Jagiellonian University
M. Smoluchowski Institute of Physics

Entropy of quantum channel in the theory of quantum information

Wojciech Roga

PhD Thesis
Atomic Optics Department
Supervisor: prof. dr hab. Karol Życzkowski

Cracow 2011

Praca współfinansowana przez Unię Europejską ze środków Europejskiego Funduszu Społecznego, Budżetu Państwa i Budżetu Województwa Małopolskiego.

This work is partially financed by the European Union from sources of the European Social Fund, the National Budget of Poland and the Budget of the Province of Małopolska.

Abstract

Quantum channels, also called quantum operations, are linear, trace preserving and completely positive transformations in the space of quantum states. Such operations describe discrete time evolution of an open quantum system interacting with an environment. The thesis contains an analysis of properties of quantum channels and different entropies used to quantify the decoherence introduced into the system by a given operation.

Part I of the thesis provides a general introduction to the subject. In Part II, the action of a quantum channel is treated as a process of preparation of a quantum ensemble. The Holevo information associated with this ensemble is shown to be bounded by the entropy exchanged during the preparation process between the initial state and the environment. A relation between the Holevo information and the entropy of an auxiliary matrix consisting of square root fidelities between the elements of the ensemble is proved in some special cases. Weaker bounds on the Holevo information are also established.

The entropy of a channel, also called the map entropy, is defined as the entropy of the state corresponding to the channel by the Jamiołkowski isomorphism. In Part III of the thesis, the additivity of the entropy of a channel is proved. The minimal output entropy, which is difficult to compute, is estimated by an entropy of a channel which is much easier to obtain. A class of quantum channels is specified, for which additivity of channel capacity is conjectured.

The last part of the thesis contains characterization of Davies channels, which correspond to an interaction of a state with a thermal reservoir in the week coupling limit, under the condition of quantum detailed balance and independence of rotational and dissipative evolutions. The Davies channels are characterized for one–qubit and one–qutrit systems.

List of papers

1. Wojciech Roga, Mark Fannes, Karol Życzkowski,

   Composition of quantum states and dynamical subadditivity,

   Journal of Physics A – Mathematical and Theoretical, **41** 035305 (15 pp) (2008).

2. Wojciech Roga, Mark Fannes, Karol Życzkowski,

   Universal bounds for the Holevo quantity, coherent information and the Jensen-Shannon divergence,

   Physical Review Letters, **105** 040505 (2010).

3. Wojciech Roga, Mark Fannes, Karol Życzkowski,

   Davies maps for qubit and qutrits,

   Reports on Mathematical Physics, **66** 311–329 (2010).

4. Wojciech Roga, Mark Fannes, Karol Życzkowski,

   Entropic characterization of quantum operations,

   International Journal of Quantum Information, **9** 1031–1045 (2011).

5. Wojciech Roga, Marek Smaczyński, Karol Życzkowski,

   Composition of Quantum Operations and Products of Random Matrices,

   Acta Physica Polonica B, **42** 1123 (18 pp) (2011).

6. Mark Fannes, Fernando de Melo, Wojciech Roga, Karol Życzkowski,

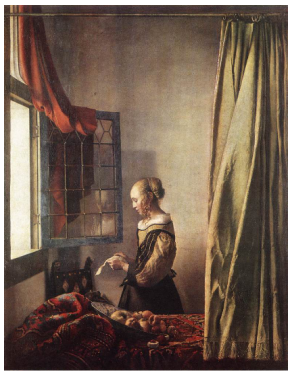
   Matrices of fidelities for ensembles of quantum states and the Holevo quantity,

   arXiv/quant-ph:1104.2271 (24 pp) (2011).

Acknowledgements

I would sincerely like to thank my supervisor Professor Karol Życzkowski for motivation and support in all the time of research and writing of this thesis. I would like to express my gratitude to Professor Mark Fannes for working together on diverse exciting projects. Special thanks to my fellow-worker and friend Fernando de Melo. It is also pleasure to thank Professor Ryszard Horodecki, Professor Paweł Horodecki, Professor Michał Horodecki and Professor Robert Alicki for many opportunities to visit National Quantum Information Centre of Gdańsk and helpful discussions. I would like to show my special gratitude to Professor Tomasz Dohnalik and Professor Jakub Zakrzewski from the Atomic Optics Department for the support and trust in me. I would like to thank my colleagues Piotr Gawron, Zbigniew Puchała, Jarosław Miszczak, Wojciech Bruzda, Łukasz Skowronek and Marek Smaczyński for fruitful collaboration.

Pracę dedykuję mojej żonie Karolinie.
This thesis is dedicated to my wife Karolina.



"Piękno przyrody jest podejrzane".
"The beauty of nature is suspicious".
Cz. Miłosz, *Uczeni* (2002)

Contents

Part I

Introduction

Chapter 1

Preliminary information

1.1 Preface

It is not easy to give a satisfactory definition of information in sense in which this word is used in everyday life. For instance one could ask, how much information is contained in an allegorical baroque painting of Vermeer. There exist, of course, many interpretations and therefore, many kinds of information concerning this picture. However, nowadays we are willing to distinguish some sort of information necessary to communicate a message independently on the interpretation. Due to our experience with computers we are used to problems how to encode the information into a string of digital symbols, transmit it and decode it in order to obtain the original message in another place. Imagine that we need to send the information contained in the Vermeer's picture. We have to encode it into digital data, transmit it to the other place and recover the picture on the screen of the receiver's computer. In a sense we send almost all the information without knowing what interpretations it may carry.

The problem rises what is the minimal amount of information measured in binary digits that enable the receiver to reliably recover the original message. In considered example we can divide the image of the Vermeer's picture into small pieces, decode colours of each piece into digital strings and transmit the description of colours one after another. However, we can also save some amount of digits when we menage to describe shapes of regions of the same colours in the picture and send only information about colours, shapes and patterns. How to do that in the most efficient way? This is a major problem for experts working on the information theory and computer graphics. Some rules of the optimal coding were used intuitively during construction of the Morse alphabet. The letters which occur in the English language more frequently are encoded by a smaller amount of symbols.

In communication and computer sciences the problem of data compression is a subject of a great importance. To what extend the data can be compressed to still remain useful? Claude Shannon worked on the problem of transmission

of messages through telecommunication channels. In 1958 he published his famous paper [1] opening the new branch of knowledge known as the theory of information. In this theory a message is composed of letters occurring with specified frequencies related to probabilities. Every letter of a message can be encoded as a string of digital units. Every digital unit can appear in one of r possible configurations. Shannon found what is the minimal average amount of digital units per symbol which encodes a given message. This smallest average number of digital units is related to the information contained in the message and is characterized by a function of the probability distribution $P = \{p_1, ..., p_n\}$ of letters, now called the *Shannon entropy*,

$$H(P) = -\sum_{i=1}^{n} p_i \log_r p_i, \tag{1.1}$$

where $0 \log_r 0 \equiv 0$, n is a number of letters and the base of the logarithm r characterizing the amount of configurations of a chosen digital unit can be chosen arbitrary. If the base is equal to 2, the unit of entropy is called *binary unit* or *bit*.

The idea of efficient coding concerns in replacing more frequent letters by means of a smaller amount of bits. Shannon treated the message as a sequence of letters generated independently according to the probability distribution P specified for a given language. The original reasoning of Shannon proceeds as follows. There are so many possible messages as the amount of typical sequences of letters with a given probability distribution in the string of length $k \to \infty$. Atypical sequences such as strings of letters a repeated k times are unlikely and are not taken into account. The amount of possible messages is given by the amount of typical sequences, which is of order of $2^{kH(P)}$ if the base of the logarithm is equal to 2. This number is justified by methods of combinatorics. Hence, every typical message of length k can be represented by a string of bits of size $kH(P)$. Therefore, the entropy $H(P)$ can be interpreted as the smallest average amount of bits per letter needed to reliably encode each typical message.

The information theory treats, as well, the information as a measure of uncertainty about the outcome of a random experiment. Looking for a function which is suitable as a measure of the uncertainty about the concrete result of experiment, provided the probabilities of all experimental outcomes are given, Shannon formulated a few postulates for the information measure [1]:

- It is a continues function of the probability distribution.
- If all events are equally likely the function of uncertainty is an increasing function of their number.
- If one of the events is split into two, the new function of uncertainty is equal to the sum of the original uncertainty and the uncertainty of the new division weighted by the probability of the divided event.

The only function which satisfies these postulates is the Shannon entropy $H(P)$. Therefore, the uncertainty or lack of information on the outcome of an experiment is the second interpretation of the entropy $H(P)$.

Taking a weaker set of axioms allows one to generalize the definition of the measure of uncertainty and to find other functions of probability vector P, which in special case converge to the Shannon entropy (1.1). For instance, Rényi introduced one parameter family of generalized entropy functions. Since, the information of an experiment consisting of two independent experiments should be given by the sum of the information gained in both experiments, the measure of information should be additive. The Shannon entropy of the joint probability distribution of two independent variables is additive. Rényi noticed [2] that the additivity of information is not equivalent to the third postulate of Shannon. However, if one replaces the third postulate by additivity of information of independent events, yet another axiom should be postulated to obtain back the Shannon's formula (1.1). This additional postulate specifies the way of calculating the mean values. If one considers the linear mean, the Shannon entropy is singled out by this set of postulates. However, other definition of the mean value also can be taken. In consequence, the new set of postulates implies an one parameter family of generalized entropy functions known as the *Rényi entropy* of order q:

$$H_q(P) = \frac{1}{1-q} \log \sum_{i=1}^{n} p_i^q. \tag{1.2}$$

Here, q denotes the free parameter depending on the definition of the average. Another generalization of entropy function was analysed by Tsallis [3, 4]. The Tsallis entropy of order q is defined as follows,

$$T_q = \frac{1}{q-1}(1 - \sum_{i}^{n} p_i^q). \tag{1.3}$$

Hence the information theory concerns entropies, however, it also investigates communication sources and communication channels which can introduce some errors to messages. Information theory defines such quantities as the *relative entropy* and the *mutual information* [1]. Using these concepts the *channel capacity* is defined. It is the maximal rate of information which can be reliably decoded after passing through the channel. The channel capacity is measured in bits per a unit of time.

The theory of quantum information, which considers quantum systems as carriers of information, should enable one to generalize the notions of classical information theory such as the channel capacity. To formulate a quantum counterpart of the Shannon concepts such as the relative entropy or channel capacity the theory of open quantum systems, quantum statistical processes, statistical operators, density matrices, partial traces and generalized measurements should be applied. In the early stage of the stochastic theory of open quantum systems, it was developed by Davies [5], and Kossakowski [6]. Moreover, other important results on accessible information transmitted through a noisy quantum channel were obtained by Holevo [7].

There are many advantages of using quantum resources to transform and transmit the information [8]. In particular, there exist a famous protocols of

superdense coding [9] of information into a quantum carrier. Furthermore, some computational problems can be solved in framework of the quantum information processing faster than classically [10–12]. Quantum world gives also new communication protocols like quantum teleportation [9, 13] which is possible due to quantum entanglement [14, 15]. In quantum case, entangled states can increase the joint capacity of two channels with respect to the sum of the two capacities [16–18]. Also a new branch of cryptography was developed due to the quantum theory [19]. Although, these new possibilities are promising, manipulation of quantum resources is difficult in practice. In particular, the quantum states carrying the information are very sensible to noise, which can completely destroy quantum information. Moreover, probabilistic nature of quantum theory does not allow us to extract uniquely the information from quantum sources. Many restrictions and laws of quantum information theory are formulated in terms of inequalities of quantum entropies. The most significant quantum entropy is the one defined by von Neumann [20, 21], which is a counterpart of the Shannon entropy. However, the other quantum entropies such like the Rényi entropy [2] or Tsallis entropy are also considered [3, 4, 22].

The issue of transmitting a classical information through a noisy channel is an important issue in the information theory [1, 23, 24]. Among problems concerning the information channels one can specify the following questions: How to encode the information in order to transmit it reliably through the channel in the most efficient way [1, 25]? What is the maximal rate of the information transmission? What is the capacity of a given communication channel [26–29]? Which states are the most resistant to errors occurring in the a channel [30, 31]? What are the efficient strategies of the error correction [32]?

Similar questions can also be formulated in the framework of quantum information theory. The quantum channels, also called *quantum operations*, are transformations in the set of states [33–36]. They describe evolution of an open quantum system interacting with an environment in discrete time.

The set of all quantum channels is still not completely understood. Merely the set of one–qubit channels is satisfactory explored [37, 38]. However, even in this simplest case some interesting problems are open. For instance, it is not clear, whether the capacity of one–qubit channels is additive [18]. Another approach to quantum channels suggests to analyse only certain special classes of them, motivated by some physical models [39–43].

Quantum channels are also formally related to measurement processes in quantum theory [35, 45]. As a measurement process changes the quantum state and in general cannot perfectly distinguish measured states, there is a fundamental restriction on the information which can be obtained from the message encoded into quantum states [7]. These restrictions are also formulated in terms of entropies.

The different aspects of quantum channels mentioned above suggest that entropies which characterize the channels play an important role in the information theory. This thesis is devoted to investigation of quantum channels and some entropies used to characterize them: the minimal output entropy [18, 39], the map entropy [46–48] and the exchange entropy [29].

1.2 Structure of the thesis

The thesis is partially based on results already published in articles [46, 49–53], which are enclosed at the end of the thesis. In a few cases some issues from these papers are discussed here in a more detailed manner. The thesis contains also some new, unpublished results and technical considerations not included in the published articles.

The structure of the thesis is the following. The thesis is divided into three parts. The first part is mostly introductory and contains a short review of the literature. This part provides basic information useful in the other parts of the thesis and fixes notation used in the entire work. In part I only the result from Section 1.6.1 concerning the Kraus representation of a complementary channels and Section 1.9 on the Kraus operators constructed for an ensemble of states are obtained by the author.

Part II contains results based on papers [46, 49, 52], not known before in the literature. However, some results not published previously are also analysed there.

Chapter 2 contains the most important result of the thesis – the inequality between the Holevo information related to an ensemble of quantum states and the entropy of the state of environment taking part in preparation of the ensemble. As the entropy of the environment can be treated equivalently as the entropy of an output of the complementary channel, or the entropy of a correlation matrix, or the entropy of a Gram matrix of purifications of mixed states, or as the entropy exchange, this relation might be considered as a new and universal result in the theory of quantum information. Consequences of this inequality have not been analysed so far. Chapter 2 contains also the discussion of the particular cases for which the inequality is saturated. This result has not been published before. Section 2.1 describes proofs of known entropic inequalities which are related to the bound on the Holevo quantity. Some new and unpublished consequences of these inequalities are presented in Section 2.1.1. Original, new results are also contained in Sections 2.2 and 2.3.

Part II contains, moreover, the conjecture on the inequality between the Holevo information of a quantum ensemble and the entropy of the matrix of square root of fidelities. Several weaker inequalities are analysed here in a greater detail than it was done in [52]. Section 3.2 presents a confirmation of the conjecture for a special class of ensembles of quantum states.

Part III of the thesis is based on the results presented in [50,51]. Article [51] described partially in Chapter 4 considers the relation between minimal output entropy and the map entropy. Section 4.2 contains a proof of additivity of the map entropy with respect to the tensor product of two maps, already published in our work [51]. These results allow us to specify a class of quantum channels for which additivity of the minimal output entropy is conjectured.

The Davies maps acting on one–qubit and one–qutrit quantum systems are analysed in Chapter 5. Conditions for the matrix entries of a quantum operation representing a Davies map are given along the lines formulated in our work [50]. Multiplicativity of the maximal output norm of one–qubit Davies maps,

entirely based on the analogical proof for bistochastic maps [54], is presented in Section 5.6. However, this result cannot be treated as a new one, since multiplicativity of the maximal output two norm was proved earlier for all one–qubit quantum channels [18]. Section 5.7 contains graphical representations of stochastic matrices of order three which correspond to quantum Davies maps, which has not been published yet.

1.3 A short introduction to quantum mechanics

The formalism of quantum mechanics can be derived from a few postulates (axioms) which are justified by experiments. The set of axioms defining the quantum theory differs depending on the author [55]. However, some features occur common in every formulation, either as axioms or as their consequences. One of such key features is the *superposition principle*. It is justified by several experimental data as interference pattern in double slit experiment with electrons or interference of a single photon in the Mach–Zender interferometer [56]. The superposition principle states that the state of a quantum system, which is denoted in Dirac notation by $|\psi\rangle$, can be represented by a coherent combination of several states $|\psi_i\rangle$ with complex coefficients a_i,

$$|\psi\rangle = \sum_i a_i |\psi_i\rangle. \tag{1.4}$$

The quantum state $|\psi\rangle$ of an N level system is represented by a vector from the complex Hilbert space $\mathcal{H}_N$. The inner product $\langle\psi_i|\psi\rangle$ defines the coefficients a_i in (1.4). The square norm of a_i is interpreted as the probability that the system described by $|\psi\rangle$ is in the state $|\psi_i\rangle$. To provide a proper probabilistic interpretation a vector used in quantum mechanics is normalized by the condition $\langle\psi|\psi\rangle = ||\psi||^2 = \sum_i |a_i|^2 = 1$.

Quantum mechanics is a probabilistic theory. One single measurement does not provide much information on the prepared system. However, several measurements on identically prepared quantum systems allow one to characterize the quantum state.

A physical quantity is represented by a linear operator called an *observable*. An observable A is a Hermitian operator, $A = A^\dagger$, which can be constructed by a set of real numbers λ_i (allowed values of the physical quantity) and a set of states $|\varphi_i\rangle$ determined by the measurement, $A = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$. The physical value corresponds to the average of the observable in the state $|\psi\rangle$,

$$\langle A\rangle_\psi = \sum_i \lambda_i |\langle\psi|\varphi_i\rangle|^2 = \langle\psi|A|\psi\rangle. \tag{1.5}$$

One can consider the situation in which a state $|\psi\rangle$ is not known exactly. Only a statistical mixture of several quantum states $|\phi_i\rangle$ which occur with probabilities p_i is given. In this case the average value of an observable has the form

$$\langle A\rangle_{\{p_i,\phi_i\}} = \sum_i p_i \langle\phi_i|A|\phi_i\rangle, \tag{1.6}$$

which can be written in terms of an operator on $\mathcal{H}_N$ called a *density matrix* $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ as

$$\langle A \rangle_{\{p_i,\phi_i\}} = \mathrm{Tr}\, \rho A. \tag{1.7}$$

A density matrix describes a so called *mixed state*. In a specific basis the density matrices characterizing an N level quantum system are represented by $N \times N$ matrices ρ which are Hermitian, have trace equal to unity and are positive. Let us denote the set of all such matrices by $\mathcal{M}_N$,

$$\mathcal{M}_N = \{\rho : \dim\rho = N, \rho = \rho^\dagger, \rho \geq 0, \mathrm{Tr}\, \rho = 1\}. \tag{1.8}$$

This set is convex. Extremal points of this set are formed by projectors of the form $|\psi\rangle\langle\psi|$ called *pure states*, which correspond to vectors $|\psi\rangle$ of the Hilbert space.

The state of composed quantum system which consists of one N_1–level system and one N_2–level system is represented by a vector of size $N_1 N_2$ from the Hilbert space which has a tensor product structure, $\mathcal{H}_{N_1 N_2} = \mathcal{H}_{N_1} \otimes \mathcal{H}_{N_2}$. Such a space contains also states which cannot be written as tensor products of vectors from separate spaces,

$$|\psi_{12}\rangle \neq |\psi_1\rangle \otimes |\psi_2\rangle. \tag{1.9}$$

and are called *entangled states*. States with a tensor product structure are called *product states*. If the state of only one subsystem is considered one has to take an average over the second subsystem. Such an operation is realized by taking the partial trace over the second subsystem and leads to a reduced density matrix,

$$\rho_1 = \mathrm{Tr}_2 \rho_{12}. \tag{1.10}$$

A density matrix describes therefore the state of an *open quantum system*.

The evolution of a normalized vector in the Hilbert space is determined by a unitary operator $|\psi'\rangle = U|\psi\rangle$. The transformation U is related to Hamiltonian evolution due to the Schrödinger equation,

$$i\hbar \frac{d}{dt}|\psi\rangle = H|\psi\rangle, \tag{1.11}$$

where H denotes the Hamiltonian operator of the system, while t represents time and $2\pi\hbar$ is the Planck constant. A discrete time evolution of an open quantum system characterized by a density operator ρ is described by a *quantum operation* which will be considered in Chapter 1.6.

According to a general approach to quantum measurement [35, 57], it can be defined by a set of k operators $\{E^i\}_{i=1}^k$ forming a *positive operator valued measure* (POVM). The index i is related to a possible measurement result, for instance the value of the measured quantity. The operators E^i are positive and satisfy the identity resolution,

$$\sum_{i=1}^{k} E^i = \mathbb{1}. \tag{1.12}$$

The quantum state is changing during the measurement process. After the measurement process that gives the outcome i as a result, the quantum state ρ is transformed into

$$\rho_i' = K^i \rho K^{i\dagger} / \operatorname{Tr}(K^i \rho K^{i\dagger}), \tag{1.13}$$

where $K^{i\dagger}K^i = E^i \geq 0$. The probability p_i of the outcome i is given by $p_i = \operatorname{Tr}(K^i \rho K^{i\dagger})$. Due to relation (1.12), the probabilities of all outcomes sum up to unity.

A quantum state characterizing a 2–level system is called *qubit* and its properties are discussed in more detail in Section 1.7.

1.4 Schmidt decomposition

The theorem known as *Schmidt decomposition* [58] provides a useful representation of a pure state of a bi–partite quantum system.

Theorem 1 (Schmidt)**.** *Any quantum state $|\psi_{12}\rangle$ from the Hilbert space composed of the tensor product of two Hilbert spaces $\mathcal{H}_1 \otimes \mathcal{H}_2$ of dimensions d_1 and d_2, respectively, can be represented as*

$$|\psi_{12}\rangle = \sum_{i=1}^{d} \lambda_i |i_1\rangle \otimes |i_2\rangle, \tag{1.14}$$

where $\{|i_1\rangle\}_{i=1}^{d_1}$ and $\{|i_2\rangle\}_{i=1}^{d_2}$ are orthogonal basis of the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, and $d = \min\{d_1, d_2\}$.

Proof. Choose any orthogonal basis $\{|\phi_1^k\rangle\}_{k=1}^{d_1}$ of $\mathcal{H}_1$ and any orthogonal basis $\{|\phi_2^j\rangle\}_{j=1}^{d_2}$ of $\mathcal{H}_2$. In this product basis, the bi–partite state $|\psi_{12}\rangle$ reads

$$|\psi_{12}\rangle = \sum_{0 \leq k \leq d_1,\ 0 \leq j \leq d_2} a_{kj} |\phi_1^k\rangle \otimes |\phi_2^j\rangle. \tag{1.15}$$

Singular value decomposition of a matrix A of size $d_1 \times d_2$ with entries a_{kj} gives $a_{kj} = \sum_i u_{ki} \lambda_i v_{ij}$. Here u_{ki} and v_{ij} are entries of two unitary matrices, while λ_i are singular values of A. Summation over indexes k and j cause changes of two orthogonal bases into

$$|i_1\rangle = \sum_k u_{ki} |\phi_1^k\rangle, \tag{1.16}$$

$$|i_2\rangle = \sum_j v_{ij} |\phi_2^j\rangle. \tag{1.17}$$

The number o nonzero singular values is not larger than the smaller one of the numbers (d_1, d_2). □

The Schmidt decomposition implies that both partial traces of any bi–partite pure state have the same nonzero part of the spectrum:

$$\mathrm{Tr}_1|\psi_{12}\rangle\langle\psi_{12}| = \sum_{i=1}^{d} \lambda_i^2 |i_2\rangle\langle i_2|, \tag{1.18}$$

$$\mathrm{Tr}_2|\psi_{12}\rangle\langle\psi_{12}| = \sum_{i=1}^{d} \lambda_i^2 |i_1\rangle\langle i_1|. \tag{1.19}$$

The Schmidt coefficients λ_i are invariant under local unitary transformations $U_1 \otimes U_2$ applied to $|\psi_{12}\rangle$. The number of non–zero coefficients λ_i is called the *Schmidt number*. Any pure state which has the Schmidt number greater than 1 is called *entangled state*. A pure state for which all Schmidt coefficients λ_i are equal to $1/\sqrt{d}$ is called a *maximally entangled state*.

Another important consequence of the Schmidt decomposition is that for any mixed state ρ there is a pure state $|\psi\rangle$ of a higher dimensional Hilbert space such that ρ can be obtained by taking the partial trace,

$$\rho = \mathrm{Tr}_1|\psi\rangle\langle\psi|. \tag{1.20}$$

Such a state $|\psi\rangle$ is called a *purification* of ρ. The Schmidt decomposition gives the recipe for the purification procedure. It is enough to take square roots of eigenvalues of ρ in place of λ_i and its eigenvectors in place of $|i_1\rangle$. Any orthogonal basis in $\mathcal{H}_2$ provides a purification of ρ, which can be written as

$$|\psi\rangle = \sum_i (U_1 \otimes \sqrt{\rho})|i_1\rangle \otimes |i_2\rangle, \tag{1.21}$$

where U_1 is an arbitrary unitary transformation and $\sqrt{\rho}|i_2\rangle = \lambda_i|i_2\rangle$.

1.5 Von Neumann entropy and its properties

Many theorems concerning the theory of quantum information can be formulated in terms of the *von Neumann entropy* [59] of a quantum state,

$$S(\rho) = -\,\mathrm{Tr}\,\rho\log\rho, \tag{1.22}$$

which is equivalent to the Shannon entropy (1.1) of the spectrum of ρ. The entropy characterizes the degree of mixing of a quantum state. Assume that ρ is a density matrix of size N. The value of $S(\rho)$ is equal to zero if and only if the state ρ is pure. It gains its maximal value $\log N$ for the maximally mixed state $\rho_* = \frac{1}{N}\mathbb{1}$ only.

Von Neumann entropy has also an important interpretation in quantum information theory, as it plays the role similar to the Shannon entropy in the classical theory of optimal compression of a message [25]. Let the letters i of the message, which occur with probabilities p_i, be encoded into pure quantum states $|\psi_i\rangle$ from the Hilbert space $\mathcal{H}_N$. Sequences of k letters are encoded into a Hilbert

space of dimension N^k. A long message can be divided into sequences of size $k \to \infty$. Among them one can distinguish sequences in typical subspaces and such which occur with negligible probability. A unitary transformation applied to the sequence of quantum systems can transmit almost all the information into a typical subspace. The space of a typical sequence has the smallest dimensionality required to encode the message reliably with negligible probability of an error. This smallest dimensionality per symbol is shown [25] to be equal to the von Neumann entropy of the state $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Therefore, quantum coding consists in taking states from the smaller subspace of dimension $2^{kS(\rho)}$ instead of a space of dimension N^k to encode the same message. If the state ρ represents completely random set of states there is no possibility to compress the message, since $S(\rho) = S(\rho_*) = \log_2 N$, where logarithm is of base 2. The entropy, therefore, describes the capability of compression of the message encoded in a given set of states, or the smallest amount of qubits needed to transmit a given message.

The von Neumann entropy, as the entropy of eigenvalues of a density matrix, describes also the uncertainty of measuring a specific state from the set of the eigenvectors. The most important properties of the von Neumann entropy are [20]:

- The von Neumann entropy is a non negative function of any ρ.
- It is invariant under unitary transformations, $S(\rho) = S(U\rho U^\dagger)$.
- It is a concave function of its argument, $\sum_{i=1}^k p_i S(\rho_i) \leq S(\sum_{i=1}^k p_i \rho_i)$, where $p_i \geq 0$ for any i and $\sum_i^k p_i = 1$.
- It is subadditive
$$S(\rho_{12}) \leq S(\rho_1) + S(\rho_2), \tag{1.23}$$
where ρ_{12} is a bi–partite state of a composite system and the partial traces read $\rho_1 = \mathrm{Tr}_2\, \rho_{12}$ and $\rho_2 = \mathrm{Tr}_1\, \rho_{12}$.
- The von Neumann entropy satisfies the relation of strong subadditivity [60],
$$S(\rho_{123}) + S(\rho_2) \leq S(\rho_{12}) + S(\rho_{23}), \tag{1.24}$$
where the state ρ_{123} is a composite state of three subsystems $(1, 2, 3)$ and the other states are obtained by its partial traces.

1.6 Quantum channels and their representations

One distinguishes two approaches to describe time evolution of an open quantum system. One of them starts from a concrete physical model defined by a given Hamiltonian which determines the Shrödinger equation (1.11) or the master equation, [45]. Solving them one may find the state of the quantum system at any moment at time. An alternative approach to the dynamics of an open quantum system relies on a stroboscopic picture and a discrete time evolution. It

starts from a mathematical construction of a quantum map, $\rho' = \Phi(\rho)$, allowed by the general laws of quantum mechanics. This approach is often used in cases in which the physical model of the time evolution is unknown. This fact justifies the name "black box" model to describe the evolution characterized by a quantum map Φ. Such a model is also considered if one wants to investigate the set of all possible operations independently on whether the physical context is specified. Main features and some representations of the map Φ, which describes a "black box" model of non–unitary quantum evolution, are given below.

The quantum map Φ describes the dynamics of a quantum system ρ which interacts with an environment. It is given by a nonunitary quantum map $\Phi : \rho \to \rho'$. Any such map is completely positive, and trace preserving [33–36]. "Complete positivity" means that an extended map $\Phi \otimes \mathbb{1}_M$, which is a trivial extension of Φ on the space of any dimension M, transforms the set of positive operators into itself. A completely positive and trace preserving quantum map is called *quantum operation* or *quantum channel*.

