an environment, in which the system and the environment become superposed (entangled). Here one can see a very strange feature of the entanglement. Compared with the correlations (often used as an analogy of the entanglement) the entanglement cannot be shared among particles freely (*Coffman-Kundu-Wootters inequalities*). If one particle starts to entangle with another, then all its previous "entanglements" decrease. It is very important to understand the behavior of the entanglement in multi-partite systems. This investigation can give us a better insight how to "coherently" control the dynamics of open quantum systems that encode the information we want to manipulate.

In this thesis we shall introduce the entanglement as a structure of quantum states having its consequences also in quantum observables and quantum dynamics. In detail we shall study the following topics: *Quantum dense coding*, *Quantum homogenization*, *Quantum version of processor*, *Entanglement and correlations in multipartite systems.* This thesis covers a few separated areas of investigation in the field of *Quantum Information Theory* that connects two different languages of physicists and computation scientists.

First two chapters introduce elements of Quantum Theory and Information Theory. They are written without any references, because they contain the general knowledge, which can be found in usual textbooks, or review articles [1]-[10]. The mathematical formalism of Quantum Theory is demonstrated on the example of the two-dimensional system called *qubit.* The notion of *information* is introduced as a measure of correlations between the outcomes of two observables. The question of classical communication is addressed at the end of the second chapter. In the whole thesis we pay attention only to finite sample spaces and finite dimensional Hilbert spaces, i.e. the difficulties of the formalism with infinite spaces are omitted. The third chapter uses the correlation properties of quantum measurements to introduce the notion of *entanglement.* The properties of the entanglement between two qubits are studied in detail and also the measure of entanglement for this case is explicitly given.

The fourth chapter concerns the problem of the dynamics of open quantum systems. We analyze the general evolution maps and time dependence of quantum states. We introduce a specific collision model between a qubit and a reservoir (*homogenization process*), that transforms the qubit into the average state of the reservoir. The dynamics plays a central role in the transmission of information. The usage of quantum states to represent (encode) the information is investigated and the derivation of the general formulas for the capacity of various communication protocols is given.

The fifth chapter continues the study of quantum dynamics, but from another point of view. We investigate the properties of the so-called *quantum processors.* These devices are defined in analogy with the classical processors, i.e. they control the evolution of one system (*data*) with the help of the second system (*program*). From the mathematical point of view the quantum processor is just a unitary transformation defined on a composite Hilbert space of program and data registers. The quantum processor can be used either in deterministic, or probabilistic regime. We will show that in the latter case a “universal” processor (= realizes all quantum operations) can be designed, whereas its deterministic counterpart does not exist. The quantum processors can be used to realize not only quantum maps, but also quantum measurements (POVM). We will show an explicit example how to exploit a quantum processor to perform the complete state reconstruction.

The last chapter turns back to the problem of entangled states, but the problem is shifted from bipartite to multi-partite systems. Hence, the question of the classification and quantification of multi-partite entanglement is open and studied. The study of the properties of multi-partite entanglement is just at the beginnings and conceptually there are many open questions. We introduce the graphical representation of the entanglement and correlations shared in composite systems between individual parties. We use this representation to divide the state space into the classes of states representable by the same graph. This is only the first step in the comprehension of the analysis of the entanglement in composite quantum systems.

Chapter 1

Basics of quantum mechanics

Never judge a theory by it's formalism.
D.Hilbert

1.1 Mathematical tools

The central notion of quantum theory is the *Hilbert space* $\mathcal{H}$ used for the description of quantum objects.

Definition.

Hilbert space is a complex linear space with defined scalar product $(.|.)$ *complete in the norm* $||\psi|| \equiv \sqrt{(\psi|\psi)}$*. The scalar product needs to satisfy the following conditions*

1. *Linearity:* $(\psi|\phi_1 + \lambda\phi_2) = (\psi|\phi_1) + \lambda(\psi|\phi_2)$
2. *Symmetry:* $(\psi|\phi) = \overline{(\phi|\psi)}$
3. *Nondegeneracy:* $(\psi|\phi) = 0 \;\; \forall\psi \;\Rightarrow\; \phi = 0$
4. *nonnegativity:* $(\psi|\psi) \geq 0$

Moreover, the Hilbert spaces used in quantum theory has to be *separable*, i.e. the orthogonal basis is countable. In this thesis we shall deal only with finite dimensional spaces, which are trivially separable.

Strictly speaking, not all elements ψ of $\mathcal{H}$ are needed in quantum theory. Define a so-called *projective* Hilbert space $\mathcal{P}(\mathcal{H})$ consisting of elements called *rays.* Two vectors ψ, ϕ represent the same ray, if $\psi = \lambda\phi$ for some $\lambda \in \mathbb{C}$. In quantum mechanics we choose such representatives of rays that are *normalized*, i.e. $(\psi|\psi) = 1$. Each element of $\mathcal{P}(\mathcal{H})$ corresponds to a *quantum state* (we exclude zero vector from this set). Although not every quantum state can be described by a ray.

By a *linear operator* we will understand a mapping $\mathbf{A} : \mathcal{H} \to \mathcal{H}$ satisfying $\mathbf{A}(\psi + \lambda\phi) = \mathbf{A}\psi + \lambda\mathbf{A}\phi$. To make such operator useful for our purposes in quantum theory we also require that the *domain of definition* $D(\mathbf{A})$ is *dense* in the Hilbert space $\mathcal{H}$ in the norm induced by scalar product. For separable spaces it means that the mapping is determined by its action on a fixed *complete orthonormal basis* $\{\psi_n\}$ $((\psi_m|\psi_n) = \delta_{mn})$ with n taking its values from the countable index set $J \subset \mathbb{N}$. The action of $\mathbf{A}$ can be extended to all elements of $\mathcal{H}$ by linearity, because each vector $\phi \in \mathcal{H}$ can be expressed in the form $\phi = \sum_n a_n\psi_n$ with $a_n := (\psi_n|\phi)$. For finite dimensions the linear operator $\mathbf{A}$ can be represented by a matrix with coefficients $(\mathbf{A})_{mn} := (\psi_m|\mathbf{A}|\psi_n)$ and to each vector ϕ it corresponds a column of complex numbers a_n.

The **norm** of the operator $\mathbf{A}$ is defined by formula

$$||\mathbf{A}|| := \sup_{\phi \in \mathcal{H}} \frac{||\mathbf{A}\phi||}{||\phi||} \tag{1.1}$$

We call an operator *bounded*, if the norm is finite, i.e. $||\mathbf{A}|| < \infty$. The set of bounded operators $\mathcal{L}(\mathcal{H})$ has the *Banach space* structure.

Definition

Banach space *is a linear space with a defined norm* $||.||$ *and complete in this norm. The norm is determined by conditions*

1. *Nonnegativity:* $||\mathbf{A}|| \geq 0$
2. *Nondegeneracy:* $||\mathbf{A}|| = 0 \quad \Rightarrow \quad \mathbf{A} = \mathbb{O}$
3. *Homogenity:* $||\lambda\mathbf{A}|| = |\lambda|.||\mathbf{A}||$
4. *Triangle inequality:* $||\mathbf{A} + \mathbf{B}|| \leq ||\mathbf{A}|| + ||\mathbf{B}||$

The set of bounded linear functionals $\mathcal{H}^*$ (*topological dual of* $\mathcal{H}$) is the set of linear mappings $f : \mathcal{H} \to \mathbb{C}$ satisfying

$$|f| := \sup_{\psi \in \mathcal{H}} \frac{|f(\psi)|}{||\psi||} < \infty$$

Let us remind that the domain of definition $D(f)$ must be dense in $\mathcal{H}$.

Riezs lemma

For each $f \in \mathcal{H}^*$ *there exists exactly one vector* $\psi_f \in \mathcal{H}$ *such that* $f(\phi) \equiv (\psi_f|\phi)$ *for all* $\phi \in D(f)$, *where* $D(f)$ *is dense in* $\mathcal{H}$.

Next, we define the *adjoint* operator $\mathbf{A}^\dagger$ corresponding to linear operator $\mathbf{A}$. We start with defining its domain of definition

$$D(\mathbf{A}^\dagger) = \{\psi \in \mathcal{H} : \phi \mapsto f_\psi(\phi) := (\psi|\mathbf{A}\phi),\ f_\psi \in \mathcal{H}^*\}$$

Applying Riezs lemma to our definition of $D(\mathbf{A}^\dagger)$ we can define the *adjoint* operator by the formula

$$\mathbf{A}^\dagger\psi := \psi_f \ \forall\psi \in D(\mathbf{A}^\dagger),\ \text{where } (\psi_f|\phi) := (\psi|\mathbf{A}\phi)\ \forall\phi \in D(\mathbf{A}) \tag{1.2}$$

In the finite dimensional case this is equivalent to

$$(\mathbf{A}^\dagger\psi|\phi) := (\psi|\mathbf{A}\phi) \equiv (\psi|\mathbf{A}|\phi) \quad \forall\psi,\phi \in \mathcal{H} \tag{1.3}$$

what means that all of the three expressions listed above are equivalent.

Maybe the most important type of the operator is the *selfadjoint* operator. The operator $\mathbf{A}$ is selfadjoint, if $\mathbf{A} = \mathbf{A}^\dagger$ for all $\psi \in D(\mathbf{A}) \equiv D(\mathbf{A}^\dagger)$. Let us consider finite dimensional Hilbert spaces. The *projective* operator is defined by condition $\mathbf{P} = \mathbf{P}^\dagger = \mathbf{P}^2$. If $(\psi, \mathbf{A}\psi) \geq 0$ for all $\psi \in \mathcal{H}$, then we say that the operator is *positive. Unitary* operator need to satisfy $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbb{1}$, where $\mathbb{1}$ is a *unit operator*, i.e. $\mathbb{1}(\psi) = \psi$ for all $\psi \in \mathcal{H}$. Unitary operators transform one complete basis onto another and preserve the scalar product, i.e. $(\psi|\phi) = (\mathbf{U}\psi|\mathbf{U}\phi)$ for all $\psi, \phi \in \mathcal{H}$. We denote by $\mathbf{A}^{-1}$ an *inverse* operator of $\mathbf{A}$. The property $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A}$ defines *normal* operators.

The *spectrum* of the normal operator $\mathbf{A}$ is a subset of complex numbers $\sigma(\mathbf{A}) \subset \mathbb{C}$ such that for each $\lambda \in \sigma(\mathbf{A})$ the inverse operator $(\mathbf{A} - \lambda\mathbb{1})^{-1}$ does not exist. In finite dimensional case all the elements of spectra are *eigenvalues*, i.e. to each $\lambda \in \sigma(\mathbf{A})$ there exists vector $\psi \in \mathcal{H}$, for which $\mathbf{A}\psi = \lambda\psi$. This vector is called an *eigenvector* ψ belonging to the eigenvalue λ. Note that the set of eigenvectors belonging to the same eigenvalue forms a closed subspace of $\mathcal{H}$. The *spectral theorem* tells that each normal operator $\mathbf{A}$ can be written in the form

$$\mathbf{A} = \sum_n \lambda_n \mathbf{P}_n \tag{1.4}$$

where $\mathbf{P}_n$ is the projector on the subspace spanned by eigenvectors ψ_{nk} belonging to the same eigenvalue λ_n, i.e. $\mathbf{P}_n = \sum_k \mathbf{P}_{\psi_{nk}}$. Using this theorem we can define the *operator function* of any normal operator

$$f(\mathbf{A}) := \sum_n f(\lambda_n)\mathbf{P}_n \tag{1.5}$$

where f is a function $f : \mathbb{C} \to \mathbb{C}$.

Let us choose a complete orthonormal basis $\{\phi_n\}$ in $\mathcal{H}$. The *trace* of the operator $\mathbf{A}$ is a mapping $\mathrm{Tr} : \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ given as

$$\mathrm{Tr}\mathbf{A} := \sum_{n=1}^{d} (\phi_n|\mathbf{A}|\phi_n) \tag{1.6}$$

and $d = \dim \mathcal{H}$. We list the properties of trace

- $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$
- $\mathrm{Tr}(\mathbf{A} + \lambda\mathbf{B}) = \mathrm{Tr}\mathbf{A} + \lambda\mathrm{Tr}\mathbf{B}$
- $\mathrm{Tr}\mathbf{A} = \mathrm{Tr}(\mathbf{UAU}^\dagger)$ for all unitary $\mathbf{U}$

The last property means that the trace operation is basis independent. Define the absolute value of the operator $|\mathbf{A}| := \sqrt{AA^\dagger}$. The set of *traceclass* operators $\mathcal{T}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$ consists of operators satisfying $\mathrm{Tr}|\mathbf{A}| < \infty$. This set $\mathcal{T}(\mathcal{H})$ form a Banach space with respect to the *trace norm* $||\mathbf{A}||_1 := \mathrm{Tr}|\mathbf{A}|$. *Density operators* form a subset $\mathcal{S}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H})$ of positive elements $\varrho \in \mathcal{T}(\mathcal{H})$ with a unit norm, i.e. $||\varrho||_1 = \mathrm{Tr}|\varrho| = \mathrm{Tr}\varrho = 1$.

Consider the scalar product defined on $\mathcal{L}(\mathcal{H})$ by relation

$$(\mathbf{A}|\mathbf{B}) := \mathrm{Tr}\mathbf{A}^\dagger\mathbf{B}. \tag{1.7}$$

The set of operators $\mathcal{T}_2(\mathcal{H})$ bounded in the induced norm $||\mathbf{A}||_2 := \mathrm{Tr}\mathbf{A}^\dagger\mathbf{A}$ has the Hilbert space structure and its elements are called *Hilbert-Schmidt* operators. Similarly, like in the case of $\mathcal{H}$ we can define the set of linear transformations $\Phi : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})$ acting on linear space of operators. *Superoperators* are specific linear mappings Φ transforming density matrices into density matrices, i.e. $\Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}_\Phi \subset \mathcal{S}(\mathcal{H})$.

Consider two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$. The *Cartesian product* of them denoted by $\mathcal{H}_1 \times \mathcal{H}_2$ consists of elements $\psi_{12} = [\psi_1, \psi_2]$ where $\psi_1 \in \mathcal{H}_1$ and $\psi_2 \in \mathcal{H}_2$. Let us define the linearity

$$[\psi_1, \psi_2] + \lambda[\phi_1, \phi_2] \equiv [\psi_1 + \lambda\phi_1, \psi_2 + \lambda\phi_2] \tag{1.8}$$

and scalar product

$$([\psi_1, \psi_2]|[\phi_1, \phi_2])_{12} := (\psi_1|\psi_2)_1 + (\phi_1|\phi_2)_2 \tag{1.9}$$

where $(.|.)_j$ is the scalar product defined on $\mathcal{H}_j$, j=1,2. The linear space $\mathcal{H}_1 \times \mathcal{H}_2$ embedded with these structures is abbreviated by $\mathcal{H}_1 \oplus \mathcal{H}_2$ and is called a *direct sum* of Hilbert spaces. In this case $\dim \mathcal{H}_1 \oplus \mathcal{H}_2 = \dim \mathcal{H}_1 + \dim \mathcal{H}_2$.

Next, define what linear combination in each of the Hilbert spaces means on $\mathcal{H}_1 \times \mathcal{H}_2$

$$[\psi_1 + \lambda\phi_1, \psi_2 + \mu\phi_2] := [\psi_1, \psi_2] + \lambda[\phi_1, \psi_2] + \mu[\psi_1, \phi_2] + \lambda\mu[\phi_1, \phi_2]. \tag{1.10}$$

The scalar product consistent with such structure is given by

$$([\psi_1,\psi_2]|[\phi_1,\phi_2]) \equiv (\psi_{12}|\phi_{12}) := (\psi_1|\phi_1)_1(\psi_2|\phi_2)_2 \tag{1.11}$$

The linearity of it means $(\psi_{12}+\lambda\phi_{12}|\xi_{12}) = (\psi_{12}|\xi_{12})+\lambda^*(\phi_{12}|\xi_{12})$. In this case the obtained structure $\mathcal{H}_1\otimes\mathcal{H}_2$ is called the *tensor product* of Hilbert spaces and $\dim\mathcal{H}_1\otimes\mathcal{H}_2 = (\dim\mathcal{H}_1)(\dim\mathcal{H}_2)$. Moreover, we are allowed to express the elements of $\mathcal{H}_1\otimes\mathcal{H}_2$ in a form of tensor product $\psi_1\otimes\psi_2 := [\psi_1,\psi_2]$. Such notation is more comfortable and the relations defined above become more natural, i.e.

$$\begin{aligned}(\psi_1+\lambda\phi_1)\otimes(\psi_2+\mu\phi_2) &= \psi_1\otimes\psi_2+\lambda\phi_1\otimes\psi_2+\mu\psi_1\otimes\phi_2+\lambda\mu\phi_1\otimes\phi_2 \\ (\psi_1\otimes\psi_2|\phi_1\otimes\phi_2) &= (\psi_1|\phi_1)_1(\psi_2|\phi_2)_2 \end{aligned}\tag{1.12}$$

Let's define *partial trace* and *partial transpose* operations. Both of the definitions will be given in the form of matrix elements. It means in a *fixed basis*. Suppose the Hilbert space composite from two Hilbert spaces $\mathcal{H}_1,\mathcal{H}_2$ as a tensor product of them with the basis $\{\phi_n\otimes\psi_k\}$, where $\{\phi_n\},\{\psi_k\}$ are complete orthonormal bases in $\mathcal{H}_1,\mathcal{H}_2$, respectively. Express the operator $\mathbf{A}$ in such basis of $\mathcal{H}_1\otimes\mathcal{H}_2$

$$(\mathbf{A})_{mn,kl} = (\phi_m\otimes\psi_k|\mathbf{A}|\phi_n\otimes\psi_l). \tag{1.13}$$

The *partial trace over the second subsystem* is a mapping $\mathrm{Tr}_2 : \mathcal{L}(\mathcal{H}_1\otimes\mathcal{H}_2)\to\mathcal{L}(\mathcal{H}_1)$ that to each operator $\mathbf{A}\in\mathcal{L}(\mathcal{H}_1\otimes\mathcal{H}_2)$ assigns an operator $\mathbf{A}_2$ on $\mathcal{H}_1$ with matrix elements

$$(\mathbf{A}_2)_{kl} := (\psi_k|\mathrm{Tr}_1\mathbf{A}|\psi_l) = \sum_{n=1}^{\dim\mathcal{H}_1}(\mathbf{A})_{nn,kl}. \tag{1.14}$$

The *partial transpose over the second subsystem* is a mapping $T_2 : \mathcal{L}(\mathcal{H}_1\otimes\mathcal{H}_2)\to\mathcal{L}(\mathcal{H}_1\otimes\mathcal{H}_2)$ defined as $\mathbf{A}\mapsto\mathbf{A}^{T_2}$ with matrix elements

$$(\mathbf{A}^{T_2})_{nm,kl} = (\mathbf{A})_{nm,lk}. \tag{1.15}$$

In the same way we can define the partial trace and partial transposition over the first subsystem. The partial trace unlike the partial transposition has one important property

$$\mathrm{Tr}_2[\mathbb{1}_1\otimes\mathbf{U}_2\mathbf{A}\mathbb{1}_1\otimes\mathbf{U}_2^\dagger] = \mathrm{Tr}_2\mathbf{A} \quad \text{for all} \quad \mathbf{U}_2\in\mathcal{U}(\mathcal{H}_2) \tag{1.16}$$

that is, the partial trace is independent of the chosen basis $\{\psi_k\}$ in $\mathcal{H}_2$, in which the operator $\mathbf{A}$ is expressed. As we said, the partial transpose does not have this property, but the eigenvalues of $\mathbf{A}^{T_2}$ are basis independent!

1.1.1 Dirac's notation and repetition

Now, we shall introduce the *Dirac notation* very often used in quantum theory. Formally, the step is very simple, but we have to be very careful with its usage in infinite dimensional cases. As we have said earlier, in this thesis **we shall work with finite dimensional Hilbert spaces.** Hence we skip these problems.

Let the element of projective Hilbert space, the ray, be denoted by *ket symbol* $|\psi\rangle\in\mathcal{H}$. We shall denote the elements of topological dual by *bra symbol* $\langle\psi|\in\mathcal{H}^*$. Here we use the Riezs theorem, which to each element of $\mathcal{H}^*$ assigns a vector from $\mathcal{H}$. For scalar product we get $(\psi|\phi) = \langle\psi|\phi\rangle$. Let $\{\psi_j\}\equiv\{|j\rangle\}$ be the orthonormal basis in $\mathcal{H}$. The operator $\mathbf{A}$ is given as $\mathbf{A} = \sum_{jk}(\mathbf{A})_{jk}|j\rangle\langle k|$ in this basis with $(\mathbf{A})_{jk} := \langle j|\mathbf{A}|k\rangle$. To each vector $|\psi\rangle$ it corresponds one dimensional projection onto subspace spanned by this vector $\mathbf{P} = |\psi\rangle\langle\psi|$, and vice versa. The completeness of the basis is equivalent to the relation $\sum_k|k\rangle\langle k| = \mathbb{1}$.

In the following we list the defined sets:

- $\mathcal{L}(\mathcal{H})$ ...... set of bounded linear operators, i.e. $||\mathbf{A}||<\infty$

- $\mathcal{L}_s(\mathcal{H})$ ...... set of selfadjoint operators, i.e. $\mathbf{A} = \mathbf{A}^\dagger$
- $\mathcal{L}_+(\mathcal{H})$ ...... set of positive operators, i.e. $\mathbf{A} = \mathbf{A}^\dagger, \mathbf{A} \geq 0$
- $\mathcal{T}(\mathcal{H})$ ...... set of traceclass operators, i.e. $\mathrm{Tr}|\mathbf{A}| < \infty$
- $\mathcal{T}_2(\mathcal{H})$ ...... set of Hilbert-Schmidt operators, i.e. $\mathrm{Tr}\mathbf{A}^\dagger\mathbf{A} < \infty$
- $\mathcal{S}(\mathcal{H})$ ...... set of density matrices, i.e. $\varrho > 0, \mathrm{Tr}\varrho = 1$
- $\mathcal{U}(\mathcal{H})$ ...... set of unitary operators, i.e. $\mathbf{U}^\dagger = \mathbf{U}^{-1}$
- $\mathcal{L}(\mathcal{T}_2)$ ...... set of superoperators, i.e. $\Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$

Some properties of these sets are of importance in the context of quantum theory, which we shall list later on. Now we only write the properties related to the set structure

$$\mathcal{S}(\mathcal{H}) \subset \mathcal{L}_+(\mathcal{H}) \subset \mathcal{L}_s(\mathcal{H}) \subset \mathcal{L}(\mathcal{H}) \quad \text{and} \quad \mathcal{T}(\mathcal{H}) \subset \mathcal{T}_2(\mathcal{H}) \subset \mathcal{L}(\mathcal{H}) \tag{1.17}$$

1.2 Minimal interpretation

By minimal interpretation we mean statements that are common for all, or at least for most of all, different interpretations of quantum theory. In each theory there are four main objects : states, observables, dynamics and symmetries. Supported by one hundred years of successful explanation of quantum phenomena, we shall not talk about the details of motivations for mathematical representation of these objects. Of course, this topic is very interesting, but also very hard and as far as the author knows, none of the existing interpretations is satisfactory. We also omit the question of what we mean by quantum objects and where the border between the classical and quantum world lies.

Let's start with the *quantum state.* As we have mentioned before, to each quantum object we can define corresponding *Hilbert space* $\mathcal{H}$ that is the central object in the theory. The dimension of the Hilbert space reflects the maximal number of mutually perfectly distinguishable quantum states in a single observation. We identify quantum states with the elements of density matrices $\mathcal{S}(\mathcal{H})$. The set $\mathcal{S}(\mathcal{H})$ contains also one dimensional projectors, that, as we know, can be written in the form $\mathbf{P}_\psi = |\psi\rangle\langle\psi|$, i.e. they correspond to a ray in the projective Hilbert space. Such states identifiable with the normalized vectors $|\psi\rangle \in \mathcal{H}$ we shall call *pure states.*

The *quantum observables* are represented by the selfadjoint operators lying in the set $\mathcal{L}_s(\mathcal{H})$. The only way how we can check the validity of the theory is compare its predictions with the experimental reality. Let $\mathbf{A}$ be a selfadjoint operator corresponding to the measurement of a quantity and ϱ be a state of the measured system. Eigenvalues of $\mathbf{A}$ determine the possible outcomes of the measurement. Since in each observation we finish with some statistics of the outcomes, the quantum theory should be able to tell us the probabilities of single events. If the system was prepared in a state ϱ the probability $P(\lambda_n, \varrho)$ of an outcome λ_n is given by formula

$$P(\lambda_n, \varrho) := \mathrm{Tr}(\varrho\mathbf{P}_n) \tag{1.18}$$

where $\mathbf{P}_n$ is the projector on the eigenspace belonging to the eigenvalue λ_n. The fact that the operators are selfadjoint reflects that we are able to observe only real numbers, i.e. eigenvalues are real. Hence, we have defined the quantities that enable us to make a comparison with real experiments. The *mean value* of the measurement $\mathbf{A}$ is

$$\langle\mathbf{A}\rangle_\varrho = \sum_n \lambda_n P(\lambda_n, \varrho) = \mathrm{Tr}(\varrho\mathbf{A}). \tag{1.19}$$

The *dynamics* of quantum theory describes the change of the state with the flow of the time t. At first we shall define the evolution of pure states by postulating the *Schrödinger equation*

$$i\hbar\frac{d}{dt}|\psi_t\rangle = \mathbf{H}|\psi_t\rangle. \tag{1.20}$$

where $\hbar = 1,00095.10^{-34} Js^{-1}$ is the *Planck constant* and $\mathbf{H}$ is a specific selfadjoint operator of energy called *Hamiltonian.* The rays $|\psi\rangle$ then evolves according to the formal rule

$$|\psi_t\rangle = \mathbf{U}_t|\psi\rangle = \exp(-\frac{i}{\hbar}\mathbf{H}t)|\psi\rangle \tag{1.21}$$

where $|\psi\rangle$ is the initial condition of the differential equation (1.20).

Each density operator can be written in infinitely many ways as a convex sum of pure states, i.e. $\varrho = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k|$. Due to the linearity of Schrödinger equation we can write for general density matrix ϱ

$$\varrho = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k| \to \varrho_t = \sum_k \lambda_k \mathbf{U}_t|\psi_k\rangle\langle\psi_k|\mathbf{U}_t^\dagger = \mathbf{U}_t\varrho\mathbf{U}_t^\dagger \tag{1.22}$$

Hence, the dynamics is given by unitary map $\mathcal{U}_t : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$

$$\mathcal{U}_t[\varrho] = \mathbf{U}_t\varrho\mathbf{U}_t^\dagger \tag{1.23}$$

satisfying the group property $\mathcal{U}_{t+s} = \mathcal{U}_t \circ \mathcal{U}_s$ for all $t, s \in \mathbb{R}$. Differentiating the above equation we get the *Heisenberg equation* for general states

$$\frac{d}{dt}\varrho_t = \frac{i}{\hbar}[\varrho_t, \mathbf{H}]. \tag{1.24}$$

We say, that the evolution of the system is generated by the Hamiltonian $\mathbf{H}$.

1.3 The simplest quantum object

The simplest nontrivial quantum object is described by two-dimensional complex Hilbert space $\mathbb{C}^2 = \mathbb{C} \times \mathbb{C}$ and is called a *qubit*, because of its importance in *quantum information processing.* The aim of this section is to make a reader more familiar with the introduced notation.

1.3.1 The set of states - quantum kinematics

We shall denote the basis of the Hilbert space $\mathcal{H}$ by *ket symbols*

$$|0\rangle \quad \text{and} \quad |1\rangle \tag{1.25}$$

that correspond to mutually orthogonal vectors normalized to one. Each pure state $|\psi\rangle$ can be expressed as a *superposition* of the basis vectors, i.e.

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{1.26}$$

where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$ to ensure the norm $||\psi|| = \langle\psi|\psi\rangle = 1$. Here we need to point out that the *bra* vector $\langle\psi|$ corresponding to ket $|\psi\rangle$ takes the form

$$\langle\psi| = \alpha^*\langle 0| + \beta^*\langle 1| \tag{1.27}$$

where α^*, β^* are the complex conjugated numbers to α, β. Hence, according to the *normalization* condition for coefficients α, β the pure states are given by three real parameters. Moreover, every complex number we can write in the *trigonometric expression*, i.e.

$$\alpha = |\alpha|e^{i\eta} \qquad \beta = |\beta|e^{i\kappa}. \tag{1.28}$$

Now the normalization means that two real numbers $a = |\alpha|, b = |\beta|$ must satisfy the condition $a^2 + b^2 = 1$. Without any loss of generality we can put

$$a = \cos\phi \qquad b = \sin\phi \tag{1.29}$$

for $\phi \in [0, 2\pi]$. Since the pure states are elements of projective Hilbert space, the vectors $|\psi\rangle$ and $k|\psi\rangle$ represent the same ray. Let $k = e^{-i\eta}$ then

$$|\psi\rangle = \cos\phi|0\rangle + \sin\phi e^{i\theta}|1\rangle \tag{1.30}$$

with $0 \leq \theta = \kappa - \eta \leq \pi$. The factor $e^{-i\eta}$ is called *global phase.* The global phase together with normalization condition implies that we need only two real parameters ϕ, θ to determine the pure states of qubit. Moreover, with these parameters from the mentioned intervals our definition of a state is unique. To each pair (ϕ, θ) corresponds one pure state. The manifold endowed with coordinates (ϕ, θ) is known as *Bloch sphere*, i.e. pure states are points on the surface of the three-dimensional sphere.

The mixed states ϱ are positive selfadjoint 2x2 matrices with unit trace, i.e.

$$\varrho = \begin{pmatrix} \frac{1}{2} + z & x - iy \\ x + iy & \frac{1}{2} - z \end{pmatrix} \tag{1.31}$$

where the positivity condition requires that $0 \leq x^2 + y^2 + z^2 \leq 1/4$. For finite dimensions the requirement of positivity is equivalent to positivity of eigenvalues of the selfadjoint operator ϱ. For 2×2 matrices $\mathbf{A}$ we have the eigenvalues read

$$\lambda_\pm = \frac{1}{2}\left(\mathrm{Tr}\mathbf{A} \pm \sqrt{(\mathrm{Tr}\mathbf{A})^2 - 4\det\mathbf{A}}\right). \tag{1.32}$$

Since for quantum states $\mathrm{Tr}\varrho = 1$ and $\det\varrho = \frac{1}{4} - x^2 - y^2 - z^2$ we get $\lambda_\pm = (1 \pm 2\sqrt{x^2 + y^2 + z^2}))/2 \geq 0$. The condition written above assures the positivity of eigenvalues.

Since the set of Hilbert-Schmidt operators $\mathcal{T}_2(\mathcal{H})$ form a *complex Hilbert space* we can express each element $\mathbf{A} \in \mathcal{T}_2(\mathcal{H})$ as a linear combination of basis elements in the following way. Choose an orthonormal basis in this space, that is a collection of operators $\Lambda_k \in \mathcal{T}_2(\mathcal{H})$ satisfying

$$\mathrm{Tr}(\Lambda_k^\dagger \Lambda_l) = \delta_{kl} \tag{1.33}$$

for $k, l = 1, \ldots, \dim\mathcal{T}_2(\mathcal{H})$, where $\dim\mathcal{T}_2(\mathcal{H}) = (\dim\mathcal{H})^2$. The general element $\mathbf{A} \in \mathcal{T}_2(\mathcal{H})$ is expressed as a linear combination

$$\mathbf{A} = \sum_k a_k \Lambda_k \quad \text{with} \quad a_k := \mathrm{Tr}(\Lambda_k^\dagger \mathbf{A}). \tag{1.34}$$

In the case of qubit the basis consists of four operators. Let us define the basis of $\sigma-$**matrices**

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.35}$$

The matrices $\frac{1}{2}\sigma_k$ are orthonormal in the required sense. Moreover the σ- matrices are all unitary. In fact, one can easily prove the validity of the following (very important) formula

$$\sigma_k \sigma_l = \delta_{kl}\mathbb{1} + i\varepsilon_{klm}\sigma_m. \tag{1.36}$$

where we put $x = 1, y = 2, z = 3$ and $k, l, m = 1, 2, 3$. The ε_{klm} denotes the completely antisymmetric symbol, i.e. $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$, $\varepsilon_{321} = \varepsilon_{132} = \varepsilon_{213} = -1$ and all the other coefficients are zero.

Expressing the general state ϱ in this basis we obtain

$$\varrho = \frac{1}{2}\mathbb{1} + x\sigma_x + y\sigma_y + z\sigma_z = \frac{1}{2}\mathbb{1} + \vec{n}.\vec{\sigma} \tag{1.37}$$

where $\vec{n} := (x, y, z)$ is the real three-dimensional vector satisfying $|\vec{n}| \leq \frac{1}{2}$.

We have found that states are in a bijective correspondence with the real three-dimensional vectors of the length less then $1/2$. Geometrically it can be described by a ball with radius $r = 1/2$. As we have said before, the pure states corresponds to surface of this ball. Indeed, the state is a projection if and only if $|\vec{n}| = 1/2$, because only in these cases

$$\varrho^2 = \frac{1}{4}\mathbb{1} + \frac{1}{2}(\vec{n}+\vec{n}).\vec{\sigma} + (\vec{n}.\vec{\sigma})(\vec{n}.\vec{\sigma}) = \frac{1}{2} + \vec{n}.\vec{\sigma} = \varrho \tag{1.38}$$

In our calculation we used the relation implied by (1.36)

$$(\vec{n}.\vec{\sigma})(\vec{m}.\vec{\sigma}) = (\vec{n}.\vec{m})\mathbb{1} + i(\vec{n}\times\vec{m}).\vec{\sigma} \tag{1.39}$$

and the fact that $\vec{n}\times\vec{n} = 0$.
Remark. *The property that the boundary of $\mathcal{S}(\mathcal{H})$ consists of pure states is typical only of the qubit case! The negative operators are arbitrarily close (in the trace norm) to the boundary of state space. For each state ϱ with one zero eigenvalue we can define an operator ϱ' with the same spectrum, only instead of the zero eigenvalue we put the eigenvalue of the new operator equal to $\varepsilon < 0$, i.e. ϱ' is negative. Then the distance between the operators ϱ and ϱ' is given by the value of ϵ which can be arbitrarily small. Consequently, each ϱ (with one zero eigenvalue) lies on the boundary and for $dim\mathcal{H} > 2$ not only pure states can have one zero eigenvalue. Hence, in general, the boundary of $\mathcal{S}(\mathcal{H})$ contains also nonpure states.*

Suppose a collection of states $\varrho_k \in \mathcal{S}(\mathcal{H})$. By *convex combination* we understand a specific linear combination

$$\varrho = \sum_k \lambda_k \varrho_k \tag{1.40}$$

with $\lambda_k \geq 0$ and $\sum_k \lambda_k = 1$. It is easy to verify that operator ϱ is again an element of state space, i.e. $\varrho \in \mathcal{S}(\mathcal{H})$. The set $\mathcal{S}(\mathcal{H})$ is *convex*, because for all states $\varrho, \xi \in \mathcal{S}(\mathcal{H})$ also their convex combinations $\lambda\varrho + (1-\lambda)\xi$ are elements of $\mathcal{S}(\mathcal{H})$. In particular, for the qubit we get

$$\lambda(\frac{1}{2}\mathbb{1} + \vec{n}.\vec{\sigma}) + (1-\lambda)(\frac{1}{2}\mathbb{1} + \vec{m}.\vec{\sigma}) = \frac{1}{2}\mathbb{1} + \vec{t}.\vec{\sigma} \tag{1.41}$$

where $\vec{t} = \lambda\vec{n} + (1-\lambda)\vec{m}$. Since $|\vec{t}| \leq \lambda|\vec{n}| + (1-\lambda)|\vec{m}| \leq 1/2$ we know that the operator defined as $\frac{1}{2} + \vec{t}.\vec{\sigma}$ corresponds to a regular quantum state of a qubit.

Notice the difference between the convex combination and superposition of pure states. They are defined on the different spaces and there is no relation between them. In fact, the pure states of a qubit are such elements of $\mathcal{S}(\mathcal{H})$, that cannot be written like a convex combination of other states. Put the projector $\mathbf{P}_\psi = |\psi\rangle\langle\psi| = \frac{1}{2}\mathbb{1} + \vec{t}.\vec{\sigma}$, i.e. $|\vec{t}| = 1/2$. Consider the possibility to write such projector $|\psi\rangle\langle\psi|$ in a way

$$|\psi\rangle\langle\psi| = \lambda\varrho + (1-\lambda)\xi. \tag{1.42}$$

for some $\varrho, \xi \in \mathcal{S}(\mathcal{H})$ and $\lambda \in [0,1]$, where ϱ, ξ are defined like before, i.e. via $\vec{n}, \vec{m}$. We shall show that the only possibility how to fulfill the condition

$$|\vec{t}|^2 = |\lambda\vec{n} + (1-\lambda)\vec{m}|^2 = \lambda^2|\vec{n}|^2 + (1-\lambda)^2|\vec{m}|^2 + 2\lambda(1-\lambda)\vec{n}.\vec{m} \tag{1.43}$$

is to put $\vec{n} = \vec{m} = \vec{t}$. The *triangle inequality* $|\vec{t}| \leq \lambda|\vec{n}| + (1-\lambda)|\vec{m}|$ implies that, if either $|\vec{n}|$, or $|\vec{m}|$, is strictly less than $\frac{1}{2}$, then also $|\vec{t}| < \frac{1}{2}$, what contradicts the fact that $\mathbf{P}_\psi$ is a pure state. In what follows $|\vec{n}| = |\vec{m}| = \frac{1}{2}$. Introducing it into the eq.(1.43) we are getting

$$\frac{1}{4} = \frac{1}{4}[1 + 2\lambda(\lambda-1)(1-\cos\theta)] \tag{1.44}$$

where $\theta \in [0,\pi]$ is an angle between the vectors $\vec{n}$ and $\vec{m}$. It is straightforward that last equation is valid if and only if $\vec{n} = \vec{m}$ ($\cos\theta = 1$). Since $\vec{t} = \lambda\vec{n} + (1-\lambda)\vec{m} = \vec{m}$ we have found that it is impossible to express pure states as a convex combination (1.42), except the $\varrho = \xi = |\psi\rangle\langle\psi|$.

1.3.2 The set of observables

Observables correspond to selfadjoint elements of $\mathcal{L}(\mathcal{H})$. Note that since the density operators are selfadjoint, too, they can also play a role of observables. The selfadjoint operators form a real linear space. It means for $\mathbf{A}, \mathbf{B} \in \mathcal{L}_s(\mathcal{H})$ also the operator $a\mathbf{A} + b\mathbf{B}$ is selfadjoint for $a, b \in \mathbb{R}$. Hence, the set $\mathcal{L}_s(\mathcal{H})$ is a real Hilbert space endowed with the *trace scalar product.* A general element of this real Hilbert space can be written in the form

$$\mathbf{A} = \begin{pmatrix} a & c-id \\ c+id & b \end{pmatrix} \tag{1.45}$$

with $a, b, c, d \in \mathbb{R}$.

Since the σ *matrices* are selfadjoint, they form an orthogonal basis in $\mathcal{L}_s(\mathcal{H})$, too. Then each element $\mathbf{A} \in \mathcal{L}_s(\mathcal{H})$ can be written in the form

$$\mathbf{A} = \alpha \mathbb{1} + \beta \vec{n}.\vec{\sigma} \tag{1.46}$$

with unit vector $|\vec{n}| = 1$ and $\alpha, \beta \in \mathbb{R}$. Comparing the two expressions for $\mathbf{A}$ we get

$$\alpha = \frac{a+b}{2} \quad \text{and} \quad \beta\vec{n} = (c, d, \frac{1}{2}(a-b)). \tag{1.47}$$

The *outcomes* of an experiment described by the selfadjoint operator $\mathbf{A}$ correspond to *eigenvalues* $\lambda_\pm$ of $\mathbf{A}$ given by Eq.(1.32), i.e. $\lambda_\pm = \alpha \pm \beta$ with $\beta = \sqrt{c^2 + d^2 + \frac{(a-b)^2}{4}}$. Since we are able to observe only real numbers, we require the eigenvalues to be real. If $\mathbf{A}$ is selfadjoint, then eigenvalues are real ($\alpha, \beta \in \mathbb{R}$), but *the opposite is not the case.* For example, the matrix $\mathbf{X} \notin \mathcal{L}_s(\mathcal{H})$

$$\mathbf{X} = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \tag{1.48}$$

with real eigenvalues $\lambda_\pm = \pm\sqrt{ab}$, if $a, b \in \mathbb{R}$ and $ab > 0$.

The corresponding eigenvectors $|\psi_\pm\rangle$ of $\mathbf{A}$ are mutually orthogonal, i.e. $\langle\psi_+|\psi_-\rangle = 0$. It means the selfadjoint operator $\mathbf{A}$ can be written in the *spectral form*

$$\mathbf{A} = \lambda_+ \mathbf{P}_+ + \lambda_- \mathbf{P}_- \quad \text{where} \quad \mathbf{P}_\pm = |\psi_\pm\rangle\langle\psi_\pm|. \tag{1.49}$$

In addition, the requirement that eigenvalues are real also requires the orthogonality of corresponding eigenvectors, we shall trivially obtain the selfadjointness of the operator $\mathbf{A}$. Why do we need the orthogonality of the eigenvectors? Here we come to the biggest problem in interpretation of quantum theory. What happens with the system after a measurement has been performed?

Consider the measurement described by the introduced operator $\mathbf{X}$. In that case the eigenvectors belonging to eigenvalues $\lambda_\pm = \pm\sqrt{ab}$ are

$$|\psi_\pm\rangle = \pm\sqrt{\frac{a}{a+b}}|0\rangle + \sqrt{\frac{b}{a+b}}|1\rangle. \tag{1.50}$$

Of course, they are not orthogonal for general $a, b > 0$, because $\langle\psi_+|\psi_-\rangle = \frac{b-a}{a+b}$. If the measurement does not destroy the system, then we can repeat the same measurement of the quantity $\mathbf{X}$ again. It is an experimental fact, that the observed outcome is always the same. If we select among the outcomes only those with the value λ_+ fixed and perform the measurement again only on such selected subensemble, then we should obtain the mean value equals to λ_+. As we have said, this means that in the second measurement we observe λ_+ with the unit probability. Since the probability $P(\lambda_+, \varrho)$ of the outcome λ_+ is equal to the mean value of the selfadjoint operator $\langle\mathbf{P}_+\rangle_\varrho$, we see that $P(\lambda_+, \varrho) = 1$ implies $\varrho = \mathbf{P}_+$. In other words, the state of the system belonging to the outcome λ_+ must correspond to eigenvector $|\psi_+\rangle$. This is known as the **projection postulate**, that is, after the measurement the state of the system jumps into the state associated with the eigenvector belonging to the observed eigenvalue. A simple consequence of this postulate is the relation $P(\lambda_-, \mathbf{P}_+) = \mathrm{Tr}\mathbf{P}_+\mathbf{P}_- = 0$. For the eigenvectors it results in the condition of orthogonality, i.e. $|\langle\psi_+|\psi_-\rangle|^2 = 0$ (the operator $\mathbf{X}$ is selfadjoint only if $a = b$, when the eigenvectors are orthogonal).

To conclude, let us stress that the selfadjointness of the operators corresponding to physical observables (=quantities) follows not only from the requirement of real spectra, but also from the validity of the projection postulate, which ensures the orthogonality of eigenvectors. The eigenvectors for the general selfadjoint operator $\mathbf{A}$ with $\vec{n} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$ coincide with the eigenvectors of the operator $\vec{n}.\vec{\sigma}$, i.e.

$$|\psi_\pm\rangle = \pm e^{-i\varphi/2}\sin\frac{\vartheta}{2}|0\rangle + e^{i\varphi/2}\cos\frac{\vartheta}{2}|1\rangle. \tag{1.51}$$

1.3.3 Quantum evolution

We mentioned that the quantum evolution is described by a unitary mapping. For qubits the general element of $\mathbf{U} \in \mathcal{U}(\mathcal{H})$ is represented by a matrix

$$\mathbf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{1.52}$$

for $a, b, c, d \in \mathbb{C}$ such that $aa^* + bb^* = 1$, $cc^* + dd^* = 1$, $a^*c + b^*d = 0$ and $c^*a + d^*b = 0$. The operators belonging to $\mathcal{U}(\mathcal{H})$ form a group with the matrix product as a group binary operation, that is $\mathbb{1} \in \mathcal{U}(\mathcal{H})$, $\mathbf{U}^{-1} \in \mathcal{U}(\mathcal{H})$ and $(\mathbf{U}_1\mathbf{U}_2)\mathbf{U}_3 = \mathbf{U}_1(\mathbf{U}_2\mathbf{U}_3)$. We shall denote this group of the 2×2 unitary matrices by $U(2)$. Moreover, it is a *Lie group*, because each element is determined by four real numbers that play the role of coordinates. To each Lie group we can define a corresponding *Lie algebra* (denoted by $u(2)$ in our case). Choose some operator $\hat{\mathbf{A}} \in \mathcal{L}(\mathcal{H})$. Define a one-parametric subset of $\mathcal{L}(\mathcal{H})$ in the following way (*exponential mapping*)

$$\mathbb{R} \ni t \mapsto \mathbf{U}_t = \exp(t\hat{\mathbf{A}}) = \mathbb{1} + \hat{\mathbf{A}} + \frac{1}{2}\hat{\mathbf{A}}^2 \ldots \in U(2) \tag{1.53}$$

The condition of unitarity of $\mathbf{U}_t$ determines the constraints on the possible operators $\hat{\mathbf{A}}$. If $\mathbf{U}_t = \exp(t\hat{\mathbf{A}})$, then $\mathbf{U}_t\mathbf{U}_t^\dagger = \mathbf{U}_t^\dagger\mathbf{U}_t = \mathbb{1}$ for all $t \in \mathbb{R}$. For small values of t we get the condition

$$\mathbb{1} = \mathbf{U}_\varepsilon\mathbf{U}_\varepsilon^\dagger = (\mathbb{1} + \varepsilon\hat{\mathbf{A}})(\mathbb{1} + \varepsilon\hat{\mathbf{A}}^\dagger) = \mathbb{1} + \varepsilon(\hat{\mathbf{A}} + \hat{\mathbf{A}}^\dagger) \tag{1.54}$$

that is, the allowed operators $\hat{\mathbf{A}}$ must be *antihermitian*, i.e. $\hat{\mathbf{A}} = -\hat{\mathbf{A}}^\dagger$. Hence, the *Lie algebra* $u(2)$ is an algebra of antihermitian operators. The Lie algebra form a real linear space endowed with the defined antisymmetric bilinear form - *commutator*. It means, if $\hat{\mathbf{A}}, \hat{\mathbf{B}} \in u(2)$, then $\alpha\hat{\mathbf{A}} + \beta\hat{\mathbf{B}} \in u(2)$ and $[\hat{\mathbf{A}}, \hat{\mathbf{B}}] := \hat{\mathbf{A}}\hat{\mathbf{B}} - \hat{\mathbf{B}}\hat{\mathbf{A}} = \hat{\mathbf{C}} \in u(2)$, too.

If $\hat{\mathbf{A}}$ is antihermitian, then the operator $\mathbf{A} := i\hat{\mathbf{A}}$ is selfadjoint (=hermitian), i.e. there is a one-to-one correspondence between antihermitian and hermitian operators. The exponential mapping then can be expressed as $\mathbf{U}_t = \exp(it\mathbf{A})$ with $\mathbf{A} \in \mathcal{L}_s(\mathcal{H})$. Note that $\mathcal{L}_s(\mathcal{H})$ is not closed according to commutator of its elements, because

$$[\mathbf{A}, \mathbf{B}] = [i\hat{\mathbf{A}}, i\hat{\mathbf{B}}] = -[\hat{\mathbf{A}}, \hat{\mathbf{B}}] = -\hat{\mathbf{C}} = i\mathbf{C} \tag{1.55}$$

is antihermitian operator. It is also not closed with respect to the product of two operators, i.e. $\mathbf{AB}$ is in $\mathcal{L}_s(\mathcal{H})$ if and only if $[\mathbf{A}, \mathbf{B}] = 0$, because $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger \neq \mathbf{AB}$ for general $\mathbf{A}, \mathbf{B} \in \mathcal{L}_s(\mathcal{H})$.

We are going to try to express the general unitary operator $\mathbf{U}$ in the basis of σ matrices. Since the unitary operator $\mathbf{U}$ is normal ($\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U}$), its eigenvectors $|\psi_\pm\rangle$ are mutually orthogonal. The eigenvalues are complex numbers of the kind $e^{i\alpha_\pm}$, because in spectral form the matrix $\mathbf{U}$ is diagonal and unitaries must preserve the scalar product, i.e. $\langle\psi_\pm|\mathbf{U}^\dagger\mathbf{U}|\psi_\pm\rangle = \lambda_\pm^*\lambda_\pm = 1$ implies $\lambda_\pm = e^{i\alpha_\pm}$ with $\alpha_\pm \in \mathbb{R}$.

It follows that we can define a unitary operator in $U(2)$ by specifying two real numbers $\alpha_\pm$ and one normalized vector $|\psi_+\rangle \equiv |\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with $\alpha, \beta \in \mathbb{C}$. In the qubit case the basis is fully determined by a single vector, because $|\psi^\perp\rangle \equiv |\psi_-\rangle = \beta^*|0\rangle - \alpha^*|1\rangle$ is unique. Put $\mathbf{U} := e^{i\alpha_+}\mathbf{P}_\psi + e^{i\alpha_-}\mathbf{P}_{\psi^\perp}$, where $\mathbf{P}_\psi = \frac{1}{2}\mathbb{1} + \vec{n}.\vec{\sigma}$ and $\mathbf{P}_{\psi^\perp} = \frac{1}{2}\mathbb{1} - \vec{n}.\vec{\sigma}$. Then for the general $\mathbf{U}$ we get

$$\mathbf{U} = \frac{e^{i\alpha_+} + e^{i\alpha_-}}{2}\mathbb{1} + \frac{e^{i\alpha_+} - e^{i\alpha_-}}{2}\vec{n}.\vec{\sigma}. \tag{1.56}$$

But the pure states in quantum theory are given up to a *global phase*. Therefore the unitary transformations $\mathbf{U}$ and $e^{i\varphi}\mathbf{U}$ are equivalent. Putting $\alpha_\pm = \varphi \pm \eta$ we obtain

$$\mathbf{U} = \cos\eta\mathbb{1} + i\sin\eta(\vec{n}.\vec{\sigma}) \tag{1.57}$$

where we have omitted the global factor $e^{i\varphi}$.

The equivalence *up to a global factor* defines new subset of unitary operators called *special unitaries* satisfying the condition $\det\mathbf{U} = e^{i\eta}e^{-i\eta} = 1$. Such operators form again a *Lie group* $SU(2)$ with a corresponding *Lie algebra* $su(2)$. The requirement $\det\mathbf{U}_t = 1$ for all t implies for elements of Lie algebra $\mathrm{Tr}\mathbf{A} = 0$. Traceless selfadjoint operators form a linear space[1] $su(2) \subset \mathcal{L}_s(\mathcal{H})$. Except $\mathbb{1}$ all the matrices from σ basis are traceless and selfadjoint and the dimension of $su(2) \subset \mathcal{T}_2(\mathcal{H})$ is three. The general element $\mathbf{A} \in su(2)$ can be written in the form

$$\mathbf{A} = x\sigma_x + y\sigma_y + z\sigma_z = \eta(\vec{n}.\vec{\sigma}) \tag{1.58}$$

with $|\vec{n}| = 1$. The element of $SU(2)$ are then $\mathbf{U} = \exp(\eta\vec{n}.\vec{\sigma})$. Using the formula (1.39) and a little calculation we obtain

$$\mathbf{U} = \cos\eta\mathbb{1} + i\sin\eta(\vec{n}.\vec{\sigma}). \tag{1.59}$$

The question is whether each element of $SU(2)$ can be expressed like $\exp(i\mathbf{A})$ for some $\mathbf{A} \in su(2)$. In the Eq.(1.57) we have the most general element of $SU(2)$. Comparing it with the last equation we can conclude that $SU(2) = \exp(i.su(2))$ is an *exponential group.* We have found that the general qubit Hamiltonian $\mathbf{H}$ generating the evolution $\mathbf{U}_t = \exp(i\mathbf{H}t)$ is an element of $su(2)$.

1.4 Quantum theory: summary

In this section we shall briefly repeat basic mathematical representations of the physical concepts of quantum theory.

1. Quantum state.

Quantum state of a physical object represents the maximum of our knowledge about the quantum system. The essential underlying concept in the determination of quantum system is the notion of the *dimension* $d = \dim\mathcal{H}$ of the associated Hilbert space $\mathcal{H}$. The dimension represents the maximal number of perfectly distinguishable quantum states in a single observation, i.e. by measuring a single outcome. The *set of states* $\varrho \in \mathcal{S}(\mathcal{H})$ is contains the selfadjoint positive elements of the set of Hilbert space operators $\mathcal{L}(\mathcal{H})$ with unit trace (*density operators*), i.e. $\varrho = \varrho^\dagger$, $\varrho \geq 0$ and $\mathrm{Tr}\varrho = 1$. The *convex structure* of the set of states implies that there exists a specific subset of *pure states*, which are identified with *extremal points* of $\mathcal{S}(\mathcal{H})$. Equivalently, one can say that the state ϱ is pure, if $\varrho = \varrho^2$ is a projection, or $\mathrm{Tr}\varrho^2 = 1$. There is one-to-one correspondence between pure states and normalized vectors $|\psi\rangle$ ($\langle\psi|\psi\rangle = 1$) from the underlying Hilbert space $\mathcal{H}$. The set of pure states can be identified with the set of unit vectors from the Hilbert space $\mathcal{H}$. Such identification enables us to speak about the *superposition principle* in quantum theory. This principle guarantees the existence of the quantum

[1] Here we use the same notation for traceless hermitian and traceless antihermitian operators, but we hope that the right usage is given by context

state $\mathbf{P}_{\alpha|\psi\rangle+\beta|\phi\rangle}$ (for $|\alpha|^2+|\beta|^2+2\mathrm{Re}[\alpha^*\beta\langle\psi|\phi\rangle]=1$), on condition the states $\mathbf{P}_\psi, \mathbf{P}_\phi$ do exist. Each quantum state can be written in the form

$$\varrho = \frac{1}{d}\mathbb{1} + \vec{n}\cdot\vec{\Theta} \tag{1.60}$$

where $\vec{\Theta} = (\Theta_1, \ldots, \Theta_{d^2-1})$ is the vector of selfadjoint traceless operators and $\vec{n}$ is a real (d^2-1)-dimensional vector such that the positivity of ϱ is assured.

2. Orthogonal (projective) measurement

The concept of *orthogonal measurement* reflects the property of the dimension of Hilbert space $\mathcal{H}$, because the number of different outcomes can be maximally d. Let us denote these outcomes by real numbers $\lambda_1, \ldots, \lambda_d$. The *projective postulate* determines that after measuring the outcome λ_j the quantum state of the physical system is described (up to the normalization) by $\varrho_j = \mathbf{P}_j\varrho\mathbf{P}_j$, where ϱ is the initial state of the quantum system and $\mathbf{P}_j$ are the projectors associated with the outcomes λ_j satisfying the normalization $\sum_j \mathbf{P}_j = \mathbb{1}$ and the property $\mathbf{P}_j\mathbf{P}_k = 0 = \delta_{jk}\mathbf{P}_j$. Projective measurements $\mathbf{A}$ are associated with *selfadjoint operators*, respectively with their (unique) *spectral decomposition*

$$\mathbf{A} = \sum_j \lambda_j \mathbf{P}_j = \kappa\mathbb{1} + \vec{m}\cdot\vec{\Theta} \tag{1.61}$$

where the last equality uses the elements introduced in the previous paragraph with an arbitrary real number κ and real vector $\vec{m}$.

3. Probability rule

The probability rule is the most important formula in the whole quantum theory, because it connects the abstract mathematical notions of states and observables with experimentally observed quantities, i.e. with the probability distributions of the measurement outcomes. Hence, the probability $P(\varrho, \lambda_j)$ of observing the result λ_j is given by the formula

$$P(\varrho, \lambda_j) = \mathrm{Tr}\varrho\mathbf{P}_j\,. \tag{1.62}$$

where ϱ is the state of the system before the measurement. The *mean value* of the quantity $\mathbf{A} = \sum_j \lambda_j\mathbf{P}_j$ in the state ϱ is given as

$$\langle\mathbf{A}\rangle_\varrho = \mathrm{Tr}\varrho\mathbf{A} = \kappa + \vec{n}.\vec{m} \tag{1.63}$$

where the last equality is valid only if operators $\{\Theta_j\}$ are orthonormal in the following sense $\mathrm{Tr}\Theta_j\Theta_k = \delta_{jk}$. It follows that the vector coefficients n_k represent the mean values of the selfadjoint operators Θ_k, i.e. $n_k = \mathrm{Tr}\varrho\Theta_k$.

4. Evolution

In quantum theory the time evolution is driven by a *Hamiltonian operator* $\mathbf{H}$ in the following sense

$$\dot{\varrho}_t = \frac{i}{\hbar}[\varrho_t, \mathbf{H}_t] \tag{1.64}$$

The Hamiltonian of the system $\mathbf{H}_t$ is a selfadjoint operator related to the energy of the system. In the case of time-independent Hamiltonian, $\mathbf{H}_t = \mathbf{H}$, the evolution is described by a one-parametric group of unitary transformations $\mathbf{U}_t = \exp(-\frac{i}{\hbar}\mathbf{H}t)$.

5. Composed systems

The description of a composed system uses the operation of the *tensor product* of Hilbert spaces. The aim of this thesis is to study consequences of this structure onto elementary objects of quantum theory, such as states, observables and dynamics. Tensor product of two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is a new Hilbert space (denoted by) $\mathcal{H}_1\otimes\mathcal{H}_2$ with elements from the Cartesian product $\mathcal{H}_1\times\mathcal{H}_2$, i.e. ordered couples $\Psi = [\psi_1, \psi_2] \in \mathcal{H}_1\times\mathcal{H}_2$ such that $\psi_1 \in \mathcal{H}_1$ and $\psi_2 \in \mathcal{H}_2$. In the Dirac notation we will write $\Psi = [\psi_1, \psi_2] = |\psi_1\rangle \otimes |\psi_2\rangle$. Moreover, we shall define new scalar product on the set $\mathcal{H}_1 \otimes \mathcal{H}_2$ by the relation

$$\langle \Psi | \Phi \rangle = \langle \psi_1 | \phi_1 \rangle \cdot \langle \psi_2 | \phi_2 \rangle \tag{1.65}$$

The dimension of the resulting Hilbert space $d = \dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = (\dim \mathcal{H}_1).\dim(\mathcal{H}_2) = d_1 d_2$. Let $\{\Theta_j\}$ be the operator basis of the space $\mathcal{L}(\mathcal{H}_1)$ and $\{\Lambda_k\}$ be the basis of $\mathcal{L}(\mathcal{H}_2)$. Then the most general operator defined on the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ takes the form

$$\mathbf{A} = \kappa \mathbb{1}_1 \otimes \mathbb{1}_2 + \vec{\alpha} \cdot \vec{\Theta} \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \vec{\beta} \cdot \vec{\Lambda} + \sum_{kl} \gamma_{kl} \Theta_k \otimes \Lambda_l \tag{1.66}$$

This operator is selfadjoint, if all the parameters $\kappa, \alpha_k, \beta_l, \gamma_{kl}$ are real. That is, the set of all quantum projective measurements $\mathcal{L}_s(\mathcal{H}_1 \otimes \mathcal{H}_2)$ form a real 16-dimensional linear space. The *local measurements* of the first subsystem have the form $\mathbf{A} \otimes \mathbb{1}_2$. Denote the set of all local operators by $\mathcal{L}_1 \subset \mathcal{L}_s(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Similarly, define the subset $\mathcal{L}_2 \subset \mathcal{L}_s(\mathcal{H}_1 \otimes \mathcal{H}_2)$ representing the local measurements of the subsystem B. Another type of measurements have the form $\mathbf{A} \otimes \mathbf{B}$. Let us denote the set of such observables by $\mathcal{L}_{12}$. They describe joint measurements of the local ones, because formally $\mathbf{A} \otimes \mathbf{B} = (\mathbf{A} \otimes \mathbb{1}_2)(\mathbb{1}_1 \otimes \mathbf{B}) = (\mathbb{1}_1 \otimes \mathbf{B})(\mathbf{A} \otimes \mathbb{1}_2)$. In classical physics the outcomes of the joint measurements are couples of single outcomes. The postulates of quantum theory imply that the outcomes are given by eigenvalues of $\mathbf{A} \otimes \mathbf{B}$, i.e. by the product of the single outcomes of $\mathbf{A}$ and $\mathbf{B}$. Physically the joint measurement $\mathbf{A} \otimes \mathbf{B}$ can be realized by two observers and each of them obtains a result of his local measurement. It means their outcomes together represent a pair of outcomes and not only one number (product of outcomes). However, the measurement of $\mathbf{A} \otimes \mathbf{B}$ can be done also in a way, where we have no knowledge about the values of $\mathbf{A}$ and $\mathbf{B}$ alone. There is nothing mysterious in it. Simply the product of any two measurements $\mathbf{X}$ and $\mathbf{Y}$ can describe two physically different situations. In one case, we realize joint observations of $\mathbf{X}$ and $\mathbf{Y}$ on the physical system, while in the second case we perform a new measurement $\mathbf{C} = \mathbf{XY}$. In both situations the operators $\mathbf{X}$ and $\mathbf{Y}$ must commute. We remind us that a product of selfadjoint operators is selfadjoint only if the operators commute. In our case $\mathbf{X} = \mathbf{A} \otimes \mathbb{1}$ and $\mathbf{Y} = \mathbb{1} \otimes \mathbf{B}$. Therefore, the product $\mathbf{A} \otimes \mathbf{B}$ can be understood in two different ways that correspond to different experimental realizations. One can say that the locality of the measurement $\mathbf{A} \otimes \mathbf{B}$ depends on the physical context. *Global measurements* are those that cannot be implemented locally (in the above sense). The generalization to more than two systems is straightforward.

Chapter 2

Basics of information theory

Probability: The Logic of Science
E.T.Jaynes

2.1 Clarifying the information

It is meaningless to ask about the absolute meaning of information, that is about the absolute amount of information contained in an object. The question *"How much information is it in ..."* is difficult and we are not able to answer it. We assume that information is a **relative notion**. Our questions should have the structure *"How much information do we have about ..., if we know that ..."*. The natural question arises, what it means, if we say "we know that ...".

Here the physical reality enters into these abstract ideas. Our knowledge (*we know that...*) is based on our **observations**, that have a strictly physical essence. Denote single observation by $\mathbf{A}$ and the whole set of observations by $\mathcal{O}$. We shall denote the outcomes of the observation $\mathbf{A}$ by a. These values determine the abstract notion of observation $\mathbf{A}$, which is a collection of outcomes. On the other hand the set of all observables $\mathcal{O}$ determines the system under investigation. This corresponds to the fact that the only way how we can learn something about the system is by performing an observation. Everything else remains hidden for us.

In the heart of every observation is a potential randomness. We cannot base our knowledge about the system on a single observation. The possible randomness of outcomes in our repeated observations requests the usage of the mathematical statistics and probabilities in our description of information gain. The characterizes the amount of our knowledge $I(\mathbf{A},\mathbf{B})$ about the observation $\mathbf{B}$ gathered from the observation of $\mathbf{A}$. In particular, if we measure (a *random event*) $a \in \mathbf{A}$, what are the possible outcomes b of the measurement $\mathbf{B}$? To answer this question we must use the probability theory and the answer is, in general, not deterministic. Without any specifications we can intuitively discriminate the situations when the information is maximal and when it is minimal. On one hand if outcome a determines the value b uniquely for each a, then $I(\mathbf{A},\mathbf{B})$ is maximal. On the other hand, if b is randomly distributed for all outcomes a, then we have no information about b, that is, information is minimal, $I(\mathbf{A},\mathbf{B}) = 0$.

The *state of a system* represents our *state of knowledge* about the given object. It enables us to calculate the probability of any outcome of every observation. Denote the state by ϱ and the probability rule by $P(a \in \mathbf{A}, \varrho)$. In a sense the *projection postulate* is valid also in classical theory, but its meaning is "trivial". If we measure the value $a \in \mathbf{A}$, then the single system is described by such ϱ_a, that $P(a \in \mathbf{A}, \varrho_a) = 1$ and $P(a' \in \mathbf{A}, \varrho_a) = 0$ for $a' \neq a$. In classical physics (unlike quantum), the measured object had been in the state ϱ_a also before the measurement was performed. That is, in classical physics the measurement does not cause a "jump evolution" from one state to another.

The *purity* of state ϱ corresponds to the existence of at least one observable $\mathbf{A}$ with the outcome a, for which the probability $P(a \in \mathbf{A}, \varrho) = 1$. In classical physics for all $\mathbf{A} \in \mathcal{O}$ the pure state has a definite value a, i.e. for all $\mathbf{A}$ there exists $a \in \mathbf{A}$ with $P(a \in \mathbf{A}, \varrho) = 1$.

In classical physics pure states form a set of *elementary events* Ω and observables are functions $\mathbf{A} : \Omega \to \mathbb{R}$. We can identify the set Ω with the set of outcomes $\mathbf{A}$, if such an outcome of observation $\mathbf{A}$ determines the results for all other observations $\mathbf{B} \in \mathcal{O}$. The standard choice of Ω is in classical mechanics given by couples of the position $\vec{r}$ and momentum $\vec{p}$. The observation $\mathbf{A} = \{1,2,3,4,5,6\}$ determines the set of *elementary events* in the case of a *die*. Suppose we are not interested in the value of the die, but only whether the number is odd, or even, or $n \leq 3$, or $n > 3$. Then the corresponding sets we are interested in are $E_1 = odd = \{1,3,5\}$, $E_2 = even = \{2,4,6\}$, $E_3 = \overline{3} = \{1,2,3\}$ and $E_4 = \overline{4} = \{4,5,6\}$. Define new observations $\mathbf{P} = \{odd, even\}$ and $\mathbf{B} = \{\overline{3}, \overline{4}\}$. The sets E_j will be called *events*. Denote the set of all possible subsets (so called the *potential set*) of Ω by $\mathcal{P}(\Omega)$. The set of all possible events $\mathcal{E}(\Omega)$ is a subset of $\mathcal{P}(\Omega)$, i.e. $\mathcal{E}(\Omega) \subset \mathcal{P}(\Omega)$.

2.1.1 Probability, statistical dependency and correlation

Kolmogorov, the father of the probability theory, simply identified the **random event = event = set** and founded the basics of modern probabilistic theory. Let Ω be the set of elementary events ω, or the *sample space*. The set of events $\mathcal{E}(\Omega)$ form a **σ-algebra**, if it possesses the following properties

1. $\emptyset, \Omega \in \mathcal{E}(\Omega)$
2. $E_j \in \mathcal{E}(\Omega)$, then $\bigcup_j E_j \in \mathcal{E}(\Omega)$
3. $E \in \mathcal{E}(\Omega)$, then $\Omega \setminus E \in \mathcal{E}(\Omega)$

The probability is a function $p : \mathcal{E}(\Omega) \to [0,1]$ satisfying

1. $p(\Omega) = 1, p(\emptyset) = 0$
2. for mutually disjoint collection $\{E_j\} \in \mathcal{E}(\Omega)$, i.e. $E_j \cap E_k = \emptyset$ for $k \neq j$, $p(\bigcup_j E_j) = \sum_j p(E_j)$
3. $p(\Omega \setminus E) = 1 - p(E)$

For finite sets Ω it is possible to define the probability as a mapping $p : \Omega \to [0,1]$. These probabilities p correspond to *classical mixtures*, i.e. general states of the classical object. *The random variable*, or observable in classical physics, is the real valued function $\mathbf{A} : \Omega \to \mathbb{R}$. The corresponding **probability distribution** $p_{\mathbf{A}}$ of $\mathbf{A}$ for the *probability space* (Ω, p) is defined as mapping $p_{\mathbf{A}} : \mathbf{A} \to [0,1]$

$$p_{\mathbf{A}}(a) := \sum_{\omega \in E(a)} p(\omega) \quad \text{where} \quad E(a) := \{\omega \in \Omega : \mathbf{A}(\omega) = a\} \tag{2.1}$$

Note, that the last expression is the *probability rule* $P(a \in \mathbf{A}, \varrho \equiv p)$ of the classical theory: each value $a \in \mathbf{A}$ determines a subset $E(a) := \{\omega \in \Omega : \mathbf{A}(\omega) = a\} \in \mathcal{E}(\Omega)$, that is, a corresponding event. *The mean value* $\langle \mathbf{A} \rangle_p$ of the random variable $\mathbf{A}$ is a *functional* defined on the set of random variables

$$\langle \mathbf{A} \rangle_p := \sum_{\omega} p(\omega) \mathbf{A}(\omega) = \sum_a a p_{\mathbf{A}}(a) \tag{2.2}$$

For two random variables $\mathbf{A}, \mathbf{B}$ we define the *joint distribution* denoted as $p_{\mathbf{AB}}(a,b)$ by the relation

$$p_{\mathbf{AB}}(a,b) := \sum_{\omega \in E(a,b)} p(\omega) \quad \text{where} \quad E(a,b) := E(a) \cap F(b) \tag{2.3}$$

with $E(a) := \{\omega \in \Omega, \mathbf{A}(\omega) = a\}$ and $F(b) := \{\omega \in \Omega, \mathbf{B}(\omega) = b\}$. Two *random events* $E, F \in \mathcal{E}(\Omega)$ are *independent*, if

$$p(E \cap F) = p(E) p(F) \tag{2.4}$$

Two *random variables* $\mathbf{A}, \mathbf{B} \in \mathcal{O}$ are *independent*, if the random events $E(a)$ and $F(b)$ are *independent* for all values $a \in \mathbf{A}$ and $b \in \mathbf{B}$. Independence is equivalent to the equation

$$p_{\mathbf{AB}}(a,b) \equiv p(E(a) \cap F(b)) = p(E(a))p(F(b)) \equiv p_{\mathbf{A}}(a)p_{\mathbf{B}}(b) \tag{2.5}$$

for all $(a,b) \in \mathbf{A} \times \mathbf{B}$. All these notions are extendible for the case of more than two variables. The collection of N observables $\mathbf{A}_k$ is independent, if the events $E_k(a_k) := \{\omega \in \Omega, \mathbf{A}_k(\omega) = a_k\}$ are independent for all the collections $(a_1, \ldots, a_N) \in \mathbf{A}_1 \times \ldots \times \mathbf{A}_N$, i.e. $p(\bigcap_k E_k(a_k)) = \prod_k p(E_k(a_k))$. The case of continuous Ω is a little more complicated, but this analysis goes beyond the scope of the present thesis.