Due to the theorem of Jamiołkowski [34] and Choi [33] the complete positivity of a map is equivalent to positivity of a state corresponding to the map by the *Jamiołkowski isomorphism*. This isomorphism determines the correspondence between a quantum operation Φ acting on N dimensional matrices and density matrix D_Φ/N of dimension N^2 which is called Choi matrix or the Jamiołkowski state

$$\frac{1}{N} D_\Phi = [\mathrm{id}_N \otimes \Phi]\big(|\phi^+ \rangle \langle \phi^+|\big), \tag{1.25}$$

where $|\phi^+ \rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |i \rangle \otimes |i \rangle$ is a maximally entangled state. The dynamical matrix D_Φ corresponding to a trace preserving operation satisfies the partial trace condition

$$\mathrm{Tr}_2 D_\Phi = \mathbb{1}. \tag{1.26}$$

The quantum operation Φ can be represented as *superoperator matrix*. It is a matrix which acts on the vector of length N^2, which contains the entries ρ_{ij} of the density matrix ordered lexicographically. Thus the superoperator Φ is represented by a square matrix of size N^2. The superoperator in some orthogonal product basis $\{|i\rangle \otimes |j\rangle\}$ is represented by a matrix indexed by four indexes,

$$\Phi_{\substack{i\,j \\ k\,l}} = \langle i| \otimes \langle j|\Phi|k\rangle \otimes |l\rangle. \tag{1.27}$$

The matrix representation of the dynamical matrix is related to the superoperator matrix by the reshuffling formula [15] as follows

$$\langle i| \otimes \langle j|D_\Phi|k\rangle \otimes |l\rangle = \langle i| \otimes \langle k|\Phi|j\rangle \otimes |l\rangle. \tag{1.28}$$

To describe a quantum operation, one may use the Stinespring's dilation theorem [61]. Consider a quantum system, described by the state ρ on $\mathcal{H}_N$, interacting with its environment characterized by a state on $\mathcal{H}_M$. The joint evolution of the two states is described by a unitary operation U. Usually it is assumed that the joint state of the system and the environment is initially not

entangled. Moreover, due to the possibility to purification the environment, its initial state is given by a pure one. The evolving joint state is therefore:

$$\omega = U\Big(|1\rangle \langle 1| \otimes \rho\Big)U^\dagger, \tag{1.29}$$

where $|1\rangle \in \mathcal{H}_M$ and U is a unitary matrix of size NM. The state of the system after the operation is obtained by tracing out the environment,

$$\rho' = \Phi(\rho) = \mathrm{Tr}_M\Big[U\big(|1\rangle \langle 1| \otimes \rho\big)U^\dagger\Big] = \sum_{i=1}^{M} K^i \rho K^{i\dagger}, \tag{1.30}$$

where the Kraus operators read, $K^i = \langle i| U |1\rangle$. In matrix representation the Kraus operators are formed by successive blocks of the first block–column of the unitary evolution matrix U. Here the state ω can be equivalently given as

$$\omega = \sum_{i,j=1}^{M} K^i \rho K^{j\dagger} \otimes |i\rangle \langle j| . \tag{1.31}$$

A transformation $\rho \to \omega$ is obtained by an isometry $F : \mathcal{H}_N \to \mathcal{H}_{NM}$, where

$$F |\phi\rangle = \sum_i (K^i |\phi\rangle) \otimes |i\rangle . \tag{1.32}$$

Due to the Kraus theorem [35] any completely positive map Φ can be written in the Kraus form,

$$\rho' = \Phi(\rho) = \sum_{i=1}^{M} K^i \rho K^{i\dagger}. \tag{1.33}$$

The opposite relation is also true, any map of the Kraus form (1.33) is completely positive.

1.6.1 Representation of a complementary channel

Consider a quantum channel Φ described by the Kraus operators K^i,

$$\Phi(\rho) = \mathrm{Tr}_M \omega = \sum_{i=1}^{M} K^i \rho K^{i\dagger}, \tag{1.34}$$

where notation from Section 1.6 is used. The channel $\tilde{\Phi}$ *complementary* to Φ is defined by

$$\tilde{\Phi}(\rho) = \mathrm{Tr}_N \omega = \sum_{i=1}^{N} \tilde{K}^i \rho \tilde{K}^{i\dagger} \tag{1.35}$$

and it describes the state of the M–dimensional environment after the interaction with the principal system ρ. One can derive the relation between operators

$\{\tilde{K}^j\}_{j=1}^N$ and $\{K^i\}_{i=1}^M$ from the last equation by substituting ω as in (1.31). This relation can be rewritten as

$$\sum_{i,j=1}^{M} (\mathrm{Tr}\, K^i \rho K^{j\dagger}) \left|i\right\rangle \left\langle j\right| = \sum_{i=1}^{N} \tilde{K}^i \rho \tilde{K}^{i\dagger}. \tag{1.36}$$

Comparison of the matrix elements of both sides gives

$$\sum_{\alpha=1}^{N} \tilde{K}^{\alpha}_{im} \rho_{mn} \tilde{K}^{\alpha\dagger}_{nj} = \sum_{\alpha=1}^{N} K^i_{\alpha m} \rho_{mn} K^{j\dagger}_{n\alpha}, \tag{1.37}$$

where matrix elements are indicated by lower indexes and the Einstein summation convention is applied. Hence, for any quantum channel Φ given by a set of Kraus operators K^i, one can define the Kraus operators $\tilde{K}^{\alpha}$ representing the complementary channel $\tilde{\Phi}$ as

$$\tilde{K}^{\alpha}_{ij} = K^i_{\alpha j}, \qquad i = 1, ..., M, \qquad j, \alpha = 1, ..., N. \tag{1.38}$$

1.7 One–qubit channels

One–qubit channels acting on density matrices of size 2 have many special features which cause that the set of these channels is well understood [37, 38, 54]. However, many properties of one–qubit maps are not shared with the quantum maps acting on higher dimensional systems. Since one–qubit quantum channels are often considered in this thesis, the following section presents a brief review of their basic properties.

A quantum two level state is called *quantum bit* or *qubit*. It is represented by a 2×2 density matrix. Any Hermitian matrix of size two can be represented in the basis of identity matrix and the three Pauli matrices $\vec{\sigma} = \{\sigma_1, \sigma_2, \sigma_3\}$,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.39}$$

One qubit state ρ decomposed in the mentioned basis is given by the formula

$$\rho = \frac{1}{2}(\mathrm{id} + \vec{r} \cdot \vec{\sigma}), \quad \vec{r} \in \mathrm{R}^3. \tag{1.40}$$

Positivity condition, $\rho \geq 0$, implies that $|\vec{r}| \leq 1$. The vector $\vec{r}$ is called the *Bloch vector*. All possible Bloch vectors representing quantum states form the *Bloch ball*. Pure one–qubit states form a sphere of radius $|\vec{r}| = 1$.

Any linear one–qubit quantum operation Φ transforms the Bloch ball into the ball or into an ellipsoid inside the ball. The channel Φ transforms the Bloch vector $\vec{r}$ representing the state ρ into $\vec{r}\,'$ which corresponds to ρ'. This transformation is described by

$$\vec{r}\,' = W\vec{r} + \vec{\kappa}. \tag{1.41}$$

Here the matrix W is a square real matrix of size 3. A procedure analogous to the singular value decomposition of the matrix W gives $W = O_1 D O_2$, where O_i represents an orthogonal rotation and D is diagonal. Up to two orthogonal rotations, one before the transformation Φ and one after it, the one–qubit map Φ can be represented by the following matrix

$$\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \kappa_1 & \eta_1 & 0 & 0 \\ \kappa_2 & 0 & \eta_2 & 0 \\ \kappa_3 & 0 & 0 & \eta_3 \end{pmatrix}. \tag{1.42}$$

The absolute values of the parameters η_i are interpreted as the lengths of the axes of the ellipsoid which is the image of the Bloch ball transformed by the map. The parameters κ_i form the vector $\vec{\kappa}$ of translation of the center of the ellipsoid with respect to the center of the Bloch ball.

Due to complete positivity of the map Φ and the trace preserving property, the vectors $\vec{\eta}$ and $\vec{\kappa}$ are subjected to several constraints. They can be derived from the positivity condition of a dynamical matrix given by [15,37]:

$$D_\Phi = \frac{1}{2} \begin{pmatrix} 1+\eta_3+\kappa_3 & 0 & \kappa_1 + i*\kappa_2 & \eta_1+\eta_2 \\ 0 & 1-\eta_3+\kappa_3 & \eta_1-\eta_2 & \kappa_1+i*\kappa_2 \\ \kappa_1 - i*\kappa_2 & \eta_1-\eta_2 & 1-\eta_3-\kappa_3 & 0 \\ \eta_1+\eta_2 & t_1 - i*\kappa_2 & 0 & 1+\eta_3-\kappa_3 \end{pmatrix}. \tag{1.43}$$

The channels which preserve the maximally mixed state are called *bistochastic* channels. The structure of one–qubit bistochastic channels is discussed in more detail in Section 4.1.

1.8 Correlation matrices

A general measurement process is described in quantum mechanics by operators forming a *positive operator valued measure* (POVM). Products of matrices $K^{i\dagger}K^i$ representing the POVM are positive and determine the identity resolution, $\mathbb{1} = \sum_{i=1}^k K^{i\dagger}K^i$. During the measurement of a quantum state ρ the output $\rho_i = \frac{K^i \rho K^{i\dagger}}{\operatorname{Tr} K^i \rho K^{i\dagger}}$ occurs with probabilities $p_i = \operatorname{Tr} K^i \rho K^{i\dagger}$. The identity resolution guarantees that $\sum_{i=1}^k p_1 = 1$.

The outcomes of a quantum measurement are not perfectly distinguishable, unless different POVM operators project on orthogonal subspaces, $K^{i\dagger}K^i K^{j\dagger}K^j = \delta_{ij} K^{i\dagger}K^i$. Probability distribution of the outcome states does not contain any information on indistinguishability of outcomes. Therefore, a better characterization of the measurement process is given by the following *correlation matrix* σ with entries

$$\sigma_{ij} = \operatorname{Tr} K^i \rho K^{j\dagger}, \qquad i,j = 1,...,k. \tag{1.44}$$

Its diagonal contains the probabilities of measurement outputs, while the off–diagonal entries are related to probabilities that the state i has been determined by the measurement as the state j. The correlation matrix depends on both, the measured state and the measurement process.

The operators K^i, satisfying $\sum_{i=1}^{k} K^{i\dagger}K^i = \mathbb{1}$, can also be treated as Kraus operators (1.30) characterizing the quantum channel, $\Phi(\rho) = \sum_{i=1}^{k} K^i \rho K^{i\dagger}$. In such an interpretation of operators K^i, the correlation matrix (1.44) is equivalent to the state of environment given by the output of the complementary channel $\tilde{\Phi}(\rho)$ specified in Eq. (1.36).

The entropy of the state σ produced by a complementary channel $\tilde{\Phi}$ is called the *exchange entropy*, since, if the initial states of the system and the environment are pure, then $S(\sigma)$ is equal to the entropy which is gained by both the state and the environment [29]. If the initial state is maximally mixed, $\rho = \rho_* = \frac{1}{N}\mathbb{1}$, where N is the dimensionality of ρ, the entropy of the output of the complementary channel is equal to the *map entropy* $S^{\max}(\Phi)$ [46] (see also discussion in Section 2.1.1),

$$S^{\max}(\Phi) = -\frac{1}{N} D_\Phi \log\left(\frac{1}{N} D_\Phi\right), \tag{1.45}$$

where the dynamical matrix D_Φ is given by Eq. (1.25). This entropy is equal to zero if Φ represents any unitary transformation. It attains the largest value $\log 2N$ for completely depolarizing channel which transform any state into the maximally mixed state. Therefore the map entropy can characterize the decoherence caused by the channel.

Due to the polar decomposition of an arbitrary non normal operator $X = HU$, we can write $K^i\rho^{1/2} = h_i U_i$, where h_i is a Hermitian matrix and U_i is unitary. One can observe that $h_i^2 = K^i \rho K^{i\dagger} = p_i \rho_i$. Therefore the entries of the correlation matrix (1.44) can be written as:

$$\sigma_{ij} = \operatorname{Tr} K^i \rho K^{j\dagger} = p_i^{\frac{1}{2}} p_j^{\frac{1}{2}} \operatorname{Tr} \rho_i^{\frac{1}{2}} U_i U_j^\dagger \rho_j^{\frac{1}{2}}. \tag{1.46}$$

As noticed above, the correlation matrix characterizing the quantum measurement can be equivalently treated as the state of an environment after evolution given by a quantum channel. The following section indicates a third possible interpretation of the correlation matrix σ. It can be formally treated as a Gram matrix of purifications of mixed states ρ_i.

Purification of a given state $\rho_i \in \mathcal{M}_N$ is given by a pure state $|\Psi_i\rangle$ (see Eq. (1.21)),

$$\operatorname{Tr}_1 |\Psi_i\rangle \langle\Psi_i| = \rho_i. \tag{1.47}$$

The purification $|\Psi_i\rangle$ of given state ρ_i can be written explicitly,

$$|\Psi_i\rangle = \sum_{r=1}^{N} \left(U_i \otimes \sqrt{\rho_i}\right) |r\rangle \otimes \left|\phi_r^i\right\rangle, \tag{1.48}$$

where $\{\left|\phi_r^i\right\rangle\}_{r=1}^N$ are eigenvectors of ρ_i. Notice that a purification of a given state ρ_i is not unique. The degree of freedom is introduced by the unitary

transformation U_i. Moreover, any purification of given state ρ_i can be given by such a form. Since eigenvectors of ρ_i denoted by $\left|\phi_r^i\right\rangle$ form an orthonormal basis in the Hilbert space, a unitary transformation V_i can transform it into the canonical basis $\{|r\rangle\}_{i=1}^N$. The purification (1.48) can be described as

$$|\Psi_i\rangle = \sum_{r=1}^{N} \left(U_i \otimes \sqrt{\rho_i} V_i\right) |r\rangle \otimes |r\rangle \,. \tag{1.49}$$

The overlap between two purifications of states ρ_i and ρ_j emerging from a POVM measurement is given by

$$|\left\langle \Psi_j \right| \left. \Psi_i \right\rangle|^2 = |\left\langle m \right| (U_j^\dagger U_i \otimes V_j^\dagger \sqrt{\rho_j}\sqrt{\rho_i} V_i) \left| m \right\rangle|^2, \tag{1.50}$$

where $|m\rangle = \sum_r |r\rangle \otimes |r\rangle$. For any two operators A and B the following relation holds, $\langle m| A \otimes B |m\rangle = \operatorname{Tr} A^\dagger B$ [62]. Hence the overlap (1.50) reads

$$|\left\langle \Psi_j \right| \left. \Psi_i \right\rangle|^2 = |\operatorname{Tr} W \sqrt{\rho_j}\sqrt{\rho_i}|^2, \tag{1.51}$$

where the unitary matrix $W = V_i U_i^\dagger U_j V_j^\dagger$. Therefore the matrix elements of σ (1.46) are equal to the scalar product of purifications of respective mixed states ρ_i and ρ_j as follows $\sigma_{ij} = \sqrt{p_i p_j}\langle \Psi_j | \Psi_i \rangle$.

1.8.1 Gram matrices and correlation matrices

In previous chapter it was shown that the correlation matrix can by defined by the set of purifications of states emerging from the quantum measurement. Therefore, the correlation matrix can be identified with the normalized Gram matrix of the purifications.

The Gram matrix is an useful tool in many fields. It can receive a geometrical interpretation, as it consists of the overlaps of normalized vectors. If vectors are real the determinant of their Gram matrix defines the volume of the parallelogram spanned by the vectors [63, 64]. The Gram matrix of the evolving pure state is analyzed in [65]. The spectrum of this matrix can determine whether the evolution is regular or chaotic.

The Gram matrix σ,

$$\sigma_{ij} = \sqrt{p_i p_j}\langle \psi_i | \psi_j \rangle \tag{1.52}$$

has the same eigenvalues as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \tag{1.53}$$

The proof of this fact [66] uses the pure state,

$$|\phi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle \otimes |e_i\rangle, \tag{1.54}$$

where states $|e_i\rangle$ form the set of orthogonal vectors. Since the state (1.54) is pure, its complementary partial traces equal to (1.52) and (1.53) have the same entropy

$$S\left([\sqrt{p_i p_j}\langle\psi_i|\psi_j\rangle]_{ij}\right) = S\left(\sum_i p_i|\psi_i\rangle\langle\psi_i|\right). \tag{1.55}$$

The entropy of the Gram matrix (1.52) can be used in quantum information theory to describe the ability of compression of quantum information [67]. The authors of [67] describe the fact that it is possible to enlarge the information transmitted by means of set of states which are pairwise less orthogonal and thus more indistinguishable. This fact encourages us to consider global properties of quantum ensemble which, sometimes, are not reduced to joint effects of each pair considered separately. In Chapter 3 some efforts will be made to define the quantity characterizing fidelity between three states.

1.9 Kraus operators constructed for an ensemble of states

The previous section concerns the ensembles $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^k$ formed by the outputs of a given quantum channel and a given input state. In the following section it will be shown that for any ensemble $\mathcal{E}$ the suitable Kraus operators K^i can be constructed and the corresponding initial state ρ can be found.

Initial state is constructed from the states of the ensemble by taking

$$\rho = \sum_{i=1}^{k} p_i U_i^\dagger \rho_i U_i, \tag{1.56}$$

where the unitary matrices U_i are arbitrary. The Kraus operators constructed for ensemble $\mathcal{E}$ and unitaries U_i are defined by

$$K^i = \sqrt{p_i \rho_i} U_i \frac{1}{\sqrt{\rho}}. \tag{1.57}$$

Notice that $K^i \rho K^{i\dagger} = p_i \rho_i$ and the Hermitian conjugation, $K^{i\dagger} = \frac{1}{\sqrt{\rho}} U_i^\dagger \sqrt{p_i \rho_i}$. Due to the choice of ρ in (1.56) the identity resolution holds,

$$\sum_{i=1}^{k} K^{i\dagger} K^i = \sum_{i=1}^{k} p_i \frac{1}{\sqrt{\rho}} U_i^\dagger \rho_i U_i \frac{1}{\sqrt{\rho}} = \mathbb{1}. \tag{1.58}$$

In the special case of $k = 2$ states in an ensemble, by choosing

$$U_2 = U_1 \frac{1}{\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}} \sqrt{\rho_1}\sqrt{\rho_2}, \tag{1.59}$$

one obtains σ_{12} equal to square root fidelity between states ρ_1 and ρ_2, as follows $\sqrt{F(\rho_1, \rho_2)} = \mathrm{Tr}\,\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}$.

In consequence of the above considerations one can say that the ensemble emerging from POVM measurement can be arbitrary and for any ensemble $\mathcal{E}$ we can construct the set of operators K^i and the corresponding initial state ρ.

1.10 Quantum fidelity

An important problem in the theory of probability is how to distinguish between two probability distributions. The so called *fidelity* is a quantity used for this purpose. Assume that $P = (p_1, p_2, ..., p_N)$ and $Q = (q_1, q_2, ..., q_N)$ are two probability distributions. The fidelity between $\mathbf{p}$ and $\mathbf{q}$ is defined as,

$$F(P,Q) = \left(\sum_{i=1}^{N} \sqrt{p_i q_i} \right)^2 . \tag{1.60}$$

This function has several properties:

- it is real,
- positive, $F(P,Q) \geq 0$,
- symmetric, $F(P,Q) = F(Q,P)$,
- smaller or equal to unity, $F(P,Q) \leq 1$.
- equal to one if and only if two distributions are the same, $(F(P,Q) = 1) \Leftrightarrow (P = Q)$.

These properties are shared by fidelities defined for quantum states given below.

Quantum counterpart of the fidelity for the pure states $|\phi_1\rangle \in \mathcal{H}_N$ and $|\phi_2\rangle \in \mathcal{H}_N$ is given by the overlap

$$F(|\phi_1\rangle, |\phi_2\rangle) = |\langle \phi_1 \mid \phi_2 \rangle|^2 . \tag{1.61}$$

A probability distribution can be considered as a diagonal density matrix. Generalization of two formulas (1.60) and (1.61) for arbitrary mixed states $\rho_1 \in \mathcal{M}_N$ and $\rho_2 \in \mathcal{M}_N$ is given by

$$F(\rho_1, \rho_2) = \left(\mathrm{Tr} \sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}} \right)^2 . \tag{1.62}$$

To show a relation to previous definitions of fidelity consider two commuting quantum states. They can be given, in the same basis, as $\rho_1 = \sum_i^N r_i |i\rangle \langle i|$, and $\rho_1 = \sum_i^N s_i |i\rangle \langle i|$. Hence the fidelity between them reads

$$\left(\mathrm{Tr} \sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}} \right)^2 = \left(\mathrm{Tr} \sqrt{\sum_{i=1}^{N} r_i s_i |i\rangle \langle i|} \right)^2 = \left(\sum_{i=1}^{N} \sqrt{r_i s_i} \right)^2 . \tag{1.63}$$

This gives a relation between fidelity between mixed quantum states (1.62) and fidelity of probability distributions which are composed by the eigenvalues of the states (1.60). Consider now pure states, $|\Psi_1\rangle, |\Psi_2\rangle \in \mathcal{H}_N \otimes \mathcal{H}_N$ such that the partial trace over the first subspace reads, $\mathrm{Tr}_1 |\Psi_i\rangle \langle\Psi_i| = \rho_i$. There exists a relation between formula (1.62) for fidelity between two mixed states and overlaps of their purifications.

Theorem 2 (Uhlmann [62])**.** *Consider two quantum states ρ_1 and ρ_2 and their purifications $|\Psi_1\rangle$ and $|\Psi_2\rangle$. Then*

$$\left(\mathrm{Tr} \sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} \right)^2 = \max_{|\Psi_1\rangle} |\langle \Psi_1 \mid \Psi_2 \rangle|^2 , \tag{1.64}$$

where the maximization is taken over all purifications $|\Psi_1\rangle$ of the state ρ_1.

Proof. The proof starts from purification formula (1.49),

$$|\Psi_i\rangle = (U_i \otimes \sqrt{\rho_i} V_i) |m\rangle , \tag{1.65}$$

where $|m\rangle$ is an unnormalized vector, $|m\rangle = \sum_{i=1}^N |r\rangle \otimes |r\rangle$. The overlap of two purifications (1.50) is given by

$$|\langle \Psi_j| \Psi_i \rangle|^2 = |\mathrm{Tr}\, W \sqrt{\rho_j}\sqrt{\rho_i}|^2, \tag{1.66}$$

where the unitary matrix $W = V_i U_i^\dagger U_j V_j^\dagger$. The maximization over purifications is equivalent to maximization over the unitary matrix W. An inequality $|\mathrm{Tr}\, A\, B| \leq \|A\|\ \mathrm{Tr}\, |B|$ provides the required lower bound

$$|\mathrm{Tr}\, W \sqrt{\rho_j}\sqrt{\rho_i}|^2 \leq \left(\mathrm{Tr}\, |\sqrt{\rho_j}\sqrt{\rho_i}|\right)^2 . \tag{1.67}$$

The upper bound is attained by the unitary matrix $W^\dagger$ equal to the unitary part of the polar decomposition of $\sqrt{\rho_j}\sqrt{\rho_i}$. This finishes the proof. □

1.10.1 Geometrical interpretation of fidelity

Consider two one–qubit states in the Bloch representation (1.40),

$$\rho_x = \frac{1}{2}(\mathrm{id} + \vec{x} \cdot \vec{\sigma}), \tag{1.68}$$

$$\rho_y = \frac{1}{2}(\mathrm{id} + \vec{y} \cdot \vec{\sigma}), \tag{1.69}$$

where $\vec{\sigma}$ is the vector of Pauli matrices (1.39). Fidelity of the pair of states ρ_x and ρ_y reads

$$F(\rho_x, \rho_y) = \frac{1}{2}(1 + \vec{x} \cdot \vec{y} + \sqrt{1 - \|\vec{x}\|^2}\sqrt{1 - \|\vec{y}\|^2}). \tag{1.70}$$

If the states ρ_x and ρ_y are both pure then $\|\vec{x}\| = \|\vec{y}\| = 1$ and the fidelity can be given by

$$F(\rho_x, \rho_y) = \cos^2 \frac{\alpha}{2}, \tag{1.71}$$

where the angle α is formed by two Bloch vectors which represent the pure states ρ_x and ρ_y at the Bloch sphere. One can use this statement to define the angle between two states as a function of the fidelity. The generalization of such an angle for arbitrary two mixed states is given by

$$A(\rho_1, \rho_2) := \arccos \sqrt{F(\rho_1, \rho_2)}. \tag{1.72}$$

It was proved [68] that such an angle satisfies the axioms of a distance and leads to a metric.

1.11 Mutual information

The goal of quantum information is to efficiently apply quantum resources for information processing. Consider the following situation. A sender transmits the letters of the message from the set $X = \{a_1, a_2, ..., a_k\}$. The letters occur with probabilities p_i, where $i = 1, ..., k$. The message is transmitted by a communication channel, which can be noisy and can change some of the letters. The receiver performs a measurement and obtains outputs Y with a possibly different probability distribution. According to the Shannon information theory [1] the amount of information contained in the message characterized by probability distribution p_i is given by the entropy $H(X) = -\sum_i p_i \log p_i$. Entropy describes the average amount of digits per letter necessary to transmit the message characterized by this probability distribution in an optimal encoding scheme.

The receiver knowing the letters Y has only a part of information contained in the original message X. The information which Y and X have in common is characterized by the *mutual information* $H(X : Y)$ defined by

$$H(X : Y) = H(X) + H(Y) - H(X, Y), \tag{1.73}$$

where $H(X, Y)$ is the Shannon entropy of the joint probability distribution of the pairs of letters, one from X and one from Y.

The errors caused by a channel can be perfectly corrected if the mutual information is equal to the entropy of the initial probability distribution. Otherwise the mutual information is bounded by the entropy of an initial distribution [8],

$$H(X : Y) \leq H(X). \tag{1.74}$$

Following properties of the mutual information hold [8]:

- Mutual information does not change $H(X : Y, Z) = H(X : Y)$ if the system Z is uncorrelated with Y.
- Mutual information does not increase if any process is made on each part, $H(X : Y) \geq H(X' : Y')$, where prime denotes the states after the transformation.

- If part of a system is discarded the mutual information decreases $H(X:Y,Z) \geq H(X:Z)$.

Mutual information can also be defined for quantum composite systems in terms of the von Neumann entropy . The definition is analogous to (1.73):

$$S(\rho_P : \rho_Q) = S(\rho_P) + S(\rho_Q) - S(\rho_{PQ}), \tag{1.75}$$

where states of subsystems are given by partial traces, for example, $\rho_P = \mathrm{Tr}_Q \rho_{PQ}$. Mutual information $S(\rho_P : \rho_Q)$ for quantum states satisfies properties analogous to these listed above for the classical mutual information $H(X,Y)$.

1.12 Holevo quantity

Holevo χ quantity (Holevo information) of the ensemble $\mathcal{E} = \{q_i, \rho_i\}_{i=1}^k$ is defined by the formula

$$\chi(\{q_i, \rho_i\}) \equiv S\left(\sum_{i=1}^k q_i \rho_i\right) - \sum_{i=1}^k q_i S(\rho_i). \tag{1.76}$$

It plays an important role in quantum information theory. As the bound on the mutual information [7], Holevo quantity is related to fundamental restriction on the information achievable from measurement allowed by quantum mechanics. It directly reflexes these features of quantum mechanics which distinguishes this theory from classical physics. In classical information theory the mutual information between the sender and the receiver is bounded only by the Shannon entropy of the probability distribution describing the original message. In the case of an ideal channel between two parts the mutual information is equal to the upper bound. In quantum case, even without any noise present during the transmission process, the mutual information is restricted by the Holevo quantity which is smaller than the entropy associated with the original message, unless the states used to encode the message are orthogonal.

The theorem of Holevo [7] is presented below together with its proof.

Theorem 3 (Holevo)**.** *Let $\{\rho_i\}_{i=1}^k$ be a set of quantum states produced with probabilities p_i from the distribution P. Outcomes of a POVM measurement performed on these states are encoded into symbols with probabilities q_j from probability distribution Q. Whichever measurement is done, the accessible mutual information is bounded from above,*

$$H(P:Q) \leq S\left(\sum_{i=1}^k p_i \rho_i\right) - \sum_{i=1}^k p_i S(\rho_i). \tag{1.77}$$

Proof. Consider a three partite state, where its parts are denoted by the letters P, Q and M

$$\omega_{PQM} = \sum_i p_i \left|i\right\rangle \left\langle i\right| \otimes \rho_i \otimes \left|0\right\rangle \left\langle 0\right|. \tag{1.78}$$

Three parts of the system P, Q and M can be associated with the preparation state, quantum systems, and the measurement apparatus respectively. The state ω_{PQM} describes the quantum system before the measurement, since the state of the apparatus is independent on the quantum states.

Assume that the state ω_{PQM} is subjected to the quantum operation acting on the subsystem QM as follows, $\Phi(\rho \otimes |0\rangle\langle 0|) = \sum_j K^j \rho K^{j\dagger} \otimes |j\rangle\langle j|$. The Kraus operators of this quantum operation form a POVM measurement since $\sum_j K^{j\dagger} K^j = \mathbb{1}$. The state after this measurement is given by

$$\omega_{P'Q'M'} = \sum_{ij} p_i |i\rangle\langle i| \otimes K^j \rho_i K^{j\dagger} \otimes |j\rangle\langle j| . \tag{1.79}$$

Properties of the mutual information listed in section 1.11 imply the key inequality of the proof:

$$S(\omega_P : \omega_Q) \geq S(\omega_{P'} : \omega M'). \tag{1.80}$$

To prove inequality (1.77) it is enough to calculate the quantities occurring in (1.80) for the state (1.78) and (1.79) respectively. Since $\omega_{PQ} = \mathrm{Tr}_M \omega_{PQM} = \sum_i p_i |i\rangle\langle i| \otimes \rho_i$, the left hand side of (1.80) is given by

$$S(\omega_P : \omega_Q) = S(\omega_P) + S(\omega_Q) - S(\omega_{PQ}) = S(\rho') - \sum_{i=1}^{k} p_i S(\rho_i), \tag{1.81}$$

where $\rho' = \sum_i p_i \rho_i$. This is the Holevo quantity which does not depend on the measurement operators K^i. To compute the right hand side of (1.80), $S(\omega_{P'} : \omega_{M'})$, consider a state (1.79). The observation that $p(x,y) = p_x p(y|x) = p_x \,\mathrm{Tr}\, K^{y\dagger} K^y \rho_x$ leads to

$$S(\omega_{P'} : \omega_{M'}) = H(P : Q), \tag{1.82}$$

where $Q = \{q_y\}_y$ and $q_y = \mathrm{Tr}\, K^y \rho' K^{y\dagger}$. This is the mutual information between the probability distributions describing the outcomes of the measurement and the original message. That finishes the proof of the Holevo bound on the mutual information of message encoded into quantum systems. □

Above theorem is one of the most important applications of the Holevo quantity. Quantum information theory uses also the Holevo quantity χ to define channel capacity. There exist several definitions of quantum capacity of a channel depending on whether the entanglement between the input states is allowed or not. In the case that quantum states in a message are not entangled the *Holevo capacity* of channel Φ is defined by

$$C_H(\Phi) = \max_{\mathcal{E}=\{p_i,\rho_i\}_{i=1}^k} \left[S\left(\sum_{i=1}^{k} p_i \Phi(\rho_i) \right) - \sum_{i=1}^{k} p_i S\left(\Phi(\rho_i)\right) \right]. \tag{1.83}$$

The Holevo quantity $\chi(\mathcal{E})$, which can be interpreted as the Holevo capacity of the identity channel, bounds the capacity C_H for any channel [8]:

$$C_H \leq \chi(\mathcal{E}). \tag{1.84}$$

Yet another application of the Holevo quantity concerns the ensembles of quantum states. Formula (1.76) can be given by the average relative entropy

$$\sum_{i=1}^{k} p_i D\left(\rho_i, \sum_{j=1}^{k} p_j \rho_j\right) = S\left(\sum_{i=1}^{k} p_i \rho_i\right) - \sum_{i=1}^{k} p_i S(\rho_i), \tag{1.85}$$

where the relative entropy is defined as $D(\rho_1, \rho_2) \equiv \operatorname{Tr} \rho_1(\log \rho_1 - \log \rho_2)$. It defines an average divergence of every state from the average state. Average (1.85) is known as the quantum Jensen Shannon divergence [69]. Its classical version, for probability measures, is considered in [70]. From mathematical point of view, the Holevo quantity can be treated as a quantity which characterizes the concavity of the entropy function.