The disjointness $E \cap F = \emptyset$ corresponds to the perfect distinguishability between the events. Note that for probability the following relation holds

$$p(E \cup F) = p(E) + p(F) - p(E \cap F) \tag{2.6}$$

Consider two events E, F corresponding to two different outcomes of the observation $\mathbf{A} = \{e, f\}$. We surely obtain one of the two possible outcomes, i.e. $P(e \text{ or } f, \varrho) = 1 = p(E \cup F)$. To achieve the unit probability on the left side of Eq.(2.6) we have to put $p(E) + p(F) = 1$ and $p(E \cap F) = 0$. In what follows the observation always defines the mutually disjoint set of events $E(a)$ corresponding to outcomes $a \in \mathbf{A}$.

The conditional probability $p_{\mathbf{B}}(b|a)$ characterizes the probability that the outcome of the observation $\mathbf{B}$ would be b providing that the outcome of $\mathbf{A}$ is a. It is defined by the *Bayes rule*

$$p_{\mathbf{B}}(b|a) = \frac{p_{\mathbf{AB}}(a,b)}{p_{\mathbf{A}}(a)} \tag{2.7}$$

for all a with $p_{\mathbf{A}}(a) \neq 0$. Writing the joint probability in the form $p_{\mathbf{AB}}(a,b) = p(E(a) \cap F(b))$ we obtain the conditional probability $p_{\mathbf{B}}(b|a) = \frac{p(E(a) \cap F(b))}{p(E(a))}$ expressed via original probability p. The conditional probability is zero if and only if $E(a) \cap F(b) = \emptyset$. If $p_{\mathbf{B}}(b|a) = p_{\mathbf{B}}(b)$ for all a, b then the observables $\mathbf{A}, \mathbf{B}$ are independent.

In the case of more than two random variables, we can define many types of *conditional probabilities*. For example, $p_{\mathbf{AB}}(a,b|c) := \frac{p_{\mathbf{ABC}}(a,b,c)}{p_{\mathbf{C}}(c)}$ is the conditional probability of the joint outcome (a,b), if we know that the outcome of the observation $\mathbf{C}$ was c. As we have said, three corresponding events $E(a), F(b), G(c)$ are independent, if $p(E(a) \cap F(b) \cap G(c)) = p_{\mathbf{ABC}}(a,b,c) = p_{\mathbf{A}}(a)p_{\mathbf{B}}(b)p_{\mathbf{C}}(c) = p(E(a))p(F(b))p(G(c))$. The conditional probability indicates a "three-partite" independence, if $p_{\mathbf{AB}}(a,b|c) = p_{\mathbf{AB}}(a,b) = p_{\mathbf{A}}(a)p_{\mathbf{B}}(b)$ for all permutations of a, b, c. That is, we can interpret the "three-partite" independence in the following way : if we know the outcome of one of the observations $\mathbf{A}, \mathbf{B}, \mathbf{C}$, then the other two observations are independent. It does not mean that these two random observables are independent without the knowledge of the outcome of the third variable (see Examples 1,2)! That is, $p_{\mathbf{AB}}(a,b) \neq p_{\mathbf{A}}(a)p_{\mathbf{B}}(b)$.

- **Example 1** (*Tripartite independence*)

In the following example we shall consider three events E, F, G. Note that each event (e.g. E) can play also a role of an observable ($\mathbf{E}$) called *characteristic function*, $\mathbf{E}(\omega) = 1$ if $\omega \in E$ and $\mathbf{E}(\omega) = 0$ if $\omega \notin E$. Put $\Omega = \{1,2,3,4,5\}$, $p(1) = p(2) = p(3) = \frac{1}{4}$, $p(4) = \frac{1}{6}$, $p(5) = \frac{1}{12}$, $E = \{1,5\}$, $F = \{2,4,5\}$, $G = \{3,4,5\}$. We have to check that $p(E \cap F \cap G) = \frac{1}{12} = p(E)p(F)p(G)$, i.e. they are independent. But $p(E \cap F) = \frac{1}{12} \neq p(E)p(F) = 16$, $p(E \cap G) = \frac{1}{12} \neq p(E)p(G) = 16$, $p(F \cap G) = \frac{1}{12} \neq p(F)p(G) = 14$. It follows they are not mutually independent. We say that the events are *bipartite correlated*, but *tripartite uncorrelated* (= independent).

- **Example 2** (*Tripartite dependence*)

Put $\Omega = \{1,2,3,4\}$, $p(i) = \frac{1}{4}$ for $i = 1,2,3,4$, $E = \{1,2\}$, $F = \{1,3\}$, $G = \{1,4\}$. Do the same as in the example before. $p(E \cap F \cap G) = \frac{1}{4} \neq p(E)p(F)p(G) = \frac{1}{8}$, but now $p(E \cap F) = \frac{1}{4} = p(E)p(F)$, $p(E \cap G) = \frac{1}{4} = p(E)p(G)$, $p(F \cap G) = \frac{1}{4} = p(F)p(G)$. Hence, in this case we have three events (random variables) *bipartite independent*, but *tripartite correlated.*

- **Example 3** (*Correlation function*)

In the last example we will show, that the *"standard" correlation function* $\mathcal{C} = |\langle \mathbf{AB} \rangle_p - \langle \mathbf{A} \rangle_p \langle \mathbf{B} \rangle_p|$ is not the best choice in order to detect the statistical dependence. Two random variables $\mathbf{A}, \mathbf{B}$ are uncorrelated (in terms of correlation function), if $\langle \mathbf{AB} \rangle_p = \langle \mathbf{A} \rangle_p \langle \mathbf{B} \rangle_p$. Put $\Omega = \{1, 2, 3\}$ and $p(i) = \frac{1}{3}$ for $i = 1, 2, 3$. Define $\mathbf{A}(1) = -1$, $\mathbf{A}(2) = -2$, $\mathbf{A}(3) = 3$, $\mathbf{B}(1) = 1$, $\mathbf{B}(2) = 4$, $\mathbf{B}(3) = 3$. It is easy to verify $\langle \mathbf{AB} \rangle_p = \langle \mathbf{A} \rangle_p \langle \mathbf{B} \rangle_p = 0$. Since the conditional probability $p(b = 1|a = -1) = p(b = 4|a = -2) = p(b = 3|a = 3) = 1$ and all the others do vanish, the information contained in $\mathbf{A}$ about $\mathbf{B}$ is maximal, that is to each $a \in \mathbf{A}$ corresponds exactly one $b \in \mathbf{B}$.

2.1.2 Measuring the information

Consider two random variables $\mathbf{A}, \mathbf{B}$. We have mentioned two bounds for information : (i) if these variables are independent, i.e. $p_{\mathbf{B}}(b|a) = p_{\mathbf{B}}(b)$ for all $(a, b) \in \mathbf{A} \times \mathbf{B}$ and $I(\mathbf{A}, \mathbf{B}) = 0$, and (ii) if the case when $I(\mathbf{A}, \mathbf{B})$ is maximal, i.e. for each $a \in \mathbf{A}$ there is $b \in \mathbf{B}$ such that $E(a) = F(b)$ and vice versa. In both cases the information is symmetric under the exchange of $\mathbf{A}$ and $\mathbf{B}$. In what follows we shall require the information to be symmetric also between these extremal cases, i.e. $I(\mathbf{A}, \mathbf{B}) = I(\mathbf{B}, \mathbf{A})$.

The occurrence of pairs (a, b) is described by joint probabilities $p_{\mathbf{AB}}(a, b)$. The probabilities $p_{\mathbf{A}}(a), p_{\mathbf{B}}(b)$ (called also **marginal**) are given by relations $p_{\mathbf{A}}(a) := \sum_{b \in \mathbf{B}} p_{\mathbf{AB}}(a, b)$ and $p_{\mathbf{B}}(b) := \sum_{a \in \mathbf{A}} p_{\mathbf{AB}}(a, b)$. The conditional probabilities are determined by formulas $p_{\mathbf{B}}(b|a) = p_{\mathbf{AB}}(a, b)/p_{\mathbf{A}}(a)$ and $p_{\mathbf{A}}(a|b) = p_{\mathbf{AB}}(a, b)/p_{\mathbf{B}}(b)$. We see that the knowledge of $p_{\mathbf{AB}}(a, b)$ is sufficient to define all the probabilities, that implicitly depend on the original probability p defined on Ω, where p represents a state of the system under consideration.

In particular, the maximum of information is achieved by "self-information" contained in $\mathbf{A}$ about $\mathbf{A}$, i.e. $I(\mathbf{A}, \mathbf{A})$. But what does such a quantity really mean? Consider we observe this quantity $\mathbf{A}$ repeatedly recording a long sequence of observed outcomes. Imagine somebody else who will measure again the same quantity $\mathbf{A}$ on each object, that we had already measured. How precisely are we able to determine the sequence of his outcomes? That is, what is the outcome of a single observation of $\mathbf{A}$ (performed by him) following the observation of $\mathbf{A}$ (performed by us) on the same copy of the system? Of course, if we once observed a, then each other measurement of $\mathbf{A}$ performed on the same object must result in a, too. It means the sequence measured in the second measurement of $\mathbf{A}$ realized in the described way can be predicted with certainty.

In general case, the information in $\mathbf{A}$ about $\mathbf{B}$ is understood in the same way. It means we measure $\mathbf{A}$ first and we try to predict the outcome of the measurement $\mathbf{B}$ based on our knowledge about the outcome of $\mathbf{A}$. Here again the observation of $\mathbf{B}$ and $\mathbf{A}$ is performed on the same single object. Consider we have recorded a sequence of outcomes of repeated measurements of $\mathbf{A}$, i.e. $\vec{a} = (a_{j_1}, \ldots, a_{j_N}) \in \mathbf{A} \times \ldots \times \mathbf{A}$. What is then the number of all possible sequences $\vec{b} = (b_{j_1}, \ldots, b_{j_N}) \in \mathbf{B} \times \ldots \times \mathbf{B}$ of the repeated observation of B, if we know the sequence of $\mathbf{A}$? This question is very similar to the original one, i.e. *How much information do we have about ..., if we know that ...?* It follows that the number of sequences $\Omega_{\mathbf{B}}(N, \vec{a})$ corresponds to the amount of information about $\mathbf{B}$, if we know the sequence of outcomes $\vec{a}$. The larger the number of sequences $\Omega_{\mathbf{B}}(N, \vec{a})$, the less information we have about the observables $\mathbf{B}$. The number $\Omega_{\mathbf{B}}(N, \vec{a})$ can be understood as a random variable $\mathbf{A}^N \to \mathbb{R}$ with the probability $p_{\mathbf{A}^N}(\vec{a})$ defined on $\mathbf{A}^N = \mathbf{A} \times \ldots \times \mathbf{A}$. The mean value of such a random variable (denoted by $\Omega_{\mathbf{B}}(N, \mathbf{A})$) will represent the information on average. For example in the case of independent observables, the value $\Omega_{\mathbf{B}}(N, \mathbf{A}) = \Omega_{\mathbf{B}}(N)$ is maximal. On the other hand, for deterministic observables the sequence can be predicted precisely, i.e. $\Omega_{\mathbf{B}}(N, \mathbf{A}) = 1$.

The fraction of numbers $\Gamma_{\mathbf{B}}(N, \mathbf{A}) := \Omega_{\mathbf{B}}(N, \mathbf{A})/\Omega_{\mathbf{B}}(N)$ represents our information gain. We shall define the information by the formula

$$I(\mathbf{A}, \mathbf{B}) = \lim_{N \to \infty} -\frac{\log \Gamma_{\mathbf{B}}(N, \mathbf{A})}{N} \tag{2.8}$$

where the logarithm is used to obtain $I(\mathbf{A}, \mathbf{B}) = 0$ for independent variables, when $\Gamma_{\mathbf{B}}(N, \mathbf{A}) = 1$. The minus ensures the positivity of information, because the defined fraction is always less then one.

• **Example 4** (*On the symmetry of information*)

Choose the observations $\mathbf{A} = \{1,2,3,4,5,6\}$ and $\mathbf{P} = \{odd, even\}$. If we measure $a \in \mathbf{A}$, then we know the outcome of the observation $\mathbf{P}$ with certainty. But, if we measure $\mathbf{P}$, then the outcome of $\mathbf{A}$ is still non-deterministic, $\Omega_{\mathbf{A}}(N,\mathbf{P}) \neq 1$. It seems that the knowledge of $\mathbf{P}$ about $\mathbf{A}$ is different than the knowledge of $\mathbf{A}$ about $\mathbf{P}$. Have we lost the symmetry of information? We have defined the information as a fraction and its value depends on the number of all possibilities. Although in this case $\Omega_{\mathbf{P}}(N) \neq \Omega_{\mathbf{A}}(N)$, it is still possible that the fractions $\Gamma_{\mathbf{P}}(N,\mathbf{A})$ and $\Gamma_{\mathbf{A}}(N,\mathbf{P})$ are equal. That is,

$$\Gamma_{\mathbf{P}}(N,\mathbf{A}) = \frac{\Omega_{\mathbf{P}}(N,\mathbf{A})}{\Omega_{\mathbf{P}}(N)} = \frac{1}{\Omega_{\mathbf{P}}(N)} = \frac{\Omega_{\mathbf{A}}(N,\mathbf{P})}{\Omega_{\mathbf{A}}(N)} = \Gamma_{\mathbf{A}}(N,\mathbf{P}). \tag{2.9}$$

We shall see that information (and also the fractions) is symmetric and there is no loophole in our consideration. As we have said, information is a *relative notion.*

Next, we shall calculate the numbers $\Omega_{\mathbf{B}}(N,\mathbf{A})$, $\Omega_{\mathbf{B}}(N)$ and $\Gamma_{\mathbf{B}}(N,\mathbf{A})$ which are defined above. Let us start with $\Omega_{\mathbf{B}}(N,\vec{a})$. Since the outcomes of the repeated observations are considered to be random and independent, the conditional probability is given by

$$p_{\mathbf{B}^N}(\vec{b}|\vec{a}) = \prod_{k=1}^{N} p_{\mathbf{B}}(b_{j_k}|a_{j_k}) \tag{2.10}$$

If the sequence $\vec{a}$ was measured, then not all sequences $\vec{b}$ are allowed, because not all of them satisfy the statistical properties. The probability of such **typical sequence** $\vec{b} \in \mathbf{B}^N$, $p_{\mathbf{B}^N}(\vec{b}|\vec{a})$, for a fixed $\vec{a}$ should approximately satisfy the relation

$$\begin{aligned} -\frac{1}{N}\log p_{\mathbf{B}^N}(\vec{b}|\vec{a}) &= -\frac{1}{N}\log \prod_{(a,b)} p_{\mathbf{B}}(b|a)^{N(a,b)} \pm \delta \\ &= -\sum_{(a,b)} p_{\mathbf{AB}}(a,b)\log p_{\mathbf{B}}(b|a) \pm \delta \\ &= H(\mathbf{B}|\mathbf{A}) \pm \delta \end{aligned} \tag{2.11}$$

where $H(\mathbf{B}|\mathbf{A}) := -\sum_{(a,b)} p_{\mathbf{AB}}(a,b)\log p_{\mathbf{B}}(b|a)$. We shall denote the set of typical sequences $\vec{b} \in \mathbf{B}^N$ by $\mathcal{T}_{\mathbf{B}}(\vec{a},N) \subset \mathbf{B}^N$. Note that the number of the sequences equals $\Omega_{\mathbf{B}}(N,\vec{a})$. This number does not depend on $\vec{a}$, because of the Eq.(2.11) that implies $\Omega_{\mathbf{B}}(N,\mathbf{A}) \equiv \Omega_{\mathbf{B}}(N,\vec{a})$ for all allowed sequences $\vec{a}$. The probability of each typical sequence is approximately the same, $p_{\mathbf{B}^N}(\vec{b}|\vec{a}) \approx 2^{-NH(\mathbf{B}|\mathbf{A})}$. The main result of the information theory is that the probability $p_{\mathbf{B}^N}(\mathcal{T}_{\mathbf{B}}(\vec{a},N)|\vec{a})$ of the set of typical sequences tends to unity with $N \to \infty$ and, consequently, the number of them equals roughly to

$$\Omega_{\mathbf{B}}(N,\mathbf{A}) = 2^{N(H(\mathbf{B}|\mathbf{A})\pm\delta)} \tag{2.12}$$

The number of all sequences $\vec{b} \in \mathbf{B}^N$ that can occur is given by typical sequences $\mathcal{T}_{\mathbf{B}}(N,\delta)$ and is defined in a similar way, but with different probability distribution $p_{\mathbf{B}^N}(\vec{b}) = \prod_k p_{\mathbf{B}}(b_{j_k})$. The condition determining such sequences $\vec{b}$ is

$$-\frac{1}{N}\log p_{\mathbf{B}}^N(\vec{b}) = -\frac{1}{N}\log \prod_{k=1}^{N} p_{\mathbf{B}}(b_{j_k}) \pm \delta \tag{2.13}$$

$$= -\sum_{b} p_{\mathbf{B}}(b)\log p_{\mathbf{B}}(b) \pm \delta \tag{2.14}$$

$$= H(\mathbf{B}) \pm \delta \tag{2.15}$$

where we used similar trick as before, i.e. $p_{\mathbf{B}}(b) = N(b)/N$ and $H(\mathbf{B}) := -\sum_b p_{\mathbf{B}}(b) \log p_{\mathbf{B}}(b)$. Hence, the number $\Omega_{\mathbf{B}}(N) = 2^{N(H(\mathbf{B}) \pm \delta)}$. The information can be expressed as

$$\begin{aligned} I(\mathbf{A}, \mathbf{B}) &= -\lim_{N \to \infty} \frac{1}{N} \log \frac{\Omega_{\mathbf{B}}(N, \mathbf{A})}{\Omega_{\mathbf{B}}(N)} & (2.16) \\ &= -\lim_{N \to \infty} \frac{1}{N} \log 2^{N(H(\mathbf{B}|\mathbf{A}) - H(\mathbf{B}) \pm \delta)} & (2.17) \\ &= H(\mathbf{B}) - H(\mathbf{B}|\mathbf{A}) & (2.18) \end{aligned}$$

where finally we put $\delta = 0$. Introducing the definitions of H functions and the relations between the mentioned probabilities we get for the information

$$I(\mathbf{A}, \mathbf{B}) = \sum_{(a,b)} p_{\mathbf{AB}}(a, b) \log \frac{p_{\mathbf{B}}(b|a)}{p_{\mathbf{B}}(b)}. \quad (2.19)$$

What about the symmetry of this function? Performing the same steps only exchanging the observables **A** and **B** we obtain

$$I(\mathbf{B}, \mathbf{A}) = \sum_{(a,b)} p_{\mathbf{AB}}(a, b) \log \frac{p_{\mathbf{A}}(a|b)}{p_{\mathbf{A}}(a)}. \quad (2.20)$$

If we use the *Bayes rule* in the form $p_{\mathbf{AB}}(a, b) = p_{\mathbf{B}}(b|a) p_{\mathbf{A}}(a) = p_{\mathbf{A}}(a|b) p_{\mathbf{B}}(b)$, the symmetry is evident.

2.2 Communication

In the previous section we introduced the notion of *information* as the measure of correlation, or degree of dependence of two random variables $\mathbf{A}, \mathbf{B}$ measured on the same object in a given state. Intuitively, this notion concerns also the possibility to communicate it among some individuals, to transfer it from sender to receiver, etc. Next we shall introduce a model of communication, in which the information characterizes the reliability of the information transfer.

The most important is to have a *source of information*, which produces information we want to communicate. Imagine a standard communication between Alice and Bob. Both Alice and Bob can be viewed as a source of information. In fact, the information arises in their minds. If their distance is small enough, they can simply talk together using their voices[1]. Increasing their mutual distance they have to use other (more sophisticated) types of communication, for example telegraphy, or telephones. In both cases they transform the original information into the sequence of *signals*, i.e. into the words, or letters, or some other symbols representing the letters. From the physical point of view different signals correspond to different states of a physical object, that carries the information from a **sender** to a **receiver**. The received signals are finally transformed back into the original message.

Hence, the communication consists of three steps:

1. **Encoding the information.** First of all, before the encoding transformation, we somehow need to read the message produced by a source. That is, we have to perform an observation **A**, with which we are able to distinguish between the possible messages. Different outcomes $a \in \mathbf{A}$ compose the set of possible messages. The encoding $\mathcal{C}$ is a transformation of these outcomes into signals, or sequences of signals, i.e. $\mathcal{C} : \mathbf{A} \to \mathbf{X}^n$, where $\mathbf{X}$ is the set of signals x which can be transmitted via the channel we use. For example, standard letters can be represented by a sequence of dots and dashes, as in the case of telegraphy (*Morse alphabet*). In the *digital encoding* we transform all letters into the sequence of eight binary digits.

[1] We overpass the possibility of explaining their feelings by sounds and gestures, and concentrate only on communicating the words.

2. **Transmission via the channel.** We consider both, the encoding and decoding transformation to be perfect. The main problem of communication is the **noise** which can arise during the transmission. In general, the interaction between the channel and the environment cannot be neglected and our description of these influences might be only probabilistic. As a result we obtain that the state evolution of the signals is described in terms of probabilities. Namely we use *conditional probabilities* $p_{\mathbf{Y}}(y|x)$ to describe the transmission, where $\mathbf{Y}$ is the set of all received signals. Note, that the number of elements in $\mathbf{X}$ and $\mathbf{Y}$ may not be the same.

3. **Decoding the received signals.** Decoding transformation $\mathcal{D}$ is in some sense inverse map to $\mathcal{C}$, i.e. $\mathcal{D} : \mathbf{Y}^m \to \mathbf{B}$, where the set $\mathbf{B}$ represents the set of all possible received messages. The received message $b \in \mathbf{B}$ is read out by performing an observation $\mathbf{B}$. If we suppose that we have one-to-one correspondence between the sets $\mathbf{B}$ and $\mathbf{A}$, i.e. between the sent and the received messages, then the problem is whether the sequences $\vec{y}$ and $\vec{x}$ represent the same message. The main question of communication is how to minimize the effect of noise in the channel.

The aim now is a little bit different than in the previous section. There we had physical system in one defined state and we were asking about the correlation between two measurements performed on the same copy of the system. Now we have again two observations $\mathbf{X}$ and $\mathbf{Y}$ realized in the same way as before, but the state of the system is changing during the transmission. However, the derived formalism can be used also in this case, because the evolution can be understood as a part of the receiver's measurement.

In particular, after an encoding the letters $x \in \mathbf{X}$ occur with some probabilities $\pi_{\mathbf{X}}(x)$. Using the conditional probability determining the channel, $p_{\mathbf{Y}}(y|x)$, for the joint probability we get

$$p_{\mathbf{XY}}(x,y) = \pi_X(x) p_{\mathbf{Y}}(y|x). \tag{2.21}$$

Having the joint probability we have everything in hands that is needed to define the information contained in observation $\mathbf{Y}$ about the preparation $\mathbf{X}$. In other words, based on the knowledge of a received sequence $\vec{y}$, we want to reconstruct the original message encoded into the sequence $\vec{x}$.

For channels we are interested in the maximal value of information they can transmit. Since the channel is given by the conditional probability $p_{\mathbf{Y}}(y|x)$, the only thing we can vary is the input signal probability $\pi_{\mathbf{X}}(x)$. The so-called **channel capacity** is then defined by formula

$$C = \max_{\pi} I(\mathbf{Y}, \mathbf{X}, \pi) = \max_{\pi} \sum_{(x,y)} p_{\mathbf{XY}}(x,y) \log \frac{p_{\mathbf{Y}}(y|x)}{\pi_{\mathbf{X}}(x)}. \tag{2.22}$$

• **Example 5** (*Classical bit*)

Consider the simplest classical object called *bit*. It is given by two elements sample space $\Omega = \{0, 1\}$. Such object is often used as signal system in the modern communication. As we have mentioned the symbols of the standard alphabet are encoded into the sequence of eight bits. The channels using such signals are called *digital*. We shall calculate the capacity of such digital channels. Put $\pi_{\mathbf{X}}(0) = \pi$ and $\pi_{\mathbf{X}}(1) = 1 - \pi$. Consider the channel is symmetric in the following sense

$$p_{\mathbf{Y}}(0|0) = p_{\mathbf{Y}}(1|1) = p \quad \text{and} \quad p_{\mathbf{Y}}(0|1) = p_{\mathbf{Y}}(1|0) = 1 - p. \tag{2.23}$$

Then for the capacity we have

$$\begin{aligned} C(p) &= \max_{\pi} \left(\pi \log \frac{1-\pi}{\pi} + p \log \frac{p}{1-p} + \log \frac{1-p}{1-\pi} \right) \qquad (2.24) \\ &= p \log p + (1-p) \log(1-p) + \max_{\pi} \left(-\pi \log \pi - (1-\pi) \log(1-\pi) \right) \\ &= \max_{\pi} H(\pi) - H(p) \end{aligned}$$

where $H(p) := -p \log p - (1-p) \log(1-p)$ is the *binary entropy function* with its maximum achieved in $p = 1/2$. For the capacity of symmetric digital channel we get

$$C(p) = \log 2 - H(p) = 1 - H(p). \tag{2.25}$$

In the limit of noiseless channel $p = 1$ we get that the capacity is equal to one, i.e. one signal bit communicate one bit of information.

To conclude this part, let us stress that the information describes the number of bits transferred via the channel per one signal object. Of course, the problem is more complicated and we have introduced here only a brief sketch of the communication problem, which is relevant in the context of the present thesis. In fact, we have considered that transmitted signals are mutually independent, i.e. that channel is *memoryless*. We used the same information function in two different contexts : as a measure of correlations between two observables performed on the same object and as transmission rate in communication.

2.3 Concept of entropy

One of the most "mysterious" physical quantity takes the name of **entropy**. It was introduced by *Clausius* in his work concerning *thermodynamics* and since then the discussions about its meaning have started. Non-decreasing of entropy represents a simple form of the *second law of thermodynamics* that determines the possible physical processes. Suppose we measure the energy of a system with the probabilities $p_{\mathbf{E}}(E_k)$ of outcomes E_k. Following *Boltzmann* the *thermodynamic entropy* is defined as

$$S = -kT \ln \Theta(N, \mathbf{E}, p) \tag{2.26}$$

where $k = 1,38.10^{-23} JK^{-1}$ is the *Boltzmann constant*, T is temperature and $\Theta(N, \mathbf{E}, p)$ is the number of all possible sequences of N repeated observations described by the probability $p_{\mathbf{E}}(E_k)$, or the number of *microstates* with the mean value of energy equal to $\langle \mathbf{E} \rangle_p$ for a given state p. This number $\Theta(N, \mathbf{E}, p)$ reflects our intuitive notion of **uncertainty** about the energy observation of the system. It means *thermodynamic entropy* measures our uncertainty about the physical system.

How to calculate $\Theta(N, \mathbf{E}, p)$ for a given probability $p_{\mathbf{E}}(E_k)$? The number of occurrences of a single outcome E_k should be about $N(E_k) = Np_{\mathbf{E}}(E_k)$. The number of all such sequences is given by formula

$$\Theta(N, \mathbf{E}, p) = \frac{N!}{\prod_k N(E_k)!}. \tag{2.27}$$

Note, that not every typical sequence defined before is an element of $\Theta(N, \mathbf{E}, p)$, because the sequences are defined by their probabilities and not by the number of occurrences of single outcomes. The number of all sequences is equal to $\Theta(N) = K^N$, where K is the number of different values of energy $\mathbf{E}$. Taking the logarithm of Eq.(2.27) and using the *Stirling approximation* we obtain (in the logarithmic sense) $\ln \Theta(N, \mathbf{E}, p) \approx NS(p_{\mathbf{E}})$, where $S(p_{\mathbf{E}}) := -\sum_k p_{\mathbf{E}}(E_k) \ln p_{\mathbf{E}}(E_k)$. For the *thermodynamic entropy* we can write

$$S = -kTN \sum_k p_{\mathbf{E}}(E_k) \ln p_{\mathbf{E}}(E_k). \tag{2.28}$$

This function measures our ability to predict the outcomes of the repeated observations of the energy. In a specific case, if there is only one possible outcome, i.e. $p_{\mathbf{E}}(E_k) = 1$ for some E_k, the *thermodynamic entropy* is zero, because our uncertainty is zero. But, if the distribution of the energy is totally random, i.e. $p_{\mathbf{E}}(E_k) = 1/K$ for all E_k, then our uncertainty about the energy is (intuitively) maximal and, of course, *thermodynamic entropy* takes its maximum $S = kTN \ln K$.

We shall generalize this notion for each probability distribution $p_{\mathbf{A}}(a)$. Define the *entropy function* by formula

$$H(\mathbf{A}) := -\sum_a p_{\mathbf{A}}(a) \log p_{\mathbf{A}}(a) \tag{2.29}$$

where we choose the logarithm with the base 2. Let us note that the entropy function of the measurement of energy and the *thermodynamic entropy* are of different values. Such quantity defined for an observation $\mathbf{A}$ represents our ability to predict its outcomes. Formally, the entropy function is the same like the one used in the definition of information.

We have mentioned that states represent our *states of knowledge* about the system. In classical physics the usage of probabilities reflects our practical inability to observe all the physical quantities of the system. It means the knowledge about the state is not complete, but still is maximal we can have. Applying the above definition of entropy to a state p we obtain $H(p) = 0$, if and only if $p(\omega) = 1$ for some $\omega \in \Omega$, i.e. if the state of the system is pure and we have full "information" about the system. It means we are able to predict outcomes of all observations with certainty. The *entropy of the state* $H(p) = -\sum_\omega p(\omega) \log p(\omega)$ reflects the *quality* of our state of knowledge. Here we used the word "information" in different context as before. By "information" we mean the characterization of our knowledge about the system, if the state p is given. Sometimes $H(p)$ is called *negative information*, because $H(p) = 0$ indicates maximal possible "information" we can have.

We can define many types of entropy functions. The *joint entropy* $H(\mathbf{A}, \mathbf{B})$ is defined for a joint probability distribution $p_{\mathbf{AB}}(a, b)$. The entropy of conditional probability $p_{\mathbf{B}}(b|a)$ with the fixed $a \in \mathbf{A}$ is defined in the standard way, i.e. $H_a(\mathbf{B}) := -\sum_b p_{\mathbf{B}}(b|a) \log p_{\mathbf{B}}(b|a)$. The *conditional entropy* is the mean value of random variables $H_a(\mathbf{B})$ averaged over the set $\mathbf{A}$, i.e. $H(\mathbf{B}|\mathbf{A}) := \sum_a p_{\mathbf{A}}(a) H_a(\mathbf{B}) = \sum_{(a,b)} p_{\mathbf{AB}}(a, b) \log p_{\mathbf{B}}(b|a)$. The equivalent definition of the conditional entropy is $H(\mathbf{B}|\mathbf{A}) = H(\mathbf{A}, \mathbf{B}) - H(\mathbf{A})$. Another important function is given by

$$H(\mathbf{A} : \mathbf{B}) = H(\mathbf{A}) + H(\mathbf{B}) - H(\mathbf{A}, \mathbf{B}) \tag{2.30}$$

and is equivalent to the *information function*, i.e. $I(\mathbf{A}, \mathbf{B}) \equiv H(\mathbf{A} : \mathbf{B})$. The collection of all the possible entropy functions is very huge and we shall not list them all.

Denote the set of probability distributions on a sample space Ω by $\mathcal{P}(\Omega)$. Let $p, q \in \mathbf{P}(\Omega)$. Then the relation

$$H(p||q) := \sum_\omega p(\omega) \log \frac{p(\omega)}{q(\omega)} \tag{2.31}$$

defines the **relative entropy** of two distributions $p, q \in \mathcal{P}(\Omega)$. Important property of this function is that $H(p||q) = 0$ if and only if $p \equiv q$. Consider the probability distribution $p_{\mathbf{AB}}(a, b)$ on the sample space $\mathbf{A} \times \mathbf{B}$. The probability distribution $p_{\mathbf{A}\times\mathbf{B}}(a \times b)$ defined on $\mathbf{A} \times \mathbf{B}$ by equation $p_{\mathbf{A}\times\mathbf{B}}(a, b) = p_{\mathbf{A}}(a) p_{\mathbf{B}}(b)$ corresponds to independent variables $\mathbf{A}$ and $\mathbf{B}$. It is easy to check that

$$H(p_{\mathbf{AB}}||p_{\mathbf{A}\times\mathbf{B}}) = \sum_{(a,b)} p_{\mathbf{AB}}(a, b) \log \frac{p_{\mathbf{AB}(a,b)}}{p_{\mathbf{A}}(a) p_{\mathbf{B}}(b)} \equiv I(\mathbf{A}, \mathbf{B}), \tag{2.32}$$

i.e. the information function can be defined via relative entropy function. The relative entropy determines the "distance" of two probability distributions [2]. Using such interpretation we can say, that the information measures the distance between the joint probability distribution and the canonical distribution corresponding to independent observations. Since the relative entropy vanishes only if $p_{\mathbf{AB}}(a, b) \equiv p_{\mathbf{A}}(a) p_{\mathbf{B}}(b)$, it follows that correlation is zero, if and only if the observations $\mathbf{A}$ and $\mathbf{B}$ are independent. As a consequence of this fact we obtain that information is really an appropriate measure of correlations.

2.4 Correlated systems

Let us suppose two physical systems. Could we say anything about the correlations between their observables? Yes, of course. We only need to know the probability of the joint outcome $p_{\mathbf{AB}}(a, b)$ of the measurements $\mathbf{A}$ realized on the first system A and $\mathbf{B}$ realized on the second system B. But what does it mean to find a joint outcome? The sample space for two systems A and B is given as Cartesian

[2] The relative entropy, unlike the distance, is not symmetric

product $\Omega = \Omega_A \times \Omega_B$ and this is also the set of joint outcomes. The first step of each observation is the *preparation process*, where we must specify, which objects correspond to the whole physical system $A + B$. The preparation process makes the question of correlations reasonable, because it allows us the repetition of the observations. For example, one can find that the daily expected value of the temperature in the South Pole is maximally correlated with the number of girls born in China per day. Nobody can take such correlations as "real" in the sense that these two events are physically related. But in this case we did not perform any joint preparation. We only read some results of some experiments and compare them *a posteriori* to find $p_{\mathbf{AB}}(a, b)$ and correlations. And it is, of course, meaningless in the physical context. We cannot expect that by increasing the temperature the population of China will decrease (or increase), too.

We often say that two systems are correlated without saying something about the observations. As we said, we only need the joint probability distribution. Let p_{AB} be the probability defined on $\Omega_A \times \Omega_B$, i.e. the classical state of the composed system $A + B$. Define the **correlation between the subsystems** A and B

$$C_p(A, B) = I(A, B) = H(A) + H(B) - H(A, B) \tag{2.33}$$

where $H(A)$ is the entropy of state $p_A(\omega_A) := \sum_{\omega_B} p_{AB}(\omega_A, \omega_B)$ and $H(B)$ is the entropy of $p_B(\omega_B) := \sum_{\omega_A} p_{AB}(\omega_A, \omega_B)$. We will say that state p_{AB} is **correlated** if and only if $C_p(A, B) \neq 0$. But, what does such formal definition really mean? In some sense the correlations of states compares two different description of the composed system. We have mentioned that the entropy of the state corresponds to the *negative information.* The correlation $C_p(A, B)$ quantifies the difference between the "information" we have in two different cases : (i) if we describe the whole system $A+B$ by p_{AB}, and (ii) if we describe only the subsystems A by p_A and B by p_B independently. It vanishes for independent subsystems. The nonzero correlations implies that two observers A and B have more information (= are less uncertain) about the whole system, if they are allowed to cooperate and communicate. From the point of view of measurements, the local measurements of subsystems for *uncorrelated states* ($C_p(A, B) = 0$) are always statistically independent. In the case of *correlated states* the subsystems "feel" each other, i.e. there exists quantities (measurements) of the subsystems which are mutually correlated.

Chapter 3

Entanglement - just two qubits

We all agree that our theory is crazy.
But is it crazy enough?
N.Bohr

3.1 Probabilities in quantum theory

In the previous the elements of two theories: Quantum Theory and Information Theory have been described. In this chapter we shall continue to connect these two theories. In both of them the concept of probability plays a very important role. The classical probabilistic theory, which is used in the Information Theory, is defined on the *sample space* Ω. In what follows, we shall assume that the set Ω contains a finite number of elements. Then the *classical states* are associated with the set of all probability distributions $\mathcal{P}(\Omega)$ and the set of random variables $\mathcal{O}$ represents the classical observables. The concept of *random event* $E \subset \Omega$ enables us to define a relation (called *probability rule*) between the theory and experimental reality. In particular, for finite sample spaces Ω each subset $E \subset \Omega$ represents a random event. The classical random variable $\mathbf{A} \in \mathcal{O}$ is determined by the set of mutually disjoint sets $E(a) \in \Omega$ such that $\bigcup_a E(a) = \Omega$. The index a distinguishes different outcomes of the observable $\mathbf{A}$. The probability of measuring the outcome a is given by the classical probability rule $P(a,p) = p(E(a)) = \sum_{j\in E(a)} p_j$, where $p : \Omega \to [0,1]$ is the classical state and j are elements of Ω.

In the Quantum Theory the probability rule is given by the equation $P(a,\varrho) = \mathrm{Tr}\varrho\mathbf{E}(a)$, where $\mathbf{E}(a)$ is a projective operator. Therefore, the quantum analogy of the random event - *quantum event* is represented by a projector $\mathbf{E}(a)$. That is, the set of quantum events is the set of all projectors that has different properties in comparison with the σ-algebra of classical events. The quantum observable is represented by the set of *mutually disjoint* (= *orthogonal*) projectors $\mathbf{E}(a)$, i.e. $\mathbf{E}(a)\mathbf{E}(a') = 0$ for $a \neq a'$ and, moreover, $\sum_a \mathbf{E}(a) = \mathbb{1}$. Hence, the operator product plays the role of the set union of the classical events. However, there is one difference between these two operations. Whereas the union of any two classical events is again a classical event, the product of two quantum events $\mathbf{EF}$ is a quantum event only if $\mathbf{E}$ and $\mathbf{F}$ commute. Simply, the product $\mathbf{EF}$ is a projection, i.e. $\mathbf{EF} = (\mathbf{EF})^\dagger = (\mathbf{EF})^2$, only if $[\mathbf{E},\mathbf{F}] = 0$.

This property is related to the so-called *uncertainty principle*, which says that in Quantum Theory for each state ϱ there exists a measurement $\mathbf{A} = \sum_a aE(a)$ with an undefined value, i.e. for all outcomes a the probabilities are less then one. For classical pure states the outcomes of all observations are determined with certainty. The Quantum Theory is not deterministic in this sense, i.e. there always exist measurements with a non-trivial statistics.

3.1.1 Entropies and correlations of quantum systems

The concept of entropy for a probability distribution is purely mathematical and therefore independent of the physical theory. Each observation $\mathbf{A} \in \mathcal{L}_s(\mathcal{H})$ determines a probability distribution and we can define its entropy by the formula

$$H(\mathbf{A}, \varrho) = \sum_a p_{\mathbf{A}}(a, \varrho) \log p_{\mathbf{A}}(a, \varrho) \tag{3.1}$$

where $p_{\mathbf{A}}(a, \varrho) = \mathrm{Tr}\mathbf{E}(a)\varrho$. Consequently, the introduced definition of the correlation between two measurements can be used also in the quantum case. In classical physics we went further. We have introduced the *entropy of a state* without any direct relevance to the performed measurements. Can we somehow generalize this notion into the field of Quantum Theory? We need to analyze, what the entropy of state means in classical physics. We have seen that it vanishes for pure states, i.e. there exists a measurement, for which an outcome is determined. Of course, in classical physics we know the outcome of each measurement, but it is not the case in quantum physics. In what follows, we can define an entropy of quantum state as a minimal entropy of the outcome probabilities (minimized over all observations), i.e.

$$S(\varrho) := \min_{\mathbf{A}} H(\mathbf{A}, \varrho). \tag{3.2}$$

The minimum is achieved for the observation $\mathbf{A}$ that commutes with the state operator ϱ, that is the probabilities of outcomes are given by the eigenvalues of ϱ. In spectral form $\varrho = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k|$ and

$$S(\varrho) = -\sum_k \lambda_k \log \lambda_k. \tag{3.3}$$

To see that this value is really minimal, let us choose a general observation $\mathbf{A}$ with the eigenprojectors $|\phi_a\rangle\langle\phi_a|$. Then the probability rule implies $P(a, \varrho) = \sum_k \lambda_k |\langle\psi_k|\phi_a\rangle|^2 = \sum_k \lambda_k p_{ka}$. Our aim is to compare the values $S(\varrho)$ and $H(\mathbf{A}, \varrho)$. Let us calculate the difference

$$\begin{aligned} H(\mathbf{A}, \varrho) - S(\varrho) &= \sum_k \lambda_k \log \lambda_k - \sum_{ka} \lambda_k p_{ka} \log(\sum_l \lambda_l p_{la}) \\ &= \sum_k \lambda_k \left(\sum_a p_{ka} \log \lambda_k - \sum_a p_{ka} \log(\sum_l \lambda_l p_{la}) \right) \\ &= \sum_{k,a} \lambda_k p_{ka} \log \frac{\lambda_k}{\sum_l \lambda_l p_{la}} \end{aligned} \tag{3.4}$$

where we used $\sum_a p_{ka} = 1$ in the second equation. Since $\log x \geq 1 - x^{-1}$ we get the following inequality

$$H(\mathbf{A}, \varrho) - S(\varrho) \geq \sum_{ka} \lambda_k p_{ka} (1 - \frac{\sum_l \lambda_l p_{la}}{\lambda_k}) = 1 - \sum_{k,a} p_{ka} \sum_l \lambda_l p_{la}. \tag{3.5}$$

Using the identity $\sum_k p_{ka} = 1$ and $\sum_{a,l} \lambda_l p_{la} = \sum_a P(a, \varrho) = 1$ we find that

$$H(\mathbf{A}, \varrho) - S(\varrho) \geq 0. \tag{3.6}$$

It means that the function $S(\varrho) = -\sum_k \lambda_k \log \lambda_k$ is really the minimal value and the observation commuting with the state ϱ is really the optimal one in this sense. Using the functional calculus we get the famous *von Neumann* formula for the *entropy of quantum state*

$$S(\varrho) = -\mathrm{Tr}\varrho \log \varrho\ . \tag{3.7}$$

In a similar way we can define the *quantum relative entropy*

$$S(\varrho||\sigma) = \mathrm{Tr}\varrho[\log\varrho - \log\sigma]. \tag{3.8}$$

Consider two quantum systems A and B described by the state $\varrho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then the *quantum joint entropy* is given as $S(\varrho_{AB}) = -\mathrm{Tr}\varrho_{AB}\log\varrho_{AB}$. The *quantum conditional entropy* can be formally defined by the formula $S(\varrho_{A|B}) := S(\varrho_{AB}) - S(\varrho_B)$, where $\varrho_A = \mathrm{Tr}_B\varrho_{AB}$. The definition of the *quantum mutual entropy* is $S(\varrho_{A:B}) := S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB})$. For the factorized states $\varrho_{AB} = \varrho_A \otimes \varrho_B$ this quantity vanishes, because $S(\varrho_A \otimes \varrho_B) = S(\varrho_A) + S(\varrho_B)$.

In what follows we shall list (without any proofs[1]) some of the properties of the entropy functions defined above.

1. *Unitary invariance:* $S(\varrho) = S(U\varrho U^\dagger)$ for all $U \in \mathcal{U}(\mathcal{H})$.
2. *Subadditivity:* $S(\varrho_{AB}) \leq S(\varrho_A) + S(\varrho_B)$
3. *Araki-Lieb inequality:* $S(\varrho_{AB}) \geq |S(\varrho_A) - S(\varrho_B)|$
4. *Strong subadditivity:* $S(\varrho_{ABC}) + S(\varrho_B) \leq S(\varrho_{AB}) + S(\varrho_{BC})$
5. *Independence:* $S(\varrho_{A:B}) = 0$ if and only if $\varrho_{AB} = \varrho_A \otimes \varrho_B$
6. *Klein's inequality:* $S(\varrho||\sigma) \geq 0$ with equality if and only if $\varrho = \sigma$
7. *Ensemble inequality:* $S(\varrho) \leq H(p_j) + \sum_j p_j S(\varrho_j)$ where $\varrho = \sum_j p_j\varrho_j$
8. *Concavity:* $S(\sum_j p_j\varrho_j) \geq \sum_j p_j S(\varrho_j)$, where $\sum_j p_j = 1$ and $p_j \geq 0$

The properties 2 and 5 assure that $S(\varrho_{A:B})$ is a good measure of bipartite correlations between two systems. Intuitively, the maximally correlated system is described by state $\varrho_{AB} = \frac{1}{d}\sum_{k=1}^{d} |k_A\rangle\langle k_A| \otimes |k_B\rangle\langle k_B|$, where $d = \min\{\dim\mathcal{H}_A, \dim\mathcal{H}_B\}$. In this case $H_{A:B} = \log d$ achieves its maximum.

The bipartite case can be directly generalized into the multi-partite case. For instance, we can say that three systems are *tripartite independent*, if the state is *three-factorizable*, i.e. $\varrho_{ABC} = \varrho_A\otimes\varrho_B\otimes\varrho_C$. The function $S(\varrho_{A:B:C}) := S(\varrho_A) + S(\varrho_B) + S(\varrho_C) - S(\varrho_{ABC})$ vanishes, if and only if the state ϱ_{ABC} is *tripartite uncorrelated*. It is a direct consequence of the identity $S(\varrho_{A:B:C}) = S(\varrho_{ABC}||\varrho_A \otimes \varrho_B \otimes \varrho_C)$ and *Klein's inequality*. The first identity follows from the equality

$$\log\varrho_A \otimes \varrho_B \otimes \varrho_C = \log\varrho_A \otimes \mathbb{1}_{BC} + \log\varrho_B \otimes \mathbb{1}_{AC} + \log\varrho_C \otimes \mathbb{1}_{AB} \tag{3.9}$$

that implies $\mathrm{Tr}\varrho_{ABC}\log\varrho_A \otimes \varrho_B \otimes \varrho_C = \mathrm{Tr}\varrho_A\log\varrho_A + \mathrm{Tr}\varrho_B\log\varrho_B + \mathrm{Tr}\varrho_C\log\varrho_C$.

In classical physics we did not satisfactorily answer the question on the meaning of correlations between two subsystems. Pure states of the classical composite system do not exhibit any correlations and the correlations between any two local observations $\mathbf{A}, \mathbf{B}$ are also zero, because $H(p_{\mathbf{AB}}) = H(p_{\mathbf{A}}) = H(p_{\mathbf{B}}) = 0$. If the state p_{AB} is not pure, but uncorrelated, i.e. $p_{AB} = p_A p_B$, then for two observables $\mathbf{A}$ of the system A and $\mathbf{B}$ of the system B we get

$$\begin{array}{rcl} p_{\mathbf{AB}}(a,b) & := & p_{AB}((E(a) \times \Omega_B) \cap (\Omega_A \times F(b))) \\ & = & p_{AB}(E(a) \times F(b)) = p_A(E(a))p_B(F(b)) = p_{\mathbf{A}}(a)p_{\mathbf{B}}(b). \end{array}$$

It implies that every pair of such observables, namely $\mathbf{A} \times \mathbf{I}_B$ and $\mathbf{I}_A \times \mathbf{B}$, are independent, if the state p_{AB} is uncorrelated. Here we denote by $\mathbf{I}$ the trivial random variable $\mathbf{I}(\omega) := 1$ for all $\omega \in \Omega$.

In the quantum case the *local observables* of the subsystems A, B have the form $\mathbf{A} \otimes \mathbb{1}_B, \mathbb{1}_A \otimes \mathbf{B}$, respectively. Consider independent qubits, i.e. $\varrho_{AB} = \varrho_A \otimes \varrho_B$. For the probabilities of outcomes of joint observations we have

$$p_{\mathbf{AB}}(a, b, \varrho_{AB}) := \mathrm{Tr}\left[(\mathbf{E}(a) \otimes \mathbb{1})(\mathbb{1} \otimes \mathbf{F}(b))\varrho_{AB}\right] = \mathrm{Tr}\left[\mathbf{E}(a) \otimes \mathbf{F}(b)\varrho_{AB}\right] \tag{3.10}$$

[1] for proofs see [8]

Note that we have used the definition of the joint outcome in analogy with the classical case, only the intersection of events is replaced by the product of events. For factorized state ϱ_{AB} we get

$$p_{\mathbf{AB}}(a,b,\varrho_A\otimes\varrho_B) = \mathrm{Tr}_A\varrho_A\mathbf{E}(a)\mathrm{Tr}_B\varrho_B\mathbf{F}(b) = p_{\mathbf{A}}(a,\varrho_A)p_{\mathbf{B}}(b,\varrho_B) \tag{3.11}$$

It means that all pairs of local observations are independent. We have obtained the same result as in the classical case. Let us conclude that independence of systems corresponds to the independence of all their local observations.

In what follows we shall use the function

$$C_\varrho(A,B) = S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB}) \tag{3.12}$$

to measure the *mutual correlations* between the systems A and B. Consider a correlated state, i.e. $\varrho_{AB} \neq \varrho_A\otimes\varrho_B$. We have mentioned that the *mutual entropy function* $S(\varrho_{A:B}) = C_\varrho(A,B)$, can be also interpreted as a distance between the joint state ϱ_{AB} and the canonical uncorrelated state $\varrho^c_{AB} = \varrho_A\otimes\varrho_B$. These two descriptions of joint systems are indistinguishable by performing local observations. Of course, if the parties are allowed to communicate and compare their results, then they are able to find differences. For correlated systems there always exists a pair of local observations, for which the joint outcomes are correlated.

3.1.2 Specific quantum correlations

The title of the whole chapter indicates that we shall deal mainly with the composite system of two qubits. Of course, we know the states of this system, its observables and dynamics. Hence, why is the case of two qubits so different in comparison with the case of a single qubit? In this section we shall introduce a relatively new notion in physics - **entanglement**. This concept is crucial in quantum mechanics, because it stands behind many purely quantum phenomena.

The general state of a two-qubit system $\varrho_{AB}\in\mathcal{S}(\mathcal{H}_A\otimes\mathcal{H}_B)$ can be expressed in the form

$$\varrho_{AB} = \frac{1}{4}\mathbb{1} + (\vec{\alpha}.\vec{\sigma})\otimes\mathbb{1}_B + \mathbb{1}_A\otimes(\vec{\beta}.\vec{\sigma}) + \sum_{k,l=1}^{3}\gamma_{kl}\sigma_k\otimes\sigma_l \tag{3.13}$$

where σ_k are the usual sigma matrices. Note again that the operators $\mathbb{1}, \sigma_k\otimes\mathbb{1}_B, \mathbb{1}_A\otimes\sigma_k, \sigma_k\otimes\sigma_l$ form an orthogonal basis of the set of operators. To define a state we need 15 real parameters $\alpha_k, \beta_k, \gamma_{kl}$, where $k,l = 1,2,3$. Of course, not each collection of 15 real numbers determine a quantum state.

The reduced states (describing the subsystems) read

$$\varrho_A = \frac{1}{2}\mathbb{1} + \vec{\alpha}.\vec{\sigma} \qquad\qquad \varrho_B = \frac{1}{2}\mathbb{1} + \vec{\beta}.\vec{\sigma}. \tag{3.14}$$

That is, the parameters α_k, β_k must satisfy the restrictions for qubits, i.e. $|\vec{\alpha}| \leq 1/2$ and $|\vec{\beta}| \leq 1/2$. The positivity conditions on ϱ_{AB} then specify the possible choices of other parameters. Uncorrelated state of two qubits is given by condition $\gamma_{kl} - \alpha_k\beta_l = 0$ for all k,l. The *correlation matrix* Γ' (with matrix elements $\gamma'_{kl} := \gamma_{kl} - \alpha_k\beta_l$) reflects all correlations between the systems A and B. The state can be expressed in the form

$$\varrho_{AB} = \varrho_A\otimes\varrho_B + \sum_{k,l=1}^{3}\gamma'_{kl}\sigma_k\otimes\sigma_l. \tag{3.15}$$

where the correlated and uncorrelated part of the state are separated. The correlations are completely described by matrix Γ'. Next, we will argue that in quantum mechanics the correlations can have a feature, which cannot be observed in the classical case.

In particular, pure states have different properties in classical and quantum case. As we said, in the quantum case we have pure states, for which the correlation matrix does not vanish. It indicates the possibility to find correlations between subsystems, if the whole system is described by a pure state, for example $|\psi\rangle_{AB}$. Such thing could never happen in the classical case, where the purity of the joint state p_{AB} implies that the subsystems are necessarily statistically independent, i.e. $p_{AB} = p_A p_B$. Consider the following pure state of two qubits

$$|\psi\rangle_{AB} = a|0\rangle_A \otimes |0\rangle_B + b|1\rangle_A \otimes |1\rangle_B \tag{3.16}$$

with $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$ This vector state can be expressed in the *operator form*

$$\varrho_{AB} = |\psi\rangle_{AB}\langle\psi| = \varrho_A \otimes \varrho_B + \sum_{kl} \gamma'_{kl}\sigma_k \otimes \sigma_l \tag{3.17}$$

where $\varrho_A = \varrho_B = |a|^2|0\rangle\langle 0| + |b|^2|1\rangle\langle 1|$, i.e. $\vec{\alpha} = \vec{\beta} = (0, 0, (|a|^2 - |b|^2)/2)$, and

$$\gamma'_{kl} := \gamma_{kl} - \alpha_k\beta_l = \frac{1}{4}\mathrm{Tr}[|\psi\rangle_{AB}\langle\psi|\sigma_k \otimes \sigma_l] - \alpha_k\beta_l = \frac{1}{4}\langle\psi|\sigma_k \otimes \sigma_l|\psi\rangle - \alpha_k\beta_l$$

After a little algebra we get the correlation matrix

$$\Gamma' = \frac{1}{4}\begin{pmatrix} ab^* + a^*b & i(ab^* - a^*b) & 0 \\ i(ab^* - a^*b) & -(ab^* + a^*b) & 0 \\ 0 & 0 & 1 - (|a|^2 - |b|^2)^2 \end{pmatrix}. \tag{3.18}$$

As a measure of correlations we will use the mutual entropy function $C(A, B) = S(\varrho_{A:B}) = S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB})$. Hence, the state $|\psi\rangle_{AB}$ is correlated except for the cases when $a = 0$ or $a = 1$. If $S(\varrho_A) = S(\varrho_B) = \log 2$, i.e. for $a = b = \sqrt{1/2}$, the correlations defined by $S(\varrho_{A:B})$ achieve the maximal value $S(\varrho_{A:B}) = 2\log 2$. In this case $\Gamma' = \frac{1}{4}\mathrm{diag}\{1, -1, 1\}$.

This example shows that the value of correlations could be larger then 1, that is the maximum allowed for classical two-bit (mixed) systems. In the classical case the state space contains only four pure states $\{00, 01, 10, 11\}$, that can be represented with quantum symbols $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, but the superposition of these symbols is completely senseless. Consider a mixture

$$\varrho_{AB}(p) = p|00\rangle_{AB}\langle 00| + (1-p)|11\rangle_{AB}\langle 11| \tag{3.19}$$

with $p \in [0, 1]$. The correlation is given by the binary entropy

$$C(A, B) = S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB}) = H_{bin}(p) \tag{3.20}$$

because $S(\varrho_A) = S(\varrho_B) = S(\varrho_{AB}) = H_{bin}(p) = -p\log p - (1-p)\log(1-p)$. The maximum is achieved for the case $p = 1/2$ (*maximally correlated state* ϱ_{max}), when $C(A, B) = 1$. In this case the correlation matrix takes the diagonal form $\Gamma'(\varrho_{max}) = \frac{1}{4}\mathrm{diag}\{0, 0, 1\}$. In quantum two-qubit state space there exist pure states (called *maximally entangled states* with $a = b = 1/\sqrt{2}$), for which the correlations can be two times larger and the correlation matrix reads $\Gamma' = \frac{1}{4}\mathrm{diag}\{1, -1, 1\}$.

However, there is still no evidence for making differences between the correlations in classical and quantum physics. The correlation matrix has the following physical meaning. Its elements are associated with the difference between the mean values

$$\Gamma'_{kl} = \frac{1}{4}\left[\langle\sigma_k \otimes \sigma_l\rangle - \langle\sigma_k\rangle\langle\sigma_l\rangle\right] \tag{3.21}$$

and therefore they represent *standard correlation functions* between the measurement σ_k of the subsystem A and the measurement σ_l of the subsystem B. For classical two-bit system there exists only one (non-trivial) pair of local measurements, that can be represented by the matrix σ_3. It means that all the other coefficients of Γ' (except Γ'_{33}) have no classical sense and must be zero in classical theory. Quantum theory of two qubits is richer. There are infinitely many pairs of measurements, but only nine of them are needed to characterize all correlation properties of quantum state. All the elements of Γ' can take values from $-1/4$ to $1/4$ (correlation coefficients). The extremal values of Γ'_{kl} indicate that the measurements are maximally correlated. The classical states of quantum system, i.e. those of the form $a|00\rangle\langle 00| + b|01\rangle\langle 01| + c|10\rangle\langle 1)| + d|11\rangle\langle 11|$, exhibit correlations only for the pair of local observables $\sigma_z \otimes \mathbb{1}$ and $\mathbb{1} \otimes \sigma_z$. For some specific states of a quantum system more pairs of local measurements can be not only correlated, but even maximally correlated. In fact, one can say, that there are more correlations in quantum than in the classical world and these "extra-correlations" got a new name - *entanglement.*

The existence of the pure state **entanglement** between two systems is related to the presence of correlations. To see the strength of these "quantum" pure state correlations we have to perform more then just a single pair of local measurements. From the mathematical point of view the entanglement arises as a result of the composition of two quantum principles: the superposition and the tensor product structure used in the description of joint quantum systems.

3.2 The entanglement

As we could see, the **entanglement** is a property of states of two systems without any direct relation to a certain choice of local quantum observations, whereas correlations have a sense even in this case. Of course, entanglement has observable effects, but one needs to perform more pairs of different measurements to be able to see them. Therefore, it is possible to use phrase *"entangled couples of observables"*, but the concept of *"entangled pair of observations"* is undefined and senseless. In this section we shall define the entanglement and we will show how it can be quantified.

We have argued that correlated states of quantum systems can be prepared by mixing the factorizable pure states together. Does it mean, that if we mix together entangled pure states, we shall obtain an *entangled mixture*? Consider a mixture of two maximally entangled states $|\psi^{\pm}\rangle_{AB} = (|00\rangle \pm |11\rangle)/\sqrt{2}$ with $p = 1/2$ (*random mixture*)

$$\varrho_{AB} = \frac{1}{2}|\psi^+\rangle\langle\psi^+| + \frac{1}{2}|\psi^-\rangle\langle\psi^-| = \frac{1}{2}\left(|00\rangle\langle 00| + |11\rangle\langle 11|\right) \tag{3.22}$$

In the previous section we constructed the same state (maximally correlated classical state) by mixing two factorizable pure states. The conclusion is, that the *entangled mixtures* could be prepared only by mixing pure entangled states, but not all such mixtures lead to an entangled mixture.

To be more precise, due to the non-existence of superposition between classical pure states, the classical region of quantum states is smaller than the set of all *separable* (equivalently *not entangled*) states of the form

$$\varrho_{AB} = \sum_k p_k \varrho_A^k \otimes \varrho_B^k \tag{3.23}$$

with $p_k \geq 0$ and $\sum_k p_k = 1$. The mentioned absence of classical superposition in classical physics implies that states are always composed like mixtures of pure and mutually orthogonal states. That is, for two bits the most general state is $a|00\rangle\langle 00| + b|01\rangle\langle 01| + c|10\rangle\langle 10| + d|11\rangle\langle 11|$. In what follows, by classical states we will understand the states, for which the spectral decomposition results in the convex combination of (mutually orthogonal) factorized states. The set of separable states is much larger and contains all convex combinations of factorized states, even the convex combinations of mutually non-orthogonal pure factorized states. In some applications the separable non-classical states can have interesting non-classical properties. In principle, the boundary between the entangled and separable states is sharp and it is only a technical problem how to determine that a given state belongs to one of these two sets.

From the mathematical point of view the property of entanglement determines a new structure in the set of joint states. This structure is related to the defined tensor product of the Hilbert spaces representing the fact that the composite system consists of physically distinguishable subsystems. Formally the Hilbert space of two qubits is only a four-dimensional complex Hilbert space. But the structure of the tensor product makes it rich and allows us to investigate the entanglement properties of this space.

3.2.1 The structure of states of two qubits

The problem of the indication of entanglement is very difficult. Suppose you have a state and you want to know whether it is separable, or not. Since the mixture of separable states is again a separable state, the set of all separable states $\mathcal{S}_{sep}$ is **convex**, unlike the set of entangled states $\mathcal{S}_{ent}$. The set of all states $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \equiv \mathcal{S}_{sep} \cup \mathcal{S}_{ent}$ is convex, too and $\mathcal{S}_{sep} \cap \mathcal{S}_{ent} = \emptyset$, because the boundary is sharp.

• PURE STATES •

A general pure state of four dimensional Hilbert space is parametrized by six independent real parameters $a, b, c \in [0, 1], \kappa, \eta, \omega \in [0, 2\pi]$

$$|\psi\rangle = pe^{i\eta}|0\rangle + qe^{i\kappa}|1\rangle + re^{i\omega}|2\rangle + s|3\rangle \tag{3.24}$$

where $s = \sqrt{1 - p^2 + q^2 + r^2}$. Introducing the tensor product we can write $|0\rangle = |0\rangle_A \otimes |0\rangle_B = |00\rangle$, $|1\rangle = |0\rangle_A \otimes |1\rangle_B = |01\rangle$, $|2\rangle = |1\rangle_A \otimes |0\rangle_B = |10\rangle$, $|3\rangle = |1\rangle_A \otimes |1\rangle_B = |11\rangle$. Note, that such relabeling of states does not correspond to any real physical action, but it reflects our knowledge, that the whole system is composite. The correspondence between these two notations is given by the possibility to write numbers in their binary form. Each binary number labels the usual basis of one qubit Hilbert space.

Next, we shall formulate a very useful and important theorem, which is widely used in the theory of entanglement .

Theorem (*Schmidt decomposition*)
Consider a pure state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ of a bipartite system. It can be written in the form

$$|\psi\rangle_{AB} = \sum_{k=1}^{d} \sqrt{\lambda_k}|\phi_k\rangle_A \otimes |\phi_k\rangle_B \tag{3.25}$$

where $\{|\phi_k\rangle_A\}$ and $\{|\phi_k\rangle_B\}$ are orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, and $d = \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}$. Moreover, the coefficients λ_k are positive and $\varrho_A = \varrho_B = \sum_k \lambda_{k=1}^{d}|\phi_k\rangle\langle\phi_k|$.

Proof. Express the general state $|\psi\rangle_{AB}$ in a **product basis** that consists of product vectors $\{|k\rangle_A \otimes |l\rangle_B\}$, where $\{|k\rangle_A\}$ is an orthonormal basis of $\mathcal{H}_A$ and $\{|l\rangle_B\}$ forms an orthonormal basis in $\mathcal{H}_B$. That is,

$$|\psi\rangle_{AB} = \sum_{k=1}^{\dim \mathcal{H}_A} \sum_{l=1}^{\dim \mathcal{H}_B} x_{kl}|k\rangle_A \otimes |l\rangle_B. \tag{3.26}$$

Formally, we are allowed to represent the states of composite systems (expressed in product bases) by suitable rectangular matrix $\mathbf{X}$ with complex elements x_{kl}. The change of a basis in one of the Hilbert spaces, for example $|k\rangle_A := u^A_{kk'}|k'\rangle_A$ in $\mathcal{H}_A$, where $u^A_{kk'}$ forms unitary matrix $\mathbf{U}_A$, corresponds to the transformation $|\psi\rangle_{AB} = \sum_{kl} x_{kl}|k\rangle|l\rangle = \sum_{kl}\sum_{k'} x_{kl}u^A_{kk'}|k'\rangle|l\rangle = \sum_{k'l} x'_{k'l}|k'\rangle|l\rangle$ where the elements $x'_{k'l} = \sum_k u^A_{k'k}x_{kl}$ determine a matrix $\mathbf{X}' = \mathbf{U}_A\mathbf{X}$. Transforming also the basis in the second Hilbert space $\mathcal{H}_B$ using the unitary transformation $\mathbf{U}_B$, we get $\mathbf{X}' = \mathbf{U}_A\mathbf{X}\mathbf{U}_B$ with $x'_{k'l'} = \sum_{kl} u^A_{k'k}x_{kl}u^B_{ll'}$. The *Schmidt theorem* says, that we are able to choose such $\mathbf{U}_A$ and $\mathbf{U}_B$, that the matrix $\mathbf{X}'$ is a diagonal matrix, i.e. $x'_{k'l'} = \sqrt{\lambda_{k'}}\delta_{k'l'}$.

To finish the proof we will use the *singular value decomposition theorem.* Each matrix $\mathbf{X}$ can be written in the so-called **polar form** $\mathbf{X} = \mathbf{U}\mathbf{E} = \mathbf{F}\mathbf{U}$ with unique positive operators $\mathbf{E} = \sqrt{\mathbf{X}^\dagger\mathbf{X}}, \mathbf{F} =$

$\sqrt{\mathbf{X}\mathbf{X}^\dagger}$ and an unitary operator $\mathbf{U}$. Since $\mathbf{E}$ is positive, we can find a unitary transformation $\mathbf{V}$, such that $\mathbf{E} = \mathbf{V}\mathbf{D}\mathbf{V}^\dagger$ and $\mathbf{D}$ is a diagonal matrix. It means, positive operators can be diagonalized and moreover the positivity implies, that the diagonal entries of $\mathbf{D}$ are nonnegative. Thus, we get the needed identity $\mathbf{X} = \mathbf{U}\mathbf{E} = \mathbf{U}\mathbf{V}\mathbf{D}\mathbf{V}^\dagger = \mathbf{W}\mathbf{D}\mathbf{V}^\dagger$. Putting $\mathbf{W} = \mathbf{U}_A$ and $\mathbf{V}^\dagger = \mathbf{U}_B$ we have finished the proof of Schmidt decomposition theorem. $\diamond$

Schmidt theorem implies that each state of two qubits can be written in the form

$$|\psi\rangle_{AB} = a|\phi\rangle_A \otimes |\phi\rangle_B + b|\phi^\perp\rangle_A \otimes |\phi^\perp\rangle_B \tag{3.27}$$

just by performing suitable local unitary transformation $\mathbf{U} = \mathbf{U}_A \otimes \mathbf{U}_B$ and, moreover, a, b are real and positive numbers. We only need to choose appropriate bases in each of the subsystems. Comparing it with eq.(3.16) we find that the correlation matrix is diagonal

$$\Gamma' = \frac{1}{4}\begin{pmatrix} 2ab & 0 & 0 \\ 0 & -2ab & 0 \\ 0 & 0 & 4a^2b^2 \end{pmatrix} \tag{3.28}$$

but the basis, in which this matrix is written has changed. Instead of σ_k we have used the operators $\mathbf{S}_k$ defined as follows $\mathbf{S}_1 := |\phi\rangle\langle\phi^\perp| + |\phi^\perp\rangle\langle\phi|$, $\mathbf{S}_2 := -i(|\phi\rangle\langle\phi^\perp| - |\phi^\perp\rangle\langle\phi|)$ and $\mathbf{S}_3 := |\phi\rangle\langle\phi| - |\phi^\perp\rangle\langle\phi^\perp|$, where $|\phi\rangle, |\phi^\perp\rangle$ represent the corresponding *Schmidt basis.* In conclusion, any pure state of two qubits can be written in the form

$$\varrho_{AB} = \varrho_A \otimes \varrho_B + \frac{ab}{2}(\mathbf{S}_1 \otimes \mathbf{S}_1 - \mathbf{S}_2 \otimes \mathbf{S}_2) + a^2b^2\mathbf{S}_3 \otimes \mathbf{S}_3. \tag{3.29}$$

The entanglement described by Γ' is parametrized and given only by one real parameter a. Hence, any real function of a can be used to measure the entanglement. To reduce the set of possible functions, we need to take into account specific properties of the entanglement. The function must be positive and it can be zero only for $a = 1$, and $a = 0$. The maximal value is achieved by the *maximally entangled states*, i.e. $a = b = \sqrt{1/2}$. Moreover, since the parameter a determines the eigenvalues of the reduced states ϱ_A, ϱ_B, we can define the *entanglement measure* with the help of these reduced states (their eigenvalues). For example,

$$\begin{aligned} E(|\psi\rangle_{AB}) &:= S(\varrho_A) = -a^2 \log a^2 - (1-a^2)\log(1-a^2) && (3.30) \\ E_2(|\psi\rangle_{AB}) &:= \det \varrho_A = a^2(1-a^2) && (3.31) \\ E_3(|\psi\rangle_{AB}) &:= 1 - \mathrm{Tr}\varrho_A^2 = a^2(1-a^2) && (3.32) \end{aligned}$$

where $\varrho_A = \mathrm{Tr}_B|\psi\rangle_{AB}\langle\psi|$. We know that the presence of correlations in pure states is purely a quantum feature. Therefore, we shall use the correlation function $C_\varrho(A, B)$

$$E(\psi_{AB}) := 2S(\varrho_A) = -2\mathrm{Tr}\varrho_A \log \varrho_A \tag{3.33}$$

as a *measure of entanglement* of pure states.

We shall see that in many applications of the information theory the entanglement plays a central role. Predominantly, the maximally entangled states are of the main importance. How many the maximally entangled states of two qubits do exist? Necessarily each of them has the Schmidt decomposition $|\psi\rangle_{AB} = (|\phi\rangle|\phi\rangle + |\phi^\perp\rangle|\phi^\perp\rangle)/\sqrt{2}$. The pure states in between the separable and the maximally entangled states will be called **partially entangled**. Fix one of the maximally entangled states, $|\psi^+\rangle_{AB} = (|00\rangle + |11\rangle)/\sqrt{2}$. Each of the two qubit pure states $|\Phi\rangle_{AB}$ can be obtained from $|\psi^+\rangle_{AB}$ by applying a transformation

$$|\Phi\rangle_{AB} = (\mathbf{A}_\Phi \otimes \mathbb{1}_B)|\psi^+\rangle_{AB} = (\mathbb{1}_A \otimes \mathbf{B}_\Phi)|\psi^+\rangle_{AB} \tag{3.34}$$

where $\mathbf{A}_\Phi, \mathbf{B}_\Phi$ are uniquely determined linear operators, i.e. $\mathbf{A}_\Phi, \mathbf{B}_\Phi \in \mathcal{L}(\mathcal{H})$. Put $\mathbf{A}_\Phi = \sqrt{2}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, then $|\Phi\rangle_{AB} = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$ and $\mathbf{B}_\Phi = \sqrt{2}\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \mathbf{A}_\Phi^T$, where ^{T} denotes the transposition in the basis $|0\rangle, |1\rangle$ determined by Schmidt decomposition of $|\psi^+\rangle_{AB}$. The state $|\Phi\rangle_{AB}$ is maximally entangled if and only if $\mathbf{A}_\Phi$ is unitary. One can find that the following important relation between the Hilbert space scalar product and the Hilbert-Schmidt operator's scalar product holds

$$\langle\Phi|\Psi\rangle = \frac{1}{2}\mathrm{Tr}(\mathbf{A}_\Phi^\dagger \mathbf{A}_\Psi). \tag{3.35}$$

Each orthogonal operator basis $\{\mathbf{S}_k\}$ (satisfying the identity $\mathrm{Tr}\mathbf{S}_k\mathbf{S}_l = 2\delta_{kl}$ for $k = 0, 1, 2, 3$), defines an orthonormal vector basis in $\mathcal{H}_A \otimes \mathcal{H}_B$.

If we put $\mathbf{S}_0 = \mathbb{1}$, then necessary all other $\mathbf{S}_k$ are traceless. Since σ-matrices are traceless and fulfill the orthogonality condition, we get an example of such operator basis. The unitarity ensures that the corresponding states are maximally entangled. Hence, we get an orthonormal basis of $\mathcal{H}_A \otimes \mathcal{H}_B$ composed of maximally entangled states . In particular,

$$|\Phi_0\rangle_{AB} = \mathbb{1} \otimes \mathbb{1}|\psi^+\rangle_{AB} \equiv |\psi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{3.36}$$

$$|\Phi_1\rangle_{AB} = \sigma_1 \otimes \mathbb{1}|\psi^+\rangle_{AB} \equiv |\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \tag{3.37}$$

$$|\Phi_2\rangle_{AB} = \sigma_2 \otimes \mathbb{1}|\psi^+\rangle_{AB} \equiv |\phi^-\rangle_{AB} = \frac{i}{\sqrt{2}}(|01\rangle - |10\rangle) \tag{3.38}$$

$$|\Phi_3\rangle_{AB} = \sigma_3 \otimes \mathbb{1}|\psi^+\rangle_{AB} \equiv |\psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)\ . \tag{3.39}$$

Any traceless operator can be expressed as a complex linear combination of $\sigma-$matrices. That is, $\mathbf{S}_k = \vec{n}_k.\vec{\sigma}$, where, in general, $\vec{n}_k$ is a complex three-dimensional vector. According to the orthogonality condition $\mathrm{Tr}\mathbf{S}_k^\dagger\mathbf{S}_l = 2\delta_{kl}$ we get $\vec{n}_k^*.\vec{n}_k = \delta_{kl}$ [cf. Eq.(1.39)]. Any collection of operators $\mathbf{S}_0 = \mathbb{1}, \mathbf{S}_k = \vec{n}_k.\vec{\sigma}$ with $k = 1, 2, 3$ and n_k mutually orthonormal three-dimensional complex vectors, determines an orthonormal basis in $\mathcal{H}_A \otimes \mathcal{H}_B$. Moreover, if the vectors $\vec{n}_k$ are real, then the corresponding basis consists of maximally entangled states, because $\mathbf{S}_k = \vec{n}_k.\vec{\sigma}$ are unitary.

• MIXED STATES •

We have already introduced a general state of two qubits, which is parametrized by 15 real numbers. For pure states the question of entanglement and separability is in a sense trivial, because we do not have any problem with classical correlations. The main problem of the quantification of entanglement is how to separate the classical correlations from the quantum correlations. We need to find such properties of states that are typical only of separable, or only of entangled ones.

For example, we know that if ϱ_{AB} is separable, then

1. $S(\varrho_A) \leq S(\varrho_{AB})$ and $S(\varrho_B) \leq S(\varrho_{AB})$
2. $\varrho_{AB}^{T_B}$, or $\varrho_{AB}^{T_A}$ is positive operator

The first necessary criterion is not a very strong one, because there are also entangled states for which these relations hold. The second criterion was introduced by *A.Perez* [18]. Let us remind that partial transposition depends on the basis, but the eigenvalues of a partially transposed operator do not. It was shown by *Horodeccy* [20] that for two qubits this criterion is also sufficient. It means that the partial transposition of an entangled state is always a negative operator, and *vice versa*.

Can we somehow quantify the entanglement between two qubits? We have shown that for pure states we can use the entropy of a reduced state. Unfortunately, this simple quantity cannot be used for mixed states, because even the factorized states can have nonvanishing entropy of the subsystems. So what to do? One approach is purely *mathematical*, where one tries to define functions that somehow diminished the part of the state, which is classically correlated. For example, we can measure the smallest "distance" between the state and the set of all separable states

$$E_{dist}(\varrho_{AB}) = \mathcal{D}(\varrho, \mathcal{S}_{sep}) := \min_{\xi_{AB}\in\mathcal{S}_{sep}} \mathcal{D}(\varrho_{AB}, \xi_{AB}) \tag{3.40}$$

where $\mathcal{D}$ may not be a distance in a strict mathematical sense. In fact, it need not be symmetric in the usual sense, but $D(\varrho_{AB}, \mathcal{S}_{sep}) = 0$ if and only if ϱ_{AB} is separable.

Another approach is *operational.* That is, we use a physical process, for which we think the entanglement is crucial. Unfortunately, till today nobody has found a process that strictly discriminates between classical and quantum correlations. The *measures of entanglement* E defined[2] must satisfy some properties that are typical of entanglement. We shall formulate them at the end of this chapter after introducing the set of all possible physical manipulations with quantum systems.

To quantify the entanglement we can exploit also the measures $E(\psi_{AB})$ defined for pure states, where the degree of entanglement can be better understood. It is reasonable to require that the measure for general states coincides with the known measure for pure states. In the two qubit case it means $E(\psi_{AB}) = S(\varrho_A)$, or $E(\psi_{AB}) = \det \varrho_A$. For mixed states, ϱ_{AB}, the entanglement measure can be defined also in the following way

$$E(\varrho_{AB}) = \min_{\varrho_{AB}=\sum_k p_k|\psi_k\rangle_{AB}\langle\psi_k|} \sum_k p_k E(\psi_k) \tag{3.41}$$

where we minimalize over all possible decompositions of the state ϱ_{AB} into convex sums of pure states. We shall call this measure the **entanglement of formation**. From the definition it is obvious that $E(\varrho_{AB})$ is zero if and only if the state ϱ_{AB} is separable.

Concurrence

For two-qubit system the problem of calculating the entanglement of formation has been solved by *W. Wootters* [23]. He introduced the function called **concurrence** C, which determines the entanglement of formation in the following sense

$$E = H_{bin}(\frac{1-\sqrt{1-C^2}}{2}) \tag{3.42}$$

To obtain the concurrence of the state ϱ_{AB} we must perform the following steps:

1. *Time reversal*

$$\varrho_{AB} \to R = \varrho\sigma_y \otimes \sigma_y \varrho^*_{AB}\sigma_y \otimes \sigma_y, \text{ where } (\varrho^*_{AB})_{kl} := \overline{(\varrho_{AB})_{kl}} \tag{3.43}$$

2. *Calculating the eigenvalues of* R: $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \sqrt{\lambda_3} \geq \sqrt{\lambda_4}$
3. *Calculating the concurrence* $C = \max\{0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}\}$

Example
Consider a state

$$\varrho = \begin{pmatrix} a & d & e & 0 \\ d^* & b & f & 0 \\ e^* & f^* & c & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.44}$$

The matrix R takes the form

$$R = \begin{pmatrix} 0 & ef^* + dc & eb + df & -2ed \\ 0 & ff^* + bc, & 2fb & -(eb+df) \\ 0 & 2cf^* & ff^* + bc & -(ef^* + dc) \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.45}$$

and it has only two non-vanishing eigenvalues

$$\lambda_\pm = ff^* + bc \pm 2\sqrt{ff^*bc} \tag{3.46}$$

[2] see for example article of V.Vedral and M.B.Plenio [14]

and consequently the *tangle* (=square of the concurrence) equals

$$\tau(\varrho) = [C(\varrho)]^2 = 2(ff^* + bc) - 2|ff^* - bc| = \min\{4bc, 4ff^*\} \tag{3.47}$$

To find the minimum let us assume that ϱ is written as the convex sum of unnormalized vectors $|\phi_j\rangle$, i.e. $\varrho = \sum_j |\phi_j\rangle\langle\phi_j|$. Then the matrix elements $|01\rangle\langle 01|, |10\rangle\langle 10|, |01\rangle\langle 10|, |10\rangle\langle 01|$ of ϱ will take the form $b = \sum_j b_j^* b_j$, $c = \sum_j c_j^* c_j$ and $f = \sum_j b_j^* c_j$, respectively. We use the definitions $b_j = \langle\phi_j|01\rangle$ and $c_j = \langle\phi_j|10\rangle$. Our task is to compare the following two terms

$$\begin{aligned} bc &= (\sum_j b_j^* b_j)(\sum_k c_k^* c_k) = |\vec{b}|^2|\vec{c}|^2 & (3.48)\\ ff^* &= (\sum_j b_j^* c_j)(\sum_j c_k^* b_k) = |\vec{b}.\vec{c}|^2 & (3.49)\end{aligned}$$

where we introduced vectors $\vec{b} = (b_1, \ldots)$ and $\vec{c} = (c_1, \ldots)$. The *Schwartz inequality* asserts $|\vec{b}|^2|\vec{c}|^2 \geq |\vec{b}.\vec{c}|^2$ and therefore we have proved the relation

$$ff^* \leq bc \tag{3.50}$$

It implies that the tangle (= square of the concurrence) of ϱ equals

$$\tau(\varrho) = C^2 = 4ff^* = 4|\langle 01|\varrho|10\rangle|^2 \tag{3.51}$$

3.3 Entanglement and positive maps

Let us denote the set of linear mappings $\Lambda : \mathcal{T}_A \to \mathcal{T}_B$ by $\mathcal{L}(\mathcal{T}_A, \mathcal{T}_B)$, where $\mathcal{T}_A, \mathcal{T}_B$ denote the sets of *Hilbert-Schmidt operators.* The element $\Lambda \in \mathcal{L}(\mathcal{T}_A, \mathcal{T}_B)$ is *positive*, if it maps positive operators of $\mathcal{T}_A$ into the positive operators of $\mathcal{T}_B$, i.e. $\mathbf{A} \geq 0 \Rightarrow \Lambda[\mathbf{A}] \geq 0$. The map Λ is called **completely positive**, if the induced map $\Lambda_n := \Lambda \otimes \mathcal{I}_n$ is positive for all $n \in \mathbb{N}$, where $\mathcal{I}_n$ is the identity map on $\mathcal{T}_C$ with $\dim \mathcal{H}_C = n$, i.e. $\mathcal{I}_n(\mathbf{C}) = \mathbf{C}$ for all $\mathbf{C} \in \mathcal{L}(\mathcal{H}_C) \equiv \mathcal{T}_C$. In what follows, we shall put $B = A$ and for convenience we shall use the notation $\mathcal{L}(\mathcal{T}_A)$ instead of $\mathcal{L}(\mathcal{T}_A, \mathcal{T}_A)$. Let $\mathcal{L}_+(\mathcal{T}_A) \subset \mathcal{L}(\mathcal{T}_A)$ be the set of all *positive maps* and $\mathcal{L}_{cp}(\mathcal{T}_A) \subset \mathcal{L}_+(\mathcal{T}_A)$ be the set of all *completely positive maps* defined on $\mathcal{L}(\mathcal{H}_A)$.

Apply a positive map $\Lambda_A \in \mathcal{L}_+(\mathcal{T}_A)$ on a separable state $\varrho_{AB} = \sum_k p_k \varrho_A^k \otimes \varrho_B^k$

$$\varrho'_{AB} = (\Lambda_A \otimes \mathcal{I}_B)[\varrho_{AB}] = \sum_k p_k(\Lambda_A[\varrho_A^k]) \otimes \varrho_B^k \tag{3.52}$$

Obviously, the obtained operator ϱ'_{AB} corresponds to a separable quantum state, $\varrho'_{AB} \in \mathcal{S}_{sep}$. The question is, whether the output operator ϱ'_{AB} corresponds to a positive operator also for initially entangled states ϱ_{AB}. If the map Λ_A is completely positive, then this operator is positive. However, for positive and non CP maps the situation is different. Therefore, we can use them to probe the entanglement. We shall see that for all positive non CP maps Λ_A the separable states remain positive, but some of the entangled states can be transformed into negative operators.

Consider the partial transposition T_A. Without any doubts this transformation is linear and positive. Transforming the maximally entangled state $|\psi^+\rangle$ by the map $T_A \otimes \mathcal{I}_B$ we get the state

$$\varrho_{AB}^{T_A} = (T_A \otimes \mathcal{I}_B)[|\psi^+\rangle\langle\psi^+|] = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.53}$$

with eigenvalues $\lambda = \{\pm 1/2\}$. It means that the operator $\varrho_{AB}^{T_A}$ is not positive and, consequently, the *partial transposition* indicates that the state $|\psi^+\rangle_{AB}$ is entangled. The *partial transposition* is an example of the positive non CP map. The relation between such maps and entanglement has been studied (see for example [20], [11], [26] and references therein) in order to detect the entangled states. Based on the result of *Woronowicz* in [19] the partial transposition can be used to completely identify the entanglement in two-qubit systems. In a sense the partial transposition is the "only" positive non CP map in this case. Unfortunately, this result does not hold for general systems [27], where the positivity of the partially transposed state does not imply that the state is separable. Let us denote the set of all positive non CP maps by $\mathcal{V}(\mathcal{T}_A) = \mathcal{L}_+(\mathcal{T}_A) \setminus \mathcal{L}_{cp}(\mathcal{T}_A)$.

Consider the set of *positive operators* $\mathbf{A} \in \mathcal{L}_{pp}(\mathcal{H}_A \otimes \mathcal{H}_B)$ *on pure product states* $\varrho_{AB} = \mathbf{P}_{\psi_A} \otimes \mathbf{Q}_{\phi_B}$, i.e. $\mathrm{Tr}(\mathbf{A}\mathbf{P}_{\psi_A} \otimes \mathbf{Q}_{\phi_B}) \geq 0$. The following set relations holds $\mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_B) \subset \mathcal{L}_{pp}(\mathcal{H}_A \otimes \mathcal{H}_B) \subset \mathcal{L}_s(\mathcal{H}_A \otimes \mathcal{H}_B)$. We have mentioned that the set of separable states $\mathcal{S}_{sep}$ is convex. The extremal points are pure factorized (= product) states It follows that the operators $\mathbf{A} \in \mathcal{L}_{pp}(\mathcal{H}_A \otimes \mathcal{H}_B)$ are positive on all separable states $\varrho_{AB} \in \mathcal{S}_{sep}$, i.e.

$$\mathrm{Tr}(\varrho_{AB}\mathbf{A}) \geq 0, \tag{3.54}$$

This characterization of separable states via elements of $\mathcal{L}_{pp}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is similar to the one with positive linear maps $\mathcal{L}_+(\mathcal{T}_A)$. Let us define the set of positive non PP operators $\mathcal{W} = \mathcal{L}_{pp}(\mathcal{H}_A \otimes \mathcal{H}_B) \setminus \mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_B)$. These operators $\mathbf{A} \in \mathcal{W}$ are called *entanglement witnesses*, because their negative mean values ($\mathrm{Tr}\varrho\mathbf{A} < 0$) indicate the entanglement. In what follows we shall show a one-to-one correpsondence between the elements of $\mathcal{V}(\mathcal{T}_A, \mathcal{T}_B)$ and $\mathcal{W}_{\mathcal{H}_A \otimes \mathcal{H}_B}$.

Take $\mathbf{A} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and define a map $\Lambda \in \mathcal{L}(\mathcal{T}_A, \mathcal{T}_B)$ by the formula

$$\Lambda[\varrho_A] := \sum_{a,a'}\sum_{b,b'} |b\rangle\langle b'|\Lambda_{bb',aa'}\varrho_{aa'} := \sum_{a,a',b,b'} \varrho_{aa'}(\mathbf{A})_{ab,a'b'}|b\rangle\langle b'| \tag{3.55}$$

where we have used the notation

$$\varrho_A := \sum_{aa'} \varrho_{aa'}|a\rangle\langle a'| \tag{3.56}$$

$$\Lambda = \sum_{a,a',b,b'} |b\rangle\langle b'|\Lambda_{bb',aa'}|a\rangle\langle a'| \tag{3.57}$$

with $\Lambda_{bb',aa'} = \mathrm{Tr}(|b\rangle\langle b'|\Lambda[|a\rangle\langle a'|]) := (\mathbf{A})_{ab,a'b'} = \langle ab|\mathbf{A}|a'b'\rangle$. This map defines the one-to-one correspondence $\mathcal{J} : \mathcal{L}(\mathcal{T}_A, \mathcal{T}_B) \to \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ between the operators and linear mappings. Note that

$$\begin{aligned}\langle ab|(\Lambda_A \otimes \mathcal{I}_B)[|\psi^+\rangle\langle\psi^+|]|a'b'\rangle &= \frac{1}{d}\sum_k \langle a|\Lambda_A[|k\rangle\langle k|]|a'\rangle\langle b|k\rangle\langle k|b'\rangle \\ &= \frac{1}{d}\langle a|\Lambda_A[|b\rangle\langle b'|]|a'\rangle = \frac{1}{d}(\mathbf{A})_{ab,a'b'}\end{aligned} \tag{3.58}$$

where $|\psi^+\rangle_{AB} = \sum_k |kk\rangle\langle kk|$ and $d = \dim \mathcal{H}_A$. As a result we get the following expression for the correspondence $\mathcal{J}$

$$\frac{1}{d}\mathbf{A} = (\Lambda_A \otimes \mathcal{I}_B)[|\psi^+\rangle\langle\psi^+|]. \tag{3.59}$$

Let us verify the positivity of the operator $\varrho_B = \Lambda_A[\varrho_A]$ providing that $\varrho_A \geq 0$. In particular, for all $|\phi\rangle \in \mathcal{H}_B$ and $\varrho_A \in \mathcal{L}_+(\mathcal{H}_A)$ we have

$$\begin{aligned}\langle\phi|\varrho_B|\phi\rangle &= \sum_{b,b'} \phi_b^*\varrho_{bb'}\phi_{b'} = \sum_{a,a',b,b'} \phi_b^*\Lambda_{bb',aa'}\varrho_{aa'}\phi_{b'} \\ &= \sum_{a,a',b,b'} \phi_b^*\phi_{b'}(\mathbf{A})_{ab,a'b'}\varrho_{aa'} \\ &= \mathrm{Tr}\mathbf{A}(\varrho_A \otimes |\phi\rangle\langle\phi|)\end{aligned} \tag{3.60}$$

It is easy to see that the linear map Λ is positive, if and only if the operator $\mathbf{A}$ is positive on pure product states. The restriction of the map $\mathcal{J}$ onto the set of positive maps transforms the set of positive maps onto the set of positive operators on pure product states, i.e. $\mathcal{J} : \mathcal{L}_+(\mathcal{T}_A, \mathcal{T}_B) \to \mathcal{L}_{pp}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is again a bijection.

The complete positivity of $\Lambda \in \mathcal{L}_{cp}(\mathcal{T}_A, \mathcal{T}_B)$ implies that for all $|\psi\rangle_{BC} \in \mathcal{H}_B \otimes \mathcal{H}_C$ and $\varrho_{AC} \in \mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_C)$ the following inequality must hold

$$0 \leq \ {}_{BC}\langle\psi|(\Lambda_A \otimes \mathcal{I}_C)[\varrho_{AC}]|\psi\rangle_{BC} = \mathrm{Tr}\left[(\mathbf{A} \otimes \mathbf{P}_+)(\varrho_{AC_1} \otimes |\psi\rangle_{BC_2}\langle\psi|)\right] \tag{3.61}$$

where the last equality follows from the Eq.(3.60). The Eq.(3.59) determines the operator $\mathbf{A} \otimes \mathbf{P}_+$ (with $\mathbf{P}_+ := |\psi^+\rangle_{C_1C_2}\langle\psi^+|$) as an image of the map $\Lambda_A \otimes \mathcal{I}_C : \mathcal{T}_{AC_1} \to \mathcal{T}_{BC_2}$ ($\mathcal{I}_C : \mathcal{T}_{C_1} \to \mathcal{T}_{C_2}$) Consequently, the operator $\mathbf{A} \otimes \mathbf{P}_+$ must be positive on pure product states, i.e. $\mathbf{A} \otimes \mathbf{P}_+ \in \mathcal{L}_{pp}(\mathcal{H}_{AC_1} \otimes \mathcal{H}_{BC_2})$. Define the operator square root $\sqrt{\mathbf{A}}$ by the condition $\sqrt{\mathbf{A}}\sqrt{\mathbf{A}} = \mathbf{A}$. Since $\mathbf{P}_+ = \mathbf{P}_+^2$ is a projector, the following identity holds $\mathbf{A} \otimes \mathbf{P}_+ = (\sqrt{\mathbf{A}} \otimes \mathbf{P}_+)(\sqrt{\mathbf{A}} \otimes \mathbf{P}_+)$. Let us rewrite $\mathrm{Tr}[(\mathbf{A} \otimes \mathbf{P}_+)|\psi\rangle_{ABC_1C_2}\langle\psi|] = \langle\psi|\mathbf{A} \otimes \mathbf{P}_+|\psi\rangle$ and put $|\psi\rangle_{ABC_1C_2} = (\sum_{a,c}\psi_{ac}|ac\rangle_{AC_1}) \otimes (\sum_{b,d}\psi'_{bd}|bd\rangle_{BC_2})$. Applying the transformation $\sqrt{\mathbf{A}} \otimes \mathbf{P}_+$ we obtain the product state

$$\begin{aligned}\sqrt{\mathbf{A}} \otimes \mathbf{P}_+|\psi\rangle_{ABC_1C_2} &= \sum_{a,b,c} \psi_{ac}\psi'_{bc}\sqrt{\mathbf{A}}|ab\rangle_{AB} \otimes |\psi^+\rangle_{C_1C_2} \\ &= (\sqrt{\mathbf{A}}|\chi\rangle_{AB}) \otimes |\psi^+\rangle_{C_1C_2}\end{aligned} \tag{3.62}$$

where $|\psi^+\rangle_{C_1C_2} = \frac{1}{\sqrt{2}}\sum_k |kk\rangle_{C_1C_2}$ and $|\chi\rangle_{AB}$ can be an entangled state of the systems A and B. Consequently, we obtain that the condition of the complete positivity (Eq.(3.61)) takes the form

$$0 \leq \ {}_{BC}\langle\psi|(\Lambda_A \otimes \mathcal{I}_C)[\varrho_{AC}]|\psi\rangle_{BC} = {}_{AB}\langle\chi|\mathbf{A}|\chi\rangle_{AB} \tag{3.63}$$

It means that the operator $\mathbf{A}$ must be positive, i.e. $\mathbf{A} \in \mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_B)$. As a result we get that $\mathcal{J} : \mathcal{L}_{cp}(\mathcal{T}_A, \mathcal{T}_B) \to \mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a bijection.

We have shown the one-to-one correspondence between the sets $\mathcal{L}_+(\mathcal{T}_A, \mathcal{T}_B), \mathcal{L}_{cp}(\mathcal{T}_A, \mathcal{T}_B)$ and $\mathcal{L}_{pp}(\mathcal{H}_A \otimes \mathcal{H}_B), \mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_B)$, respectively. Consequently, also the following one-to-one correspondence holds $\mathcal{J} : \mathcal{V}(\mathcal{T}_A, \mathcal{T}_B) \to \mathcal{W}(\mathcal{H}_A \otimes \mathcal{H}_B)$, i.e. the positive non CP maps can be used to characterize the entanglement in the same way like *entanglement witnesses*. The separability of the state ϱ_{AB} is equivalent to the positivity of the operator $(\Lambda_A \otimes \mathcal{I}_B)[\varrho_{AB}]$ for all linear maps $\Lambda_A \in \mathcal{V}(\mathcal{T}_A, \mathcal{T}_B)$. We said, that for two qubits the *partial transposition* is a generic form of all elements of $\mathcal{V}(\mathcal{H}_A, \mathcal{H}_B)$. Therefore, the separability of two-qubit state is equivalent to the positivity of partial transposition.

3.4 Quantum operations

In this section we shall review all the possible manipulations we are able to perform on quantum systems. In particular, any quantum operation results in the state transformation of the system. That is, mathematically the operations are described by mappings $\mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$. Any operation can be divided into the following four *elementary operations:*

1.Measurements

Measuring the observable $\mathbf{A} = \sum_a a\mathbf{E}(a) \in \mathcal{L}_s(\mathcal{H})$ we obtain the following map $\mathcal{M}_\mathbf{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$

$$\varrho \mapsto \varrho' = \mathcal{M}_\mathbf{A}[\varrho] = \sum_a (\mathrm{Tr}\mathbf{E}(a)\varrho)\mathbf{E}(a) \ . \tag{3.64}$$

2.Unitary transformations

$$\varrho \mapsto \varrho' = \mathcal{U}[\varrho] = \mathbf{U}\varrho\mathbf{U}^\dagger \ . \tag{3.65}$$

3.Addition of uncorrelated ancilla

Consider the following situation. We have a quantum system ϱ we want to manipulate. Consider we have also some other quantum systems called *ancilla* prepared in the state ξ_A. We can put these two systems together and apply measurements, or unitary transformation on both of them. That is, we realize a mapping $\mathbf{A}_\xi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H} \otimes \mathcal{H}_A)$

$$\varrho \mapsto \varrho' = \Lambda_\xi[\varrho] := \varrho \otimes \xi_A \ . \tag{3.66}$$

4.Tracing out the subsystem
Consider the reverse situation. We have a composed system $A + B$ described by the joint state $\varrho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The system (of our interest) A is described by the reduced state ϱ_A. That means, we can define a map $\Lambda_B : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_A)$ by the relation

$$\varrho_{AB} \mapsto \varrho_A = \Lambda_B[\varrho_{AB}] := \mathrm{Tr}_B \varrho_{AB} \tag{3.67}$$

All these quantum operations together form a general quantum operation, which can be realized on quantum systems.

It is easy to see that all these maps $\mathcal{M}_\mathbf{A}, \mathcal{U}, \Lambda_\xi, \Lambda_B$ are completely positive and tracepreserving. The question is, whether each tracepreserving completely positive map $\Lambda : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$ can be written as a sequence of these elementary quantum operations, or not.

In the previous section we have shown the one-to-one correspondence between the set of completely positive maps $\Lambda_\mathbf{A} \in \mathcal{L}_{cp}(\mathcal{T}_A, \mathcal{T}_B)$ and positive operators $\mathbf{A} \in \mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_B)$ given by the relation $(\Lambda_\mathbf{A} \otimes \mathcal{I}_A)[\mathbf{P}_+] = \mathbf{A}$. Consider $\mathbf{A} = \mathbf{P}_\psi$, where $|\psi\rangle_{AB} = \sum_{a,b} \psi_{ab} |a\rangle \otimes |b\rangle$ is a vector in $\mathcal{H}_{AB}$. Then the induced completely positive map Λ_ψ reads

$$\begin{aligned} (\varrho'_B)_{bb'} &= \langle b|\Lambda_\psi[\varrho_A]|b'\rangle = \sum_{a,a'} \varrho_{aa'} \psi_{ab} \psi^*_{a'b'} \\ &= \sum_{a,a'} (\mathbf{M})_{ba} \varrho_{aa'} (\mathbf{M}^\dagger)_{a'b'} = (\mathbf{M} \varrho_A \mathbf{M}^\dagger)_{bb'} \end{aligned} \tag{3.68}$$

where $\varrho_A = \sum_{a,a'} \varrho_{aa'} |a\rangle_A \langle a'| \in \mathcal{L}(\mathcal{H}_A)$ and $\mathbf{M} = \sum_{a,b} \psi_{ab} |b\rangle\langle a| : \mathcal{H}_A \to \mathcal{H}_B$. As a result we get that the linear map Λ_ψ corresponding to each vector $|\psi\rangle_{AB}$ takes the form

$$\Lambda_\psi[\varrho_A] = \mathbf{M} \varrho_A \mathbf{M}^\dagger. \tag{3.69}$$

Because the set of positive operators $\mathcal{L}_+(\mathcal{H}_A \otimes \mathcal{H}_B)$ is *convex*, each positive element $\mathbf{A}$ can be expressed as a convex sum of extreme points (i.e. operators corresponding to pure states $\mathbf{P}_\psi$). Consider $\mathbf{A} = \sum_k p_k \mathbf{P}_{\psi_k}$. Then the corresponding completely positive map $\Lambda_\mathbf{A}$ is given by formula

$$\Lambda_\mathbf{A}[\varrho_A] = \sum_k \mathbf{M}_k \varrho_A \mathbf{M}_k^\dagger \tag{3.70}$$

where $\mathbf{M}_k := \sum_{a,b} \sqrt{p_k} \psi^k_{ab} |b\rangle\langle a|$ are linear bounded operators $\mathcal{H}_A \to \mathcal{H}_B$. Thus, we have obtained the most general form of completely positive maps. The tracepreservity of $\Lambda_\mathbf{A}$, i.e. $\mathrm{Tr}\Lambda_\mathbf{A}[\varrho] = \mathrm{Tr}\varrho$, implies that $\sum_k \mathbf{M}_k^\dagger \mathbf{M}_k = \mathbb{1}_A$. Let us denote the set of all tracepreserving completely positive maps $\Lambda : \mathcal{T}_A \to \mathcal{T}_B$ by $\mathcal{L}_{tcp}(\mathcal{T}_A, \mathcal{T}_B)$.

In what follows we will see that every tracepreserving completely positive map $\Lambda \in \mathcal{L}_{tcp}(\mathcal{T}_A, \mathcal{T}_B)$ can be composed from the elementary quantum operations. Consider the map $\Lambda_\mathbf{A}$ given by the collection of operators $\mathbf{M}_k : \mathcal{H}_A \to \mathcal{H}_B$ satisfying $\sum_k \mathbf{M}_k^\dagger \mathbf{M}_k = \mathbb{1}_A$. Introduce the third Hilbert space $\mathcal{H}_C$ with $\dim \mathcal{H}_B = \dim \mathcal{H}_C$. Let $|0\rangle_{BC} \in \mathcal{H}_B \otimes \mathcal{H}_C$ be a pure state. It means we add an uncorrelated ancilla BC to obtain a state $\varrho_A \otimes |0\rangle_{BC}\langle 0|$. Define the linear operator $\mathbf{U} : \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \to \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ by the formula

$$\mathbf{U}|a\rangle_A \otimes |0\rangle_{BC} = \sum_k (\mathbf{M}_k |a\rangle_A) \otimes |k\rangle_{AC} \tag{3.71}$$

where $|a\rangle_A$ is a basis in $\mathcal{H}_A$ and $|k\rangle_{AC}$ is an orthonormal basis in $\mathcal{H}_A \otimes \mathcal{H}_C$. The transformation $\mathbf{U}$ restricted onto the subspace spanned by vectors $\{|a\rangle_A \otimes |0\rangle_{BC}\}_a$ preserves the scalar product, because

$$\begin{aligned}
{}_{ABC}\langle a'0|\mathbf{U}^\dagger\mathbf{U}|a0\rangle_{ABC} &= \sum_{k,k'} {}_{AC}\langle k'|k\rangle_{AC}\ {}_{B}\langle a'|\mathbf{M}_k^\dagger\mathbf{M}_k|a\rangle_A \\
&= \sum_{k,k'} \delta_{kk'} \langle a'|\mathbf{M}_k^\dagger\mathbf{M}_k|a\rangle_A =_A \langle a'|(\sum_k \mathbf{M}_k^\dagger\mathbf{M}_k)|a\rangle_A \\
&= \delta_{aa'} =_{ABC} \langle a'0|a0\rangle_{ABC}
\end{aligned} \tag{3.72}$$

It means the collection of mutually orthonormal vectors is transformed onto another collection of mutually orthonormal vectors. The action of the operator $\mathbf{U}$ can be extended to the whole Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ in a unitary way, i.e. it can preserve the scalar product for all states. Applying this unitary transformation onto the state $\varrho_A \otimes |0\rangle_{BC}\langle 0|$ we get the state

$$\varrho_A \otimes |0\rangle_{BC}\langle 0| \mapsto \sum_{k,k'} \mathbf{M}_{k'}^\dagger \varrho_A \mathbf{M}_k \otimes |k'\rangle_{AC}\langle k| \tag{3.73}$$

Tracing over the systems A and C we obtain the desired map $\Lambda : \mathcal{T}_A \to \mathcal{T}_B$

$$\varrho_A \mapsto \sum_{k,k'} \mathbf{M}_{k'}^\dagger \varrho_A \mathbf{M}_k \mathrm{Tr}_{AC} |k'\rangle\langle k| = \sum_k \mathbf{M}_k^\dagger \varrho_A \mathbf{M}_k = \Lambda_{\mathbf{A}}[\varrho_A] \tag{3.74}$$

In conclusion, we have obtained that any completely tracepreserving map Λ can be carried out like a composition of three elementary maps:

$$\Lambda \equiv \Lambda_{AC} \circ \mathcal{U}_{ABC} \circ \mathcal{A}_{|0\rangle_{BC}} \tag{3.75}$$

3.4.1 Local manipulations

In the context of *entanglement* the local manipulations are of importance . Let us suppose two spatially separated subsystems A and B. The local operations are maps of the form $\Lambda \otimes \mathcal{I}_B$, or $\mathcal{I}_A \otimes \Lambda$. It is important that these classes of operations commute. It means that Bob (Alice) does not know about the existence of the second party. And, of course, they cannot say anything about their mutual correlations, or about the shared amount of entanglement.

One must allow the parties to communicate to enable them to find the possible correlations and entanglement. We have mentioned that in a sense the communication can help to "simulate" the *local joint* observations and they are able to reconstruct their joint state. The question is, whether the possibility of cooperation via the exchange of classical information could create, or destroy the existing correlations and entanglement.

To create the entanglement we need to perform an operation that superpose the product states. Since the local operations are of tensor product form, they are not able to create the entanglement. It is straightforward that the classical communication does not change this property. Hence, the local quantum operations powered by classical communication (*LQCC*) cannot increase the shared entanglement. Using such operations we can only concentrate the entanglement. Consider we have a number N of copies of the state ϱ_{AB}, i.e. Alice and Bob share N pairs of equally prepared qubits. That is, the whole system consists of $2N$ qubits and the corresponding state is $\varrho_{AB} = \bigotimes_{j=1}^N \varrho_{A_j B_j} = \varrho_{AB}^{\otimes N}$. By the *concentration of entanglement* we mean the possibility to transform this state into k EPR pairs ($k < N$) by LQCC, i.e. $\Lambda : \varrho_{AB}^{\otimes N} \mapsto \mathbf{P}_+^{\otimes k}$, where $\mathbf{P}_+$ is maximally entangled state.

The case of classical correlations is different. On one side we cannot create correlations by local operations, but if we have an opportunity to communicate, then the correlations can be established. Consider two systems in the uncorrelated state $\varrho_{AB} = |00\rangle\langle 00|$. Creation of the correlations means the possibility of mixing. We can use the sum of tensor product operations to do it. Let Alice perform a measurement in $|\pm x\rangle_A$ basis. Then she transmits the outcome of her measurement to Bob and

Bob performs unitary transformation $\mathbf{U}_{\pm}$. It corresponds to a linear map $\Lambda[\varrho_{AB}] = p_{+}|+\rangle\langle+| \otimes \mathbf{U}_{+}|0\rangle\langle 0|\mathbf{U}_{+}^{\dagger} + p_{-}|-\rangle\langle-| \otimes \mathbf{U}_{-}|0\rangle\langle 0|\mathbf{U}_{-}^{\dagger}$, where $p_{\pm} = \mathrm{Tr}\varrho_A \mathbf{P}_{\pm x}$ and $\mathbf{U}_{\pm}|0\rangle_B = |\pm\rangle_B$. For the pure state $|00\rangle$ we get the state $\varrho'_{AB} = \frac{1}{2}(|++\rangle\langle++| + |--\rangle\langle--|)$, that is, the maximally correlated state.

3.5 Summary

3.5.1 State space

general form		$\varrho_{AB} = \varrho_A \otimes \varrho_B + \sum_{kl} \gamma'_{kl} \sigma_k \otimes \sigma_l$
factorized (product) states	$\gamma'_{kl} = 0$	$\varrho_{AB} = \varrho_A \otimes \varrho_B$
correlated states	$\gamma'_{kl} \neq 0$	$\varrho_{AB} \neq \varrho_A \otimes \varrho_B$
classical states $\mathcal{S}_{cl}$		ϱ_{AB} has factorized eigenstates
separable states $\mathcal{S}_{sep}$		$\varrho_{AB} = \sum_k p_k \varrho_A^k \otimes \varrho_B^k$
entangled states $\mathcal{S}_{ent}$		$\varrho_{AB} \neq \sum_k p_k \varrho_A^k \otimes \varrho_B^k$

3.5.2 Entanglement properties

1. *Equivalence:* $E(\mathbf{U}_A \otimes \mathbf{U}_B \varrho_{AB} \mathbf{U}_A^{\dagger} \otimes \mathbf{U}_B^{\dagger}) = E(\varrho_{AB})$
2. *LQCC:* Entanglement between two parties **cannot increase**, if the parties are allowed to perform local operations and communicate classical information, only.
3. *Sharpness:* $E(\varrho_{AB}) = 0$ if and only if the state ϱ_{AB} is separable
4. *Subadditivity:* $E(\varrho_{AB} \otimes \sigma_{CD}) \leq E(\varrho_{AB}) + E(\sigma_{CD})$
5. *Pure states:* $E(|\psi\rangle_{AB}\langle\psi|) = S(\varrho_A)$ where $\varrho_A = \mathrm{Tr}_B|\psi\rangle_{AB}\langle\psi|$

3.5.3 Conclusion

In this chapter we have introduced the notion of entanglement and we have showed its structural meaning. Thus, what is the entanglement? It was defined as a pure quantum characteristics of physical systems, that cannot be found in the classical domain. The reason is in different mathematical structures representing the objects of both theories. We started this chapter by comparing the notions of entanglement and correlations. We mentioned that it is not clear what such correlations represent and therefore it is also questionable what the entanglement is.

This chapter did not concern all aspects of entanglement [3]. We demonstrated properties of entanglement mainly of two-qubit systems. Most of the results can be simply generalized for the case of more dimensional subsystems, except the uniqueness of the partial transposition. We left untouched the problem of entanglement measures, which is still an open challenge. We only mentioned some possibilities how to quantify the entanglement for pure states, where everything is "clear", and we introduced the concurrence (and tangle) that measures the entanglement between two qubits. Finally we qualitatively described the behavior of entanglement under quantum operations and formulated the properties it needs to satisfy. The whole chapter has been only a brief introduction into the huge research area concerning the entanglement, which meaning and role is still not satisfactorily understood and explained. We will turn back to the problem of entanglement in the last chapter, where we will be interested in the behavior of multi-partite entanglement.

[3] A nice review of the problem of entanglement one can be found in [11].

Chapter 4

Quantum dynamics and communication

By convention there is colour, ...
but in reality there are atoms and space
Democritus

Dynamics is a rule that describes a change of states of a system with the flow of time, i.e. the time dependence of quantum states. Each dynamical transformation $\mathcal{E}$ of states caused by dynamics has to be an element of the set $\mathcal{L}(\mathcal{T}_2)$ (= completely positive tracepreserving maps) that transforms the set of states $\mathcal{S}(\mathcal{H})$ onto a subset $\mathcal{F}_{\mathcal{E}} \subset \mathcal{S}(\mathcal{H})$. In closed systems, i.e. systems that do not interact with an environment, the dynamics should be *reversible*. Because of the reversibility it has to be a bijection, i.e. $\mathcal{F}_{\mathcal{E}} = \mathcal{S}(\mathcal{H})$. Moreover, the requirement of the mean values ($\langle \mathbf{P}_\psi \rangle_\varrho$) preservation (for all $|\psi\rangle \in \mathcal{H}$ and all $\varrho \in \mathcal{S}(\mathcal{H})$)

$$\langle \mathbf{P}_\psi \rangle_\varrho = \langle \mathcal{E}[\mathbf{P}_\psi] \rangle_{\mathcal{E}[\varrho]} \tag{4.1}$$

implies (*Wigner theorem*) that $\mathcal{E}$ is either *unitary*, or *antiunitary*. The evolution as a function of time t will be denoted by $\mathcal{E}_t$. We expect that for time $t = 0$ the evolution is trivial, i.e. $\mathcal{E}_0 = \mathcal{I}$ and from the continuity of the evolution in time it follows that evolution of closed systems must be *unitary*. However, *open systems* can evolve also in a different way as we will see later.

4.1 Dynamics of composite systems

Dynamics of two qubits is described by elements $\mathcal{U}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Similarly to the case of observables we can define two basic subsets of unitaries: *local unitary transformations* $\mathcal{U}_{loc}(\mathcal{H}_A \otimes \mathcal{H}_B) \equiv \mathcal{U}_{loc}$ having the form $\mathbf{U}_{AB} = \mathbf{U}_A \otimes \mathbf{U}_B$ and *global unitary transformations*, i.e. $\mathbf{U}_{AB} \neq \mathbf{U}_A \otimes \mathbf{U}_B \in \mathcal{U}_{AB}(\mathcal{H}_A \otimes \mathcal{H}_B) \equiv \mathcal{U}_{AB}$. Any unitary transformation $\mathbf{U}$ can be written in the form $\mathbf{U} = \exp(i\mathbf{A})$ with $\mathbf{A} \in \mathcal{L}_s(\mathcal{H})$. For example, for two qubits the general unitary transformation can be expressed as

$$\mathbf{U}_{AB} = \exp\left[i \left(s\mathbb{1}_A \otimes \mathbb{1}_B + (\vec{\alpha}.\vec{\sigma}) \otimes \mathbb{1}_B + \mathbb{1}_A \otimes (\vec{\beta}.\vec{\sigma}) + \sum_{kl} \gamma_{kl}\sigma_k \otimes \sigma_l \right) \right] \tag{4.2}$$

Unitary transformation is local only if the associated generator has the form $\mathbf{A} \otimes \mathbb{1} + \mathbb{1} \otimes \mathbf{B}$. Then

$$\mathbf{U}_{AB} = \mathbf{U}_A \otimes \mathbf{U}_B = e^{i\mathbf{A}} \otimes e^{i\mathbf{B}} = e^{i[\mathbf{A}\otimes\mathbb{1}+\mathbb{1}\otimes\mathbf{B}]} \tag{4.3}$$

where the last equation is easy to get by applying the *Baker-Haussdorf formula* for commuting operators $\mathbf{X}, \mathbf{Y}$, $e^{\mathbf{X}+\mathbf{Y}} = e^{\mathbf{X}}e^{\mathbf{Y}}$. Note, that $\mathbf{U}_A \otimes \mathbf{U}_B = e^{i\mathbf{A}} \otimes e^{i\mathbf{B}} \neq e^{i\mathbf{A}\otimes\mathbf{B}}$. Therefore, not every local observable induces local unitary transformation. Physically, local unitary transformations describe the evolution of mutually non-interacting systems.

As we have already mentioned in Chapter II the time evolution of a closed quantum system is fully determined by a specific operator $\mathbf{H}$ called Hamiltonian. The equation of motion is the Schrödinger equation with the formal solution

$$\varrho_{AB}(t) = \mathcal{U}_t[\varrho_{AB}] := e^{-it\mathbf{H}}\varrho_{AB}e^{it\mathbf{H}} \tag{4.4}$$

where we put the Planck constant equals to unity, i.e. $\hbar = 1$. That is, the time evolution of the isolated composite system is represented by *one parametric group* of unitary transformations $\mathcal{U}_t$ generated by Hamiltonian $\mathbf{H}_{AB}$.

4.1.1 Reduced dynamics

In this section we shall derive how the states of the subsystem ϱ_A transform providing that their joint system evolves in a unitary way. In principle, there are many situations where we are interested only in the state of one subsystem (say A). For example, the evolution of open systems, or coherent quantum manipulations (as we shall see later). The second system is still present, but for us it is experimentally unreachable (we are not able to manipulate, or control it), or we are simply not interested in its behavior. The transformation of the system A alone is obtained by the partial trace rule

$$\varrho_A' = \mathcal{E}[\varrho_A] = \mathrm{Tr}_B\mathbf{U}_{AB}\varrho_{AB}\mathbf{U}_{AB}^\dagger \tag{4.5}$$

where ϱ_{AB} was the initially prepared state of both systems and $\varrho_A = \mathrm{Tr}_B\varrho_{AB}$ was the original state of the subsystem A only.

In these general settings the map $\mathcal{E}$ can be written as the composition of the following three transformations: (i) $\mathcal{P} : \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_{AB})$, (ii) $\mathcal{U} : \mathcal{S}(\mathcal{H}_{AB}) \to \mathcal{S}(\mathcal{H}_{AB})$, (iii) $\mathcal{T} : \mathcal{S}(\mathcal{H}_{AB}) \to \mathcal{S}(\mathcal{H}_A)$. The map $\mathcal{P}$ is not well defined. It assigns a state $\varrho_{AB} = \mathcal{P}[\varrho_A] \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ to each state $\varrho_A \in \mathcal{S}(\mathcal{H}_A)$, but there are many possibilities, in which such an assignment can be done, only the validity of the relation $\mathrm{Tr}_B\varrho_{AB} = \varrho_A$ is required. The transformation $\mathcal{U}$ is the unitary evolution of a composite system $A + B$ and finally, the map $\mathcal{T}$ corresponds to a partial trace, which describes the removing of the second system from our further consideration. In a sense $\mathcal{P}$ is an inverse operation to the partial trace $\mathcal{T}$, but as we know the partial trace is not a bijection. Let the composite system be in a state $\varrho_{AB} = \sum R_{kl,\alpha\beta}|k\rangle\langle l| \otimes |\alpha\rangle\langle\beta|$. The general unitary transformation can be expressed in the following form (see the next chapter) $\mathbf{U} = \sum_{\mu,\nu}\mathbf{A}_{\mu\nu} \otimes |\mu\rangle\langle\nu|$, where $\mathbf{A}_{\mu\nu} = \langle\mu|\mathbf{U}|\nu\rangle$. Then the mapping $\mathcal{E}$ reads

$$\mathcal{E}[\varrho_A] = \sum_{kl\alpha\beta\mu} R_{kl,\alpha\beta}\mathbf{A}_{\mu\alpha}|k\rangle\langle l|\mathbf{A}_{\beta\mu}^\dagger \tag{4.6}$$

There is one problem with the equation above. The same state $\varrho = \sum_{kl}\varrho_{kl}|k\rangle\langle l|$ with $\varrho_{kl} = \sum_\alpha R_{kl,\alpha\alpha}$ can be obtained from many different joint states ϱ_{AB} and the mapping $\mathcal{E}$ strongly depends on such choice. Therefore, in order to obtain a well defined dynamics of our subsystem (valid for all the states ϱ_A) we need to define the mapping $\mathcal{P}$ in a consistent way. *Linearity* is quite a general requirement on any dynamics of quantum systems [34]. In order to preserve the linearity of $\mathcal{E} = \mathcal{T} \circ \mathcal{U} \circ \mathcal{P}$ we need to define the assignment $\mathcal{P}$ to be linear. Note that $\mathcal{U}$ and $\mathcal{T}$ are already linear. Since we demand the validity of $\mathcal{P}$ for any unitary transformation $\mathcal{U}$, the only possible choice [33] is the mapping $\mathcal{P} : \varrho_A \mapsto \varrho_A \otimes \xi_B$, where ξ_B is a fixed state of the subsystem B. Therefore, it is standard to use an initially factorized state of the systems A and B, when speaking about dynamics of open systems. Such assumption (of initially uncorrelated particles) is even experimentally quite acceptable.

Consider that $\mathbf{U}_{AB}$ is given and the initial state is uncorrelated, i.e. $\varrho_{AB} = \varrho_A \otimes \varrho_B$ Then the induced mapping

$$\varrho_A' = \mathcal{E}[\varrho_A] = \mathrm{Tr}_B\mathbf{U}_{AB}\varrho_A \otimes \varrho_B\mathbf{U}_{AB}^\dagger = \sum_{kl}\mathbf{M}_{kl}\varrho_A\mathbf{M}_{kl}^\dagger \tag{4.7}$$

where (in spectral form) $\varrho_B = \sum_k \lambda_k |k\rangle\langle k|$ and $\mathbf{M}_{kl} = \sqrt{\lambda_k}{}_B\langle l|\mathbf{U}_{AB}|k\rangle_B$. This linear map $\mathcal{E}$ has been obtained in the following three steps: (i) adding an uncorrelated ancilla in the state ϱ_B, (ii) realizing the unitary transformation $\mathbf{U}_{AB}$ and (iii) tracing out the ancilla system B, i.e. $\mathcal{E} = \Lambda_B \circ \mathcal{U} \circ \Lambda_{\varrho_B}$. According to our discussion in the previous chapter it follows that the evolution mapping $\mathcal{E}$ is necessarily a completely positive and tracepreserving map. The representation of $\mathcal{E}$ by a collection of operators $\{\mathbf{M}_k\}$ will be called *Kraus representation.* In what follows by a *general dynamical map* $\mathcal{E}$ will be understood any mapping satisfying the following conditions

1. *Linearity:* $\mathcal{E}[(1-\lambda)\varrho_1 + \lambda\varrho_2] = (1-\lambda)\mathcal{E}[\varrho_1] + \lambda\mathcal{E}[\varrho_2]$
2. *Tracepreservity:* $\mathrm{Tr}\mathcal{E}[\varrho] = \mathrm{Tr}\varrho = 1$
3. *Complete positivity:* $\mathcal{E} \otimes \mathcal{I}_n \geq 0$ for all $n \in \mathbb{N}$

The complete positivity reflects physical situations, such that the system is entangled with another system, before it starts to interact with the ancilla system B. The complete positivity of $\mathcal{E}$ ensures that the whole state of all three systems remains positive, that is, it remains still a quantum state. In other words, if $\mathcal{E}$ is not completely positive, then we are not able to find an induced unitary operation acting on the joint Hilbert space of ancilla and the system under consideration. We shall call the evolution maps $\mathcal{E}$ *superoperators.* Let us list the basic classification of possible types of superoperators:

- **1.Unitary superoperators.**
  This class of superoperators corresponds to the set of all unitary operators $\mathbf{U}$ defined on the original Hilbert space $\mathcal{H}$. The relation $\mathcal{U}[\varrho] = \mathbf{U}\varrho\mathbf{U}^\dagger$ defines the unitary transformations as mappings on the set of states $\mathcal{S}(\mathcal{H})$. The set of unitary transformations $\mathcal{U}$ possesses all the group properties according to composition of mappings. Only unitary superoperators are invertible. Therefore, no other subset od superoperators (which is not a subset of unitary transformations) can form a group. The number of fixed points, i.e. elements ϱ of $\mathcal{S}(\mathcal{H})$, such that $\mathcal{U}[\varrho] = \varrho$, is infinite. namely, the fixed points are given by eigenvectors of $\mathbf{U}$ in the following sense. Every convex combination of corresponding eigenvectors is a fixed point. We say that the set of fixed points form a *simplex* with eigenvectors as extremal points. If the spectrum of $\mathbf{U}$ is degenerated, then this set is more complicated.

- **2.Unital superoperators.**
  This class is characterized by its action on the maximally mixed state of $\mathcal{S}(\mathcal{H})$, i.e. $\frac{1}{d}\mathbb{1}$. If the transformation $\mathcal{E}$ preserves a total mixture, i.e. $\mathcal{E}[\frac{1}{d}\mathbb{1}] = \frac{1}{d}\mathbb{1}$, then $\mathcal{E}$ is called *unital.* It means that such superoperators have one specific fixed point. In Kraus representation $\mathcal{E}[\varrho] = \sum_k \mathbf{M}_k\varrho\mathbf{M}_k^\dagger$ the unitality is equivalent to the condition $\sum_k \mathbf{M}_k\mathbf{M}_k^\dagger = \mathbb{1}$. It means that also the mapping $\mathcal{E}'[\varrho] = \sum_k \mathbf{M}_k^\dagger\varrho\mathbf{M}_k$ describes a regular evolution, but it is not an inverse map to $\mathcal{E}$. Sometimes the unital maps are called *bistochastic.* They form a convex subset of the set of all superoperators.

- **3.Contractive superoperators.**
  Let $D(.,.)$ be a *distance function* defined on the set of all states $\mathcal{S}(\mathcal{H})$. For example, we can use the *trace distance* $D(\varrho,\varrho') := \mathrm{Tr}|\varrho - \varrho'|$. Each superoperator $\mathcal{E}$ satisfying the property $D(\mathcal{E}[\varrho],\mathcal{E}[\varrho']) \leq kD(\varrho,\varrho')$ with $k < 1$ for all $\varrho,\varrho' \in \mathcal{S}(\mathcal{H})$ will be called *contractive.* As a consequence we get that $\mathcal{E}$ has a unique fixed point ϱ_f and repeated iterations $\mathcal{E}^N = \mathcal{E} \circ \ldots \circ \mathcal{E}$ transform arbitrary initial state ϱ into this fixed state ϱ_f, i.e. $\mathcal{E}^N[\varrho] \to \varrho_f$ for all $\varrho \in \mathcal{S}(\mathcal{H})$ with $N \to \infty$. We shall see later that the set of contractive superoperators is convex, too. Moreover, it can be shown [30] that this set is *dense* in the set of all superoperators.

The evolution described by $\mathcal{E}$ of a system A is physically caused by its interactions $\mathbf{U}_{AE}$ with the external system usually called *environment* E. In the case when $\mathcal{E}$ is unitary the interaction should obviously take the trivial form $\mathbf{U}_{AB} = \mathbf{U}_A \otimes \mathbf{U}_B$, but it is not always the case. It means the system A and B can interact and still the evolution can be unitary. Here one needs to specify what we mean by an *interaction.* We will say that two systems interact one with another, if the overall unitary evolution is not local, i.e. if the generating Hamiltonian does not have the form $\mathbf{H}_{AB} \neq \mathbf{H}_A \otimes \mathbb{1} + \mathbb{1} \otimes \mathbf{H}_B$. It is easy to verify that the transformation $\mathbf{U}_{AB} = \sum_k \mathbf{V}_k \otimes |k\rangle\langle k|$ (with $\mathbf{V}_k^{-1} = \mathbf{V}_k^\dagger$ and $|k\rangle$ being a basis of $\mathcal{H}_B$) is unitary and obviously non-local. If the environment B is initially in the specific state $|k\rangle$, then system A evolves according to mapping $\varrho_A \to \mathbf{V}_k \varrho_A \mathbf{V}_k^\dagger$, i.e. in the unitary way. In the case of unitary evolutions the purity of quantum states is preserved, but in all other cases the purity can be changed in both directions. It is possible to evolve a system into the total mixture, or *vice versa*, the total mixture can be transformed into a pure state.

4.2 One qubit's superoperators

4.2.1 Left-right formalism

We shall introduce the so-called *left-right (matrix) representation* of superoperators. Let $\Lambda \in \mathcal{L}_{tcp}(\mathcal{T},\mathcal{T})$ and $\varrho \in \mathcal{S}(\mathcal{H}) \subset \mathcal{T}_2(\mathcal{H})$. We remind us that states form a closed convex subset of the Hilbert space $\mathcal{T}_2(\mathcal{H})$ endowed with scalar product $(\varrho|\sigma) := \mathrm{Tr}\varrho^\dagger\sigma$. For convenience, we shall denote the elements of $\mathcal{T}_2(\mathcal{H})$ by vector kets $|\varrho)$. In this notation the superoperators take a form of matrices with elements $\mathcal{E}_{mn} = (\Theta_m|\mathcal{E}|\Theta_n) = \mathrm{Tr}(\Theta_m^\dagger \mathcal{E}[\Theta_n])$, where the Hilbert space operators (="vectors") $|\Theta_m)$ form an orthonormal basis in $\mathcal{T}_2(\mathcal{H})$, i.e. $(\Theta_m|\Theta_n) = \delta_{mn}$.

In what follows, we shall choose the orthonormal basis of a qubit system consisting of σ-matrices, i.e. $\Theta_0 = \frac{1}{\sqrt{2}}\mathbb{1}, \Theta_1 = \frac{1}{\sqrt{2}}\sigma_x, \Theta_2 = \frac{1}{\sqrt{2}}\sigma_y, \Theta_3 = \frac{1}{\sqrt{2}}\sigma_z$. Remind that σ-matrices do not form an orthonormal basis, because $\mathrm{Tr}\sigma_k\sigma_l = 2\delta_{kl}$, but the operators Θ_k do. Then the elements of $\mathcal{S}(\mathcal{H})$ in this basis have the form $|\varrho) = \frac{1}{2}\mathbb{1} + \vec{n}.\vec{\sigma} = (1/\sqrt{2}, n_x, n_y, n_z) \equiv (n_0, \vec{n})$ and we will often use $\vec{n}$ to denote the state. In this basis $|\vec{n}| = 1/2$. Let us evaluate the element of general superoperator $\mathcal{E}$. Consider the realization of the transformation $\mathcal{E}$ by operators $\{\mathbf{M}_k\}_k$. Then

$$\begin{aligned}
\mathcal{E}_{0n} &:= \frac{1}{2}\sum_k \mathrm{Tr}\sigma_0 \mathbf{M}_k \sigma_n \mathbf{M}_k^\dagger = \frac{1}{2}\mathrm{Tr}(\sum_k \mathbf{M}_k^\dagger \mathbf{M}_k)\sigma_n = \delta_{0n} \\
\mathcal{E}_{m0} &:= \frac{1}{2}\sum_k \mathrm{Tr}\sigma_m \mathbf{M}_k \sigma_0 \mathbf{M}_k^\dagger = \frac{1}{2}\mathrm{Tr}\sigma_m(\sum_k \mathbf{M}_k \mathbf{M}_k^\dagger) \equiv (2\vec{e})_m \\
\mathcal{E}_{mn} &:= \frac{1}{2}\sum_k \mathrm{Tr}\sigma_m \mathbf{M}_k \sigma_n \mathbf{M}_k^\dagger \equiv E_{mn} \qquad (4.8)
\end{aligned}$$

where we defined the new real vector $\vec{e}$ and real (if operators Θ_k are hermitian) matrix E. For the matrix corresponding to the superoperator $\mathcal{E}$ we have

$$\mathcal{E} = \begin{pmatrix} 1 & \vec{0}^T \\ 2\vec{e} & E \end{pmatrix} \qquad (4.9)$$

where $\vec{0}^T$ denotes the transposed zero vector. From the mathematical point of view the superoperators (in this basis) correspond to matrices associated with affine transformations on a real (three-dimensional) vector space. Of course, not each affine matrix can be a superoperator, because superoperators must be linear completely positive and tracepreserving. The linearity follows from the matrix representation. The tracepreservity is trivially ensured by the affine form, i.e. the first row implies the preservation of the first element of density vector. We note that (in the chosen operator basis) $\mathrm{Tr}\varrho = 2n_0 = 1$. After the action of $\mathcal{E}$ the state $\vec{n}$ evolves according to the rule

$$\vec{n} \mapsto \vec{n}' = \vec{e} + E\vec{n}\,. \qquad (4.10)$$

In the standard form $\varrho = \frac{1}{2}\mathbb{1} + \vec{n}.\vec{\sigma} \mapsto \varrho' = \mathcal{E}[\varrho] = \frac{1}{2}\mathbb{1} + (\vec{e} + E\vec{n}).\vec{\sigma}$. To ensure that the final operator corresponds to a density matrix the vector $|\vec{n}'| = |\vec{e} + E\vec{n}| \leq 1$ for all $|\vec{n}| \leq 1$. This condition puts non-trivial requirements on the choices of E and $\vec{e}$. However, this condition just reflects the positivity of the transformation $\mathcal{E}$. It is not easy to exhibit the complete positivity in this left-right formalism.

The fixed points $\vec{f}$ are given by solving the equation $\vec{e} + E\vec{f} = \vec{f}$. Consequently, the formal solution is $\vec{f} = (I - E)^{-1}\vec{e}$.

It is simple to rewrite the unitary superoperators $\mathcal{U}[\varrho] = \mathbf{U}\varrho\mathbf{U}^\dagger$ in this matrix form. We get

$$\mathcal{U} = \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & U \end{pmatrix}. \tag{4.11}$$

where U is an orthogonal rotation in three-dimensional real vector space. There is one-to-one correspondence between the special unitary transformations $SU(2)$ and special orthogonal rotations $SO(3)$, but this bijection is valid only for two-dimensional Hilbert space. Such a bijection does not hold in general. In particular, any unitary transformation $\mathbf{U} \in SU(2)$ can be written as $\mathbf{U}_{\alpha,\vec{r}} = \exp(i\alpha\vec{r}.\vec{\sigma})$, where $\alpha \in \mathbb{R}$ and $\vec{r}$ is a normalized three-dimensional real vector. Then the corresponding matrix U of the superoperator $\mathcal{U}$ takes the form of the special orthogonal rotation $O_{2\alpha,\vec{r}}$, which is associated with the rotation around the axis determined by $\vec{r}$ by the angle 2α. It is easy to verify the relation

$$U_{kl} = \frac{1}{2}\mathrm{Tr}\left[\sigma_k \mathbf{U}_{\alpha,\vec{r}}\sigma_l \mathbf{U}^\dagger_{\alpha,\vec{r}}\right] = \delta_{kl}\cos 2\alpha + (1-\cos 2\alpha)r_k r_l - \varepsilon_{klm}r_m \sin 2\alpha = O_{2\alpha,\vec{r}}\,. \tag{4.12}$$

The unitary transformation $\mathbf{U}$ of one qubit has two mutually orthogonal eigenvectors. The corresponding eigenstates are $\vec{n}$ and $-\vec{n}$ with the norm $|\vec{n}| = 1/2$. Since $\vec{e} = \vec{0}$ we get the condition for fixed points $U\vec{f} = \vec{f}$. Trivially for eigenevectors $U(\pm\vec{n}) = \pm\vec{n}$, i.e. they are fixed points of the unitary map $\mathcal{U}$. Moreover, an arbitrary convex combination of them, $\vec{f} = \lambda\vec{n} - (1-\lambda)\vec{n}$, is a fixed point of $\mathcal{U}$, too. Let us remind the *Bloch ball* representation of the states of a qubit. Mutually orthogonal eigenvectors correspond to antipode points on the boundary of the Bloch ball, i.e. on the surface of the *Bloch sphere.* The line connecting these two points represents the convex set of all fixed points of $\mathcal{U}$. Since the eigenvectors are opposite, this line contains also the total mixture $\vec{m} = \frac{1}{2}(\vec{n} - \vec{n}) = (0,0,0)$. Unitary transformations $\mathcal{U}$ leave the points lying on this line unaffected. For unital maps we obtain the superoperator $\mathcal{E}$ of the same form, because in this case $\sum_k \mathbf{M}_k\mathbf{M}_k^\dagger = \mathbb{1}$, too. Therefore, $\vec{e} = \vec{0}$. But now, U is not an orthogonal rotation.

Let us consider a d dimensional Hilbert space $\mathcal{H}$. and let Θ_k be the orthonormal basis of the Hilbert-Schmidt space $\mathcal{T}_2(\mathcal{H})$, i.e. $k = 0,\dots,d^2-1$ and $(\Theta_k|\Theta_l) = \delta_{kl}$. The constraint that $\Theta_0 = \frac{1}{\sqrt{d}}\mathbb{1}_d$ ensures that the superoperator $\mathcal{E}$ will have the same form as for the qubit. In this notation the states $\mathcal{S}(\mathcal{H})$ have the form $\varrho = \frac{1}{d}\mathbb{1}_d + \vec{n}\cdot\vec{\Theta}$, where all Θ_k are traceless. If, moreover, Θ_k are hermitian, then $\vec{n}$ is a real d^2-1 dimensional vector. The basis of $\mathcal{T}_2(\mathcal{H})$ with traceless and hermitian operators Θ_k (except the case of Θ_0) always exists, because they form generators of the *Lie algebra $su(d)$*. And we know that such Lie algebra forms a d^2-1 dimensional real linear space. In analogy with the qubit we can define the action of the superoperator $\mathcal{E}$ in a matrix form $\mathcal{E}[\varrho] = \mathcal{E}[\vec{n}] = \vec{n}' = \vec{e} + E\vec{n}$, where E and $\vec{e}$ are defined as before (see eq.(4.8)). Again, we get the superoperator as a specific affine map on vectors $\vec{n} \in \mathbb{R}^{d^2-1}$. Let us find the most general unitary map $\mathcal{E}$ with respect to the *Hilbert-Schmidt* scalar product between two elements σ, ϱ given by

$$\begin{aligned}(\varrho\,|\,\xi) &= \mathrm{Tr}(\frac{1}{d}\mathbb{1} + \vec{n}.\vec{\Theta})(\frac{1}{d}\mathbb{1} + \vec{m}.\vec{\Theta}) = \frac{1}{d^2}\mathrm{Tr}\mathbb{1} + \sum_{k,l=1}^{d^2-1} n_k m_l \mathrm{Tr}\Theta_k\Theta_l \\ &= \frac{1}{d} + \vec{n}.\vec{m}\end{aligned} \tag{4.13}$$

where we used $\mathrm{Tr}\Theta_k\Theta_l = \delta_{kl}$. Then $(\mathcal{E}[\varrho]\,|\,\mathcal{E}[\xi]) = \frac{1}{d} + (\vec{e} + E\vec{n}).(\vec{e} + E\vec{m})$ and

$$(\vec{e} + E\vec{n}).(\vec{e} + E\vec{m}) = \vec{n}.\vec{m} \tag{4.14}$$

is the condition of unitarity of $\mathcal{E}$. Note that $\vec{n}.\vec{m}$ is the standard scalar product in $\mathbb{R}^{d^2-1}$. For unitary mappings $\mathcal{U}$ the induced vector $\vec{e} = \vec{0}$ and the unitarity condition implies $U^{-1} = U^T$, because the space of vectors $\vec{n}$ is real. As a result, to each unitary transformation $\mathcal{U}$ on the Hilbert space $\mathcal{H}$ it corresponds an orthogonal rotation U defined on $\mathbb{R}^{d^2-1}$. However, the opposite implication (except qubit) does not hold, because of the inequality $u(d) = d^2 < \frac{(d^2-2)(d^2-1)}{2} + 1 = o(d^2-1)$ that compares dimensions of unitary transformations and orthogonal transformations.

We turn back to the qubit case. It is now easy to illustrate the actions of unitary superoperators on the Bloch ball. They correspond to three-dimensional rotations of this ball around the axis determined by the line connecting the eigenvectors of the original operator $\mathbf{U}$. Since the basis of $\sigma-$matrices is not orthonormal, but $\mathrm{Tr}\sigma_k\sigma_l = 2\delta_{kl}$, for the Hilbert-Schmidt scalar product in this basis we obtain

$$(\varrho|\xi) = \frac{1}{4}\mathrm{Tr}\mathbb{1} + \sum_{k,l=1}^{3} n_k m_l \mathrm{Tr}\sigma_k\sigma_l = \frac{1}{2} + 2\vec{n}.\vec{m} \tag{4.15}$$

Let the distance $D(.,.)$ be given by the *trace distance*, i.e.

$$D(\varrho,\xi) := \mathrm{Tr}|\varrho - \xi| = \mathrm{Tr}|(\vec{n}-\vec{m}).\vec{\sigma}| = \mathrm{Tr}|\vec{r}.\vec{\sigma}| \tag{4.16}$$

where we put $\vec{r} = \vec{n} - \vec{m}$. Note that the operator $\vec{r}.\vec{\sigma}$ is selfadjoint and its eigenvalues are $\lambda_\pm = \pm|\vec{r}|$. It means $D(\varrho,\xi) = 2|\vec{r}|$. If we transform both states ϱ, σ according to the map $\mathcal{E} = (\vec{e}, E)$ the distance between them changes

$$D(\mathcal{E}[\varrho],\mathcal{E}[\xi]) = 2|\vec{e} + E\vec{n} - \vec{e} - E\vec{m}| = 2|E\vec{r}|. \tag{4.17}$$

It means that the contractivity of $\mathcal{E}$ is equivalent to the contractivity of the transformation E.