The Holevo information will be the main object considered in Part II of this thesis.

Part II

Bounds on the Holevo quantity

Chapter 2

Holevo quantity and the correlation matrix

In the following chapters several inequalities for the Holevo information (Holevo quantity) will be given. It is well-known [8] that the Shannon entropy of the probability vector $P = \{p_1, ..., p_k\}$ is an upper bound for the Holevo quantity of an ensemble $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^k$:

$$\chi(\mathcal{E}) \ \leq \ H(P).$$

Since the Holevo quantity forms a bound on accessible mutual information, the difference between entropy of probability vector $H(P)$ and the Holevo quantity specifies how the chosen set of density matrices differs from the ideal code, which can be decoded perfectly by the receiver. The upper bound on the Holevo quantity can be used for estimating this difference. One of the estimation for the Holevo quantity is presented in the following section.

As discussed in Section 1.8 the correlation matrix σ can be equivalently interpreted in several ways. If the set of the Kraus operators K^i defines a quantum channel, $\Phi(\rho) = \sum_{i=1}^k K^i \rho K^{i\dagger}$, the correlation matrix σ characterizes the output state of the complementary channel, $\sigma = \tilde{\Phi}(\rho)$, or the state of the environment after the quantum operation. As mentioned in Section 1.8.1, σ defines also the Gram matrix of purifications of the states $\{\rho_i\}_{i=1}^k$. The entropy $S(\sigma)$ is related to the exchange entropy or the entropy which the environment gains during a quantum operation provided the initial state of the environment is pure. In the following analysis a quantum channel $\Phi(\rho) = \sum_i K^i \rho K^{i\dagger}$ is treated as a device preparing an ensemble of quantum states $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^k$, where

$$p_i = \operatorname{Tr} K^i \rho K^{i\dagger}, \qquad \text{and} \qquad \rho_i = \frac{K^i \rho K^{i\dagger}}{\operatorname{Tr} K^i \rho K^{i\dagger}}. \tag{2.1}$$

The described situation is illustrated in Fig. 2.1.

Independently of the interpretation of the Kraus operators K^i the following theorem proved in [49] holds.

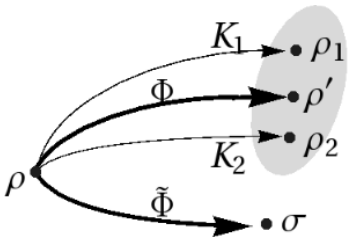


Figure 2.1: A quantum channel Φ represents a device preparing the ensemble of quantum states $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^2$. The average of this ensemble is denoted as $\rho' = \Phi(\rho) = \sum_{i=1}^2 K^i \rho K^{i\dagger}$. The complementary channel $\tilde{\Phi}$ transforms an initial state ρ into the state σ of the environment.

Theorem 4. *Let $\sum_{i=1}^k K^{i\dagger} K^i = \mathbb{1}$ be the identity decomposition and ρ an arbitrary quantum state. Define the probability distribution $p_i = \operatorname{Tr} K^i \rho K^{i\dagger}$ and a set of density matrices $\rho_i = \frac{K^i \rho K^{i\dagger}}{\operatorname{Tr} K^i \rho K^{i\dagger}}$. The Holevo quantity $\chi(\{\rho_i, p_i\})$ is bounded by the entropy of the correlation matrix, $\sigma = \sum_{i,j=1}^k \operatorname{Tr} K^i \rho K^{j\dagger} |i\rangle\langle j|$:*

$$\chi(\{\rho_i, p_i\}) = S\big(\sum_{i=1}^k p_i \rho_i\big) - \sum_{i=1}^k p_i S(\rho_i) \leq S(\sigma) \leq H(P), \tag{2.2}$$

where $H(P)$ is the Shannon entropy of the probability distribution $P = \{p_1, ..., p_k\}$.

Proof. The right hand side of the inequality: $S(\sigma) \leq H(P)$, is a consequence of the majorization theorem, see e.g. [15]. Since the probability vector P forms a diagonal of a correlation matrix, we have $S(\sigma) \leq S(\operatorname{diag}(\sigma)) = H(P)$. The left hand side of the inequality (2.2) is proved due to the strong subadditivity of the von Neumann entropy [60]. The multipartite state ω_{123} is constructed in such a way that entropies of its partial traces are related to specific terms of (2.2).

The multipartite state ω_{123} is constructed by using an isometry $F|\phi\rangle = \sum_{i=1}^k |i\rangle \otimes |i\rangle \otimes K^i |\phi\rangle$. The state $\omega_{123} = F\rho F^\dagger$ is given explicitly by the formula

$$\omega_{123} = F\rho F^\dagger = \sum_{i,j=1}^k |i\rangle\langle j| \otimes |i\rangle\langle j| \otimes K^i \rho K^{j\dagger}. \tag{2.3}$$

States of the subsystems ω_i are given by partial traces over the remaining subsystems, for example, $\omega_1 = \operatorname{Tr}_{23} \omega_{123}$ and so on.

Let us introduce the following notation $A_{ij} = K^i \rho K^{j\dagger}$. In this notation the quantities from the Theorem 4 take the form $p_i = \operatorname{Tr} A_{ii}$ and $\rho_i = A_{ii}/p_i$. Notice that

$$S(\omega_{12}) = S(\sigma), \tag{2.4}$$

$$S(\omega_3) = S\big(\sum_{i=1}^k p_i \rho_i\big). \tag{2.5}$$

Moreover

$$-\sum_{i=1}^{k} p_i S(\rho_i) = \sum_{i=1}^{k} \operatorname{Tr} A_{ii} \log A_{ii} - \sum_{i=1}^{k} \operatorname{Tr}(A_{ii}) \log \operatorname{Tr}(A_{ii})$$
$$= S(\omega_1) - S(\omega_{23}). \tag{2.6}$$

The strong subadditivity relation in the form which is used most frequently

$$S(\omega_{123}) + S(\omega_2) \leq S(\omega_{12}) + S(\omega_{23}) \tag{2.7}$$

does not lead to the desired form (2.2). However, due to the purification procedure and the fact that a partial trace of a pure state has the same entropy as the complementary partial trace, inequality (2.7) can be rewritten in an alternative form [21]:

$$S(\omega_3) + S(\omega_1) \leq S(\omega_{12}) + S(\omega_{23}). \tag{2.8}$$

This inequality applied to the partial traces of the state (2.3) proves Theorem 4. □

For an ensemble of pure states $\rho_i = |\psi_i\rangle\langle\psi_i|$, the left hand side of (2.2) consists of the term $S(\sum_i p_i |\psi_i\rangle\langle\psi_i|)$ only. The correlation matrix σ in the case of pure states is given by the Gram matrix. Due to the simple observation (1.55), the left inequality (2.2) is saturated in case of any ensemble $\mathcal{E}$ consisting of pure states only.

Using a different method an inequality analogous to Theorem 4 has been recently proved in [71] for the case of infinite dimension. It can be also found in [72] in context of quantum cryptography. The authors analyse there the security of a cryptographic key created by using so called 'private qubits'. In such a setup an inequality analogous to (2.2) appears as a bound on the information of the eavesdropper.

2.1 Other inequalities for the Holevo quantity

Methods similar to that used to prove Theorem 4 can be applied to prove other useful bounds.

Proposition 1. *Consider a POVM measurement characterized by operators* $\sum_{i=1}^{k} K^{i\dagger} K^i = \mathbb{1}$ *which define the outcome states,* $\rho_i = \frac{K^i \rho K^{i\dagger}}{\operatorname{Tr} K^i \rho K^{i\dagger}}$ *and their probabilities,* $p_i = \operatorname{Tr} K^i \rho K^{i\dagger}$. *The average entropy of the output states is smaller than entropy of the initial state,*

$$\sum_{i=1}^{k} p_i S(\rho_i) \leq S(\rho). \tag{2.9}$$

Proof. Due to the fact that the transformation F in Eq. (2.3) is an isometry, the three-partite state ω_{123} has the same nonzero spectrum as the initial state ρ. Hence ω_{123} and ρ have the same entropy. Due to equality (2.6) and the Araki–Lieb inequality [76]:

$$S(\omega_1) - S(\omega_{23}) \leq S(\omega_{123}), \tag{2.10}$$

one completes the proof of Proposition 1. □

Note that concavity of entropy implies also another inequality $\sum_{i=1}^{k} p_i S(\rho_i) \leq S(\rho') = S(\sum_{i=1}^{k} p_i \rho_i)$. Proposition 1 has been known before [77] as the *quantum information gain*.

Definition of the channel capacity (1.83) encourages one to consider bounds on the Holevo quantity for the concatenation of two quantum operations. Treating the probabilities p_i and states ρ_i as the outputs from the first channel one can replace maximization over $\mathcal{E} = \{\rho_i, p_i\}_{i=1}^{k}$ in (1.83) by maximization over the initial state ρ and the quantum operation Φ_1. The strategy similar to that used in Theorem 4 allows us to prove the following relations.

Proposition 2. *Consider two quantum operations:* $\Phi_1(\rho) = \sum_{i=1}^{k_1} K_1^i \rho K_1^{i\dagger}$ *and* $\Phi_2(\rho) = \sum_{i=1}^{k_2} K_2^i \rho K_2^{i\dagger}$. *Define* $p_i = \operatorname{Tr} K_1^i \rho K_1^{i\dagger}$ *and* $\rho_i = \frac{K_1^i \rho K_1^{i\dagger}}{\operatorname{Tr} K_1^i \rho K_1^{i\dagger}}$. *The following inequality holds:*

$$S\big(\Phi_2 \circ \Phi_1(\rho)\big) - \sum_{i=1}^{k_1} p_i S\big(\Phi_2(\rho_i)\big) \leq S(\Phi_1(\rho)) - \sum_{i=1}^{k_1} p_i S(\rho_i). \tag{2.11}$$

Proof. Let us consider the four–partite state:

$$\omega'_{1234} = \sum_{n,l=1}^{k_1} \sum_{i,j=1}^{k_2} |i\rangle\langle j| \otimes |nn\rangle\langle ll| \otimes K_2^i K_1^n \rho K_1^{l\dagger} K_2^{j\dagger}, \tag{2.12}$$

where $|nn\rangle \equiv |n\rangle \otimes |n\rangle$, and the strong subadditivity relation in the form

$$S(\omega'_{124}) + S(\omega'_4) \leq S(\omega'_{14}) + S(\omega'_{24}). \tag{2.13}$$

Notice that

$$\begin{aligned} S(\omega'_4) &= S(\Phi_2 \circ \Phi_1(\rho)), \\ S(\omega'_3) - S(\omega'_{24}) &= -\textstyle\sum_i p_i S\big(\Phi_2(\rho_i)\big), \\ S(\omega'_{14}) &= S(\textstyle\sum_{i,j=1}^{k_2} |i\rangle\langle j| \otimes K_2^i \Phi_1(\rho) K_2^{j\dagger}) = S(\Phi_1(\rho)). \end{aligned}$$

The third equality is due to the fact that an isometry, $F_2|\phi\rangle = \sum_{i=1}^{k_2} |i\rangle \otimes K_2^i |\phi\rangle$, does not change the nonzero part of spectrum. This property is also used to justify the following equation

$$S(\omega'_3) - S(\omega'_{124}) = -\sum_{i=1}^{k_1} p_i S(\rho_i). \tag{2.14}$$

Substituting these quantities to the strong subadditivity relation (2.13) we finish the proof. □

Inequality 2.11 is known [8] as the property that the Holevo quantity decreases under a quantum operation $\chi(p_i, \rho_i) \geq \chi(p_i, \Phi(\rho_i))$.

Consider notation used in the proof of Proposition 2. Concavity of the entropy gives

$$\sum_{i=1}^{k_1} p_i S\Big(\Phi_2(\rho_i)\Big) = \sum_{i=1}^{k_1} p_i S\Big(\sum_{j=1}^{k_2} q_j \rho_{ij}\Big) \geq \sum_{i=1}^{k_1}\sum_{j=1}^{k_2} p_i q_j S(\rho_{ij}). \tag{2.15}$$

where $\rho_{ij} = \frac{K_2^j K_1^i \rho K_1^{i\dagger} K_2^{j\dagger}}{\operatorname{Tr} K_2^j K_1^i \rho K_1^{i\dagger} K_2^{j\dagger}}$ and probabilities $p_i q_j = \operatorname{Tr} K_2^j K_1^i \rho K_1^{i\dagger} K_2^{j\dagger}$. Using Theorem 4 and concavity of entropy (2.15) one proves:

Proposition 3. *Consider two quantum operations:* $\Phi_1(\rho) = \sum_{i=1}^{k_1} K_1^i \rho K_1^{i\dagger}$ *and* $\Phi_2(\rho) = \sum_{i=1}^{k_2} K_2^i \rho K_2^{i\dagger}$. *Define* $p_i = \operatorname{Tr} K_1^i \rho K_1^{i\dagger}$ *and* $\rho_i = \frac{K_1^i \rho K_1^{i\dagger}}{\operatorname{Tr} K_1^i \rho K_1^{i\dagger}}$. *The following inequality holds:*

$$S\Big(\Phi_2 \circ \Phi_1(\rho)\Big) - \sum_{i=1}^{k_1} p_i S\Big(\Phi_2(\rho_i)\Big) \leq S(\sigma_{II}), \tag{2.16}$$

where the output of the complementary channel to $\Phi_2 \otimes \Phi_1$ *is denoted as* $\sigma_{II} = \widetilde{\Phi_2 \circ \Phi_1}(\rho)$.

2.1.1 Some consequences

This section provides three applications of theorems proved in Sections 2 and 2.1. One of them concerns the *coherent information*. This quantity is defined for a given quantum operation Φ and an initial state ρ as follows [73]

$$I_{coh}(\Phi, \rho) = S\big(\Phi(\rho)\big) - S\big(\tilde{\Phi}(\rho)\big), \tag{2.17}$$

where $\tilde{\Phi}(\rho)$ is the output state of the channel complementary to Φ. To some extent, coherent information in quantum information theory plays a similar role to mutual information in classical information theory. It is known [8] that $I_{coh}(\Phi, \rho) \leq S(\rho)$. That is a relation similar to (1.74). Moreover, it has been shown that only if $I_{coh}(\Phi, \rho) = S(\rho)$ the process Φ can be perfectly reversed. In this case the perfect quantum error correction is possible [73]. The coherent information is also used to define the *quantum capacity* of a quantum channel [74]

$$C_Q(\Phi) = \max_{\rho} I_{coh}(\Phi, \rho). \tag{2.18}$$

The definition of the coherent information (2.17) can be formulated alternatively [73] by means of an extended quantum operation $\Phi \otimes \mathrm{id}$ acting on a purification $|\psi\rangle \in \mathcal{H}_2 \otimes \mathcal{H}_3$ of an initial state, $\rho = \operatorname{Tr}_3 |\psi\rangle\langle\psi|$. This fact is

justified as follows. The purification of ρ determines as well the purification $\Omega_{123} \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ of the state $\omega \in \mathcal{H}_1 \otimes \mathcal{H}_2$ in (1.29),

$$\Omega_{123} = U_{12} \otimes \mathrm{id}_3 \Big(|1\rangle\langle 1|_1 \otimes |\psi\rangle\langle\psi|_{23} \Big) U_{12}^\dagger \otimes \mathrm{id}_3 \,. \tag{2.19}$$

The partial trace over the environment (subspace $\mathcal{H}_1$) reads

$$\Omega_{23} = [\Phi \otimes \mathrm{id}] \left(|\psi\rangle\langle\psi| \right). \tag{2.20}$$

It has the same entropy as the partial trace over the second and third subspace, $\Omega_1 = \sigma$, which is a state of environment after evolution,

$$S(\sigma) = S([\Phi \otimes \mathrm{id}] \left(|\psi\rangle\langle\psi| \right)), \tag{2.21}$$

and $S(\sigma) = S(\tilde{\Phi}(\rho))$.

Coherent information (2.17) can be written as

$$I_{coh}(\Phi, \rho) = S(\mathrm{Tr}_3 \Omega_{23}) - S(\Omega_{23}). \tag{2.22}$$

The classical counterpart of the coherent information can be defined by using the Shannon entropy instead of the von Neumann entropy and probability vectors instead of density matrices in Eq. (2.22). The classical coherent information is always negative, since the entropy of a joint probability distribution cannot be smaller than its marginal distribution.

Inequalities proved in Theorem 4 and Proposition 1 together provide the following bound on the coherent information,

$$I_{coh}(\Phi, \rho) \leq \sum_{i=1}^{k} p_i S(\rho_i) \leq S(\rho), \tag{2.23}$$

where $p_i = \mathrm{Tr}\, K^i \rho K^{i\dagger}$ and $\rho_i = K^i \rho K^{i\dagger}/p_i$ are defined by Kraus representations of the channel, $\Phi(\rho) = \sum_{i=1}^{k} K^i \rho K^{i\dagger}$. The equality between coherent information and the entropy of initial state $S(\rho)$ guarantees that Φ is reversible. Inequality (2.23) implies a similar, weaker statement: only if the following equality holds $\sum_{i=1}^{k} p_i S(\rho_i) = S(\rho)$, the quantum operation Φ can be reversed.

Another consequence of inequalities proved in Section 2.1 concerns the so called *degradable channels*. These channels are considered in quantum information theory in the context of their capacity [42]. A channel Φ_{deg} is called *degradable* if there exists a channel Ψ such that $\Psi \circ \Phi_{deg} = \tilde{\Phi}_{deg}$. Substituting the degradable channel $\Phi_1 = \Phi_{deg}$ and the additional channel $\Phi_2 = \Psi$ to inequality in Proposition 2 one obtains a lower bound for the average entropy of $\Psi(\rho_i)$, where ρ_i are output states from the channel Φ_{deg},

$$0 \leq \sum_{i=1}^{k} p_i S(\rho_i) - I_{coh}(\Phi_{\mathrm{deg}}, \rho) \leq \sum_{i=1}^{k} p_i S\big(\Psi(\rho_i)\big), \tag{2.24}$$

where $I_{coh}(\Phi,\rho) = S\big(\Phi_{deg}(\rho)\big) - S\big(\tilde{\Phi}_{deg}(\rho)\big)$. The left inequality is due to inequality (2.23). Therefore Proposition 2 provides some characterization of the channel Ψ which is associated with a degradable channel.

The third application of propositions from Section 2.1 is given as follows. The Jamiołkowski isomorphism [34] gives a representation of a quantum map Φ which acts on N dimensional system by a density matrix on the extended space of size N^2. This state can be written as:

$$\sigma_\Phi = [\mathrm{id}\otimes\Phi]\big(\,|\phi^+\rangle\langle\phi^+|\,\big), \tag{2.25}$$

where $|\phi^+\rangle = \frac{1}{\sqrt{N}}\sum_{i=1}^N |i\rangle\otimes|i\rangle$ is the maximally entangled state. A rescaled state $D_\Phi = N\sigma_\Phi$ is called the *dynamical matrix*. In the special case, if the initial state is maximally mixed, $\rho = \frac{1}{N}\mathbb{1}$, the entropy of the correlation matrix σ written in (2.21) is equal to the entropy of the dynamical matrix.

A quantum map Φ can by defined using its Kraus representation (1.30). Since the Kraus representation is not unique [15], one can associate many different correlation matrices with a given quantum operation Φ depending on both, the initial state and the set of Kraus operators. However the entropy of the dynamical matrix D_Φ is invariant under different decompositions. This entropy characterizes the quantum operation and is called the *entropy of a map* [49], denoted by $S^{\mathrm{map}}(\Phi)$ as defined in Eq. (1.45).

Due to Theorem 4 the entropy of a map has the following interpretation. It determines an upper bound on the Holevo quantity (1.76) for a POVM measurement defined by the Kraus operators of Φ if the initial state is maximally mixed $\rho = \rho_* = \frac{1}{N}\mathbb{1}$. Moreover, the entropy of a map is an upper bound for the Holevo quantity for POVM given by any set of Kraus operators $\{K^{i\dagger}K^i\}$ which realize the same quantum operation Φ,

$$\max_{\{K^i\}} \quad \chi\Big(\{p_i = \mathrm{Tr}\,K^i\rho_* K^{i\dagger},\quad \rho_i = \frac{K^i\rho_* K^{i\dagger}}{\mathrm{Tr}\,K^i\rho_* K^{i\dagger}}\}\Big) \le S(\Phi), \tag{2.26}$$

where $\rho' = \Phi(\rho) = \sum_{i=1}^k K^i\rho K^{i\dagger}$.

Proposition 3 provides also an alternative lower bound for the entropy of composition of two quantum maps given by Theorem 3 in [46]. The inequality for the entropy of composition of two maps can be now stated as

$$0 \le \mathrm{MAX}\Big\{S(\Phi_2\circ\Phi_1(\rho_*)) - \sum_{i=1}^k p_i S(\Phi_2(\rho_i)),\ S(\Phi_1) + \Delta\Big\} \le S(\Phi_2\circ\Phi_1), \tag{2.27}$$

where $\Delta = S\big(\Phi_2\circ\Phi_1(\rho_*)\big) - S\big(\Phi_1(\rho_*)\big)$ and $\Phi(\rho) = \sum_{i=1}^k p_i\rho_i$. The lower bound proved in our earlier paper [46] could be smaller than 0. The improved bound is always greater than 0 due to concavity of entropy.

2.2 Discussion on the Lindblad inequality

Lindblad [75] proved an inequality which relates the von Neumann entropy of a state ρ, its image $\rho' = \Phi(\rho) = \sum_{i=1}^k p_i\rho_i$ and the entropy of the correlation

matrix σ equal to the output state of the complementary channel $\sigma = \tilde{\Phi}(\rho)$,

$$|S(\rho') - S(\rho)| \leq S(\sigma) \leq S(\rho') + S(\rho). \tag{2.28}$$

Another two Lindblad inequalities are obtained by permuting the states ρ, ρ' and σ in this formula. The proof of Lindblad proceeds in a similar way to the proof of Theorem 4. It involves a bi–partite auxiliary state $\omega'' = \sum_{i,j=1}^{k} |i\rangle\langle j| \otimes K_i \rho K_j^\dagger$, where the identity $S(\rho) = S(\omega'')$ is due to an isometry similar to F in (2.3). The Araki–Lieb inequality [76], $|S(\rho_1) - S(\rho_2)| \leq S(\rho_{12})$ applied to ω'' proves the left hand side inequality of (2.28), while the subadditivity relation $S(\rho_{12}) \leq S(\rho_1) + S(\rho_2)$ applied to ω'' proves the right hand side inequality of (2.28).

Inequalities from Theorem 4 and Proposition 1

$$S(\rho') - \sum_{i=1}^{k} p_i S(\rho_i) \leq S(\sigma), \tag{2.29}$$

$$\sum_{i=1}^{k} p_i S(\rho_i) \leq S(\rho) \tag{2.30}$$

use a three–partite auxiliary state $\omega = \sum_{i,j=1}^{k} |ii\rangle\langle jj| \otimes K_i \rho K_j^\dagger$. As in the case of the Lindblad inequality (2.28), the identity $S(\rho) = S(\omega)$ holds due to isometry. The strong subadditivity relation applied to ω proves inequality (2.29), while the Araki–Lieb inequality applied for ω proves inequality (2.30). Notice that an extension of the auxiliary state and application of the strong subadditivity relation allows one to use the average entropy to new inequalities for interesting quantities: the entropy of the initial state, the entropy of the output state of a quantum channel $\rho' = \Phi(\rho)$ and the entropy of the output state of the complementary channel $\tilde{\Phi}(\rho)$.

In the case $S(\rho') \geq S(\rho)$ (e.g. for any bistochastic operations) the result (2.29) gives a better lower constraints for $S(\sigma)$ than the Lindblad bound (2.28). In this case

$$S(\rho') - S(\rho) \leq S(\rho') - \sum_{i=1}^{k} p_i S(\rho_i) \leq S(\sigma), \tag{2.31}$$

due to Prop. 1. However, if $S(\rho') \leq S(\rho)$ the result of Lindblad can be more precise depending on the values of $S(\rho)$, $S(\rho')$ and the average entropy $\sum_{i=1}^{k} p_i S(\rho_i)$. In consequence, due to Lindblad inequality (2.28) and the inequality (2.26) one obtains another lower bound for the entropy of a map:

$$\mathrm{MAX}\Big\{ \log(N) - S(\Phi(\rho_*)),\ \max_{\{K^i\}} \chi\Big(p_i = \mathrm{Tr}\, K^i \rho_* K^{i\dagger},\ \rho_i = \frac{K^i \rho_* K^{i\dagger}}{\mathrm{Tr}\, K^i \rho_* K^{i\dagger}}\Big)\Big\} \leq S^{\mathrm{map}}(\Phi), \tag{2.32}$$

where $\rho' = \Phi(\rho) = \sum_{i=1}^{k} K^i \rho K^{i\dagger} = \sum_{i=1}^{k} p_i \rho_i$.

2.3 Inequalities for other entropies

Inequality (2.2) uses the strong subadditivity relation in the form (2.8) which is a specific feature of the von Neumann entropy. Relation (2.8) can be equivalently

formulated in terms of relative von Neumann entropies.

The relative von Neumann entropy $D(\rho_1, \rho_2)$ is defined as follows

$$D(\rho_1, \rho_2) = \mathrm{Tr}\, \rho_1 \big[\log \rho_1 - \log(\rho_2) \big] \tag{2.33}$$

and is finite for $\rho_2 \in \mathrm{supp}(\rho_1)$, otherwise it becomes infinite.

Monotonicity of relative entropy states that for any three–partite quantum state ω_{123} and its partial traces the following inequality holds:

$$D(\omega_{23}, \omega_2 \otimes \omega_3) \leq D(\omega_{123}, \omega_{12} \otimes \omega_3). \tag{2.34}$$

It is an important and nontrivial property of the von Neumann entropy [60], [78]. Monotonicity of the von Neumann entropy (2.34) rewritten using the definition (2.33) leads to the strong subadditivity relation:

$$S(\omega_{123}) + S(\omega_3) \leq S(\omega_{13}) + S(\omega_{23}). \tag{2.35}$$

Complementary partial traces of any multipartite pure state have the same entropy. This fact can be applied to purifications of ω_{123}. Therefore, relation (2.35) is equivalent to (2.8) which can be applied to the specific three–partite state (2.3)

$$\omega_{123} = \sum_{i,j=1}^{k} |i\rangle\langle j| \otimes |i\rangle\langle j| \otimes K_i \rho K_j^\dagger \tag{2.36}$$

and used to prove the upper bound on the Holevo quantity in terms of a correlation matrix $\chi \leq S(\sigma)$. Hence, inequality (2.2) is a consequence of the monotonicity of the relative von Neumann entropy.

Monotonicity of entropy holds also for some generalized entropies e.g. Tsallis entropies of order $0 \leq \alpha < 1$ [79] or Rényi entropies of order $0 \leq q \leq 2$ [80]. Direct generalization of $\chi \leq S(\sigma)$ is not so easy, since the key step in the proof was the strong subadditivity form (2.8). In case of generalized entropies such a form cannot be obtained from the monotonicity of relative entropy.

The Holevo quantity can be expressed by the relative entropy. Consider the state (2.36) and the notation: $K^i \rho K^{i\dagger} = p_i \rho_i$, and $\sum_{i=1}^{k} p_i \rho_i = \rho'$. The relative entropy reads:

$$\begin{aligned}
& D(\omega_{23}, \omega_2 \otimes \omega_3) = && (2.37)\\
= \;& \mathrm{Tr}\, \omega_{23} \log \omega_{23} - \mathrm{Tr}\, \omega_{23} \log \omega_2 - \mathrm{Tr}\, \omega_{23} \log \omega_3 && (2.38)\\
= \;& \sum_{i=1}^{k} \mathrm{Tr}\, p_i \rho_i \log p_i \rho_i - \sum_{i=1}^{k} p_i \log p_i - \mathrm{Tr}\, \rho' \log \rho' && (2.39)\\
= \;& \sum_{i=1}^{k} p_i\, \mathrm{Tr}\, \rho_i \log \rho_i - \mathrm{Tr}\, \rho' \log \rho' && (2.40)\\
= \;& S(\rho') - \sum_{i=1}^{k} p_i S(\rho_i) = \sum_{i=1}^{k} p_i D(\rho_i, \rho') = \chi. && (2.41)
\end{aligned}$$

The equality between the Holevo quantity and relative entropy holds also for the Tsallis entropies of any order q

$$T_\alpha(\rho) = \frac{1}{1-\alpha}\Big[1 - \mathrm{Tr}\,\rho^\alpha\Big], \tag{2.42}$$

where the relative Tsallis entropy D_α^T of order α is defined as [79]

$$D_\alpha^T(\rho_1, \rho_2) = \frac{1}{\alpha-1}\Big[1 - \mathrm{Tr}\,\rho_1^\alpha \rho_2^{1-\alpha}\Big]. \tag{2.43}$$

It is now possible to compute the Tsallis–like generalized relative entropy D_α^T between a bipartite state ω_{23} and the product of its partial traces which leads to the *generalized Holevo quantity* χ_α^T. If one considers the state (2.36)

$$\begin{aligned} D_\alpha^T(\omega_{23}, \omega_2 \otimes \omega_3) &= \frac{1}{\alpha-1}\Big[1 - \mathrm{Tr}\,\omega_{23}^\alpha (\omega_2 \otimes \omega_3)^{1-\alpha}\Big] && (2.44)\\ &= \frac{1}{\alpha-1}\Big[1 - \sum_{i=1}^k \mathrm{Tr}(p_i\rho_i)^\alpha p_i^{1-\alpha} \rho'^{1-\alpha}\Big] && (2.45)\\ &= \sum_{i=1}^k p_i \frac{1}{\alpha-1}(1 - \mathrm{Tr}\,\rho_i^\alpha \rho'^{1-\alpha}) && (2.46)\\ &= \sum_{i=1}^k p_i D_\alpha^T(\rho_i, \rho') \equiv \chi_\alpha^T. && (2.47) \end{aligned}$$

In a similar way we can work with the Rényi entropy $S_q^R(\rho) = \frac{1}{1-\alpha}\log[\mathrm{Tr}\,\rho^\alpha]$. The corresponding relative Rényi entropy reads [81]

$$D_q^R(\rho_1, \rho_2) = \frac{1}{q-1}\log \mathrm{Tr}[\rho_1^q \rho_2^{1-q}] \tag{2.48}$$

and the *Rényi–Holevo quantity* is given by

$$\chi_q^R = \frac{1}{q-1}\log \mathrm{Tr}(\sum_i p_i \rho_i^q)^{1/q}. \tag{2.49}$$

Equality between the generalized Rényi–Holevo quantity (2.49) and the Rényi relative entropy (2.48) holds if relative entropy concerns partial traces of (2.36) and the state $\rho'' = (\sum_i p_i \rho_i^q)^{1/q}$ as follows

$$\chi_q^R = D_q^R(\omega_{23}, \omega_2 \otimes \rho''). \tag{2.50}$$

The Holevo quantity (2.50) is smaller than $D_q^R(\omega_{23}, \omega_2 \otimes \omega_3)$ [81].