Another important property of the trace distance is its non-increasing under tracepreserving completely positive maps $\mathcal{E}$, i.e.

$$D(\mathcal{E}[\varrho],\mathcal{E}[\xi]) \leq D(\varrho,\xi). \tag{4.18}$$

To prove this inequality we need to consider, that the operator $\varrho - \xi$ is still selfadjoint. Every selfadjoint operator can be expressed in the spectral form. Since it is not necessarily positive, we define two positive operators $\mathbf{X}$ and $\mathbf{Y}$ such that $\varrho - \xi = \mathbf{X} - \mathbf{Y}$ and $[\mathbf{X},\mathbf{Y}] = 0$. It means that the eigenvalues of $\mathbf{X}$ are positive eigenvalues of $\varrho - \xi$, and $\mathbf{Y}$ has originally negative eigenvalues of $\varrho - \xi$, but taken with positive sign. Then the trace distance $\mathrm{Tr}|\varrho - \xi| = \mathrm{Tr}|\mathbf{X} - \mathbf{Y}| = \mathrm{Tr}\mathbf{X} + \mathrm{Tr}\mathbf{Y}$. Without absolute value we have $0 = \mathrm{Tr}(\varrho - \xi) = \mathrm{Tr}\mathbf{X} - \mathrm{Tr}\mathbf{Y}$. Hence $D(\varrho,\xi) = \mathrm{Tr}|\varrho - \xi| = 2\mathrm{Tr}\mathbf{X}$. Moreover, we can write

$$D(\varrho,\xi) = 2\max_{\mathbf{P}}\mathrm{Tr}[\mathbf{P}(\varrho - \xi)] = 2\mathrm{Tr}\mathbf{X} \tag{4.19}$$

where $\mathbf{P} = \mathbf{P}^\dagger = \mathbf{P}^2$ is a projector onto a subspace of $\mathcal{H}$. The proof is given in the following steps

$$\begin{aligned} D(\varrho,\xi) &= 2\mathrm{Tr}\mathbf{X} = 2\mathrm{Tr}\mathcal{E}[\mathbf{X}] \geq 2\mathrm{Tr}\mathbf{P}\mathcal{E}[\mathbf{X}] \\ &\geq 2\mathrm{Tr}\mathbf{P}(\mathcal{E}[\mathbf{X}] - \mathcal{E}[\mathbf{Y}]) = 2\mathrm{Tr}\mathbf{P}(\mathcal{E}[\varrho] - \mathcal{E}[\xi]) \\ &= D(\mathcal{E}[\varrho],\mathcal{E}[\xi]) \end{aligned} \tag{4.20}$$

where in the second equality the tracepreservity of $\mathcal{E}$ we used.

Using this distance we can show that the set of contractive superoperators is a convex subset of the whole set of superoperators. In particular, giving two contractive superoperators $\mathcal{E}_1$ and $\mathcal{E}_2$, then also $\mathcal{E} = \lambda\mathcal{E}_1 + (1-\lambda)\mathcal{E}_2$ is a contractive superoperator. Since the trace distance is induced by the trace norm, i.e. $D(\varrho,\xi) \equiv ||\varrho - \xi||_1 := \mathrm{Tr}|\varrho - \xi|$, we can write

$$\begin{aligned} D(\mathcal{E}[\varrho],\mathcal{E}[\xi]) &= ||\mathcal{E}[\varrho] - \mathcal{E}[\xi]||_1 = ||\lambda\mathcal{E}_1[\varrho - \xi] + (1-\lambda)\mathcal{E}_2[\varrho - \xi]||_1 \\ &\leq \lambda||\mathcal{E}_1[\varrho] - \mathcal{E}_1[\xi]||_1 + (1-\lambda)||\mathcal{E}_2[\varrho] - \mathcal{E}_2[\xi]||_1 \\ &\leq [\lambda k_1 + (1-\lambda)k_2]D(\varrho,\xi) = kD(\varrho,\xi) \end{aligned}$$

where $k = \lambda k_1 + (1-\lambda)k_2 < 1$ and $k_1 < 1, k_2 < 1$ are the contractive coefficients for $\mathcal{E}_1, \mathcal{E}_2$, respectively. As a result, we get the convexity of the set of all contractive superoperators. Moreover, the fixed point $\vec{f}$ of $\mathcal{E}$ is given as the sum of fixed points $\vec{f_1}, \vec{f_2}$ of superoperators $\mathcal{E}_1, \mathcal{E}_2$, i.e. $\vec{f} = \lambda \vec{f_1} + (1-\lambda)\vec{f_2}$.

Next, we shall show, that also the composition $\mathcal{E}_1 \circ \mathcal{E}_2$ of two contractive maps $\mathcal{E}_1, \mathcal{E}_2$ is again a contractive map. We have argued (for the qubit) that the contractivity of a map $\mathcal{E}$ is equivalent to the contractivity of an induced matrix E. The composition of two affine maps results in the affine transformation with $\vec{e} = \vec{e_1} + E_1\vec{e_2}$ and $E = E_1E_2$. That is, to show the contractivity of $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2$ it is enough to prove the contractivity of the matrix E_1E_2. Since in general $|E_1E_2\vec{r}| \leq ||E_1E_2|| \cdot |\vec{r}| \leq ||E_1|| \cdot ||E_2|| \cdot |\vec{r}|$ and $||E_1|| = k_1 < 1$, $||E_2|| = k_2 < 1$, we obtain that also $||E|| \leq k_1k_2 < 1$ and $|E\vec{r}| < |\vec{r}|$. Therefore, the composition of two contractive maps is again contractive. To specify a new fixed point $\vec{f}$ is quite a difficult problem and according to the knowledge of the author such solution is not known.

4.2.2 Examples

1.Pauli superoperator

Define the set of operators $\mathbf{M}_0 = \sqrt{1-p}\mathbb{1}, \mathbf{M}_k := \sqrt{p_k}\sigma_k$ where $p = p_1 + p_2 + p_3$. They trivially satisfy the normalization condition. In the left-right form we get for the *Pauli superoperator* $\mathcal{P}$

$$\mathcal{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-2(p_2+p_3) & 0 & 0 \\ 0 & 0 & 1-2(p_1+p_3) & 0 \\ 0 & 0 & 0 & 1-2(p_1+p_2) \end{pmatrix}. \tag{4.21}$$

It means that the initial state $\varrho = \vec{n}$ is transformed into the state $\varrho' = \vec{n}' = ((1-2(p_2+p_3))n_1, (1-2(p_1+p_3))n_2, (1-2(p_1+p_2))n_3)$. The first column of matrix $\mathcal{P}$ indicates, that *Pauli superoperators* are *unital.* Their action corresponds to a contraction and a rotation of the Bloch ball, but the center (total mixture) remains preserved. Note that not each superoperator $\mathcal{E}$ with the diagonal (with respect to some basis) matrix E and a zero vector $\vec{e} = \vec{0}$ is the *Pauli superoperator* (for example see *amplitude damping channel*). The condition for fixed points implies $(p_2+p_3)f_1 = 0, (p_1+p_3)f_2 = 0, (p_1+p_2)f_3 = 0$. In specific cases, if at most one $p_j = 0$, the only fixed point is the total mixture, i.e. $\vec{f} = \vec{m} = (0,0,0)$. In this case the Pauli superoperator is also contractive, since $D(\mathcal{P}[\varrho], \mathcal{P}[\xi]) \leq kD(\varrho, \xi)$ with $k = 1 - 2\min\{p_1+p_2, p_1+p_3, p_2+p_3\} < 1$. In the specific case when $p_1 = p_2 = p_3 \equiv q \in [0, 1/3]$ and $p = 1-3q$, the Pauli superoperator is called *depolarizing superoperator.* Such superoperator describes the pure (and symmetric) contraction of the Bloch sphere with the contractivity parameter $1-4q$ and the final state is $\vec{n}' = (1-4q)\vec{n}$. If we put $p_1 = p_2 = 0$ and $p_3 = p$, we get the so-called *x-Pauli superoperator.*

2.Phase damping superoperator

Let us introduce this class of superoperators $\mathcal{C}_p$ by the relation

$$\varrho \mapsto \mathcal{C}_p[\varrho] = (1-p)\varrho + p\begin{pmatrix} \varrho_{00} & 0 \\ 0 & \varrho_{11} \end{pmatrix} = \begin{pmatrix} \varrho_{00} & (1-p)\varrho_{01} \\ (1-p)\varrho_{10} & \varrho_{11} \end{pmatrix} \tag{4.22}$$

where ϱ_{jk} are the elements of the input state ϱ. It is easy to see that the iterations of this map transform each state into its diagonal form in a preferred basis (in our case $|0\rangle, |1\rangle$), i.e. $\varrho' = \mathcal{C}_p^N[\varrho] \to \mathrm{diag}(\varrho_{00}, \varrho_{11})$. This evolution is typical of *decoherence processes*, because the quantumness hidden in the superposition principle vanishes together with the off-diagonal elements in a given basis. In the left-right form we obtain

$$\mathcal{C}_p = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-p & 0 & 0 \\ 0 & 0 & 1-p & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.23}$$

Comparing this matrix with the previous example we see, that if we put $p_3 = p/2$ and $p_1 = p_2 = 0$ in the *Pauli superoperator* matrix $\mathcal{P}$, then we will get the *phase damping superoperator* with $\mathbf{M}_0 = \mathbb{1}\sqrt{1-p/2}$ and $\mathbf{M}_1 = \sqrt{p/2}\sigma_z$ being the corresponding operators in the *Kraus representation.* Using the previous example we can conclude, that $\mathcal{C}_p$ is unital and the fixed points read $\vec{f} = (0,0,z)$ for all $z \in [-1/2, 1/2]$. It reflects the fact that diagonalized matrices do not decohere. Let us note that the presented phase damping superoperator is exactly the z-Pauli superoperator and causes the diagonalization in the eigenbasis of the operator σ_z.

3.Amplitude damping superoperator

Previous example of the Pauli superoperator described the evolution, in which pure states evolved into mixtures. The amplitude damping superoperator $\mathcal{A}$ describes in a sense the "opposite" transformation, where the mixed states evolve into the pure state $|0\rangle\langle 0|$. If we define the superoperator $\mathcal{A}_p$ by Kraus operators $\mathbf{M}_1 = |0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1|$ and $\mathbf{M}_2 = \sqrt{p}|0\rangle\langle 1|$, we obtain the map

$$\mathcal{A}_p = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & \sqrt{1-p} & 0 \\ p & 0 & 0 & 1-p \end{pmatrix} \tag{4.24}$$

that serves our purposes. Applying such transformation on $\vec{n}$ we obtain $\vec{n}' = (n_1\sqrt{1-p}, n_2\sqrt{1-p}, p/2 + (1-p)n_3)$, because $\mathcal{A} = (\vec{a}, A)$ with the vector $\vec{a} = (0,0,p/2)$. Since

$$\begin{aligned} D(\mathcal{A}_p[\varrho], \mathcal{A}_p[\xi]) &= 2|A\vec{r}| = 2|(r_1\sqrt{1-p}, r_2\sqrt{1-p}, (1-p)r_3)| \\ &\leq 2\sqrt{1-p}|\vec{r}| = \sqrt{1-p}D(\varrho, \xi) \end{aligned} \tag{4.25}$$

and $\sqrt{1-p} < 1$ (if $p \neq 0$), we can conclude that amplitude damping superoperator is *contractive.* The unique fixed point $\vec{f}$ is given by the equation $\vec{a} + A\vec{f} = (\sqrt{1-p}f_1, \sqrt{1-p}f_2, \frac{p}{2} + (1-p)f_3) = (f_1, f_2, f_3)$. To fulfill such condition we need to put $f_1 = f_2 = 0$ and then $f_3 = 1/2$. Hence, the fixed point of $\mathcal{A}_p$ is the state

$$\varrho_f = \frac{1}{2}\mathbb{1} + \vec{f}.\vec{\sigma} = \frac{1}{2}(\mathbb{1} + \sigma_3) = |0\rangle\langle 0| \tag{4.26}$$

and the *Banach theorem* implies that the iterations of the superoperator $\mathcal{A}_p$ map any initial state into this fixed point. Thus, finally the whole Bloch ball transforms into the single pure state $|0\rangle\langle 0|$.

To each state ξ there exist many contractive superoperators $\mathcal{C}_\xi$ with the state ξ as a fixed point, i.e. $\mathcal{C}_\xi^N[\varrho] \to \xi$ for all states $\varrho \in \mathcal{S}(\mathcal{H})$, if $N \to \infty$. We have seen that some of the *Pauli superoperators* ensure the convergence to the total mixture. Hence, in that cases we can put $\mathcal{C}_{\frac{1}{2}\mathbb{1}} = \mathcal{P}$. The example of *amplitude damping superoperators* describes the possibility how to transform all the states ϱ into a pure state $|\psi\rangle\langle\psi|$. If we change $|0\rangle \mapsto |\psi\rangle$ and $|1\rangle \mapsto |\psi^\perp\rangle$, we obtain the amplitude damping superoperator $\mathcal{A}_p(\psi)$ transforming every state ϱ into the state $|\psi\rangle$. It means $\mathcal{C}_\psi = \mathcal{A}_p(\psi)$. The above example of the evolution of all states to a single pure state (ground state) $|0\rangle\langle 0|$, i.e. any transformation $\mathcal{C}_{|0\rangle\langle 0|}$, can be used to describe the *exponential decay* process.

4.3 Quantum dynamical semigroups

In this section we shall pay attention to the explicit time evolution of quantum states. At any time t the state transformation of the system corresponds to completely-positive trace-presereving (CPT) map $\mathcal{E}_t$. That is, the time dependency of the evolution is associated with a trajectory (one-parametric subset) drawn in the space of all possible CPT maps. Consequently, any one-parametric subset could be in principle considered as time evolution, but it is not the case. We know that for the isolated systems the time evolution is driven by the *Schrödinger* (or *von Neumann*) equation

$$\dot{\varrho}_t = i[\mathbf{H}, \varrho_t] \tag{4.27}$$

where $\mathbf{H}$ is the Hamiltonian of the system. Note that the existence of such equation of motion is an independent postulate of quantum theory. As a result of this postulate (solving this equation) we obtain that any time evolution is represented by one-parametric set of unitary transformations $\mathcal{U}_t$ with the following properties:

1. $\lim_{t\to\infty}\mathcal{U}_t = \mathcal{I}$
2. $\mathcal{U}_t\mathcal{U}_s = \mathcal{U}_{t+s}$ for all $t, s \in \mathbb{R}$

These two properties mean that the evolution $\mathcal{U}_t$ of isolated systems corresponds to a one-parametric group of unitary transformations.

The evolution of open systems is caused by the interaction of the system with its surroundings, but together these systems evolve in a unitary manner. This assumption results in the following time evolution

$$\varrho_t = \mathcal{E}_t[\varrho] = \mathrm{Tr}_B\mathcal{U}_t[\varrho_{AB}] \qquad (4.28)$$

where ϱ_{AB} is the initial state of the system and the environment such that $\varrho = \mathrm{Tr}_B\varrho_{AB}$. The overall unitary evolution $\mathcal{U}_t$ is induced by the hamiltonian $\mathbf{H} = \mathbb{1}\otimes\mathbf{H}_B + \mathbf{H}_A\otimes\mathbb{1} + \mathbf{H}_{int}$. Note that the initial state ϱ_{AB} is really arbitrary. However, to be able to describe the evolution $\mathcal{E}_t$ only in terms of CPT maps, we have to assume that the initial state is factorized. i.e. $\varrho_{AB} = \varrho\otimes\xi$.

The properties of $\mathcal{U}_t$ imply that $\mathcal{E}_t$ begins with the identity, i.e. $\lim_{t\to 0+}\mathcal{E}_t = \mathcal{I}$, but the second (group) property of $\mathcal{U}_t$ is valid for $\mathcal{E}_t$ only in specific cases. Obviously, it is valid whenever $\mathcal{E}_t$ is unitary during the whole evolution, i.e. $\mathcal{E}_t[\varrho] = e^{-i\mathbf{H}t}\varrho e^{i\mathbf{H}t}$ for all t. In general, the inverse transformation $\mathcal{E}_t^{-1}$ is ill defined, because its derivation from $\mathcal{U}_t^{-1}$ is based on the assumption that the initial state is correlated. On the other hand, we can derive the inverse transformation without any relevance to the system B. We can define $\mathcal{E}_t^{-1}$ by the relation $\mathcal{E}_t^{-1}\mathcal{E}_t = \mathcal{E}_t\mathcal{E}_t^{-1} = \mathcal{I}$. In terms of left-right formalism this corresponds to finding an inverse matrix.

In the orthonormal operator basis Θ_k the time evolution is described by the matrix $\mathcal{E}_t$ with time-dependent coefficients. For the specific choice of basis with $\Theta_0 = \frac{1}{\sqrt{d}}\mathbb{1}$ the first row of the matrix is fixed, i.e. $\mathcal{E}_{00}(t) = 1, \mathcal{E}_{0k}(t) = 0$ for all t. The question of the existence of the inverse matrix is equivalent to the property $\det\mathcal{E}_t = \det E_t \neq 0$. Due to the continuity of $\mathcal{E}_t$ the determinant is a smooth function of time and cannot vanish suddenly. Only for countable number of times t_k it can happen that $\det\mathcal{E}_{t_k} = 0$, but these points can be excluded from the evolution (set of zero measure). However, the main problem is not the reversibility of the matrix $\mathcal{E}$, but the physical reversibility of the evolution. Note, that matrices $\mathcal{E}_t$ can be mathematically understood as transformations of the selfadjoint operators, but if we restrict their action only to quantum states $\mathcal{S}(\mathcal{H})$, then the inverse transformation $\mathcal{E}_t^{-1}$ can easily transform density operators into unphysical quantum states.

In the Bloch sphere for a qubit ($d = 2$) the time evolution corresponds either to rotations (unitary transformations), or to compressions of the sphere combined with rotations and shifts. The whole space of selfadjoint operators with unit trace covers the whole three-dimensional real vector space. That is, the inverse transformation of $\mathcal{E}$ exists in most cases. On the other hand the inverse of compression is expansion and the expansion of Bloch sphere results outside the original boundaries of Bloch sphere. Therefore, inverse transformations lead to unphysical situations and they are "physically senseless". Only the rotations (unitary transformations) can be reversed. As a result we get that only for unitary $\mathcal{E}_t$ their restriction to the set of states $\mathcal{S}(\mathcal{H})$ can be reversed. Consequently, the general time evolution of open systems is irreversible. To make it reversible one needs to use the environment as well and make the inverse operation $\mathcal{U}_t^{-1}$ on a larger system.

Likewise the classical case the irreversibility reflects our lack of knowledge about the details of the evolution. Irreversibility will disappear if we consider also the environment in our description, or if we define the inverse transformations $\mathcal{E}^{-1}$ only on the subsets $\mathcal{S}_\mathcal{E} = \mathcal{E}[\mathcal{S}(\mathcal{H})]$ of the state space $\mathcal{S}(\mathcal{H})$. Then the evolution will be invertible, whenever the inverse matrix exists. Specific quantum operations $\mathcal{E}$ for which the inverse matrices $\mathcal{E}^{-1}$ do not exist are those that map the whole state space $\mathcal{S}(\mathcal{H})$ into a single fixed point, i.e. $\mathcal{E} : \varrho \mapsto \xi$ for all ϱ.

For the open systems we often use the concept of a dynamical semigroup, which reflects the notion of irreversibility in the following sense. Instead of a group structure of the time evolution the set of superoperators $\mathcal{E}_t$ possesses a semigroup structure. That is, the inverse of $\mathcal{E}_t$ may not be contained in this set. This requirement relaxes the second property of one-parametric group and defines the one-parametric semigroup

1. $\lim_{t\to 0+} \mathcal{E}_t = \mathcal{I}$
2. $\mathcal{E}_t\mathcal{E}_s = \mathcal{E}_{t+s}$ for all $t, s \geq 0$

It can be shown that general equations of motion generating this type of dynamics has the *Lindblad form* [35]:

$$\dot{\varrho}_t = \mathcal{G}[\varrho_t] = i[\mathbf{H}, \varrho_t] + \sum_{\alpha,\beta} c_{\alpha\beta} \left([\Theta_\alpha, \varrho_t\Theta_\beta^\dagger] + [\Theta_\alpha\varrho_t, \Theta_\beta^\dagger]\right) \tag{4.29}$$

where $\Theta_\alpha = \Theta_\alpha^\dagger,$ $\mathrm{Tr}\Theta_\alpha = 0$, $\mathrm{Tr}\Theta_\alpha\Theta_\beta = \delta_{\alpha\beta}$ for $\alpha, \beta = 1, \ldots, d^2 - 1$ and the coefficients $c_{\alpha\beta}$ form positive hermitian matrix.

Differentiating the time evolution represented in the left-right form, i.e. $\varrho_t = \mathcal{E}_t[\varrho_0]$, we formally obtain the following left-right expression for the generators $\mathcal{G}$

$$\dot{\varrho}_t = \dot{\mathcal{E}}_t[\varrho_0] = \dot{\mathcal{E}}_t\mathcal{E}_t^{-1}[\varrho_t] \quad \Rightarrow \quad \mathcal{G} = \dot{\mathcal{E}}_t\mathcal{E}_t^{-1} \tag{4.30}$$

In the specific (but standard) choice of basis ($\Theta_0 = \mathbb{1}$ and for $k = 1, \ldots, d^2 - 1$ the operators Θ_k possess the properties mentioned above, i.e. they are traceless, selfadjoint and form an orthonormal basis of the operator space) these generators take the form

$$\mathcal{G} = \begin{pmatrix} 0 & \vec{0}^T \\ \vec{g} & G \end{pmatrix} \tag{4.31}$$

In what follows we will discuss the case of qubit in detail.

4.3.1 Qubit

The qubit is the most simple example of a quantum system when the dimension of the corresponding Hilbert space $\mathcal{H}$ is two, i.e. $\dim(\mathcal{H}) = 2$. The vector space of all hermitian operators $A : \mathcal{H} \to \mathcal{H}$ is four-dimensional $N^2 = 4$ and every basis has four elements $\sigma_0, \sigma_1, \sigma_2, \sigma_3$. According to the previous section instead of considering an arbitrary basis we will use a very specific one. The first element of the basis is the identity operator $\sigma_0 = \mathbb{1}$ and $\sigma_j, j = 1, 2, 3$ are Pauli operators that are traceless, selfadjoint and mutually orthogonal, i.e. $\mathrm{Tr}\sigma_j\sigma_k = 2\delta_{jk}$ (for all $j, k = 0, 1, 2, 3$). In order to normalize the Pauli operators properly, we must put $\Theta_j = \sigma_j/\sqrt{2}$ $(j = 0, 1, 2, 3)$. However, the Pauli operators are very well known and extensively used. Therefore, we shall use them too. If we redefine the operator scalar product

$$(A|B) = \frac{1}{2}\mathrm{Tr}A^\dagger B \tag{4.32}$$

then the general state of the qubit takes the form

$$\varrho = \frac{1}{2}\mathbb{1} + \sum_j \alpha_j\sigma_j \leftrightarrow \vec{v}_\varrho = \begin{pmatrix} 1/2 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \tag{4.33}$$

with $\alpha_j = \frac{1}{2}\mathrm{Tr}\varrho\sigma_j$ and everything above remains valid.

Consequently, any evolution of qubits, which has the properties of a dynamical semigroup, can be written as

$$\dot{\varrho}_t = -i[\mathbf{H}, \varrho_t] + \frac{1}{2}\sum_{j,k=1}^{3} c_{jk}\left([\sigma_j, \varrho_t\sigma_k] + [\sigma_j\varrho_t, \sigma_k]\right) \tag{4.34}$$

where $\mathbf{H} = \sum_{j=1}^{3} h_j\sigma_j$ and h_j are real parameters. Moreover, the hermitian matrix c_{jk} can be rewritten as $c_{jk} = d_{jk} - ie_{jk}$, where $d_{jk} = \frac{1}{2}(c_{jk}+c_{kj})$ is a real symmetric matrix and $e_{jk} = i\frac{1}{2}(c_{jk}-c_{kj})$ is a real antisymmetric matrix. The differential equation Eq.(4.34) takes the form

$$\begin{pmatrix} 0 \\ \dot{\alpha}_1 \\ \dot{\alpha}_2 \\ \dot{\alpha}_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix} \begin{pmatrix} 1/2 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \tag{4.35}$$

where the coefficients g_{jk} of the generator $\mathcal{G}$ can be calculated directly from the Eq.(4.34) and can be expressed in terms of the parameters c_{jk} and h_j as follows

$$g_{jk} = (\sigma_j|\mathcal{G}|\sigma_k) = \frac{1}{2}\mathrm{Tr}\sigma_j\mathcal{G}[\sigma_k] \tag{4.36}$$

which reads

$$g_{jk} = 2\sum_l \epsilon_{jkl}h_l + \frac{1}{2}(c_{kj} + c_{jk}) - \sum_l c_{ll}\delta_{jk} \tag{4.37}$$

$$g_{k0} = 2i\sum_{jl} \epsilon_{jlk}c_{\alpha\beta} \tag{4.38}$$

More explicitly, the matrix $\mathcal{G}$ can be expressed using d's and e's

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 4e_{23} & -2d_{22} - 2d_{33} & 2d_{12} - 2h_3 & 2d_{13} + 2h_2 \\ 4e_{31} & 2d_{12} + 2h_3 & -2d_{11} - 2d_{33} & 2d_{23} - 2h_1 \\ 4e_{12} & 2d_{31} - 2h_2 & 2d_{32} + 2h_1 & -2d_{11} - 2d_{22} \end{pmatrix}$$

One can derive also the inverse relations. If the generator $\mathcal{G}$ is given in the left-right form then the coefficients d's, e's and h's can be expressed as functions of g's

$$h_1 = \frac{g_{32}-g_{23}}{4} \quad h_2 = \frac{g_{13}-g_{31}}{4} \quad h_3 = \frac{g_{21}-g_{12}}{4}$$

$$e_{23} = \frac{g_{10}}{4} \quad e_{31} = \frac{g_{20}}{4} \quad e_{12} = \frac{g_{30}}{4}$$

$$\begin{aligned} d_{11} &= \frac{-g_{22}-g_{33}+g_{11}}{4} & d_{12} &= \frac{g_{12}+g_{21}}{4} \\ d_{22} &= \frac{-g_{11}-g_{33}+g_{22}}{4} & d_{23} &= \frac{g_{23}+g_{32}}{4} \\ d_{33} &= \frac{-g_{11}-g_{22}+g_{33}}{4} & d_{13} &= \frac{g_{13}+g_{31}}{4} \end{aligned}$$

For a qubit we have derived the relation between two different expressions for generators. Using this relation we can switch between these two expressions at any time. It is important to note here that this relation enables us to exhibit not only dynamical semigroups, but also time evolution of qubit in the Lindbland-like form. In this general case the coefficients $c_{\alpha\beta}$ will not be constant, but will depend on time. In this way we can derive more general master equations describing not only Markovian dynamics.

4.3.2 Example

Consider the controlled dynamics, where qubit is coupled to a harmonic oscillator and the Hamiltonian evolution is given by the following rule

$$\mathbf{U}_t = \sum_{j=0}^{\infty} e^{i\omega_j t\sigma_z} \otimes |j\rangle\langle j| \tag{4.39}$$

Then the evolution $\mathcal{E}_t$ of the qubit in the left-right form is given as

$$\mathcal{E}_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & f(t) & g(t) & 0 \\ 0 & -g(t) & f(t) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.40}$$

where $f(t) = \sum_j |\alpha_j|^2 \cos 2\omega_j t$, $g(t) = \sum_j |\alpha_j|^2 \sin 2\omega_j t$ and the initial state of the harmonic oscillator is $|\Xi\rangle = \sum_j \alpha_j |j\rangle$. Calculating the generator $\mathcal{G}$ of this dynamics we can use the derived formula $\mathcal{G} = \dot{\mathcal{E}}_t \mathcal{E}_t^{-1}$ to obtain

$$\mathcal{G} = \frac{1}{f^2+g^2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & f\dot{f}+g\dot{g} & -\dot{f}g+\dot{g}f & 0 \\ 0 & +\dot{f}g-\dot{g}f & f\dot{f}+g\dot{g} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.41}$$

and consequently, the master equation reads

$$\dot{\varrho}_t = i\frac{\dot{g}f-\dot{f}g}{2(f^2+g^2)}[\sigma_z,\varrho_t] - \frac{f\dot{f}+g\dot{g}}{2(f^2+g^2)}(\sigma_z\varrho_t\sigma_z - \varrho_t) \tag{4.42}$$

4.4 Quantum homogenization process

Suppose, we have a qubit in an unknown state ϱ_S and many qubits prepared in a state ξ. Let the number of them be N and call them *reservoir R*. We shall call the single qubit in the state ϱ_S a *system qubit*. The system qubit evolution is driven by its interactions with the qubits from the reservoir and is described by the unitary transformation $\mathbf{U}_{SR}$. Suppose the mutual interaction is composed of bipartite qubit-qubit interactions. It means, the system qubit interacts always only with a single qubit from the reservoir. Thus, the interaction has the form $\mathbf{U}_{SR} = \mathbf{U}_{Sj} \otimes \mathbb{1}_{R\backslash j}$, where j denotes the $j-$th qubit from the reservoir and $R\backslash j$ is the reservoir without the j-th qubit. If the total number of interactions between the chosen qubit of the reservoir and the system qubit one at most, then the evolution of the system is described by the superoperator $\mathcal{E}_\xi^k$, where k is the number of realized two-qubit interactions and $\mathcal{E}_\xi$ is the superoperator induced by two-qubit unitary transformation $\mathbf{U}_{Sj} \in \mathcal{U}(\mathcal{H}_S \otimes \mathcal{H}_j)$. It means

$$\mathcal{E}_\xi^k[\varrho_S] = \varrho_S^{(j)} = \mathrm{Tr}_j \mathbf{U}_{Sk} \varrho_S^{(k-1)} \otimes \xi_k \mathbf{U}_{Sk}^\dagger \tag{4.43}$$

where $\varrho_S^{(k-1)}$ is the state of the system qubit after the $(k-1)$-th interaction. Note, that a *collision model* always stands behind the powers of a map $\mathcal{E}^k$. To be able to realize powers of any superoperator $\mathcal{E}$ we need to have the ancilla prepared in a factorized state. The unitary transformation (describing the interaction process) must have the form of sequentially applied unitaries between a single object from the ancilla and the evolving system (see Fig.3.1).

Let us now formulate the homogenization problem: After the evolution we want the system qubit to be described by the "same" state as the original reservoir qubits. Moreover, we do not want to change the states of the qubits in reservoir. Mathematically, we are interested in transformation

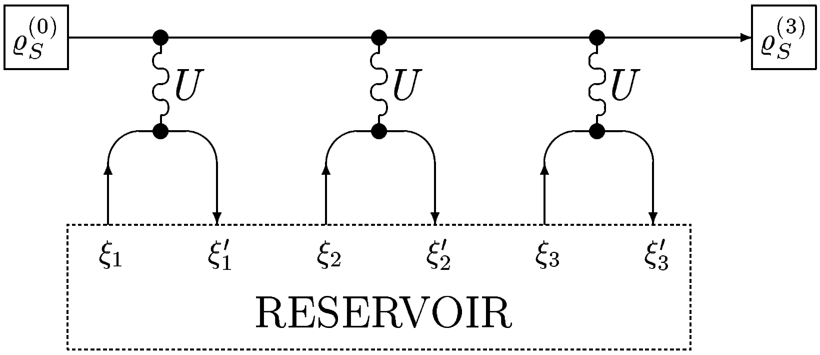

Figure 4.1: Collision model of quantum homogenization.

$\varrho\otimes\xi\otimes\ldots\otimes\xi \to \xi\otimes\xi\otimes\ldots\otimes\xi$. But due to the well known fact of the impossibility of quantum cloning [31] such transformation is forbidden. Thus, we shall formulate the problem in a different way. The mentioned sequence of two-qubit unitary transformations induces the superoperator $\mathcal{E}^k_\xi[\varrho_S]$. After this transformation we demand that the system is transformed into the state ξ_S, but simultaneously the states of reservoir qubits change only little. It means that we will take into account some inaccuracy. Let us characterize our approximation by a real parameter $\delta > 0$. Then by the *homogenization* we mean the following

$$D(\mathcal{E}^k_\xi[\varrho],\xi) < \delta \qquad \text{and} \qquad \forall j \;\; D(\xi'_k,\xi) < \delta \tag{4.44}$$

where $D(.,.)$ is a distance function and $\xi'_k = \mathrm{Tr}_S[\mathbf{U}_{Sk}\varrho_S \otimes \xi \mathbf{U}^\dagger_{Sk}]$ is the state of k-th qubit from the reservoir after its interaction with the system qubit. Note that qubits in the reservoir are labeled by its order in the sequence of interactions. The unitary transformation $\mathbf{U}_{SR}$ of the whole system takes the form $\mathbf{U}_{SR} = \mathbf{U}_N \ldots \mathbf{U}_1$ with $\mathbf{U}_k = \mathbf{U}_{Sk} \otimes \mathbb{1}_{R\setminus k}$. Our aim is to find the unitary transformations that lead the system to become "homogenized" with the reservoir.

We assume that the homogenization is independent of the initial state of the system qubit (ϱ) as well of the initial state of qubits in reservoir (ξ). Moreover, if the system qubit is in the same state as the reservoir qubits, then the transformation $\mathbf{U}$ acts as follows

$$\mathrm{Tr}_S(\mathbf{U}\xi \otimes \xi \mathbf{U}) = \xi \tag{4.45}$$
$$\mathrm{Tr}_j(\mathbf{U}\xi \otimes \xi \mathbf{U}) = \xi \tag{4.46}$$

Let us first discuss the case of pure states. If ξ represents a pure state then the condition (4.45) says that $\mathbf{U}\xi\otimes\xi\mathbf{U}^\dagger = \xi\otimes\xi_1$, where ξ_1 needs to be determined. However from the second condition (4.46) it follows that $\mathbf{U}\xi\otimes\xi\mathbf{U}^\dagger = \xi_2\otimes\xi$ where ξ_2 is unknown. Putting the last two results together we obtain that

$$\mathbf{U}\xi\otimes\xi\mathbf{U}^\dagger = \xi\otimes\xi, \tag{4.47}$$

for any ξ representing a pure state. From here it follows that the unitary transformation $\mathbf{U}$ acting on the joint Hilbert space $\mathcal{H}\otimes\mathcal{H}$ must be of the form

$$\mathbf{U} : |\psi\rangle\otimes|\psi\rangle \to e^{i\varphi}|\psi\rangle\otimes|\psi\rangle, \tag{4.48}$$

where the parameter φ is independent of the state $|\psi\rangle$. Therefore, the action of the unitary transformation on the symmetric subspace of $\mathcal{H}\otimes\mathcal{H}$ is fixed up to a phase factor $e^{i\varphi}$. Neither of the two conditions (4.45) and (4.46) nor the condition (4.48) tells us anything about the action of the unitary transformation $\mathbf{U}$ on the antisymmetric subspace of $\mathcal{H} \otimes \mathcal{H}$. This means that the action of $\mathbf{U}$ on the antisymmetric subspace is arbitrary. However, in the case of qubits the antisymmetric subspace is one-dimensional, and we can proceed further. Because the antisymmetric subspace is one-dimensional and invariant under the action of the unitary transformation $\mathbf{U}$, we have

$$\mathbf{U}(|\psi\rangle|\psi^\perp\rangle - |\psi^\perp\rangle|\psi\rangle) = e^{i\theta}\left(|\psi\rangle|\psi^\perp\rangle - |\psi^\perp\rangle|\psi\rangle\right), \tag{4.49}$$

where θ is a constant depending on $\mathbf{U}$. Now the transformation $\mathbf{U}$ is given by the equations (4.48) and (4.49) up to two constants φ and θ. If we define the unitary operator $\mathbf{U}'$ to be

$$\mathbf{U}' = \exp^{i(-\theta-\varphi)/2}\mathbf{U},$$

then equations (4.48) and (4.49) give us

$$\begin{aligned} |\psi\rangle|\psi\rangle &\overset{\mathbf{U}'}{\rightarrow} e^{i(\varphi-\theta)/2}|\psi\rangle|\psi\rangle \\ |\psi\rangle\psi^\perp\rangle - |\psi^\perp\rangle|\psi\rangle) &\overset{\mathbf{U}'}{\rightarrow} e^{i(\theta-\varphi)/2}\left(|\psi\rangle|\psi^\perp\rangle - |\psi^\perp\rangle|\psi\rangle\right). \end{aligned}$$

In a more convenient way we can express this unitary transformation $\mathbf{U}'$ with the help of the **swap operation**

$$\mathbf{S} = |00\rangle\langle 00| + |11\rangle\langle 11| + |01\rangle\langle 10| + |10\rangle\langle 01|. \tag{4.50}$$

Applying it on factorized operators we get

$$\mathbf{S}\varrho \otimes \xi \mathbf{S}^\dagger = \xi \otimes \varrho. \tag{4.51}$$

If we define $\eta := (\varphi - \theta)/2$, we can rewrite the unitary transformation $\mathbf{U}'$ in the form

$$\mathbf{U}' \equiv \mathbf{P}(\eta) = \cos\eta \mathbb{1} + i \sin\eta \mathbf{S}. \tag{4.52}$$

We shall call the unitary transformations $\mathbf{P}(\eta)$ as the *partial swap*. We can conclude that in the case of qubits, the partial swap is the *only* possible operator that (for certain values of η) satisfies the conditions of homogenization (4.45) and (4.46).

After a single interaction facilitated by the partial swap we find the system qubit and the first reservoir qubit in the states

$$\begin{aligned} \varrho_S^{(1)} &= \mathcal{E}_\xi[\varrho_S^{(0)}] = c^2\varrho_S^{(0)} + s^2\xi + ics[\xi, \varrho_S^{(0)}] \qquad (4.53) \\ \xi_1' &= s^2\varrho_S^{(0)} + c^2\xi + ics[\varrho_S^{(0)}, \xi], \qquad (4.54) \end{aligned}$$

where we put $c := \cos\eta$ and $s := \sin\eta$. Applying the sequence of k partial swap operations we obtain the reccurence relation

$$\begin{aligned} \varrho_S^{(k)} &= c^2\varrho_S^{(k-1)} + s^2\xi + ics[\xi, \varrho_S^{(k-1)}] \qquad (4.55) \\ \xi_k' &= s^2\varrho_S^{(k-1)} + c^2\xi + ics[\varrho_S^{(k-1)}, \xi]. \qquad (4.56) \end{aligned}$$

Denote $\varrho \equiv \vec{n}$ and $\xi \equiv \vec{t}$. Since in the left-right form the partial swap induces the map $\vec{n} \rightarrow \vec{n}' = c^2\vec{n} + s^2\vec{t} - 2cs\vec{t} \times \vec{n}$, the superoperator $\mathcal{E}_\xi$ can be represented by the following matrix

$$\mathcal{E}_\xi = \mathcal{E}_{\vec{t}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2s^2t_x & c^2 & 2cst_z & -2cst_y \\ 2s^2t_y & -2cst_z & c^2 & 2cst_x \\ 2s^2t_z & 2cst_y & -2cst_x & c^2 \end{pmatrix}. \tag{4.57}$$

Hence $\mathcal{E}_{\vec{t}} = (s^2\vec{t}, T)$. Let us calculate the modification of the state distance caused by the evolution $\mathcal{E}_\xi$

$$\begin{aligned} D^2(\mathcal{E}_\xi[\varrho], \mathcal{E}_\xi[\tau]) &= 4|T\vec{r}|^2 = 4(c^4|\vec{r}|^2 + 4s^2c^2|\vec{t}\times\vec{r}|) \\ &= 4|\vec{r}|^2c^2(c^2 + 4s^2|\vec{t}|^2\cos\beta) \\ &\leq 4c^2|\vec{r}|^2 = c^2D^2(\varrho,\tau) \end{aligned} \tag{4.58}$$

where $\beta \leq \pi$ is the angle between vectors $\vec{t}$ and $\vec{r}$. The inequality holds due to the facts that $|\vec{t}|^2 \leq 1/4$ and $c^2 + s^2\cos\beta \leq 1$. Since we have shown that $D(\mathcal{E}_\xi[\varrho], \mathcal{E}_\xi[\tau]) \leq cD(\varrho, \tau)$ and $c < 1$ (except the case of $\eta = 0$), it follows that $\mathcal{E}_\xi$ is a *contractive superoperator*. We need to find the fixed point ϱ_f of our superoperator. The condition $s^2\vec{t} + c^2\vec{f} - 2cs\vec{t}\times\vec{f} = \vec{f}$ is fulfilled for $\vec{f} = \vec{t}$. That is, the fixed point of superoperator $\mathcal{E}_\xi$ is the initial state of the reservoir qubits ξ and the convergence of the state of system qubit to the state ξ is assured by the *Banach theorem*.

Of course, not every partial swap operation satisfies the original conditions of homogenization (4.44) with the defined approximation δ. We have ensured the convergence of the system qubit, but we have not considered the states of reservoir qubits yet. Taking into account the success (parametrized by δ) of the homogenization process, we need to adjust the parameters η of the partial swap operation and the number of required interactions N_δ (see [52]).

The discussed process of homogenization describes a physical realization of contractive maps for each quantum state ξ. We showed the uniqueness of the partial swap family of unitary transformations. Namely, the partial swap is the only map satisfying the *trivial* homogenization (4.47). In this section we have defined the family of contractive superoperators $\mathcal{E}_\xi$ realizable by the same interaction.

We remind us that *amplitude damping superoperator* $\mathcal{A}_p(\psi)$ describes a very similar evolution. It is also contractive with pure states $|\psi\rangle$ being fixed points. But the unitary transformation $\mathbf{U}$ leading to the superoperator $\mathcal{A}_p(\psi)$ depends on this pure fixed state $|\psi\rangle$, unlike the homogenization transformation, which is state (ξ) independent.

4.4.1 Quantum homogenization as a continuous process

Obviously, the dynamics of the homogenization process is sequential. It consists of subsequent collisions. Anyway, after each collision the state transformation is given by the CPT map $\mathcal{E}_\xi^k$. The set of all mappings $\mathcal{E}_k := \mathcal{E}_\xi^k$ (the whole homogenization process) forms a one-parameter discrete semigroup, if the identity $\mathcal{I} = \mathcal{E}_0$ is added. It is an open question whether the continuous time parameter can be introduced in a way that the discretized one-parameter semigroup becomes continuous. In fact any collision model (arbitrary interaction) determines certain discretized one-parameter semigroup of CPT maps, but not all of them can be made smooth. For instance, the pure swap ($\eta = \pi/2$) generates a semigroup consisting only of two maps, $\mathcal{I}$ and $\mathcal{E}_\xi$, because $\mathcal{E}_\xi^k = \mathcal{I}$ for $k > 1$. Therefore, under the conditions that the discrete semigroup contains large number of superoperators and, moreover, the first of them is “close” to trivial map $\mathcal{I}$, it could be possible to introduce a time parameter t and define a continuous dynamical semigroup. This smoothed version of the semigroup coincides with the former one in the time points $t = k\tau$, where τ is the time of the duration of collision.

Note that the derivation of the dynamical semigroup is always tricky and requires approximations, introduction of the coarse-grained time scale, etc. It reflects the fact that semigroup description of the evolution is not “valid” in details, but it is reasonably exact. The descriptions of exponential decay, or spontaneous photon emission, or Pauli master equation are the elementary examples of the usage of semigroups. The semigroup approximation is generally valid in cases of weak interactions between the system and the reservoir, when the reservoir rapidly forgets and seems to be memoryless. The derivation of the master equations is usually done in the so-called Markovian approximation, but we are not going into details of this large subject of investigation.

Let us get back to the problem of homogenization with a very small value of the parameter η, i.e. the interaction is very weak. The figure shows the simulations of the homogenization process in three different situations of the reservoir state, but always diagonal in the σ_z basis. The initial state

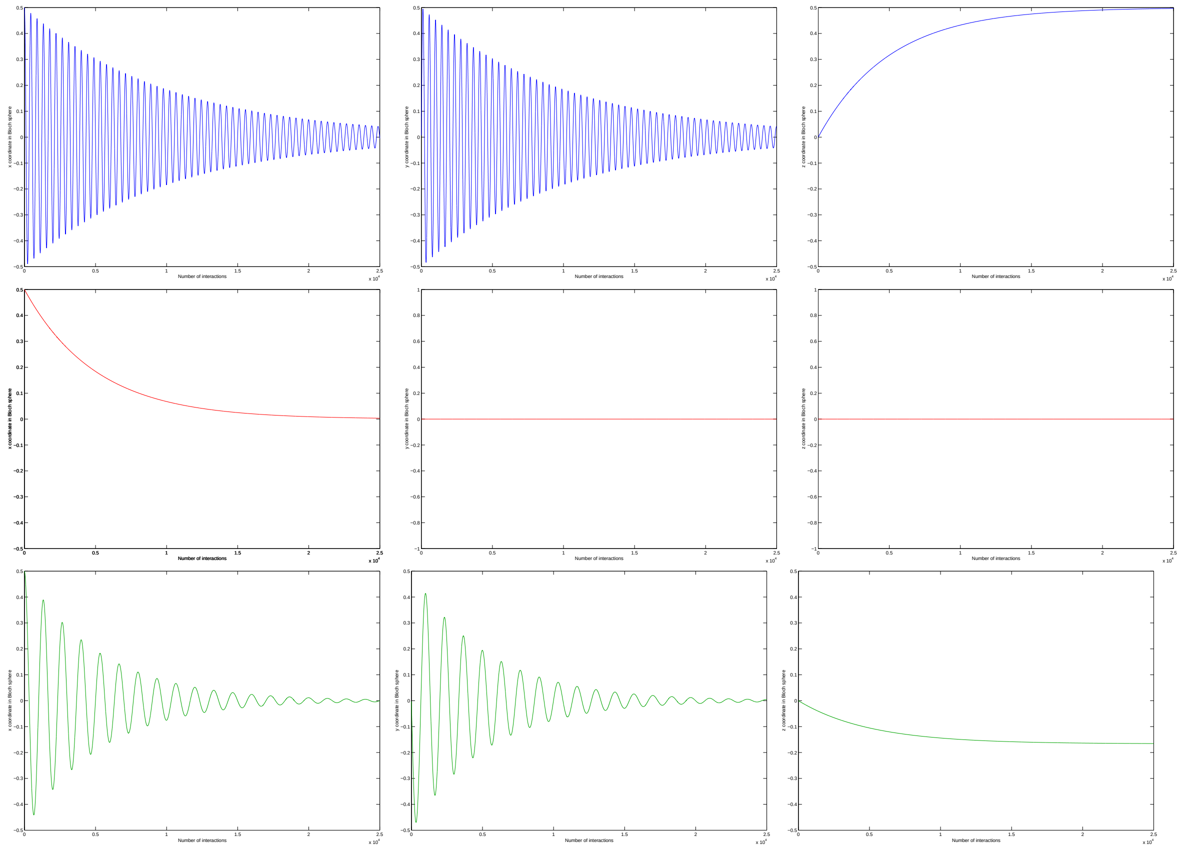

Figure 4.2: Homogenization process with 25000 qubits in reservoir and the parameter η adjusted to the value 0.001. Initial state of the system qubit is fixed to be $|x\rangle = |0\rangle + |1\rangle$. The top row corresponds to the case of reservoir with qubits in state $\xi = |0\rangle$. The middle row represents the case of $\xi = \frac{1}{2}\mathbb{1}$ and the bottom row corresponds to $\xi = \frac{1}{3}|0\rangle\langle 0| + \frac{2}{3}1\rangle\langle 1|$. The columns represent the x,y,z axis, respectively, in the Bloch sphere picture.

of the system qubit is fixed to be $|x\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ for all the cases. The columns represent the x, y, z axis in the Bloch sphere picture and rows correspond to different choices of the reservoir state. Qualitatively the behavior can be understood in the following way. The frequency of the damped oscillations around the x and y axis depend on the initial difference of the z components of the states ϱ_S and ξ. Note that the differences in x and y components are fixed for all studied examples. On the other hand the damping itself seems to be independent of this difference. The question was, whether the system state exponentially tends to the state of the reservoir, but as we have shown it is not the case.

The general homogenization map $\mathcal{E}_\xi$ (4.57) can be always transformed into the simplified form

$$\mathcal{E}_\xi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c^2 & 2csw & 0 \\ 0 & -2csw & c^2 & 0 \\ 2s^2w & 0 & 0 & c^2 \end{pmatrix}. \tag{4.59}$$

where the chosen operator basis is $\mathbf{S}_k$ in which $\xi = \frac{1}{2}\mathbb{1} + w\mathbf{S}_3$. Note that the partial swap has the same form in all bases. The nth power of the evolution map $\mathcal{E}_\xi^n$ can be written as

$$\mathcal{E}_\xi^n = \begin{pmatrix} 1 & 0\ 0 & 0 \\ 0 & c^n \mathbf{A}^n & 0 \\ 2w(1-c^{2n}) & 0\ 0 & c^{2n} \end{pmatrix} \tag{4.60}$$

where $\mathbf{A}$ is a squared 2x2 matrix. The above form helps us in the qualitative analysis of the homogenization process. We see that there is always a possibility to choose coordinates of the Bloch sphere so that in one of them the process is exponential $r_3(n) = (1-c^{2n})w + c^{2n} r_3$. One can introduce the continuous time by putting $n = t/\tau$ with τ representing the duration of the single collision. Using the standard algebraic relations we can write $r_3(t) = (1-e^{2t\ln c/\tau})w + e^{2t\ln c/\tau} r_3$, where $\ln c < 0$, because $c < 1$, and therefore the parameter r_3 tends to a fixed value w exponentially as e^{-t/T_1} with the rate $T_1 = \frac{\tau \ln(1/c)}{2}$. The other two state parameters, r_1 and r_2, evolve according to the map $c^n \mathbf{A}^n$. The following two relations $r_1^2(n) + r_2^2(n) = (r_1^2 + r_2^2)(1 - s^2(1-4w^2))^n c^n$ and $0 \leq s^2(1-4w^2) < 1$ imply that the norm of the vector $(r_1(n), r_2(n))$ vanishes, and consequently so do the coefficients $r_1(n) \to 0$ and $r_2(n) \to 0$, too. The transformation of the initial vector (r_1, r_2) into the zero vector can be considered as exponential $e^{-t/T_2}(r_1, r_2)$ where the time is introduced in the same way as before, i.e. $n = t/\tau$ but for a new relaxation time $T_2 = \tau \ln[1/c(1 - s^2(1-4w^2))]$. From the physical point of view these two times are associated with the "decay relaxation rate" (T_1) and with the "decoherence relaxation rate" (T_2). Both the relaxation times are in the following relation

$$\frac{T_1}{T_2} = \frac{\ln(1/c)}{2\ln[1/c(1 - s^2(1-4w^2))]} \geq \frac{1}{2} \tag{4.61}$$

where the last inequality is in accordance with the general formula that states $T_1 \geq \frac{1}{2}T_2$. The equality is saturated for the value $w = \pm 1/2$, i.e. when the initial state ξ of the reservoir qubits is pure. As a result we have obtained that the diagonal elements of the system qubit density matrix (associated with r_3) expressed in the basis $\mathbf{S}_k$ achieve a fixed value faster than the off-diagonal elements (associated with r_1, r_2) manage to vanish (see the pictures).

The oscillatory behavior is hidden in the form of the matrix $\mathbf{A}$. One can easily find out the explicit time evolution of the system state parametrized by vector $\vec{r} = (r_1, r_2, r_3)$

$$\begin{aligned} r_1(t) &= Re^{-t/T_2}\cos(\Omega t + \phi) \\ r_2(t) &= Re^{-t/T_2}\sin(\Omega t + \phi) \\ r_3(t) &= r_3(0)e^{-t/T_1} + w(1 - e^{-t/T_1}) \end{aligned} \tag{4.62}$$

where the parameters T_1, T_2 have been already determined, $R = \sqrt{r_1^2(0) + r_2^2(0)}$ and $\phi = \arctan(r_1(0)/r_2(0))$. The coordinates r_1 and r_2 oscillate with the same frequency Ω, because we have shown that the evolution of the length of the vector (R) is exponential. The frequency Ω could depend on the initial states of the reservoir as well as of the system qubit. In particular, the parameters of the matrix $\mathbf{A}$ can be rewritten into the form

$$\mathbf{A} = \sqrt{c^2 + 4s^2w^2} \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix} \tag{4.63}$$

with $\omega = \arctan(2ws/c)$. In the derivation of the above form we have used the identity $\cos\arctan x = (1+x^2)^{-1/2}$. Having such expression of the matrix $\mathbf{A}$ one can easily evaluate its powers and the continuous time $n = t/\tau$ can be introduced. Obviously,

$$\mathbf{A}^n = (c^2 + 4s^2w^2)^{n/2} \begin{pmatrix} \cos(\omega n) & \sin(\omega n) \\ -\sin(\omega n) & \cos(\omega n) \end{pmatrix} \to \mathbf{A}_t = (c^2 + 4s^2w^2)^{t/2\tau} \begin{pmatrix} \cos(\Omega t) & \sin(\Omega t) \\ -\sin(\Omega t) & \cos(\Omega t) \end{pmatrix} \tag{4.64}$$

As a result we have found that the frequency $\Omega = \omega/\tau = \frac{1}{\tau}\arctan(2ws/c)$ does not depend on the initial state of the system qubit. For values of $\eta \to \pi/2$ ($s/c \to \infty$) this frequency tends to infinity.

The explicit form of time-continuous quantum homogenization reads

$$\mathcal{E}_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-t/T_2}\cos(\Omega t) & e^{-t/T_2}\sin(\Omega t) & 0 \\ 0 & -e^{-t/T_2}\sin(\Omega t) & e^{-t/T_2}\cos(\Omega t) & 0 \\ 2w(1-e^{-t/T_1}) & 0 & 0 & e^{-t/T_1} \end{pmatrix} \tag{4.65}$$

Using the derived formulas, we can write explicitly the master equation that drives the homogenization process. In particular, the generator takes the form

$$\mathcal{G} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\Omega/T_2 & 0 \\ 0 & \Omega/T_2 & 0 & 0 \\ 2w/T_1 & 0 & 0 & -1/T_1 \end{pmatrix} \tag{4.66}$$

where

$$\Omega = \arctan(2ws/c)/\tau \tag{4.67}$$

$$T_1 = \tau \ln(1/c)/2 \tag{4.68}$$

$$T_2 = \tau \ln[1/(c^2+4s^2w^2)] \tag{4.69}$$

The main point is that the generator $\mathcal{G}$ is independent of time t and therefore it can generate the evolution governed by the Lindblad master equation.

4.5 Quantum generalized measurements

Except the mentioned form of dynamics, quantum theory contains also a kind of state transformations that cannot be described in a unitary way, and, moreover, they are in deep contrast with the determinism of quantum evolution. In this section we shall briefly mention the description of the measurement process in quantum theory. There are two possibilities how to understand any state transformation occurring during the measurement

1. *Ensemble state transformation.* On one side to make the measurement valuable it has to be repeated (in principle) infinitely many times, i.e. on huge number of equally prepared physical systems. One can consider that measurement transforms the ensemble of particles initially described by a state ϱ_{in} into a new state ϱ_{out}. The information gain of this process can be associated with the change of the description of this ensemble. For instance, after performing the measurement associated with σ_z the ensemble of particles initially described by a pure state $|\psi\rangle = a|0\rangle + b|1\rangle$ will be in the mixture $\varrho_{out} = |a|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|$. Note that states $|0\rangle, |1\rangle$ are eigenstates of σ_z. If we use the concept of purity of quantum states to exhibit the quality of our description, then the measurements understood in this way raise impurities in our knowledge.

2. *Individual system state transformation.* On the other hand, we can make the so-called *post-selection*, i.e. we can divide the final ensemble of particles according to the observed outcomes. In this way we transform the original ensemble into smaller ones and the measurement process is described by one-to-many stochastic map. The size of these subensembles is given by the probability of finding the associated outcomes. The description of the state transformation of an individual particle depends on the chosen model of measurement and does not follow from the basic principles of quantum theory. However, the usage of the projection postulate enables us to speak also about the states of individual particles. Using the same example as before, the original state $|\psi\rangle$ evolves either into the state $|0\rangle$, or $|1\rangle$.

So far we have not mentioned the mathematical description of the most general observation one can perform in quantum theory. Without going into details, let us define the *positive operator valued* *measure* (POVM) by the set of positive operators $\mathbf{X} = \{\mathbf{X}_b\}$, such that $\sum_b \mathbf{X}_b = \mathbb{1}$. The real numbers b label the outcomes of the measurement $\mathbf{X}$. The relation between the abstract mathematical objects and experimental reality is given by the following probability rule. The outcome b occurs with a probability $P(b, \varrho) := \mathrm{Tr}\mathbf{X}_b\varrho$. Unlike the case of *orthogonal* measurements, when the operators $\mathbf{X}_b$ were orthogonal (mutually commuting) projections, for the general measurement we cannot use the projection postulate to describe the state of the system after the observation. The *Neumark theorem* connects the generalized measurements (POVM) with the orthogonal measurements in the following sense. For any POV measure defined by $\mathbf{X}_b : \mathcal{H} \to \mathcal{H}$ there exists a Hilbert space $\overline{\mathcal{H}}$ ($\mathcal{H} \subset \overline{\mathcal{H}}$) and projection valued measure $\mathbf{E}_b : \overline{\mathcal{H}} \to \overline{\mathcal{H}}$ such that for all b and for all $|\psi\rangle \in \mathcal{H}$

$$\mathbf{X}_b|\psi\rangle = \Pi[\mathbf{E}_b]|\psi\rangle \tag{4.70}$$

where $\Pi : \overline{\mathcal{H}} \to \mathcal{H}$ is the orthogonal projection of $\overline{\mathcal{H}}$ onto $\mathcal{H}$. That is, any POVM can be represented by an orthogonal measurement in a larger Hilbert space $\overline{\mathcal{H}}$ and $\mathbf{X}_b = \Pi[\mathbf{E}_b]$, but such representation is only formal and non-unique. Therefore, such reduction of POV measures to orthogonal measurements does not diminish the need for POV measures in the description of physical systems. This relation between POV measures and projective measurements allows us to speak about the state transformation of the measured system in the process of measurement. To different enlargements we will refer as to different realizations of the same POVM with different state transformations.

To summarize, the definition of POVM via positive operators $\mathbf{X}_b$ enables us to calculate the outcome probabilities p_b (b labels the result of measurement) by the well known trace rule $p_\varrho(b) = \mathrm{Tr}\varrho\mathbf{X}_b$, where ϱ is the state of the system before the measurement. However, the concept of POVM does not tell us anything about the transformation of the state caused by the observation. In fact, this transformation $\varrho \to \varrho_b$ depends on the particular realization of the whole measurement process. Only for the so-called projective measurements the *projection postulate* determines the canonical state transformation. If we use the picture supported by Neumark theorem, i.e. the POVM is a projective measurement on a larger Hilbert space, then we can use the projection postulate to determine the state transformation also for the case of POVMs. Note, that Eq.(4.70) does not mean that state of the system after the measurement is described by unnormalized state $\mathbf{X}_b|\psi\rangle$.

4.5.1 Von Neumann-Lüders measurements

We will widely use the von Neumann-Lüders implementations of POVMs. In this case the state transformation can be described by the set of operators $\mathbf{M}_b$, for which $\mathbf{M}_b^\dagger\mathbf{M}_b = \mathbf{X}_b$ and $\sum_b \mathbf{M}_b^\dagger\mathbf{M}_b = \mathbb{1}$. Then the state transformation is given by the well-known formula

$$\varrho \to \varrho_b = \frac{\mathbf{M}_b\varrho\mathbf{M}_b^\dagger}{\mathrm{Tr}\varrho\mathbf{M}_b^\dagger\mathbf{M}_b} \tag{4.71}$$

Note that the formal relation $\mathbf{M}_b = \sqrt{\mathbf{X}_b}$ does not determine the operator $\mathbf{M}_b$ uniquely. Any set of operators $\mathbf{M}_b' = \mathbf{U}_b\mathbf{M}_b$ represents the same POVM, since $(\mathbf{M}_b')^\dagger\mathbf{M}_b' = \mathbf{M}_b^\dagger\mathbf{U}_b^\dagger\mathbf{U}_b\mathbf{M}_b = \mathbf{M}_b^\dagger\mathbf{M}_b = \mathbf{X}_b$. Let us compare the purity (defined as $P(\varrho) = 1 - \mathrm{Tr}\varrho^2$) of the original state ϱ and the output state ϱ_b, i.e.

$$P(\varrho) = 1 - \mathrm{Tr}\varrho^2 \qquad \textit{versus} \qquad P(\varrho_b) = 1 - \mathrm{Tr}\varrho_b^2 = 1 - \frac{\mathrm{Tr}(\mathbf{X}_b\varrho)^2}{(\mathrm{Tr}\mathbf{X}_b\varrho)^2} \tag{4.72}$$

For pure state $\varrho = |\psi\rangle\langle\psi|$ we obtain that $P(\varrho_b) = 0 = P(\varrho)$. Note that in this case $\mathrm{Tr}(\mathbf{X}_b\varrho)^2 = |\langle\psi|\mathbf{X}_b|\psi\rangle|^2 = (\langle\psi|\mathbf{X}_b|\psi\rangle)^2 = (\mathrm{Tr}\mathbf{X}_b\varrho)^2$. Hence, the pure states remain pure (for any outcome b) after this type of measurements. In particular, von Neuman-Lüders measurements induce the vector state (possibly non-linear) transformation

$$|\psi\rangle \to \frac{\mathbf{M}_b|\psi\rangle}{\sqrt{\langle\psi|\mathbf{X}_b|\psi\rangle}} \tag{4.73}$$

But, let us note that not every measuring process is of this type (see section VI.6).

Von Neumann-Lüders measurement can be realized as a projective measurement in the following way. Let M denote the number of different outcomes, i.e. number of operators $\mathbf{X}_b$, or equivalently $\mathbf{M}_b$. Let us choose an ancillary system with the dimension equal to this number, i.e. $\dim \mathcal{H}' = M$, and let the vectors $|b\rangle$ $(b = 0, \ldots, M-1)$ form a basis of $\mathcal{H}'$. Define the unitary transformation by its action on product states $|\psi\rangle \otimes |0\rangle$

$$\mathbf{U}(|\psi\rangle \otimes |0\rangle) = \sum_b \mathbf{M}_b|\psi\rangle \otimes |b\rangle \tag{4.74}$$

Suppose we perform a projection $\mathbf{E}_b = \mathbb{1} \otimes |b\rangle\langle b|$, which is one-dimensional in the sense of ancilla system $\mathcal{H}'$. When an outcome b is observed then the projection postulate implies that the state of the joint system is

$$\frac{\mathbf{E}_b\mathbf{U}|\psi\rangle \otimes |0\rangle}{\sqrt{\langle 0|\langle\psi|\mathbf{U}^\dagger\mathbf{E}_b\mathbf{U}|\psi\rangle|0\rangle}} = \frac{\mathbf{M}_b|\psi\rangle \otimes |b\rangle}{\sqrt{\langle\psi|\mathbf{M}_b^\dagger\mathbf{M}_b|\psi\rangle}} \tag{4.75}$$

Tracing out the ancillary system we obtain the above measurement transformation (4.73).

If we are not interested in particular results of the measurement, but only in the state transformation of the whole ensemble caused by the process of measurement, then von Neumann-Lüders measurement mapping is described by a completely tracepreserving linear map (superoperator)

$$\varrho \mapsto \sum_b \sqrt{\mathbf{X}_b}\varrho\sqrt{\mathbf{X}_b} \tag{4.76}$$

It is easy to verify that operators $\mathbf{M}_b = \sqrt{\mathbf{X}_b}$ are normalized, i.e. $\sum_b \mathbf{M}_b^\dagger\mathbf{M}_b = \sum_b \mathbf{X}_b = \mathbb{1}$. Moreover, the induced superoperator is unital, because this measurement process does not transform the ensemble described by total mixture. Sometimes this ensemble transformation is associated with the *pre-measurement* process and the measurement alone is related with a *value objectification* of the measurement outcome, in which the stochasticity (and non-linearity) takes place.

4.6 Quantum channels and capacities

We have already described the communication via classical channels characterized by conditional probability $p_{\mathbf{B}}(b|a)$ and defined the capacity of classical channels. The *quantum channel* is nothing else than a tracepreserving completely positive linear map $\mathcal{E}$. It describes the noise occurring during the transmission of signals encoded into states of quantum systems. We shall use the quantum channels for transmitting classical information. Suppose we have a source of classical information expressed by the probability $\pi(a)$, where each a represents a message we want to encode into the quantum state ϱ_a. The *encoding procedure* $\mathcal{C}$ is the mapping from the classical set of messages represented by the *alphabet* $\mathcal{A}$ into the states of quantum system $\mathcal{S}(\mathcal{H})$

$$\{\pi_a, a\} \to \{\pi_a, \varrho_a\} = \varrho = \sum_a \pi_a\varrho_a \tag{4.77}$$

Thus, the encoding process $\mathcal{C}$ corresponds to the superoperator mapping, but the alphabet $\mathcal{A}$ represents the classical system with the states given by probability distributions on $\mathcal{A}$, i.e. $\mathcal{S}(\mathcal{A}) = \mathcal{P}(\mathcal{A})$. The decoding process $\mathcal{D}$ is understood in the same way, i.e. $\mathcal{D} : \mathcal{S}'(\mathcal{H}) \to \mathcal{B}$, where $\mathcal{B}$ is the alphabet of the output messages and $\mathcal{S}'(\mathcal{H}) := \mathcal{E}[\mathcal{S}(\mathcal{H})]$ is the image of the set of states after the action of the channel $\mathcal{E}$.

In particular, the output states of the channel read

$$\varrho_a' = \mathcal{E}[\varrho_a] = \sum_k \mathbf{C}_k\varrho_A\mathbf{C}_k^\dagger. \tag{4.78}$$

In what follows we shall consider that the repetitive input of the channel is described by state $\varrho_{\vec{a}} = \varrho_{a_1} \otimes \ldots \otimes \varrho_{a_N}$ with the probability $\pi_{\vec{a}} = \pi_{a_1} \ldots \pi_{a_N}$. It means that the encoding procedure is *memoryless.* Moreover, we shall assume that the channel is also *memoryless*, i.e. $\mathcal{E}[\varrho_{a_1} \otimes \ldots \otimes \varrho_{a_N}] = \mathcal{E}[\varrho_{a_1}] \otimes \ldots \otimes \mathcal{E}[\varrho_{a_N}]$. The main problem in quantum case of communication is the decoding procedure, because we need to perform a suitable measurement, unlike the classical case, where the measurement is considered to be perfect.

Let us describe the encoding $\mathcal{C}$ by the probabilities π_a, the channel by the superoperator $\mathcal{E}$ and the decoding by POVM $\{\mathbf{X}_b\}$. For the conditional probabilities we have

$$p(b|a) = \mathrm{Tr}(\mathbf{X}_b \mathcal{E}[\varrho_a]) \tag{4.79}$$

and we can use the formulas derived for the Shannon information $I(\mathcal{A}, \mathcal{B}, \pi, \mathcal{E})$ to compute the transmission rates of quantum channels.However, now the information depends also on the decoding measurement, i.e. $I(\mathcal{A}, \mathcal{B}, \pi, \mathbf{X}_b, \mathcal{E})$. For the capacity of the channel $\mathcal{E}$ transmitting the classical information we get

$$\begin{aligned} C(\mathcal{E}) &= \max_{\pi, \mathbf{X}} I(\mathcal{A}, \mathcal{B}, \pi, \mathbf{X}_b, \mathcal{E}) \qquad (4.80) \\ &= \max_{\pi, \mathbf{X}} \left[\sum_{a,b} \pi_a \mathrm{Tr}(\mathbf{X}_b \mathcal{E}[\varrho_a]) \log \frac{\mathrm{Tr}(\mathbf{X}_b \mathcal{E}[\varrho_a])}{\sum_c \pi_c \mathrm{Tr}(\mathbf{X}_b \mathcal{E}[\varrho_c])} \right] \end{aligned}$$

where the maximum is taken over all the possible decoding POVMs and all the possible encodings associated with the probability distributions π_a. *A.S.Holevo* proved in [28], and [29], that this quantity is equivalent to the so-called *Holevo information*

$$C(\mathcal{E}) = \max_{\pi} \left[S(\mathcal{E}[\overline{\varrho}]) - \sum_a \pi_a S(\mathcal{E}[\varrho_a]) \right] \tag{4.81}$$

where $\overline{\varrho} := \sum_a \pi_a \varrho_a$ is the *average state* and $S(\varrho) = -\mathrm{Tr}\varrho \log \varrho$ is the *von Neumann entropy.* In this formula we need to perform the maximalization only over all encodings, what makes this formula more convenient for calculations. Sometimes the capacity of quantum channels is defined also by performing the maximalization over the choice of signal quantum states ϱ_a, but we suppose encoding $\mathcal{C}$ is given by fixed states ϱ_a.

4.6.1 Capacity of qubit channels

Now we shall calculate the capacities of qubit channels. It means we will encode messages into the sequences of qubits that are transmitted via the quantum channel $\mathcal{E}$. Moreover we shall assume, that the information source produces information encoded in classical bits. It means $\mathcal{A} = \{0, 1\}$. We shall not consider the possibility that the alphabet $\mathcal{A}$ contains more elements. The states at the output of the channel are described by states $\varrho_0' = \mathcal{E}[\varrho_1] = \vec{n}$ and $\varrho_1' = \mathcal{E}[\varrho_1] = \vec{m}$. The average state $\overline{\varrho}' = \mathcal{E}[\overline{\varrho}] = \pi_0 \vec{n} + \pi_1 \vec{m}$. Since we know the eigenvalues of the qubit states explicitly, we do not have any problem with calculating the capacity

$$C = \max_{\pi} [S(\pi_0 \vec{n} + \pi_1 \vec{m}) - \pi_0 S(\vec{n}) - \pi_1 S(\vec{m})] \tag{4.82}$$

In what follows we will limit ourselves only to **unitary encodings** $\mathcal{C}$. Consider that Alice obtains a state $\varrho = \vec{n}$. and she is allowed to choose between unitary transformations $\mathbf{U}_a$ to encode the message a into the state $\varrho_a = \mathbf{U}_a \varrho \mathbf{U}_a^\dagger$. Since the encoding transformation is unitary it does not change the eigenvalues and entropy, i.e. $S(\vec{n}_a) = S(\vec{n})$. We know that the eigenvalues of the density operators of a qubit depend only on the length of vector $\vec{n}$, namely $\lambda_\pm = \frac{1}{2} \pm |\vec{n}|$. In our case $|\vec{n}| = |\vec{m}| = m$. Since in our case $\mathcal{A}$ represents the classical bit and $a = 0, 1$, the eigenvalues of the average state $\pi_0 \vec{n} + \pi_1 \vec{m}$

are $\overline{\lambda}_\pm = \frac{1}{2} \pm m\sqrt{1+2\pi_0\pi_1(1-\cos\phi)}$ where ϕ is the angle between vectors $\vec{n}$ and $\vec{m}$. Putting these facts together in Eq.(4.82) the condition of the extreme $\frac{\partial C}{\partial \pi_0}$ requires that the following equality holds

$$-\frac{m(2\pi_0-1)(1-\cos\phi)}{2\sqrt{1+2\pi_0(1-\pi_0)(1-\cos\phi)}}\log\frac{1+m\sqrt{1+2\pi_0(1-\pi_0)(1-\cos\phi)}}{1-m\sqrt{1+2\pi_0(1-\pi_0)(1-\cos\phi)}} = 0.$$

It implies that the maximum is achieved for $\pi_0 = \pi_1 = 1/2$ and we get

$$C = S(m\sqrt{(1+\cos\phi)}/2) - S(m) \tag{4.83}$$

with $S(m) := -(\frac{1}{2}+m)\log(\frac{1}{2}+m) - (\frac{1}{2}-m)\log(\frac{1}{2}-m)$.

In case of the *unitary encoding* the length of vectors is fixed. For a given length m the maximum capacity is obtained for the angle $\phi = 180°$, when

$$C = \log 2 - S(m) = 1 - S(\varrho). \tag{4.84}$$

For pure states ($m = 1/2$) the vectors $\vec{n}$ and $\vec{m}$ correspond to mutually orthogonal states and the capacity $C = 1$ *bit* is the maximal capacity of the qubit channel.

To obtain the maximal capacity of a qubit noiseless channel with the unitary encoding $\mathcal{C}$ Alice needs to perform unitary transformations $\mathbf{U}_0$ and $\mathbf{U}_1$ that generate the mutually orthogonal states in pure case, i.e.

$$\langle\psi|\mathbf{U}_1\mathbf{U}_0|\psi\rangle = 0 \tag{4.85}$$

for all $|\psi\rangle \in \mathcal{H}$. For a fixed state $|\psi\rangle$ it is not difficult to find such operators $\mathbf{U}_0, \mathbf{U}_1$, but in general case it is impossible, because this task is equivalent to the problem of universal NOT machine. And we know, that it is impossible to perform transformation $|\psi\rangle \to |\psi^\perp\rangle$ by a single unitary operation for all $|\psi\rangle$. Let us conclude that, if Alice knows the basis $|\psi\rangle, |\psi^\perp\rangle$, in which the original states ϱ are diagonal, then she can apply the unitaries $\mathbf{U}_0 = \mathbb{1}$ and $\mathbf{U}_1 = \mathbf{S}_1 = |\psi\rangle\langle\psi^\perp| + |\psi^\perp\rangle\langle\psi|$ to obtain signal states

$$\varrho_0 = |\alpha|^2|\psi\rangle\langle\psi| + |\beta|^2|\psi^\perp\rangle\langle\psi^\perp| \tag{4.86}$$

$$\varrho_1 = |\beta|^2|\psi\rangle\langle\psi| + |\alpha|^2|\psi^\perp\rangle\langle\psi^\perp| \tag{4.87}$$

and the capacity equals $C = 1 + |\alpha|^2\log|\alpha|^2 + |\beta|^2\log|\beta|^2$.

Note, that for noisy channels the orthogonality of input states need not achieve the maximum of the transmitted information [36].

4.6.2 Blind encoding

By a blind encoding we will understand the following problem. Alice has two possibilities: she either knows the states on which she performs the encoding, or not. To the second (negative) case we shall refer as to the blind encoding. The main question is, how Alice could make her encoding independent of the initial (unknown) state and what the maximum information she can transmit to Bob is. We can consider that Bob knows the state of the qubit and he knows the transformations Alice is going to perform. That is, the state of the qubit plays the role of public key. To make this protocol work the decoding transformations should not have the property of covariance with respect to the initial state. But the security is not going to be discussed in this section.

Let us denote the unitary encoding transformations of Alice by $\mathbf{U}_a$. Let us express the initial state as $\varrho = \frac{1}{2}\mathbb{1} + \vec{n}.\vec{\sigma}$. The signal states $\varrho_a = \mathbf{U}_a\varrho\mathbf{U}_a^\dagger$ are then given by the three-dimensional vectors $\vec{n}_a$. For the capacity we can write

$$C(\varrho) = \max_\pi\left[S(\sum_a \pi_a\varrho_a)\right] - S(\varrho) \tag{4.88}$$

where we used the equality $S(\varrho_a) = S(\varrho)$ for all a. To maximize the capacity the state $\overline{\varrho} := \sum_a \pi_a \varrho_a$ must maximize the entropy, i.e. $\overline{\varrho} = \frac{1}{2}\mathbb{1}$. In the previous section we could see that it is impossible to fulfill this condition, if we use digital alphabet, i.e. $a = 0, 1$. We have achieved this maximum only with the state-dependent unitary encoding.

It may sound counter-intuitive that the maximum can be achieved, if we use a more-valued alphabet, i.e. more unitary transformations. Let us choose the following four unitary transformations $\mathbf{U}_0 = \mathbb{1}, \mathbf{U}_k = \vec{n}_k.\vec{\sigma}$ for $k = 1, 2, 3$ and $\vec{n}_k.\vec{n}_l = \delta_{kl}$. That is the vectors $\vec{n}_k$ form an orthonormal basis in three-dimensional real vector space. Let us put $\pi_a = 1/4$ for all values a and calculate the state $\overline{\varrho}$

$$\begin{aligned}
\frac{1}{4}\sum_a \mathbf{U}_a \varrho \mathbf{U}_a^\dagger &= \frac{1}{4}\left[\varrho + \sum_k (\vec{n}_k.\vec{\sigma})\varrho(\vec{n}_k.\vec{\sigma})^\dagger\right] \\
&= \frac{1}{2}\mathbb{1} + \frac{1}{4}\left[\vec{n}.\vec{\sigma} + \sum_k (\vec{n}_k.\vec{\sigma})(\vec{n}.\vec{\sigma})(\vec{n}_k.\vec{\sigma})^\dagger\right] \\
&= \frac{1}{2}\mathbb{1} + \frac{1}{4}\left[\vec{n}.\vec{\sigma} + \sum_k (\vec{n}_k.\vec{\sigma})(\vec{n}.\vec{n}_k\mathbb{1} + i(\vec{n} \times \vec{n}_k).\vec{\sigma})\right] \\
&= \frac{1}{2}\mathbb{1} + \frac{1}{4}\left[\vec{n}.\vec{\sigma} + \sum_k ((\vec{n}_k.\vec{n})(\vec{n}_k.\vec{\sigma}) - (\vec{n}_k \times (\vec{n} \times \vec{n}_k)).\vec{\sigma}\right] \\
&= \frac{1}{2}\mathbb{1} + \frac{1}{4}\left[2\sum_k (\vec{n}_k.\vec{n})(\vec{n}_k.\vec{\sigma}) - 2(\vec{n}.\vec{\sigma})\right] = \frac{1}{2}\mathbb{1} \qquad (4.89)
\end{aligned}$$

where we have used the following identities

$$\begin{aligned}
(\vec{n}.\vec{\sigma})(\vec{m}.\vec{\sigma}) &= (\vec{n}.\vec{m})\mathbb{1} + i(\vec{n} \times \vec{m}).\vec{\sigma} && (4.90) \\
\vec{a} \times (\vec{b} \times \vec{c}) &= (\vec{a}.\vec{c})\vec{b} - (\vec{a}.\vec{b})\vec{c} && (4.91) \\
\sum_k (\vec{n}_k.\vec{n})(\vec{n}_k.\vec{\sigma}) &= \sum_{k\alpha\beta} (n_k)^\alpha n^\alpha (n_k)^\beta \sigma_\beta = \vec{n}.\vec{\sigma} && (4.92) \\
\sum_k (n_k)^\alpha (n_k)^\beta &= \delta^{\alpha\beta} \quad \text{completeness} && (4.93)
\end{aligned}$$

and $(n_k)^\alpha$ denotes the α-th component of the vector $\vec{n}_k$.

From the above calculation it follows that if Alice uses four unitary operations $\mathbf{U}_0, \mathbf{U}_k = \vec{n}_k.\vec{\sigma}$ and the source produces messages $a = 0, 1, 2, 3$ with equal probabilities, then the capacity achieves its maximum

$$C(\varrho) = 1 - S(\varrho) \qquad (4.94)$$

We stress once again that the universality (i.e. covariance with respect to the choice of the state ϱ) cannot be achieved with digital alphabet, but with four-character alphabet it is possible.

4.6.3 Noisy channels

The most trivial example of a noisy channel is the unitary one, which can be regarded as noiseless, because its action does not change the structure of the input signal alphabet given by C and preserves the entropy of signal states. The *unital channels* (nonunitary) cause the entropy of signal states to decrease. *Contractive channels* may cause that entropy increases, but they have only one fixed point. It means that if we use the transmission via contractive channels, the output signal states are all closer to this fixed point. If the contractivity parameter is small, then the output states are in a very small vicinity of this fixed point. It follows that channels with more than one fixed point are of importance. If we use these fixed points as input signal states, then the channel acts like noiseless. Of course, such channels make sense only if the fixed points are not too close.

4.7 Quantum dense coding

In this section we shall describe a specific communication protocol, where the entanglement shared between Alice and Bob plays an important role. It was discovered by *Bennett & Wiesner* [42]. Let us briefly describe their example.

Suppose Alice and Bob share a pair of two qubits in a maximally entangled state $|\psi\rangle_{AB} = |\psi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$. In the section 4.2.1 we showed that it is possible to find maximally entangled states which form a basis. Moreover, they can be generated by local unitary transformations applied on one of them. In particular, operators $\sigma_k \otimes \mathbb{1}_B$ acting on $|\psi^+\rangle$ give us the set mutually orthogonal states $|\psi^\pm\rangle, |\phi^\pm\rangle$. Let Alice perform unitary encoding $\mathcal{C}$ on her qubit by choosing one of these unitaries. She prepares one of the mutually orthogonal states. We know that there exists a measurement distinguishing the mutually orthonormal states from a basis with certainty. In this case, the so-called *Bell measurement* is an *global type* of observable. It means neither Bob, nor Alice can distinguish the states locally. Alice must send the prepared qubit to Bob and Bob will realize the *Bell measurement* on both qubits to obtain one of the four possible outcomes. Each outcome of Bob's decoding $\mathcal{D}$ corresponds to a definite operation of Alice's encoding $\mathcal{E}$. Thus, Alice and Bob can communicate 2 bits of information by using entanglement and qubit noiseless channel. Remind us that in the previous consideration the qubit carried maximally one bit of information and now it transmits two classical bits. We shall call this procedure a *dense coding.*

The dense coding strategy is interesting also from the point of security. Since for every maximally entangled state $|\psi\rangle$ the Alice's qubit is described by state $\text{Tr}_A|\psi\rangle\langle\psi| = \frac{1}{2}\mathbb{1}$, the eventual eavesdropper catching signal qubits sent via channel will always find them in the total mixture. Thus, he is unable to say anything about the sent message, because Bob's qubit plays a role of decoding key. Without this key the sent qubit carries no information.

4.7.1 Dense coding with partially entangled states

We have seen that the dense coding protocol is based on the property of maximally entangled states. Namely, it is based on the possibility of preparation of the basis consisting of maximally entangled states just by performing local unitary operations on the Alice's side. The encoding $\mathcal{C}$ used by Alice is unitary. Of course, the choice of encoding is not unique. As we showed (Section 4.2.1), each collection of unitary operators $\mathbf{S}_k$ satisfying $\text{Tr}\mathbf{S}_j\mathbf{S}_k = 2\delta_{jk}$ will do the job.

Now we shall try to generalize the mentioned scheme for partially entangled states. We start with pure states written in the Schmidt basis

$$|\psi\rangle_{AB} = a|00\rangle + b|11\rangle \tag{4.95}$$

where $\{|0\rangle, |1\rangle\}$ is an arbitrary basis of the corresponding Hilbert space. Our aim is to find such choice of four unitary operations $\mathbf{U}_k \otimes \mathbb{1}_B$ (unitary encoding), for which the capacity achieves its maximum. Moreover, we assume, that the choice is independent of the initial state $|\psi\rangle_{AB}$. Without the loss of generality we can put $\mathbf{U}_0 = \mathbb{1}_A$. Let us denote states forming our alphabet by $|\psi_k\rangle = \mathbf{U}_k \otimes \mathbb{1}_B|\psi\rangle$. In the limit of $|\psi\rangle$ is maximally entangled, we must obtain $\langle\psi_k|\psi_l\rangle = \text{Tr}\mathbf{U}_k^\dagger\mathbf{U}_l = 2\delta_{kl}$. In section 4.2.1 we mentioned that this is possible only if $\mathbf{U}_0 = \mathbb{1}$ and $\mathbf{U}_k \equiv \mathbf{S}_k = \vec{n}_k.\vec{\sigma}$ with $\vec{n}_k$ mutually orthonormal vectors in three-dimensional real vector space.

Consider ϱ_{AB} is a general state on which Alice realizes her encoding strategy $\mathcal{E} = \{\mathbf{S}_k \otimes \mathbb{1}_B\}$. Since operators $\mathbf{S}_k$ satisfy the same relations as σ_k-matrices we can express the general state in the operator basis $\mathbf{S}_k$

$$\varrho_{AB} = \frac{1}{4}\mathbb{1}_{AB} + \vec{\alpha}.\vec{\mathbf{S}} \otimes \mathbb{1} + \mathbb{1} \otimes \vec{\beta}.\vec{\mathbf{S}} + \sum_{a,b=1}^{3} \gamma_{ab}\mathbf{S}_a \otimes \mathbf{S}_b. \tag{4.96}$$

and it is much simpler to calculate the alphabet states ϱ_k for $k = 1, 2, 3$

$$\varrho_k \quad = \quad (\mathbf{S}_k \otimes \mathbb{1})\varrho(\mathbf{S}_k^\dagger \otimes \mathbb{1})$$

$$
\begin{aligned}
&= \frac{1}{4}\mathbb{1} + \sum_{c=1}^{3} \alpha_c \mathbf{S}_k \mathbf{S}_c)\mathbf{S}_k^\dagger \otimes \mathbb{1} + \mathbb{1} \otimes \vec{\beta}.\vec{\sigma} + \sum_{a,b=1}^{3} \gamma_{ab} \mathbf{S}_k \mathbf{S}_a \mathbf{S}_k^\dagger \otimes \mathbf{S}_b \\
&= \frac{1}{4}\mathbb{1} + \mathbb{1} \otimes \vec{\beta}.\vec{\sigma} + [2\alpha_k \mathbf{S}_k - \vec{\alpha}.\vec{\mathbf{S}}] \otimes \mathbb{1} + 2\sum_{b=1}^{3} \gamma_{kb} \mathbf{S}_k \otimes \mathbf{S}_b - \sum_{a,b=1}^{3} \gamma_{ab} \mathbf{S}_a \otimes \mathbf{S}_b
\end{aligned}
$$

where we used the identities

$$
\begin{aligned}
\mathbf{S}_a \mathbf{S}_k^\dagger &= \mathbf{S}_a \mathbf{S}_k = (\vec{n}_a.\vec{\sigma})(\vec{n}_k.\vec{\sigma}) \\
&= \vec{n}_a.\vec{n}_k \mathbb{1} + i(\vec{n}_a \times \vec{n}_k).\vec{\sigma} = \delta_{ak}\mathbb{1} + i\varepsilon_{akm}\mathbf{S}_m \qquad (4.97) \\
\mathbf{S}_k \mathbf{S}_a \mathbf{S}_k &= 2\delta_{ak}\mathbf{S}_k + \delta_{kk}\mathbf{S}_a. \qquad (4.98)
\end{aligned}
$$

We remind us that $\vec{n}_k$ are mutually orthonormal three-dimensional vectors which implies that the vector product of two of them is proportional to the third one.

Since the maximum of the capacity is achieved for maximally entangled states with $\pi_k = 1/4$, we shall use such encoding also for general case. For the average state $\overline{\varrho}$ we get

$$
\overline{\varrho} = \frac{1}{4}\sum_{k=0}^{3} \varrho_k = \frac{1}{4}\mathbb{1} + \mathbb{1}_A \otimes \vec{\beta}.\vec{\mathbf{S}} = \frac{1}{2}\mathbb{1}_A \otimes \varrho_B. \qquad (4.99)
$$

where $\varrho_B = \mathrm{Tr}\varrho_{AB} = \frac{1}{2}\mathbb{1} + \vec{\beta}.\vec{\mathbf{S}}$. Finally, we find out that the capacity is equal to

$$
C = S(\overline{\varrho}) - S(\varrho_{AB}) = 1 + S(\varrho_B) - S(\varrho_{AB}) \qquad (4.100)
$$

because $S(\varrho_A \otimes \varrho_B) = S(\varrho_A) + S(\varrho_B)$ and $S(\frac{1}{2}\mathbb{1}_A) = 1$.

4.7.2 Remarks on dense coding

Remark 1. *Maximal capacity*
It is still an open issue whether we can increase the capacity of the dense coding by utilizing larger set of coding unitaries. We address this question in the next section concerning the qudits.

Remark 2. *The asymmetry of dense coding*
Let us note, that the obtained capacity of noiseless qubit channel using the dense coding strategy is not symmetric with respect to the exchange of Bob and Alice. Suppose the same situation as before, i.e. Alice and Bob share a pair of qubits in a state ϱ_{AB}. If Bob sends the messages to Alice (using dense coding strategy) we will obtain

$$
C_{B\to A} = 1 + S(\varrho_A) - S(\varrho_{AB}) \neq 1 + S(\varrho_B) - S(\varrho_{AB}) = C_{A\to B} \qquad (4.101)
$$

since in general $S(\varrho_A) \neq S(\varrho_B)$. In fact, there is no reason to expect the equality, since Bob and Alice use different signal states.

Remark 3. *Beating the classical bound*
It is well known that classical single-bit communication channel cannot be used to transmit more than one bit of information per usage. That is, the capacity with classical resources is always less or equal to one. However, as shown above, in the quantum case single-qubit channel can be used in that way and the transmission rate is greater than the classical bound. We usually use the dense coding to exhibit the nonclassical property of quantum theory. In what follows we will refer to this capacity of dense coding protocol as C^{dense}.

Remark 4. *Correlations are crucial*
To compare our result (4.100) with the previous example (4.84) of the capacity of noiseless qubit channel without using dense coding strategy, we see that always $C^{dense} \geq C^{normal}$. Both of these

scenarios use a unitary encoding $\mathcal{C}$ in order to generate the input signals according to information produced by information source. The choice of the scheme to be used depends on Alice and Bob. It means that the comparison of these two quantities is not only formal, but reflects real possibilities. In particular, the difference

$$C_{A\to B}^{dense} - C_{A\to B}^{normal} = C_{B\to A}^{dense} - C_{B\to A}^{normal} = S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB}) = C(A,B) \tag{4.102}$$

tells us that not the entanglement, but correlations are crucial in dense coding scenario.

Remark 5. *Dense coding in classical settings*

Consider the following classical scenario. Alice performs unitary encoding on total mixtures. It means the states she is getting from the source are uniformly distributed. The source produces either bit value 0, or bit value 1 with probabilities $p(0) = p(1) = 1/2$. Obviously, Alice is unable to encode any information, because on the Bob's side the bits will have random distribution, too. But let the source produce two bits described by the joint classical state $p(00) = p(11) = 1/2$ and $p(01) = p(10) = 0$. One of them is sent to Bob and the second one to Alice. Note, that Alice's bit is in the same state as before, i.e. is randomly distributed. Let the Alice's encoding consist of two operations. If she wants to send the message 0, she will do nothing, and if the message is 1, then she performs classical NOT on her bit. That is, if she obtains a bit of the value j, then she will send a bit of the value $j \oplus k$, where k is the message she wants to send. Bob will receive her bit and compare it with the original one he obtained from the source. If their values coincide, then he knows Alice sent him message 0. If he finds difference between the bits, then he will know that Alice performed NOT and the sent message is 1. Thus, Alice and Bob can communicate using the dense coding strategy. In this classical picture it is more illustrative, that Bob's bit plays a role of a secure key. Formally, the classical and quantum situations are the same. The main difference is on the sender (Alice's) site, because in the classical case Alice may perform only two operations on the classical bit at all. In quantum settings, Alice can choose among infinitely many unitary operations.

Remark 6. *Information flow*

In both the classical and quantum scenarios of dense coding the really transmitted particles from Alice to Bob were in the total mixture irrespective of the encoding transformation selected by Alice. That is the particle sent via the channel has carried no information. Obviously one can ask: *How is the information transmitted from Alice to Bob?* Where is it hidden during the transmission? The answer is transparent - the information is "hidden" in correlations (quantum or classical) between communicating parties. Consequently, any correlations can provide us a "safe", where the information is stored and only person who has the "key" can enter and read the information. It seems that the main difference between the two discussed protocols is in the process how the information behaves during the transmission. In the standard protocol it is attached to the transmitted particle, whereas in the dense coding it is partially hidden in mutual correlations. One can say, that the example of dense coding reflects the non-locality of the information. As we shall discuss in the next paragraph, the non-local transfer can be utilized for secure communication.

Remark 7. *Security*

In this remark we are going to open the question of security of the dense coding (correlation-assisted) communication protocol. Before starting we must somehow determine the conditions, under which the communication is considered to be secure. By secure we intuitively understand those protocols, in which the information is hidden in such a way that no other users except for legitimate can read the message. In what follows we will not care about the detection of various attacks of a potential eavesdropper (Eve). We assume that our eavesdropper does not have any ambition to stay in total illegality. Her aim is only to acquire the secret message and simply to intercept all the information sent by Alice, but she does not bother about her presence being detected. Eve just breaks the channel and intercepts all qubits sent by Alice. It is clear that if Alice and Bob use the standard communication protocol, then Eve can acquire the same amount of information as Bob, i.e. $I_E = I_B = 1 - S(\varrho)$ where ϱ is the state of the system that Alice obtained from the source. We will quantify the security $\mathcal{S}$ of the protocol by this direct eavesdropper attack as the fraction of the information that remains hidden to Eve, i.e. $\mathcal{S} = (I_B - I_E)/I_B = [C - 1 + S(\varrho)]/C$, where C denotes the capacity (information rate) of the transmition between Alice and Bob. As a result we find that the standard communication protocol is totally insecure, i.e. $\mathcal{S} = 0$. From the definition it is clear that security can vary from zero to one, i.e. $\mathcal{S} = 1$ indicates the highest security, when Eve acquires no information, $I_E = 0$.

Earlier we have discussed two specific cases of correlation-assisted communication protocols, in which the transmitted qubit (bit) is in a total mixture. In the classical scenario it is more illustrative that Bob's bit plays a role of a secure key. In fact, the classical correlation-assisted coding protocol is equivalent to the use of the *Vernam code* between Alice and Bob. The Vernam code is known to be unconditionally secure and therefore the communication is as secure as it can be. To send a message written in n bits we use the Vernam code represented by a string of n bits known only by Alice and Bob. In our setting it means that the source produces pairs of bits of the same value. One qubit of each pair is sent to Alice and the second one to Bob. That is, initially the ensemble of pairs of bits is described by the maximally correlated state. Alice and Bob have the same register of n bits. Alice transforms her register according to the message she wants to send by performing a logical (mod2) addition of the message with the register of bits obtained from the source. Then the encrypted message is sent to Bob, who can decode it using his register of bits by the same logical (mod2) addition. In this way Bob obtains the original message. It is clear that the correlation-assisted communication protocol with *maximally* correlated state is equivalent to the communication with the established Vernam code and the security of such communication is maximal. In this case we obtain $\mathcal{S} = [C^{dense}(\varrho_{AB}) - 1 + S(\varrho_A)]/C^{dense}(\varrho_{AB}) = [S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB})]/C^{dense}(\varrho_{AB}) = C_\varrho(A,B)/C^{dense}(\varrho_{AB})$. That is the security of the correlation-assisted coding communication protocol is proportional to mutual correlations $C_\varrho(A,B)$ between Alice and Bob. This is apparent because in the dense communication protocols the information is transferred via correlations between Alice and Bob and therefore it is unreachable by Eve. To achieve the *maximal* level of security it is enough to utilize correlated bi-partite states such that Alice's particles are in a maximally mixed state described by the density operator $\varrho_A = \frac{1}{2}\mathbb{1}$. In such case the information acquired by Eve vanishes, i.e. $I_E = 1 - S(\frac{1}{2}\mathbb{1}) = 0$. Formally, the classical and quantum communication schemes are the same, only the existence of the potential eavesdropper is more difficult to detect in classical case. The problem of much clever attacks (in the sense of Eve's detectability) is out of the scope of this thesis.

4.8 Coding with qudits

Qudits are d-dimensional quantum objects. Following the same line of reasoning like for the qubit case, our aim is to show the possibility of the choice of unitaries $\mathbf{U}_a$, for which the following equality holds

$$\overline{\varrho} = \sum_a \pi_a \mathbf{U}_a \varrho \mathbf{U}_a^\dagger = \frac{1}{d}\mathbb{1} \tag{4.103}$$

because in this case the capacity of the noiseless qudit channel approaches its maximum. Note that for qudits the capacity reads

$$C = \max_\pi [S(\overline{\varrho})] - S(\varrho)\ . \tag{4.104}$$

The following lemma [17] will be useful.

Lemma. *Consider d^2 operators $\mathbf{F}_1, \ldots, \mathbf{F}_{d^2}$ on a d-dimensional Hilbert space $\mathcal{H}$. Then the following conditions are equivalent*

1. $\mathrm{Tr}\mathbf{F}_x^\dagger \mathbf{F}_y = \delta_{xy}$

2. $\sum_x \mathbf{F}_x \mathbf{A} \mathbf{F}_x^\dagger = \mathrm{Tr}(\mathbf{A})\mathbb{1}$ *for all* $\mathbf{A}$

Proof: The condition 1 simply says that the operators $\mathbf{F}_x$ form an orthonormal basis with respect to the scalar product $(\mathbf{A}|\mathbf{B}) = \mathrm{Tr}\mathbf{A}^\dagger\mathbf{B}$. Therefore, the following completeness relation holds $\sum_x |\mathbf{F}_x)(\mathbf{F}_x| = \mathcal{I}$, where we used the notation $|\mathbf{A})$ to distinguish elements of the operator's Hilbert space from the standard ket vectors $|\psi\rangle$ of the original Hilbert space $\mathcal{H}$. Note that $\mathcal{I}$ denotes the identity on the operator Hilbert space $\mathcal{L}(\mathcal{H})$, i.e. $\mathcal{I}[\mathbf{A}] = \mathbf{A}$ for all operators $\mathbf{A}$. Then the following sequence of equalities can be done

$$(\mathbf{P}_\psi|\mathbf{P}_\phi) = \sum_x (\mathbf{P}_\psi|\mathbf{F}_x)(\mathbf{F}_x|\mathbf{P}_\phi) = \sum_x \langle\psi|\mathbf{F}_x|\psi\rangle\langle\phi|\mathbf{F}_x^\dagger|\psi\rangle = \langle\psi|\left[\sum_x \mathbf{F}_x|\psi\rangle\langle\phi|\mathbf{F}_x^\dagger\right]|\phi\rangle \qquad (4.105)$$

Because the left side of this equation can be written as $(\mathbf{P}_\psi|\mathbf{P}_\phi) = \langle\psi|\left[\langle\phi|\psi\rangle\mathbb{1}\right]|\phi\rangle$ and the equality must be valid for any $|\psi\rangle, |\phi\rangle$, we get the operator identity

$$\sum_x \mathbf{F}_x|\psi\rangle\langle\phi|\mathbf{F}_x^\dagger = \langle\psi|\phi\rangle\mathbb{1} = \mathrm{Tr}(|\psi\rangle\langle\phi|)\mathbb{1} \qquad (4.106)$$

On the other hand any operator $\mathbf{A}$ can be written as a linear superposition of the operators of the type $|\psi\rangle\langle\phi|$. Therefore, the same equality holds for any operator $\mathbf{A}$, i.e.

$$\sum_x \mathbf{F}_x\mathbf{A}\mathbf{F}_x^\dagger = \mathrm{Tr}(\mathbf{A})\mathbb{1}\,, \qquad (4.107)$$

and this equality ends the proof. ◇

To use this lemma for our purposes we have to change the conditions a little. Since we need that operators $\mathbf{F}_a = \mathbf{U}_a$ are unitary, the first condition cannot be satisfied, because $\mathrm{Tr}\mathbf{U}_a\mathbf{U}_a^\dagger = d$ instead of one, i.e. the unitary basis is not properly normalized. However, this "complication" can be easily undergone by introducing operators $\mathbf{F}_a := \frac{1}{\sqrt{d}}\mathbf{U}_a$. Using this notation the second condition of the lemma takes the form

$$\frac{1}{d^2}\sum \mathbf{U}_a\mathbf{A}\mathbf{U}_a^\dagger = \frac{1}{d}(\mathrm{Tr}\mathbf{A})\mathbb{1} \qquad (4.108)$$

which is exactly the second condition, i.e. this condition must not be changed. Assuming that $\pi_a = 1/d$ for all a and comparing the last equality with the Eq.(4.103) we see that mutually orthogonal unitary transformations enable us to perform a blind unitary encoding with the maximal transmission rate, i.e.

$$C = S(\overline{\varrho}) - S(\varrho) = \log_2 d - S(\varrho) \qquad (4.109)$$

Using the same reasoning like for the qubit case we derive that the capacity of the noiseless qudit channel using the dense coding strategy is given by the formula

$$C_{dense} = \max_\pi[S(\overline{\varrho}_{AB})] - S(\varrho_{AB}) = \log_2 d + S(\varrho_B) - S(\varrho_{AB}) \qquad (4.110)$$

because $\overline{\varrho}_{AB} = \sum_a (\mathbf{U}_a \otimes \mathbb{1})\varrho_{AB}(\mathbf{U}_a \otimes \mathbb{1})^\dagger = \frac{1}{d}\mathbb{1} \otimes \varrho_B$.

Remark 1 *Maximal capacity*

The obtained result is very useful and completes the question of the maximality of the dense coding strategy. In particular, we opened the question whether a larger set of unitaries can lead to the increasing of the transmission rates of dense coding scenario. But, it is impossible to find such set of local transformations that are mutually orthogonal ($\mathrm{Tr}\mathbf{U}_a\mathbf{U}_b^\dagger = d\delta_{ab}$) and consequently the average state $\overline{\varrho}_{AB}$ cannot be proportional to identity.

The derived formulas for capacities with qudits allow us to generalize directly all the previous results.

Chapter 5

Quantum processors

In theory there is no difference
between theory and practise,
but in practise, there is.
S.D.Poe

5.1 Programming the evolution

Coherent control over individual quantum systems is one of the most exciting achievements in physics in the last decade. The possibility to control quantum dynamics has far reaching consequences for quantum technologies, and in particular for quantum computing. In the theory of coherent quantum control it is assumed that the control of the dynamics is realized via external ("classical") parameters (such as intensity of laser pulses, or duration of the interaction). These external parameters can be viewed as classical information available to experimentalists who use them to achieve the desired control.

In this chapter we will study different type of quantum control. We will assume that the information about the quantum dynamics of the system under consideration is not represented by classical external parameters, but rather is encoded in the state of another quantum system. Generally speaking we want to design programmable quantum device (quantum processor) which would allow us to simulate completely positive maps (i.e. quantum mechanical process) on quantum systems.

Schematically the classical computer is a device with a fixed piece of hardware called the *processor* ("black box") with the input and the output represented by the *register of bits*. Part of this register is associated with the input data and the rest bits encode the program we want to implement. That is, the register of bits is divided into *data register* and *program register*. The processor action is fixed and causes the transformation of the *data register* according to a *program* written in the *program register*. In accordance with this picture of classical processor we shall study its quantum version.

By programs in quantum settings we will understand completely positive tracepreserving linear maps. The Kraus representation theorem implies that any such map can be realized as a unitary transformation **G** acting on a larger Hilbert space providing that the ancillary system is in the state $|0\rangle\langle 0|$. If we change the state of the ancilla, then (in general) the system evolves in a different way. That is, the states of the ancilla system determine the evolution of the system. This evolution is induced by the fixed unitary transformation **G** that describes the interaction between the system and the ancilla. In accordance with the classical case the transformation **G** can be considered as quantum processor and the ancilla system can be associated with the program register. The system represents the data register.

In what follows by the processor we will understand a fixed unitary transformation **G** acting on two systems: data system and program system. The input of the quantum processor will be described by a factorized state $\varrho_d \otimes \xi_p$, where the state ξ_p encodes the program. The output is described by a general state of $\mathcal{S}(\mathcal{H}_d \otimes \mathcal{H}_p)$. From the mathematical point of view the set of processors is the set of all unitary transformations defined on the Hilbert space with the defined tensor product structure. Different types of processors with respect to this structure will be investigated in the next section.

There are two basic questions

- Given a set of programs $\mathcal{C}$ we want to implement. Does there exist a processor $\mathbf{G}$ that is able to implement this set of programs?
- Given a processor $\mathbf{G}$. What is set of all implementable programs $\mathcal{C}$?

5.2 Basic formalism

5.2.1 Pure program states

Let $|\psi\rangle_d$ be the input state of the data register, $|\Xi\rangle_p$ be the input program state and $\mathbf{G}$ be the unitary operator that describes the action of the array of quantum gates. If $\{|j\rangle_p | j = 1, \ldots N\}$ is a basis for the space of program states, then we have that

$$\mathbf{G}(|\psi\rangle_d \otimes |\Xi\rangle_p) = \sum_{j=1}^{N} |j\rangle_p \, {}_p\langle j|\mathbf{G}(|\psi\rangle_d \otimes |\Xi\rangle_p). \tag{5.1}$$

If we define the operator $\mathbf{A}_j(\Xi)$, which acts on the data register, by

$$\mathbf{A}_j(\Xi)|\psi\rangle_d = \, {}_p\langle j|\mathbf{G}(|\psi\rangle_d \otimes |\Xi\rangle_p), \tag{5.2}$$

then we have that

$$\mathbf{G}(|\psi\rangle_d \otimes |\Xi\rangle_p) = \sum_{j=1}^{N} \mathbf{A}_j(\Xi)|\psi\rangle_d \otimes |j\rangle_p. \tag{5.3}$$

This means that the output density matrix of the data register is given by

$$\rho_d^{out} = \sum_{j=1}^{N} \mathbf{A}_j(\Xi)|\psi\rangle_d \, {}_d\langle\psi|\mathbf{A}_j^\dagger(\Xi). \tag{5.4}$$

The operator $\mathbf{A}_j(\Xi)$ depends on the program state, but it can be expressed in terms of operators that do not. Define the operators

$$\mathbf{A}_{jk} = \mathbf{A}_j(|k\rangle) = \, {}_p\langle j|\mathbf{G}|k\rangle_p, \tag{5.5}$$

where $|k\rangle$ is one of the basis states we have chosen for the space of program states. We have that for any program state $|\Xi\rangle$

$$\mathbf{A}_j(\Xi) = \sum_{k=1}^{N} {}_p\langle k|\Xi\rangle_p \mathbf{A}_{jk}. \tag{5.6}$$

This means that the operators $\mathbf{A}_{jk}$ completely characterize the processor in the case of pure program states. We shall call these operators the basis operators for the processor.

These operators have the following property,

$$\sum_{j=1}^{N} \mathbf{A}_{jk_1}^\dagger \mathbf{A}_{jk_2} = \sum_{j=1}^{N} \langle k_1|\mathbf{G}^\dagger|j\rangle\langle j|\mathbf{G}|k_2\rangle = \mathbb{1}\delta_{k_1 k_2}. \tag{5.7}$$

In order to prove this it is enough to consider the decomposition $\sum_j |j\rangle\langle j| = \mathbb{1}_p$.

An obvious question to ask at this point is whether any set of operators satisfying Eq. (5.7) corresponds to a quantum processor. The following construction allows us to show that this is the case [2]. Given a set of N^2 operators acting on $\mathcal{H}_d$, we can construct an operator, $\mathbf{G}$, acting on the product space $\mathcal{H}_d \otimes \mathcal{H}_p$, where $\mathcal{H}_p$ is an N-dimensional space with basis $\{|k\rangle_p | k = 1, \ldots N\}$. We set

$$\mathbf{G} = \sum_{j,k=1}^{N} \mathbf{A}_{jk} \otimes |j\rangle_{p\,p}\langle k|. \tag{5.8}$$

It is now necessary to verify that $\mathbf{G}$ constructed in this way is unitary. Noting that

$$\mathbf{G}^\dagger = \sum_{j,k=1}^{N} \mathbf{A}_{jk}^\dagger \otimes |k\rangle_{p\,p}\langle j|, \tag{5.9}$$

we see that Eq. (5.7) implies that $\mathbf{G}^\dagger \mathbf{G} = \mathbb{1}$, so that $\mathbf{G}$ preserves the length of vectors and is unitary.

It is possible to express the basis operators for closely related processors in terms of each other. For example, if $\{\mathbf{B}_{jk} | j, k = 1, \ldots N\}$ are the basis operators for $\mathbf{G}^\dagger$, then from Eq. (5.9) we see that $\mathbf{B}_{jk} = \mathbf{A}_{kj}^\dagger$. If $\mathbf{G}_1$ and $\mathbf{G}_2$ are two processors (unitary operators) with basis operators $\{\mathbf{A}_{jk}^{(1)} | j, k = 1, \ldots N\}$ and $\{\mathbf{A}_{jk}^{(2)} | j, k = 1, \ldots N\}$, respectively, then the basis operators, $\mathbf{C}_{jk}$, for the processor corresponding to the operator $\mathbf{G}_1 \mathbf{G}_2$ are

$$\mathbf{C}_{jk} = \sum_{n=1}^{N} \mathbf{A}_{jn}^{(1)} \mathbf{A}_{nk}^{(2)}. \tag{5.10}$$

This follows immediately if both $\mathbf{G}_1$ and $\mathbf{G}_2$ are expressed in the form given in Eq. (5.8) and then multiplied together. If we apply this equation to the case $\mathbf{G}_1 = \mathbf{G}$ and $\mathbf{G}_2 = \mathbf{G}^\dagger$, and note that $\mathbf{G}\mathbf{G}^\dagger = \mathbb{1}$, we have that

$$\sum_{j=1}^{N} \mathbf{A}_{k_1 j} \mathbf{A}_{k_2 j}^\dagger = \mathbb{1}\delta_{k_1 k_2}. \tag{5.11}$$

It is clearly possible to generalize Eq. (5.10) to the case when there is a product of more than two operators.

5.2.2 General program states

Let the program is represented by a mixed state $\varrho_p = \sum_{kl} R_{kl} |k\rangle\langle l|$. Then for the induced mapping we have

$$\begin{aligned}\varrho_d^{out} &= \sum_{klmn} R_{kl} \mathbf{A}_{mk} \varrho_d^{in} \mathbf{A}_{nl}^\dagger \mathrm{Tr} |m\rangle\langle n| \\ &= \sum_{klm} R_{kl} \mathbf{A}_{mk} \varrho_d \mathbf{A}_{ml}^\dagger. \end{aligned} \tag{5.12}$$

Earlier we have found that the class of all possible superoperators $\mathcal{C}_G$ realizable deterministically by using some fixed processor $\mathbf{G}$ is fully given by the set of operators $\mathbf{A}_{jk}$ defined by Eq.(5.5). The number of these operators is equal to the $[dim(\mathcal{H}_p)]^2$ and in some sense they create a basis for the class $\mathcal{C}_G$. We have seen that the mapping associated with the processor $\mathbf{G}$ maps the space of program states $\mathcal{S}(\mathcal{H}_p)$ onto the class of superoperators $\mathcal{C}_G$, that is a subset of all possible superoperators.

5.2.3 Correspondence between pure and mixed program states

Let us investigate the case of mixed program states in more detail. In particular, we will address the question whether it is possible for a given class of superoperators $\mathcal{C}_G$ which can be realized by a *fixed* processor $\mathbf{G}$, to find another processor $\mathbf{G}'$ that realize all superoperators in $\mathcal{C}_G$ only by using pure states? We know that each state can be purified, but the purification is not unique [8]. Let us define the purification in the following way

$$\varrho_p = \sum_k \lambda_k |\chi_k\rangle\langle\chi_k| \quad\longrightarrow\quad |\Phi\rangle_{p'} = \sum_k \sqrt{\lambda_k}|\chi_k\rangle_p \otimes |k\rangle\,, \tag{5.13}$$

where ϱ_p is written in spectral decomposition. We define the action of the new processor by the identity

$$\mathbf{G}' := \mathbf{G} \otimes \mathbb{1}\,. \tag{5.14}$$

The dimension of the program system $\mathcal{H}_{p'}$ is equal to $2M$, where M is the dimension of $\mathcal{H}_p$. Now we have to check whether by using only pure program states of the processor $\mathbf{G}'$ we shall obtain the same class of superoperators $\mathcal{C}_G$.

Consider the state ϱ_p and its purification (5.13) $|\Phi\rangle_{p'}$ (i.e. $\varrho_{p'} = |\Phi\rangle\langle\Phi|$) . Then we have to analyze the validity of the equality

$$\mathrm{Tr}_p \mathbf{G}\varrho_d \otimes \varrho_p \mathbf{G}^\dagger = \mathrm{Tr}_{p'} \mathbf{G}'\varrho_d \otimes \varrho_{p'} \mathbf{G}'^\dagger \tag{5.15}$$

for all ϱ_d. The right-hand side of this equation can be rewritten as

$$\begin{aligned}
\mathrm{Tr}_{p'} \mathbf{G}'\varrho_d \otimes \varrho_{p'} \mathbf{G}'^\dagger &= \mathrm{Tr}_{p'}\left[\sum_{kl} \sqrt{\lambda_k\lambda_l}\left(\mathbf{G}\varrho_d \otimes |\chi_k\rangle\langle\chi_l|\mathbf{G}^\dagger\right) \otimes |k\rangle\langle l|\right] \\
&= \sum_{kl} \sqrt{\lambda_k\lambda_l}\mathrm{Tr}_p\left[\left(\mathbf{G}\varrho_d \otimes |\chi_k\rangle\langle\chi_l|\mathbf{G}^\dagger\right)\delta_{kl}\right] \\
&= \mathrm{Tr}_p\left[\mathbf{G}\varrho_d \otimes \left(\sum_k \lambda_k|\chi_k\rangle\langle\chi_k|\right)\mathbf{G}^\dagger\right] \\
&= \mathrm{Tr}_p \mathbf{G}\varrho_d \otimes \varrho_p \mathbf{G}^\dagger\,,
\end{aligned} \tag{5.16}$$

which proves Eq. (5.15). This result allows us to "mimic" mixed program states for a given processor by introducing a larger program system "P'" and a new processor mapping $\mathbf{G}' = \mathbf{G} \otimes \mathbb{1}$.

Let us summarize:

- We can mimic each mixed program state by a pure pure program by using the mapping $\mathbf{G}' = \mathbf{G} \otimes \mathbb{1}$ as a processor.
- For a given processor $\mathbf{G}$ the class of all possible superoperators $\mathcal{C}_G$ is fully given by the operators $\mathbf{A}_{jk}$ that corresponds to some fixed orthonormal basis $|j\rangle_p$ of $\mathcal{H}_p$.
- For any two superoperators Ψ and Φ given by the processor mapping $\mathbf{G}$ and pure states $|\Psi\rangle_p$ and $|\Phi\rangle_p$ the identity

$$\sum_k \mathbf{M}_k^\dagger \mathbf{N}_k = \langle\Phi|\Psi\rangle\mathbb{1} \tag{5.17}$$

holds, where $\mathbf{M}_k = \langle k|\mathbf{G}|\Phi\rangle$ and $\mathbf{N}_k = \langle k|\mathbf{G}|\Psi\rangle$ are any operators representing the given superoperators.

5.2.4 Equivalent processors

We shall regard two processors, $\mathbf{G}_1$ and $\mathbf{G}_2$ as essentially equivalent if one can be converted into the other by inserting *fixed* unitary gate arrays at the input and output of the program register, that is if

$$\mathbf{G}_2 = (\mathbb{1}_d \otimes \mathbf{U}_{p1})\mathbf{G}_1(\mathbb{1}_d \otimes \mathbf{U}_{p2}), \tag{5.18}$$

where $\mathbf{U}_{p1}$ and $\mathbf{U}_{p2}$ are unitary transformations on the program space. If this equation is satisfied, then the processors defined by the two gate arrays will perform the same set of operations on data

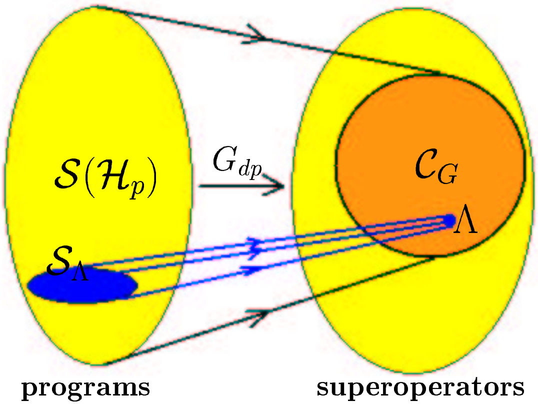

Figure 5.1: We figure the relation between the set of program states $\mathcal{S}(\mathcal{H}_p)$ and the set of all data superoperators induced by the action of the processor mapping $\mathbf{G}_{dp}$. The image of the programs form the subset $\mathcal{C}_G$ of all superoperators. Moreover, it can be seen that the induced map is not bijective, since to a superoperator Λ there exists the subset $\mathcal{S}_\Lambda \subset \mathcal{S}(\mathcal{H}_p)$ containing programs that implements Λ.

states, but the program states required to perform a given operation are different, and the outputs of the program registers will be different as well. If Eq. (5.18) holds, then for the basis operators $\mathbf{A}_{jk}^{(i)}$ $(i = 1, 2)$ associated with the two processors we have

$$\mathbf{A}_{jk}^{(2)} = \sum_{m,n=1}^{N} (\mathbf{U}_{p1})_{jm}(\mathbf{U}_{p2})_{nk}\mathbf{A}_{mn}^{(1)}. \tag{5.19}$$

Therefore, we can regard two processors whose set of operators $\mathbf{A}_{jk}^{(i)}$ are related by the above equation as equivalent.

A special case of this type of equivalence occurs when the two processors are simply related by a change of the basis in the program space, i.e. when $\mathbf{U}_{p1} = \mathbf{U}_{p2}^{-1}$. It is possible to derive conditions that the basis operators of the two processors must satisfy if the processors are to be equivalent in this more restricted sense. These follow from the fact that the trace is independent of the basis in which it is taken. If $\mathbf{U}_{p1} = \mathbf{U}_{p2}^{-1}$, then $\mathrm{Tr}_p(\mathbf{G}_1) = \mathrm{Tr}_p(\mathbf{G}_2)$, which implies that

$$\sum_{j=1}^{N} \mathbf{A}_{jj}^{(1)} = \sum_{j=1}^{N} \mathbf{A}_{jj}^{(2)}. \tag{5.20}$$

We also have that $\mathrm{Tr}_p(\mathbf{G}_1^n) = \mathrm{Tr}_p(\mathbf{G}_2^n)$, which for the case $n = 2$ gives us

$$\sum_{j,k=1}^{N} \mathbf{A}_{jk}^{(1)}\mathbf{A}_{kj}^{(1)} = \sum_{j=1}^{N} \mathbf{A}_{jk}^{(2)}\mathbf{A}_{kj}^{(2)}. \tag{5.21}$$

Clearly, by taking higher values of n we can derive more complicated equivalence conditions.

5.2.5 Measuring the output program register

We now make a measurement

$$\mathbf{M} = \sum_a a\mathbb{1}_d \otimes \mathbf{Q}_a \tag{5.22}$$

of the program system, and, in particular, we shall be interested in the influence of such measurement on the data system. The operators $\mathbf{Q}_a$ are projections, i.e. $\mathbf{Q}_a = \mathbf{Q}_a^\dagger = \mathbf{Q}_a^2$. The measurement will be called successful if we obtain the eigenvalue $a = 1$. Once we obtain the outcome reflecting our success, the resulting state (we assume that the realized measurement is orthogonal) of the data system is given by the following rule (*projection postulate*)

$$\tilde{\rho}_d^{(out)} = \frac{1}{p(s)} \sum_{j=1}^{N} \sum_{j'=1}^{N} \mathbf{A}_j(\Xi)|\psi\rangle_d\, {}_d\langle\psi|\mathbf{A}_{j'}^\dagger(\Xi)\, {}_p\langle j'|\mathbf{Q}|j\rangle_p. \tag{5.23}$$

The normalization factor $p(s)$ plays the role of the success probability (finding the result $a = 1$). It is given by relation

$$p(s) = \sum_{j=1}^{N} \sum_{j'=1}^{N} {}_d\langle\psi|\mathbf{A}_j^\dagger(\Xi)\mathbf{A}_{j'}(\Xi)|\psi\rangle_d\, {}_p\langle j|\mathbf{Q}|j'\rangle_p \tag{5.24}$$

In the case when $\mathbf{Q} = \mathbb{1}_p$ (no measurement) we obtain the Kraus representation of the quantum operations caused by the unitary map $\mathbf{G}$. In this case $\langle j|\mathbf{Q}|j'\rangle_p = \delta_{jj'}$ implies $p(s) = 1$ ($\sum_j \mathbf{A}_j^\dagger(\Xi)\mathbf{A}_j(\Xi) = \mathbb{1}_d$) and $\tilde{\rho}_d^{(out)} = \sum_j \mathbf{A}_j(\Xi)|\psi\rangle_d\langle\psi|\mathbf{A}_j^\dagger(\Xi)$.

If we use the definition of $\mathbf{G}$ via operators $\mathbf{A}_{jk}$ the success probability can be rewritten in the form

$$p(s) = \sum_{jj'kk'} \alpha_k^* \alpha_{k'} Q_{jj'} \langle\psi|\mathbf{A}_{jk}^\dagger \mathbf{A}_{j'k'}|\psi\rangle \tag{5.25}$$

with $Q_{jj'} := \langle j|\mathbf{Q}|j'\rangle$ and $\alpha_j = \langle j|\Xi\rangle$.

Consider that the data register is initially in a pure state $|\psi\rangle$. In what follows we will show that the output data register (after measuring an outcome) is still described by a pure state. Therefore, the whole transformation (induced by processor plus measurement) can be understood as a linear map $\mathbf{A}$ acting on the Hilbert space of the data register $\mathcal{H}_d$. In fact, the unitary transformation followed by a measurement is associated with the set of operators $\{\mathbf{M}_a\}$ (a labels different outcomes), i.e. with the POV measure on $\mathcal{H}_d$. That is, each quantum processor can be used to implement generalized quantum measurements described by POVMs. We shall turn back to this problem in the Section VI.7. Anyway, when the outcome a is observed then the data register evolves according to transformation

$$\varrho_{in} \to \varrho_{out} = \frac{\mathbf{M}_a \varrho_{in} \mathbf{M}_a^\dagger}{\mathrm{Tr}\varrho_{in}\mathbf{F}_a} \tag{5.26}$$

with $\mathbf{F}_a = \mathbf{M}_a^\dagger \mathbf{M}_a$ being positive operators. Calculating the trace of the squared output state we obtain

$$\mathrm{Tr}\varrho_{out}^2 = \mathrm{Tr}\left[\frac{\mathbf{M}_a\varrho_{in}\mathbf{M}_a^\dagger\mathbf{M}_a\varrho_{in}\mathbf{M}_a^\dagger}{[\mathrm{Tr}\varrho_{in}\mathbf{F}_a]^2}\right] = \frac{\mathrm{Tr}(\varrho_{in}\mathbf{F}_a)^2}{[\mathrm{Tr}\varrho_{in}\mathbf{F}_a]^2} \tag{5.27}$$

Assuming that $\varrho_{in} = |\psi\rangle\langle\psi|$ is in a pure state the above equation gives us that $\mathrm{Tr}\varrho_{out}^2 = 1$. It proves our statement that the output data register is in a pure state providing that the input state is pure. **Note.** This statement has an interesting consequence. It can be applied also to orthogonal measurement of the joint data+program system described by a pure state $|\Omega_{in}\rangle_{dp} = |\psi\rangle_d \otimes |\Xi\rangle_p$. Consequently, the output state $|\Omega_a\rangle_{dp}$ must be pure. The subscript a labels the outcomes of the measurement. Apply the same reasoning only to data register initially in a state $|\psi\rangle_d$ the output must be pure state, too. Let us denote this output data state by $|\psi_a\rangle$. Note that the measurement understood on data register is not orthogonal, but it is described by POV measure. Nevertheless, combining these two results we obtain that the output joint state $|\Omega_a\rangle_{dp}$ must be factorized, i.e. $|\Omega_a\rangle_{dp} = |\psi_a\rangle_d \otimes |\phi_a\rangle_p$. That is, any measurement of the joint system providing that the systems are initially pure and factorized results in a pure factorized state. Note that entangled states can be obtained only if the initial states are not pure, or not factorized. This result does not depend on the choice of the unitary transformation $\mathbf{G}$ (before the measurement) and on the choice of the performed measurement as well.

There are several questions that could be stated. Essentially, we have three free "parameters" (we assume that the processor $\mathbf{G}$ is fixed): the input data state, the input program state and the choice of the measurement. Employing these three tools in different configurations we obtain different sets of questions and problems. On the other hand our motivation is to study the possible transformations of the data system. Therefore, the input data state is arbitrary. This decreases the number of parameters we would like to control. We are allowed to prepare an initial program state and choose an arbitrary measurement in order to realize some transformation of the data system. To this problem we will refer as to *probabilistic implementation* of quantum processor and it will be discussed in more details latter. In the next sections we shall investigate the above picture without performing the measurements, i.e. *deterministic implementation*, where the only free parameter is the initial state of the program system.

5.3 Classes of processors

In this section we will introduce several classes of quantum processors. We do not have a handle how to characterize all possible processors so the classes presented below are probably incomplete. Anyway, they give us a deeper insight into what we can do with programmable quantum processors and what are the resources (in terms of properties of program states) required. We will address the question of the properties of the set of realizable programs $\mathcal{C}_G$ by a processors $\mathbf{G}$ belonging to the listed classes of processors.

5.3.1 Local processors

This is the most trivial class of processors defined by the relation $\mathbf{G} = \mathbf{U}_d \otimes \mathbf{U}_p$. These processors are only able to implement single unitary transformation. That is, the set $\mathcal{C}_G$ contains only a single element, namely $\mathbf{U}_d$.

Any unitary transformation can be defined by its eigenvectors and eigenvalues. For example the unitaries belonging to local processors are given by the basis defined in $\mathcal{H}_d$ and $\mathcal{H}_p$, separately. In the following we shall define different types of processors based on the type of their eigenvectors. We start with the case when all the eigenvectors are separable, i.e. $|\psi_\mu\rangle_{dp} = |\phi_\mu\rangle_d \otimes |\chi_\mu\rangle_p$.

5.3.2 U-processors

To obtain the so-called *U-processors* we fix the basis in the program Hilbert space $\mathcal{H}_p$. Thus the eigenvectors are given by

$$|\psi_\mu\rangle_{dp} = |\psi_{ak}\rangle_{dp} = |\phi_{ak}\rangle_d \otimes |k\rangle_p. \tag{5.28}$$

In this case the operators $\mathbf{A}_{jk} = \delta_{jk}\mathbf{U}_j$ are unitary. We remind ourselves that the dimension of $\mathcal{H}_p$ is equal to the number of unitary operators we are able to perform by such processor. This means that the processor can be defined by the relation

$$\mathbf{G}(|\psi\rangle_d \otimes |j\rangle_p) = (\mathbf{U}_j|\psi\rangle_d) \otimes |j\rangle_p \tag{5.29}$$

where the set $\{|j\rangle_p\}_j$ form a basis in $\mathcal{H}_p$. The equation (5.29) holds for all data states $|\psi\rangle_d$. For a general pure program state $|\Phi\rangle_p = \sum_j \alpha_j|j\rangle_p$ the encoded superoperator Φ is given by the expression $\Phi[\varrho_d] = \sum_j |\alpha_j|^2\mathbf{U}_j\varrho_d\mathbf{U}_j^\dagger$. In the case of mixed program state $\varrho_p = \sum_{jk} R_{jk}|j\rangle\langle k|$ the data system evolves according to the equation $\Phi[\varrho_d] = \sum_j R_{jj}\mathbf{U}_j\varrho_d\mathbf{U}_j^\dagger$. Comparing these two different cases we can conclude, that we can always mimic the mixed program state by the pure one. For this purpose it is enough to set $\alpha_j = \sqrt{R_{jj}}$. Hence, for this type of processors it is enough to consider pure program states without any loss of generality. In other words the class of all possible superoperators $\mathcal{C}_G$ is fully given by pure program states. Since for all program states $|\Phi\rangle_p$

$$\Phi[\frac{1}{d}\mathbb{1}] = \sum_j |\alpha_j|^2 \mathbf{U}_j \frac{1}{d}\mathbb{1}\mathbf{U}_j^\dagger = \frac{1}{d}\mathbb{1} \tag{5.30}$$

one can see that each element of $\mathcal{C}_G$ is *unital.*

Example: `CNOT` gate

In accordance with the classical controlled `NOT` gate transforming the bits in the following way

$$00 \to 00 \quad 10 \to 10 \quad 01 \to 11 \quad 11 \to 01 \tag{5.31}$$

the quantum version of the `CNOT` operation is given by the relation

$$\mathtt{CNOT} = \mathbb{1} \otimes |0\rangle\langle 0| + \sigma_x \otimes |1\rangle\langle 1| \tag{5.32}$$

As a result of this transformation the first qubit (data register) is transformed with respect to the state of the second qubit (program register). In particular, if program qubit is prepared in the state $|0\rangle$, then the data qubit remains unchanged. If the program qubit is in the state $|1\rangle$, then the data qubit evolves according to unitary map σ_x. The general program state ξ_p induces the evolution known as *x-Pauli channel,* i.e.

$$\varrho_d \to \varrho_d' = |\alpha|^2 \varrho_d + |\beta|^2 \sigma_x \varrho_d \sigma_x \tag{5.33}$$

with $|\alpha|^2 = \langle 0|\xi_p|0\rangle$ and $|\beta|^2 = 1 - |\alpha|^2$. That is, set $\mathcal{C}_G$ coincideswith the set of x-Pauli channels.

5.3.3 Y-processors

Another possibility how to choose a set of separable eigenvectors is to fix a basis $|k\rangle_d$ in $\mathcal{H}_d$, i.e.

$$|\psi_\mu\rangle_{dp} = |a\rangle_d \otimes |\chi_{ak}\rangle_p. \tag{5.34}$$

In this case the processor is given as $\mathbf{G} = \sum_a |a\rangle_d\langle a| \otimes \mathbf{U}_a$. Assuming the orthonormal basis $\{|\chi_k\rangle_p\}$ in $\mathcal{H}_p$ we obtain

$$\begin{aligned}\mathbf{A}_{jk} &= {}_p\langle\chi_j|\mathbf{G}|\chi_k\rangle_p = \sum_a |a\rangle\langle a|\langle\chi_j|\mathbf{U}_a|\chi_k\rangle \\ &= \sum_a (\mathbf{U}_a)_{jk}|a\rangle\langle a|.\end{aligned} \tag{5.35}$$

We can also check the unitality, that is

$$\begin{aligned}\sum_j \mathbf{A}_{jk_1}\mathbf{A}_{jk_2}^\dagger &= \sum_j\sum_{ab}(\mathbf{U}_a)_{jk_1}(\mathbf{U}_b^\dagger)_{jk_2}|a\rangle\langle a|b\rangle\langle b| \\ &= \sum_{ja}(\mathbf{U}_a)_{jk_1}(\mathbf{U}_a^\dagger)_{jk_2}|a\rangle\langle a| \\ &= \delta_{k_1k_2}\sum_a |a\rangle\langle a| = \delta_{k_1k_2}\mathbb{1}\end{aligned} \tag{5.36}$$

We have found that the Y-processors realize *only* unital superoperators. In particular, consider general program state ξ. Then the data register evolves according to relation

$$\varrho_d \to \varrho_d' = \sum_{a,a'} |a\rangle\langle a'|\langle a|\varrho_d|a'\rangle \mathrm{Tr}_p(\xi \mathbf{U}_{a'}^\dagger \mathbf{U}_a) \tag{5.37}$$

Obviously, if we put $\varrho_d = \mathbb{1}$, then $\varrho_d' = \mathbb{1}$ as well.

5.3.4 Intersection of U-type and Y-type

It is obvious that local processors belong to the intersection of U and Y processors. However, in this paragraph we will show that also another type of processors can be found in this intersection.

The CNOT processor used in our previous example has an interesting property. Namely, it belongs to the intersection of the U-processors and Y-processors, but it cannot be written as a tensor product, i.e.

$$\mathtt{CNOT} = \sum_{\kappa} \mathbf{U}_j \otimes |j\rangle\langle j| = \sum_{\kappa} |\kappa\rangle\langle\kappa| \otimes \mathbf{V}_\kappa \neq \tilde{\mathbf{U}}_d \otimes \tilde{\mathbf{V}}_p \tag{5.38}$$

where $j = 0, 1$, $\mathbf{U}_0 = \mathbb{1}, \mathbf{U}_1 = \sigma_x$, $\kappa = +, -$ and $\mathbf{V}_+ = \mathbb{1}, \mathbf{V}_- = \sigma_z$. Note that vectors $|\pm\rangle$ are defined as $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. It is easy to see that elementary operations $\{\mathbf{U}_j\}$ ($\{\mathbf{V}_\kappa\}$ as well) commute. In general, let us assume that the unitary mapping $\mathbf{G}$ acting on two systems (processor) can be expressed in the form

$$\mathbf{G} = \sum_j \mathbf{U}_j \otimes |j\rangle\langle j| \tag{5.39}$$

and that the unitary operators $\{\mathbf{U}_j\}$ are pairwise commutative ($[\mathbf{U}_j, \mathbf{U}_{j'}] = 0$ for all j, j'). Then this transformation can be written also in the form

$$\mathbf{G} = \sum_\kappa |\kappa\rangle\langle\kappa| \otimes \mathbf{V}_\kappa \tag{5.40}$$

where vectors $\{|\kappa\rangle\}$ form an eigenbasis of each of the unitaries $\mathbf{U}_j$. On the other hand, the vectors $\{|j\rangle_p\}$ are eigenvectors common for all unitary transformations $\mathbf{V}_\kappa$. These facts follows directly from the following sequence of implications. Mutual commutativity of the set $\{\mathbf{U}_j\}$ means that the each of them can be written in the spectral form $\mathbf{U}_j = \sum_\kappa e^{i\alpha_{j\kappa}}|\kappa\rangle\langle\kappa|$, where $\{|\kappa\rangle\}$ is a fixed orthogonal basis of the Hilbert space $\mathcal{H}_d$. Introducing this into the expression of $\mathbf{G}$ we obtain

$$\mathbf{G} = \sum_{j\kappa} e^{i\alpha_{j\kappa}}|\kappa\rangle\langle\kappa| \otimes |j\rangle\langle j| \tag{5.41}$$

It follows then that $\mathbf{V}_\kappa = \sum_j e^{i\alpha_{j\kappa}}|j\rangle\langle j|$ are unitary transformations. As a result we can formulate the following implication

Lemma

Pairwise commutativity of the operators $\{\mathbf{U}_j\}$ *(or* $\{\mathbf{V}_\kappa\}$*) is a sufficient condition for* $\mathbf{G}$ *to be a member of the intersection of U- and Y-processors.*

We note that it is not clear whether there exists some other type of processors with separable eigenvectors. Next we will discuss classes for which the eigenvectors can be entangled.

5.3.5 U′ - processors

Let us consider a simple modification of the U-processor. Using this new processor (which we will call the U′-processor) we are able to implement unitary transformations, too. But in this case, the eigenvectors are no longer separable. Let us define the mapping $\mathbf{G}$ by the relation

$$\mathbf{G} = \sum_k \mathbf{U}_k \otimes |k\rangle_p\langle k'| \tag{5.42}$$

where $\{|k\rangle_p\}$ and $\{|k'\rangle_p\}$ are complete orthonormal bases of the program space $\mathcal{H}_p$. We remind us that if $|k\rangle \equiv |k'\rangle$ for all k, then we get the $\mathbf{U}$-processor. Firstly, we need to show that Eq. (5.42) really defines a unitary transformation. The following calculation

$$\begin{aligned}
\mathbf{G}\mathbf{G}^\dagger &= \sum_{k,l} \mathbf{U}_k\mathbf{U}_l^\dagger \otimes |k\rangle\langle k'|l'\rangle\langle l| = \sum_k \mathbf{U}_k\mathbf{U}_k^\dagger \otimes |k\rangle_p\langle k| \\
&= \mathbb{1}_d \otimes \sum_k |k\rangle_p\langle k| = \mathbb{1}_{dp} = \mathbf{G}^\dagger\mathbf{G}
\end{aligned} \tag{5.43}$$

confirms the unitarity of $\mathbf{G}$. What will happen if we use $|\Phi\rangle_p$ as the state of the program register? We obtain

$$\mathbf{G}|\psi\rangle_d \otimes |\Phi\rangle_p = \sum_k \alpha_k(\mathbf{U}_k|\psi\rangle_d) \otimes |k\rangle_p \tag{5.44}$$

where $|\Phi\rangle_p = \sum_k \alpha_k|k'\rangle_p$, i.e. $\alpha_k := \langle k'|\Phi\rangle$. Evidently, preparing the program register in the state $|k'\rangle_p$ the induced superoperator is unitary. Namely,

$$\varrho_d \to \mathbf{U}_k\varrho_d\mathbf{U}_k^\dagger. \tag{5.45}$$

Let us call these program states as *elementary*. Unlike the U-processor, the program register transforms as $|k'\rangle_p \to |k\rangle_p$, and therefore the eigenstates are not separable. But the action of $\mathbf{G}$ on the elements $|\psi\rangle_d \otimes |k'\rangle_p$ preserves the factorability of these states. It means that elementary program states do not entangle the program and the data. Only in such situation we are able use the processor $\mathbf{G}$ to implement unitary transformations.

For general program register $\varrho_p = R_{kl}|k'\rangle\langle l'|$ we obtain the map

$$\Phi[\varrho_d] = \mathrm{Tr}_p\mathbf{G}\varrho_d \otimes \varrho_p\mathbf{G}^\dagger = \sum_k R_{kk}\mathbf{U}_k\varrho_d\mathbf{U}_k^\dagger. \tag{5.46}$$

This means that the set of realizable superoperators is the same as in the case of the U-processor with $|k\rangle = |k'\rangle$. From this point of view, there is no difference between the U-processors and the U$'$-processors. Nevertheless, U$'$ processors specify two different bases of the program register: the basis of the input (elementary) program states $\{|k'\rangle_p\}$ and the basis associated with the output states of the program register $\{|k\rangle_p\}$. If we compare the operators $\mathbf{A}_{jk}$ and $\mathbf{A}'_{jk}$ associated with these bases, respectively, we find

$$\mathbf{A}_{jk} = \langle j|\mathbf{G}|k\rangle = \sum_m \langle j|m\rangle\langle m'|k\rangle\mathbf{U}_m = \langle j'|k\rangle\mathbf{U}_j, \tag{5.47}$$

and

$$\mathbf{A}'_{jk} = \langle j'|\mathbf{G}|k'\rangle = \sum_m \langle j'|m\rangle\langle m'|k'\rangle\mathbf{U}_m = \langle j'|k\rangle\mathbf{U}_k. \tag{5.48}$$

Evidently, the operators in these bases take different forms, i.e. $\mathbf{A}_{jk} \neq \mathbf{A}'_{jk}$. We remind us that for the U-processor we have $\langle j'|k\rangle = \langle j|k\rangle = \delta_{jk}$, because the two bases coincides. Note that two bases are different, if they are composed from the same set of vectors, but ordered in different way. For instance bases $\{|\psi\rangle, |\phi\rangle, |xi\rangle\}$ and $\{|\xi\rangle, |\psi\rangle, |\phi\rangle\}$ are different.

5.3.6 Y$'$-processors

In what follows we shall generalize the Y-type of processors in the similar way as we have generalized the U-processors. It means we define the unitary transformation by

$$\mathbf{G} = \sum_a |a\rangle_d\langle a'| \otimes \mathbf{U}_a \tag{5.49}$$

where again $\{|a\rangle_d\}$ and $\{|a'\rangle_d\}$ are two complete orthonormal bases in $\mathcal{H}_d$. We only verify the unitality of induced superoperators. For the operators $\mathbf{A}_{jk}$ we obtain

$$\mathbf{A}_{jk} = \langle j|\mathbf{G}|k\rangle = \sum_a |a\rangle\langle a'|(\mathbf{U}_a)_{jk}. \tag{5.50}$$

The unitality condition

$$\begin{aligned}\sum_j \mathbf{A}_{jk_1}\mathbf{A}^\dagger_{jk_2} &= \sum_{j,a,b} |a\rangle\langle a'|b'\rangle\langle b|(\mathbf{U}_a)_{jk_1}(\mathbf{U}^\dagger_b)_{k_2j} \\ &= \sum_{j,a} |a\rangle\langle a|(\mathbf{U}^\dagger_a)_{k_2j}(\mathbf{U}_a)_{jk_1} = \delta_{k_1k_2}\sum_a |a\rangle\langle a| \\ &= \delta_{k_1k_2}\mathbb{1}_d \end{aligned} \tag{5.51}$$

is fulfilled and therefore the implemented superoperators are again unital.