The monotonicity of relative entropy for three considered types of generalized entropies: von Neumann entropy, Tsallis entropy of order $0 \le \alpha < 1$ and Rényi entropy of order $0 \le q \le 2$ gives

$$\chi \le D(\omega_{123}, \sigma \otimes \rho'), \tag{2.51}$$

$$\chi_\alpha^T \le D_\alpha^T(\omega_{123}, \sigma \otimes \rho'), \tag{2.52}$$

$$\chi_q^R \le D_q^R(\omega_{123}, \sigma \otimes \rho'). \tag{2.53}$$

These relations state that the Holevo quantity is bounded by the relative entropy between the joint state of the quantum system and its environment and the states of these subsystems taken separately.

In case of von Neumann entropy, inequality (2.51) can be written explicitly as

$$\chi \leq S(\sigma) + S(\rho') - S(\rho). \tag{2.54}$$

Notice that ρ is an initial state and $S(\rho) = S(\omega_{123})$ due to isometry transformation, $F : \rho \to \omega_{123}$. Relation (2.54) joints entropies of the initial state, the final state, the state of the environment and the Holevo quantity in a single formula. Inequality (2.54) which can be rewritten as

$$S(\rho) \leq S(\sigma) + \sum_{i=1}^{k} p_i S(\rho_i) \tag{2.55}$$

gives a finer bound than that provided by the Lindblad inequality: $S(\rho) \leq S(\sigma) + S(\rho')$. Inequality (2.54) can be written as $\chi \leq S(\sigma) + Y$, where $|Y| = |S(\rho') - S(\rho)| \leq S(\sigma)$, due to one of the Lindblad inequalities. In some cases this inequality confines the relation (2.2).

2.4 Searching for the optimal bound

The state σ can be defined for a triple consisting of a probability distribution, set of k density matrices of size N and a set of k unitary matrices, $\{p_i, \rho_i, U_i\}_{i=1}^k$. Every triple (p_i, ρ_i, U_i) defines uniquely the pure state $|\psi_i\rangle$ which is the purification of state ρ_i as follows

$$|\psi_i\rangle = \sum_{r=1}^{N} (U_i \otimes \sqrt{\rho_i} V_i)|e_r\rangle \otimes |e_r\rangle \tag{2.56}$$

as shown in (1.49). The Holevo quantity depends only on $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^k$. Therefore, Theorem 4 can be reformulated as follows:

Theorem 5. *For any ensemble $\{p_i, \rho_i, U_i\}_{i=1}^k$ the Holevo quantity is bounded by the entropy of the correlation matrix σ minimized over all unitary matrices U_i*

$$\chi(\{p_i, \rho_i\}) = S(\sum_{i=1}^{k} p_i \rho_i) - \sum_{i=1}^{k} p_i S(\rho_i) \leq \min_{\{U_i\}} S(\sigma) = \min_{\{U_i\}} S(\sum_{i=1}^{k} p_i |\psi_i\rangle\langle\psi_i|), \tag{2.57}$$

where $|\psi_i\rangle = \sum_{r=1}^{N} (U_i \otimes \sqrt{\rho_i} V_i)|e_r\rangle \otimes |e_r\rangle$ and $\sigma_{ij} = \sqrt{p_i p_j}\, \mathrm{Tr}\, \sqrt{\rho_i}\sqrt{\rho_j} U_j^\dagger U_i$.

The last equality of (2.57) holds since the correlation matrix σ can be represented as the Gram matrix of purifications of ρ_i. It is known that for any Gram matrix equality (1.55) holds.

Finding minimization of $S(\sigma)$ over unitaries is not an easy problem in general. In the following chapter the problem will be solved for the ensemble of $k = 2$ states, and the solution is written in terms of square root of the fidelity between both states. A conjecture that the matrix of the square roots of fidelities also bounds the Holevo quantity for ensembles of $k = 3$ states will be formulated and some weaker bounds will be proved in the next section.

2.4.1 Optimal bound for two matrices

The tightest upper bound on the Holevo quantity occurring in Theorem 5 is obtained by taking minimum of $S(\sigma)$ over the set of unitaries. This is equivalent to the POVM which minimizes the correlation matrix among all POVM which give the same output states. For two output states ρ_1 and ρ_2 occurring with probabilities $(\lambda, 1-\lambda)$ the correlation matrix is given by

$$\sigma = \begin{pmatrix} \lambda & \sqrt{\lambda(1-\lambda)}\,\mathrm{Tr}\,\sqrt{\rho_1}\sqrt{\rho_2}U_2^\dagger U_1 \\ \sqrt{\lambda(1-\lambda)}\,\mathrm{Tr}\,\sqrt{\rho_2}\sqrt{\rho_1}U_1^\dagger U_2 & 1-\lambda \end{pmatrix}. \qquad (2.58)$$

Its entropy is the lowest, if the absolute values of the off–diagonal elements are the largest. As has been shown in Eq. (1.67) the expression $\mathrm{Tr}\,\sqrt{\rho_1}\sqrt{\rho_2}U_2^\dagger U_1$ attains its maximum over unitary matrices at the value

$$\sqrt{F_{12}} = \mathrm{Tr}\,\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}, \qquad (2.59)$$

where for brevity we use F_{12} instead of $F(\rho_1, \rho_2)$. This quantity is equal to the square root fidelity (1.62). Therefore the correlation matrix of the smallest entropy can be rewritten in terms of the square root fidelity,

$$\sigma_{min} = \begin{pmatrix} \lambda & \sqrt{\lambda(1-\lambda)}\sqrt{F_{12}} \\ \sqrt{\lambda(1-\lambda)}\sqrt{F_{12}} & 1-\lambda \end{pmatrix}. \qquad (2.60)$$

2.5 Jensen Shannon Divergence

Minimal entropy of the correlation matrix characterizing an ensemble of two density matrices is related to the distance between them in the set of density matrices. If the probability distribution in (2.60) is uniform, $\lambda = 1/2$, the square root of the von Neumann entropy of σ_{min} forms a metric [53]. It is called the *entropic distance* $D_E(\rho_1, \rho_2)$

$$D_E(\rho_1, \rho_2) = \sqrt{S(\sigma_{min})}, \qquad \sigma_{min} = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{F(\rho_1,\rho_2)} \\ \sqrt{F(\rho_1,\rho_2)} & 1 \end{bmatrix}. \qquad (2.61)$$

Inequality (2.57) provides the relation between this metric and another one defined by means of the *Jensen–Shannon Divergence*. The Jensen–Shannon Divergence $JSD(\{\alpha_\nu P_\nu\})$ has been initially defined [69], [82] as the divergence of classical probability distributions P_ν occurring with probabilities α_ν

$$JSD(\{\alpha_\nu P_\nu\}) = H\Big(\sum_\nu \alpha_\nu P_\nu\Big) - \sum_\nu \alpha_\nu H(P_\nu) = \sum_\nu \alpha_\nu H(P_\nu||\bar{P}) \qquad (2.62)$$

where $H(P)$ denotes the Shannon entropy of the probability distribution P, $H(P_\nu||\bar{P})$ is the relative entropy between P_ν and $\bar{P}$, while the average probability distribution reads $\bar{P} = \sum_\nu \alpha_\nu P_\nu$.

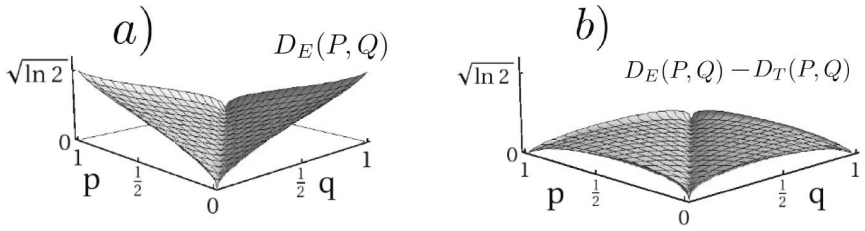


Figure 2.2: *a*) The entropic distance D_E (2.61) for two 2–point probability distributions $P = (p, 1-p)$ and $Q = (q, 1-q)$. *b*) The difference between the entropic distance D_E and the transmission distance D_T (2.64).

The square root of the Jensen-Shannon divergence between two probability distributions P and Q,

$$JSD(P||Q) = \frac{1}{2}H(P||M) + \frac{1}{2}H(P||M), \tag{2.63}$$

where $M = \frac{1}{2}(P+Q)$, forms a metric in the set of classical probability distributions [82], [83] called the *transmission distance* $D_T(P,Q)$,

$$D_T(P,Q) = \sqrt{JSD(P||Q)}. \tag{2.64}$$

A probability distribution can be considered as a diagonal density matrix. Therefore, Eq. (2.57) in Theorem 5 demonstrates a relation between functions of two distances in the set of diagonal density matrices. Fig. 2.2 and Fig. 2.3 shows the comparison between these two distances for exemplary probability distributions.

A quantum counterpart of the Jensen–Shannon divergence, in fact coinciding with the Holevo quantity, was also considered [69], [82]. Inequality (2.57) provides thus an upper bound on the quantum Jensen–Shannon divergence.

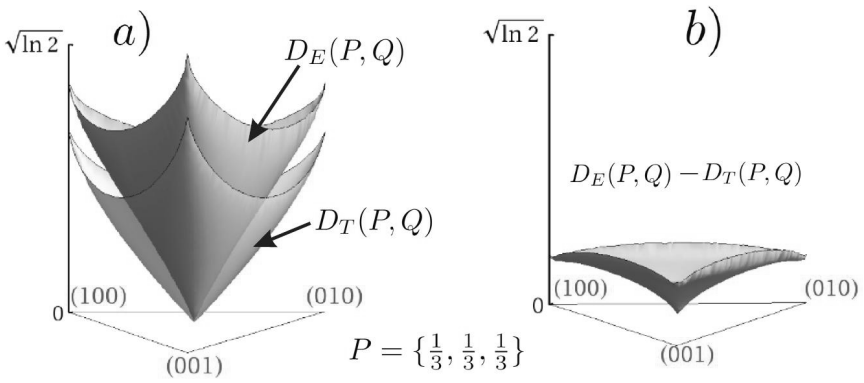


Figure 2.3: *a*) The entropic distance D_E (2.61) and the transmission distance D_T (2.64) for two probability distributions, $P = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $Q = (q_1, q_2, q_3)$ which is arbitrary distribution of dimension 3 represented by a point in the simplex – the base of the figure. *b*) The difference between the entropic distance D_E and the transition distance D_T for the same distributions P and Q.

Chapter 3

Conjecture on three–fidelity matrix

The minimization problem for the entropy of the correlation matrix (2.58) has been solved for an ensemble consisting of $k = 2$ quantum states. In this case the solution is given by the square root fidelity matrix. In the case of $k = 3$ states in the ensemble the optimization over the set of three unitary matrices is more difficult. Our numerical tests support the following conjecture, which is a generalization of the bound found for the case of $k = 2$.

Conjecture 1. *For an ensemble of $k = 3$ quantum states, $\{p_i, \rho_i\}_{i=1}^3$ the entropy of the square root fidelity matrix $G_{ij} = \sqrt{p_i p_j}\sqrt{F(\rho_i, \rho_j)}$ gives the upper bound on the Holevo quantity,*

$$\chi(\{p_i, \rho_i\}) \leq S\left(\begin{bmatrix} p_1 & \sqrt{p_1 p_2}\sqrt{F_{12}} & \sqrt{p_1 p_3}\sqrt{F_{13}} \\ \sqrt{p_2 p_1}\sqrt{F_{21}} & p_2 & \sqrt{p_2 p_3}\sqrt{F_{23}} \\ \sqrt{p_3 p_1}\sqrt{F_{31}} & \sqrt{p_3 p_2}\sqrt{F_{32}} & p_3 \end{bmatrix}\right), \tag{3.1}$$

where fidelity between two quantum states reads $F_{ij} = F(\rho_i, \rho_j) = \left(\mathrm{Tr}\sqrt{\sqrt{\rho_i}\rho_j\sqrt{\rho_i}}\right)^2$.

It has been shown [84], [52] that the matrix G containing square root fidelities is positively semi–defined for $k = 3$. However, the square root fidelity matrix is in general not positive for $k > 3$. Numerical tests provide several counterexamples for positivity of G for $k > 3$, even in case of an ensemble of pure states. Note that the matrix G is not a special case of the correlation matrix σ, which is positive by construction.

Theorem 4 implies that Conjecture 1 holds for ensembles containing three pure states. Inequality (2.2) is in this case saturated as discussed in section 2. Square root fidelity matrix G is obtained from the Gram matrix of given pure states by taking modulus of its matrix entries. Taking modulus of entries of a positive 3×3 matrix does not change neither the trace nor the determinant of the matrix. Only the second symmetric polynomial of the eigenvalues is growing. Since the entropy is a monotonic increasing function of the second symmetric

polynomial [67], the entropy of the square root fidelity matrix G is larger than the entropy of the Gram matrix and therefore it is also larger than the Holevo quantity.

3.1 A strategy of searching for a proof of the conjecture

The proof of Theorem 4 consist of two steps. In the first step one has to find suitable multipartite state. In the second step the strong subadditivity relation of entropy has to be applied for the constructed multipartite state. The same strategy will be used searching for the proof of Conjecture 1 or for proving other weaker inequalities.

For the purpose of obtaining the Holevo quantity from suitable terms of the strong subadditivity relation, the multipartite state ω should have a few features:

- it is a block matrix which is positive,
- blocks on the diagonal should contain states ρ_i multiplied by probabilities p_i,
- traces of off-diagonal blocks should give square root fidelities, or some smaller numbers if one aims to obtain a weaker bound.

The following matrix satisfies above conditions,

$$X = \left[\begin{array}{ccc|ccc|ccc} p_1\rho_1 & 0 & 0 & 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & p_2\rho_2 & 0 & 0 & 0 & * \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & * & 0 & 0 & 0 & p_3\rho_3 \end{array}\right], \tag{3.2}$$

where in place of $*$ one can put any matrix, provided the matrix X remains positive. If in place of $*$ one substitutes zeros, the strong subadditivity relation implies the known formula that $\chi(\{p_i, \rho_i\}) \leq S(\{p_i\})$. Examples presented in the next section use described strategy to prove some entropic inequalities for the Holevo quantity.

The main problem is to find a suitable positive block matrix. In order to check positivity the Schur complement method [85] is very useful.

Lemma 1 (Schur)**.** *Assume that A is invertible and positive matrix, then*

$$X = \begin{bmatrix} A & B \\ B^\dagger & C \end{bmatrix} \tag{3.3}$$

is positive if and only if $\mathsf{S} = C - B^\dagger A^{-1} B$ *is positive semi–definite:*

$$A > 0 \;\; => \;\; (X > 0 \;\; <=> \;\; \mathsf{S} \geq 0). \tag{3.4}$$

The matrix S is called the Schur complement.

3.1.1 Three density matrices of an arbitrary dimension

The strategy mentioned in the previous section will be used to prove the following

Proposition 4. *For a three states ensemble* $\{p_i, \rho_i\}_{i=1,2,3}$ *the following bound for the Holevo quantity* χ *holds*

$$\chi(p_i, \rho_i) \leq S\left(\begin{bmatrix} p_1 & \sqrt{p_1 p_2}\sqrt{F_{12}}/b & \sqrt{p_1 p_3}\sqrt{F_{13}}/b \\ \sqrt{p_2 p_1}\sqrt{F_{21}}/b & p_2 & \sqrt{p_2 p_3}\sqrt{F_{23}}/b \\ \sqrt{p_3 p_2}\sqrt{F_{31}}/b & \sqrt{p_3 p_2}\sqrt{F_{32}}/b & p_3 \end{bmatrix}\right), \tag{3.5}$$

where $b \geq 2$.

Proof. It will be assumed that considered density matrices $\{\rho_i\}_{i=1}^3$ are invertible. After [106] the square root of the product of two density matrices $\sqrt{\rho\sigma}$ will be defined as follows:

$$\sqrt{\rho\sigma} \equiv \rho^{1/2}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\rho^{-1/2} = \sigma^{-1/2}\sqrt{\sigma^{1/2}\rho\sigma^{1/2}}\sigma^{1/2}. \tag{3.6}$$

In this notation the fidelity between two states ρ_i and ρ_j can be written as:

$$F_{ij} = F(\rho_i, \rho_j) = \left(\mathrm{Tr}\sqrt{\rho_i^{1/2}\rho_j\rho_i^{1/2}}\right)^2 = (\mathrm{Tr}\sqrt{\rho_i\rho_j})^2. \tag{3.7}$$

Formula (3.7) can be generalized for non-invertible matrices [52].

One can use the Schur complement Lemma 1 to prove positivity of the block matrix:

$$X = \begin{bmatrix} \rho_1 & \sqrt{\rho_1\rho_2} \\ \sqrt{\rho_2\rho_1} & \rho_2 \end{bmatrix}. \tag{3.8}$$

In this case the matrices A and S, which enter the Lemma 1, take the form: $A = \rho_1$, assume that it is invertible, and $\mathsf{S} = \rho_2 - \sqrt{\rho_2\rho_1}\rho_1^{-1}\sqrt{\rho_1\rho_2}$. Notice that

$$\rho_2 - \mathsf{S} = \sqrt{\rho_2\rho_1}\rho_1^{-1}\sqrt{\rho_1\rho_2}\rho_1\rho_1^{-1} \tag{3.9}$$

$$= \sqrt{\rho_2\rho_1}\sqrt{\rho_2\rho_1}\rho_1^{-1} = \rho_2, \tag{3.10}$$

therefore in the case of matrix (3.8), $\mathsf{S} = 0$ and $X > 0$. Hence the following

matrix Y is also positive:

$$Y = \begin{bmatrix} \frac{1}{2}\rho_1 & 0 & 0 & \frac{1}{2}\sqrt{\rho_1\rho_2} & 0 & 0 & 0 & 0 & \\ 0 & \frac{1}{2}\rho_1 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}\sqrt{\rho_1\rho_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2}\sqrt{\rho_2\rho_1} & 0 & 0 & \frac{1}{2}\rho_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}\rho_2 & 0 & 0 & \frac{1}{2}\sqrt{\rho_2\rho_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2}\sqrt{\rho_3\rho_1} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}\rho_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}\sqrt{\rho_3\rho_2} & 0 & 0 & \frac{1}{2}\rho_3 \end{bmatrix}. \tag{3.11}$$

Using strong subadditivity as described in section 3.1 to the multipartite state $\mathrm{Tr}_2 Y$ extended by some rows and columns of zeros, one proves inequality (3.5) for $b = 2$. To prove relation (3.5) for $b \geq 2$ a small modification of matrix (3.11) is needed. The off–diagonal elements can be multiplied by the number $0 \leq r \leq 1$ without changing the positivity of the block matrix. □

3.1.2 Three density matrices of dimension 2

Proposition 4 can be amended for the case of 2×2 by decreasing the parameter b to the value at least $\sqrt{3}$.

Proposition 5. *For an ensemble of three states of size two, $\{p_i, \rho_i\}_{i=1}^3$ one has*

$$\chi(p_i, \rho_i) \leq S\left(\begin{bmatrix} p_1 & \sqrt{p_1 p_2}\sqrt{F_{12}}/b & \sqrt{p_1 p_3}\sqrt{F_{13}}/b \\ \sqrt{p_2 p_1}\sqrt{F_{21}}/b & p_2 & \sqrt{p_2 p_3}\sqrt{F_{23}}/b \\ \sqrt{p_3 p_2}\sqrt{F_{31}}/b & \sqrt{p_3 p_2}\sqrt{F_{32}}/b & p_3 \end{bmatrix}\right) \tag{3.12}$$

with $b \geq \sqrt{3}$.

Proof. The main task in the proof is to show that the block matrix

$$Y = \begin{bmatrix} p_1\rho_1 & \sqrt{p_1 p_2}\sqrt{\rho_1\rho_2}/b & \sqrt{p_1 p_3}\sqrt{\rho_1\rho_3}/b \\ \sqrt{p_2 p_1}\sqrt{\rho_2\rho_1}/b & p_2\rho_2 & \sqrt{p_2 p_3}\sqrt{\rho_2\rho_3}/b \\ \sqrt{p_3 p_2}\sqrt{\rho_3\rho_1}/b & \sqrt{p_3 p_2}\sqrt{\rho_3\rho_2}/b & p_3\rho_3 \end{bmatrix} \tag{3.13}$$

is positive for $b \geq \sqrt{3}$ as well as the analogous matrix enlarged by adding rows and columns of zeros in order to have a matrix of the form (3.2). The Schur complement method described in section 3.1 will be used, where:

$$A = \begin{bmatrix} \mathbb{1} & 0 \\ 0 & p_1\rho_1 \end{bmatrix}, \qquad C = \begin{bmatrix} p_2\rho_2 & \sqrt{p_2 p_3}\sqrt{\rho_2\rho_3}/b \\ \sqrt{p_2 p_3}\sqrt{\rho_3\rho_2}/b & p_3\rho_3 \end{bmatrix}, \tag{3.14}$$

$$B = \begin{bmatrix} 0 & 0 \\ \sqrt{p_1 p_2}\sqrt{\rho_1\rho_2}/b & \sqrt{p_1 p_3}\sqrt{\rho_1\rho_3}/b \end{bmatrix}, \qquad B^\dagger = \begin{bmatrix} 0 & \sqrt{p_1 p_2}\sqrt{\rho_2\rho_1}/b \\ 0 & \sqrt{p_1 p_3}\sqrt{\rho_3\rho_1}/b \end{bmatrix}. \tag{3.15}$$

Due to the fact that A is positive one needs to prove the positivity of $\mathsf{S} = C - B^\dagger A^{-1} B$:

$$\mathsf{S} = \begin{bmatrix} p_2\rho_2(1-\frac{1}{b^2}) & \sqrt{p_2p_3}(\sqrt{\rho_2\rho_3}/b - \sqrt{\rho_2\rho_1}\rho_1^{-1}\sqrt{\rho_1\rho_3}/b^2) \\ \sqrt{p_2p_3}(\sqrt{\rho_3\rho_2}/b - \sqrt{\rho_3\rho_1}\rho_1^{-1}\sqrt{\rho_1\rho_2}/b^2) & p_3\rho_3(1-\frac{1}{b^2}) \end{bmatrix}. \tag{3.16}$$

To prove positivity of (3.13) the Schur complement S should be positive. One can apply the Schur complement Lemma second time to the matrix S. Positivity condition required by Lemma 1 enforces that

$$b^2(b^2-3)\rho_1 + by \geq 0, \tag{3.17}$$

where $y = \sqrt{\rho_1\rho_2}\rho_2^{-1}\sqrt{\rho_2\rho_3}\rho_3^{-1}\sqrt{\rho_3\rho_1} + h.c.$ For 2×2 matrices one can assume without lost of generality that $\frac{1}{\sqrt{\rho_1}} y \frac{1}{\sqrt{\rho_1}} \geq 0$. It is so because the matrix $\sqrt{\rho_1}^{-1}\sqrt{\rho_1\rho_2}\rho_2^{-1}\sqrt{\rho_2\rho_3}\rho_3^{-1}\sqrt{\rho_3\rho_1}\sqrt{\rho_1}^{-1}$ is a unitary matrix and its determinant is equal to 1, therefore its eigenvalues are two conjugate numbers. The matrix $\frac{1}{\sqrt{\rho_1}} y \frac{1}{\sqrt{\rho_1}}$, which consists of sum of the unitary matrix and its conjugation, is proportional to identity. If it is negative one can change $\sqrt{\rho_1\rho_2}$ into $-\sqrt{\rho_1\rho_2}$ and $\sqrt{\rho_2\rho_1}$ into $-\sqrt{\rho_2\rho_1}$ in (3.13). Transformation changing the sign does not act on the final result because off-diagonal blocks do not take part in forming the Holevo quantity and in the case of 3×3 matrices we can take modulus of each element of the matrix without changing its positivity.

Let us take $y = 0$ in the positivity condition (3.17). This condition implies $b \geq \sqrt{3}$. Knowing that (3.13) is a positive matrix, the rest of the proof of (3.12) goes like in section 3.1. □

3.1.3 Fidelity matrix for one–qubit states

In previous section some bounds on the Holevo quantity were established. These bounds are weaker than the bound postulated by Conjecture 1, since decreasing the off–diagonal elements of a matrix one increases its entropy. In previous proposition the square root fidelities were divided by numbers greater than 1. In the following section the squares of the off–diagonal elements of the matrix G in (3.1) will be taken. For such modified matrices the following proposition holds for an arbitrary number of k states in the ensemble.

Proposition 6. *Consider the ensemble $\{\rho_i, p_i\}_{i=1}^k$ of arbitrary number k of **one-qubit** states and their probabilities. The Holevo information $\chi(\{p_i,\rho_i\})$ is bounded by the entropy of the auxiliary state ς which acts in the k - dimensional Hilbert space,*

$$\chi(\{p_i,\rho_i\}) \leq S(\varsigma), \tag{3.18}$$

where $\varsigma_{ij} = \sqrt{p_ip_j}\,(\mathrm{Tr}\,\sqrt{\rho_i\rho_j})^2 = \sqrt{p_ip_j}\,F(\rho_i,\rho_j)$.

Proof. A positive block matrix W is constructed in the following way:

$$W = \begin{bmatrix} M_1 & 0 & ... & 0 \\ ... & ... & ... & ... \\ M_K & 0 & ... & 0 \end{bmatrix} \begin{bmatrix} M_1^\dagger & ... & M_K^\dagger \\ 0 & ... & 0 \\ ... & ... & ... \\ 0 & ... & 0 \end{bmatrix}, \tag{3.19}$$

where $M_i = \sqrt{p_i}(A_i, B_i)$ are block vectors of size 2×4 and $A_i = \rho_i$ and $B_i = \sqrt{\det\rho_i}\mathbb{1}$ are sub–blocks of size 2×2. The blocks of the block matrix W read

$$W_{ij} = \sqrt{p_i p_j}(\rho_i\rho_j + \sqrt{\det\rho_i\rho_j}\mathbb{1}). \tag{3.20}$$

This formula can be compared with an expression for the square root of any 2×2 positive matrix X

$$\sqrt{X} = \frac{(X + \sqrt{\det X}\mathbb{1})}{\mathrm{Tr}\sqrt{X}}. \tag{3.21}$$

Therefore the block matrix (3.19) is given by

$$W_{ij} = \sqrt{p_i p_j}\sqrt{\rho_i\rho_j}\,\mathrm{Tr}\sqrt{\rho_i\rho_j}. \tag{3.22}$$

The matrix W is positive by construction. Partial trace of this matrix gives matrix of fidelities (without square root). The rest of the proof of Proposition 6 goes in analogy to proofs analysed in Section 3.1. □

This proposition holds for one-qubit states only since we applied relation (3.21), which holds for matrices of dimension $d = 2$.

The fidelity matrix $\varsigma_{ij} = \sqrt{p_i p_j}\,(\mathrm{Tr}\,\sqrt{\rho_i\rho_j})^2 = \sqrt{p_i p_j}\,F_{ij}$ is not positive for a general k and general dimensionality of ρ_i. However the fidelity matrix is positive and bounds the Holevo quantity in the case of an ensemble containing an arbitrary number of pure quantum states of an arbitrary dimension. This is shown in the following proposition.

Proposition 7. *Let $\{|\varphi_j\rangle\}$ be a set of vectors, then*

$$\chi(\{p_i, \rho_i\}) \leq S(\mathsf{F}), \tag{3.23}$$

where $\mathsf{F}_{ij} = \sqrt{p_i p_j}|\langle\varphi_i|\varphi_j\rangle|^2$.

Proof. Introduce a complex conjugation $\varphi \mapsto \overline{\varphi}$ by taking complex conjugations of all coordinates of the state in a given basis. Hence for any choice of φ, ψ one has

$$\langle\varphi|\psi\rangle = \langle\overline{\psi}|\overline{\varphi}\rangle. \tag{3.24}$$

The matrix $F_2 := [F(\rho_i, \rho_j)^2]_{ij}$ can be rewritten as

$$\begin{aligned}[|\langle\varphi_i|\varphi_j\rangle|^2]_{ij} &= [\langle\varphi_i|\varphi_j\rangle\langle\varphi_j|\varphi_i\rangle]_{ij} \\ &= [\langle\varphi_i|\varphi_j\rangle\langle\overline{\varphi_i}|\overline{\varphi_j}\rangle]_{ij} \\ &= [(\langle\varphi_i| \otimes \langle\overline{\varphi_i}|)(|\varphi_j\rangle \otimes |\overline{\varphi_j}\rangle)]_{ij}.\end{aligned} \tag{3.25}$$

This last matrix is the Gram matrix of the set of product states $\{|\varphi_j\rangle \otimes |\overline{\varphi_j}\rangle\}_{j=1}^k$ and therefore is positively defined.

The next part of the proof continues according to the scheme presented in Section 3.1. We use the multipartite state

$$\omega = \sum_{ij} \sqrt{p_i p_j}\langle\varphi_i|\varphi_j\rangle|ii\rangle\langle jj| \otimes |\varphi_i\rangle\langle\varphi_j|. \tag{3.26}$$

Its positivity is shown by taking the partial trace of the Gram matrix

$$\tilde{\omega} = \sum_{ij} |ii\rangle\langle jj| \otimes |\varphi_i\rangle \otimes |\bar{\varphi}_i\rangle\langle\varphi_j| \otimes \langle\bar{\varphi}_j|. \tag{3.27}$$

The proof is completed by considering partial traces of the state ω and using the strong subadditivity relation. □

3.1.4 Special case of the correlation matrix

The previous propositions use the strategy from the proof of Theorem 4 and apply it to positive block matrices which are not necessary related to the correlation matrices. Construction of multipartite states allows one to obtain the matrices containing fidelities after a partial trace. The following section deals again with the correlation matrices $\sigma_{ij} = \sqrt{p_i p_j}\,\mathrm{Tr}\sqrt{\rho_i}\sqrt{\rho_j}U_j^\dagger U_i$. Since the Holevo quantity does not depend on unitaries U_i, these matrices can be chosen in such a way that the three–diagonal of σ consists of the square fidelity matrices, $\sigma_{ij} = \sqrt{F(\rho_i,\rho_j)}$, where $|i-j| \leq 1$. This construction is used in the following proposition.

Proposition 8. *Consider an ensemble $\{\rho_i, p_i\}_{i=1}^k$ consisting of arbitrary number k of invertible states of an **arbitrary dimension**. The Holevo information $\chi(\{p_i,\rho_i\})$ is bounded by the exchange entropy $S(\sigma)$,*

$$\chi(\{p_i,\rho_i\}) \leq S(\sigma), \tag{3.28}$$

where the correlation matrix σ is given by:

$$\sigma_{ii} = p_i, \tag{3.29}$$

$$\sigma_{ij} = \sqrt{p_i p_j}(\mathrm{Tr}\sqrt{\rho_i\rho_j}), \quad \text{iff} \quad |i-j| = 1, \tag{3.30}$$

and the upper off-diagonal matrix elements, where $(j-i) > 1$, read:

$$\sigma_{ij} = \sqrt{p_i p_j}\ \mathrm{Tr}\sqrt{\rho_j\rho_{j-1}}\ \frac{1}{\rho_{j-1}}\ \sqrt{\rho_{j-1}\rho_{j-2}}\ \frac{1}{\rho_{j-2}}\ \ \ldots\ \ \frac{1}{\rho_{i+1}}\ \sqrt{\rho_{i+1}\rho_i}, \tag{3.31}$$

while lower off diagonal satisfy $\sigma_{ij} = \bar{\sigma}_{ji}$.

The matrix σ has a layered structure presented here for $k = 4$,

$$\sigma = \begin{bmatrix} p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \\ 0 & 0 & 0 & p_4 \end{bmatrix} + \begin{bmatrix} 0 & f_{12} & 0 & 0 \\ f_{21} & 0 & f_{23} & 0 \\ 0 & f_{32} & 0 & f_{34} \\ 0 & 0 & f_{43} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & f_{13}^{(2)} & f_{14}^{(3)} \\ 0 & 0 & 0 & f_{24}^{(2)} \\ f_{31}^{(2)} & 0 & 0 & 0 \\ f_{41}^{(3)} & f_{42}^{(2)} & 0 & 0 \end{bmatrix} \tag{3.32}$$

with entries of this matrix equal to σ_{ij} specified in Proposition 8.

Proof. Consider a correlation matrix:

$$\sigma_{ij} = \sqrt{p_1 p_3}\,\mathrm{Tr}\sqrt{\rho_i}\sqrt{\rho_j}U_j^\dagger U_i \tag{3.33}$$

where unitaries U_i are chosen in such a way that elements $\sigma_{i\ i\pm1}$ are square root fidelities: $\sqrt{F_{i\ i\pm1}} = \mathrm{Tr}\sqrt{\rho_i\ \rho_{i\pm1}}$. Hence

$$U_j^\dagger = V_{j-1,j}^\dagger\ U_{j-1}^\dagger, \tag{3.34}$$

where $V_{j-1,j}^\dagger$ is the unitary matrix from the polar decomposition,

$$\sqrt{\rho_i}\sqrt{\rho_j} = |\sqrt{\rho_i}\sqrt{\rho_j}|V_{i,j} = \sqrt{\rho_i^{1/2}\rho_j\rho_i^{1/2}}V_{i,j}. \tag{3.35}$$

Here the Hermitian conjugated unitary matrix $V_{i,j}^\dagger$ reads:

$$V_{i,j}^\dagger = \frac{1}{\sqrt{\rho_j}}\frac{1}{\sqrt{\rho_i}}\sqrt{\rho_i^{1/2}\rho_j\rho_i^{1/2}}. \tag{3.36}$$

The first unitary U_1 can be chosen arbitrarily. The recurrence relation (3.34) allows one to obtain formula (3.31). □

To analyse properties of the matrix σ consider, for example, the matrix element σ_{13}.

$$\begin{aligned}\sigma_{13} &= \sqrt{p_1p_3}\,\mathrm{Tr}\,\sqrt{\rho_1}\sqrt{\rho_3}\,U_3^\dagger U_1\\ &= \sqrt{p_1p_3}\,\mathrm{Tr}\,\sqrt{\rho_1}\sqrt{\rho_3}\,V_{2,3}^\dagger U_2^\dagger U_1\\ &= \sqrt{p_1p_3}\,\mathrm{Tr}\,\sqrt{\rho_1}\sqrt{\rho_3}\,V_{2,3}^\dagger V_{1,2}^\dagger.\end{aligned} \tag{3.37}$$

Using Eq. (3.36) one obtains

$$\begin{aligned}\sigma_{13} =&\sqrt{p_1p_3}\,\mathrm{Tr}\,\sqrt{\rho_1}\sqrt{\rho_3}\frac{1}{\sqrt{\rho_3}}\frac{1}{\sqrt{\rho_2}}\sqrt{\rho_2^{1/2}\rho_3\rho_2^{1/2}}\frac{1}{\sqrt{\rho_2}}\frac{1}{\sqrt{\rho_1}}\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\\ =&\ \sqrt{p_1p_3}\,\mathrm{Tr}\,\frac{1}{\sqrt{\rho_2}}\sqrt{\rho_2^{1/2}\rho_3\rho_2^{1/2}}\sqrt{\rho_2}\frac{1}{\rho_2}\frac{1}{\sqrt{\rho_1}}\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\sqrt{\rho_1}\\ =&\ \sqrt{p_1p_3}\,\mathrm{Tr}\,\sqrt{\rho_3\rho_2}\frac{1}{\rho_2}\sqrt{\rho_2\rho_1},\end{aligned} \tag{3.38}$$

that gives the matrix element σ_{13} of (3.31). The assumption that the matrices are invertible is used in (3.36) where the unitary matrix of the polar decomposition of $\sqrt{\rho_i}\sqrt{\rho_j}$ is given explicitly. However, the same strategy of the proof leads to analogous proposition involving non–invertible matrices. Only the equations (3.36) and (3.31) are changed in this case.

3.1.5 Hierarchy of estimations

One can compare average values of entropies from Conjecture 1 and Propositions 4, 6 and 8. The average values are situated on the scale in which the Holevo quantity is set to 0 and the entropy $S(P)$ of probability distribution is set to unity. The variable $\frac{x-\chi}{S(P)-\chi}$ is used, where x is replaced by the entropy of

respective state. The standard deviations are also computed. The probability distributions are generated according to the Dirichlet measure, while the set of $k = 3$ density matrices is chosen randomly according to the Hilbert–Schmidt measure [86] on the set of density matrices of size 2.

- $< \chi >= 0$
- $< S_{fid} >= 0.176 \pm 0.065$, where S_{fid} corresponds to the entropy from Conjecture 1.
- $< S_{layered} >= 0.193 \pm 0.087$, where $S_{layered}$ corresponds to the entropy from Proposition 8 for $k = 3$ states in the ensemble.
- $< S_{fid^2} >= 0.37 \pm 0.13$, where S_{fid^2} corresponds to the entropy from Proposition 6 for $k = 3$ states in the ensemble.
- $< S_{fid/b} >= 0.750 \pm 0.015$, where $S_{fid/b}$ corresponds to the entropy from Proposition 4.
- $< S(P) >= 1$.

For an ensemble of $k = 3$ one–qubit states Conjecture 1 is the strongest, as it gives on average the lowest bound, while among the statements proved in Propositions 4, 6 and 8 the tightest bound (on average) is provided by Proposition 8.

3.2 Fidelity bound on the Holevo quantity for a special class of states

Although, Conjecture 1 has been confirmed in several numerical tests, it has been proved so far for the set of pure states (Section 3) only. The aim of the following section is to prove that the square root fidelity matrix bounds the Holevo quantity for a restricted set of states. It will be shown that for one–qubit states among which two are pure and one is mixed and for the uniform probability distribution, $\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$, Conjecture 1 holds.

Proposition 9. *Consider $k = 3$ one–qubit states ρ_i among which two are pure $\rho_1 = |\phi_1\rangle\langle\phi_1|$, $\rho_2 = |\phi_2\rangle\langle\phi_2|$, and the state ρ_3 is mixed. The square root fidelity matrix G for these states and the uniform distribution $P = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$ bounds the Holevo quantity,*

$$\chi(\{\rho_i, p_i\}) \leq S\left(\frac{1}{3}\begin{bmatrix} 1 & \sqrt{F_{12}} & \sqrt{F_{13}} \\ \sqrt{F_{21}} & 1 & \sqrt{F_{23}} \\ \sqrt{F_{31}} & \sqrt{F_{32}} & 1 \end{bmatrix}\right), \tag{3.39}$$

where $F_{ij} = (\mathrm{Tr}\sqrt{\sqrt{\rho_i}\rho_j\sqrt{\rho_i}})^2$.

The proof goes as follows. First proper parameters characterizing three states will be chosen. After that the formulas for the left and right side of

inequality (3.39), which are functions of two variables only, will be given. The fact that one of these functions is greater than the other is shown graphically.

Notice that the left hand side of Eq. (3.39) depends only on the lengths of the Bloch vectors which represent the mixed state ρ_3 and the average state $\bar{\rho} = \frac{1}{3}(\rho_1 + \rho_2 + \rho_3)$ inside the Bloch ball. The same average $\bar{\rho}$ can be realized by many triples $\{\rho_1, \rho_2, \rho_3\}$ where ρ_1, ρ_2 are pure and ρ_3 is mixed of given length of the Bloch vector. The family of such triples is parametrized by two numbers α and β as shown in Fig. 3.1. The points B, D, E denote the following states: $B \to \rho_3$ which is mixed, $D \to \rho_1 = |\phi_1\rangle\langle\phi_1|$ and $E \to \rho_2 = |\phi_2\rangle\langle\phi_2|$, while $A \to \bar{\rho}$ represents the average state. The vector $\vec{OA}$ of length a denotes the Bloch vector of the average state $\bar{\rho}$, the vector $\vec{OB}$ of length b characterizes the mixed state ρ_3. The position of the vector $\vec{OB}$ with respect to $\vec{OA}$ can be parametrized by an angle α. These two vectors, $\vec{OB}$ and $\vec{OA}$, determine, but not uniquely, two pure states from the same triple characterized by $\vec{OD}$ and $\vec{OE}$. Equivalently one can rotate the vectors $\vec{OD}$ and $\vec{OE}$ by an angle β around the axis $\vec{OC}$ and obtain pure states denoted by F and G. The ratio $|AB| : |AC|$ is equal to $2:1$ because in this case the average A is the barycenter of three points B, D and E or a triple B, F and G. The method of obtaining the points C, D, E, F and G, when a, b and α are given, is presented in Appendix 1. Given a pair of parameters (a, b) distinguishes the family of triples $\{|\phi_1\rangle\langle\phi_1|, |\phi_2\rangle\langle\phi_2|, \rho_3\}$ characterized by two angles α and β. The range of α is given by condition $|OC| \leq 1$, it is

$$\begin{cases} \frac{1}{2}\sqrt{9a^2 - 6b\cos(\alpha)a + b^2} \leq 1 \\ 0 \leq \alpha \leq \pi, \end{cases} \tag{3.40}$$

while the range of β is $(0, \pi)$. Left hand side of Eq. (3.39) depends only on

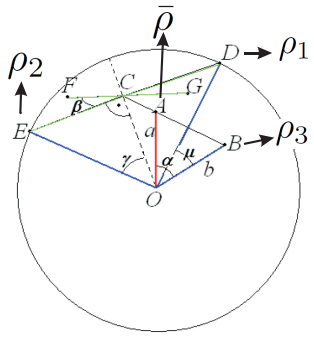


Figure 3.1: The Bloch representation of the three states $\{\rho_1 = |\phi_1\rangle\langle\phi_1|, \rho_2 = |\phi_2\rangle\langle\phi_2|, \rho_3\}$ and the parameters used in the proof of Proposition 9 are presented schematically in the Bloch ball. Two angles (α, β) characterize all possible triples $\{\rho_1, \rho_2, \rho_3\}$ if parameters (a, b) are fixed.

the lengths a and b, and is independent of the concrete realization of the triple. Therefore to prove (3.39) for given a and b one has to find minimum of the

entropy of the square root fidelity matrix over all triples parametrized by the angles α and β.

The entropy of the square–root fidelity matrix defined $G_{ij} = \sqrt{p_i p_j}\sqrt{F_{ij}}$ in Eq. (3.39) is a function of roots of the characteristic polynomial:

$$(\frac{1}{3} - \lambda)^3 + p(\frac{1}{3} - \lambda) + q = 0, \tag{3.41}$$

where

$$p = -(F_{12} + F_{13} + F_{23})/9 \tag{3.42}$$
$$q = 2\sqrt{F_{12}F_{13}F_{23}}/27. \tag{3.43}$$

The parameter p determines the second symmetric polynomial s_2 of eigenvalues of the square root fidelity matrix G

$$s_2 = \frac{1}{9}\sum_{i<j}(1 - F_{ij}). \tag{3.44}$$

The roots of equation (3.41) are equal to:

$$\lambda_k = \frac{1}{3} + 2\sqrt{\frac{-p}{3}}\cos\Big[\Big(\frac{1}{3}\arccos(\frac{3q}{2p}\sqrt{\frac{3}{-p}}) + k\frac{2\pi}{3}\Big)\Big], \tag{3.45}$$

where $k = 1, ..., 3$.

The entropy of the square root fidelity matrix is a function of p and q, which determine the second symmetric polynomial of eigenvalues (3.44) and the third symmetric polynomial is in this case equal to the determinant of the 3×3 matrix $G_{ij} = \sqrt{p_i p_j}\sqrt{F_{ij}}$. The von Neumann entropy is a monotonically increasing function of all symmetric polynomials of eigenvalues [67]. The parameter q is a function of (a, b, α, β), while parameter p depends only on a and b which is shown in following lemma:

Lemma 2. *For any triple of two pure and one mixed state of an arbitrary dimension the sum of fidelities depends only on the purity of the mixed state and the barycenter of the ensemble.*

Proof. Denote by $\bar{\rho}$ the barycenter of a mixed state ρ and two pure states, $|\phi_1\rangle$, $|\phi_2\rangle$,

$$\bar{\rho} = \frac{1}{3}\rho + \frac{1}{3}|\phi_1\rangle\langle\phi_1| + \frac{1}{3}|\phi_2\rangle\langle\phi_2|. \tag{3.46}$$

The purity of $\bar{\rho}$ is given by

$$\operatorname{Tr}\bar{\rho}^2 = \frac{1}{9}\Big(\operatorname{Tr}\rho^2 + 2 + 2\langle\phi_1|\rho|\phi_1\rangle + 2\langle\phi_2|\rho|\phi_2\rangle + 2|\langle\phi_1|\phi_2\rangle|^2\Big). \tag{3.47}$$

After reordering the terms one gets

$$F_{12} + F_{13} + F_{23} = \frac{1}{2}(9\operatorname{Tr}\bar{\rho}^2 - \operatorname{Tr}\rho^2 - 2), \tag{3.48}$$

where

$$F_{12} = |\langle\phi_1|\ \phi_2\rangle|^2\,, \tag{3.49}$$

$$F_{23} = \langle\phi_2|\,\rho\,|\phi_2\rangle\,, \tag{3.50}$$

$$F_{13} = \langle\phi_1|\,\rho\,|\phi_1\rangle\,. \tag{3.51}$$

Since $\mathrm{Tr}\,\bar{\rho}^2 = \frac{1}{2}(1+a^2)$ and $\mathrm{Tr}\,\rho^2 = \frac{1}{2}(1+b^2)$, the parameter p defined in (3.42) does not depend on the angles α and β. This completes the proof of Lemma 2. □

The parameter p and the second symmetric polynomial (3.44) does not depend on the angles α and β. Therefore, for given a and b, the entropy of the square root fidelity matrix attains its minimum over α and β for minimal value of the determinant of G, since the entropy is an increasing function of the determinant. The determinant is given by

$$\det\left(\frac{1}{3}\begin{bmatrix} 1 & \sqrt{F_{12}} & \sqrt{F_{13}} \\ \sqrt{F_{21}} & 1 & \sqrt{F_{23}} \\ \sqrt{F_{31}} & \sqrt{F_{32}} & 1 \end{bmatrix}\right) = \frac{1}{27}\left(1 + 2\sqrt{F_{12}F_{13}F_{23}} - (F_{12}+F_{13}+F_{23})\right). \tag{3.52}$$

It is the smallest for the smallest value of the parameter q which is the function (3.43) of the off–diagonal elements of the matrix. During computations of the minimal value of q another lemma will be useful:

Lemma 3. *Among triples of one–qubit states which realize the same barycenter, where one state is mixed of a given purity and two others are pure, the product of three pairwise fidelities is the smallest if three states and the average lie on the plane containing the great circle of the Bloch ball, i.e.* $\beta = 0$.

Proof. The function $f(a,b,\alpha,\beta) = F_{12}F_{13}F_{31}$ is given explicitly in Appendix 2 based on Appendix 1. For given a, b and α this function has minimum only at $\beta = 0$ and equivalently for $\beta = \pi$. □

In consequence, searching for the minimum of the entropy of the square root fidelity matrix we can restrict our attention to the case $\beta = 0$. In fact, for our purpose it suffices to take the specific value of α which is shown in the following lemma.

Lemma 4. *Among triples of one–qubit states which realize the same barycenter, in which one state is mixed of given purity and two others are pure, the product of three pairwise fidelities is the smallest when two pure states are symmetric with respect to the mixed state i.e.* $\beta = 0$ *and* $\alpha = 0$ *or* $\alpha = \pi$.

Proof. The function $f_0(a,b,\alpha,\beta=0) = F_{12}F_{13}F_{31}$ is given directly in Appendix 1. It has only one minimum at $\alpha = 0$ but in certain cases, depending on a and b, the value on the edge of variable range, i.e. at $\alpha = 0$ or $\alpha = \pi$ is smaller. □

3.2.1 Proof of the fidelity bound

To prove inequality (3.39) the smallest entropy of the square root fidelity matrix for three states consistent with the left hand side of this inequality should be found. Entropy is a function of four parameters, (a, b, α, β). The left hand side of (3.39), which is the Holevo quantity depends on two parameters (a, b) as follows

$$\chi = S\left(\frac{1}{2}\begin{bmatrix} 1+a & 0 \\ 0 & 1-a \end{bmatrix}\right) - \frac{1}{3}S\left(\frac{1}{2}\begin{bmatrix} 1+b & 0 \\ 0 & 1-b \end{bmatrix}\right). \tag{3.53}$$

For given parameters a and b lemmas 1, 2 and 3 allows us to find specific α and β for which minimization of right hand side of (3.39) is obtained. One can fix $\alpha = 0$ or $\alpha = \pi$ and $\beta = 0$. That means, that minimal entropy of the square root fidelity G over the angles is obtained if the three states $\{\rho_1 = |\phi_1\rangle\langle\phi_1|, \rho_2 = |\phi_2\rangle\langle\phi_2|, \rho_3\}$ are lying on the great circle and the two pure states are symmetric with respect to the mixed state. In this case the matrix G is characterized by two parameters, $F = F_{12} = F_{23}$ and b. Here F is the fidelity between the pure state ρ_1 and the mixed state ρ_3 whereas b characterize the length of the Bloch vector of the mixed state ρ_3. The matrix G reads

$$G = \frac{1}{3}\begin{pmatrix} 1 & \sqrt{F} & |\frac{2F-1}{b}| \\ \sqrt{F} & 1 & \sqrt{F} \\ |\frac{2F-1}{b}| & \sqrt{F} & 1 \end{pmatrix}, \tag{3.54}$$

where F is a function of b, such that $F(b) = \frac{1}{2}(1 - bc)$, and c is the length of the Bloch vector representing the barycenter of two pure states ρ_1 and ρ_2. The fidelity F is equal to $1/2$ if b tends to 0. The parameter c determines also the projection of the Bloch vector of the pure state ρ_1 on the Bloch vector of the mixed state ρ_3. The absolute value $|c|$ is equal to the square root fidelity between the two pure states. The range of variables are $0 \leq b \leq 1$ and $\frac{1}{2}(1-b) \leq F \leq \frac{1}{2}(1+b)$.

Considered case is shown in Fig. 3.2. There are two surfaces – functions of two parameters F and b. The lower surface represents the Holevo quantity χ, and the upper surface denotes the entropy of the square root fidelity matrix (3.54). The surface $S(G)$ lies always above χ and is composed of two smooth functions characterizing cases in which all vectors lay on the same semicircle or pure states and the mixed state belong to the opposite semicircles.

Fig. 3.2 suggests that in the case of three pure states, $b = 1$, laying on the same semicircle the inequality is saturated, $\chi = S(G)$. In this case, $F \geq 1/2$, the rank of the square root fidelity matrix is equal to 2, and the nonzero eigenvalues are $(1 \pm a)/2$, where $a = (4F - 1)/3$ is the length of the Bloch vector of the average state $\bar{\rho}$. In general we have $a = \frac{1}{3}(b + 2\frac{2F-1}{b})$. In case of $\chi = S(G)$ the Holevo quantity is equal to the entropy of the average state $\bar{\rho}$. This finishes the proof of Proposition 9.

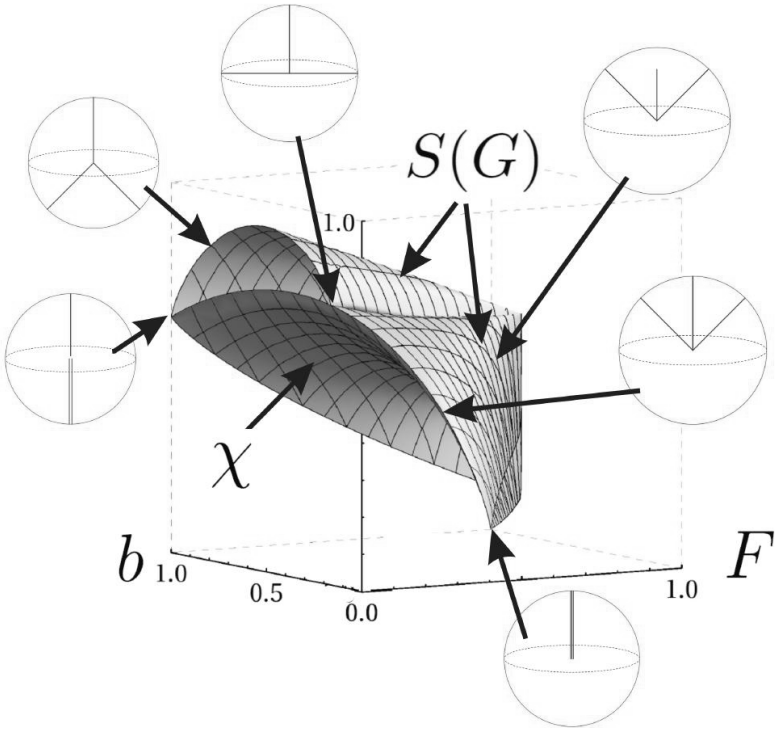


Figure 3.2: Evidence in favour of Proposition 9. The Holevo quantity as function of two variables: fidelity F between the pure state ρ_1 and the mixed state ρ_3, and the length b of the Bloch vector characterizing the state ρ_3. The upper surface representing the square root fidelity matrix G is composed of two smooths parts. Every circle represents schematically the Bloch ball with exemplary positions of Bloch vectors characterizing three states $\{\rho_1 = |\phi_1\rangle\langle\phi_1|, \rho_2 = |\phi_2\rangle\langle\phi_2|, \rho_3\}$ of the ensemble.

Part III

Minimal output entropy and map entropy

Chapter 4

Entropies for one-qubit channels

The question on additivity of the channel capacity is one of the most interesting problems in quantum information theory [40]. Shor showed [39] that this problem has several equivalent formulations. One of them concerns the minimal output entropy,

$$S^{\min}(\Phi) = \min_{\rho} S(\Phi(\rho)). \tag{4.1}$$

In the case of one–qubit channel the minimal output entropy is the entropy of a state characterized by point on the ellipsoid, which is the image of the Bloch sphere, the closest to this sphere. The pure state which is transformed into a state of the minimal entropy is called minimizer.

For any setup in which minimal output entropy is additive the quantum channel capacity is additive as well. Additivity implies that an entangled state cannot increase capacity of two channels with respect to the sum of their capacities taken separately. The additivity conjecture can also be formulated as a statement that capacity of two channels is minimized for a product state.

The conjecture was confirmed in many special cases. For instance, additivity holds, if one of the channels is arbitrary and the second one is: bistochastic one–qubit map [87], a unitary transformation [40], generalized depolarizing channel [41], entanglement breaking channel [88], very noisy map [89] and others. A useful review on this subject was written by Holevo [90]. Different strategies of proving the additivity conjecture are analyzed there. For a recent relation on the additivity conjecture see also [18].

Also counterexamples to the additivity conjecture have been found. One of them was presented by Hastings [16]. He found the lower bound for the output entropy of some channels when the input was a product state. Next he estimated the output entropy for a maximally entangled input. Due to such estimations it was shown that the entangled state decreases channel capacity below the value achievable for product states.

The proof of Hastings used pairs of complementary channels. His argument was not constructive and works in high dimensional spaces. Further counterexamples for the additivity hypothesis presented in [17] are constructive.

It is still an open question, whether the additivity holds for an arbitrary one–qubit channel. Originally, the hypothesis on additivity of minimal output entropy was formulated for the von Neumann entropy. One of the approaches to the problem uses a one–parameter family of entropies, called Rényi entropies characterized by a parameter q,

$$S_q(\rho) := \frac{1}{1-q} \log \operatorname{Tr} \rho^q. \tag{4.2}$$

Calculations are sometimes easier when the Rényi entropies are considered. The quantity S_q tends to the von Neumann entropy in the limit $q \to 1$. Additivity of the minimal output Rényi entropy has been proved only in some range of the parameter q depending on the channels considered [18, 41, 87].

Although the Rényi entropy is sometimes computationally more feasible, finding minimum over entire set of quantum states is still a hard problem. One of the ideas how to omit this difficulty tries to use some relations between minimal output entropy and other quantities which are easier to calculate. In the following chapter the Rényi entropy of a map (the map entropy) will be used to estimate the minimal output entropy. Map entropy (entropy of a map) is defined by the entropy of the Choi-Jamiołkowski state (1.28) corresponding to the map. This quantity is easy to obtain. Numerical tests presented in Fig. 4.2, 4.4, 4.5 show that there is no simple functional relation between the map entropy and the minimal output entropy. Nevertheless being aware of the structure of the set of quantum maps projected on the plane $(S_q^{\max}, S_q^{\min})$ can be useful. Knowledge of entropies of maps at the boundaries of the allowed set can be used to estimate the minimal output entropy by the entropy of the map.

4.1 Structure of the set of Pauli channels

Quantum channels which preserve the maximally mixed state are called bistochastic. All bistochastic one–qubit channels can be represented as a convex combination of the identity matrix $\sigma_0 = \mathbb{1}$ and three Pauli matrices $\sigma_{i=1,2,3}$ (1.39)

$$\Phi_{\vec{p}}(\rho) = \sum_{i=0}^{3} p_i \sigma_i \rho \sigma_i, \qquad \sum_{i=0}^{3} p_i = 1, \qquad \forall_i p_i \geq 0. \tag{4.3}$$

Bistochastic one–qubit quantum operations are thus called *Pauli channels*. The structure of the set of all Pauli channels forms a regular tetrahedron Δ_3 as shown in Fig. 4.1*a*. There are many channels characterized by the points of tetrahedron which can be obtained from other channels following a unitary transformation. Our considerations are often restricted to the asymmetric tetrahedron K (see Fig. 4.1*b*) which is a subset of Δ_3. All maps in Δ_3 can be obtained from channels of K by concatenation these channels with unitary

transformations. The set K is formed by the convex combination of four vectors $\vec{p}$ from (4.3), $A = (0,0,0,0)$, $B = (1/2,1/2,0,0)$, $C = (1/3,1/3,1/3,0)$, and $D = (1/4,1/4,1/4,1/4)$. Extremal lines of the asymmetric tetrahedron

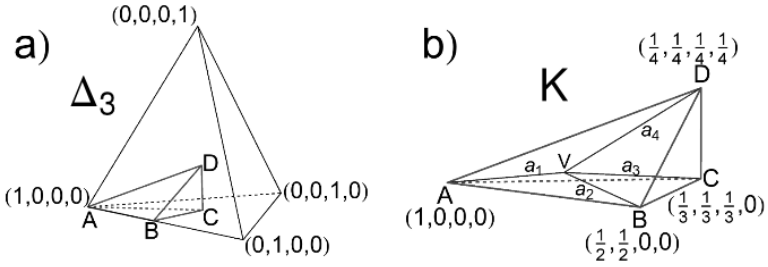


Figure 4.1: The structure of one–qubit bistochastic quantum operations corresponds to the regular tetrahedron Δ_3. This figure is spanned by four extremal vectors $\vec{p}$ from formula (4.3). Symmetries of the tetrahedron allow us to distinguish the asymmetric set K inside Δ_3. Any vector $\vec{p}$ characterizing a Pauli channel can be obtained by permutation of elements of vectors from K.

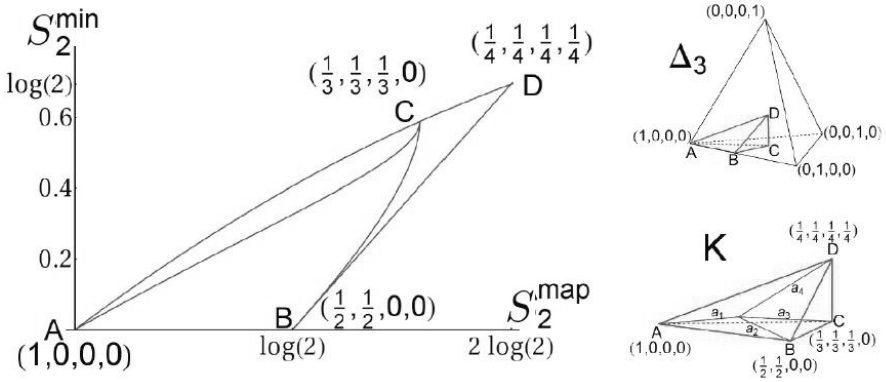


Figure 4.2: Lines AB, BD and AD, which correspond to the edges of asymmetric tetrahedron K form the boundaries of the entire set of Pauli matrices projected on the plane $(S_2^{\max}, S_2^{\min})$.

correspond to the following families of maps: AB - dephasing channels, BD - classical bistochastic maps, AD and CD - depolarizing channels. The families mentioned above are also shown in Fig. 4.2 which presents boundaries of the set of all one–qubit bistochastic channels projected onto the plane $(S_2^{\max}, S_2^{\min})$. A following proposition proved in [51] characterizes this projection.

Proposition 10. *Extremal lines of asymmetric tetrahedron correspond to boundaries of the set of all bistochastic one–qubit maps on the plot* $(S_2^{\max}, S_2^{\min})$.

4.2 Depolarizing channels

Fig. 4.4 and Fig. 4.5 show the projection of the Pauli channels on the plane $(S_q^{\max}, S_q^{\min})$ with parameter q different than 2. Comparison of these figures with Fig. 4.2 shows that the structure of the set of channels on the plane $(S_q^{\max}, S_q^{\min})$ is the simplest in case of the Rényi entropy of order $q = 2$. In this case, the depolarizing channels form one of the edges of the set of all quantum one–qubit maps projected onto the plane. Indeed the following theorem proved in [51] confirms the special role of depolarizing channels in the set of all quantum channels acting on states of arbitrary dimension N.