Unlike the U and Y-processors both the U′- and the Y′-processors have entangled eigenvectors. Moreover, if we denote by $\mathbf{G}'$ the U′-processor and by $\mathbf{G}$ the U-processor, then we can write

$$\mathbf{G}'_{dp} = \mathbf{G}_{dp}(\mathbb{1}_d \otimes \mathbf{U}_p) \tag{5.52}$$

where $\mathbf{U}_p = |k\rangle\langle k'|$ and similarly for the Y-processors. That is, these processors are the same up to local unitary transformation performed on the program register (see the discussion in Sec. VI.1.4.). Let us denote by $\mathcal{U}, \mathcal{Y}$ the sets of U-processors and Y-processors, respectively, and by $\mathcal{U}', \mathcal{Y}'$ the sets of U′-processors and Y′-processors. Since prime processors are generalizations of the non-prime processors, we can write

$$\mathcal{U} \subset \mathcal{U}' \quad \text{and} \quad \mathcal{Y} \subset \mathcal{Y}'. \tag{5.53}$$

The sets of implementable superoperators $\mathcal{C}_G, \mathcal{C}_{G'}$ for $\mathbf{G} \in \mathcal{U}, \mathbf{G}' \in \mathcal{U}'$ coincide. More precisely, to each processor $\mathbf{G}$ the processors $\mathbf{G}' = \mathbf{G}(\mathbb{1}_d \otimes \mathbf{U}_p)$ realize the same set of superoperators, i.e. $\mathcal{C}_G \equiv \mathcal{C}_{G'}$. The case of Y-processors is different. Let $\mathbf{G} \in \mathcal{Y}, \mathbf{G}' \in \mathcal{Y}'$. Then the sets of superoperators $\mathcal{C}_G$ and $\mathcal{C}_{G'}$ (with $\mathbf{G}' = \mathbf{G}(\mathbf{U}_d \otimes \mathbb{1}_p)$) are different, since to an arbitrary superoperator $\Phi \in \mathcal{C}_G$ there corresponds a superoperator $\Phi \circ \mathcal{U}_d \in \mathcal{C}_{G'}$ where $\mathcal{U}_d[\varrho_d] = \mathbf{U}_d\varrho_d\mathbf{U}'_d$.

If we define the suitable relations of equivalence among the processors then, in the mathematical sense one can say that the non-prime processors are *quotient sets* of the prime processors, that is

$$\mathbf{G} \sim_p \mathbf{G}', \quad \text{if} \quad \mathbf{G} = \mathbf{G}'(\mathbb{1} \otimes \mathbf{U}_p) \;\Rightarrow\; \mathcal{U} = \mathcal{U}'|_{\sim_p} \tag{5.54}$$

$$\mathbf{G} \sim_d \mathbf{G}', \quad \text{if} \quad \mathbf{G} = \mathbf{G}'(U_d \otimes \mathbb{1}) \;\Rightarrow\; \mathcal{Y} = \mathcal{Y}'|_{\sim_d} \tag{5.55}$$

In this case for the general processor $\mathbf{G}$ the processors $\mathbf{G}'_r = \mathbf{G}(\mathbb{1}\otimes\mathbf{U}_p)$, or $\mathbf{G}'_l = (\mathbb{1}\otimes\mathbf{U}_p)\mathbf{G}$, realize the same set of superoperators, i.e. $\mathcal{C}_{G'_r} \equiv \mathcal{C}_{G'_l} \equiv \mathcal{C}_G$. On the other hand the processors $\mathbf{G}'_r = \mathbf{G}(U_d \otimes \mathbb{1})$, or $\mathbf{G}'_l = (\mathbf{U}_d \otimes \mathbb{1})\mathbf{G}$ implement the superoperators $\Phi \circ \mathcal{U}_d$, or $\mathcal{U}_d \circ \Phi$, where $\Phi \in \mathcal{C}_G$ and $\mathcal{U}_d$ are defined above. Unlike in the previous case, the sets $\mathcal{C}_{G'_r} \neq \mathcal{C}_{G'_l}$, since, in general, $[\mathcal{U}_d, \Phi] \neq 0$. This result is obvious, because a unitary transformation performed on the program register cannot affect the state of the data system, but the unitary transformation of the data register changes the action of the prime processors.

5.3.7 Nonunital processors

All the discussed classes of processors possess one common feature. The sets of all implementable quantum operations $\mathcal{C}_\mathbf{G}$ contain only unital maps. To complete our discrimination of processors with respect to structure of eigenvectors and the type of programs we introduce the class of nonunital processors which contains all other processors which do not belong to the mentioned classes. However, we shall not go into the details.

5.3.8 Covariant processors

Another class of processors that may be of interest are *covariant* processors. These have the property that if the processor maps the input data state $\varrho_{in} = |\psi\rangle_d\, {}_d\langle\psi|$, which we shall assume is a qudit, onto the output density matrix ρ_{out}, then it maps the input state $\mathbf{U}|\psi\rangle_d$ onto the output density matrix $\mathbf{U}\rho_{out}\mathbf{U}^{-1}$, for all $\mathbf{U} \in \mathcal{G}$, where $\mathcal{G}$ is a subgroup of $SU(D)$, for some set of program states $\mathcal{S}$. This relation implies that if $|\Xi\rangle \in \mathcal{S}$, then the operators $\mathbf{A}_j(\Xi)$ satisfy the relation

$$\sum_{j=1}^{N} \mathbf{U}\mathbf{A}_j(\Xi)\varrho_{in}\mathbf{A}_j^\dagger(\Xi)\mathbf{U}^{-1} = \sum_{j=1}^{N} \mathbf{A}_j(\Xi)\mathbf{U}\varrho_{in}\mathbf{U}^{-1}\mathbf{A}_j^\dagger(\Xi), \tag{5.56}$$

for all $\mathbf{U} \in \mathcal{G}$. Let us now consider the case $\mathcal{G} = SU(D)$. If we take $|\psi\rangle_d$ to be a member of an orthonormal basis and then sum both sides of the above equation over all basis elements, we find[1]

$$\sum_{j=1}^{N} \mathbf{U}\mathbf{A}_j(\Xi)\mathbf{A}_j^\dagger(\Xi)\mathbf{U}^{-1} = \sum_{j=1}^{N} \mathbf{A}_j(\Xi)\mathbf{A}_j^\dagger(\Xi). \tag{5.57}$$

Because this holds for all $\mathbf{U} \in SU(D)$, Schur's Lemma implies that

$$\sum_{j=1}^{N} \mathbf{A}_j(\Xi)\mathbf{A}_j^\dagger(\Xi) = c\,\mathbb{1}, \tag{5.58}$$

where c is a constant. Taking the trace of both sides of Eq.(5.58) we find

$$\mathrm{Tr}\left(\sum_{j=1}^{N} \mathbf{A}_j(\Xi)\mathbf{A}_j^\dagger(\Xi)\right) N = c\,\mathrm{Tr}(\mathbb{1}) = c\,N, \tag{5.59}$$

so that $c = 1$. Because this relation holds for any program state, we have that

$$\sum_{j=1}^{N} \mathbf{A}_{jk_1}\mathbf{A}_{jk_2}^\dagger = \delta_{k_1k_2}\mathbb{1}. \tag{5.60}$$

This last equation defines the so-called *unital* superoperators that leaves the total mixture unaffected. Note that this is not the same constraint as Eq. (5.7) because it must hold for any processor.

We showed that if the processor is covariant, then all superoperators implemented by it are unital. But the opposite implication does not hold. Let us assume a processor that is able to implement some unitary operation $\mathbf{V}$. Then from Eq. (5.56) we obtain the condition $\mathbf{U}\mathbf{V}|\psi\rangle_d = \mathbf{V}\mathbf{U}|\psi\rangle_d$ for all $\mathbf{U}$ and $|\psi\rangle_d$, that is $[\mathbf{U}, \mathbf{V}] = 0$ for all $\mathbf{U}$. Applying Schurr's lemma we get $\mathbf{V} = \mathbb{1}_d$. Therefore we conclude that the processor which is able to perform a nontrivial *unitary* transformation $\mathbf{V}$ cannot be covariant for an arbitrary program state. As a result we get that processors from the U$'$ class cannot be covariant with respect to the group $\mathcal{G} = SU(D)$. The question of the existence of nontrivial covariant processor remains still open.

On the other hand, we should note that for a restricted set of program states (i.e. the subset of implementable superoperators) the given processor can behave in a covariant fashion. As an example, let us consider a processor provided by the *quantum information distributor* (see Section VI.5.2). The program state of this device consists of two qubits and the data state is one qubit. The unitary operator, $\mathbf{G}$ can be implemented by a sequence of four `CNOT` gates. A `CNOT` gate acting on qubits j and k, where j is the control bit and k is the target bit, is described by the operator

$$\mathbf{D}_{jk}|m\rangle_j|n\rangle_k = |m\rangle_j|m \oplus n\rangle_k, \tag{5.61}$$

[1] In fact, we note that this is equivalent to consider $\varrho_{in} = \frac{1}{d}\mathbb{1}_d$, Here is the point where the implication "covariance implies" cannot be reversed, and we loose the equivalence relation between the following equations and Eq. (5.56)! Moreover, since all unitary operations performed on a total mixture leave the total mixture unchanged, it is not surprising that the covariance requires unitality. Consequently, we would not need the arguments bellow.

where $m, n = 0$ or 1, and the addition is modulo 2. If we denote the data qubit as qubit 1 and the two program qubits as qubits 2 and 3, then the operator $\mathbf{G}$ for this processor is

$$\mathbf{G} = \mathbf{D}_{31}\mathbf{D}_{21}\mathbf{D}_{13}\mathbf{D}_{12}. \tag{5.62}$$

We remind us that such defined $\mathbf{G}$ is the U-processor, which cannot be covariant in the previous sense.

Let the set of implementable superoperators be specified by the following states of the program register

$$|\Xi\rangle = \alpha|\Xi_{00}\rangle_{23} + \beta|\Phi\rangle_{23}, \tag{5.63}$$

where

$$\begin{aligned} |\Xi_{00}\rangle &= \frac{1}{\sqrt{2}}(|0\rangle_2|0\rangle_3 + |1\rangle_2|1\rangle_3) \\ |\Phi\rangle &= \frac{1}{\sqrt{2}}|0\rangle_2(|0\rangle_3 + |1\rangle_3), \end{aligned} \tag{5.64}$$

and α and β are real, and $\alpha^2 + \beta^2 + \alpha\beta = 1$.

If the data register at the input is described by the state vector ϱ_{in}, then at the output of the processor we find the data register in the state

$$\rho_{out} = (1 - \beta^2)\varrho_{in} + \frac{\beta^2}{2}\mathbb{1}. \tag{5.65}$$

From this equation, it is clear that the processor is covariant with respect to $SU(2)$ if the program state is given by Eq. (5.63). We say, that the *subset of superoperators* $\mathcal{F} \subset \mathcal{C}_G$ (determined by the processor $\mathbf{G}$ and the subset of program states $\{|\Xi\rangle\}$) *is covariant*, if for any unitary transformation, $\mathbf{U} \in SU(D)$, the relation

$$\Phi[\mathbf{U}\varrho_{in}\mathbf{U}^\dagger] = \mathbf{U}\Phi[\varrho_{in}]\mathbf{U}^\dagger \tag{5.66}$$

holds for every $\Phi \in \mathcal{F}$ and for all ϱ_d. Equivalently, we can say that the processor $\mathbf{G}$ is *$\mathcal{F}$-covariant.* In the left-right formalism the above condition corresponds to the commutativity of matrices $\mathcal{U}$ (associated with unitaries $\mathbf{U}$) and $\Phi \in \mathcal{F}$.

5.3.9 Maximal processors

We have mentioned that given processor $\mathbf{G}$ determines the set of all programs $\mathcal{C}_G$. From the mathematical point of view, it induces mapping $\mathcal{G} : \mathcal{S}(\mathcal{H}_p) \to \mathcal{L}_{tcp}(\mathcal{H}_d)$ from the set of states of program register $\mathcal{S}(\mathcal{H}_p)$ into the set of all trace-preserving completely positive linear maps (superoperators) of the states of data register $\mathcal{L}_{tcp}(\mathcal{H}_d)$. The image of this mapping is the set of all realizable programs by the given processor, i.e. $\mathcal{C}_G = \mathcal{G}[\mathcal{S}(\mathcal{H}_p)] \subset \mathcal{L}_{tcp}(\mathcal{H}_d)$. Till now we have studied the following problem: Given a processor $\mathbf{G}$. What is the image of the program state space $\mathcal{C}_G$?

Next we shall investigate the question of the existence of such *maximal processors* $\mathbf{G}$, for which the induced mapping $\mathcal{G}$ is *injective.* That is, each program state encodes different superoperator and the set $\mathcal{C}_G$ contains as many elements as it can providing that the size of the program space is fixed. We have seen that for U$'$-processors the mixed program states are irrelevant, i.e. the mapping $\mathcal{G}$ is not injective in this case. For Y$'$-processors the injectivity of $\mathcal{G}$ is questionable, but it should depend on the specific processor $\mathbf{G}$.

The question on the existence of maximal processors is not so trivial as it may seem. Consider the following example that answer this question in the positive way. In the previous chapter we discussed the process of homogenization. In this process the main result was that the evolution of the system depends on the initial state of the reservoir. For different states ξ the system evolves differently. In more details, it evolves into the state ξ given by the reservoir. Take as a processor $\mathbf{G}$ the partial swap operation, i.e.

$$\mathbf{G} = \cos\eta\mathbb{1} + \sin\eta\mathbf{S} \tag{5.67}$$

where $\mathbf{S} = \sum_{jk} |kj\rangle\langle jk|$ is the swap operation. The program and the data register are associated with single qudit. To prove that different states of the program ξ_p generate different superoperators we shall use the contractivity of the induced mappings. However, in the previous chapter we proved the contractivity only for qubits and now we need to generalize our previous proof.

The induced superoperator is given as

$$\mathcal{E}_\xi[\varrho_d] = c^2\varrho_d + s^2\xi + ics[\varrho_d, \xi] \tag{5.68}$$

with the standard notation $c = \cos\eta, s = \sin\eta$. Let us introduce the distance function $D(\varrho, \sigma) := \sqrt{\mathrm{Tr}(\varrho - \sigma)^2}$. For the given distance we obtain

$$\begin{aligned} D^2(\mathcal{E}_\xi[\varrho], \mathcal{E}_\xi[\varrho']) &= \mathrm{Tr}(c^2(\varrho - \varrho') + ics[\varrho - \varrho', \xi])^2 \\ &= c^4\mathrm{Tr}(\varrho - \varrho')^2 - c^2s^2\mathrm{Tr}[\varrho - \varrho', \xi]^2 \\ &= c^4D^2(\varrho, \varrho') - 2c^2s^2\mathrm{Tr}\{(\varrho - \varrho')\xi\}^2 \\ &\quad +2c^2s^2\mathrm{Tr}(\varrho - \varrho')^2\xi^2 \\ &= c^4D^2(\varrho, \varrho') + 2c^2s^2\sum_k \lambda_k^2\langle k|(\varrho - \varrho')^2|k\rangle \\ &\quad -2c^2s^2\sum_{k,m}\lambda_k\lambda_m|\langle m|\varrho - \varrho'|k\rangle|^2 \\ &= c^4D^2(\varrho, \varrho') \\ &\quad +s^2c^2\sum_{k,m}(\lambda_k - \lambda_m)^2|\langle m|\varrho - \varrho'|k\rangle|^2 \\ &\leq c^4D^2(\varrho, \varrho') \end{aligned} \tag{5.69}$$

where we have used the expression $\sum_{k,m}(\lambda_k - \lambda_m)^2|\langle m|\varrho - \varrho'|k\rangle|^2 \geq 0$. In this calculation we have assumed $\xi = \sum_k \lambda_k|k\rangle\langle k|$ to be written in its spectral decomposition. The last inequality in Eq. (5.69) implies that the superoperator $\mathcal{E}_\xi$ is contractive. Since evidently $\mathcal{E}_\xi[\xi_d] = \xi_d$ is the fixed point of the superoperator $\mathcal{E}_\xi$ and the Banach theorem ensures its uniqueness, it follows that $\mathcal{E}_\xi \neq \mathcal{E}_{\xi'}$ (if $\xi \neq \xi'$). Therefore, the partial swap processors belongs to the class of maximal processors. Different program states ξ_p encodes different superoperators $\mathcal{E}_\xi$. In the context of homogenization problem the above proof enable us to generalize the homogenization process also to qudits.

The crucial step in the proof of the "maximality" of processors is to show the difference between the superoperators. To be sure that two superoperators are different we need some canonical way of their expression in which each superoperator has a unique form. The expression via Kraus operators is not unique, but the left-right expression is.

5.4 Processor design

In the previous sections we have studied sets of superoperators that a given processor can perform. We would now like to turn the problem around and suppose that we have a given set of superoperators, and our aim is to construct a processor that will be able to execute them. Here we will ask more modest question: Under what circumstances we are able to find a processor $\mathbf{G}$ that will perform some set of superoperators? In particular, suppose that we have the superoperators Λ_θ, where the parameter θ varies over some (possibly continuous) range, and that these operators have a Kraus representation $\{\mathbf{B}_j(\theta)|j = 1, \dots M\}$ such that

$$\Lambda_\theta[\varrho] = \sum_{j=1}^{M}\mathbf{B}_j(\theta)\varrho\mathbf{B}_j^\dagger(\theta). \tag{5.70}$$

Our aim is to find a unitary operator, G, and a set of program states $|\Xi_\theta\rangle_p$, where

$$\Lambda_\theta[\varrho_d] = \mathbf{G}(\varrho_d \otimes |\Xi_\theta\rangle_{p\,p}\langle\Xi_\theta|)\mathbf{G}^\dagger. \tag{5.71}$$

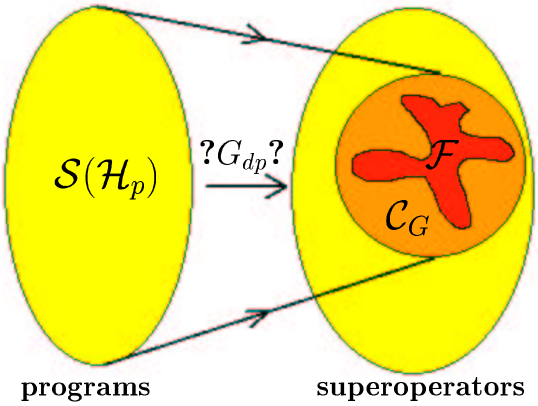

Figure 5.2: Given the set $\mathcal{F}$ of superoperators. We figure the problem of finding the processor G_{dp} that enables us to implement every superoperator $\Lambda \in \mathcal{F}$.

The operators $\mathbf{A}_j(\Xi)$ that represent the action of the processor on the data states when the program state is $|\Xi\rangle$, are now functions of θ and we shall denote them as $\mathbf{A}_j(\theta)$. Our processor then transforms the input data state ϱ_d into the output state, $\varrho_d^{(out)}$

$$\rho_d^{(out)} = \sum_{j=1}^{N} \mathbf{A}_j(\theta)\varrho_d\mathbf{A}_j^\dagger(\theta). \tag{5.72}$$

We shall find the condition that the operators in the Kraus representation of the superoperator must satisfy in order for such processor to exist. We first note that the operators $\{\mathbf{A}_j(\theta)|j = 1,\dots N\}$ constitute also a Kraus representation of the superoperator Λ_θ. The Kraus representation of a superoperator is not unique. Any two different Kraus representations of the same superoperator, $\{\mathbf{B}_j|j = 1,\dots M\}$ and $\{\mathbf{C}_j|j = 1,\dots N\}$, where $N \geq M$, are related as [2]

$$\mathbf{C}_j = \sum_{k=1}^{N} U_{kj}\mathbf{B}_k, \tag{5.73}$$

where the coefficients U_{kj} form a unitary matrix. It is understood that if $N > M$, then zero operators are added to the set $\{\mathbf{B}_j|j = 1,\dots M\}$ so that the two sets of operators have the same cardinality.

In what follows we will present some processors which simulate specific quantum channels and can be programmed with the help of quantum program registers.

1. Universal quantum processor

Universality of devices is a very valuable and desired feature. Universal quantum processor should be able to implement any superoperator. The first step to achieve this objection is to implement any unitary transformation. If we succeed, then the implementation of all superoperators is straightforward. The Kraus theorem implies that any superoperator can be imagined as a unitary transformation on d^2-dimensional Hilbert space, where $d = \dim \mathcal{H}_d$. Therefore, if we are able to implement any unitary transformation on d^2 Hilbert space, then we are able to realize also any superoperator on our d-dimensional data register. This problem was firstly considered by Nielsen and Chuang [44].

Consider that such universal processor exists. If $|\psi\rangle_d$ is the state of the data register and $|\Xi_U\rangle_p$ is the state of the program register that implements the operation $\mathbf{U}$, the processor carries out the transformation

$$\mathbf{G}(|\psi\rangle_d \otimes |\Xi_U\rangle_p) = \mathbf{U}|\psi\rangle_d \otimes |\Xi'_{U,\psi}\rangle_p \tag{5.74}$$

where $|\Xi'_{U,\psi}\rangle_p$ is the state of the program register after the transformation $\mathbf{G}$ has been carried out. The subscripts U and ψ indicate that this state can depend on both the operation $\mathbf{U}$ and data state $|\psi\rangle_d$. The linearity of $\mathbf{G}$ implies that $|\Xi'_{U,\psi}\rangle$ is ψ-independent, i.e. $|\Xi'_U\rangle$.

Consider two programs $|\Xi_U\rangle$ and $|\Xi_V\rangle$ that cause operations $\mathbf{U}$ and $\mathbf{V}$, respectively, to act on data register. This implies that

$$\begin{aligned} \mathbf{G}(|\psi\rangle_d \otimes |\Xi_U\rangle_p) &= \mathbf{U}|\psi\rangle_d \otimes |\Xi'_U\rangle_p && (5.75)\\ \mathbf{G}(|\psi\rangle_d \otimes |\Xi_V\rangle_p) &= \mathbf{V}|\psi\rangle_d \otimes |\Xi'_V\rangle_p && (5.76) \end{aligned}$$

The unitarity of G implies that

$$_p\langle\Xi_V|\Xi_U\rangle_p = \ _d\langle\psi|\mathbf{V}^{-1}\mathbf{U}|\psi\rangle_d \ _p\langle\Xi'_V|\Xi'_U\rangle_p \qquad (5.77)$$

and if $_p\langle\Xi'_V|\Xi'_U\rangle_p \neq 0$, then

$$_d\langle\psi|\mathbf{V}^{-1}\mathbf{U}|\psi\rangle_d = \frac{_p\langle\Xi_V|\Xi_U\rangle_p}{_p\langle\Xi'_V|\Xi'_U\rangle_p} \qquad (5.78)$$

The left-hand side of this equation depends on $|\psi\rangle_d$ while the right-hand side does not. The only way this can be true is if

$$\mathbf{V}^{-1}\mathbf{U} = e^{i\phi}\mathbb{1} \qquad (5.79)$$

for some real ψ. This means that operators $\mathbf{V}$ and $\mathbf{U}$ are same up to a phase. If we want these operators to be different, then necessarily $_p\langle\Xi'_V|\Xi'_U\rangle_p = 0$, which implies that $_p\langle\Xi_V|\Xi_U\rangle_p = 0$, too. Therefore, the states of program registers corresponding to different unitary transformations must be orthogonal. This implies that the dimension of the program register must be greater or equal to the number of different unitary transformations that can be performed on data register.

Alternatively, this can be proven by using the following statement (see Sec. VI.2). If two superoperators $\Phi \leftrightarrow \{\mathbf{M}_\mu\}$, $\Psi \leftrightarrow \{\mathbf{N}_\mu\}$ are realized with the same processor $\mathbf{G}$, then

$$\sum_\mu \mathbf{M}^\dagger_\mu \mathbf{N}_\mu = \langle\Phi|\Psi\rangle\mathbb{1} \qquad (5.80)$$

with the operators $\mathbf{M}_\mu = \langle\mu|\mathbf{G}|\Phi\rangle$ and $\mathbf{N}_\mu = \langle\mu|\mathbf{G}|\Psi\rangle$. Unitary transformations $\mathbf{U}, \mathbf{V}$ contains only one operator in their Kraus representation, i.e. $\mathbf{U}$ and $\mathbf{V}$ respectively. If they are implementable by the same processor $\mathbf{G}$, then necessarily $\mathbf{V}^{-1}\mathbf{U} = c\mathbb{1}$ where c is some constant. The unitarity of $\mathbf{U} \neq \mathbf{V}$ lead us to the solution that the parameter c must be equal to zero, i.e. $c = 0$. This parameter c corresponds to the scalar product between the associated states of the program register $(|\Xi_U\rangle, |\Xi_V\rangle)$. Hence, we obtain the same result as before. Program states must be orthogonal.

Consequently, no universal quantum processor exists, since its existence implies that the program space is uncountable and therefore the corresponding Hilbert space is not separable. One can use the generalized eigenfunctions (δ-distributions) to represent the desired programs, but the physical relevance of such device is questionable [50].

2. Phase-damping channel.

We mentioned that phase damping channel is a specific example of *Pauli superoperators.* It is given by operators $\mathbf{M}_0(p) = \sqrt{1-p}\mathbb{1}$ and $\mathbf{M}_1(p) = \sqrt{p}\sigma_z$ where both, the σ_z and $\mathbb{1}$ are unitary operators. Hence for the mapping $\mathcal{C}(p)$ we have

$$\mathcal{C}(p)[\varrho_d] = (1-p)\mathbb{1}\varrho_d\mathbb{1} + p\sigma_z\varrho_d\sigma_z^\dagger. \qquad (5.81)$$

Simply we can define the processor (*U-type*) by the equation

$$\mathbf{G}_{dp}|\phi\rangle_d \otimes |k\rangle_p = (\mathbf{U}_k|\phi\rangle_d) \otimes |k\rangle_p\,, \qquad (5.82)$$

where $k = 0, 1$ and $U_0 = \mathbb{1}, U_1 = \sigma_3$. The state $|\Phi(p)\rangle_p = \sqrt{p}|0\rangle_p + \sqrt{1-p}|1\rangle_p$ represent the program state, in which the required transformation $\mathcal{P}(p)$ is encoded. It means we are able to encode the parameter p into the program state and we can execute the whole one parameter set of superoperators $\mathcal{C}(p)$.

3. Amplitude damping channel

We have already mentioned this type of superoperator given by the parameter p and the operators $\mathbf{M}_0(p) = |0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1|, \mathbf{M}_1(p) = \sqrt{p}|0\rangle\langle 1|$. For a fixed p such mapping can be realized by introducing one qubit and apply the transformation

$$\begin{aligned}\mathbf{G}_{dp} &= |00\rangle\langle 00| + |11\rangle\langle 11| \\ &+\sqrt{1-p}(|10\rangle\langle 10| - |01\rangle\langle\ 01|) \\ &+\sqrt{p}(|01\rangle\langle 10| + |10\rangle\langle 01|)\end{aligned} \tag{5.83}$$

Unfortunately such realization doesn't serve our purposes, because the processor mapping itself depends on the value of the parameter p, i.e. $\mathbf{G}_{dp} = \mathbf{G}_{dp}(p)$, and the program state is fixed, which is not the situation we wanted. In our scenario we want to encode the value p into the state of program register and simulate the amplitude damping by the given processor.

So the question under consideration is: Are we able to find a processor independent of p which executes the whole set of superoperators $\mathcal{P}(p)$? Of course, we can do it for a *finite* number of values p. As we have already mentioned, for any two superoperators $\mathcal{P}_\Phi \leftrightarrow \{\mathbf{N}_\mu\}$ and $\mathcal{P}_\Psi \leftrightarrow \{\mathbf{M}_\mu\}$ implementable by the processor mapping $\mathbf{G}_{dp}$ and pure states $|\Psi\rangle_p$ and $|\Phi\rangle_p$ holds the identity

$$\sum_\mu \mathbf{M}_\mu^\dagger \mathbf{N}_\mu = \langle\Psi|\Phi\rangle\mathbb{1} \tag{5.84}$$

To satisfy this property for any two parameters p, p', we must let $\mathbf{N}_1 = 0, \mathbf{N}_2 = 0, \mathbf{N}_3 = \mathbf{M}_0(p), \mathbf{N}_4 = \mathbf{M}_1(p)$ to be Kraus operators of the program $\mathcal{P}(p)$ and $\mathbf{N}_1' = \mathbf{M}_0(p'), \mathbf{N}_2' = \mathbf{M}_1(p'), \mathbf{N}_3' = 0, \mathbf{N}_4' = 0$ to represent the program $\mathcal{P}(p')$. Then trivially $\sum_k \mathbf{N}_k^\dagger \mathbf{N}_k' = 0$ and the identity (5.84) holds. In principle, we can construct the processor $\mathbf{G}_{dp}$ realizing both superoperators for two different values of p, p'. In order to realize the uncountable set of parameters p we need the Hilbert space with uncountable basis. (We remind us the case of universal processor).

However, the question of a realization of all amplitude damping channels is still open, because to single superoperator there exists many different Kraus representations. What we now must do is to try to find a Kraus representation for this channels that does satisfy Eq.(5.84). In particular we assume that

$$\mathbf{C}_\nu(p) = \sum_{\mu=0}^{N-1} U_{\mu\nu}(p)\mathbf{M}_\mu(p) \tag{5.85}$$

where coefficients $U_{\mu\nu}(p)$ form an $N \times N$ unitary matrix, and $\mathbf{M}_\mu(p) = 0$ for $\mu = 2, \ldots, N-1$. Note that N represents the number of Kraus operators representing the same channel (fixed p) as operators $\mathbf{M}_0(p), \mathbf{M}_1(p)$. In addition we demand

$$\sum_\nu \mathbf{C}_\nu^\dagger(p_1)\mathbf{C}_\nu(p_2) = f(p_1, p_2)\mathbb{1}\,, \tag{5.86}$$

where $f(p_1, p_2)$ is a function whose magnitude is less than or equal to one. We will show that there is no such Kraus representation with N finite. If the last equation is to hold, then the coefficients of $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ must be the same.Inserting the explicit expressions for $\mathbf{C}_\nu(p)$ in terms of $\mathbf{B}_\mu(\theta)$, this condition becomes

$$(1 - \sqrt{(1-p_1)(1-p_2)})\sum_\nu U_{0\nu}^*(p_1)U_{0\nu}(p_2) = \sqrt{p_1 p_2}\sum_\nu U_{1\nu}^*(p_1 0 U_{1\nu}(p_2) \tag{5.87}$$

We can now make use of the fact that the rows of a unitary matrix constitute orthonormal vectors and the Schwarz inequality to show that the magnitude of the sum on the right-hand side of this equation is less than or equal to one. This gives us that

$$|\sum_{\nu} U^*_{0\nu}(p_1)U_{0\nu}(p_2)| \leq \frac{\sqrt{p_1p_2}}{1-\sqrt{(1-p_1)(1-p_2)}} \tag{5.88}$$

Now we need the following lemma

Lemma. If the set $\{|v_j\rangle\}$ contains N vectors of length 1, and $|\langle v_j|v_k\rangle| < 1/(N-1)$, then the vectors $\{|v_j\rangle\}$ are linearly independent.

Proof. The vectors are linearly dependent, if there are constants c_j, at least some of them are nonzero, such that $\sum_j c_j|v_j\rangle = 0$. Taking the inner product of both sides with $|v_k\rangle$ the assumptions of the lemma imply that

$$|c_k| = |\sum_{j;j\neq k} c_j\langle v_k|v_j\rangle| < \frac{1}{N-1}\sum_{j;j\neq k}|c_j| \tag{5.89}$$

Summing both sides of the above inequality over k gives us that

$$\sum_k |c_k| < \frac{1}{N-1}\sum_k\sum_{j;j\neq k}|c_j| = \sum_k |c_k| \tag{5.90}$$

which is clearly impossible. Therefore the vectors must be linearly independent. ⋄

This lemma can be applied to the first row of the unitary matrix $U(p)$, which we can think of as an N-component normalized vector, which we shall call $u_0(p)$. What we will show is that we can find arbitrarily many of these vectors whose inner products can be made arbitrarily small. The lemma then implies that these vectors are linearly independent, but this contradicts the fact that they lie in an N-dimensional space. Hence, there must be an infinite number of Kraus operators, and the program space must be infinite dimensional.

In order to study the inner products of the vectors $u_0(p)$ for different values of p, we need to examine the function appearing on the right-hand side of Eq.(5.88)

$$g(p_1,p_2) = \frac{\sqrt{p_1p_2}}{1-\sqrt{(1-p_1)(1-p_2)}} \tag{5.91}$$

Using the fact that if $0 \leq p \leq 1$, then $\sqrt{1-p} \leq 1-\theta/2$. we have that

$$g(p_1,p_2) \leq \frac{2\sqrt{p_1p_2}}{p_1+p_2-(p_1p_2/2)} \tag{5.92}$$

Finally noting that for $p_1, p_2 \in [0,1]$

$$\frac{p_1+p_2}{p_1+p_2-(p_1p_2/2)} \leq \frac{4}{3} \tag{5.93}$$

we derive the upper bound

$$g(p_1,p_2) \leq \frac{8\sqrt{p_1p_2}}{3(p_1+p_2)} \tag{5.94}$$

Consider now that we would like to realize the set of channels parametrized by values $q_n = [1/(16M^2)]^n$, where $n = 1, 2, \ldots$ and M is any positive integer. Putting $p_1 = q_n$ and $p_2 = q_m$ where $m > n$ the inner product is bounded by

$$g(p_1,p_2) \leq \frac{8}{3}\frac{1}{(4M)^{m-n}} \tag{5.95}$$

Therefore the set of vectors $\{u_0(q_m) : m = 1, \dots M\}$ have pairwise inner product whose magnitudes are less that $1/M$, and, therefore, they are linearly independent. As these vectors have N components, if we choose $M > N$ we have a contradiction, because the number of linearly independent vectors in N-dimensional space must be less than, or equal to N. This implies that the number of Kraus operators is infinite, and the amplitude-damping channel cannot be realized by a finite quantum processor.

In summary, for amplitude damping channel $\mathcal{A}_p$ the condition (5.84) cannot hold for every pair of parameters p, p'. Consequently, it means that it is impossible to encode the parameter p (rate of the damping) into a state of the program register.

5.5 Probabilistic implementation

Till now we have described the deterministic implementation, it means with probability equal one we were able to implement a set of superoperators $\mathcal{C}_G$ using the processor $\mathbf{G}_{dp}$. We have seen that no universal processor in this scenario can exist. In this section we will investigate what will happen, if we perform a measurement $\mathbf{M}$ of the output program register. Of course, the realization of the measurement will bring probabilistic features in our description and our implementation becomes to be characterized by the *probability of success.*

In particular, we are mainly interested in the realization of unitary mappings, i.e. only in cases when $\varrho_d^{(out)} = \mathbf{U}\varrho_d\mathbf{U}^\dagger$. Let us assume that a projection $\mathbf{Q}$ is only one-dimensional and projects onto the state $|Q\rangle$. Define the numbers $q_j := \langle Q|j\rangle$. Then the Eq.(5.23) can be rewritten in the form

$$\tilde{\varrho}_d^{(out)} = \left[\sum_j \frac{q_j}{\sqrt{p(s)}}\mathbf{A}_j(\Xi)\right] |\psi\rangle\langle\psi| \left[\sum_{j'} \frac{q_{j'}^*}{\sqrt{p(s)}}\mathbf{A}_{j'}(\Xi)^\dagger\right]$$

That is, the output state $\varrho_d^{(out)}$ can be written in the following form $\tilde{\varrho}_d^{(out)} = |\psi^{(out)}\rangle\langle\psi^{(out)}|$, where

$$|\psi^{(out)}\rangle_d = \left[\sum_j \frac{q_j}{\sqrt{p(s)}}\mathbf{A}_j(\Xi)\right] |\psi\rangle_d = \mathbf{A}(\Xi, Q, \psi)|\psi\rangle_d \tag{5.96}$$

Note that $\mathbf{A}(\Xi, Q, \psi)$ is not necessarily a linear operation, because it is ψ-dependent.

Define a ψ-independent linear operator $\tilde{\mathbf{A}}(\Xi, Q) := \sum_j q_j \mathbf{A}_j(\Xi)$. Then the probability of success reads

$$p(s) = \langle\psi|\tilde{\mathbf{A}}(\Xi, Q)^\dagger \tilde{\mathbf{A}}(\Xi, Q)|\psi\rangle \tag{5.97}$$

Further analysis will focus on the conditions under which it is possible to make this probability independent of the initial data states $|\psi\rangle_d$, when the performed transformation is indeed linear. Such requirement means $\langle\psi|\tilde{\mathbf{A}}^\dagger\tilde{\mathbf{A}}|\psi\rangle = const$ for all states $|\psi\rangle \in \mathcal{H}_d$. Let us fix the vector $|\psi\rangle$. Then all the other vectors can be obtained by the unitary transformation $\mathbf{V}$ acting on this vector, i.e. $|\psi'\rangle = \mathbf{V}|\psi\rangle$. Then the condition "for all vectors $|\psi\rangle$" can be rewritten in the following way

$$\langle\psi|\mathbf{V}^\dagger\tilde{\mathbf{A}}^\dagger\tilde{\mathbf{A}}\mathbf{V}|\psi\rangle = \langle\psi|\tilde{\mathbf{A}}^\dagger\tilde{\mathbf{A}}|\psi\rangle \quad \forall\mathbf{V} \tag{5.98}$$

Since such condition should be independent of the state $|\psi\rangle$, then the above equation represents the commutation relation

$$[\tilde{\mathbf{A}}^\dagger\tilde{\mathbf{A}}, \mathbf{V}] = 0 \qquad \forall\mathbf{V} \tag{5.99}$$

Finally, the Schurr lemma implies that the operator $\tilde{\mathbf{A}}^\dagger\tilde{\mathbf{A}}$ must be proportional to identity operator $\mathbb{1}$. The solution has the form $\tilde{\mathbf{A}} = k\mathbf{U}$, where k is a complex number ($|k|^2 = p(s)$) and $\mathbf{U}$ is a unitary operation. That is, we have obtained some restrictions on our choice of the initial program state and projection (measurement).

On the other side the data system evolves in the following way $|\psi\rangle \mapsto \frac{1}{\sqrt{p(s)}}\tilde{\mathbf{A}}(\Xi, Q)|\psi\rangle$. The requirement of the unitarity of such transformation leads us to the condition

$$\frac{1}{p(s)}\tilde{\mathbf{A}}^\dagger\tilde{\mathbf{A}} = \mathbb{1} \tag{5.100}$$

what is the same result as we have derived from the ψ-independence condition, i.e. $\tilde{\mathbf{A}}^\dagger\tilde{\mathbf{A}} = p(s)\mathbb{1}$. As a result we can formulate the following theorem:

Theorem. *The implemented operation* $\mathbf{A} = \frac{1}{p(s)}\tilde{\mathbf{A}}$ *is unitary if and only if the probability of success p(s) is independent of the initial state of the data register.*

This theorem limits the possible choices of couples $(\Xi, \mathbf{Q})$. In other words not all program states $|\Xi\rangle$ and not all measurements $\mathbf{Q}$ cause the unitary transformation of the data system. Next two types of questions can be asked:

- Given (fixed) a measurement projection $\mathbf{Q}$. For which program states $|\Xi\rangle$ the data system evolves in a unitary way? Let us denote by $\mathcal{P}_Q$ the set of all such program states.
- Given a program state $|\Xi\rangle$. Does there exist a projection $\mathbf{Q}$, such that the performed transformation of the data system is unitary? Denote the set of all such projections by $\mathcal{M}_\Xi$.

Next we shall be interested in cases when a measurement $\mathbf{Q}$ is fixed and the probability of success $p(s)$ for all programs in $\mathcal{P}_Q$ is constant. That is, it is not only ψ-independent (the transformation is unitary), but also Ξ-independent (for $|\Xi\rangle \in \mathcal{P}_Q$). With these requirements one can investigate also the relation between the size of the programs $\mathcal{P}_Q$ and the value of the success probability $p(s)$. The last question is the question on the optimality in the sense of the mutual relation "$p(s)$ *versus* $\mathcal{P}_Q$". That is we have to aims. On one hand we would like to realize as much unitaries as it is possible, but on the other hand we would like to preserve the probability large enough.

We are not able to give here an explicit form of projections $\mathbf{Q}$ and programs $\mathcal{P}_Q$ in these most general settings (with general processor). Consider therefore a very specific case, when $q_j = \sqrt{p(s)}$ for all values of j (note that $p(s) < 1$). The realized transformation takes the form

$$\mathbf{U} = \sum_j \mathbf{A}_j(\Xi) \tag{5.101}$$

We have to show the existence of the one-dimensional projection $\mathbf{Q}$ with the required property $Q_{jj'} = \langle j|\mathbf{Q}|j'\rangle = q_j^* q_j' = p(s)$ for all values of indices j, j'. The property $\mathbf{Q}^2 = \mathbf{Q}$ implies

$$(\mathbf{Q}^2)_{jj'} = \sum_k (\mathbf{Q})_{jk}(\mathbf{Q})_{kj'} = \sum_{k=1}^N p(s)^2 = Np(s)^2 = p(s) = (\mathbf{Q})_{jj'}$$

where N denotes the dimension of the program Hilbert space $\mathcal{H}_p$. The equality $Np(s)^2 = p(s)$ can be satisfied only when $p(s) = 1/N$. For example the state $|Q\rangle = \frac{1}{\sqrt{N}}\sum_j |j\rangle$ satisfies such property.

The goal we have achieved is that the success probability with one-dimensional projector can be made constant and in our case it equals $1/N$. If we use a general program state $|\Xi\rangle$ and apply the measurement $\mathbf{Q}$ with $q_j = 1/N$, then the data transforms according to map

$$\sum_j \frac{1}{N\sqrt{p(s)}}\mathbf{A}_j(\Xi) \tag{5.102}$$

which is not linear and the probability of success ($p(s) \neq 1/N$) depends on the initial data state $|\psi\rangle$. The set of all allowed program states $|\Xi\rangle = \sum_k \alpha_k|k\rangle \in \mathcal{P}_Q$ is determined by the condition

$$\sum_{jj'kk'} \alpha_k\alpha_{k'}\mathbf{A}_{jk}\mathbf{A}^\dagger_{j'k'} = \mathbb{1} \tag{5.103}$$

where the operators $\mathbf{A}_{jk} := \langle j|\mathbf{G}|k\rangle$ are fully determined by the processor mapping $\mathbf{G}$. Of course, the explicit form of these program states strongly depends on the particular processor $\mathbf{G}$.
Note. Any one-dimensional projection $\mathbf{Q} = |Q\rangle\langle Q|$ can be represented by a constant matrix $Q_{jj'} = 1/N$. It is always possible to find a basis $\{|k\rangle\}$ in which the vector $|Q\rangle$ takes the form $|Q\rangle = \frac{1}{\sqrt{N}}\sum_k |k\rangle$. Therefore, our above discussion is valid for a general one-dimensional projection. However, the operators $\mathbf{A}_{jk}$ must be defined with respect to this new basis.

5.5.1 Probability of success

Consider a general U-processor $\mathbf{G} = \sum_j \mathbf{U}_j \otimes |j\rangle\langle j|$. Let us denote by $|\phi_a\rangle$ the basis in which the measurement $\mathbf{M} = \sum_a \mathbb{1} \otimes |\phi_a\rangle\langle\phi_a|$ is performed. The output state of the data register after the action of the processor $\mathbf{G}$ can be expressed in the following (very convenient) way

$$|\Omega'\rangle_{dp} = \sum_a \left[\sum_j \langle\phi_a|j\rangle\alpha_j \mathbf{U}_j|\psi\rangle_d\right] \otimes |\phi_a\rangle \tag{5.104}$$

$$= \sum_a \mathbf{A}(a)|\psi\rangle_d \otimes |\phi_a\rangle_p \tag{5.105}$$

where the program state is $|\Xi\rangle_p = \sum_j \alpha_j|j\rangle$ and we define the operator $\mathbf{A}(a) := \sum_j \langle\phi_a|j\rangle\alpha_j\mathbf{U}_j$. The probability of success (i.e. measuring the outcome a) then reads

$$p(s) = \langle\psi|\mathbf{A}^\dagger(a)\mathbf{A}(a)|\psi\rangle = ||\sum_j V_{aj}\alpha_j\mathbf{U}_j|\psi\rangle_d||^2 \tag{5.106}$$

$$= \sum_j |V_{aj}|^2|\alpha_j|^2 + \sum_{j,j':j\neq j'} V^*_{aj'}V_{aj}\alpha^*_{j'}\alpha_j\langle\psi|\mathbf{U}^\dagger_{j'}\mathbf{U}_j|\psi\rangle \tag{5.107}$$

where the coefficients $V_{aj} := \langle\phi_a|j\rangle$ form a unitary transition matrix between the two bases. We have divided the success probability into two parts. One of them is clearly ψ-independent and the second one not. However (as we shall see in the subsequent section) it is still possible (in special cases) to extract an independent part from this second term.

Note that the projection $\mathbf{Q} = |\phi_a\rangle\langle\phi_a|$ is one-dimensional. It follows that $(\mathbf{Q})_{jj'} = V^*_{aj}V_{aj'}$. In the specific case when $Q_{jj'} = 1/N$ ($N = \dim\mathcal{H}_p$) (now $|j\rangle$ is the basis in which $\mathbf{A}_{jk} = \delta_{jk}\mathbf{U}_j$) we have that $V_{aj} = 1/\sqrt{N}$ for all j. Note that in this case we project onto the vector $|\phi_a\rangle = |Q\rangle = \frac{1}{\sqrt{N}}\sum_j |j\rangle$. Such choice of the measurement outcome implies that

$$p(s) = \frac{1}{N}\langle\psi|\mathbf{A}^\dagger\mathbf{A}|\psi\rangle \tag{5.108}$$

and the initial state $|\psi\rangle_d$ evolves into the state $|\psi'\rangle_d$ given by

$$|\psi\rangle_d \to |\psi'\rangle_d = \frac{\mathbf{A}|\psi\rangle_d}{\sqrt{{}_d\langle\psi|\mathbf{A}^\dagger\mathbf{A}|\psi\rangle_d}} \tag{5.109}$$

where we used the notation $\mathbf{A} = \sum_j \alpha_j\mathbf{U}_j$ with $\alpha_j = \langle j|\Xi\rangle$. As we have already mentioned before, whenever the realized transformation is unitary, the probability of success becomes ψ-independent. Moreover, in this specific case it is also Ξ-independent and equal $p(s) = 1/N$.

Another choice can be done so that $V_{aj} = \delta_{aj}$. It corresponds to the case when the measurement basis coincides with the basis of elementary programs, i.e. $\{|\phi_a\rangle\} = \{|j\rangle\}$. In this case we obtain ψ-independent success probability

$$p(s) = |\alpha_a|^2 \tag{5.110}$$

and the realized transformation is always

$$|\psi\rangle_d \to |\psi'\rangle_d = \mathbf{U}_a|\psi\rangle_d \tag{5.111}$$

whatever the initial program has been chosen. That is, the program register evolves in the same way irrespective on the choice of the program state. However, the probabilities of success depend on the particular state $|\Xi\rangle_p$.

5.5.2 Example: `CNOT` gate

Consider again the `CNOT` processor. Since it belongs to the intersection of the Y-processors and U-processors, it follows that its action can be expressed as

$$\begin{aligned} |\psi\rangle_d \otimes |+\rangle_p &\to |\psi\rangle_d \otimes |+\rangle & (5.112)\\ |\psi\rangle_d \otimes |-\rangle_p &\to \sigma_z|\psi\rangle_d \otimes |-\rangle & (5.113)\end{aligned}$$

That is, we have the identity (see Eq.(5.38))

$$\mathtt{CNOT} = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \sigma_x = \mathbb{1} \otimes |+\rangle\langle +| + \sigma_z \otimes |-\rangle\langle -| \tag{5.114}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. If we apply the general theory (discussed above) onto the `CNOT` gate, then we find out that the choice of the projection reads $|Q\rangle = |0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$. It is an easy exercise to show that introducing the `CNOT` gate into this general evaluation one reproduces the original example studied by Vidal and Cirac [45]. They showed that in this settings the program states $|\alpha\rangle = \frac{1}{\sqrt{2}}(e^{i\alpha}|0\rangle + e^{-i\alpha}|1\rangle)$ encodes unitary operations $\mathbf{U}_\alpha = exp(i\alpha\sigma_z)$.

However, the basis of the program Hilbert space can be chosen arbitrarily and our results are still applicable. Let us assume the basis $|0\rangle_p, |1\rangle_p$ and the measurement projection $|Q\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle$. What kind of transformations does the `CNOT` gate implement in this settings? Again we have $Q_{jj'} = 1/2$. In such case the operators are given by the relation $\mathbf{A}_j(\Xi) := \langle j|\mathtt{CNOT}|\Xi\rangle$. The direct calculation leads us to operation

$$\mathbf{U} = \frac{1}{\sqrt{2p(s)}}\sum_j \mathbf{A}_j(\Xi) = \frac{\Xi_0 + \Xi_1}{\Xi_0 + \Xi_1}\mathbb{1} = \mathbb{1} \tag{5.115}$$

where the probability of success $p(s) = \frac{1}{2}|\Xi_0 + \Xi_1|^2$ and $\Xi_j := \langle j|\Xi\rangle$. We see that the implemented transformation is always trivial. On the other hand, we have no restriction on the program state $|\Xi\rangle$, but the probability of success $p(s)$ depends on the particular choice.

There can be an interesting question whether for each program state there exists a projection (measurement) such that the probability of success is ψ-independent, i.e. whether each program state can be used to implement a unitary transformation. The above projection onto the vector $|+\rangle$ answer this question in a positive, but trivial way (the implemented transformation is always trivial). Therefore, we will pay our attention to this problem and we would like to know how many transformations can be implemented on condition that the program state is fixed and projections $\mathbf{Q}$ are varied. One can consider general measurement projecting onto the vector $|Q\rangle_p = \cos\eta|+\rangle + e^{i\phi}\sin\eta|-\rangle$. Then the success probability reads

$$\begin{aligned} p(s) &= |a|^2\cos^2\eta + |b|^2\sin^2\eta + \\ &\quad \langle\psi|\sigma_z|\psi\rangle\cos\eta\sin\eta\{ab^*e^{i\phi} + a^*be^{-i\phi}\} \end{aligned} \tag{5.116}$$

where the program system is initially in the state $|\Xi\rangle = a|+\rangle + b|-\rangle$. If we put $a = \cos\xi, b = e^{i\varphi}\sin\xi$, then the requirement of ψ-independence implies $\cos(\phi - \varphi) = 0$, i.e. $\phi - \varphi = \pm\pi/2$, or $\cos\eta = 1$, or $\sin\eta = 0$. Note, that the last two possibilities corresponds to trivial cases of projecting onto vectors

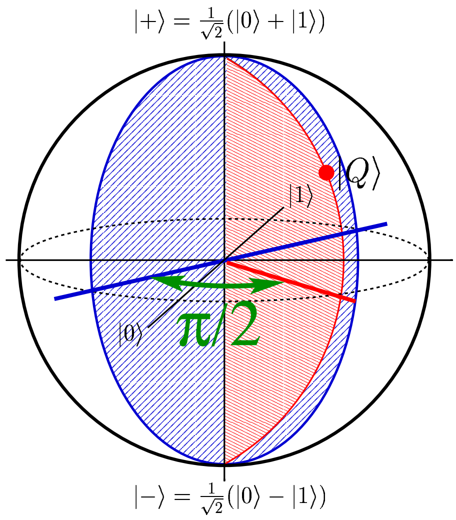
$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$
$|Q\rangle$
$|1\rangle$
$|0\rangle$
$\pi/2$
$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$

Figure 5.3: The figure represents the so-called Bloch sphere, i.e. the graphical representation of the state space of a single qubit. In this case it corresponds to the set of all quantum programs of the `CNOT` processor. You can see the relation between the choice of the projection $|Q\rangle$ and the set of all programs (line around the sphere), which encodes unitary transformations, i.e. $\mathcal{P}_Q$. Obviously, the plane corresponding to allowed programs is perpendicular to the plane containing the projections $|Q\rangle$. In our case the vectors $|\pm\rangle$ plays the role of poles and the points in equatorial determines the value of the relative phases ϕ, φ.

$|\pm\rangle$. As we know, in that case all program states are allowed and all of them encode the same unitary transformation $\mathbb{1}$, or σ_z, respectively.

The condition $\phi - \varphi = \pm\pi/2$ is much more of interest. It says that the programs with the same value of the relative phase φ can encode unitary transformations providing that we choose the projection onto $|Q\rangle = \cos\eta|+\rangle + e^{i(\pi/2+\varphi)}\sin\eta|-\rangle$ with an arbitrary angle η. Hence, for a fixed program state $|\Xi\rangle$ there exists a continuum of projections $\mathcal{M}_\Xi$. On the other hand, if we fix the projection $\mathbf{Q}$ (the relative phase ϕ is fixed), then only those programs $|\Xi\rangle \in \mathcal{P}_Q$ encode some unitary transformations, for which the phase is adjusted according to relation $\phi = \pm\pi/2 + \varphi$. As a result we have found a very simple relation between the projections and allowed programs $\mathcal{P}_Q$, namely

$$p(s) \text{ is } \psi\text{−independent} \Leftrightarrow \phi - \varphi = \pm\pi/2 \tag{5.117}$$

Next we look at the unitary transformations we are implementing. Providing that $p(s)$ is ψ-independent (i.e. $\phi - \varphi = \pm\pi/2$) the implemented unitary transformation has the form

$$\mathbf{U} = \frac{a\cos\eta\mathbb{1} + be^{-i\phi}\sin\eta\sigma_z}{\sqrt{|a|^2\cos^2\eta + |b|^2\sin^2\eta}} = \frac{|a|\cos\eta\mathbb{1} \pm i|b|\sin\eta\sigma_z}{\sqrt{|a|^2\cos^2\eta + |b|^2\sin^2\eta}}$$

where we used $e^{i(\phi-\varphi)} = e^{\pm i\pi/2} = \pm i$. It is clear that for the fixed program state the number of unitaries form a one-parametric continuum parametrized by η determining the projection. Conversely, for a fixed projection $\mathbf{Q}$ the set of allowed programs $\mathcal{P}_Q$ is again one-parametric and characterized by ξ ($|a| = \cos\xi$). However, the probability of success $p(s)$ depends on the choice of the program state $|\Xi\rangle$ except the case when $\cos^2\eta = \sin^2\eta = 1/2$. This case corresponds to the original example derived in [45]. We have shown that this is the only case (of measurement) when the success probability is also Ξ- independent (for all allowed programs). For all other choices of measurement the success $p(s)$ depends on the program state $|\Xi\rangle \in \mathcal{P}_{\mathbf{Q}}$.

5.5.3 U-processors

Let us see how this mechanism might work for a general U-processor. For this type of processors it is possible to find such basis of the program register $\{|j\rangle\}$ in which $\mathbf{A}_{jk} = \delta_{jk}\mathbf{U}_j$ where $\mathbf{U}_j$ are unitary operators. We then have that

$$\mathbf{G}(|\psi\rangle_d \otimes |\Xi\rangle_p) = \sum_j \alpha_j \mathbf{U}_j |\psi\rangle_d \otimes |j\rangle_p. \tag{5.118}$$

with $\alpha_j := \langle j|\Xi\rangle$. In order to complete the procedure we apply the fixed (independent of the program state) projection $\mathbb{1}_d \otimes \mathbf{Q}_p$ to the right-hand side.

We want this processor to implement the one-parametric set of operator $\mathbf{U}_\mu = \sum_{j=1}^{M} z_j(\mu)\mathbf{U}_j$ on $|\psi\rangle_d$, for some coefficients $z_j(\mu)$. It implies that the following equality holds

$$\frac{1}{\sqrt{p(s)}}(\mathbb{1}_d \otimes \mathbf{Q}_p)G(|\psi\rangle_d \otimes |\Phi_\mu\rangle_p) = \mathbf{U}_\mu|\psi\rangle_d \otimes |v_\mu\rangle_p, \tag{5.119}$$

where $|v_\mu\rangle_p$ is a unit vector lying in the range of $\mathbf{Q}_p$ and

$$p(s) = \|\sum_{j=1}^{N} \kappa_j(\mu)\mathbf{U}_j|\psi\rangle\|^2 \tag{5.120}$$

is the success probability of carrying out the transformation. Let us remind that $\kappa_j(\mu) = \langle\phi_a|j\rangle\langle j|\Xi_\mu\rangle$ and we put $\mathbf{Q} = |\phi_a\rangle\langle\phi_a|$. Since $\mathbf{Q}$ is one-dimensional the vectors $|v_\mu\rangle$ and $|Q\rangle$ coincide.

Two-dimensional program space

It is useful at this point to look at a simple generalization of the previous example. Suppose that we have a U processor where the program space is two dimensional. The vectors $|t\rangle_p$ and $|u\rangle_p$ are orthonormal program vectors, and $|t\rangle_p$ causes the unitary operation $\mathbf{T}$ to be performed on the data state, while $|u\rangle_p$ causes the unitary operation $\mathbf{U}$ to be performed. The action of our processor is then

$$\mathbf{G}(|\psi\rangle_d \otimes |\Phi\rangle_p) = \mathbf{T}|\psi\rangle_d \otimes |t\rangle_p\langle t|\Phi\rangle + \mathbf{U}|\psi\rangle_d \otimes |u\rangle_p\langle u|\Phi\rangle. \tag{5.121}$$

The μ dependence of Φ is understood, but not explicitly indicated. In this case our projection operator can only be one-dimensional, so we set $\mathbf{Q}_p = |w_\perp\rangle_p\,{}_p\langle w_\perp|$, for some unit vector $|w_\perp\rangle_p$ in the program space. Note that if $|w\rangle_p$ is the unit vector orthogonal to $|w_\perp\rangle_p$, then the projection onto $|w\rangle_p$ corresponds to failure. Our object is to take the input data state, $|\psi\rangle_d$, and to create the output state $(\cos\mu\mathbf{T} + \sin\mu\mathbf{U})|\psi\rangle_d$ (we assume, that $\cos\mu\mathbf{T} + \sin\mu\mathbf{U}$ is an unitary transformation). Looking at Eq. (5.119), we see that we will accomplish our goal if

$$\mathbf{Q}_p|u\rangle_p\langle u|\Phi\rangle = z\mathbf{Q}_p|t\rangle_p\langle t|\Phi\rangle \tag{5.122}$$

where we use the definition $z = \tan\mu$. Since $\mathbf{Q}_p = |w_\perp\rangle\langle w_\perp|$ the above equation implies that the vector $|w\rangle$ must be parallel to the vector

$$|\xi\rangle_p = |u\rangle_p\langle u|\Phi\rangle - z|t\rangle_p\langle t|\Phi\rangle, \tag{5.123}$$

From this we can infer that

$$\langle t|\Phi\rangle = -\frac{\langle t|w\rangle}{z\langle u|w\rangle}\langle u|\Phi\rangle. \tag{5.124}$$

The normalized program vector, $|\Phi\rangle_p$, can then be written as

$$|\Phi\rangle_p = \left[\frac{1}{|\langle u|w\rangle|^2 + (|\langle t|w\rangle|/|z|)^2}\right]^{1/2}(|u\rangle_p\langle u|w\rangle - \frac{1}{z}|t\rangle_p\langle t|w\rangle). \tag{5.125}$$

The failure probability, p_f, is just the square of the norm of the projection onto $|w\rangle_p$ of the output vector of the processor (see Eq. (5.121)), or we can insert our results into the general formula (5.120) to obtain

$$\begin{aligned} p(f) &= \sum_{j=u,t} |\langle j|\Phi\rangle|^2 |\langle j|w\rangle|^2 \\ &\quad + \langle u|\Phi\rangle\langle\Phi|t\rangle\langle t|w\rangle\langle w|u\rangle\langle\psi|\mathbf{T}^\dagger\mathbf{U}|\psi\rangle \\ &\quad + \langle t|\Phi\rangle\langle\Phi|u\rangle\langle u|w\rangle\langle w|t\rangle\langle\psi|\mathbf{U}^\dagger\mathbf{T}|\psi\rangle \end{aligned} \tag{5.126}$$

Introducing the relation (5.124) the failure probability can be rewritten as follows

$$p(f) = \sum_{j=u,t} |\langle j|\Phi\rangle|^2 |\langle j|w\rangle|^2 - z|\langle\Phi|t\rangle|^2 |\langle u|w\rangle|^2 \langle\psi|\mathbf{T}^\dagger\mathbf{U} + \mathbf{U}^\dagger\mathbf{T}|\psi\rangle \tag{5.127}$$

To make this probability independent of the initial state of the data system, we need to choose the operators $\mathbf{U}, \mathbf{T}$ in a way that either $\mathbf{T}^\dagger\mathbf{U} + \mathbf{U}^\dagger\mathbf{T} = 0$, or $\mathbf{T}^\dagger\mathbf{U} + \mathbf{U}^\dagger\mathbf{T} = \mathbb{1}$.

Let us now make such a specific choice of operators $\mathbf{T} = \vec{n}.\vec{\sigma}, \mathbf{U} = \vec{m}.\vec{\sigma}$, such that $\mathbf{T}^\dagger\mathbf{U} + \mathbf{U}^\dagger\mathbf{T} = 0$. In this case it is easy to see that $\mathbf{T}^\dagger\mathbf{U} + \mathbf{U}^\dagger\mathbf{T} = 2\vec{m}.\vec{n}\mathbb{1}$. This expression vanishes whenever $\vec{n} \perp \vec{m}$ and in that case the performed transformation is unitary. The probability of success (failure) is independent of the state of data system. Note that different choices of $\mathbf{T}$ and $\mathbf{U}$ require different processors. The implemented operations are $\cos\mu(\vec{n}.\vec{\sigma}) + \sin\mu(\vec{m}.\vec{\sigma})$.

The failure probability is obviously ψ-independent, but is it possible to make it also μ-independent? Let us denote $a := |\langle u|w\rangle|^2, b := |\langle u|\Phi\rangle|^2$, then the equation (5.124) and the normalization of $|\Phi\rangle, |w\rangle$ implies

$$1 - b = \frac{1-a}{az^2} b \tag{5.128}$$

The probability of failure reads

$$p(f) = |\langle u|\Phi\rangle|^2 |\langle u|w\rangle|^2 + |\langle t|\Phi\rangle|^2 |\langle t|w\rangle|^2 = ba + (1-b)(1-a)$$

It follows that we can express b as a function of a, i.e. $b = az^2/(1 + a(z^2 - 1))$. It corresponds to a relation between projections and programs (cf. section about CNOT). Note that the parameter a represents the choice of the measurement projection, whereas the parameter b is associated with the state of the program register. Introducing this expression into the above equation for the failure probability we obtain

$$p(f) = \frac{a^2z^2 + (1-a)^2}{1 + a(z^2 - 1)} \tag{5.129}$$

That is, in general $p(f)$ depends on the value of the parameter z (μ). Only with a specific choice of projections, when $a = |\langle u|w\rangle|^2 = 0, 1, 1/2$ (in this cases $b = 0, 1, z^2/(1 + z^2)$) the probability of failure (success, as well) becomes to be μ-independent. In the mentioned cases it takes values $p(f) = 1, 1, 1/2$ (respectively). But for instance, in the case $a = 1/3$ we have $p(f) = \frac{1}{3}\frac{z^2+4}{z^2+2}$. For us the result, when $a = 1/2$ is of interest, because then the set of performed transformations is largest (in comparison with choices $a = 0, 1$). In fact, if $a = 0, 1$ then for all program states the processor performs either $\mathbf{T}$, or $\mathbf{U}$ (respectively).

Let us note that the overlap between program states and basis state $b = |\langle u|\Phi\rangle|^2$ depends on the parameter z (or equiv. μ). This behavior is nothing unexpected. Indeed, it is obvious that the program $|\Phi\rangle$ encoding the unitary transformation $\cos\mu\mathbf{T} + \sin\mu\mathbf{U}$ should depend on the parameter μ. We remind us, that due to the more convenient notation we omitted to label program vector by the index μ. On the other hand if we fix a program state $|\Phi\rangle$ (i.e. value b), then the set of all possible transformations will depend on the measurements we perform. And the μ-dependency will be hidden in the choices of projections $\mathbf{Q}$.

We have shown another example of probabilistic implementation of quantum processor that can be used to implement one-parametric set of unitaries $\mathbf{U}_\mu = \cos\mu\mathbf{T} + \sin\mu\mathbf{U}$ with probability $p(s) = 1/2$. The program states have the following explicit form

$$|\Phi_\mu\rangle_p = \cos\mu|t\rangle + \sin\mu|u\rangle \tag{5.130}$$

and the successful implementation is associated with the projection onto the vector

$$|w_\perp\rangle = \frac{1}{\sqrt{2}}(|u\rangle + |t\rangle) \tag{5.131}$$

The verification of this result is straightforward. Note that this one-parametric set do not possess group properties. That is, we have shown different example of the implementation of one-parametric set of unitary transformations as it was in CNOT case. It is easy to see (by following the same steps) that the second possibility, when $\mathbf{T}\mathbf{U}^\dagger + \mathbf{U}\mathbf{T}^\dagger = \mathbb{1}$, does not result in the success probability independent of the parameter μ.

5.5.4 Quantum distributor machine

In this section we shall investigate another important type of quantum processor which can be used like universal NOT machine, or universal COPY machine [46]. We shall not define the processor mapping $\mathbf{G}_{dp}$ directly by unitary transformation, but as the sequence of specific (elementary) quantum gates. We must first introduce the basis element of our network called **generalized** CNOT. Originally, the CNOT gate is defined for two qubits by the formula

$$\mathbf{D}_{ab} = \sum_{k,m=0}^{1} |k\rangle_a\langle k| \otimes |m \oplus k\rangle_b\langle m|. \tag{5.132}$$

where a, b labels the pair of qubits on which the gate is applied. In principle, one can also introduce an operator $\mathbf{D}_{ab}^\dagger$ defined as

$$\mathbf{D}_{ab}^\dagger = \sum_{k,m=0}^{1} |k\rangle_a\langle k| \otimes |m \ominus k\rangle_b\langle m|. \tag{5.133}$$

In the case of qubits these two operators are equal, but this will not be the case when we generalize the CNOT operation to Hilbert spaces whose dimension is larger than 2. In particular, we can generalize the operator $\mathbf{D}$ for dimension $N \equiv \dim\mathcal{H}_d > 2$ by defining

$$\mathbf{D}_{ab} = \sum_{k,m=0}^{N-1} |k\rangle_a\langle k| \otimes |(m+k)\mathrm{mod}\,N\rangle_b\langle m|, \tag{5.134}$$

which implies that

$$\mathbf{D}_{ab}^\dagger = \sum_{k,m=0}^{N-1} |k\rangle_a\langle k| \otimes |(m-k)\mathrm{mod}\,N\rangle_b\langle m|. \tag{5.135}$$

From this definition it follows that the operator $\mathbf{D}_{ab}$ acts on the basis vectors as

$$\mathbf{D}_{ab}|k\rangle|m\rangle = |k\rangle|(k+m)\mathrm{mod}\,N\rangle \tag{5.136}$$

Now we see that for $N > 2$ the two operators $\mathbf{D}$ and $\mathbf{D}^\dagger$ do differ; they describe conditional shifts in opposite directions. Therefore the generalizations of the CNOT operator to higher dimensions are just **conditional shifts**.

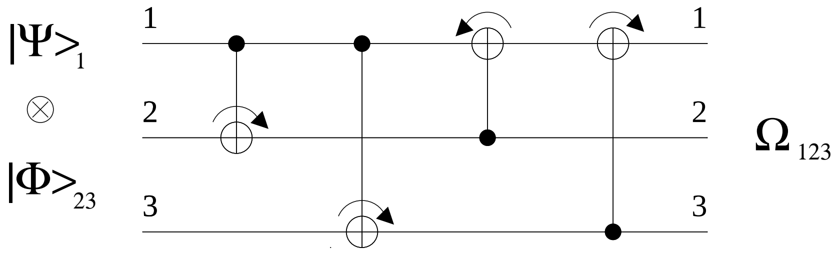

Figure 5.4: A logic network for the universal quantum processor as given by the unitary transformation (5.138).

In accordance with the definition of quantum computational network (quantum processor) discussed in the previous section, we assume the network for the quantum processor to be[2]

$$\mathbf{G}_{dp} = \mathbf{G}_{123} = \mathbf{D}_{31}\mathbf{D}_{21}^{\dagger}\mathbf{D}_{13}\mathbf{D}_{12} \ . \tag{5.137}$$

The data register consists of system 1 and the program register of systems 2 and 3. Name the physical representatives of N dimensional Hilbert spaces by **qudits**. The output state of the three qudit system, after the four conditional shifts are applied, reads

$$|\Omega\rangle_{123} = \mathbf{D}_{31}\mathbf{D}_{21}^{\dagger}\mathbf{D}_{13}\mathbf{D}_{12}|\Psi\rangle_1|\Xi\rangle_{23} \ . \tag{5.138}$$

A graphical representation of the logical network (5.138) with the conditional shift gates $\mathbf{D}_{ab}$ in Fig. 5.4.

Our first aim is to investigate the type of this quantum processor. The sequence of four operators acting on the basis vectors gives $|n\rangle_1|m\rangle_2|k\rangle_3$ as

$$\mathbf{G}_{123}|n\rangle_1|m\rangle_2|k\rangle_3 = |(n-m+k)\mathrm{mod}\,N\rangle_1\;|(m+n)\mathrm{mod}\,N\rangle_2\;|(k+n)\mathrm{mod}\,N\rangle_3 \ .$$

We now turn to the fundamental program states. A basis consisting of maximally entangled two-particle states (the analogue of the Bell basis for spin-$\frac{1}{2}$ particles) is given by

$$|\Xi_{mn}\rangle = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}\exp\Big(i\frac{2\pi}{N}mk\Big)|k\rangle|(k-n)\mathrm{mod}\,N\rangle\,, \tag{5.139}$$

where $m,n = 0,\dots,N-1$. If $|\Xi_{mn}\rangle_p$ is the initial state of the program register, and $|\Psi\rangle = \sum_j \alpha_j|j\rangle_d$ (here, as usual, $\sum_j|\alpha_j|^2 = 1$) is the initial state of the data register, then follows that

$$\begin{aligned}\mathbf{G}_{123}|\Psi\rangle_1|\Xi_{mn}\rangle_{23} &= \sum_{jk}\frac{\alpha_j}{\sqrt{N}}\exp\frac{2\pi ikm}{N}\mathbf{G}_{123}|j\rangle|k\rangle|k-n\rangle \\ &= \sum_{jk}\frac{\alpha_j}{\sqrt{N}}\exp\frac{2\pi ikm}{N}|j-n\rangle|k+j\rangle|k+j-n\rangle \\ &= \sum_{jk}\alpha_j\exp\frac{-2\pi ijm}{N}|j-n\rangle|\Xi_{mn}\rangle \\ &= (\mathbf{U}^{(mn)}|\Psi\rangle)|\Xi_{mn}\rangle,\end{aligned} \tag{5.140}$$

[2] the choice is supported by [31, 32]

where we have introduced the notation

$$\mathbf{U}^{(mn)} = \sum_{s=0}^{N-1} \exp \frac{-2i\pi sm}{N} |s-n\rangle\langle s|. \tag{5.141}$$

The last equality implies that $\mathbf{G}_{123}$ is the $U-processor$. It implements the set of unitary operators $\mathbf{U}^{(mn)}$, which form an orthogonal basis of Hilbert-Schmidt operators, because

$$\text{Tr}\left[(U^{(m'n')})^{\dagger}U^{(mn)}\right] = N\delta_{mm'}\delta_{nn'}. \tag{5.142}$$

The programs realizable deterministically are in some sense generalizations of *Pauli superoperators* for qudits. In fact, in the case of qubits we get $\mathbf{U}_{(00)} = \mathbb{1}, \mathbf{U}^{(01)} = \sigma_1, \mathbf{U}^{(10)} = \sigma_2, \mathbf{U}^{(11)} = \sigma_3$ and the corresponding set of realizable superoperators $\mathcal{C}_P = \mathcal{P}(p_1, p_2, p_3)$ is the set of *Pauli superoperators* (see previous chapter).