Theorem 6. *Depolarizing channels have the smallest map Rényi entropy $S_2^{\max}$ among all channels with the same minimal output Rényi entropy $S_2^{\min}$.*

The family of depolarizing channels is represented in the plane $(S_2^{\max}, S_2^{\min})$ by the continuous line on the entire range of $S_2^{\max}$. The minimal output entropy of a depolarizing channel Λ_N acting on $\mathcal{M}_N$ is the following function of the map entropy

$$S_2^{\min}\Big(S_2^{\max}(\Lambda_N)\Big) = -\log\Big(\frac{1 + N\mathrm{e}^{-S_2^{\max}(\Lambda_N)}}{N+1}\Big). \tag{4.4}$$

This is a monotonously increasing function from 0 to $\log N$. Therefore the following theorem holds.

Theorem 7. *Depolarizing channels have the greatest minimal output Rényi entropy $S_2^{\min}$ among all maps of the same Rényi entropy of a map $S_2^{\max}$.*

One can try to use the extremal position of depolarizing channels to estimate the minimal output entropy of some channels. In the case of Hastings' counterexample for the additivity conjecture the author showed that due to a maximally entangled input state one can obtain smaller output entropy of the product of two channels than in the case of any product state taken as an input. Let us estimate the Rényi $q = 2$ output entropy for a product channel when the input is maximally entangled. Following proposition proved in [51] presents one of estimations.

Proposition 11. *For any entropy $\mathcal{S}$ which is subadditive the following inequality holds*

$$|\mathcal{S}^{\max}(\Phi_1) - \mathcal{S}^{\max}(\Phi_2)| \leq \mathcal{S}\Big([\Phi_1 \otimes \Phi_2](|\phi_+\rangle\langle\phi_+|)\Big) \leq \mathcal{S}^{\max}(\Phi_1) + \mathcal{S}^{\max}(\Phi_2), \tag{4.5}$$

where $|\phi_+\rangle\langle\phi_+|$ is a maximally entangled state.

Proof. The proof starts form the Lindblad inequality [75], which is based on the subadditivity of the von Neumann entropy,

$$\big|\mathcal{S}(\rho) - \mathcal{S}\big(\varsigma(\Phi,\rho)\big)\big| \leq \mathcal{S}(\Phi(\rho)) \leq \mathcal{S}(\rho) + \mathcal{S}\big(\varsigma(\Phi,\rho)\big), \tag{4.6}$$

where $\varsigma(\Phi,\rho) = [\mathrm{id}\otimes\Phi]\Big(|\phi\rangle\langle\phi|\Big)$ and $|\phi\rangle$ is a purification of ρ as in (2.21). The entropy of this state, $S(\varsigma(\Phi,\rho))$, is the exchange entropy which does not depend

on the choice of purification [29]. The state ς defined for a channel Φ and the maximally mixed state $\rho_* = \mathbb{1}/N$ is equal to the normalized dynamical matrix of Φ (1.25),

$$\varsigma(\Phi, \rho_*) = \sigma_\Phi = \frac{1}{N} D_\Phi, \tag{4.7}$$

The entropy of this state defines $\mathcal{S}^{\rm map}(\Phi)$. Since the map Φ is trace preserving, the condition $\mathrm{Tr}_2 \sigma_\Phi = \frac{1}{N}\mathbb{1}$ holds, see (1.26). Apply Lindblad formula (4.6) to the state

$$[\Phi_1 \otimes \Phi_2](|\psi_+\rangle\langle\psi_+|) = [\Phi_1 \otimes \mathrm{id}]\left([\mathrm{id} \otimes \Phi_2](|\psi_+\rangle\langle\psi_+|)\right), \tag{4.8}$$

where $|\psi_+\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^N |i\rangle \otimes |i\rangle$ is the maximally mixed state which is a purification of ρ_*. Expression (4.6) applied to this state gives

$$\begin{aligned} \left|\mathcal{S}^{\rm map}(\Phi_2) - \mathcal{S}\big(\varsigma(\Phi_1 \otimes \mathrm{id},\, \sigma_{\Phi_2})\big)\right| &\leq \mathcal{S}\big((\Phi_1 \otimes \Phi_2)(|\psi_+\rangle\langle\psi_+|)\big) \\ &\leq \mathcal{S}^{\rm map}(\Phi_2) + \mathcal{S}\big(\varsigma(\Phi_1 \otimes \mathrm{id},\, \sigma_{\Phi_2})\big). \end{aligned} \tag{4.9}$$

The exchange entropy $\mathcal{S}\big(\varsigma(\Phi_1 \otimes \mathrm{id},\, \sigma_{\Phi_2})\big)$ is the same as $\mathcal{S}\big(\varsigma(\Phi_1,\, \mathrm{Tr}_2\, \sigma_{\Phi_2})\big)$ since a purification of σ_{Φ_2} is as well the purification of $\mathrm{Tr}_2\, \sigma_{\Phi_2}$ and the exchange entropy does not depend on a purification. Due to the trace preservation formula, $\mathrm{Tr}_2\, \sigma_{\Phi_2} = \rho_*$, the state $\varsigma(\Phi_1,\, \mathrm{Tr}_2\, \sigma_{\Phi_2}) = \varsigma(\Phi_1,\, \rho_*) = \sigma_{\Phi_1}$ which completes the proof. □

Proposition 11 is applicable for any entropy which is subadditive. The Rényi entropy of order $q = 2$ is not subadditive, however, it is a function of the Tsallis entropy of order 2 for which the subadditivity holds. Therefore Proposition 11 can be used to estimate the output Rényi $q = 2$ entropy of a product channel if the input state is maximally entangled. The following inequality corresponds to Rényi $q = 2$ version of the lower bound in (4.5),

$$-\log\Big(1 - |e^{-S_2^{\rm map}(\Phi_1)} - e^{-S_2^{\rm map}(\Phi_2)}|\Big) \leq S_2\Big((\Phi_1 \otimes \Phi_2)(|\psi_+\rangle\langle\psi_+|)\Big). \tag{4.10}$$

It is possible to find channels Φ_1 and Φ_2 such that the left hand side of (4.10) is greater than $S_2^{\min}$ of depolarizing channel Λ, which has the same map entropy as $S_2^{\rm map}(\Phi_1 \otimes \Phi_2)$. Notice that for any two channels the map entropy of their tensor product is characterized by the following result.

Proposition 12. *The Rényi map entropy $S_q^{\rm map}$ is additive with respect to tensor product of quantum maps for any parameter $q \geq 0$:*

$$S_q^{\rm map}(\Phi_1 \otimes \Phi_2) = S_q^{\rm map}(\Phi_1) + S_q^{\rm map}(\Phi_2). \tag{4.11}$$

Proof. The map entropy $S_q^{\rm map}(\Phi)$ is defined as the entropy of normalized dynamical matrix D_Φ. The matrix representation of $D_{\Phi_1 \otimes \Phi_2}$ is related to superoperator matrix of the quantum operation $\Phi_1 \otimes \Phi_2$, due to formula (1.28). Using explicit calculations on matrix elements one can show that $D_{\Phi_1 \otimes \Phi_2}$ is unitarily

equivalent with $D_{\Phi_1} \otimes D_{\Phi_2}$. That implies the additivity of the map entropies, since the quantum Rényi entropy of any order of a given state is a function of its spectrum.

Consider a set of N-dimensional matrices equipped with the Hilbert-Schmidt inner product

$$\langle A|B\rangle_{\mathbf{h}} := \operatorname{Tr} A^\dagger B. \tag{4.12}$$

In this space the matrix units $\big\{|i\rangle\langle j| \;\big|\; i,j = 1,2,\ldots,N\big\}$ form an orthonormal basis. The elements of this basis are denoted by $|i\rangle\langle j| := |ij\rangle_{\mathbf{h}}$. A quantum operation Φ is represented by a matrix $\hat{\Phi}$:

$$\langle ij|\hat{\Phi}|k\ell\rangle_{\mathbf{h}} = \operatorname{Tr}\Big(|j\rangle\langle i|\,\Phi(|k\rangle\langle\ell|)\Big), \tag{4.13}$$

hence

$$\Phi(|k\rangle\langle\ell|) = \sum_{i,j} \langle ij|\hat{\Phi}|k\ell\rangle_{\mathbf{h}}\,|i\rangle\langle j|. \tag{4.14}$$

Due to the reshuffling procedure (1.28), the entries of the dynamical matrix D_Φ read

$$\langle ab|D_\Phi|cd\rangle_{\mathbf{h}} = \langle ac|\hat{\Phi}|bd\rangle_{\mathbf{h}}. \tag{4.15}$$

The entries of $D_{\Phi_1\otimes\Phi_2}$ are obtained by using unnormalized maximally entangled state $|\Psi_+\rangle := \sum_{i,\ell} |i\ell\rangle \otimes |i\ell\rangle$ according to definition (1.25) as follows

$$\begin{aligned}\langle abcd|D_{\Phi_1\otimes\Phi_2}|efgh\rangle &= \langle abcd|\big[(\Phi_1\otimes\Phi_2)\otimes \mathrm{id}\big]\big(|\Psi_+\rangle\langle\Psi_+|\big)|efgh\rangle \\ &= \sum_{i,\ell,j,m} \langle abcd|\big[(\Phi_1\otimes\Phi_2)(|i\ell\rangle\langle jm|)\otimes |i\ell\rangle\langle jm|\big]|efgh\rangle.\end{aligned} \tag{4.16}$$

Now expression (4.14) is used and the matrix elements of $D_{\Phi_1\otimes\Phi_2}$ read

$$\langle abcd|D_{\Phi_1\otimes\Phi_2}|efgh\rangle = \sum_{\alpha,\beta,\gamma,\delta} \langle\alpha\beta|\widehat{\Phi_1}|ij\rangle_{\mathbf{h}}\,\langle\gamma\delta|\widehat{\Phi_2}|ij\rangle_{\mathbf{h}}\,\langle abcd|\alpha\gamma i\ell\rangle\,\langle\beta\delta jm|efgh\rangle. \tag{4.17}$$

Since $\langle abcd|\alpha\gamma i\ell\rangle$ is expressed in terms of Kronecker deltas $\delta_{a\alpha}\delta_{b\gamma}\delta_{ci}\delta_{d\ell}$ and $\langle\beta\delta jm|efgh\rangle$ analogously, the summation over the Greek indexes gives,

$$\begin{aligned}\langle abcd|D_{\Phi_1\otimes\Phi_2}|efgh\rangle &= \langle ac|D_{\Phi_1}|eg\rangle\,\langle bd|D_{\Phi_2}|fh\rangle \\ &= \langle acbd|D_{\Phi_1}\otimes D_{\Phi_2}|egfh\rangle.\end{aligned} \tag{4.18}$$

The matrix $D_{\Phi_1\otimes\Phi_2}$ is related to $D_{\Phi_1}\otimes D_{\Phi_2}$ by a unitary matrix $U = \sum_{a,b,c,d} |abcd\rangle\langle acbd|$. Therefore both matrices have the same eigenvalues and the same entropies. □

Since minimal output entropy of a depolarizing channel Λ is a function of its map entropy (4.4), the estimation on the left hand side of (4.10) can be made in terms of such Λ for which $S_2^{\max}(\Lambda) = S_2^{\max}(\Phi_1) + S_2^{\max}(\Phi_2)$. As a result of this estimation one obtains condition on the pair of channels, for which a maximally

mixed input state does not decrease the output entropy below the smallest value obtained by the product input state,

$$1-\frac{MN+1}{MN}|e^{-S_2^{\max}(\Phi_1)}-e^{-S_2^{\max}(\Phi_2)}| \leq e^{-S_2^{\max}(\Phi_1\otimes\Phi_2)} = e^{-\left[S_2^{\max}(\Phi_1)+S_2^{\max}(\Phi_2)\right]}, \tag{4.19}$$

where Φ_1 acts on $\mathcal{M}_N$ and Φ_1 on $\mathcal{M}_M$. Fig. 4.3 presents the region defined

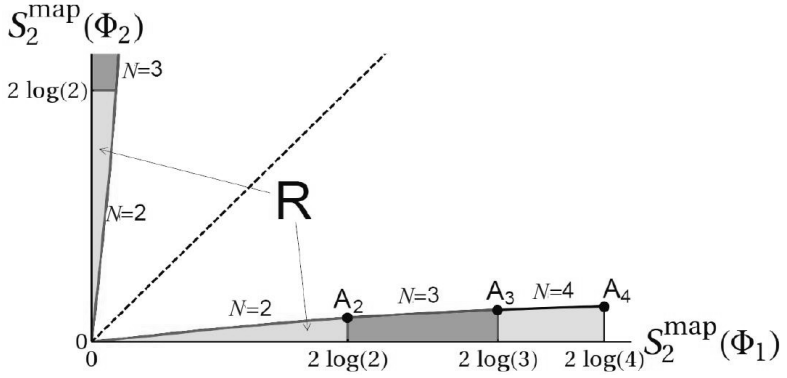


Figure 4.3: Colored parts of the figure denote the region described by inequality (4.19). This region contains pairs of maps characterized by their map entropy for which the additivity is conjectured. The dotted line contains the pairs of complementary channels. The region is enlarged if a larger dimensions are considered.

by (4.19). Such a set is not empty and contains maps, for which $S_2^{\max}(\Phi_1) << S_2^{\max}(\Phi_2)$ or $S_2^{\max}(\Phi_2) << S_2^{\max}(\Phi_1)$. The dotted line represents the set of complementary channels for which both map entropies are equal. This set contains the channels breaking the conjecture of additivity of minimal output entropy according to the proof of Hastings. The region defined by (4.19) does not intersect the set. It was also shown [40], [89] that additivity holds if one of the channels is unitary or if one of the channels is very noisy. These both cases are covered by condition (4.19). These examples support formulation of

Conjecture 2 ([51]). *The additivity of minimal output Rényi $q = 2$ entropy holds for pair of channels satisfying inequality (4.19).*

Recent literature does not answer the question, whether the additivity conjecture is broken for low dimensional channels and the Rényi entropy of order $q = 2$. Our Conjecture 2 suggests for which pairs of channels finding a counterexample of additivity is unlikely. Conjecture 2 uses the map entropy and is formulated for the Rényi entropy of order 2, for which the theorem about extremal position of the depolarizing channels was proved. This is the key theorem which allows us to derive estimations (4.10) and (4.19). Numerical tests (Fig. 4.4 and 4.5) suggest that the depolarizing channels are not situated at the boundary of the set of all channels in the plane $(S_q^{\max}, S_q^{\min})$ for $q \leq 2$,

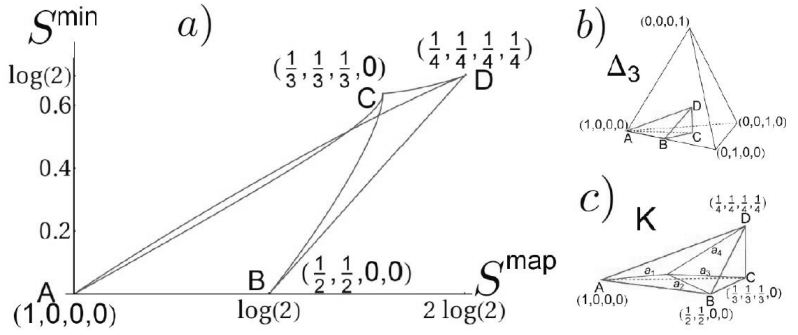


Figure 4.4: The set of the Pauli channels projected on the plane spanned by the map entropy S^{map} and the minimal output entropy $S^{\min}$. The von Neumann entropies are considered. Solid curves correspond to the edges of the asymmetric tetrahedron K. The curve AD characterizing the family of depolarizing channels does not belong to the boundary of the set.

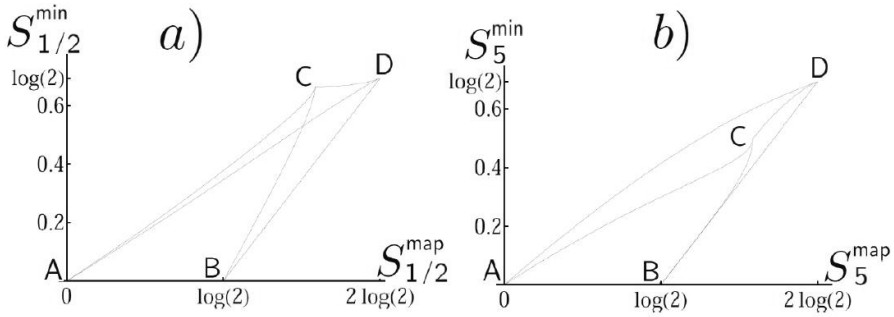


Figure 4.5: As in Fig. 4.4*a*: projection of the set of Pauli channels onto the plane spanned by the Rényi entropy of a map S_q^{map} and the minimal Rényi output entropy $S_q^{\min}$ obtained for *a*) $q = 1/2$ and *b*) $q = 5$.

while their extremal position could be confirmed in case $q \geq 2$. Nevertheless, the Rényi entropy is a smooth function of q. Therefore, a conjecture similar to Conjecture 2 may hold also for other values of the Rényi parameter q.

4.3 Transformations preserving minimal output entropy

In previous chapter the set of one–qubit quantum operations was considered in context of the plot $(S_q^{\max}, S_q^{\min})$. One could ask, whether the family of maps lying at the same vertical or horizontal line can be characterized. The following section gives a partial answer to this question. Transformations of one qubit maps which preserve the minimal output entropy will be considered. Such a transformation changes the quantum channel and moves the corresponding point in the plane $(S_q^{\max}, S_q^{\min})$ along a given horizontal line. In the following section we consider the geometrical picture of one–qubit maps acting on the set of pure states. One–qubit quantum operation transforms the Bloch ball into an ellipsoid inside the ball. A transformation of quantum operation which changes the lengths of the axes of the ellipsoid and their orientation and leaves the minimal output entropy unchanged will be studied.

Consider the superoperator matrix of a one–qubit quantum operation:

$$\Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \overline{\Phi_{12}} & \Phi_{14} \\ \overline{\Phi_{21}} & \Phi_{22} & \overline{\Phi_{32}} & \Phi_{24} \\ \overline{\Phi_{21}} & \Phi_{32} & \overline{\Phi_{22}} & \overline{\Phi_{34}} \\ 1-\Phi_{11} & -\Phi_{12} & -\overline{\Phi_{12}} & 1-\Phi_{14} \end{pmatrix}. \tag{4.20}$$

Parameters Φ_{11} and Φ_{14} are real, the complex conjugation of Φ_{ij} is denoted by $\overline{\Phi_{ij}}$. The form (4.20) guarantees that the dynamical matrix of Φ is Hermitian and the trace preserving condition (1.26) is satisfied.

Assume that the quantum operation Φ_1 has the output entropy minimizer at the point

$$\rho_p = \begin{pmatrix} p & \sqrt{p(1-p)} \\ \sqrt{p(1-p)} & 1-p \end{pmatrix}. \tag{4.21}$$

Such an assumption is not restrictive since one can always treat the operation Φ_1 as a concatenation of a given operation with a unitary rotation which does not change the minimal output entropy. The quantum operation (4.20) acting on a pure state

$$\rho_{in} = \begin{pmatrix} a & \sqrt{a(1-a)} \\ \sqrt{a(1-a)} & 1-a \end{pmatrix} \tag{4.22}$$

gives an output state

$$\begin{aligned} \rho_{out} \;=\; & a \begin{pmatrix} \Phi_{11} & \Phi_{21} \\ \overline{\Phi_{21}} & 1-\Phi_{11} \end{pmatrix} + (1-a) \begin{pmatrix} \Phi_{14} & \Phi_{24} \\ \overline{\Phi_{24}} & 1-\Phi_{14} \end{pmatrix} + \\ & + \sqrt{a(1-a)} \begin{pmatrix} 2\Re(\Phi_{12}e^{i\phi}) & \Phi_{22}e^{i\phi} + \overline{\Phi_{32}}e^{-i\phi} \\ \overline{\Phi_{22}}e^{-i\phi} + \Phi_{32}e^{i\phi} & -2\Re(\Phi_{12}e^{i\phi}) \end{pmatrix}, \end{aligned} \tag{4.23}$$

which attains the minimum entropy if $a = p$.

- Transformation changing the lengths of the axes of the ellipsoid.

  Consider a quantum operation $\Phi_{ellipsoid}$, which transforms the Bloch ball into such an ellipsoid that the end of its longest axis touches the Bloch sphere in the "North Pole",

$$\Phi_{ellipsoid} = \begin{pmatrix} 1 & 0 & 0 & 1-\eta 3 \\ 0 & \frac{\eta_1+\eta_2}{2} & \frac{\eta_1-\eta_2}{2} & 0 \\ 0 & \frac{\eta_1-\eta_2}{2} & \frac{\eta_1+\eta_2}{2} & 0 \\ 0 & 0 & 0 & \eta_3 \end{pmatrix}. \tag{4.24}$$

  Suitable rotations of the Bloch ball before and after the action of $\Phi_{ellipsoid}$ guarantees that the point of contact with the Bloch sphere is the minimizer of Φ_1. Therefore the concatenation of $\Phi_1 \cdot \Phi_{rotation} \cdot \Phi_{ellipsoid} \cdot \Phi_{rotation}$ has the same minimal output entropy and the same minimizer that Φ_1. The rotation operation is given by

$$\Phi_{rotation} = \begin{pmatrix} p & -\sqrt{(1-p)p} & -\sqrt{(1-p)p} & 1-p \\ \sqrt{(1-p)p} & p & p-1 & -\sqrt{(1-p)p} \\ \sqrt{(1-p)p} & p-1 & p & -\sqrt{(1-p)p} \\ 1-p & \sqrt{(1-p)p} & \sqrt{(1-p)p} & p \end{pmatrix}, \tag{4.25}$$

  where p is defined by the minimizer of output entropy for Φ_1. This transformation changes the lengths of axes of the ellipsoid but it does not change the point at the ellipsoid which is the closest to the Bloch sphere. In other words, this transformation does not change the directions of the axes of the image of Φ_1 into the Bloch ball, but only their lengths.

- Transformation changing directions of the axis.

  The next transformation changes directions of axes of an ellipsoid but preserves the entropy minimizer. In particular, if the image of the minimizer is on the longest axis of an ellipsoid, after the transformation the point which is the closest to the Bloch sphere is no longer on the main axis of the ellipsoid.

  Entropy of an output state (4.23) is a function of its determinant. The minimum of the determinant determines the minimum of the entropy. Consider a transformation which preserves the value of the determinant and compute its derivative in a minimizer. It is useful to introduce the compact notation of Eq. (4.23):

$$\rho_{out} = aA + (1-a)B + \sqrt{a(1-a)}C, \tag{4.26}$$

  where matrices A, B and C correspond to the matrices (4.23). Consider a transformation $\Phi_1 \to \Phi_1 + \Phi_{direction}$. The output of $\Phi_1 + \Phi_{direction}$ is

given by

$$\rho' = a(A + \frac{1}{2}\frac{\sqrt{1-p}}{\sqrt{p}}X) + (1-a)(B + \frac{1}{2}\frac{\sqrt{p}}{\sqrt{1-p}}X) + \sqrt{a(1-a)}(C - X), \tag{4.27}$$

where X is a matrix, which is hermitian and has trace equal to zero. Moreover, the matrix X satisfies the condition guaranteeing that $\Phi_1 + \Phi_{direction}$ is completely positive. The state ρ' coincides with (4.26) if $a = p$. Moreover, the derivative of formula (4.26) with respect to a is the same as the derivative of Eq. (4.27) at the point $a = p$. Therefore, the determinants of (4.26) and (4.27) are the same and the derivative at $a = p$ is equal to zero. A proper choice of parameters in X guarantees that there is a minimum at point $a = p$. Hence both maps, Φ_1 and $\Phi_1 + \Phi_{direction}$ have the same minimal output entropy.

The part $\Phi_{direction}$ can be characterized by two parameters (t, n),

$$\Phi_{direction} = \frac{1}{2}\begin{pmatrix} \sqrt{\frac{1-p}{p}}\, t & -t & -t & \sqrt{\frac{p}{1-p}}\, t \\ i\sqrt{\frac{1-p}{p}}\, n & -i\, n & -i\, n & i\sqrt{\frac{p}{1-p}}\, n \\ -i\sqrt{\frac{1-p}{p}}\, n & i\, n & i\, n & -i\sqrt{\frac{p}{1-p}}\, n \\ -\sqrt{\frac{1-p}{p}}\, t & t & t & -\sqrt{\frac{p}{1-p}}\, t \end{pmatrix}. \tag{4.28}$$

Such a form guarantees that the output state of $\Phi_1 + \Phi_{direction}$ is given by Eq. (4.27).

The map Φ_2 of the same minimal output entropy as Φ_1 obtained by joint action of three transformations, $\Phi_{rotation}$, $\Phi_{ellipsoid}$ and $\Phi_{direction}$, on Φ_1 can be given by:

$$\Phi_2 = \Phi_1 \Phi_{rotation} \cdot \Phi_{ellipsoid} \cdot \Phi_{rotation}^T + \Phi_{direction}. \tag{4.29}$$

We are not able to prove that this transformation contains all possibilities of obtaining maps with the same minimal output entropy as a given one, however, the transformation is characterized by 5 parameters and also 5 parameters are needed to have all different (up to one rotation) ellipsoids tangent to the sphere on its inner side in a given point. Three parameters are associated with the lengths of axes $|\eta_1|, |\eta_2|, |\eta_3|$, while two parameters define the direction of the longest axis n, t.

Above considerations introduce a 5-parameter transformation of a quantum map $\Phi_1 \to \Phi_2$. The transformation preserves the minimal output entropy. Therefore, it determines the family of maps which are situated at the same horizontal line of the plot $(S^{\max}, S^{\min})$. Characterization of the family of quantum maps parametrized by the minimal output entropy can be useful to further investigations of relations between $S^{\min}$ and $S^{\max}$ and their consequences.

Chapter 5

Davies maps for qubits and qutrits

Explicit description of general continuous dynamics of an open quantum system is difficult in practice. Exact formulas describing the time evolution are known in some special cases only. One of the cases in which the problem can be solved uses the assumption of a week coupling [91] of a low dimensional quantum system interacting with much bigger reservoir in the thermal equilibrium. Such an interaction changes only the state of the system whereas the state of the environment remains unchanged. By analogy to the classical process, in which the evolution of a state does not depend on the history, such an evolution is called a *Markov process*.

However, while analysing the continuous evolution of the input state, sometimes there is no need to know the entire time evolution since only the output state is relevant. The "black box" description is useful in such cases. A "black box" acts like an evolution discrete in time and can be described using completely positive maps, represented as matrices of superoperators.

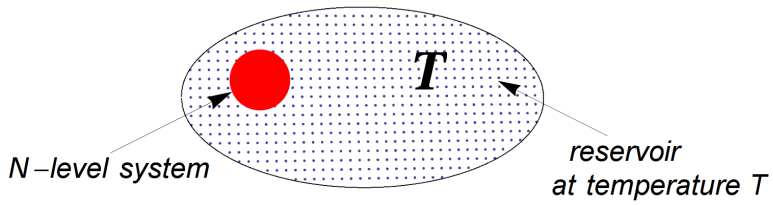


Figure 5.1: Model of a quantum N–level system characterized by Hamiltonian H interacting with a much larger environment in a thermal equilibrium at temperature T.

The following chapter distinguishes a concrete class of physical processes described by a *Davies map* [92]. Such a process is compatible with the interaction

of a quantum state with an environment in a given temperature, see Fig. 5.1. Due to a suitable choice of the entries of a superoperator matrix Φ and relations between them one can say whether some continuous time evolution is described by a given discrete quantum map. The solution concerns the maps acting on one–qubit, $N = 2$, and one–qutrit, $N = 3$. In the case of one-qubit maps we determine the state which is the most resistant on Davies channels. It will be shown that the maximal output 2–norm of Davies maps is additive with respect to the tensor product of two such maps.

5.1 Quantum Markov process

The quantum Markov process is characterized by quantum maps belonging to the one-parameter completely positive semigroup, $\Phi_t = \exp \mathcal{G}t$, where $\mathcal{G}$ denotes a generator and positive parameter t is associated with time.

The most general form of the generator of a completely positive semigroup was given by Gorini, Kossakowski, Sudarshan [93] and Lindblad [43]. It can be written as

$$\mathcal{G} = i\delta + \mathcal{L}, \tag{5.1}$$

where δ, given by the commutator with the effective Hamiltonian of the system $\delta : \rho \to [\rho, H]$, describes the unitary part of the evolution. The dissipative part $\mathcal{L}$ has the Lindblad form

$$\mathcal{L} : \rho \to \sum_{\alpha} \Big(K^{\alpha} \rho K^{\alpha\dagger} - \frac{1}{2} \{ K^{\alpha\dagger} K^{\alpha}, \rho \} \Big), \tag{5.2}$$

where $\{A, B\} = AB + BA$ is anticommutator, while operators K^{α} can be associated with the Kraus representation of the quantum operation.

Deciding whether a given superoperator matrix belongs to the completely positive semigroup was shown [94] to be a problem 'NP' hard with respect to the dimension N. Nevertheless, some additional assumptions allow one to characterize matrices from completely positive semigroups at least for a few low dimensions. In following chapter, such a solution will be given for $N = 2$, and $N = 3$, under additional conditions: independence of unitary and dissipative parts of the evolution and the detailed balance condition. These three conditions define the so–called Davies maps [92]. Sometimes the uniqueness of the invariant state is also added to the definition.

5.2 Characterization of the model

Consider a quantum N - level system characterized by the Hamiltonian in its eigenbasis,

$$H = \sum_{i=1}^{N} \epsilon_i |i\rangle\langle i|. \tag{5.3}$$

Assume that such a system is weekly coupled to the environment of a given temperature T, see Fig. 5.1. An interaction with the environment preserves one invariant state, which is the Gibbs state

$$\rho_\beta = \frac{1}{\mathcal{Z}} \exp(-\beta H), \tag{5.4}$$

where $\mathcal{Z} = \sum_{i=1}^{N} \exp(-\beta \epsilon_i)$ is a partition function and $\beta = \frac{1}{kT}$. Here k represents the Boltzmann constant. A quantum map Φ satisfies the *detailed balance* condition if it is Hermitian with respect to the scalar product defined by the Gibbs state

$$\operatorname{Tr} \rho_\beta A \Phi^*(B) = \operatorname{Tr} \rho_\beta \Phi^*(A) B, \tag{5.5}$$

where A and B are arbitrary observables and Φ^* the quantum operation in the Heisenberg picture. Detailed description of this condition can be found in [96].

The name "detailed balance" was taken from the theory of stochastic processes. Detailed balance means that in an equilibrium state any two levels of the evolving system remain in an equilibrium: the rate of transition from the level i to j and the transition rate from j to i are equal. Mathematical formula describing this fact reads

$$\mathcal{F}_{ij} p_i = \mathcal{F}_{ji} p_j, \tag{5.6}$$

where $\mathcal{F}_{ij}$ are entries of a stochastic transition matrix and p_i represent the components of the invariant probability vector.