Now we would like to examine which transformations we can perform on the state in the data register by using a program consisting of a linear combination of the vectors $|\Xi_{mn}\rangle$ followed by the action of the processor $\mathbf{G}_{123}$ and a subsequent measurement $\mathbf{X}_{dp} = \mathbb{1}_d \otimes \mathbf{X}_p$ of the program register. Any operator $\mathbf{A} \in \mathcal{T}_2(\mathcal{H})$ can be expressed in terms of them

$$\mathbf{A} = \sum_{m,n=0}^{N-1} q_{mn}\mathbf{U}^{(mn)} \tag{5.143}$$

with $q_{mn} = \frac{1}{N}\text{Tr}\left[\left(\mathbf{U}^{(mn)}\right)^{\dagger}\mathbf{A}\right]$. Equations (5.142) and (5.143) imply that

$$\sum_{m,n=0}^{N-1} |q_{mn}|^2 = \frac{1}{N}\text{Tr}(\mathbf{A}^{\dagger}\mathbf{A}) = 1. \tag{5.144}$$

The most general program vector is given by

$$|v_{\mathbf{A}}\rangle_{23} = \sum_{m,n=0}^{N-1} q_{mn}|\Xi_{mn}\rangle_{23}. \tag{5.145}$$

The subscript $\mathbf{A}$ indicates that this state could be used for the implementation of the operation $\mathbf{A}$. Application of the processor to the input state $|\Psi\rangle_1|v_{\mathbf{A}}\rangle_{23}$ yields the output state

$$|\Omega\rangle_{123} = \sum_{mn} q_{mn}\mathbf{U}^{(mn)}|\Psi\rangle_1 \otimes |\Xi_{mn}\rangle_{23}. \tag{5.146}$$

To obtain the final result we perform a projective measurement $\mathbf{Q}$ of the program register onto the vector $|Q\rangle_{23}$ with the property $Q_{jj'} = 1/N^2$, i.e.

$$|Q\rangle_{23} = \frac{1}{N}\sum_{m,n=0}^{N-1} |\Xi_{mn}\rangle_{23} \tag{5.147}$$

If the outcome of the measurement is positive, then we get the required transformation $\mathbf{A} = \sum_{mn} q_{mn}\mathbf{U}^{(mn)}$ acting on an unknown, arbitrary input state $|\Psi\rangle_1$ and the probability $p(s)$ of achieving the desired result is the same as the probability of successfully implementing the transformation $\mathbf{A}$

$$p(s) = \frac{1}{N^2}\langle\Psi|\mathbf{A}^{\dagger}\mathbf{A}|\Psi\rangle \tag{5.148}$$

The derived theorem implies that the transformation $\mathbf{A}$ is unitary only if the probability of success is ψ-independent. Otherwise the data register evolves in the way

$$|\Psi\rangle \to |\Psi'\rangle = \frac{\mathbf{A}|\Psi\rangle}{\sqrt{\langle\Psi|\mathbf{A}^\dagger\mathbf{A}|\Psi\rangle}} \tag{5.149}$$

Hence, only those program states are allowed, for which the operation $\mathbf{A} = \sum_{mn} q_{mn}\mathbf{U}^{(mn)}$ is unitary. In this case the probability of success is program independent and equals $1/N^2$. Otherwise, the data register transforms in a way determined by the operation $\mathbf{A}$, but the transformation is Ψ-dependent in non-linear way and the probability of its success as well.

Nevertheless, as a result we obtain that we are able to perform any unitary transformation of the data register with the probability equals $1/N^2$. This result follows from the fact that the each unitary transformation can be written via the operators $\mathbf{U}_{mn}$ with the coefficients satisfying the normalization condition $\sum_{mn}|q_{mn}|^2 = 1$. Therefore, there exists state $|v_\mathbf{A}\rangle_{23}$ that encodes such transformation. Moreover, we obtained that the success probability is independent of the choice of such unitary transformation. Consequently, universal probabilistic quantum processor does exist.

5.6 Probability amplification

In our previous discussion we have introduced one specific example of the measurement $\mathbf{Q}$, for which the success probability is Ξ-independent for all program states in $\mathcal{P}_Q$. The transformation encoded in the state $|\Xi\rangle$ is realized with $p(s) = 1/N$, where N is the dimension of the program Hilbert space $\mathcal{H}_p$. If we restrict the set of unitary operators (i.e. states $|\Xi\rangle$) we would like to implement, then it is possible to beat this general limit and increase $p(s)$. We know that input states $|\Xi\rangle = \sum_{j=1}^D \alpha_j|j\rangle$ ($D < N$) encode transformations $U = \sum_{j=1}^D \alpha_j\mathbf{U}_j$. However, since the program space is effectively only D dimensional, we can use a projection $|Q\rangle = \frac{1}{\sqrt{D}}\sum_{j=1}^D |j\rangle$ to realize these transformations. It means that the probability of success is amplified to $p(s) = 1/D$.

Is it possible to make the probability much better? Till now the projection associated with the successful implementation were always one-dimensional and therefore the probability was limited by the dimension of program register. Because the dimension is a natural number, it follows that the success probability is always less than 1/2, i.e. $p(s) \leq 1/2$. To increase the probability we need to associate a multi-dimensional projection $\mathbf{Q}$ with the successful realization. As we shall see the generalization of the previous investigation is not so straightforward and not so trivial as it may seem. In particular, we shall see that in the case of multi-dimensional (in terms of program Hilbert space) projections $\mathbf{Q}$ the resulting state of the data register may not be pure and the realized transformations cannot be associated with linear operators on data Hilbert space $\mathcal{H}_d$. Note, that this result is not contradictory to the proven fact that in von Neumann-Lüders realizations of POVMs pure states transform into pure states. Simply, in this case the projective measurement performed on joint data+program system does not induce a von Neumann-Lüders measurement. Only, if the measurement can be represented by a non-degenerate selfadjoint operator, the induced data measurement is of this type.

Let us see the reasons for such statement. The joint state is transformed into an unnormalized state

$$|\psi\rangle \otimes |\Xi\rangle \to |\Omega_a\rangle = \sum_m \alpha_m \mathbf{A}_m(\Xi)|\psi\rangle \otimes |\phi_m\rangle \tag{5.150}$$

where $\mathbf{A}_m(\Xi) = \langle\phi_m|\mathbf{G}|\Xi\rangle$ and $\alpha_m = \langle\phi_m|\Xi\rangle$. We use a basis in which the measurement is performed, i.e. $\mathbf{Q}_a = \sum_m |\phi_m\rangle\langle\phi_m|$ and $m = 1, \ldots K_a < \dim\mathcal{H}_p$ and $\sum_a K_a = \dim\mathcal{H}_p$. That is, $\mathbf{Q}_a$ is K_a-dimensional projection. As a result we get a superposition of pure factorized states, i.e. potentially entangled state $|\Omega_a\rangle$. Only if $\mathbf{A}_m = \mathbf{A}$ for all m, the resulting state $|\Omega_a\rangle$ is factorized and the data system evolves into a pure state. But if not, then tracing out the program register leads us to a mixed output data state

$$\varrho_d' = \frac{\sum_m |\alpha_m|^2 \mathbf{A}_m(\Xi)\varrho_d \mathbf{A}_m^\dagger(\Xi)}{\mathrm{Tr}\varrho\mathbf{F}_a} \tag{5.151}$$

with $\mathbf{F}_a = \sum_m |\alpha_m|^2 \mathbf{A}_m^\dagger(\Xi)\mathbf{A}_m(\Xi)$. In this case the POVM operators $\mathbf{F}_a$ are well defined, but the state transformation is not associated with an linear operator on the data Hilbert space $\mathcal{H}_d$, but rather with a map $\mathcal{M}_a$

$$\mathcal{M}_a[\varrho] = \sum_m \mathbf{M}_m \varrho \mathbf{M}_m^\dagger \tag{5.152}$$

where $\mathbf{M}_m = \alpha_m \mathbf{A}_m(\Xi)/\sqrt{\mathrm{Tr}\varrho\mathbf{F}_a}$. That is, generally the measurement can be represented by a set of maps $\mathcal{M}_a$ such that a convex combination of these maps, i.e. $\mathcal{M} = \sum_a p_a \mathcal{M}_a$, is a trace-preserving completely positive linear transformation. Note that the particular map $\mathcal{M}_a$ is linear only if the probability $p_a = \mathrm{Tr}\varrho\mathbf{F}_a$ is ϱ-independent. To summarize, let us formulate conditions under which the transformation $\mathcal{M}_a$ is unitary. Obviously,the following equality must hold

$$(\mathbb{1} \otimes \mathbf{Q}_a)\mathbf{G}(|\psi\rangle \otimes |\Xi\rangle) = |\psi'\rangle \otimes |\Xi'\rangle \tag{5.153}$$

where $|\psi'\rangle = \mathbf{U}|\psi\rangle$. As a result we have found that for a multi-dimensional projection the implemented operations cannot be unitary except the case when $\mathbf{A}_m = \mathbf{A}$ for all m. Therefore the aim is to find such decomposition of the projection, in which this condition holds. This is equivalent to introduce one-dimensional projections for which the realized transformations are the same.

5.6.1 Case studies

I. Cirac and Vidal proposal

What we would like to do is to find a way of increasing the probability of success. In order to get an idea of how this can be accomplished, first we shall examine a method due to Vidal and Cirac. They were interested in implementation of one parameter group of transformations, $\mathbf{U}(\alpha) = \exp(i\alpha\sigma_3)$ on a single qubit. They first do this using a single `CNOT` gate, the data qubit being the control and the program the target. The program state is

$$|\alpha\rangle = \frac{1}{\sqrt{2}}(e^{i\alpha}|0\rangle + e^{-i\alpha}|1\rangle), \tag{5.154}$$

and the probability of success is 1/2. This example was described in the section concerning the `CNOT` gate. In order to increase the probability of performing $\mathbf{U}(\alpha)$ they consider a more complicated network consisting of three qubits. Qubit 1 is the data qubit, and qubits 2 and 3 are the program qubits. The network consists of two gates, first the `CNOT` gate, with qubit 1 as the control and qubit 2 as the target, followed by a `TOFFOLI` gate, with qubits 1 and 2 as controls and qubit 3 as the target, i.e.

$$\mathbf{G} = (\mathtt{TOFFOLI})(\mathtt{CNOT} \otimes \mathbb{1}) = |0\rangle\langle 0| \otimes \mathbb{1}_p + |1\rangle\langle 1| \otimes \mathbf{S}_p \tag{5.155}$$

where $S_p = |0\rangle\langle 1| \otimes \mathbb{1} + |1\rangle\langle 0| \otimes \sigma_x$. It is easy to see that $\mathbf{S}_p$ is a unitary operator and therefore $\mathbf{G}$ is the *Y-processor.* This network is able to implement $\mathbf{U}(\alpha)$ with probability 3/4. It turns out that this is also the U-processor. This can be seen, if we consider the program vectors

$$|\Xi_z\rangle = \frac{1}{2}[|00\rangle + z^2|01\rangle + z|10\rangle + z^3|11\rangle] \tag{5.156}$$

where $z = \pm 1, \pm i$. After a little algebra one can find that

$$\mathbf{G} = \sum_{z=\pm 1,\pm i} \mathbf{B}(z) \otimes |\Xi_z\rangle\langle\Xi_z| \tag{5.157}$$

where $\mathbf{B}(z)$ are unitary operators of the form $\mathbf{B}(z) := |0\rangle\langle 0| + z|1\rangle\langle 1|$.

It is easy to check that among the four unitary transformations $\mathbf{B}(z)$ only two of them are linearly independent. In particular, we can write

$$\mathbf{B}(z) = \frac{1+z}{2}\mathbb{1} + \frac{1-z}{2}\sigma_z \tag{5.158}$$

Moreover these four operators form an algebra closed in multiplication and adjoint operation, i.e. $\mathbf{B}(z)\mathbf{B}(z') = \mathbf{B}(zz')$ and $\mathbf{B}(z)^\dagger = \mathbf{B}(z^*)$. It follows that these four operators form a group, since $\mathbf{B}(z)^{-1} = \mathbf{B}(z)^\dagger = \mathbf{B}(z^*)$. As a result we obtain that the four operators $B(z)$ are linearly dependent and commute. This suggests that U-processors in which the operators are linearly dependent or commute (i.e. belong to the intersection of U and Y processors) can be used to boost the probability of performing certain sets of operations.

Let us see how their example works in more details. Consider the initial state of three qubits

$$|\psi\rangle_d \otimes |\alpha\rangle \otimes |2\alpha\rangle = |\psi\rangle \otimes \frac{1}{2}\left(e^{i3\alpha}|00\rangle + e^{-i\alpha}|01\rangle + e^{i\alpha}|10\rangle + e^{i3\alpha}|-11\rangle\right) \tag{5.159}$$

In the $\{|\Xi_a\rangle\}$ basis the program vector state is given as $|\Xi\rangle = \sum_a \alpha_z|\Xi_z\rangle$ with coefficients

$$\begin{aligned}
\alpha_1 = \langle\Xi|\Xi_1\rangle &= \frac{1}{4}(e^{i3\alpha} + e^{-i\alpha} + e^{i\alpha} + e^{i3\alpha}) \\
\alpha_{-1} = \langle\Xi|\Xi_{-1}\rangle &= \frac{1}{4}(e^{i3\alpha} + e^{-i\alpha} - e^{i\alpha} - e^{i3\alpha}) \\
\alpha_i = \langle\Xi|\Xi_i\rangle &= \frac{1}{4}(e^{i3\alpha} - e^{-i\alpha} - ie^{i\alpha} + ie^{i3\alpha}) \\
\alpha_{-i} = \langle\Xi|\Xi_{-i}\rangle &= \frac{1}{4}(e^{i3\alpha} - e^{-i\alpha} + ie^{i\alpha} - ie^{i3\alpha})
\end{aligned}$$

After applying the processor action $\mathbf{G}$ followed by the measurement projecting onto vectors $|00\rangle, |01\rangle, , |10\rangle, |11\rangle$ the resulting states are given as follows

$$\begin{aligned}
\mathbb{1}\otimes\mathbf{Q}_{00} &\rightarrow \frac{1}{2}([\alpha_1\mathbf{B}(1) + \alpha_{-1}\mathbf{B}(-1) + \alpha_i\mathbf{B}(i) + \alpha_{-i}\mathbf{B}(-i)]|\psi\rangle)\otimes|00\rangle = \frac{e^{i2\alpha}}{2}\mathbf{U}_\alpha|\psi\rangle\otimes|00\rangle \\
\mathbb{1}\otimes\mathbf{Q}_{01} &\rightarrow \frac{1}{2}([\alpha_1\mathbf{B}(1) + \alpha_{-1}\mathbf{B}(-1) - \alpha_i\mathbf{B}(i) - \alpha_{-i}\mathbf{B}(-i)]|\psi\rangle)\otimes|01\rangle = \frac{e^{-i2\alpha}}{2}\mathbf{U}_\alpha|\psi\rangle\otimes|01\rangle \\
\mathbb{1}\otimes\mathbf{Q}_{10} &\rightarrow \frac{1}{2}([\alpha_1\mathbf{B}(1) - \alpha_{-1}\mathbf{B}(-1) + i\alpha_i\mathbf{B}(i) - i\alpha_{-i}\mathbf{B}(-i)]|\psi\rangle)\otimes|10\rangle = \frac{1}{2}\mathbf{U}_\alpha|\psi\rangle\otimes|10\rangle \\
\mathbb{1}\otimes\mathbf{Q}_{11} &\rightarrow \frac{1}{2}([\alpha_1\mathbf{B}(1) - \alpha_{-1}\mathbf{B}(-1) - i\alpha_i\mathbf{B}(i) + i\alpha_{-i}\mathbf{B}(-i)]|\psi\rangle)\otimes|11\rangle = \frac{1}{2}\mathbf{U}_{-3\alpha}|\psi\rangle\otimes|11\rangle
\end{aligned}$$

Consequently, if we realize a multi-dimensional projection $\mathbf{Q} = \mathbf{Q}_{00} + \mathbf{Q}_{01} + \mathbf{Q}_{10}$ the data register evolves into the state $\mathbf{U}_\alpha|\psi\rangle$ and this result occurs with the probability $p(s) = 3/4$, i.e. the probability of the implementation of one-parametric set of unitaries $\mathbf{U}_\alpha$ has been amplified.

The same amount of the amplification can be obtained also in a different way. Due to the fact that using the `CNOT` gate the wrong outcome results in the state $\mathbf{U}_{-\alpha}|\psi\rangle$, it is possible to correct the wrong result by using the `CNOT` processor again, but with a new program register prepared in the state $|2\alpha\rangle$. In this case the probability of success can be calculated in the following way $p(s) = \frac{1}{2} + \frac{1}{2}.\frac{1}{2} = \frac{3}{4}$. This *feedback probabilistic scenario* uses always the same processor, whereas the previous way of amplification requires the usage of a new processor. Cirac and Vidal shows that both these scenarios can be used to further amplification and lead us to probability $p = 1 - 1/2^n$ where n is the number of program qubits. If $n \to \infty$ the probabilistic implementation started to be deterministic. On the other hand, it is known that infinite number of qubits is associated with an inseparable (=unphysical) Hilbert space. That is, there is no contradiction with the proven fact that no continuous set of unitaries can be implemented with a single quantum processor in a deterministic manner. However, the result that amplification can be done arbitrarily close to unity is of importance. It will be nice to generalize this result for any continuous set of unitary transformations.

5.6.2 Conditioned loops

The first aim of our study will be the question whether for all single qubit rotations the feedback control (*conditioned loops*) results in the probability amplification. We have shown that any single qubit unitary transformation can be implemented with the probability $p = 1/4$ by using quantum distributor machine (`QDM`) as the processor. In this case the success is associated with one-dimensional projection onto the vector $|0\rangle|+\rangle$. In what follows we shall explicitly show how to correct the cases of wrong results, i.e. projections onto one of the vectors $|0\rangle|-\rangle, |1\rangle|+\rangle, |1\rangle|-\rangle$.

The `QDM` processor is given by relation

$$\mathbf{G} = \sum_{j=0}^{3} \sigma_j \otimes |\Xi_j\rangle\langle\Xi_j| \tag{5.160}$$

where σ_j are standard σ-matrices with $\sigma_0 = \mathbb{1}$. The basis program vectors $|\Xi_j\rangle$ form the standard Bell basis, i.e.

$$\begin{aligned} |\Xi_0\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \qquad & |\Xi_x\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ |\Xi_z\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \qquad & |\Xi_y\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \end{aligned}$$

The general program state $|\Xi(\vec{\mu})\rangle_p$ encoding the unitary transformation $\mathbf{U}_{\vec{\mu}} = \exp(i\vec{\mu}.\vec{\sigma}) = \cos\mu\mathbb{1} + i\sin\mu\frac{\vec{\mu}}{\mu}.\vec{\sigma}$ ($\mu = |\vec{\mu}|$) is given by

$$|\Xi(\vec{\mu})\rangle_p = \cos\mu|\Xi_0\rangle + i\frac{\sin\mu}{\mu}(\mu_x|\Xi_x\rangle + \mu_y|\Xi_y\rangle + \mu_z|\Xi_z\rangle) \tag{5.161}$$

All other vectors are irrelevant for probabilistic implementation. Performing the mentioned measurement in the program basis $|0+\rangle, |0-\rangle, |1+\rangle, |1-\rangle$ results in the ensuing unitary transformations

$$|0\rangle \otimes |+\rangle \quad : \quad |\psi\rangle_d \to [\cos\mu\mathbb{1} + \frac{i\sin\mu}{\mu}(\mu_x\sigma_x + \mu_y\sigma_y + \mu_z\sigma_z)]|\psi\rangle_d = \mathbf{U}_{\vec{\mu}}|\psi\rangle_d \tag{5.162}$$

$$|0\rangle \otimes |-\rangle \quad : \quad |\psi\rangle_d \to [\cos\mu\mathbb{1} + \frac{i\sin\mu}{\mu}(-\mu_x\sigma_x - \mu_y\sigma_y + \mu_z\sigma_z)]|\psi\rangle_d = \sigma_z\mathbf{U}_{\vec{\mu}}\sigma_z|\psi\rangle_d \tag{5.163}$$

$$|1\rangle \otimes |+\rangle \quad : \quad |\psi\rangle_d \to [\cos\mu\mathbb{1} + \frac{i\sin\mu}{\mu}(\mu_x\sigma_x - \mu_y\sigma_y - \mu_z\sigma_z)]|\psi\rangle_d = \sigma_x\mathbf{U}_{\vec{\mu}}\sigma_x|\psi\rangle_d \tag{5.164}$$

$$|1\rangle \otimes |-\rangle \quad : \quad |\psi\rangle_d \to [\cos\mu\mathbb{1} + \frac{i\sin\mu}{\mu}(-\mu_x\sigma_x + \mu_y\sigma_y - \mu_z\sigma_z)]|\psi\rangle_d = \sigma_y\mathbf{U}_{\vec{\mu}}\sigma_y|\psi\rangle_d \tag{5.165}$$

where we have used the identity $\sigma_j\sigma_k\sigma_j = -\sigma_k$ if $k \neq j$. Using the above notation the action of the `QDM` can be expressed in the form

$$|\psi\rangle_d \otimes |\Xi(\vec{\mu})\rangle_p \to \frac{1}{2}\left(\sum_{j=0}^{3} \sigma_j\mathbf{U}_{\vec{\mu}}\sigma_j|\psi\rangle_d \otimes |\tilde{j}\rangle_p\right) \tag{5.166}$$

where vectors $\{|\tilde{j}\rangle_p\}$ form the basis of $\mathcal{H}_p$ associated with the realized measurement.

That is, each outcome of the measurement indicates different unitary transformation. According to a specific result we can use the same processor again to correct the wrongly transformed data register and, consequently, improve the success probability. In particular, in the case of the result j a new program register needs to encode a correcting transformation $\mathbf{U}_j^{(1)} = \mathbf{U}_{\vec{\mu}}\sigma_j\mathbf{U}_{\vec{\mu}}^{\dagger}\sigma_j$. The probability of implementing the unitary transformation using one conditioned loop is given as $p(s) = \frac{1}{4} + 3\frac{1}{16} = \frac{7}{16}$. Using more and more conditioned loops the success probability is given by $p(s) = \sum_{j=1}^{n} \frac{1}{4^j}3^{j-1} =$

$\frac{1}{4}\sum_j(\frac{3}{4})^j = \frac{1}{4}\frac{1-(3/4)^n}{1/4} = 1-(3/4)^n$ converges to unity, i.e. $p(s) \to 1$ as the number of conditioned loops n goes to infinity. For instance, thirty conditioned loops result in negligible probability of failure $p(f) = 10^{-4}$.

It seems that there is one important difference between examples of QDM and CNOT. In the case of QDM the particular results determine the new states of program register, but for CNOT processor the state $(|\alpha\rangle|2\alpha\rangle|4\alpha\rangle|6\alpha\rangle\ldots)$ can be prepared before the computation starts. However, in both these cases the next state of the program register depends on the measured outcome. In fact, if for CNOT we find out the positive outcome, then we cannot use the new program register in the state $|2\alpha\rangle_p$. We must either stop the whole process, or proceed with program state $|\alpha = 0\rangle_p$ which encodes the identity operation on data system.

The example of Cirac and Vidal shows that we are able to replace the feedback scenario with a probabilistic scenario by using different processors. A big open problem is whether the same replacement can be done in general, or at least for the case of QDM.

Note. Let us consider the QDM processor. For general program state encoding a unitary transformation, i.e. $\mathbf{F}_{0+} = \mathbf{U}_{0+}$ is unitary, it follows that for all results the data system is transformed in a unitary way, because $\mathbf{U}_j = \sigma_j\mathbf{F}_{0+}\sigma_j$ is unitary, if $\mathbf{F}_{0+}$ is. Moreover in some cases, some of these unitary transformations $\mathbf{U}_j$ coincides. For instance, if $\mathbf{F}_{0+} = e^{i\alpha\sigma_x}$ then $\sigma_x\mathbf{F}_{0+}\sigma_x = \mathbf{F}_{0+}$ and $\sigma_y\mathbf{F}_{0+}\sigma_y = \sigma_z\mathbf{F}_{0+}\sigma_z = \mathbf{F}_{+0}^\dagger = e^{-i\alpha\sigma_x}$. As a result we have found that for a certain subclasses of unitary transformations $\mathbf{U}_\alpha = e^{i\alpha\sigma_x}$ (the same holds for sets $e^{i\alpha\sigma_y}$ and $e^{i\alpha\sigma_z}$) the probability of implementation equals $p(s) = 1/2$. It means that the probability of success is greater than $1/4$. That is, using the QDM processor three one-parametric sets of unitaries can be implemented with probability $1/2$.

5.7 POVM realization

Till now we were interested in the possibility of controlling the implementation of quantum maps. But the same scheme can be used to realize POVM. In fact, any measurement of the program system determines a non-demolishing POVM of the data system. The general POVM is given by a set of positive operators $\mathbf{F}_a$ that sum up to one, i.e. $\sum_a \mathbf{F}_a = \mathbb{1}$. These operators are chosen in such a way that the frequencies of outcomes a (for a performed experiment) satisfy the following probability rule

$$p_\varrho(a) = \mathrm{Tr}(\varrho\mathbf{F}_a) \tag{5.167}$$

for all states ϱ. Using the same scenario the probabilistic quantum processor can be used to implement generalized measurements described by POVM. Note, that in this case we are not interested in a particular state transformation of the data system, because the crucial point in measurements are probabilities $p_\varrho(a)$. These probabilities (understood as linear mappings on the set of data states) completely determine the given POVM. That is, two POVMs associated with collections of positive operators $\{\mathbf{F}_a\}$ and $\{\mathbf{F}'_a\}$ are different, if $p_\varrho(a) \neq p_\varrho(a)$ for some state ϱ and some outcome a.

In our settings (U processor) the probability of finding the result a is given as

$$p_\varrho(a) = \mathrm{Tr}[\sum_{jj'} \mathbf{Q}^{(a)}_{jj'}\Xi_{jj'}\mathbf{U}_j\varrho\mathbf{U}^\dagger_{j'}] \tag{5.168}$$

where $\mathbf{M} = \sum_a a\mathbb{1}\otimes\mathbf{Q}^{(a)}$ is a projective measurement and initial program state is described by a density matrix Ξ with coefficients $\Xi_{jj'} := \langle j|\Xi|j'\rangle$. Comparing last two equations we obtain the explicit form of POVM operators

$$\mathbf{F}_a = \sum_{jj'} \mathbf{Q}^{(a)}_{jj'}\Xi_{jj'}\mathbf{U}^\dagger_{j'}\mathbf{U}_j \tag{5.169}$$

Since $\sum_a \mathbf{F}_a = \sum_{jj'}(\sum_a \mathbf{Q}^{(a)}_{jj'})\Xi_{jj'}\mathbf{U}^\dagger_{j'}\mathbf{U}_j$ and $\sum_a \mathbf{Q}^{(a)}_{jj'} = \delta_{jj'}$, it is obvious that $\sum_a \mathbf{F}_a = \mathbb{1}$. The positivity of each $\mathbf{F}_a$ follows from the construction, because the probabilities $p(a) \geq 0$ for any state ϱ.

It is enough to realize that $p(a) = \mathrm{Tr}\mathbf{F}_a\varrho = \langle\psi|\mathbf{F}_a|\psi\rangle$ providing that ϱ is a pure state. Each program state implements a POVM. However, let us look what kind of POVMs can be encoded in program states. If the implemented transformations are unitary for all outcomes a, then the probabilities are data-independent. In other words the probabilities $p_\varrho(a) = p(a)$ are not functions of initial data states. It means that such POVMs does not satisfy the main purpose of any measurement, i.e. they are useless in a state discrimination. The aim of observations is that we would like to gain information about the system, or more precisely, about the state of the system. In the mentioned case the information gain vanishes. On the other hand, only in the cases, when this information gain is zero, we can accomplish a unitary transformation. It follows that there is a relation between the unitarity of physical processes and the amount of information acquired from these processes. This result is intuitively clear. Unitary transformations cannot give us any information about the system. If we acquire some new information, it means that we have influenced the system in a non-unitary way. Therefore all states that are not interesting from the point of view of the implementation of unitary transformations, starting to be interesting from the point of view of POVM realizations. And as we can see in the `CNOT` example the set of such states is much larger.

A specific type of POVMs are those that enable us to completely distinguish between all possible states. In this case the state reconstruction (and the information gain) is complete. An example of such POVM is any collection of d^2 positive operators Λ_j (summing to identity) forming a basis of the operator space. That is any state ϱ can be written as a linear combination $\varrho = \sum_j \varrho_j\Lambda_j$. Probability of finding the result a is then given by relation

$$p_\varrho(a) = \mathrm{Tr}\varrho\Lambda_a = \sum_j \varrho_j \mathrm{Tr}\Lambda_j\Lambda_a = \sum_j \varrho_j \mathbf{L}_{ja} \tag{5.170}$$

where we have used the definition $\mathbf{L}_{ja} = \mathrm{Tr}\Lambda_j\Lambda_a$. The problem of the state reconstruction then reduces to solving the system of linear equations $p_a = \sum_a \varrho_j \mathbf{L}_{ja}$, where the numbers ϱ_j are unknown. Let us see how such "complete POVM" of a single qubit can be realized with the help of `QDM` processor.

5.7.1 `QDM` processor

Consider a general program state $|\Xi\rangle = \sum_j \alpha_j|\Xi_j\rangle_p$ and let us assume the same measurement as before, i.e. associated with the program space basis $\{|0+\rangle, |0-\rangle, |1+\rangle, |1-\rangle\}$. In this case

$$\begin{aligned}
\mathbf{F}_{0+} &= \frac{1}{4}\sum_{j,k=0}^{3} \alpha_j^*\alpha_k\sigma_j\sigma_k = \frac{1}{4}\mathbb{1} + \frac{1}{4}(\alpha_0^*\vec{\alpha} + \alpha_0\vec{\alpha^*} + i\vec{\alpha^*}\times\vec{\alpha}).\vec{\sigma} && (5.171)\\
\mathbf{F}_{0-} &= \sigma_z\mathbf{F}_{0+}\sigma_z && (5.172)\\
\mathbf{F}_{1+} &= \sigma_x\mathbf{F}_{0+}\sigma_x && (5.173)\\
\mathbf{F}_{1-} &= \sigma_y\mathbf{F}_{0+}\sigma_y && (5.174)
\end{aligned}$$

where we used $\vec{\alpha^*} = (\alpha_1, \alpha_2, \alpha_3)$ and $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$. Probability of finding the result $a\epsilon$ ($a = 0, 1$ and $\epsilon = +, -$) then reads

$$p(a\epsilon) = \mathrm{Tr}\varrho\mathbf{F}_{a\epsilon} = \mathrm{Tr}\varrho_{a\epsilon}\mathbf{F}_{0+} = \frac{1}{4} + \frac{1}{2}\vec{m}_{a\epsilon}.(\alpha_0^*\vec{\alpha} + \alpha_0\vec{\alpha^*} + i\vec{\alpha^*}\times\vec{\alpha}) \tag{5.175}$$

where $\varrho = \varrho_{0+} = \frac{1}{2}\mathbb{1} + \vec{m}.\vec{\sigma}$, $\varrho_{0-} = \frac{1}{2}\mathbb{1} + \sigma_z\vec{m}.\vec{\sigma}\sigma_z = \frac{1}{2}\mathbb{1} + \vec{\sigma}.\vec{m}_{0-}$, $\varrho_{1+} = \frac{1}{2}\mathbb{1} + \sigma_x\vec{m}.\vec{\sigma}\sigma_x = \frac{1}{2}\mathbb{1} + \vec{\sigma}.\vec{m}_{1+}$, and $\varrho_{1-} = \frac{1}{2}\mathbb{1} + \sigma_y\vec{m}.\vec{\sigma}\sigma_y = \frac{1}{2}\mathbb{1} + \vec{\sigma}.\vec{m}_{1-}$. We are using two different abbreviations: double index $a\epsilon$ and simple index j, where the relation among these two notations is obvious

$$\begin{aligned}
&0+ \leftrightarrow 0 \equiv 0 \qquad && 0- \leftrightarrow 3 \equiv z\\
&1+ \leftrightarrow 1 \equiv x \qquad && 1- \leftrightarrow 2 \equiv y
\end{aligned}$$

It is easy to verify that in the case of the implementation of unitary transformations the above rule gives us a constant distribution $p(a\epsilon) = 1/4$ (for all $\vec{m}$), because in this case $\alpha_0 = \cos\mu$ and $\vec{\alpha} = \frac{i\sin\mu}{\mu}(\mu_x, \mu_y, \mu_z)$ and, consequently, $\mathbf{F}_{0+} = \frac{1}{4}\mathbb{1}$.

In the basis of σ matrices the operators $\mathbf{F}_{a\epsilon}$ can be expressed via vectors $\vec{v}_{a\epsilon}$. The linear dependency of these vectors is equivalent to finding a nontrivial solution, $(a, b, c, d) \neq (0, 0, 0, 0)$, of the equation

$$a\vec{v}_{0+} + b\vec{v}_{0-} + c\vec{v}_{1+} + d\vec{v}_{1-} = \vec{0}\,. \tag{5.176}$$

Direct calculation shows the explicit form of the vectors

$$\begin{aligned} \vec{v}_{0+} &= (w, x, y, z) \\ \vec{v}_{0-} &= (w, -x, -y, z) \\ \vec{v}_{1+} &= (w, x, -y, -z) \\ \vec{v}_{1-} &= (w, -x, y, -z) \end{aligned} \tag{5.177}$$

where w, x, y, z stand for suitable expressions. The requirement of a nontrivial solution of the above equation (5.176) implies that none of the numbers w, x, y, z equals zero. In fact, the value of the w is fixed to be $1/4$ for all allowed operators. The three-dimensional vector $\vec{r} = (x, y, z)$ is given by the parameters of the program state $\alpha_0, \vec{\alpha}$. The explicit relation reads

$$\vec{r} = \frac{1}{4}(\alpha_0^*\vec{\alpha} + \alpha_0\vec{\alpha}^* + i\vec{\alpha}^* \times \vec{\alpha}) \tag{5.178}$$

Consider a particular example. Let us assume that the coefficients α_j are real. Then the vector $\vec{r}$ can be written in a simpler form $\vec{r} = 2\alpha_0\vec{\alpha}$. In this case

$$\mathbf{F}_{0+} = \frac{1}{2}\left[\frac{1}{2}\mathbb{1} + \alpha_0\vec{\alpha}.\vec{\sigma}\right] = \frac{1}{2}\varrho_{0+} \tag{5.179}$$

where the operator ϱ_{0+} can be associated with the quantum state if $|\alpha_0\vec{\alpha}| \leq 1/2$. The same holds for all $\mathbf{F}_{a\epsilon} = \frac{1}{2}\varrho_{a\epsilon}$. Moreover, if $|\alpha_0\vec{\alpha}| = 1/2$, then the associated operators $\varrho_{a\epsilon}$ are one-dimensional projections and therefore represent pure states. In this case the realized POVM has a nice graphical representation. In particular, the measurement projections are associated with points on the Bloch sphere. Due to the normalization of the program vector $|\Xi\rangle = \sum_j \alpha_j|\Xi_j\rangle$ ($|\alpha_0|^2 + |\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 = 1$) the condition $|\alpha_0\vec{\alpha}| = 1/2$ implies $\alpha_0 = 1/\sqrt{2}$ and $|\vec{\alpha}| = 1/\sqrt{2}$. For example, one can choose[3] $\vec{\alpha} = \frac{1}{\sqrt{6}}(1, 1, 1)$. In this particular case the operators reads

$$\begin{aligned} \mathbf{F}_{0+} = \frac{1}{2}\varrho_{0+} &= \frac{1}{2}\left[\frac{1}{2}\mathbb{1} + \frac{1}{2\sqrt{3}}(\sigma_x + \sigma_y + \sigma_z)\right] \leftrightarrow \vec{n}_{0+} = \frac{1}{2\sqrt{3}}(1, 1, 1) \\ \mathbf{F}_{0-} = \frac{1}{2}\varrho_{0-} &= \frac{1}{2}\left[\frac{1}{2}\mathbb{1} + \frac{1}{2\sqrt{3}}(-\sigma_x - \sigma_y + \sigma_z)\right] \leftrightarrow \vec{n}_{0-} = \frac{1}{2\sqrt{3}}(-1, -1, 1) \\ \mathbf{F}_{1+} = \frac{1}{2}\varrho_{1+} &= \frac{1}{2}\left[\frac{1}{2}\mathbb{1} + \frac{1}{2\sqrt{3}}(\sigma_x - \sigma_y - \sigma_z)\right] \leftrightarrow \vec{n}_{1+} = \frac{1}{2\sqrt{3}}(1, -1, -1) \\ \mathbf{F}_{1-} = \frac{1}{2}\varrho_{1-} &= \frac{1}{2}\left[\frac{1}{2}\mathbb{1} + \frac{1}{2\sqrt{3}}(-\sigma_x + \sigma_y - \sigma_z)\right] \leftrightarrow \vec{n}_{1-} = \frac{1}{2\sqrt{3}}(-1, 1, -1) \end{aligned} \tag{5.180}$$

The vectors $\vec{n}_{a\epsilon}$ represents the coordinates of pure states $\varrho_{a\epsilon}$ in the Bloch sphere picture. The (POVM) measurement consists of four projections forming a tetrahedron[4] corners on Bloch sphere (see Figure). Note that angles between these vectors equal $\eta = \arccos(1/3) = 109.47^\circ$. It is easy to check that these

[3] Note that vectors with zero entries mean that the operators $\mathbf{F}_{a\epsilon}$ are mutually linearly dependent and therefore the realized POVMs in these cases cannot be complete.

[4] Similar geometric picture is valid also for the structure of the methane molecule

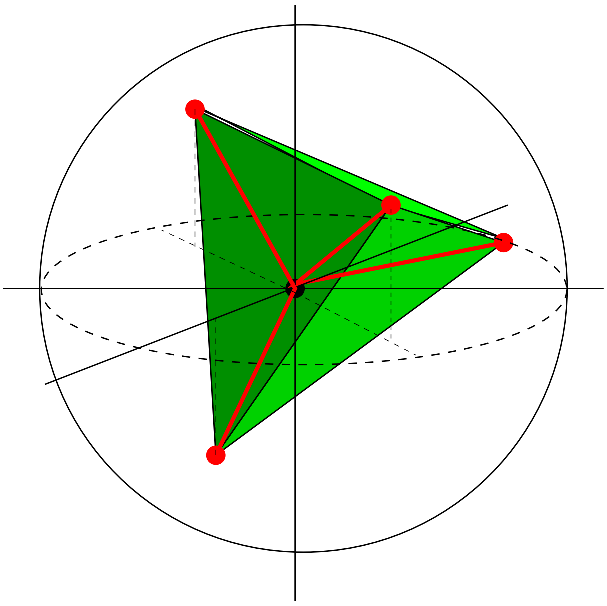

Figure 5.5: Each point represents different projection operator which is associated with the elements of complete POVM encoded in the initial program state $|\Xi\rangle = \frac{1}{\sqrt{2}}|\Xi_0\rangle + \frac{1}{2\sqrt{3}}(|\Xi_1\rangle + |\Xi_2\rangle + \Xi_3\rangle)$.

four operators are not mutually orthogonal (in the operator sense), i.e. $\mathrm{Tr}\mathbf{F}_j\mathbf{F}_k \neq 0$ for $j \neq k$. It can be also shown that using `QDM` as the processor and the same measurement the implemented POVM cannot contain mutually orthogonal projections. Whatsoever, we have shown the explicit example how to exploit the `QDM` processor in realization of the complete (nondemolition) POVM. Of course, this POVM is not unique and we can perform many different complete POVMs. The result that this processor enables us to perform the full state reconstruction is valuable on its own.

In what follows we will consider the same processor `QDM` but use the measurement $\mathbf{M} = \mathbb{1} \otimes \sum_j |\Xi_j\rangle\langle\Xi_j|$ where $|\Xi_j\rangle$ are the elementary program states encoding the transformations σ_j on data register. After observing the result j the whole system transforms in the following way

$$|\psi\rangle \otimes |\Xi\rangle \to \mathbb{1} \otimes |\Xi_j\rangle\langle\Xi_j|\mathbf{G}_{\mathtt{QDM}}|\psi\rangle \otimes |\Xi\rangle = \alpha_j\sigma_j|\psi\rangle \otimes |\Xi_j\rangle \tag{5.181}$$

In this case the resulting state is $\sigma_j|\psi\rangle$ without any dependency on the choice of the program state, only the probabilities of finding the particular result depends on such choice, $p(j) = |\alpha_j|^2$. From the point of view of POVM realization the class of realized POVM consists of the operators $\mathbf{F}_j = |\alpha_j|^2\mathbb{1}$. That is, the performed POVM is always trivial and cannot be used for state discrimination.

5.8 Deterministic loops

We have already described and analyzed two basic ways how to use quantum processor in order to realize quantum maps: deterministic and probabilistic regime. We discussed the possibility how to increase the probability of the implementation using the conditioned loops, i.e. using the same processor more than twice. To make our analysis complete in this section we will study the case of deterministic loops, i.e. repetitive usage of the processor without performing a measurement.

From the other point of view, if we have a processor that can realize a certain set of superoperators, and we wish to realize a greater set, it seems that we have to construct a new, bigger processor. Of course, this will work, but it might be enough to use the original processor more than once and repeat its action without using the new program register. Thus, the output is fed back into the input This procedure can be repeated many times, and it can considerably enlarge the set of operations that a given processor can perform.

If we use the processor again, then instead of the processor $\mathbf{G}$ we use the processor $\mathbf{G}^2$. In general, the set of induced quantum operations $\mathcal{C}_{\mathbf{G}}$ $\mathcal{C}_{\mathbf{G}^2}$ can be not only different, but also can have different size. However, the possibility that $\mathcal{C}_{\mathbf{G}^2} > \mathcal{C}_{\mathbf{G}}$ is not exactly the goal we meant. Instead of the set $\mathcal{C}_{\mathbf{G}^2}$ we are interested in the size of the set $\mathcal{C}_G^2 := \mathcal{C}_{\mathbf{G}} \cup \mathcal{C}_{\mathbf{G}^2}$ and to this set we shall refer as to the class of realizable superoperators by the processor $\mathbf{G}$ in two loops. In general, for n loops the quantum operations form a set $\mathcal{C}_{\mathbf{G}}^n = \mathcal{C}_{\mathbf{G}} \cup \ldots \cup \mathcal{C}_{\mathbf{G}^n}$ and our aim is to investigate the properties of these sets. There is a hidden assumption in the formulation of this task. In order to be able to perform all superoperators contained in $\mathcal{C}_{\mathbf{G}}^n$, we must be able to control the number of realized loops n. Hence, some additional *loop counter* is required. Its existence is a nontrivial problem and it is very similar to a *halt problem*, but so far we will assume its existence.

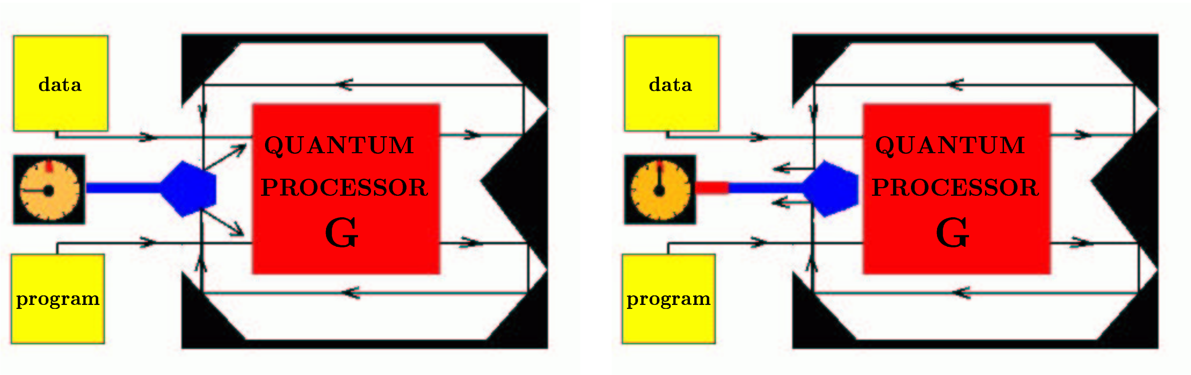

Figure 5.6: How to finish the loops?

The realized program $\Phi_{\xi,n}$ depends not only on the initial state of the program register ξ_p but also on the number of loops n. In what follows we shall focus our attention onto the implementation of unitary transformations. Obviously, $\Phi_{\xi,n}$ is unitary only if $\Phi_{\xi,n-1}$ is unitary. Since, our aim is to enlarge the set of realizable unitaries, we shall start our investigation with the use of U'-type processors, which are able to implement a countable set of unitaries in a single loop. The realization of the unitary transformation implies that the output registers are described by a factorable state, namely $\Phi_{\xi,j}[\varrho] \otimes \xi_j$, where the lower index j labels the output program state after the j-th loop (usage of the processor). Let us denote by $\mathcal{B} = \{|k\rangle\}$ and $\mathcal{B}' = \{|k'\rangle\}$ input and output bases of the U$'$-processor. Hence using the program register in the state $|k\rangle$ the data register transforms according to unitary transformation $\mathbf{U}_k$ and the output program register is described by the state $|k'\rangle$, i.e.

$$G|\psi\rangle_d \otimes |k\rangle_p = (U_k|\psi_k\rangle_d) \otimes |k'\rangle_p \,. \tag{5.182}$$

If the state $|k'\rangle$ belongs also to the basis $\mathcal{B}$, then the new action of the processor causes unitary transformation, too. But if $|k'\rangle \notin \mathcal{B}$, then the performed transformation is not unitary and new loops will not realize unitary transformations. Therefore, let us assume, that the bases $\mathcal{B}$ and $\mathcal{B}'$ contains the same elements only the ordering of them is different. That is, the sets $\mathcal{B}$ and $\mathcal{B}'$ are ordered and one can be obtained by permuting the elements of the other. Let us fix the basis $\mathcal{B}$ to be the set on which the permutations are defined. Each bijection $\sigma : \mathcal{B} \to \mathcal{B}$ represents some permutation and $\sigma\mathcal{B} = \mathcal{B}'$. In particular, the action of the processor can be written in the form

$$\mathbf{G}(|\psi\rangle \otimes |k\rangle) = \mathbf{U}_k|\psi\rangle \otimes |\sigma(k)\rangle \tag{5.183}$$

After n loops we have

$$\mathbf{G}^n(|\psi\rangle \otimes |k\rangle) = (\mathbf{U}_{\sigma^{n-1}(k)} \ldots \mathbf{U}_{\sigma(k)}\mathbf{U}_k|\psi\rangle) \otimes |\sigma^n(k)\rangle \,, \tag{5.184}$$

i.e. the state of the program register after the j-th loop is given by relation $|k_j\rangle = |\sigma^j(k)\rangle$.

For instance, the U-processors are specified by the trivial permutation $\sigma = \mathrm{id}_{\mathcal{B}}$. In such case the set $\mathcal{C}^n_{\mathcal{G}}$ contains only the powers of elementary unitaries $\mathbf{U}^n_k$, because $\mathbf{U}_{\sigma(k)} = \mathbf{U}_{\mathrm{id}_{\mathcal{B}}(k)} = \mathbf{U}_k$. In the case of nontrivial permutations σ we are able to implement certain sequences of elementary unitaries, i.e. $\mathbf{U}_{n,k} = \mathbf{U}_{k_n} \dots \mathbf{U}_{k_1}$. Of course, the sequence is fully determined by the chosen program state $|k\rangle$ and the number of loops n. As a result, we obtain that the set of implementable unitary transformations is greater, if one uses the loops. The maximal number of all implementable unitaries equals to $n \times \dim \mathcal{H}_p$, where n is the number of loops. Very interesting is the question of the covering properties of the set of realizable unitaries in the set of all unitaries. In other words, it is an open question whether this subset is dense, or not. If yes, then in practice we are able to (approximately) realize any unitary transformation in this scenario.

5.8.1 Example: single-qubit processors with loops.

In order to understand the potential of the multiple application of the same processor, let us consider a simple example when both the data and the program are represented by a qubit. Hence, the number of realizable unitaries in one loop equals two. Let us denote these unitaries by $\mathbf{U}$ and $\mathbf{V}$, i.e. $\mathbf{G} = \mathbf{U} \otimes |0\rangle\langle 1| + \mathbf{V} \otimes |1\rangle\langle 0|$. Then the implementable unitary transformations take one of the following forms

$$\mathbf{W}_{|0\rangle,n} = \mathbf{U}^{s_n}(\mathbf{VU})^{n/2} \tag{5.185}$$

$$\mathbf{W}_{|1\rangle,n} = \mathbf{V}^{s_n}(\mathbf{UV})^{n/2} \tag{5.186}$$

where $s_n = |\sin(n\pi/2)|$, $n/2$ is the integer division, i.e. $5/2 = 2$ and $1/2 = 0$, and $|j\rangle$ represents the choice of the initial program state. Since general unitary transformation can be written as an exponential of a hermitian operator, it follows that $\mathbf{VU} = e^{i\mathbf{A}}$ and $\mathbf{UV} = e^{i\mathbf{B}}$. In the case of a qubit, every hermitian operators can be expressed in a convenient form via *Pauli matrices*, i.e. $\mathbf{A} = \vec{r}\cdot\vec{\sigma}$ and $\mathbf{B} = \vec{s}\cdot\vec{\sigma}$. In the *Bloch-sphere representation* each unitary transformation corresponds to a 3D rotation of this sphere around its center. The vectors $\vec{r}, \vec{s}$ represent the axis of the rotations, respectively. Their norms, $|\vec{r}|, |\vec{s}|$, are proportional to angles of rotations described by $e^{i\mathbf{A}}, e^{i\mathbf{B}}$, respectively. The parameter $n/2$ represents the number of performed rotations around the fixed axis by a fixed angle. We can conclude that the points of the Bloch sphere transforms in jumps. The transformations $e^{i\mathbf{A}n/2}$ and $e^{i\mathbf{B}n/2}$ draws circles around the Bloch ball with respect to the initial state of the data system. Trivially, the rotations U and V preserve the shapes of these circles and consequently the set of transformations $\mathbf{U}e^{i\mathbf{A}n/2}$ represents another (transformed) circle. To illustrate the action of the processor, we can say that the initial state of the data register, represented as a point on the Bloch sphere is mapped into two orbits (circles) on the sphere. After each action of the processor (i.e. after each loop) the point representing the data state moves (jumps) from one orbit to the other.

The fact that the set of transformations is dense is exhibited by the property that each initial point travels (approximately) around the whole Bloch sphere. In our case, starting from any point we are able to cover maximally two circles (orbits) on the Bloch sphere. The rate of covering of these orbits depends on the angle of single rotations (one jump caused by one action of the processor).

Let us generalize the previous scheme by introducing a larger program space. Let us denote its dimension by d. In this case the realized unitaries are

$$\mathbf{W}_{|k\rangle,n} = \mathbf{U}_{\sigma^j(k)} \dots \mathbf{U}_k e^{i\mathbf{A}_k n/d} \tag{5.187}$$

Likewise before, we can use the Bloch ball picture to illustrate the transformation of the state of the data register. Starting from any given point (state) in each loop we jump among one of the d circles. If we want to cover the whole sphere, then the number of circles must tend to infinity together with the dimension of the program system. To be able to realize all unitaries we must increase two parameters: the dimension of program space and the number of loops. Note one possible disadvantage of considering the loops. On one hand for specific transformations the realization is exact and fast, but in general the realization strongly depends on the number of loops and therefore long time is required. In all previous regimes, the time plays no explicit role and every process took us the same time.

5.9 Conclusion

The field of the so-called quantum processor is quite large, because the formulation of the problem is common for many other seemingly unrelated areas. The results from this field can be useful in measurement theory, in decoherence problem, and mainly in quantum dynamics of open systems. Basicly the study of quantum processors is equivalent to the study of the quantum dynamics. In this chapter we have concentrated on different settings in which a fixed unitary transformation plays a central role. Let us now summarize briefly this chapter.

Quantum processor is nothing else as a fixed unitary transformation acting on two different systems, which are called (in accordance with computer terminology) data and program register. It is commonly believed that the most general quantum evolution can be associated with a unitary one realized on a larger system (*Kraus theorem*). The effect of nonunitarity of the evolution of the systems is caused by mutual interactions between the system and its surrounding. In a sense, the properties of the environment determines the evolution of the systems and *vice versa*. The aim of the quantum processor is to control and determine dynamics of systems in a programmable way by using existing fixed unitary transformation. We have discussed two basic regimes: deterministic and probabilistic (or stochastic). We used these regimes to implement quantum maps and also generalized measurements (described by POVM). In particular, we studied the properties of the realization of quantum maps of a single qubit. We also studied the inverse problem, whether any set of quantum maps can be realized in the framework of these models. We showed that in a probabilistic way the universal quantum processor can be constructed. Moreover, by applying conditioned loops the probability of successful realization of any map can be made arbitrarily close to one.

There is clearly a lot more to do here. One question is how to improve the probability of success in a similar way as it was done by Vidal et al. in the case of one-parametric group $\mathbf{U}_\alpha = \exp(i\alpha\sigma_x)$ by using larger quantum processor. In other words to find out some relation between the conditioned implementation and unconditioned one with a new quantum processor. This is an object of the future study and even the existence of the solution is an open question.

Chapter 6

Multipartite entanglement

The present is yours.
The future, for which I really worked is mine.
N.Tesla

6.1 On the bipartite entanglement and correlations

We shall differentiate among the following types of bipartite states:

State	Name	Correlation	Entanglement
$\varrho = \varrho_A \otimes \varrho_B$	factorized state	$C_\varrho(A,B) = 0$	$E(\varrho) = 0$
$\varrho = \sum_k p_k \varrho_k^A \otimes \varrho_k^B$	separable state	$C_\varrho(A,B) \geq 0$	$E(\varrho) = 0$
$\varrho = \sum_k p_k \varrho_k^A \otimes \varrho_k^B$	correlated state	$C_\varrho(A,B) \neq 0$	$E(\varrho) = 0$
$\varrho \neq \sum_k p_k \varrho_k^A \otimes \varrho_k^B$	entangled state	$C_\varrho(A,B) \neq 0$	$E(\varrho) \neq 0$

where

$$C_\varrho(A,B) = S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB}) \tag{6.1}$$

is the measure of correlation between the systems A and B. The function $E(\varrho)$ stands for a measure of entanglement. We have introduced a new (probably non-standard) subset of quantum states called *correlated states.* Note that the notion of a correlated state means that $C_\varrho(A,B) > 0$, but $E(\varrho_{AB}) = 0$. That is, correlated states are not entangled, but only "correlated". It does not mean, that entangled states are uncorrelated in the standard sense (of $C_\varrho(A,B)$). In fact, the *correlation* is a statistical notion independent of the physical theory we use. Therefore, in a sense the notions like *classical*, or *quantum correlations* can be used only to label the theory in which the correlations are studied. However, we shall use these names to distinguish subsets of quantum states, for which the correlations have different "quality". In particular, one can require that these two correlations should sum up to *total correlations* measured for instance by the function $C_\varrho(A,B)$. Moreover, the classes of states can be made more subtle, if we define also the *classical states* of quantum systems, and consequently, we shall require that only these states contain *"pure" classical correlations.* Then there exist separable states, for which the correlations are not only classical, but also quantum (without entanglement). Let us remind the definition of the *classical states* of quantum systems. The state ϱ_{AB} is *classical*, if in its spectral decomposition $\varrho_{AB} = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k|$ the pure states $|\psi_k\rangle_{AB}$ are factorized. All these definitions are disputable and without an introduction of the measure of a "classical" correlation, or a "quantum" correlation, this discussion does not have a big sense. The main point of the definition of the "classical" (or "quantum" correlations) is to shift the border between the quantumness and the classicality beyond the existing line of separable and entangled states.

To understand deeper the concept of entanglement we must clarify its relation to the existing concept of correlations. Elementary steps to leading further in the analysis of this phenomenon are the creation (or destruction) of the entanglement and its detection. In the next paragraph we will pay attention to the creation of entanglement and correlations.

6.1.1 The creation of entanglement and correlations

Hamiltonian dynamics of classical theory is *deterministic* in the sense that pure states remain pure during the whole evolution. In quantum theory the Hamiltonian dynamics is associated with unitary transformations. Due to the nonexistence of correlated pure states in classical theory, it follows that during the Hamiltonian dynamics no correlation can arise. However, in quantum case there exist entangled pure states, for which $\mathbf{C}_\varrho(A,B)\neq 0$. Therefore the creation of the correlations (from uncorrelated states) governed by a unitary (Hamiltonian) evolution is possible, because any two pure states are related by (many) unitary transformations. Moreover, we can say that the genesis of entanglement is conditioned by the presence of interactions between the systems. Due to the reversibility of Hamiltonian dynamics, the same holds also for the entanglement destruction. It is an interesting question on the capability of a given quantum operation (not only unitary) to create (or destroy) the entanglement, but this problem is out of the scope of this thesis.

The creation of entanglement (or correlations) from the initially mixed state is a little bit more complicated. The following example shows that a unitary transformation can be used to create correlations without entangling the systems. Consider the factorized state of two qubits $|0\rangle\langle 0|\otimes\frac{1}{2}\mathbb{1}$. Applying a global unitary transformation given by $|00\rangle\to(|00\rangle+|11\rangle)/\sqrt{2}$ and $|01\rangle\to(|00\rangle-|11\rangle)/\sqrt{2}$, we obtain the state $\varrho_{AB}=\frac{1}{2}(|00\rangle\langle 00|+|11\rangle\langle 11|)$, which is known as the maximally correlated state. For this state $C_\varrho(A,B)=1$, but entanglement vanishes. This quantum example cannot be used to answer the same question stated in the field of classical theory, because it transforms factorized states into entangled (non-classical) ones. Anyway, the same result can be obtained by using unitary transformation $|00\rangle\to|00\rangle$ and $|01\rangle\to|11\rangle$, i.e. $|0\rangle\langle 0|\otimes\frac{1}{2}\mathbb{1}\to\frac{1}{2}(|00\rangle\langle 00|+|11\rangle\langle 11|)$. This unitary transformation can be understood as a classical reversible transformation with two classical bits. Therefore, the correlations can be created by reversible (Hamiltonian) dynamics in both theories.

What is interesting is that the reversible (unitary) dynamics of quantum systems cannot create an arbitrary amount of the entanglement between the systems. Its value is limited by the initial state, namely by its purity (or entropy). This follows from the fact that unitary transformations preserve the joint entropy of states (purity), but the entropy of subsystems can be changed. In particular, the maximally entangled states ($E=2$) are pure (entropy equals zero), but the total mixture (state with maximal entropy and zero entanglement) is preserved under unitary transformations. It means that starting with states that have maximal joint entropy $S=2$, no entanglement can be created by unitary evolution. On the other hand any pure state can be transformed into maximally entangled state. These are the two extremes of the mutual relation between the entanglement and the joint entropy.

If we look at the classical systems, then the production of the correlations does depend on the entropy of the initial state. However, the relation is completely different, because the correlation is maximal for the state (of two bits) with the entropy $S(\varrho_{AB})=1$, when the total mixture has entropy $S=2$. Therefore, the correlations raise together with the initial entropy up to the value $S=1$ and then they get down to zero, when the initial state is the total mixture. But such behavior is valid only in classical theory, because in quantum theory the maximally entangled states are also maximally correlated ($C_\varrho=2$). In quantum theory the correlations decrease, if the joint entropy increases, and they vanish only for total mixture. In particular, the correlation function of two qubit system equals $C_\varrho(A,B)=S(\varrho_A)+S(\varrho_B)-S(\varrho_{AB})=2-S(\varrho_{AB})$, where the last equality is valid, because it is possible to find such states, for which the subsystems are described by total mixtures for all values of the joint entropy. For instance, Werner states $|\mu\rangle=\mu|EPR\rangle\langle EPR|+\frac{1-\mu}{4}\mathbb{1}$ are of this type. By unitary transformation we can always transform the given state into one of the Werner states parametrized by μ. Of course, during this process the mutual correlations (entanglement) between the subsystems can change. The answer to the question whether Werner states maximize also the entanglement (for a fixed joint entropy), depends on the entanglement measure we use [55].

We have seen that there are differences between the concepts of entanglement and correlations. Therefore, it is often misleading to use analogies. However, it is often very instructive to have a correlation analogy in hands. Most of the examples, in which the entanglement plays a central role, have their "correlation" (classical) analogy. These analogies help us to better understand and specify the properties that make the entanglement so interesting.

6.2 Classes of multipartite entanglement

The mathematical definition of the bipartite entanglement is clear. In the same way we can define the entangled state of a composite N-partite system associated with the Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_N$. The *factorized states* are all states of the form $\varrho_{1\ldots N} = \varrho_1 \otimes \ldots \otimes \varrho_N$. In accordance with the definition of the bipartite entanglement we say that the state $\varrho_{1\ldots N}$ is **entangled**, if it cannot be expressed as a convex sum of factorized states, i.e.

$$\varrho_{1\ldots N} = \sum_k p_k \varrho_1^k \otimes \ldots \otimes \varrho_N^k. \tag{6.2}$$

Otherwise, the state will be called *separable*. These definitions divide all quantum states into two basic groups of states: entangled and separable (= not entangled).

Likewise the bipartite case, a pure state is entangled whenever it contains some correlations, i.e. it is not factorized. We can use the function

$$C_\varrho(A_1, \ldots, A_N) = \sum_j S(\varrho_{A_j}) - S(\varrho_{A_1,\ldots,A_N}) \tag{6.3}$$

to quantify the "total" amount of correlations in quantum states. For pure states the second term (entropy of the state of the whole system) vanishes and the degree of entanglement (of pure states) can be identified with the presence of correlations, i.e. it is given as the sum of entropies of all subsystems $E = C_\psi(A_1, \ldots, A_N) = \sum_j S(\varrho_{A_j})$. This correspondence between entanglement and correlation $C_\varrho(A, B)$ is inappropriate in the case of mixed states. Simply, the function $C_\varrho(A, B)$ measures the correlations and does not discriminate separable and entangled mixed states. However, the *entanglement of formation* can be defined by the following formula

$$E_f(\varrho) = \min_{\varrho = \sum_k p_k |\psi_k\rangle\langle\psi_k|} \sum_k p_k C_{\psi_k}(A_1, \ldots, A_N) \tag{6.4}$$

The existence of the multi-tensor structure makes the state space more complicated and interesting. For instance, it is commonly accepted that there exist at least two different (nontrivial) types of pure state entanglement shared between three qubits. The extremal examples (and typical representatives) of these two classes are the *GHZ state* and *W state*.

The Greenberger-Horne-Zeillinger state $|GHZ\rangle$ is in a sense a generalization of the EPR pair of two qubits. In particular

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \otimes |1\rangle) \tag{6.5}$$

The entanglement in this state cannot be reduced to entanglement shared between two qubits. In fact, all three couples of qubits are described by the maximally correlated state $\frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$, which is obviously not entangled. Sometimes we say that GHZ contains only a *pure three-partite entanglement.*

The second type of the three-partite entanglement is exhibited by the W state

$$|W\rangle = \frac{1}{\sqrt{3}}(|0\rangle \otimes |0\rangle \otimes |1\rangle + |0\rangle \otimes |1\rangle \otimes |0\rangle + |1\rangle \otimes |0\rangle \otimes |0\rangle) \tag{6.6}$$

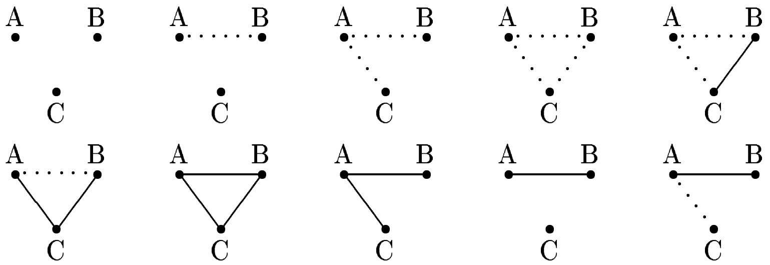

Figure 6.1: Figure describes ten possible configurations (with respect to shared correlations) of three qubits. Solid lines correspond to existing entanglement and dotted lines represent the possible correlations, if no entanglement is shared.

Unlike the GHZ state in this case all the couples of qubits are entangled. Moreover, the shared entanglement between each pair has the same value. Under the symmetry requirements it can be shown that the value of shared bi-qubit entanglement is also maximal (in terms of the concurrence it equals $C = 2/3$).

6.2.1 Graph representation

In this section we shall study the entanglement properties with respect to all bipartite splittings. That is, we will use all bipartite entanglements to characterize the multipartite entanglement. The most simple type of investigation is based on the analysis of the amount of entanglement shared between each pair of particles. In this case one can draw a simple graph representation of quantum states (see Figure), where vertices correspond to particles and lines exhibit the presence of entanglement between the particles. In cases when the vertices are not connected by a line the two particles are not entangled. Roughly speaking, entangled graph is a graphical representation of the bipartite entanglement shared between N particles. However, the couple of non-entangled particles can be still correlated, i.e. the correlation function $C_\varrho(A,B) \neq 0$ can be non-zero even if the entanglement vanishes. The existence of purely correlated (not entangled) pairs will be drawn in by dotted lines.

6.2.2 Three qubits

In this paragraph we will analyze the simplest multipartite example of three qubits to see explicitly the state representation of all graphs. The three qubits can form one of the following entanglement-correlation arrangements. If we are interested in both: the correlations and the entanglement, there exist ten different graphs (see Figure). In the case, when we consider only the entanglement the number of graphs is reduced to four. It is obvious that the set of quantum states can be divided into subsets determined by these graphs, but the question is, whether for each configuration this subset is nonempty. Thus the existence of the quantum state ϱ representing a given graph is questionable.

Consider that only the entanglement is taken into account, i.e. we have four types of graphs (with zero, one, two and three lines). The problem of the existence of such mixed states has been already solved in [25]. We shall not repeat the arguments given there, but rather we shall study the existence of the pure state representation of these graphs. Obviously, the trivial case of the graph with no entanglement is fulfilled by factorized states $|\psi_{ABC}\rangle = |\psi\rangle_A \otimes |\psi\rangle_B \otimes |\psi\rangle_C$. However, also the entangled *GHZ state* $|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ belongs to this class of states. The so-called *W* *state* $|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |100\rangle + |010\rangle)$ is a typical representative of the graph with all the three lines, i.e. in this case each pair of qubits shares the entanglement. The state with a single entangled pair $|\psi\rangle_{ABC} = |\psi\rangle_{AB} \otimes |\psi\rangle_C$ represents the configuration with a single line.

The last case of the graph depicting two lines might sound (in some sense) counterintuitive. It corresponds to a system of three particles, where one of them is entangled to others, but the other two are not mutually entangled. Let us give a simple example that such situation can occur in the framework of classical correlations. Consider three random variables, where the third of them is simply a summation of the first two. In the Dirac notation the state of three bits can be written as

$$\varrho_{ABC} = \sum_{kl} p_{kl} |k\rangle\langle k| \otimes |l\rangle\langle l| \otimes |k \oplus l\rangle\langle k \oplus l| \tag{6.7}$$

where the first two qubits are statistically independent, i.e. $p_{kl} = p_k p_l$ and the state of the first two qubits is given by formula $\varrho_{AB} = (\sum_k p_k |k\rangle\langle k|) \otimes (\sum_l p_l |l\rangle\langle l|)$. The states of the other two couples of qubits is obviously correlated, because the reduced state $\varrho_{AC} = \sum_{kl} p_k p_l |k\rangle\langle k| \otimes |k \oplus l\rangle\langle k \oplus l|$ cannot be written in a factorized form. Note that $k, l = 0, 1$ and $1 \oplus 1 = 0$. However, the entanglement is quite far from being a correlation and therefore one cannot expect that similar property will hold also for entanglement.

Consider the following state of three qubits

$$|\Xi\rangle_{ABC} = \frac{1}{\sqrt{2}} (|0\rangle|0\rangle|+\rangle + |1\rangle|+\rangle|1\rangle) \tag{6.8}$$

where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. It is easy to check that in this case $\varrho_{BC} = \frac{1}{2}(|0+\rangle\langle 0 + | + | + 1\rangle\langle 1 + |)$, i.e. the systems B and C are not entangled. Direct calculation shows that the reamining two pairs of qubits are entangled. This state is only a special example of the general class of states that we shall study later on. Note that, if one considers four qubits, then this kind of graph can be obtained much easier. Simply imagine that the system consists of two Bell pairs. Take a couple of the qubits (not described by a Bell state) and denote these two qubits as a single system A. Then the system A is entangled to other two systems (qubits B and C), respectively and, moreover, these two systems (B and C) are not entangled. Let us now summarize the results in the following table

Graph	State	E_{AB}	E_{AC}	E_{BC}
0 lines	$\lvert\psi\rangle_A \otimes \lvert\phi\rangle_B \otimes \lvert\xi\rangle_C$	0	0	0
	$\lvert GHZ\rangle_{ABC}$	0	0	0
1 line	$\lvert EPR\rangle_{AB} \otimes \lvert\psi\rangle_C$	1	0	0
2 lines	$\lvert\Xi\rangle_{ABC}$	1/2	1/2	0
3 lines	$\lvert W\rangle_{ABC}$	1/3	1/3	1/3

where E_{jk} denotes the entanglement between jth and kth qubit ($j, k = A, B, C$). Note that the values are only illustrative.