5.3 Matrix representation of Davies maps

One qubit map in the "black box" description is represented by a superoperator matrix. It is a matrix acting on the vector formed by the entries of a density matrix ordered in a single column. A superoperator Φ represents a Davies map, if the following conditions are satisfied.

- The map Φ is completely positive.

  This condition is guaranteed if the Choi–Jamiołkowski matrix D_Φ (1.25) of the map is positive. One has to reshuffle the elements of the matrix Φ according to (1.28) and check positivity of the resulting dynamical matrix $D_\Phi = \Phi^R$.

- Superoperator Φ belongs to the semigroup of completely positive maps.

  This is equivalent to existence of a generator $\mathcal{G}$ of the Lindblad form (5.2) and the parameter $t \geq 0$ such that $\mathcal{G}t = \log \Phi$. Knowing the logarithm of Φ one has to determine whether it is of the Lindblad form. It was shown in [95] that if the Choi-Jamiołkowski matrix of a given generator is positive in the subspace orthogonal to the maximally entangled state, then the generator can be written in the Lindblad form.

  It is not a trivial task to write an analytical expression for the logarithm of a given matrix if its dimension is greater than two. Such a problem for

3×3 stochastic matrices is discussed in the last section of the following chapter.

- Since the rotational part of the evolution is independent of the dissipative (contractive) part, the structure of the superoperator is restricted to the block diagonal form. Off–diagonal elements of the density matrix are just multiplied by numbers, while the diagonal elements can be mixed between themselves. More detailed discussion on this property is given in Section 5.7.

- The detailed balance condition introduces further restrictions on the elements of the block acting on the diagonal part of the density matrix. This block is a stochastic matrix, the entries of which satisfy Eq. (5.6).

Since now, only the dissipative part of the evolution will be considered. Due to the above conditions the dissipative part of the generator of the one–qubit Davies maps can be written as

$$\mathcal{L}_{\alpha,\lambda,p} = \begin{pmatrix} -\alpha & 0 & 0 & \alpha\frac{p}{1-p} \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ \alpha & 0 & 0 & -\alpha\frac{p}{1-p} \end{pmatrix}, \tag{5.7}$$

while the corresponding superoperator acting on two-dimensional states (in the Hamiltonian basis) has the form

$$\Phi_{a,c,p} = \begin{pmatrix} 1-a & 0 & 0 & a\frac{p}{1-p} \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ a & 0 & 0 & 1-a\frac{p}{1-p} \end{pmatrix}. \tag{5.8}$$

Here, p is a function of temperature, $p = \left(1+\exp\left(-\frac{\epsilon}{(kT}\right)\right)^{-1}$, which determines the invariant state

$$\Phi_{a,c,p}(\rho_*) = \rho_* = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}. \tag{5.9}$$

Notice that (5.8) has a block diagonal form which is a consequence of independence of rotational and contractive evolution. This is also equivalent to independence of changes in diagonal and off–diagonal entries of a density matrix. The detailed balance condition (5.6) implies the form of the outer block in Eq. (5.8). One–qubit Davies maps form a three-parameter family characterized by (a, c, p), where p is a function of the temperature. Conditions that such a matrix is an element of the semigroup of completely positive maps introduce the following restrictions on the parameters (a, c, p):

$$a + p < 1, \qquad 0 < c < \sqrt{1 - \frac{a}{1-p}}. \tag{5.10}$$

Equality $\Phi = \exp \mathcal{L}t$ allows one to write explicit formulas for time dependence of parameters a and c,

$$a = (1-p)\Big(1 - \exp(-At)\Big), \qquad c = \exp(-\Gamma t), \tag{5.11}$$

where A and Γ are parameters such that $A \geq \frac{1}{2}\Gamma \geq 0$. The entire paths of the semigroup are showed in Fig. 5.2

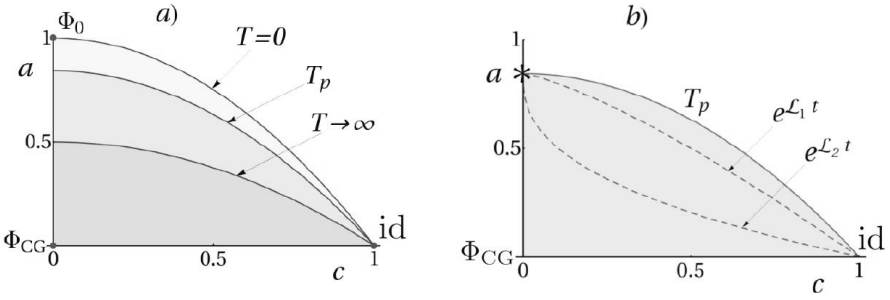


Figure 5.2: Panel a) contains the region of parameters (a, c) allowed by relation (5.10) and describing the one–qubit Davies maps. The upper border lines are also drawn for different temperature T. Panel b) shows the region allowed for a given temperature T. The lines describe two semigroup corresponding to two different randomly chosen generators $\mathcal{L}_1$ and $\mathcal{L}_2$. The extremal lines corresponding to the solid line on panel a describe the semigroup with the smallest ratio of decoherence to the damping rate. Maps $\mathrm{id}, \Phi_0, \Phi_{CG}$ are the identity channel, completely depolarizing and coarse graining channel respectively.

One–qubit Davies map can be written using the Bloch parametrization (1.42)

$$\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \eta_1 & 0 & 0 \\ 0 & 0 & \eta_1 & 0 \\ \kappa_3 & 0 & 0 & \eta_3 \end{pmatrix}. \tag{5.12}$$

where $|\eta_i|$ denote the lengths of axes of the ellipsoid and $\vec{\kappa}$ is the translation vector. These parameters are related to the parameters (a, c, p)

$$\begin{aligned} \eta_1 = c \geq 0, \qquad & \eta_3 = 1 - \frac{a}{1-p} \geq 0, \\ \kappa_1 = \kappa_2 = 0, \qquad & \kappa_3 = a\frac{2p-1}{1-p} \geq 0. \end{aligned} \tag{5.13}$$

The image of the set of pure states under an action of one–qubit Davies map is shown in Fig. 5.3. The image of the Bloch ball forms an ellipsoid with rotational symmetry. Fig. 5.3 presents the image of an exemplary one–qubit Davies map for which $\eta_1 \geq \eta_3$, however conditions (5.10) admits also the case $\eta_3 \geq \eta_1$.

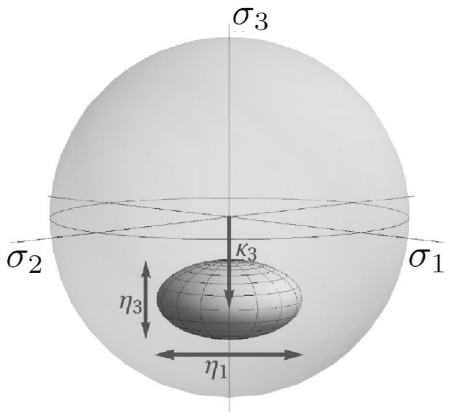


Figure 5.3: Ellipsoid obtained by an action of a one–qubit Davies channel on the Bloch sphere. The channel is characterized by parameters $(\eta_1, \eta_3, \kappa_3)$ defined in (5.13).

5.4 Physical examples

Qubit maps of the structure similar to (5.12) were analysed before in context of quantum optics. The unitary evolution is induced by the laser field, while the dissipative dynamics is caused by an interaction with the environment. The state of a two level atom is characterized by the Bloch vector (x, y, z), where z represents the difference between the diagonal entries of a density matrix equal to the inversion of populations of the atomic levels. Variables x and y are associated with the atomic dipole operators. The evolution in this set has been defined by means of variables describing the decay rate τ_1 of the coherences and the rate τ_2 of attaining the equilibrium state. These parameters correspond to the variables considered in the Section 5.3, $\eta_1 = \exp(-t/\tau_1)$ and $\eta_3 = \exp(-t/\tau_3)$ which are related to squeezing of the axes of the ellipsoid. Formula (5.10) corresponds to the relation between the decay rates:

$$\tau_1 \leq 2\tau_3. \tag{5.14}$$

This relation was obtained by analysing a concrete physical model of the evolution of the two level system by means of Bloch equations [97]. The one–qubit operations (5.12) were also studied by [98].

5.5 Minimal output entropy of Davies maps

In context of transmission of quantum information, it is natural to ask, which pure states are the most resistant with respect to the changes caused by the Davies maps. The answer depends on the selected measure of decoherence. Such a measure can be described, for example, by means of some matrix norm of the

output state maximized over the input states. Among quantities measuring the decoherence, the minimal output entropy is of special importance because some questions concerning the channel capacity, such as additivity problem, can be related with similar problem written in terms of minimal output entropy. The minimal output entropy is related to the maximal norm of the output state if the input is pure.

Since a Davies map has rotational symmetry, the minimizer can be chosen to be a real state:

$$\rho = \begin{pmatrix} \mu & \nu \\ \nu & 1-\mu \end{pmatrix}, \tag{5.15}$$

where $\nu^2 = (1-\mu)\mu$ since the state is pure. After an action of the operation (5.8) this state is transformed into

$$\rho' = \begin{pmatrix} (1-a)\mu + b(1-\mu) & c\nu \\ c\nu & a\mu + (1-b)(1-\mu) \end{pmatrix}, \tag{5.16}$$

where $b = ap/(1-p)$. Computing the eigenvalues and minimizing the entropy over μ one can characterize the minimizer in two cases:

- If $c^2 \leq (1-a-b)(1-2b)$ the minimizer is characterized by $\mu = 0$ and it forms an eigenstate of the Hamiltonian H.
- If $c^2 \geq (1-a-b)(1-2b)$ the minimizer is characterized by

$$\mu = \frac{(a+b-1)(2b-1) - c^2}{2(a+b-1)^2 - 2c^2}. \tag{5.17}$$

  It is no longer the eigenvalue of the Hamiltonian, however, after some time of the evolution $t \gg 0$ the second case changes into the first one and the minimizer is a state $\mathrm{diag}(0,1)$. This is an eigenstate of the Hamiltonian. The situation that the minimizer is in the vector $\mathrm{diag}(0,1)$ reminds the classical evolution of two–dimensional vector governed by the stochastic matrix. In this case the extremal vector like $(0,1)$ is the minimizer of the Shannon entropy of the output.

5.6 Multiplicativity of maximal output norm of one–qubit Davies maps

As discussed in the introduction to Chapter 4, the question of additivity of minimal output von Neumann entropy with respect to the tensor product of quantum operations is one of the most interesting problem in quantum information theory. This problem can be equivalently stated in terms of channel capacity. In general, the conjecture on additivity of channel capacity is false, however, there is still an interesting problem, for which class of maps the conjecture can be confirmed.

Recent studies of the problem use the notion of the Rényi entropy of order q. This entropy tends to the von Neumann version as $q \to 1$. The problem of additivity of minimal output Rényi q entropy is directly related to multiplicativity of the maximal output Schatten q–norm. This norm is defined as

$$\|X\|_q^S = (\mathrm{Tr}\,|X|^q)^{1/q}, \tag{5.18}$$

where $|X| = \sqrt{X^\dagger X}$. Maximal Schatten q norm of a quantum map Φ is:

$$\|\Phi\|_q^S := \max_\rho (\mathrm{Tr}\,|\Phi(\rho)|^q)^{1/q}, \tag{5.19}$$

where maximization is taken over the entire set of density matrices ρ. The Rényi entropy of order q of a state ρ can be defined as follows [107]

$$S_q(\rho) = \frac{q}{1-q} \log \|\rho\|_q^S. \tag{5.20}$$

Due to logarithm in this formula the multiplicativity of maximal q–norm is equivalent to the additivity of minimal output entropy $S_q^{\min}$.

In this section, multiplicativity of operator 2–norm induced by the Euclidean vector norm will be proved for the quantum one–qubit Davies maps. This vector induced norm is not related to the Rényi entropy by such an elegant formula like it is in the case for Schatten norm, however, it is a bit easier to calculate than the Schatten counterpart. These particular results support the general solution for multiplicativity problem for Schatten 2–norm which implies the additivity property for minimal output Rényi entropy of order 2 and which has been already proved for general one–qubit quantum operations [99] (see also [18]).

5.6.1 Outline of the proof of multiplicativity

The Euclidean norm (2–norm) of a vector $x = (x_1, ..., x_n)$ is defined as:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}. \tag{5.21}$$

This vector norm induces the 2–norm of an operator A:

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}. \tag{5.22}$$

One of the property of this norm (see [100]) is that $\|A\|_2$ is equal to square root of the spectral radius of $A^\dagger A$ or equivalently to the greatest singular value of the matrix A,

$$\|A\|_2 = \left[r(A^\dagger A)\right]^{1/2}, \tag{5.23}$$

where a spectral radius $r(A^\dagger A) = \max_i |\xi_i|$ and ξ_i are eigenvalues of $A^\dagger A$. In this section the maximal two norm of the output of a quantum map $\Phi : \mathcal{M}_N \to \mathcal{M}_N$ will be considered

$$M_\Phi = \max_{\rho \in \mathcal{M}_N} \|\Phi(\rho)\|_2 = \max_{A \geq 0} \frac{\|\Phi(A)\|_2}{\operatorname{Tr} A}. \tag{5.24}$$

One can ask, whether the maximal two–norm is multiplicative in a sense:

$$M_{\Phi \otimes \Omega} = M_\Phi M_\Omega. \tag{5.25}$$

It will be shown that if Φ is one–qubit Davies map and Ω is an arbitrary quantum map acting on N–dimensional state the multiplicativity holds.

The idea of the proof of the theorem given below is borrowed from the paper of King and Ruskai [54]. These authors prove an analogical theorem about a bistochastic quantum map Φ. They noted that the same proof holds as well for stochastic one–qubit maps. Here we will present an explicit calculations for the case of Davies maps with $|\eta_3| \leq |\eta_1|$.

Theorem 8. *Let $\Phi : \mathcal{M}_2 \to \mathcal{M}_2$ be an one–qubit Davies map and $\Omega : \mathcal{M}_N \to \mathcal{M}_N$ be an arbitrary quantum map. The maximal two norm of the output is multiplicative:*

$$M_{\Phi \otimes \Omega} = M_\Phi M_\Omega. \tag{5.26}$$

In this section the sketch of the proof will be given, while some details of the calculation will be presented in the next section. In order to present the proof we need to introduce the following set. An arbitrary density matrix on $\mathcal{H}_2 \otimes \mathcal{H}_N$ can be written as a block matrix

$$\rho = \begin{pmatrix} \rho_1 & \gamma \\ \gamma^\dagger & \rho_2 \end{pmatrix}, \tag{5.27}$$

where $\rho_1,\ \rho_2,\ \gamma$ are $N \times N$ matrices and the trace condition $\operatorname{Tr}(\rho_1 + \rho_2) = 1$ is satisfied. The output state of the product of two quantum operations $\Phi \otimes \Omega$, can be described by:

$$\big(\Phi \otimes \Omega\big)(\rho) = \begin{pmatrix} P & L \\ L^\dagger & Q \end{pmatrix}. \tag{5.28}$$

Here Φ denotes an one–qubit operation, while the map Ω acts on $\mathcal{M}_N$. Also other block matrices will occur and their positivity will play an important role during the proof. The Schur complement lemma [101] ensures positivity of block matrices, see Lemma 1, Section 3.1.

To demonstrate additivity (5.26) we shall analyse the inequality $M_{\Phi \otimes \Omega} \geq M_\Phi M_\Omega$ which is almost immediate since the equality is attained by a product of states which maximize output norm of each map. Because the entire set of states

is larger, it contains product and entangled states, the result of maximizing over the entire set can give only a better result. Therefore to prove multiplicativity of maximal output 2–norm with respect to the tensor product of two maps it is enough to show that

$$z \geq M_\Phi M_\Omega \Rightarrow z\mathbb{1} - \big(\Phi \otimes \Omega\big)(\rho) \geq 0. \tag{5.29}$$

Insert the block matrix form (5.28) to (5.29). Due to the Schur complement lemma the right hand side of (5.29) is positive if and only if

$$L(z\mathbb{1} - P)^{-1}L^\dagger \leq \ z\mathbb{1} - Q. \tag{5.30}$$

Notice that this inequality holds if

$$\big\|LL^\dagger\big\|_2 \leq (z - \|P\|_2)(z - \|Q\|_2), \tag{5.31}$$

since using the general property $P \leq \|P\|$ one gets:

$$\begin{aligned} L(z1 - P)^{-1}L^\dagger \ &\leq \ L(z - \|P\|_2)^{-1}L^\dagger \ \leq \ \big\|LL^\dagger\big\|_2 (z - \|P\|_2)^{-1} && (5.32)\\ &\leq \ (z - \|Q\|_2) \ \leq \ z1 - Q. && (5.33) \end{aligned}$$

Therefore the positivity of $(z - \|P\|_2)$ and $(z - \|Q\|_2)$ and inequality (5.31) are the only relations needed to prove Theorem 8. These relations will be proved in the next section for the case of Φ being an arbitrary one–qubit Davies map with $|\eta_3| \leq |\eta_1|$.

5.6.2 Details of the proof of multiplicativity

Proof. of Theorem 8. It is necessary to find the specific form of M_Φ, P an Q in (5.28), then to check positivity of $(M_\Phi M_\Omega - \|P\|_2)$ and $(M_\Phi M_\Omega - \|Q\|_2)$, and finally to prove (5.31). Let us restrict our considerations to the case of Davies maps Φ, for which $\eta_3^2 \leq \eta_1^2$ in (5.12) as discussed in Section 5.6.1.

- Maximal 2–norm of the output, M_Φ.

  Use the Bloch parametrization of Φ as in (5.12), let it act on the Bloch vector $(1, x, y, z)^\dagger$, where x, y, z are real. Moreover $x^2 + y^2 + z^2 = 1$ guarantees restriction to pure states. It is enough to take pure input state because the 2–norm is convex on the set of density matrices and it attains maximum at the boundary of the set. The spectral radius of the square of the output state reads according to (5.23):

  $$\sqrt{[r(\Phi(\rho)^\dagger\Phi(\rho))]} = \frac{1}{2}\big(1 + \sqrt{(\kappa_3 + z\eta_3)^2 + (1 - z^2)\eta_1^2}\big). \tag{5.34}$$

  Since the image of the Davies map has rotational symmetry, there are no parameters x and y in this formula. Second derivative of the function (5.34) with respect to z is negative under the condition: $\eta_3^2 \ \leq \ \eta_1^2$. Therefore function (5.34) has a maximum:

  $$M_\Phi = \frac{1}{2}\left(1 + \sqrt{\eta_1^2 + \frac{\kappa_3^2\eta_1^2}{\eta_1^2 - \eta_3^2}}\right). \tag{5.35}$$

- Output of a product map.

  Now the explicit form of matrices P, Q and L of the output state (5.28) will be given. Consider an one–qubit input state. A vector $(1, x, y, z)^\dagger$ corresponds to the density matrix:

$$\rho = \frac{1}{2} \begin{pmatrix} 1+z & x+iy \\ x-iy & 1-z \end{pmatrix}. \tag{5.36}$$

  Its image with respect to a Davies map (5.12) reads:

$$\Phi(\rho) = \frac{1}{2} \begin{pmatrix} 1+z\eta_3+\kappa_3 & \eta_1(x+iy) \\ \eta_1(x-iy) & 1-z\eta_3-\kappa_3 \end{pmatrix}. \tag{5.37}$$

  In the analogous way the initial state in a space $\mathcal{M}_{2N}$ can be given by (5.27)

$$\rho = \frac{1}{2} \begin{pmatrix} \rho_1+\rho_2+\hat{z} & \hat{x}+i\hat{y} \\ \hat{x}-i\hat{y} & \rho_1+\rho_2-\hat{z} \end{pmatrix}, \tag{5.38}$$

  where $\hat{z} = \rho_1 - \rho_2$ and $\hat{x} - i\hat{y} = 2\gamma$ are $N \times N$ matrices. The output state of a map $\Phi \otimes \mathbb{1}$ is:

$$(\Phi \otimes \mathrm{id})(\rho) = \tag{5.39}$$

$$\begin{pmatrix} \frac{1}{2}\big(\rho_1+\rho_2+\eta_3(\rho_1-\rho_2)+\kappa_1(\rho_1+\rho_2)\big) & \eta_1\gamma \\ \eta_1\gamma^\dagger & \frac{1}{2}\big(\rho_1+\rho_2-\eta_3(\rho_1-\rho_2)-\kappa_3(\rho_1+\rho_2)\big) \end{pmatrix}.$$

  Finally the matrices P, Q and L are defined by comparison of suitable blocks of two block matrices:

$$(\Phi \otimes \Omega)(\rho) = \begin{pmatrix} P & L \\ L^\dagger & Q \end{pmatrix} \tag{5.40}$$

$$= \begin{pmatrix} \frac{1}{2}\Omega\big(\rho_1+\rho_2+\eta_3(\rho_1-\rho_2)+\kappa_3(\rho_1+\rho_2)\big) & \eta_1\Omega(\gamma) \\ \eta_1\Omega(\gamma)^\dagger & \frac{1}{2}\Omega\big(\rho_1+\rho_2-\eta_3(\rho_1-\rho_2)-\kappa_3(\rho_1+\rho_2)\big) \end{pmatrix}.$$

- Multiplicativity.

  One can use the property $\|\Omega(\rho)\|_2 \leq \mathrm{Tr}(\rho) M_\Omega$ (5.24) to show that $(M_\Phi M_\Omega - \|P\|_2)$ is positive. It is so if

$$\frac{1}{2} M_\Omega \Big(1 + \sqrt{\eta_1^2 + \frac{\kappa_3^2 \eta_1^2}{\eta_1^2 - \eta_3^2}}\Big) > \frac{1}{2} M_\Omega \Big(\mathrm{Tr}(\rho_1+\rho_2) + \eta_3(\rho_1-\rho_2) + \kappa_3(\rho_1+\rho_2)\Big). \tag{5.41}$$

Notice that $\mathrm{Tr}(\rho_1+\rho_2)=1$. To show that the above inequality is true, it is sufficient to prove:

$$\sqrt{\eta_1^2+\frac{t^2\eta_1^2}{\eta_1^2-\eta_3^2}}>\eta_3+\kappa_3. \tag{5.42}$$

Taking the square of both sides one gets the expression:

$$\big(\kappa_3\eta_3-(\eta_1^2-\eta_3^2)\big)^2>0. \tag{5.43}$$

This implies that $(M_\Phi M_\Omega-\|P\|_2)>0$. In a similar way we prove the positivity of $(M_\Phi M_\Omega-\|Q\|_2)$. The last step is to prove inequality (5.31).

Consider a positive block matrix $\big(\mathbb{1}\otimes\Omega\big)(\rho)=\begin{pmatrix}\Omega(\rho_1) & \Omega(\gamma)\\ \Omega(\gamma^\dagger) & \Omega(\rho_2)\end{pmatrix}$. Assume that $\Omega(\rho_1)>0$ (if $\Omega(\rho_1)\geq 0$ one can add $\epsilon\mathbb{1}$ to ρ_1 and eventually take the limit $\epsilon\to 0$). Due to the inequality $\Omega(\rho_1)\leq\|\Omega(\rho_1)\|_2$ one can write

$$\langle v|\Omega(\gamma)\Omega(\gamma)^\dagger|v\rangle\leq\|\Omega(\rho_1)\|_2\langle v|\Omega(\gamma)\Omega(\rho_1)^{-1}\Omega(\gamma)^\dagger|v\rangle. \tag{5.44}$$

Due to the Schur complement lemma we have $\Omega(\rho_2)\geq\Omega(\gamma)\Omega(\rho_1)^{-1}\Omega(\gamma)^\dagger$ and therefore,

$$\|\Omega(\rho_1)\|_2\langle v|\Omega(\gamma)\Omega(\rho_1)^{-1}\Omega(\gamma)^\dagger|v\rangle\leq\|\Omega(\rho_1)\|_2\langle v|\Omega(\rho_2)|v\rangle\leq\|\Omega(\rho_1)\|_2\|\Omega(\rho_2)\|_2. \tag{5.45}$$

Hence the inequality $\|\Omega(\gamma)\Omega(\gamma)^\dagger\|_2\leq\|\Omega(\rho_1)\|_2\|\Omega(\rho_2)\|_2$ holds. This inequality together with definition (5.24) implies

$$\|\Omega(\gamma)\Omega(\gamma)^\dagger\|_2\leq M_\Omega^2\,\mathrm{Tr}\,\rho_1\,\mathrm{Tr}\,\rho_2. \tag{5.46}$$

Denote $\mathrm{Tr}(\rho_1)$ by x. To prove inequality (5.31) it is enough to show that the second inequality holds in the expression below

$$\big\|LL^\dagger\big\|_2=\eta_1^2\,\|\Omega(\gamma)\|_2^2\leq\eta_1^2x(1-x)M_\Omega^2\leq(M_\Phi M_\Omega-\|P\|_2)(M_\Phi M_\Omega-\|Q\|_2), \tag{5.47}$$

and this is true if

$$\eta_1^2x(1-x)M_\Omega^2\leq\frac{1}{4}M_\Omega^2\Big[\eta_1^2+\frac{\kappa_3^2\eta_1^2}{\eta_1^2-\eta_3^2}-(\eta_3(2x-1)+\kappa_3)^2\Big]. \tag{5.48}$$

This inequality can be shown by taking the function which is the difference between the right hand side and the left hand side. The second derivative of this function is equal to $2(\eta_1^2-\eta_3^2)$. Therefore whenever $(\eta_1^2>\eta_3^2)$ the difference is a convex function which has minimum at 0. That finishes the proof of the last inequality. Therefore inequality (5.31) holds and it proves Theorem 8.

□

In the case $|\eta_1| \leq |\eta_3|$ the proof goes analogously. The maximal output norm (5.35) has in this case a simpler form, since the maximizer is a pure state described by the Bloch vector $(x = 0, y = 0, z = 1)$. The specific form of the Davies map was used in this proof in (5.35) when the formula of the maximal output norm was computed and in formula (5.39). Moreover, positivity of κ_3 is used in (5.42).

5.7 Davies maps acting on qutrits

In this chapter a characterization of the Davies maps for qutrits, $N = 3$, will be given. Going to higher dimensions demands more abstract and systematic approach than in the case of one–qubit maps. The entire evolution consists of the unitary part and the dissipative part and such is the structure of the generator $\mathcal{G} = i\delta + \mathcal{L}$. The unitary evolution δ is governed by the Hamiltonian which in its eigenbasis has a form $H = \sum_{i=1}^{3} \epsilon_i |i\rangle\langle i|$, where $\epsilon_1 > \epsilon_2 > \epsilon_3$. Differences of energies $\{\omega_{ij} = \epsilon_i - \epsilon_j\}$ are called Bohr frequencies. They are eigenvalues of the unitary part of the evolution, δ given by $\rho \to [H, \rho]$, while the eigenvectors of δ are $|i\rangle\langle j|$ for $i, j = 1, 2, 3$. Assume that the set of Bohr frequencies is not degenerated beside the zero frequency case, ω_{ii}. The subspace related to the zero frequency is 3–dimensional. Since the dissipative part $\mathcal{L}$ of the evolution commutes with the unitary part, it has the same eigenvectors and therefore it does not couple the non-degenerated subspaces. Thus the off diagonal entries of a density matrix are not mixed with the diagonal ones, if the matrix is written in the eigenbasis of the Hamiltonian.

Like in the case of one–qubit maps only the dissipative part of the evolution will be analysed. An one–qutrit Davies map has a structure

$$\Phi = \begin{pmatrix} 1-\mathcal{F}_{21}-\mathcal{F}_{31} & 0 & 0 & 0 & \mathcal{F}_{12} & 0 & 0 & 0 & \mathcal{F}_{13} \\ 0 & \mu_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_1 & 0 & 0 & 0 & 0 & 0 \\ \mathcal{F}_{21} & 0 & 0 & 0 & 1-\mathcal{F}_{12}-\mathcal{F}_{32} & 0 & 0 & 0 & \mathcal{F}_{23} \\ 0 & 0 & 0 & 0 & 0 & \mu_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_3 & 0 \\ \mathcal{F}_{31} & 0 & 0 & 0 & \mathcal{F}_{32} & 0 & 0 & 0 & 1-\mathcal{F}_{13}-\mathcal{F}_{23} \end{pmatrix}, \tag{5.49}$$

where $\mathcal{F}_{21}, \mathcal{F}_{31}, \mathcal{F}_{32}$ and μ_1, μ_2, μ_3 parametrize the map. The off–diagonal elements are related by the detailed balance formula

$$\mathcal{F}_{ij} p_j = \mathcal{F}_{ji} p_i. \tag{5.50}$$

Here p_i determine the invariant Gibbs state (5.4). The Choi-Jamiołkowski ma-

trix of (5.49) preserves the same structure:

$$D_\Phi = \frac{1}{3}\begin{pmatrix} 1-\mathcal{F}_{31}-\mathcal{F}_{21} & 0 & 0 & 0 & \mu_1 & 0 & 0 & 0 & \mu_2 \\ 0 & \mathcal{F}_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathcal{F}_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathcal{F}_{21} & 0 & 0 & 0 & 0 & 0 \\ \mu_1 & 0 & 0 & 0 & 1-\mathcal{F}_{32}-\mathcal{F}_{21} & 0 & 0 & 0 & \mu_3 \\ 0 & 0 & 0 & 0 & 0 & \mathcal{F}_{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{F}_{31} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{F}_{32} & 0 \\ \mu_2 & 0 & 0 & 0 & \mu_3 & 0 & 0 & 0 & 1-\mathcal{F}_{32}-\mathcal{F}_{31} \end{pmatrix}. \tag{5.51}$$

The generator and its Choi-Jamiołkowski matrix have also the same structure.