"Correlated" graphs

The division of the state space according to these four graphs is not sufficient in order to say that states belonging to one subset have the same quality (or type) of entanglement. Obviously, the factorized states and GHZ state have completely different entanglement properties. Therefore, we will add dotted lines into the graphs that correspond to correlated, but not entangled states. Let us denote different graphs by (j, k) , where j represents the number of entangled (full) lines and k represents the number of dotted lines. The following table shows the relation between the graphs with "only entangled" lines and "correlated" graphs (j, k).

"entangled" graph	"correlated" graph
zero lines	(0,0) (0,1) (0,2) (0,3)
one line	(1,0) (1,1) (1,2)
two lines	(2,0) (2,1)
three lines	(3,0)

Our aim is to find a pure state representatives of all correlated graphs. Using the results of the previous paragraph the state representation of types $(0,0)$, $(0,3)$, $(1,0)$ and $(3,0)$ is already known (see the table bellow). The existence of all other graphs is an open problem that we are going to solve now. First consider the graphs with one dotted line, i.e. $(k,1)$ type. Let us assume that qubits B and C are in a separable correlated state

$$\varrho_{BC} = \lambda|0\psi\rangle\langle 0\psi| + (1-\lambda)|\phi 1\rangle\langle\phi 1| \tag{6.9}$$

where $|\psi\rangle_C$ and $|\phi\rangle_B$ are arbitrary vectors. Next we would like to purify this state by adding only one qubit. This requirement of the single qubit purification implies that the rank of the state ϱ_{AB} must be less or equal to 2, i.e. $rank(\varrho_{BC}) \leq 2$. Therefore, the above state is the most general we can consider. Its single qubit purification takes the form

$$|\Psi\rangle_{ABC} = \sqrt{\lambda}|00\psi\rangle + e^{i\theta}\sqrt{1-\lambda}|1\phi 1\rangle \tag{6.10}$$

The reduced states of the particular couples read

$$\begin{aligned}
\varrho_{AB} &= \mathrm{Tr}_C|\Psi\rangle\langle\Psi| = \lambda|00\rangle\langle 00| + (1-\lambda)|1\phi\rangle\langle 1\phi| + e^{-i\theta}\langle 1|\psi\rangle\sqrt{\lambda(1-\lambda)}|00\rangle\langle 1\phi| + c.c. && (6.11)\\
\varrho_{AC} &= \mathrm{Tr}_B|\Psi\rangle\langle\Psi| = \lambda|0\psi\rangle\langle 0\psi| + (1-\lambda)|11\rangle\langle 11| + e^{-i\theta}\langle 0|\phi\rangle\sqrt{\lambda(1-\lambda)}|0\psi\rangle\langle 11| + c.c. && (6.12)\\
\varrho_{BC} &= \lambda|0\psi\rangle\langle 0\psi| + (1-\lambda)|\phi 1\rangle\langle\phi 1| && (6.13)
\end{aligned}$$

Note that the phase factor $e^{i\theta}$ can be included into the complex parameters of states $|\phi\rangle_B = a|0\rangle + b|1\rangle$ and $|\psi\rangle_C = \alpha|0\rangle + \beta|1\rangle$.

Graph $(1,2)$

Instead of solving the entanglement properties of $\varrho_{AB}, \varrho_{AC}$ in general, let us make the following simplification. Let us assume that $|\psi\rangle_C = |0\rangle_C$ and $|\phi\rangle_B$ is arbitrary. In this parametrization we get

$$\begin{aligned}
\varrho_{AB} &= \lambda|00\rangle\langle 00| + (1-\lambda)|1\phi\rangle\langle 1\phi| && (6.14)\\
\varrho_{AC} &= \lambda|00\rangle\langle 00| + (1-\lambda)|11\rangle\langle 11| + a\sqrt{\lambda(1-\lambda)}|00\rangle\langle 11| + c.c. && (6.15)\\
\varrho_{BC} &= \lambda|00\rangle\langle 00| + (1-\lambda)|\phi 1\rangle\langle\phi 1| && (6.16)
\end{aligned}$$

It is easy to see that the states ϱ_{AB} and ϱ_{BC} are not entangled. To see whether ϱ_{AC} is entangled or not we check the positivity of partially transposed matrix $\varrho_{AC}^{T_C}$. Direct calculation determines that the eigenvalues of $\varrho_{AC}^{T_C}$ are $\{\lambda, 1-\lambda, |a|, -|a|\}$. Due to the negativity of eigenvalue $-|a|$ (if $a \neq 0$) we can conclude that the qubits A and C are entangled. As a result we obtain that states

$$|Z_{1,2}\rangle = \lambda|000\rangle\langle 000| + a(1-\lambda)|101\rangle + b(1-\lambda)|111\rangle \tag{6.17}$$

(with $a \neq 0$ and $|a|^2 + |b|^2 = 1$) represent the graph $(1,2)$. In the special case of $a = 1$, this state corresponds to graph $(1,0)$, because the state $\varrho_{AC} = \varrho_{AC}^2$ is pure and entangled, and the qubit B (in the state $|0\rangle$) becomes to be factorized from the systems A and C.

Graph $(2,1)$

We use the formulas derived for the concurrence of specific states (3.44) in order to calculate the value of the entanglement in states ϱ_{AB} and ϱ_{AC} Applying a local unitary transformation these states can be transformed into the generic form (3.44). It is enough to change the labels $0 \leftrightarrow 1$ to obtain the desired form

$$\varrho'_{AB} = \lambda|10\rangle\langle 10| + (1-\lambda)|0\phi\rangle\langle 0\phi| + \langle 1|\psi\rangle\sqrt{\lambda(1-\lambda)}|10\rangle\langle 0\phi| + c.c. \tag{6.18}$$

$$\varrho'_{AC} = \lambda|0\psi'\rangle\langle 0\psi'| + (1-\lambda)|10\rangle\langle 10| + \langle 0|\phi\rangle\sqrt{\lambda(1-\lambda)}|0\psi'\rangle\langle 10| + c.c. \tag{6.19}$$

where $|\psi'\rangle = \alpha|1\rangle + \beta|0\rangle$. The matrix elements $|01\rangle\langle 10|$ completely characterize the amount of entanglement. As a result we obtain the following values for the shared entanglement

$$C_{AB} = |b\beta|\sqrt{\lambda(1-\lambda)} \tag{6.20}$$

$$C_{AC} = |a\alpha|\sqrt{\lambda(1-\lambda)} \tag{6.21}$$

where $\alpha = \langle 0|\psi\rangle$, $\beta = \langle 1|\psi\rangle$, $a = \langle 0|\phi\rangle$ and $b = \langle 1|\phi\rangle$. In the specific case of $\beta = 0$ we get the previous example of the graph $(1,2)$. In particular, in this case $C_{AB} = 0$ and $C_{AC} = |a|\sqrt{\lambda(1-\lambda)}$. In conclusion, the states of the type (6.10) can be represented by the graph $(2,1)$, if $a, b, \alpha, \beta \neq 0$. Let us denote all such states by $|Z_{2,1}\rangle$.

Other graphs

The three of the remaining types of graphs contain a single factorized (uncorrelated) pair, let us say that the BC one, i.e. $\varrho_{BC} = \varrho_B \otimes \varrho_C$. In order to fulfil the condition that the rank of the operator can be maximally two, one of the states (say ϱ_C) must be pure. To preserve the purity of the whole three-qubit system the couple AB has to be pure, too. Consequently, there can exist only one line in the graph representation of such states. Therefore the graphs $(1,1)$, $(0,2)$ and $(2,0)$ cannot be associated with pure states of a three-qubit system. The last graph $(0,1)$ does not exist, because it contains two factorized pairs and from the previous it follows that the the dotted line between A and B must be described by a pure state, but the non-entangled pure states are always factorized.

Graphs with mixed states

To complete our discussion of three-qubit states let us briefly analyze the mixed states representation of graphs. It is obvious that all types $(0,k)$ exist. The problematic might be the graph $(0,2)$, but the example given above (see Eq.(6.7)) proves its existence. Except the graphs $(1,1)$ and $(2,0)$ the existence of all the other has been proved, because they exist for pure states. Consider the following mixture

$$\varrho_{ABC} = \lambda|EPR\rangle_{AB}\langle EPR| \otimes \frac{1}{2}\mathbb{1} + (1-\lambda)\frac{1}{2}\mathbb{1} \otimes |EPR\rangle_{BC}\langle EPR| \tag{6.22}$$

and calculate the reduced density operators

$$\begin{aligned} \varrho_{AB} &= \lambda|EPR\rangle\langle EPR| + (1-\lambda)\tilde{\mathbb{1}} && (6.23)\\ \varrho_{AC} &= \tilde{\mathbb{1}} && (6.24)\\ \varrho_{BC} &= (1-\lambda)|EPR\rangle\langle EPR| + \lambda\tilde{\mathbb{1}} && (6.25)\\ & && (6.26) \end{aligned}$$

where $\tilde{\mathbb{1}} := \frac{1}{4}\mathbb{1}$ is the state describing the total mixture and $|EPR\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. The state ϱ_{AC} fulfils the condition that systems A and C are factorized. Moreover, the states ϱ_{AC} and ϱ_{BC} belong to the family of the so-called *Werner states* $|\mu\rangle = \mu\mathbf{P}_{EPR} + (1-\mu)\tilde{\mathbb{1}}$, where $\mathbf{P}_{EPR}$ is the projection onto the maximally entangled state $|EPR\rangle$. These states are entangled, if the parameter $\mu > 1/3$. In our case we are looking for λ from the intersection of the intervals $\lambda > 1/3$ and $1-\lambda > 1/3$, i.e. $1/3 < \lambda < 2/3$. In this case both couples AB and BC are entangled, i.e. the state ϱ_{ABC} can be associated with the graph $(2,0)$. Moreover, if λ is outside the allowed region, then only one of the pairs is entangled and the other one is correlated. For instance, if $\lambda < 1/3$, then the state ϱ_{AB} is not entangled, but the correlation function $C_\varrho(A,B) = 2 - S(\varrho_{AB})$ does not vanish. The state ϱ_{BC} is entangled, because in this case $\mu = 1-\lambda > 1/3$.

We have shown that all the graphs can be represented by a mixed state of the three-qubit state. Let us summarize our results (together with the pure state case) in the following table.

Graph	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(2,0)	(2,1)	(3,0)
Mixed State	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Pure State	yes	no	no	yes	yes	no	yes	no	yes	yes
Example	$\|000\rangle$	—	—	$\|GHZ\rangle$	$\|0\rangle\|EPR\rangle$	—	$\|Z_{1,2}\rangle$	—	$\|Z_{2,1}\rangle$	$\|W\rangle$

6.2.3 Arbitrary pure states of three-partite systems

In this section we will see that the pure states of three qubits and pure states of arbitrary three particles have different properties from the point of view of their graph representation. Namely, we are interested in the possibility to construct graphs of the type $(0,1)$, $(0,2)$, $(1,1)$ and $(2,0)$ by pure state of general three-partite system. These graphs cannot be done by a pure three-qubit states.

Obviously, the graph $(0,1)$ cannot exist, because it requires that one party is completely factorized from the joint system. On the other side the purity of the joint state implies that the other two systems must be described by a pure state and therefore they must not be correlated without being entangled. Both graphs $(1,1)$ and $(0,2)$ contain a subsystem (say A) which is correlated (not entangled) with the system B, but uncorrelated with the system C, i.e.

$$\varrho_{AC} = (\sum_j p_j |j\rangle_A\langle j|) \otimes (\sum_k q_k |k\rangle_C\langle k|) \tag{6.27}$$

Due to the orthogonality of the basis $|j\rangle_A \otimes |k\rangle_C$ the joint pure state can be written in the Schmidt basis

$$|\Psi\rangle_{ABC} = \sum_{jk} \sqrt{p_j q_k} |j\rangle_A \otimes |jk\rangle_B \otimes |k\rangle_C \tag{6.28}$$

where vectors $|jk\rangle_B$ form an orthonormal basis of $\mathcal{H}_B$. Tracing out the system A we obtain the state

$$\varrho_{BC} = \sum_{j,k,k'} p_j \sqrt{q_k q_{k'}} |jk\rangle_B\langle jk'| \otimes |k\rangle_C\langle k'| \tag{6.29}$$

The basis of the Hilbert space $\mathcal{H}_B$ is labelled by the double index (jk). Formally we can introduce a tensor structure into this Hilbert space by putting $|jk\rangle_B = |j\rangle_{B_1} \otimes |k\rangle_{B_2}$. Using such notation the state ϱ_{BC} can be rewritten into the form

$$\varrho_{BC} = (\sum_j p_j |j\rangle_{B_1}\langle j|) \otimes \sum_{k,k'} \sqrt{q_k q_{k'}} |kk\rangle_{B_2C}\langle k'k'| = \varrho_{B_1} \otimes |\psi\rangle_{B_2C}\langle\psi| \tag{6.30}$$

where $|\psi\rangle = \sum_k \sqrt{q_k}|k\rangle_{B_2} \otimes |k\rangle_C$. Obviously, the entanglement between the system B and system C is governed by the correlations contained in the pure state $|\psi\rangle_{B_2C}$, i.e. all correlations are associated with entanglement. The same result can be derived also for the state ϱ_{AB}. In conclusion, the graphs with one uncorrelated and one correlated line ($(1,1)$ and $(0,2)$) cannot be designed by pure states of arbitrary three systems.

The last graph $(2,0)$ can be easily constructed in the following say. The joint system will consist of four qubits A_1, A_2, B and C. We have already mentioned that the state

$$|\Psi\rangle_{ABC} = |EPR\rangle_{A_1B} \otimes |EPR\rangle_{A_2C} \tag{6.31}$$

is of the type $(2,0)$, because the parties B and C are described by the factorized state $\tilde{\mathbb{1}}$.

The following table summarizes the results obtained for arbitrary three-partite systems. We have found that in comparison with the three-qubit state the graph $(2,0)$ can be designed by a pure state of general three-partite system.

Graph	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(2,0)	(2,1)	(3,0)
Mixed State	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Pure state	yes	no	no	yes	yes	no	yes	yes	yes	yes

6.2.4 Multipartite systems

The number of graphs associated with multipartite states increases rapidly. Each (uncorrelated) graph is given by the set of double indices $I = \{kj\}$, where (kj) corresponds to existing line between the vertices k and j, i.e. the particles A_k and A_j are entangled. The question of the mixed state representation of a given graph (without the dotted lines) has been solved by *W.Dür* [41]. He proposed a multi-qubit state

$$\varrho_I = \frac{1}{M} \sum_I x_{kj} |\psi_{kj}\rangle\langle\psi_{kj}| \tag{6.32}$$

where $|\psi_{kj}\rangle = |\psi^+\rangle_{kj} \otimes |00\ldots\rangle_{rest}$, $M = \sum x_{kj}$ and $|\psi^+\rangle_{kj}$ is maximally entangled state of two qubits k and j. It is easy to verify that such mixture of entangled states has the required properties. In particular, the concurrence $C_{mn} = x_{mn}/M$ if $mn \in I$ and $C_{mn} = 0$ if $mn \notin I$. The same question for the pure states has been solved by *M.Plesch* [57].

Introducing the correlated lines into the multipartite graphs the state space will be divided into much more subsets. The question of the physical representation of all these graphs is very interesting, but it is a content of another work.

6.3 Local unitary equivalence

Entanglement has divided the state space of composite quantum systems into two basic subsets: entangled and nonentangled states. Given a state ϱ. An elementary question, whether this state is entangled or not, has not been solved satisfactorily and only partial results are known. However, the story continuous and modified questions are being solved with the objective to better understand the structure of the quantum state space. From the physical point of view the entangled states can have different *"entanglement properties"*. Of course, the differences depend on the situation we are interested in. For instance, two states can be represented by the same graph (without dotted lines), but the first one is factorized and the other one is entangled, or one of them is pure and the second one is mixed.

In this section we shall pay attention to a basic property of entanglement. If two states ϱ and ξ of multipartite systems contain the same amount of entanglement, then they must be *locally unitarily equivalent*, i.e. $\varrho = \mathbf{U}\xi\mathbf{U}^\dagger$ with $\mathbf{U} = \mathbf{U}_{A_1} \otimes \ldots \otimes \mathbf{U}_{A_N}$. Our aim is to find a criterion that would enable us to answer the question whether two states are locally unitarily equivalent (LUE), or not. In what follows we are going to show several approaches how to deal with this mathematical problem.

6.3.1 Invariants of local unitary transformations

Firstly we will investigate the invariants of local unitary transformations, i.e. necessary conditions of LUE. After that we shall try to reverse the implication to find out the sufficient conditions.

Let us simplify the problem a little bit . It is easy to see that the solution of the bipartite LUE will enable us to answer also the multipartite LUE problem. Simply, we can always divide a multipartite system into a bipartite. Then apply a yes/no criterion. The positive answer means that we can split both subsystems into new subsystems and apply the criterion again. The negative answer means that the states are not LUE. The bipartite splittings are being applied recursively until the system is divided into original particles. The last positive answer gives us the result that states are LUE. The described algorithm enables us to solve the multipartite LUE problem just by using the bipartite criterion. Let us now list some invariants of bipartite local unitary transformations

1. *Unitary equivalence.* It is well known that the necessary and sufficient condition for two states to be unitarily equivalent is that they have the same set of eigenvalues together with their multiplicity. The multiplicity of a given eigenvalue λ_a is represented by the dimensionality of the associated projector $\mathbf{P}_j$ ($\mathbf{Q}_j$) in spectral decomposition $\varrho = \sum_j \lambda_j \mathbf{P}_j$ ($\xi = \sum_k \mu_k \mathbf{Q}_k$). Let us denote by the set of eigenvalues of the density matrix ϱ $eig(\varrho) = \{\lambda_1, \ldots\}$. Then

$$\xi = \mathbf{U}\varrho\mathbf{U}^\dagger, \quad \text{iff} \quad eig(\xi) \equiv eig(\varrho) \quad \& \quad \dim \mathbf{P}_j = \dim \mathbf{Q}_j \tag{6.33}$$

2. *Unitary equivalence of the reduced states.* The reduced density matrices ϱ_X and ϱ_Y change under the action of the local unitary transformation $\mathbf{U}_X \otimes \mathbf{U}_Y$ as follows: for all $\mathbf{U}_X$ and $\mathbf{U}_Y$

$$\begin{aligned} \xi_X &= \mathrm{Tr}_Y \left[\mathbf{U}_X \otimes \mathbf{U}_Y \varrho \mathbf{U}_X^\dagger \otimes \mathbf{U}_Y^\dagger\right] = \mathbf{U}_X \varrho_X \mathbf{U}_X^\dagger \, , \\ \xi_Y &= \mathrm{Tr}_X \left[\mathbf{U}_X \otimes \mathbf{U}_Y \varrho \mathbf{U}_X^\dagger \otimes \mathbf{U}_Y^\dagger\right] = \mathbf{U}_Y \varrho_Y \mathbf{U}_Y^\dagger \, . \end{aligned} \tag{6.34}$$

It is obvious that under the action of the local unitary transformation the reduced states of ϱ and ξ are unitarily equivalent, i.e. the sets of eigenvalues coincide

$$\begin{aligned} eig(\varrho_X) &\equiv eig(\xi_X)\,, \\ eig(\varrho_Y) &\equiv eig(\xi_Y)\,. \end{aligned} \tag{6.35}$$

3. *Unitary equivalence of the reduced eigenprojectors.* Let $\mathbf{U} = \mathbf{U}_X \otimes \mathbf{U}_Y$ and $\varrho = \sum_k \lambda_k \mathbf{P}_k$, then $\xi = \mathbf{U}\varrho\mathbf{U}^\dagger = \sum_k \lambda_k \mathbf{Q}_k$, where $\mathbf{P}_k$ and $\mathbf{Q}_k = \mathbf{U}\mathbf{P}_k\mathbf{U}^\dagger$ are eigenprojectors belonging to the eigenvalue λ_k. Since the unitary transformation under the consideration is local, i.e. $\mathbf{U} = \mathbf{U}_X \otimes \mathbf{U}_Y$, we obtain the implication

$$\begin{aligned} \xi = \mathbf{U}\varrho\mathbf{U}^\dagger \qquad \text{then} \qquad & \dim(\mathbf{P}_k) = \dim(\mathbf{Q}_k) \\ & eig(\mathrm{Tr}_Y \mathbf{P}_k) \equiv eig(\mathrm{Tr}_Y \mathbf{Q}_k) \\ & eig(\mathrm{Tr}_X \mathbf{P}_k) \equiv eig(\mathrm{Tr}_X \mathbf{Q}_k)\,, \end{aligned} \tag{6.36}$$

where $\dim(\mathbf{P}_k)$ denotes the dimension of the projection $\mathbf{P}_k$ defined as the number of its non-zero eigenvalues.

4. *Partial transposition.* We shall use the following operator identity

$$(\mathbf{K}_X \otimes \mathbf{L}_Y \varrho \mathbf{M}_X \otimes \mathbf{N}_Y)^{\mathrm{Tr}_Y} = \mathbf{K}_X \otimes \mathbf{N}_Y^T \varrho^{T_Y} \mathbf{M}_X \otimes \mathbf{L}_Y^T \tag{6.37}$$

to prove that under the action of local unitary transformation also the partially transposed operators are related by a local unitary transformation. Introducing the local unitary transformation $\mathbf{U}_X \otimes \mathbf{U}_Y$ we obtain

$$\begin{aligned} (\xi)^{T_Y} &= (\mathbf{U}_X \otimes \mathbf{U}_Y \varrho \mathbf{U}_X^\dagger \otimes \mathbf{U}_Y^\dagger)^{T_Y} \\ &= \mathbf{U}_X \otimes (\mathbf{U}_Y^\dagger)^T \varrho^{T_Y} \mathbf{U}_X^\dagger \otimes \mathbf{U}_Y^T \\ &= \mathbf{W}\varrho^{T_Y}\mathbf{W}^\dagger \end{aligned} \tag{6.38}$$

where $\mathbf{W} = \mathbf{U}_X \otimes (\mathbf{U}_Y^T)^\dagger$ is a local unitary operator, because the transposed unitary operator is still unitary. And this proves our proposition, i.e. if two states are locally unitarily equivalent, then also their partial transpositions are. It follows that we can apply the above properties to verify the LUE of partially transposed operators in order to obtain the LUE of original states.

6.3.2 Bipartite LUE sufficient condition

For pure bipartite states the condition of local unitary equivalence is trivial. We know, that each pure bipartite state can be written in the Schmidt-decomposed form

$$\begin{aligned} |\psi\rangle &= \sum_k \sqrt{\lambda_k} |k\rangle_X \otimes |k\rangle_Y \\ |\phi\rangle &= \sum_k \sqrt{\mu_k} |k'\rangle_X \otimes |k'\rangle_Y. \end{aligned} \tag{6.39}$$

From the equality $\lambda_k = \mu_k$ (for all k) the validity of the second invariant property follows. That is, the reduced states contain the same eigenvalues, namely $\{\lambda_k\}$. The first invariant property is fulfilled trivially for pure states. Moreover, let us define local unitaries by equations $U_X|k\rangle_X = |k'\rangle_X$ and $U_Y|k\rangle_Y = |k'\rangle_Y$. Then the states (6.39) are locally unitary equivalent and $|\phi\rangle = U_X \otimes U_Y|\psi\rangle$.

Theorem 1

Let us consider two pure states $|\psi\rangle$ and $|\phi\rangle$ of a bi-partite system. These states are locally unitary equivalent if and only if their coefficients in the Schmidt decomposition are equal, i.e. two sets of eigenvalues of reduced states coincides.

Unfortunately, the Theorem 1 is valid only for pure states and cannot be used for arbitrary impure states. In what follows we will show a simple counter example. The idea behind this example is that we cannot create entanglement by local unitary transformation and that the spectral decomposition of a density matrix is unique. It means the separable eigenstate in spectral decomposition cannot evolve into entangled one.

Example 1.*(Counterexample)*

Let us consider two states $\varrho_{XY}, \sigma_{XY}$ of a bi-partite system $\mathcal{H}_{\mathcal{X}} \otimes \mathcal{H}_{\mathcal{Y}}$ where $dim\mathcal{H}_Y = 2$ and $dim\mathcal{H}_Y = 8$. It means the system of four qubits is divided into a one-qubit and a three-qubit subsystems. Let the set of eigenvalues be the same

$$eig(\varrho_{XY}) = eig(\sigma_{XY}) = \left\{ \frac{1}{4}, \frac{3}{8}, \frac{5}{16}, \frac{1}{16} \right\} \tag{6.40}$$

and the corresponding eigenvectors of ϱ_{XY} are

$$\begin{pmatrix} \frac{1}{\sqrt{4}} \left(|0,1\rangle + |0,2\rangle + |0,4\rangle + |1,0\rangle \right) \\ |1,7\rangle \\ |0,5\rangle \\ \frac{1}{\sqrt{3}} (|1,1\rangle + |1,2\rangle + |1,4\rangle) \end{pmatrix} \tag{6.41}$$

and for σ_{XY} we have

$$\begin{pmatrix} |1,7\rangle \\ \frac{1}{\sqrt{3}} (|0,7\rangle + |1,3\rangle + |0,5\rangle) \\ |0,0\rangle \\ |0,1\rangle \end{pmatrix} . \tag{6.42}$$

It is easy that for these states also the sets of eigenvalues of the reduced states coincide, i.e.

$$\begin{aligned} eig(\varrho_X) &= eig(\sigma_X) = \left\{ \frac{1}{2}, \frac{1}{2} \right\} \\ eig(\varrho_Y) &= eig(\sigma_Y) = \left\{ \frac{3}{8}, \frac{5}{16}, \frac{1}{4}, \frac{1}{16} \right\} \end{aligned} \tag{6.43}$$

So for the state under considerations the conditions of the Theorem 1 are satisfied. On the other hand, in the contradiction with the general properties of local unitaries, we see the violation of the property that the entanglement creation is prohibited by local unitary transformation. In other words, the condition of the same set of eigenvalues of the reduced eigenprojectors (invariant 3) is not fulfilled.

It is not clear whether the invariants presented in previous section provide the final answer for the problem of the local equivalence for mixed states. To see the difficulty of the problem we formulate the following theorem

Theorem 2

Suppose two non-degenerate states ϱ and σ of a composite system $X+Y$ are given by the relations

$$\varrho = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k| \, ; \qquad \sigma = \sum_k \mu_k |\phi_k\rangle\langle\phi_k| . \tag{6.44}$$

where $\lambda_k \neq \lambda_l$, $\mu_k \neq \mu_l$ for $k \neq l$. Let us express the vectors $|\psi_1\rangle$ and $|\phi_1\rangle$ in their Schmidt bases and fix these two bases on $\mathcal{H}_X \otimes \mathcal{H}_Y$. If for each k, $\lambda_k = \mu_k$ and the coefficients of the corresponding eigenvectors $|\psi_k\rangle = \sum_{m,n} \alpha^k_{mn} |mn\rangle$ and $|\phi_k\rangle = \sum_{m,n} \beta^k_{mn} |m'n'\rangle$ (where $|mn\rangle, |m'n'\rangle$ are the previously mentioned fixed bases) coincide, i.e. $\alpha^k_{mn} = \beta^k_{mn}$, then the states ϱ and σ are locally unitary equivalent. The local unitary transformation is given by $\mathbf{U}_X|m\rangle = |m'\rangle$ and $\mathbf{U}_Y|n\rangle = |n'\rangle$.

From the construction it is clear that the theorem is valid, but in some sense its content is trivial. We have mentioned this theorem only to illustrate how difficult the problem of LUE is.

6.3.3 Geometric analysis of equivalence classes

In this section we will calculate the number of parameters that must coincide to be sure that two states are LUE. From the mathematical point of view the local unitary transformations determine equivalence classes on the manifold of quantum states $\mathcal{S}(\mathcal{H})$. Since states belonging to the same equivalence class are mutually reachable by a local unitary transformation, we say that classes form *orbits* of the action of the group of local unitary transformations $U_{loc} \subset U(D)$. Let $d_j = \dim \mathcal{H}_j$ be the dimension of Hilbert space of the system A_j. Then the quantum state of the composite system $\varrho \in \mathcal{S}(\mathcal{H})$ is parametrized by $D^2 - 1$ real numbers, where $D = d_1 \dots d_N$ is the dimension of composite Hilbert space $\mathcal{H} = \mathcal{H}_{A_1} \otimes \dots \otimes \mathcal{H}_{A_N}$. It means that the manifold of states is a subset of the real space $\mathbb{R}^{D^2-1}$. The action of the group element $\mathbf{U} \in U_{loc}$ on the space $\mathcal{S}(\mathcal{H})$ is given by the *conjugation* $\mathbf{U}[\varrho] := \mathbf{U}\varrho\mathbf{U}^\dagger$.

For instance, in the Bloch-sphere representation of qubit states the orbits of all unitary transformations correspond to the spheres with different radiuses. In that case the ball represent the manifold of qubit's states and the orthogonal rotations $SO(3)$ of this ball correspond to actions of unitaries $SU(2)$. Note, that the irrelevance of the global phase allowed us to consider $SU(D)$ transformations instead of $U(D)$, since $U(D) = SU(D) \times U(1)$. In more dimensional cases the action of an unitary transformation $\mathbf{U} \in SU(D)$ represents an orthogonal rotation $SO(D^2 - 1)$, but the converse statement does not hold. That is, the set of representatives of $SU(D)$ in $SO(D^2 - 1)$ form just a specific subgroup of $SO(D^2 - 1)$.

Local transformations form a Lie group. The associated Lie algebra contains operators of the form $i\mathbf{A}_1 \otimes \mathbb{1}_2 \otimes \dots \otimes \mathbb{1}_N, i\mathbb{1}_1 \otimes \mathbf{A}_2 \otimes \dots \otimes \mathbb{1}_N, \dots, i\mathbb{1}_1 \otimes \dots \otimes \mathbf{A}_N$, where $\mathbf{A}_j$ are hermitian operators acting on the Hilbert space $\mathcal{H}_j$. Choose a basis $\mathbf{X}_l$ in the Lie algebra $u(d_1 \times \dots \times d_N)$. Then the general local unitary transformation $\mathbf{U}$ can be expressed as $\mathbf{U} = e^{i\mathbf{A}}$, where $\mathbf{A} = \sum_l \alpha_l \mathbf{X}_l$. It is well known in differential geometry, that each element of the Lie algebra $\mathbf{A}$ determines the vector field $\vec{\mathcal{V}}_{\mathbf{A}}$ on the manifold $\mathcal{S}(\mathcal{H})$ by the relation

$$\vec{\mathcal{V}}_{\mathbf{A}}[f(\varrho)] := \frac{\partial}{\partial t}[f(e^{it\mathbf{A}}\varrho e^{-it\mathbf{A}})]|_{t=0} \tag{6.45}$$

where $f : \mathcal{S}(\mathcal{H}) \to \mathbb{C}$. The span of vector fields at the point $\varrho \in \mathcal{S}(\mathcal{H})$ associated with the whole Lie algebra forms a tangent space to the orbit of local unitaries and so the dimension of this tangent space gives us the orbit's dimension at this point. That is, the number of local parameters. To calculate the dimension of LUE classes at the fixed point ϱ, we need to investigate the linear dependency of vector fields associated with the basis elements of the Lie algebra $u(d_1 \times \dots \times d_N)$.

For example, consider the system of two qubits. Let us make the following choice of the basis

$$\mathbf{X}_k^{(1)} = \sigma_k \otimes \mathbb{1}_2 \quad \text{and} \quad \mathbf{X}_k^{(2)} = \mathbb{1}_1 \otimes \sigma_k. \tag{6.46}$$

The general state can be expressed in the form $\varrho = \frac{1}{4}\mathbb{1} + a_k\mathbf{X}_k^{(1)} + b_k\mathbf{X}_k^{(2)} + \gamma_{kl}\sigma_k \otimes \sigma_l$ and the associated vector fields read

$$\begin{aligned} \vec{\mathcal{V}}_{\mathbf{X}_k^{(1)}} &= \sum_{j,m} \varepsilon_{kjm}\left(a_j\frac{\partial}{\partial a_m} + \sum_l \gamma_{jl}\frac{\partial}{\partial \gamma_{ml}}\right) & (6.47) \\ \vec{\mathcal{V}}_{\mathbf{X}_k^{(2)}} &= \sum_{j,m} \varepsilon_{kjm}\left(b_j\frac{\partial}{\partial b_m} + \sum_l \gamma_{lj}\frac{\partial}{\partial \gamma_{lm}}\right). & (6.48) \end{aligned}$$

Consider firstly the collection of vectors $\vec{\mathcal{V}}_{\mathbf{X}_k^{(1)}}$. The condition $\sum_k \alpha_k \vec{\mathcal{V}}_{\mathbf{X}_k^{(1)}} = 0$ holds for all ϱ only if $\alpha_k = 0$. That is, vectors $\vec{\mathcal{V}}_{\mathbf{X}_k^{(1)}}$ are mutually linearly independent. The same result holds for the second collection of vector fields $\vec{\mathcal{V}}_{\mathbf{X}_k^{(2)}}$. Mutually these two collections are also independent. Hence, the dimension of the orbit of a *generic* density operator of two qubits is 6 and the number of non-local parameters is $15 - 6 = 9$. We have seen that the dimension of the orbit in the general case was equal to the number of generators of the associated Lie algebra $su(2) \times su(2)$, because for a general state ϱ the associated vector fields are linearly independent.

In the general case of the composite system of arbitrary subsystems with arbitrary dimension the situation is similar [40]. The basis of the Lie algebra of local unitary transformations is given by operators $\mathbf{X}_l^{(j)} = \Theta_l^{(j)} \otimes \mathbb{1}_{N\setminus j}$. The general quantum state takes the form

$$\varrho_{1\ldots N} = \frac{1}{D}\mathbb{1} + \sum_{j=1}^{N} \sum_{l=1}^{d_j^2-1} a_l^{(j)} \mathbf{X}_l^{(j)} + \sum_{l_1,\ldots,l_N} \Gamma_{l_1,\ldots,l_N} \Theta_{l_1}^{(1)} \otimes \ldots \otimes \Theta_{l_N}^{(N)}. \tag{6.49}$$

Similar consideration as in the two qubit case show, that the associated vector fields $\vec{\mathcal{V}}_{\mathbf{X}_l^{(j)}}$ are linearly independent. Hence, the dimension of orbits for the general case is $d_1^2 + \ldots + d_N^2 - N$, that is, the number of generators of $u(d_1 \times \ldots \times d_N)$. Consequently the number of non-local parameters $\mathcal{N}$ equals

$$\mathcal{N} = \prod_j d_j^2 - \sum_j d_j^2 + N - 1\ . \tag{6.50}$$

We have derived the following theorem.

Theorem

General state ϱ of a multipartite system is characterized by $D^2 - 1$ real parameters, where $D = d_1 \ldots d_N$ and $d_j = \dim \mathcal{H}_j$. The number

$$\mathcal{N}_L = \sum_{j=1}^{N} (d_j^2 - 1) \tag{6.51}$$

of them is preserved under the action of local unitary transformations and the rest

$$\mathcal{N} = \prod_j d_j^2 - \sum_j d_j^2 + N - 1 \tag{6.52}$$

parameters characterize the non-local properties of states.

To answer, whether two given states belong to the same orbit, we need to compare their local parameters, that uniquely determine the orbits. Any invariant of local unitaries must be a function of these parameters. How easy, but we did not say, which of the parameters are local. And this problem is the aim of the future investigation [37, 38, 39, 40].

6.3.4 Another equivalence relations

In different applications different equivalence relations can be of interest. Let us denote by $\mathcal{T}$ the group of transformations that factorizes the state space into the equivalence classes. In the quantum theory of entanglement one can meet with the following types of sets $\mathcal{T}$: all unitary transformations $\mathcal{T} = \mathcal{U}(\mathcal{H}) = U(D) = SU(D)$, local unitary transformations $\mathcal{T} = \mathcal{U}_{loc} = U(d_1) \times \ldots \times U(d_N)$ (see [37]-[40]), stochastic local transformations (SLOCC) $\mathcal{T} = SL(d_1) \times \ldots \times SL(d_N)$ (see [41]) and local operations with classical communications (LOCC).

It is known (see Ref[41]) that according to SLOCC each pure state of three qubits must belong into one of the six equivalence classes (orbits)
1. A-B-C factorizable states, i.e. $|\psi\rangle_{ABC} = |000\rangle_{ABC}$
2. AB-C factorizable states, i.e. $|\psi\rangle_{ABC} = |\phi\rangle_{AB} \otimes |0\rangle_C$
3. A-BC factorizable states, i.e. $|\psi\rangle_{ABC} = |\phi\rangle_{BC} \otimes |0\rangle_A$
4. AC-B factorizable states, i.e. $|\psi\rangle_{ABC} = |\phi\rangle_{AC} \otimes |0\rangle_B$

5. W type states, i.e. $W = |010\rangle + |001\rangle + |100\rangle$
6. GHZ type states, $GHZ = |000\rangle + |111\rangle$
We remind us that SLOCC operations are given by invertible linear transformations of the type $\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}$ with $\det \mathbf{A} = \det \mathbf{B} = \det \mathbf{C} = 1$. Note that the local unitary transformations cannot change the graph representation of the state, but the SLOCC transformations can. In fact, the divisions of the state space made by graphs and by SLOCC are not compactible.

6.4 Quantification of multipartite entanglement

Several times we have mentioned that the measures of bipartite entanglement are known, but their evaluation is usually very difficult. Therefore, we cannot expect that the multipartite measures will be different. However, the task is not only to quantify the entanglement, but rather to verify and identify some features of multipartite systems. In this section we shall study the multipartite entanglement in multi-qubit physical systems.

6.4.1 Three-partite entanglement and correlations

Not only the question of the three-partite entanglement, but also the field of three-partite correlations have not still be satisfactorily investigated. For example, consider that Alice, Bob and Carol share three qubits described by GHZ state. Let them know the individual qubit states only. It is obvious that in this case they do not have any information about the entanglement or correlations of the joint state. But let them know also the bipartite states. Are they able to say anything about the shared entanglement? In the case of GHZ state no pair shares entanglement, so it seems that it is reasonable to conclude that their three qubits contain no entanglement at all. And, of course, this is not true, because we know that the overall state is pure and the couples are mutually correlated. Any occurrence of correlations in pure states implies that the state is not factorized, but entangled. The aim of this section is to understand the mutual relations between the bipartite and three-partite correlations and entanglement. Can the entanglement be shared in an arbitrary way, or not?

The example of GHZ state provides us with a curious phenomenon of three-partite entanglement. Despite of the nonexistence of bipartite entanglement the joint state is entangled. We shall refer to this property as to *intrinsic three-partite entanglement*, because it cannot be reduced (in any sense) into the bipartite one. The various functionals (entropy, purity, determinant, concurrence) defined on quantum states will be important in our analysis. For instance, one can derive the following inequality [22].

CKW inequality

To calculate the tangle τ_{AB} (= square of the concurrence) we need to calculate the eigenvalues ($\lambda_1^2 \geq \lambda_2^2 \geq \lambda_3^2 \geq \lambda_4^2$) of the operator $R_{AB} = \varrho_{AB}(\sigma_y \otimes \sigma_y)\varrho_{AB}(\sigma_y \otimes \sigma_y)$. In particular, $\tau_{AB} = [\max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}]^2$. We remind us that the entanglement of formation is connected with the tangle by the formula $E_f(\varrho_{AB}) = H_{bin}(\frac{1}{2} + \frac{1}{2}\sqrt{1-\tau_{AB}})$ where $H_{bin}(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function. For the special cases, when ϱ_{AB} is pure, the tangle is proportional to the determinant of the reduced state ϱ_A, i.e. $\tau_{AB} = 4\det\varrho_A$. In our case the fact that the state of three qubits is pure implies that the matrix R_{AB} can have maximally two nonzero eigenvalues, i.e.

$$\tau_{AB} \quad = \quad (\lambda_1 - \lambda_2)^2 = \lambda_1^2 + \lambda_2^2 - 2\lambda_1\lambda_2 = \mathrm{Tr}R_{AB} - 2\lambda_1\lambda_2 \leq \mathrm{Tr}R_{AB} \tag{6.53}$$

The same can be done with τ_{AC} to get the inequality $\tau_{AC} \leq \mathrm{Tr}R_{AC}$. Combining these two inequalities we obtain

$$\tau_{AB} + \tau_{AC} \leq \mathrm{Tr}R_{AB} + \mathrm{Tr}R_{AC} \tag{6.54}$$

Let us calculate the matrix R_{AB} for a general three-qubit pure state $|\Psi\rangle = \sum_{j,k,l}\psi_{jkl}|jkl\rangle$. Introducing the antisymmetric matrix with elements $\epsilon_{01} = -\epsilon_{10} = 1$ and $\epsilon_{00} = \epsilon_{11} = 0$, the trace of the R_{AB} reads

$$\mathrm{Tr}R_{AB} = \sum \psi_{jkl}\psi^*_{mnl}\epsilon_{mm'}\epsilon_{nn'}\psi^*_{m'n'p}\psi_{j'k'p}\epsilon_{j'j}\epsilon_{k'k} \tag{6.55}$$

because the elements of the matrix $\sigma_y \otimes \sigma_y$ can be expressed via the antisymmetric matrix $(\sigma_y \otimes \sigma_y)_{mn,m'n'} = \langle mn|\sigma_y \otimes \sigma_y|m'n'\rangle = \epsilon_{mm'}\epsilon_{nn'}$. Using the identity $\epsilon_{nn'}\epsilon_{k'k} = \delta_{nk'}\delta_{n'k} - \delta_{nk}\delta_{n'k'}$ the above equation can be rewritten as

$$\mathrm{Tr}R_{AB} = 2\det\varrho_A - \mathrm{Tr}\varrho_B^2 - \mathrm{Tr}\varrho_C^2 \tag{6.56}$$

Because the following identity $\mathrm{Tr}\varrho^2 = 1 - 2\det\varrho$ holds for two-dimensional matrices with unit trace, we get

$$\mathrm{Tr}R_{AB} = 2(\det\varrho_A + \det\varrho_B - \det\varrho_C) \tag{6.57}$$

By the symmetry also

$$\mathrm{Tr}R_{AC} = 2(\det\varrho_A + \det\varrho_C - \det\varrho_B) \tag{6.58}$$

Combining the last two equations we obtain the CKW inequality

$$\tau_{AB} + \tau_{AC} \le 4\det\varrho_A \tag{6.59}$$

The right hand side of this inequality can be used as a measure of entanglement shared between the qubit A and the rest two qubits, i.e. $\tau_{A(BC)} \equiv [C_{A(BC)}]^2 := 4\det\varrho_A$.

This inequality together with its generalization to multi-qubit systems takes name by their inventors *Coffman, Kundu and Wootters*, i.e. *CKW inequality.* The validity of this inequality has been proved for three-qubit system and it was conjectured that it should be satisfied also for multi-qubit system. Let us denote by C_{jk} the concurrence between the jth and kth qubit, and by $C_{j,\bar{j}}$ the concurrence between the jth qubit and rest of the composite system. Let us now formulate the CKW conjecture:

Conjecture(*Coffman,Kundu, Wootters*)
Consider the system composed from N qubits in a pure multi-qubit state. Then
(a) the inequality

$$\sum_{k,k\neq j} C_{jk}^2 \le C_{j,\bar{j}}^2 \tag{6.60}$$

holds for each qubit j
(b) the difference

$$\Delta_j = C_{j,\bar{j}}^2 - \sum_{k,k\neq j} C_{jk}^2 \tag{6.61}$$

has the same value for all j.

Note that the (b) part of the conjecture did not appear in the original formulation of the conjecture. However, for three qubits this part is also fulfilled. Moreover, if one wants to use this difference in order to quantify the *intrinsic entanglement* (= the entanglement shared in multipartite systems which is not shared in bipartite way), then it is reasonable to require such property.

In what follows we shall analyze the states of pure three-qubit states with respect to CKW inequalities. The $|GHZ\rangle$ state corresponds to graph with no lines, i.e. the bipartite concurrencies vanish, but the entanglement between each qubit and the rest qubit pair does not, i.e. $C_{j,\bar{j}} > 0$ and consequently $\Delta_j > 0$. Therefore, it seems that the difference Δ_j can be used to measure three-partite entanglement. On the other hand the second typical representative of three-partite entangled states is the $|W\rangle$ state, for which one can easily verify that $C_{AB} = C_{AC} = C_{BC} = 2/3$ and, consequently, the CKW inequality is saturated, because $\tau_A = 8/9 = C_{AB}^2 + C_{AC}^2$.

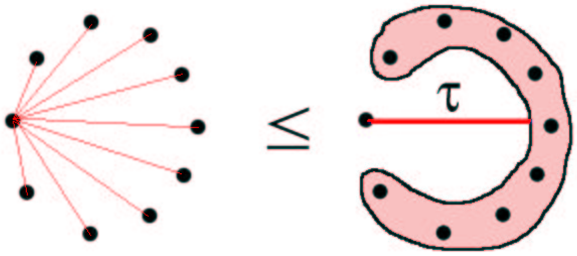

Figure 6.2: CKW inequality

6.5 Saturation of CKW inequalities

6.5.1 Basic observations

At the end of the original paper [22] it is stated that states in the subspace covered by the basis $\{|1\rangle_j \otimes |0^{\otimes N-1}\rangle_{\bar{j}}\}$ saturates the CKW inequalities. The question of the most general state for which the saturation holds is well defined mathematical problem. Unfortunately, the solution is far from being known. In what follows we will give arguments why any state of the form

$$|\Psi\rangle = \alpha_0|0^{\otimes N}\rangle + \sum_{j=1}^{N} \alpha_j|1\rangle_j \otimes |0^{\otimes(N-1)}\rangle \tag{6.62}$$

saturates the CKW inequalities.

The bipartite density matrices of the state under consideration takes the form

$$\varrho = \begin{pmatrix} a & d & e & 0 \\ d^* & b & f & 0 \\ e^* & f^* & c & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.63}$$

and only one element determines the concurrence of such state, namely

$$C^2(\varrho) = 4ff^* \,. \tag{6.64}$$

It follows that only the matrix element standing with the $|01\rangle\langle 10|$ term will be important for us. Direct calculations lead us to the following value $f = \alpha_j\alpha_k^*$ for the pair of qubits (jk). The corresponding square of the concurrence between j-th and k-th qubit equals

$$C_{jk}^2 = 4|\alpha_j|^2|\alpha_k|^2 \,. \tag{6.65}$$

In the next step we evaluate the tangle between the j-th qubit and the rest of the system. The state of single qubit is described by matrix

$$\varrho_j = \begin{pmatrix} |\alpha_0|^2 + \sum_{k\neq j}|\alpha_k|^2 & \alpha_0\alpha_j^* \\ \alpha_j\alpha_0^* & |\alpha_j|^2 \end{pmatrix} \tag{6.66}$$

Now it is easy to check that

$$\tau_j = 4\det\varrho_j = 4|\alpha_j|^2\sum_{k\neq j}|\alpha_k|^2 = \sum_{k\neq j} C_{jk}^2 \tag{6.67}$$

and therefore the CKW inequalities are saturated, i.e. $\Delta_j = \tau_j - \sum_{k\neq j} C_{jk}^2 = 0$ for all j.

Next we will show that part (b) of the CKW conjecture is no longer valid for multi-qubit pure states (for $N > 3$). For instance, consider the state of four qubits $|\Phi\rangle = |0\rangle \otimes |GHZ\rangle = \frac{1}{\sqrt{2}}(|0000\rangle + |0111\rangle)$. In such case it is easy to show that $\Delta_1 = 0$, since the first qubit is separated from other three. But, if we evaluate Δ_2 we obtain a nonvanishing value. In this state all bipartite concurrencies are zero, i.e. $C_{jk}^2 = 0$, but the tangles are different. For the second qubit its tangle equals $\tau_2 = 1$. It means that $\Delta_2 = \tau_2 = 1 \neq \Delta_1$. It proves that the part (b) of the conjecture does not hold for more than three qubits.

The last result a little decreases the valuability of the difference Δ_j for the characterization of the multi-partite entanglement. In what follows we will assume that a state ϱ saturates CKW inequality only if $\Delta_j = 0$ for all values of j. Our class of states $|\Psi\rangle$ can be generalized in the following way. Any state of multi-qubit system of the form

$$|\Psi\rangle = |0^{\otimes(N-n)}\rangle \otimes |\Psi_n\rangle \tag{6.68}$$

saturates the CKW inequality. We used the notation, where $|\Psi_n\rangle$ is the state defined in Eq.(6.62) for n qubits, whereas the state $|\Psi\rangle$ describes N qubits. The saturation of such state can be easily verifies. Next consider the factorized state of two states of the form (6.62), i.e.

$$|\Psi\rangle = |\Psi_n\rangle \otimes |\Psi_{N-n}\rangle \tag{6.69}$$

Denote the set of n qubits by symbol A and the rest qubits by B. Such states again saturate the CKW inequality. The proof is straightforward. We can divide the set of N qubits in as many parties as we want. If each of them is described by the state of the form (6.62), or it is in factorized state, then the whole state saturates CKW inequality, because all its sub-parties do so. Another kind of generalization uses the concept of local unitary equivalence between two states. Since tangles and concurrencies are invariant under local unitary transformations, it follows that also the difference Δ_j is invariant. Therefore, if one of the states (say $|\Psi\rangle$) saturates CKW, then also the second one ($|\Phi\rangle$) does.

In conclusion each state of N qubits

$$|\Psi\rangle = |\Psi_{n_1}\rangle \otimes \ldots \otimes |\Psi_{n_M}\rangle \tag{6.70}$$

where $|\Psi_{n_j}\rangle$ are states locally unitary equivalent to state (6.62) saturates CKW inequalities. The question of the general class of states, for which CKW saturation holds, we left open.

6.5.2 Multipartite entanglement in collision processes

A. Homogenization

Within the context of our investigation it is very natural to ask, what is the nature of the entanglement created during the process of homogenization. In this section we will address several questions related to this issue. Let us consider a specific initial state of the system and the reservoir: $|\psi\rangle_0 = |1\rangle$ and $|\xi\rangle_j = |0\rangle$. Note that the partial swap $\mathbf{P}_\eta$ is invariant according the local unitary transformations of the type $U \otimes U$. That is, it takes the same form in any basis (of a single qubit) we choose. It follows that our further calculations remain valid for any mutually orthogonal states, not only for $|0\rangle$ and $|1\rangle$ and this makes our calculations much simpler.

With the given initial conditions, we easily find the state vector describing the whole system after n interactions:

$$|\Psi_n\rangle = c^n|1\rangle_0 \otimes |0\rangle^{\otimes N} + \sum_{l=1}^{n} |1\rangle_l \otimes |0\rangle^{\otimes N_{\bar{l}}} \left[isc^{l-1}(c+is)^{N-l}\right]$$

The state $|0\rangle^{\otimes N_{\bar{l}}}$ denotes a state in which all qubits except the lth one are in the state $|0\rangle$. For a general pure system state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ the input state of the whole system $|\psi\rangle \otimes |0^{\otimes N}\rangle$ evolves after n interactions into the state

$$|\Omega_n\rangle = \alpha|0^{\otimes(N+1)}\rangle + \beta|\Psi_n\rangle \tag{6.71}$$

In what follows we will find out the explicit expression for all bi-qubit and single qubit states during the process of homogenization and then apply the definitions of concurrence and tangles to find the value of shared entanglement. It is easy to see that the joint pure state $|\Omega_n\rangle$ is of the form (6.62), for which the values of all concurrencies can be given explicitly.

$$\tau_{jk}^{(n)} = [C_{jk}^{(n)}]^2 = \begin{cases} 0 & \text{for} \quad n < k \leq N \\ 4|\beta|^4 s^4 c^{2(j+k-2)} & \text{for} \quad k \leq n \leq N \end{cases} \tag{6.72}$$

$$\tau_{0k}^{(n)} = [C_{0k}^{(n)}]^2 = \begin{cases} 0 & \text{for} \quad n < k \leq N \\ 4|\beta|^4 s^2 c^{2(n+k-1)} & \text{for} \quad k \leq n \leq N \end{cases} \tag{6.73}$$

$$\tau_{j}^{(n)} = [C_{j,\bar{j}}^{(n)}]^2 = \begin{cases} 0 & \text{for} \quad n < j \leq N \\ 4|\beta|^4 s^2 c^{2(j-1)}(1 - s^2 c^{2(j-1)}) & \text{for} \quad j \leq n \leq N \end{cases} \tag{6.74}$$

$$\tau_{0}^{(n)} = [C_{0,\bar{0}}^{(n)}]^2 = 4|\beta|^4 c^{2n}(1 - c^{2n}) \tag{6.75}$$

Moreover, the form of the state guarantees that during the whole process of homogenization the CKW inequalities are saturated, i.e. $\Delta_j = 0$.

These results show that system qubit acts as a mediator of entanglement between the reservoir qubits, which have never interacted directly. It is obvious that the later the two reservoir qubits interact with the system qubit, the smaller the degree of their mutual entanglement is. Nevertheless, this value is constant and does not depend on the subsequent evolution of the system qubit (i.e., it does not depend on the number of interactions n). On the other hand between the system qubit and j-th reservoir qubit (j is arbitrary) the entanglement monotically decreases with the number of interaction steps.

B. CNOT-ization

In what follows we will use the same model of particle-reservoir dynamics, only instead of the partial `SWAP` we shall use the partial `CNOT`, i.e.

$$P_\eta = \cos\eta\mathbb{1} + i\sin\eta\texttt{CNOT} = c\mathbb{1} + is\texttt{CNOT} \tag{6.76}$$

Note that the action of the `CNOT` operation can be understood in two nonequivalent ways. In the first case the system qubit plays the role of the control qubit and in the second case the reservoir qubits are the control ones.

Let us start with the first, i.e. $\texttt{CNOT} = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \sigma_x$. The joint system of the control and target qubits evolves according to rule

$$\begin{aligned} \varrho \otimes \xi \quad \mapsto \quad & c^2 \varrho \otimes \xi + s^2\{\varrho_{00}P_{00} \otimes \xi + \varrho_{11}P_{11} \otimes \sigma_x\xi\sigma_x\} \\ & + s^2\{\varrho_{01}P_{01} \otimes \xi\sigma_x + \varrho_{10}P_{10} \otimes \sigma_x\xi\} \\ & + isc\{[\varrho, P_{00}] \otimes \xi + [P_{11} \otimes \sigma_x, \varrho \otimes \xi]\} \end{aligned} \tag{6.77}$$

with $P_{jk} = |j\rangle\langle k|$ and $\varrho_{jk} = \langle j|\varrho|k\rangle$. Then for the subsystems it follows

$$\varrho_S^{(n)} = \begin{pmatrix} \varrho_{00} & \kappa^n \varrho_{01} \\ \overline{\kappa}^n \varrho_{10} & \varrho_{11} \end{pmatrix} \tag{6.78}$$

with

$$\kappa := ce^{i\eta} + se^{-i\eta}\langle\sigma_x\rangle_\xi \tag{6.79}$$

where $\langle\sigma_x\rangle_\xi := \mathrm{Tr}\xi\sigma_x$ denotes the mean value of the x-spin component and

$$\xi'_j = c^2\xi + s^2\varrho_{00}\xi + s^2\varrho_{11}\sigma_x\xi\sigma_x \tag{6.80}$$

Hence, in this case the CNOTization leads the state of the system into its diagonal form, i.e. the system qubit decoheres. On the other hand the qubits in the reservoir are described by the same physical state $\xi' = \xi'_j$ in each step.

Next, let us assume that reservoir qubits are the controlled ones, i.e. $\texttt{CNOT} = \mathbb{1}\otimes|0\rangle\langle 0| + \sigma_z\otimes|1\rangle\langle 1|$. In this case the evolution leads the qubits into the states

$$\varrho_S^{(n)} = c^2\varrho + s^2\xi_{00}\varrho + s^2\xi_{11}\sigma_x\varrho\sigma_x \tag{6.81}$$

and

$$\xi' = \begin{pmatrix} \xi_{00} & \kappa\xi_{01} \\ \overline{\kappa}\xi_{10} & \xi_{11} \end{pmatrix} \tag{6.82}$$

with $\kappa = se^{i\eta}\langle\sigma_x\rangle_\varrho + ce^{-i\eta}$.

Consider the initial state $|\Omega_0\rangle = (\alpha|0\rangle + \beta|1\rangle)\otimes|0^{\otimes N}\rangle$. If the reservoir qubits play the role of control qubits, then the evolution is trivial, i.e. $|\Omega_n\rangle = |\Omega_0\rangle$. In the second settings, when system qubit is the controlled one, the state $|\Omega_n\rangle = (\alpha|0^{\otimes(n+1)}\rangle + \beta|1^{\otimes(n+1)}\rangle)\otimes|0^{\otimes(N-n)}\rangle$, i.e. the resulting state is of GHZ type. Therefore, we can conclude that the "type" of entanglement created in this process is completely different in comparison with the homogenization. Simply, in this case all bipartite states remain separable (contain no entanglement) and the difference $\Delta_j = 4|\alpha|^2|\beta|^2$ for $j \leq n$. For $j > n$ the difference vanishes.

Conclusion

As we have already mentioned the problems studied in this thesis can be divided into four groups:

1. *Quantum processors*
2. *Quantum dense coding*
3. *Quantum homogenization*
4. *Multi-partite entanglement*

Quantum processors were analyzed in the Chapter V., where we explicitly constructed the processor (`QDM`) that implements all unitary transformations in the probabilistic regime. The same processor was exploited to realize the complete state reconstruction of the data system. We indicated several classes of quantum processors with similar features (*U-processors, Y-processors, covariant processors, maximal processors, etc.*). We investigated the possibilities of deterministic implementation of one-parametric sets of quantum maps, where we have found a useful relation that must hold between any two quantum maps realizable by the same quantum processor. We used this relation to show that there cannot exist a *universal quantum processor*, and that the rate of the *amplitude damping channel* cannot be controlled by a fixed quantum processor. Finally, also the case of the probability amplification was studied. We have shown that *conditioned loops* are possible and they enable us to improve the probability. However, we have left open the question, whether such conditioned loops can be replaced by the action of a new processor (with a larger program space) that does the same job.

The second problem of the quantum dense coding continues and extends my Diploma Thesis, where I was studied this quantum communication protocol for different types of quantum qubit channels in detail. In this thesis I have paid attention only to noiseless quantum channels, because I have been interested in the general properties of the entanglement and correlations in the information transmition. I have found that the quantum dense coding can be understood as a generalization of the only secure protocol called *one-time pad.* Also the relation between the derived capacities and the measure of entanglement has been given in a clear way. The formulas for the qubit channel capacities have been generalized into the case of qudit noiseless channel.

The quantum homogenization is a physical process originally introduced in [52]. We have found many interesting properties of this simple model of the interaction between a single qubit and a reservoir of qubits. This single qubit evolves according to the model of homogenization and finally all qubits are described by the same state, which is equal to the original state of the qubits in the reservoir. We have shown that only the partial swap interaction satisfies the conditions given by the model. Different aspects of the homogenization can be found in different chapters of the thesis. The model was introduced in the Chapter IV. Homogenization process is discrete. However, we have shown that it is possible to introduce continuous time to describe it. Moreover, this continuous extension possesses semigroup properties. Partial swaps belong also to a specific class of processors, for which each program determines different quantum operation, i.e. there is no redundancy in the program space of such processor. Namely, each state ξ of the reservoir determines a contractive map with the fixed point ξ. Also in the last chapter the homogenization has its representative. The study of the entanglement properties of the states generated by the homogenization process leads us to the following result. During the whole process, the shared entanglement shows similar features, namely, the CKW inequalities are saturated.

The entanglement of the multi-partite system is the central notion of the last chapter. The meaning of the entanglement in composite system is still a very vague concept. In this chapter we introduced the graph representation of quantum states, where also correlations (without entanglement) was involved. We paid attention to the analysis of three-partite systems in order to show, which of the graphs could be represented by quantum states. Very few results are known about the multi-partite entanglement. The mentioned CKW inequality demonstrates that entanglement in multi-partite pure states cannot be shared freely. The problem of local unitary equivalence has been solved only partially. However, its difficulty is famous. So far, only a few features of multi-partite entanglement have been studied. The relation between the graph representation and entanglement properties of multi-partite system is trivial, but it can be that some deeper relations could exist. In these days the graph theory is not applicable into the problem of entanglement and we use graphs only as illustrations of different entanglement configurations. To introduce a reasonable graph theory one needs to specify the properties of entanglement, some new inequalities, etc. This is an open problem left for the future investigation.

Bibliography

[1] A.Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993)

[2] J.Preskill, *Quantum theory of Information and Computation* see `http://www.theory.caltech.edu/people/preskill`

[3] J.Pišút, V.Černý, V.Gomolčák, *Úvod do kvantovej mechaniky* (Alfa, Bratislava, 1983)

[4] B.Riečan, T.Neubrunn, *Teória miery*, (Veda, Bratislava, 1992)

[5] P.Bóna, *"Extended Quantum mechanics"*, *Acta Phys. Slovaca* **50**, (2000)

[6] D.Blochincev,*Osnovy kvantovoj mechaniki*, (Nauka, Moskva, 1967)

[7] J.Blank, P.Exner, M.Havlíček, *Lineární operátory v kvantové fyzice*, (Karolinum, Praha, 1993)

[8] M.A.Nielsen and I.L.Chuang *Quantum Computation and Quantum information* (Cambridge University Press, Cambridge, 2000)

[9] R.F.Streater, *"Classical and quantum probability"*, *J.Math.Ph.* **41**, 3556-3603 (2000)

[10] A.S.Holevo, *Probabilistic and statistical aspects of quantum theory*, (North-Holland publishing company, 1982)

[11] G.Alber, T.Beth, M.Horodecki, P.Horodecki, R.Horodecki, M.Rötteler, H.Winfurter, R.F.Werner, A.Zelineger *Quantum information-an introduction to basic theoretical concepts and experiments* Verlag, Berlin, 2001 (Springer Tracts in Modern Physics vol.173)

[12] John S. Bell, *"On the Einstein-Podolski-Rosen paradox"*, *Phsysics*, **1**, 195-200 (1964)

[13] R.F.Feynman, *Kvantovomechaniceskoje EVM*, *Uspekhi Phys. Nauk* **149**:4 671-688 (1986)

[14] V.Vedral and M.B.Plenio, *"Entanglement measures and purification procedures"*, *Phys.Rev. A* **57**, 1619-1633 (1998)

[15] K.G.H.Vollbrecht and R.F.Werner, *"Why two qubits are special"*, LANL preprint archive `quant-ph/9910064`

[16] K.G.H.Vollbrecht, R.F.Werner, *"Entanglement measures under symmetry"*, LANL preprint archive `quant-ph/0010095`

[17] R.F.Werner, *"All teleprotation and dense coding schemes"*, LANL preprint archive `quant-ph/0003070`

[18] A.Peres, *"Separability criterion for density matrices"*, *Phys.Rev.Lett.* **77** (1996), LANL preprint archive `quant-ph/9604005`

[19] S.L.Woronovicz, *"Positive maps of low dimensional matrix algebras"*, *Rep.Math.Phys.* **10**, 165-183, (1976)
S.L.Woronowicz, *"Nonextendible positive maps"*, *Commun.Math.Ph.* **51**, 243-282 (1976)

[20] M.Horodecki, P.Horodecki, R.Horodecki, *"Separability of mixed states: necessary and sufficient conditions"*, *Phys.Lett. A* **223**, 1-8 (1996), LANL preprint archive quant-ph/960538

[21] P.Horodecki and A.Ekert, *"Direct detection of quantum entanglement"*, LANL archive quantu-ph/0111064

[22] V. Coffman, J. Kundu, W. K.Wootters, *"Distributed entanglement"*, *Phys.Rev. A* 61, 052306 (2000).

[23] W. K. Wootters, *"Entanglement of formation of an arbitrary state of two qubits". Phys. Rev. Lett.* **80**, 2245 (1998).

[24] M. Koashi, V. Bužek, and N. Imoto, Phys. Rev. A **62**, 050302(R) (2000),

[25] W. Dür, *"Entanglement molecules"*, Phys. Rev. A **63**, 020303 (R) (2001), LANL archive quant-ph/0006105

[26] M.J.Donald, M.Horodecki, O.Rudolph, *"The uniqueness theorem for entanglement neasures"* LANL preprint archive quant-ph/0105017

[27] B.M.Terhal, *"A family of indecomposable positive linear maps based on entangled quantum states"*, LANL preprint archive quant-ph/0010091

[28] A.S.Holevo, *IEEE Trans. Inf. Theory* **44**, 269 (1998).

[29] A.S.Holevo, *"Quantum coding theorems"*, *Russian Math. Surveys* **53**:6 1295-1331 (1998), *Uspekhi Mat. Nauk* **53**:6 193-230 (1998)

[30] M.Raginsky, *"Strictly contractive quantum channels and physically realizable quantum computers"*, *Phys.Rev.A* **65**, 032306 (2002), LANL preprint archive quant-ph/0105141

[31] V. Bužek and M. Hillery, *"Quantum copying: Beyond the No-Cloning Theorem" Phys. Rev. A* **54**, 1844 (1996).

[32] V. Bužek, S. Braunstein, M. Hillery, and D. Bruß, *Phys. Rev. A* **56**, 3446 (1997).

[33] Peter Štelmachovič and Vladimír Bužek, *"Dynamics of open quantum systems initially entangled with enviroment: Beyond the Kraus representation"*, *Phys.Rev.A* **64** (2001), LANL preprint archive quant-ph/0108136

[34] Mário Ziman and Peter Štelmachovič, *"Quantum Theory: kinematics, linearity and no-signaling condition"*, LANL preprint archive quantu-ph/0211149

[35] G. Lindblad, *On the generators of Quantum Dynamical Semigroups*, *Commun. Math. Ph.* **48**, 119-130 (1976)

[36] Ch.Fuchs, *"Nonorthogonal quantum states can maximize classical information capacity"*, *Phys.Rev.Lett.* **79**, 1162-1165 (1997), LANL preprint archive quant-ph/9703043

[37] A.Sudbery, *"On local invariants of pure three-qubit states" J.Phys.A* **34**, 643-652 (2001), LANL preprint archive quant-ph/0001116

[38] H.Barnum and N.Linden, *"Monotones and invariants for multi-particle quantum states"*, LANL preprint archive quant-ph/0103155

[39] M.Kus and K.Zyckowski, *"Geometry of entangled states"*, *Phys.Rev.A* **63**, 032307 (2001), LANL preprint archive `quant-ph/0006068`

[40] N.Linden, S.Popescu and A.Sudbery, *"Non-local properties of multiparticle density matrices"*, *Phys.Rev.Lett* **83**, 243-247 (1999), LANL preprint archive `quant-ph/9801076`

[41] W.Dür, G.Vidal and J.I.Cirac, *"Three qubits can be entangled in two inequivalent ways"*, *Phys.Rev. A* **62** (2000), LANL preprint archive `quant-ph/0005115`

[42] C.H.Bennett and S.Wiesner, *Phys.Rev.Lett.* **69**, 2881 (1992)

[43] M.Ziman, V.Bužek, *"Equally distant partially entangled alphabet states for quantum channels"*, *Phys.Rev* **62** (2000), LANL preprint archive `quant-ph/0009075`

[44] M. A. Nielsen and I. L. Chuang, *"Programmable quantum gate arrays"*, *Phys. Rev. Lett.* **79**, 321 (1997).

[45] G. Vidal, L. Masanes, and J.I. Cirac, *"Storing quantum dynamics in quantum states: A stochastic programmable gate"*, *Phys.Rev.Lett.* **88** 047905 (2002), Los Alamos arXiv `quant-ph/0102037.`

[46] S. Braunstein, V. Bužek, and M. Hillery, *"Quantum Information Distributor:Quantum network for symmetric and anti-symmetric cloning in arbitrary dimension and continuous limit"*, *Phys. Rev. A* **63**, 052313 (2001).

[47] M.Hillery, V.Bužek, M.Ziman, *"Programmable quantum gate arrays"*, *Forstschritte der Physik* **49**, 987-992 (2001)

[48] M.Hillery, V.Bužek, M.Ziman, *"Probabilistic implementation of quantum processors"*, *Phys.Rev A* **65**, (2001) LANL preprint archive `quant-ph/0106088`

[49] M.Ziman, V.Bužek, P.Štelmachovič, *"On the local unitary equivalence of states of multipartite systems"*, *Forstschritte der Physik* **49**, 1123-1131 (2001), LANL preprint archive `quant-ph/0107016`

[50] A. Yu. Vlasov, *"Aleph-QP:Universal Hybrid quantum processors"*, LANL preprint archive `quant-ph/0205074`

[51] V.Scarani, M.Ziman, P.Štelmachovič, N.Gisin, V.Bužek, *"Thermalizing quantum machines:Dissipation and Entanglement"*, *Phys. Rev. Lett.* **88**,097905 (2002), LANL preprint archive `quant-ph/0110088`

[52] M.Ziman, P.Štelmachovič, V.Bužek, M.Hillery, V.Scarani, N.Gisin, *"Dilluting the quantum information"*, *Phys.Rev A* **65**, 042105 (2002), LANL preprint archive `quant-ph/0110164`

[53] M.Hillery, M.Ziman, V.Bužek, *"Implementation of quantum maps by programmable quantum processors"*, *Phys.Rev A* **66**, 042302 (2002), LANL preprint archive `quant-ph/0205050`

[54] Mário Ziman and Vladimír Bužek, *"Correlation-assisted communication"*, to appear in *Phys.Rev.A* (2003), LANL preprint archive `quant-ph/0205078`

[55] T.Wei, K.Nemoto, P.Goldbart, P.Kwiat, W.Munro, F.Verstraete, *"Maximal entanglement versus entropy for mixed states"*, *Phys.Rev A* **67**, 022110 (2003),

[56] M.Ziman, P.Štelmachocič,V.Bužek, *"Quantum homogenization:Saturation of CKW inequalities"*, *J.Optics B* (2003)

[57] M.Plesch and V.Buzek, *"Entangled graphs:Bipartite entanglement in multi-qubit systems"*, *Phys. Rev. A* **67** 012322 (2003), LANL preprint archive `quant-ph/0211020`