Block of the superoperator Φ of the Davies quantum operation which is related to zero frequency space is a 3×3 stochastic matrix

$$\mathcal{F} = \begin{pmatrix} 1-\mathcal{F}_{31}-\mathcal{F}_{21} & \mathcal{F}_{12} & \mathcal{F}_{13} \\ \mathcal{F}_{21} & 1-\mathcal{F}_{21}-\mathcal{F}_{32} & \mathcal{F}_{23} \\ \mathcal{F}_{31} & \mathcal{F}_{32} & 1-\mathcal{F}_{13}-\mathcal{F}_{23} \end{pmatrix}, \tag{5.52}$$

where $\mathcal{F}_{32}, \mathcal{F}_{31}, \mathcal{F}_{21} \geq 0$. Due to the definition of the quantum detailed balance condition (5.5) the Davies map is Hermitian with respect to scalar product $\langle X, Y \rangle_\beta := \operatorname{Tr} \rho_\beta^{-1} X^\dagger Y$ and therefore it has a real spectrum. Moreover, the spectrum is positive, since there is real logarithm of the matrix Φ represented the Davies map. The positivity of the zero frequency block implies that

$$\begin{gathered} \mathcal{F}_{32} + \mathcal{F}_{31} + \mathcal{F}_{21} \leq 1, \\ 3 - 4(\mathcal{F}_{32} + \mathcal{F}_{31} + \mathcal{F}_{21}) + 3(\mathcal{F}_{32}\mathcal{F}_{31} + \mathcal{F}_{31}\mathcal{F}_{21} + \mathcal{F}_{21}\mathcal{F}_{32}) \geq 0. \end{gathered} \tag{5.53}$$

The question considered in this chapter concerns explicit analytical relations for entries of the superoperator (5.49), which imply that the superoperator represents a Davies map. One of the condition for Φ is that there exists an exponential form

$$\mathcal{F} = e^{Lt}. \tag{5.54}$$

Operator L is the zero frequency part of the contractive part of the generator of completely positive Davis semigroup. It is parameterized as follows:

$$L = \begin{pmatrix} -L_{21}-L_{31} & L_{12} & L_{13} \\ L_{21} & -L_{12}-L_{32} & L_{23} \\ L_{31} & L_{32} & -L_{13}-L_{23} \end{pmatrix}. \tag{5.55}$$

This is only the zero frequency block which satisfies the detailed balance condition. The entire dissipative part of the generator is represented by a 9×9 matrix. Its Choi matrix has on diagonal elements L_{32}, L_{31}, L_{21}. Since the Choi state of the generator has to be positive on the subspace perpendicular to the maximally entangled state we need to require that $L_{32}, L_{31}, L_{21} \geq 0$.

In the next section an explicit calculation of the logarithm of a stochastic matrix of order three (5.52) is presented.

5.7.1 Logarithm of a stochastic matrix of size three

To compute analytically the logarithm of a positive matrix (5.52) one may relay on the following construction. As matrix $\mathcal{F}$ has the eigenvalues $\{1, x+y, x-y\}$, where

$$\begin{array}{l} x = \frac{1}{2}(\mathrm{Tr}\,\mathcal{F} - 1), \\ y = \frac{1}{2}\sqrt{2\,\mathrm{Tr}\,\mathcal{F}^2 - (\mathrm{Tr}\,\mathcal{F})^2 + 2\,\mathrm{Tr}\,\mathcal{F} - 3}, \end{array} \tag{5.56}$$

the logarithm has the form

$$\log(\mathcal{F}) = \log\Big[U \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+y & 0 \\ 0 & 0 & x-y \end{pmatrix} U^{-1}\Big], \tag{5.57}$$

where U is a unitary matrix which transforms $\mathcal{F}$ into its diagonal form. Let us evaluate $\log(\mathcal{F})$ without computing the matrix U explicitly. One can write that

$$\log(\mathcal{F}) = \frac{1}{2}\Big[\log(x^2-y^2)Z^2 + \log(\frac{x+y}{x-y})Z\Big], \tag{5.58}$$

where

$$Z = U \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} U^{-1}. \tag{5.59}$$

The matrix $\mathcal{F}$ can be given in terms of Z:

$$\mathcal{F} = 1 - Z^2 + xZ^2 + yZ. \tag{5.60}$$

This relation allows one to compute $yZ = (\mathcal{F}-1) - (x-1)Z^2$. Formula for Z^2 can be calculated by taking the square of this equation and using the fact that $Z^4 = Z^2$ and that $Z^2(\mathcal{F}-1) = (\mathcal{F}-1)Z^2 = (\mathcal{F}-1)$. The last formula holds since operator $(\mathcal{F}-1)$ is defined in the subspace for which Z^2 is the identity,

$$Z^2 = \frac{(\mathcal{F}-1)\,[(\mathcal{F}-1) - 2(x-1)]}{y^2 - (x-1)^2}. \tag{5.61}$$

Therefore the logarithm of the matrix $\mathcal{F}$ can be expressed according to Eq. (5.58). By comparing a suitable entries of $\log(\mathcal{F})$ with the parameters of L, one gets the parameter L_{21} as a function of $(\mathcal{F}_{32}, \mathcal{F}_{31}, \mathcal{F}_{21})$,

$$L_{21} = y_2(1-x-y)\log(x-y) - (y_1(1-x+y)\log(x+y)), \tag{5.62}$$

where x and y are given by Eq. (5.56) and

$$y_1 := 2y - \mathcal{F}_{12} - \mathcal{F}_{21} + \mathcal{F}_{13} - \mathcal{F}_{31} + \mathcal{F}_{23} - \mathcal{F}_{32} + 2\frac{\mathcal{F}_{23}\mathcal{F}_{31}}{\mathcal{F}_{21}}, \tag{5.63}$$

$$y_2 := 4 - 2y - \mathcal{F}_{12} - \mathcal{F}_{21} + \mathcal{F}_{13} - \mathcal{F}_{31} + \mathcal{F}_{23} - \mathcal{F}_{32} + 2\frac{\mathcal{F}_{23}\mathcal{F}_{31}}{\mathcal{F}_{21}}. \tag{5.64}$$

The set of points $\{\mathcal{F}_{23}, \mathcal{F}_{13}, \mathcal{F}_{12}\}$, which defines the set of symmetric bistochastic matrices from the dynamical semigroup, is shown in Fig. 5.4 and Fig. 5.5 denoted by E. This set is inside the set of all bistochastic 3×3 matrices which is denoted by D. The boundaries of the set are stated by the constraints $L_{21}, L_{31}, L_{32} \geq 0$.

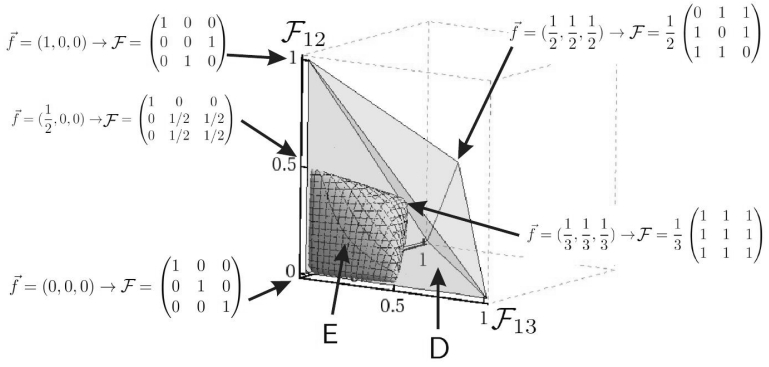


Figure 5.4: The set E of 3×3 bistochastic matrices $\mathcal{F}$ (5.52), which form the zero frequency block of the Davies channel Φ (5.49) under condition $T \to \infty$, is represented by the vector of the off–diagonal elements $\vec{f} = \{\mathcal{F}_{12}, \mathcal{F}_{13}, \mathcal{F}_{23}\}$. The set E is inside the set of all bistochastic 3×3 matrices D. The characteristic points are denoted by $\vec{f}$ and the corresponding matrix $\mathcal{F}$.

Expression (5.62) allows one to check that the set of stochastic matrices belonging to the semigroup of completely positive maps with the detailed balance condition is not convex. Consider two exemplary points which lie near the border of the cross-section and belong to the set ($L_{21} \geq 0$): $\{0.5, 0, 0\}$ and $\{0.22744, 0.22744, 0.04512\}$. Their convex combination does not belong to the set. Therefore the set of Davies map is not convex. Fig. 5.5 presents the cross-section of the set of bistochastic matrices which form the zero frequency part of the Davies map represented in the space of parameters $\mathcal{F}_{21}, \mathcal{F}_{31}, \mathcal{F}_{32}$. Fig. 5.5*b* plots a non–convex cross–section of the set E by the plane M.

In order to obtain a full characterization of the Davies map for qutrits, not only its zero frequency par have to be analysed. One needs to take into consideration also the complete positivity condition and the condition on the semigroup related to the Choi–Jamiołkowski matrices of the superoperator and its generator. These conditions allow us to specify the matrix entries μ_i from Eq. (5.49).

In this way the full characterization of the Davies channels for one–qubit and one–qutrits is provided.

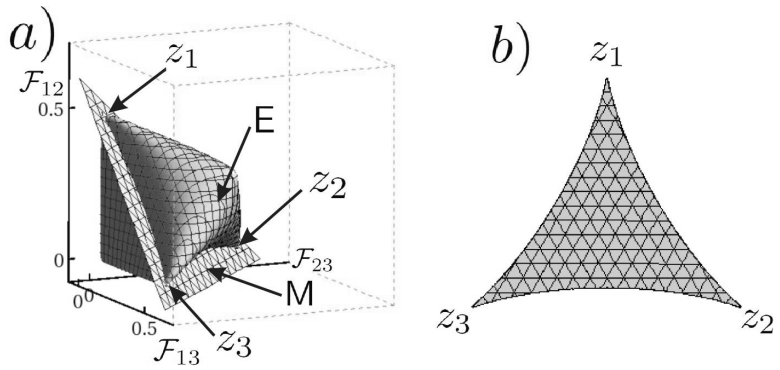


Figure 5.5: *a*) The set E of 3×3 bistochastic matrices $\mathcal{F}$ (5.52), which form the zero frequency block of the Davies channel Φ (5.49) under condition $T\to\infty$, is represented by the vector of the off–diagonal elements $\{\mathcal{F}_{12},\mathcal{F}_{13},\mathcal{F}_{23}\}$. The set E is cut by the plane $\mathsf{M}:\mathcal{F}_{12}+\mathcal{F}_{13}+\mathcal{F}_{23}=\frac{1}{2}$. The cross–section is presented in Panel *b*) and shows that the set M is not convex.

Chapter 6

Concluding remarks and open problems

The aim of this thesis was to investigate quantum channels on different levels of generality and using different approaches. For instance, general properties of quantum channels were considered in Chapter 2, while some particular classes of one–qubit and one–qutrit quantum channels were analysed in Part III of these thesis. The Davies maps motivated by a specific physical model were studied in Chapter 5. Some useful characteristics of a quantum channel are provided by different kinds of entropies. Among them we used the minimal output entropy, the entropy of a map, the entropy of an environment which takes part in an evolution described by a channel. Apart of the standard von Neumann entropy which is the quantum counterpart of the Shannon entropy, the quantum Rényi and Tsallis entropies were also applied.

In Part II the universal entropic inequality for an arbitrary ensemble of quantum states is proved for the von Neumann entropy. This part of the thesis treats a quantum channel as a device preparing a quantum ensemble. The Holevo quantity of this ensemble is shown to be bounded by the entropy of an environment, used in the preparation process. The state of the environment after a quantum operation Φ is equivalent to the output of the complementary channel $\tilde{\Phi}$.

One can define *selfcomplementary* channels for which $\Phi(\rho) = \tilde{\Phi}(\rho)$ for any ρ. Relation (1.38) between the Kraus operators of Φ and the Kraus operators of $\tilde{\Phi}$ is useful to specify selfcomplementary channels. Since the coherent information (2.17) of such channels is equal to 0, the same holds also for the quantum channel capacity (2.18). Identification of selfcomplementary channels, as well as investigation of their properties are worth to be studied in future.

Chapter 3 contains the conjecture which establishes a relation between the Holevo quantity, and the matrix of fidelities. This leads to a geometric characterization of the states in the ensemble. The bound on the Holevo quantity proved in Chapter 2 can be also related to other notions of quantum information

theory, such as the *quantum discord* [102], [103] which measures the quantum correlations in a two–partite system.

The study of quantum channels is an important task of the modern theory of quantum information. For example, the problem of additivity of the channel capacity, or equivalently, additivity of the minimal output entropy remains open even for channels acting on a single qubit. Results presented in this thesis could be further developed to investigate the additivity conjecture for different classes of quantum channels.

Some results of Chapter 4 concern general properties of quantum channels. For instance we proved the additivity of the map entropy (4.11), and Theorem 6 establishing the extremal position of depolarizing channels in the set of all channels characterized by the Rényi entropies $S_2^{\min}(\Phi)$ and $S_2^{\max}(\Phi)$. These results allow us to pose Conjecture 2 specifying pairs of maps for which the additivity of channel capacity may hold.

In Part III, some specific types of channels are investigated. Properties of one–qubit channels are analysed in Chapter 4. Some transformations on one–qubit quantum channels defined in Section 4.3 lead to new results on the characterization of the set of quantum channels in the plane $(S^{\min}(\Phi), S^{\max}(\Phi))$. The aim of this analysis is to find some conditions that enable one to estimate the minimal output entropy, which is difficult to compute, by the entropy of the map easy to calculate.

The Davies channels, which correspond to a concrete physical model, are studied in Chapter 5. Superoperators of the Davies maps are specified in the case of one–qubit maps and one–qutrit maps. The question whether the channel capacity of the Davies maps is additive is still open, although, Davies maps acting on N–level system compose the set of only $d = N^2 - 1$ dimensions, while the set of all quantum operations acting on system of the same N has $N^2(N^2-1)$ dimensions.

The quantum information theory is a modern field of science which creates an environment for new future applications and opens new paths for development of technology. Quantum channels, which describe any possible evolution of a quantum state, play an important role in possible applications. Quantum channels describe decoherence caused by the interaction with an environment. Knowledge of their properties allows one to choose the most efficient quantum protocols for a given purpose. Theoretical investigations uncover new possibilities, new laws and fundamental restrictions on processing of quantum information.

The classical theory of information began with investigations on communication in a given language through given technological tools. However, very fast, the laws of information became treated as fundamental properties of nature. Therefore, studies in the field of quantum information are so exciting.

Appendix 1

In Appendix we analyze ensembles of three one–qubit states $\{\rho_1 = |\phi_1\rangle\langle\phi_1|, \rho_2 = |\phi_2\rangle\langle\phi_2|, \rho_3\}$ and provide calculations related to Fig. 3.1 necessary to prove Lemma 4 in Section 3.2.

The Bloch vector characterizing the average states can be given by

$$\vec{OA} = a(0,0,1). \tag{6.1}$$

The Bloch vector representing the mixed state ρ_3 is parameterized by an angle α

$$\vec{OB} = b(0, \sin\alpha, \cos\alpha). \tag{6.2}$$

The vector $\vec{OC}$ is chosen in such a way that the ratio $|CA| : |AB|$ is $1 : 2$. Therefore one has

$$\vec{OC} = \frac{1}{2}(3\vec{OA} - \vec{OB}). \tag{6.3}$$

The point C is in the center of the interval DE, between two pure states $|\phi_1\rangle$ and $|\phi_2\rangle$ characterized by the points D and E. Both vectors $\vec{OD}$ and $\vec{OE}$ form with vector $\vec{OC}$ the angle γ so that

$$\cos\gamma = |OC|. \tag{6.4}$$

This is in turn the square root of the fidelity $|\langle\phi_1|\phi_2\rangle|^2$, because the angle γ is half of the angle between two pure states,

$$F_{23} = \cos^2\gamma. \tag{6.5}$$

The fidelity between two one–qubit states represented by Bloch vectors $\vec{x}$ and $\vec{y}$ reads

$$F = \frac{1}{2}(1 + \vec{x}\cdot\vec{y}). \tag{6.6}$$

The scalar product of $\vec{OB}$ and $\vec{OD}$ is equal to:

$$\vec{OB}\cdot\vec{OD} = b\cos(\mu+\gamma-\gamma) = b\Big[\cos(\mu+\gamma)\cos\gamma + \sin(\mu+\gamma)\sin\gamma\Big]. \tag{6.7}$$

Hence

$$F_{12} = \frac{1}{2}(1 + \vec{OC}\cdot\vec{OB} + b\sqrt{1 - \frac{(\vec{OC}\cdot\vec{OB})^2}{b^2\vec{OC}\cdot\vec{OC}}}\sqrt{1 - \vec{OC}\cdot\vec{OC}}). \tag{6.8}$$

The third fidelity F_{13} can by obtained using Lemma 2. For $\beta = 0$ the product of three fidelities used in Lemma 4 is a function $f_0(a,b,\alpha,\beta=0)$ given by

$$f_0(a,b,\alpha,\beta=0) = F_{12}F_{13}F_{23} = \tag{6.9}$$

$$\tfrac{1}{64}\left(9a^2 - 6b\cos\alpha a + b^2\right)\left(b^2 - 3a\cos\alpha b - 2\right)^2 \tag{6.10}$$

$$+\tfrac{1}{64}\left(9a^2b^2\left(9a^2 - 6b\cos\alpha a + b^2 - 4\right)\sin^2\alpha\right). \tag{6.11}$$

Appendix 2

In this appendix we present computations necessary to prove Lemma 3. It is convenient to change the basis such that the vector $\vec{OB}$ (see Fig. 3.1) is transformed into

$$\vec{OB}' = b(0,0,1). \tag{6.12}$$

Denote the angle $\nu := \mu + \gamma$, where μ is the angle between $\vec{OB}$ and $\vec{OD}$. The vectors $\vec{OD}$ and $\vec{OE}$ in the new basis can be obtained by rotating the state $(0,0,1)$ around the axis x by angles:

$$\vec{OD}' = U_x(\mu)(0,0,1), \tag{6.13}$$

$$\vec{OE}' = U_x(\mu + 2\gamma)(0,0,1). \tag{6.14}$$

The vectors $\vec{OG}'$ and $\vec{OF}'$ are obtained by rotating the above vectors around the axis $U_x(\nu)(0,0,1)$ by angle β. Such a rotation can be defined as an action of a unitary matrix U on vectors (6.13). The unitary matrix is given by

$$U = U_z(-\frac{\pi}{2})\, U_y(\nu)\, U_z(\beta)\, U_y^\dagger(\nu)\, U_z^\dagger(-\frac{\pi}{2}), \tag{6.15}$$

where the rotation matrices read

$$U_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, \quad U_y(\alpha) = \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix},$$

$$U_z(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

One can use formula (6.6) to calculate the product of three fidelities for three considered states, $\vec{OB}', \vec{OG}'$ and $\vec{OF}'$ as a function of the angle β.

$$\begin{aligned} f &= F_{12}F_{13}F_{23} \qquad (6.16)\\ &= \frac{1}{16}\cos^2\gamma\left((\cos\mu + \cos(2\gamma+\mu) + 2)^2 - \cos^2\beta(\cos\mu - \cos(2\gamma+\mu))^2\right). \end{aligned}$$

The product of three pairwise fidelities attains its minimum at $\beta = 0$ as stated in Lemma 3.

Bibliography

[1] C. Shannon, *A Mathematical Theory of Communication*, The Bell System Technical Journal, **27** 379–423, 623–656 (1948).

[2] A. Rényi, *On measures of information and entropy*, Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability, 547–561 (1960).

[3] C. Tsallis, *Possible generalization of Boltzmann-Gibbs statistics*, J. Stat. Phys., **52** 479-487 (1988).

[4] A. Plastino, A. R. Plastino, *Tsallis Entropy and Jaynes' Information Theory Formalism*, Brazilian Journal of Physics, **29** 50-60 (1999).

[5] E. Davies, *Quantum stochastic processes*, Commun. Math. Phys., **15** 277–306 (1970).

[6] A. Kossakowski, *On quantum statistical mechanics of non-Hamiltonian systems*, Rep. Math. Phys., **3** 247–274 (1972).

[7] A. Holevo, *Bounds for the quantity of information transmitted by a quantum communication channel*, *Prob. Inf. Transm.* **9** 177–83 (1973).

[8] M. Nielsen, I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge (2000).

[9] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, Phys. Rev. Lett., **70** 1895–1899 (1993).

[10] D. Deutsch, R. Jozsa, *Rapid solutions of problems by quantum computation*, Proceedings of the Royal Society of London A, **439** 553–558 (1992).

[11] Grover L.K. *A fast quantum mechanical algorithm for database search*, Proceedings, 28th Annual ACM Symposium on the Theory of Computing, 212–219 (1996).

[12] P. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J.Sci.Statist.Comput., **26** 1484–1509 (1997).

[13] M. Horodecki, P. Horodecki, and R. Horodecki, *General teleportation channel, singlet fraction, and quasidistillation*, Phys. Rev. A, **60** 1888–1898 (1999).

[14] R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, *Quantum entanglement*, Rev. Mod. Phys., **81** 865–942 (2009).

[15] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press, Cambridge (2006)

[16] M. Hastings, *Superadditivity of communication capacity using entangled inputs*, Nature Physics, **5** 255–257 (2009).

[17] M. Horodecki, *On Hastingsâ counterexamples to the minimum output entropy additivity conjecture*, Open Systems & Information Dynamics, **17** 31–52 (2010) .

[18] C. King, *Remarks on the Additivity Conjectures for Quantum Channels*, in: Entropy and the Quantum, eds. R. Sims, D. Ueltschi, Contemporary Mathematics, **529** 177–188, University of Arizona (2010).

[19] C. Bennett, G. Brassard, *Quantum Cryptography: Public Key Distribution and Coin Tossing*, Proceedings of IEEE International Conference on Computers Systems and Signal Processing, Bangalore India, 175–179 (1984).

[20] A. Wehrl, *General properties of entropy*, Rev. Mod. Phys., **50** 221–260 (1978).

[21] M. B. Ruskai, *Inequalities for Quantum Entropy: A Review with Conditions for Equality*, J. Math. Phys., **43** 4358–4375 (2002).

[22] F. Nielsen, R. Nock, *On Rényi and Tsallis entropies and divergences for exponential families*, arXiv:1105.3259, (2011).

[23] C. Shannon, *Communication in the presence of noise*, Proc. Institute of Radio Engineers, **37** 10–21. (1949).

[24] R. Hartley, *Transmission of Information*, Bell System Technical Journal, **7** 535–563 (1928).

[25] B. Schumacher, *Quantum coding*, Phys. Rev. A, **51** 2738–2747 (1995).

[26] E. Desurvire, *Classical and Quantum Information Theory. An Introduction for the Telecommunication Scientists*, Cambridge University Press, Cambridge (2009).

[27] A. Holevo, *The capacity of quantum channel with general signal states*, IEEE Trans. Info. Theory, **44** 269–273 (1998).

[28] B. W. Schumacher and M. Westmoreland, *Sending classical information via noisy quantum channels*, Physical Review A, **56** 131–138 (1997).

[29] B. Schumacher, *Sending entanglement through noisy quantum channels*, Phys. Rev. A, **54** 2614–2628 (1996).

[30] C. Fuchs, *Nonorthogonal quantum states maximize classical information capacity*, Phys. Rev. Lett., **79** 1162–1165 (1997).

[31] M. Hayashi, H. Imai, K. Matsumoto, M. B. Ruskai, T. Shimono, *Qubit channels which require four inputs to achieve capacity: Implication for additivity conjectures*, Quantum Inf. Comput., **5** 13–31 (2005).

[32] E. Knill, R. Laflamme, A. Ashikhmin, H. Barnum, L. Viola and W. H. Zurek, *Introduction to Quantum Error Correction*, arXiv:quant-ph/0207170 (2002).

[33] M.-D. Choi. *Completely positive linear maps on complex matrices*, Linear Algebra and Its Applications, **10** 285â-290 (1975).

[34] A. Jamiołkowski, *Linear transformations which preserve trace and positive semidefiniteness of operators*, *Rep. Math. Phys.* **3** 275 (1972).

[35] K. Kraus. *States, Effects and Operations: Fundamental Notions of Quantum Theory*, Springer-Verlag, Berlin (1983).

[36] J. De Pillis, *Linear transformations which preserve hermitian and positive semidefinite operators*, Pacific J. Math., **23** 129–137 (1967).

[37] A. Fujiwara and P. Algoet, *Affine parametrization of quantum channels.* Phys. Rev. A, **59** 3290–3294 (1999).

[38] M. B. Ruskai, S. Szarek, E. Werner, *An analysis of completely-positive trace-preserving maps on* $\mathcal{M}_2$, Linear Algebra and its Applications, **347** 159–187 (2002).

[39] P. Shor, *Equivalence of additivity questions in quantum information theory*, Commun. Math. Phys., **246** 453–472 (2004).

[40] G. Amosov, A. Holevo, and R. Werner, *Additivity/multiplicativity problems for quantum communication channels*, Quantum Communication, Computing, and Measurement, **3** 3–10 (2001).

[41] C. King, *The capacity of the quantum depolarizing channel*, IEEE Transactions on Information Theory, **49** 221–229 (2003).

[42] T. S. Cubitt, M B. Ruskai, G. Smith, *The structure of degradable quantum channels*, *J. Math. Phys.* **49** 102104 (27 pp) (2008).

[43] G. Lindblad, *On the generators of quantum dynamical semigroups*, Commun. Math. Phys., **48** 119–130 (1976).

[44] K. Kraus, *General state changes in quantum theory*, *Ann. Phys.* , **64** 311–35 (1971).

[45] H.-P. Breuer, F. Petruccione, *The theory of open quantum systems*, Clarendon Press, Oxford (2006).

[46] W. Roga, M. Fannes, K. Życzkowski, *Composition of quantum states and dynamical subadditivity*, J. Phys. A. Math. Theor., **41** 035305 (15 pp) (2008).

[47] F. Verstraete, H. Verschelde, *On quantum channels*, arXiv:quant-ph/0202124 (2002).

[48] M. Ziman, *Incomplete quantum process tomography and principle of maximal entropy*, Phys. Rev. A, **78** 032118 (8 pp) (2008)

[49] W. Roga, M. Fannes, K. Życzkowski, *Universal Bounds for the Holevo Quantity, Coherent Information, and the Jensen-Shannon Divergence*, *Phys. Rev. Lett.*, **105** 040505 (4 pp) (2010).

[50] W. Roga, M. Fannes, K. Życzkowski, *Davies maps for qubit and qutrits*, Rep. Math. Phys., **66** 311–329 (2010).

[51] W. Roga, M. Fannes, K. Życzkowski, *Entropic characterization of quantum operations*, International Journal of Quantum Information, **9** 1031–1045 (2011).

[52] M. Fannes, F. de Melo, W. Roga, K. Życzkowski *Matrices of fidelities for ensembles of quantum states and the Holevo quantity*, arXiv/quant-ph:1104.2271 (2011).

[53] W. Roga, M. Smaczyński, K. Życzkowski, *Composition of Quantum Operations and Products of Random Matrices*, Acta Physica Polonica B, **42** 1123 (18 pp) (2011).

[54] C. King and M. B. Ruskai, *Minimal Entropy of States Emerging from Noisy Quantum Channels*, IEEE Trans. Info. Theory., **47** 192–209 (2001).

[55] M. Agrawal, *Axiomatic / Postulatory Quantum Mechanics, in Fundamental Physics in Nano-Structured*, Stanford University (2008).

[56] C. Hong, Z. Ou, L. Mandel, *Measurement of subpicosecond time intervals between 2 photons by interference*, Phys. Rev. Lett., **59** 2044–2046 (1987).

[57] E. B. Davies, *Quantum Theory of Open Systems*, Academic Press, London (1976).

[58] E. Schmidt, *Zur Theorie der linearen und nicht linearen Integralgleichungen*, Math. Ann, **63** 433–466 (1907).

[59] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik (Mathematical Foundations of Quantum Mechanics)*, Springer, Berlin (1955).

[60] E. H. Lieb and M.B. Ruskai, *Proof of the strong subadditivity of quantum mechanical entropy*, J. Math. Phys., **14** 1938–1941 (1973).

[61] W. Stinespring, *Positive Functions on C*-algebras*, Proc. Amer. Math. Soc., **6** 211–216, (1955).

[62] A. Uhlmann, *The "transition probability" in the state space of a *-algebra*, Rep. Math. Phys., **9** 273–279 (1976).

[63] S. Kokkendorf, *Gram matrix analysis of finite distance spaces in constant curvature*, Discrete Comput. Geom., **31** 515–543 (2004).

[64] B. Bahr, B. Dittrich, *Regge calculus from a new angle*, New Journal of Physics, **12** 033010, (10 pp) (2010).

[65] R. Alicki, M. Fannes, *Quantum dynamical systems*, Oxford University Press (2001).

[66] R. Jozsa, J. Schlienz, *Distinguishability of states and von Neumann entropy*, Phys. Rev. A, **62** 012301 (11 pp) (2000).

[67] G. Mitchison, R. Jozsa, *Towards a geometrical interpretation of quantum–information compression*, Phys. Rev. A, **69** 032304 (6 pp) (2004).

[68] W. Wootters, *Statistical distance and Hilbert space*, Phys. Rev. D., **23** 357–362 (1981).

[69] B. Fuglede, F. Topsøe, *Jensen-Shannon divergence and Hilbert space embedding*, IEEE International Symposium on Iinformation Theory, Proceedings, 31–31 (2004).

[70] F. Topsøe, *Some inequalities for information divergence and related measures of discrimination*, IEEE Trans. Inform. Theory, **46** 1602–1609 (2000).

[71] A. Holevo, M. Sirokov, *Mutual and coherent information for infinite-dimensional quantum channels*, Problems of information transmission, **46** 201–218 (2010).

[72] K. Horodecki, M. Horodecki, P. Horodecki, J. Oppenheim, *General paradigm for distilling classical key from quantum states*, IEEE Transactions on Information Theory, **55** 1898–1929 (2009).

[73] B. W. Schumacher, M. A. Nielsen, *Quantum data processing and error correction*, Phys. Rev. A, **54** 2629–2635 (1996).

[74] S. Lloyd, *Capacity of the noisy quantum channel*, Phys. Rev. A, **55** 1613–1622 (1997).

[75] G. Lindblad, *Quantum entropy and quantum measurements*, in Quantum Aspects of Optical Communication, eds. C. Bendjaballah et al., Lecture Notes in Physics, **378** 71–80, Springer-Verlag, Berlin (1991).

[76] H. Araki, E. Lieb, *Entropy inequalities*, Comm. Math. Phys. **18** 160–170 (1970).

[77] G. Lindblad, *An entropy inequality for quantum measurements*, Commun. Math. Phys., **28** 245–249 (1972).

[78] A. Uhlmann, *Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory*, Commun. Math. Phys., **54** 21–32 (1977).

[79] S. Furuichi, K. Yanagi, and K. Kuriyama *Fundamental properties of Tsallis relative entropy*, J. Math. Phys., **45** 4868 (10 pp) (2004).

[80] F. Hiai, M. Mosonyi, D. Petz, *Monotonicity of f-divergences: A review with new results*, arXiv/math-phys: 1008.2529 (2008).

[81] R. Konig, S. Wehner, *A Strong Converse for Classical Channel Coding Using Entangled Inputs*, Phys. Rev. Lett., **103** 070504 (4 pp) (2009).

[82] J. Briët, P. Herremoës *Properties of classical and quantum Jensen-Shannon divergence*, *Phys. Rev. A*, **79** 052311 (11 pp) (2009).

[83] D. M Endres, J. E. Schindelin, *A new metric for probability distributions*, IEEE Trans. Inf. Theory, **49** 1858–1860 (2003).

[84] M. Fannes and D. Vanpeteghem, *A three state invariant*, arXiv:quant-ph/0402045 (2002).

[85] R. Bhatia, *Positive Definite Matrices*, Princeton University Press, Princeton (2007).

[86] K. Życzkowski, H.-J. Sommers, *Hilbert–Schmidt volume of the set of mixed quantum states*, J. Phys. A, **36** 10115–10130 (2003).

[87] C. King, *Additivity for unital qubit channels*, J. Math. Phys., **43** 4641 (13 pages) (2002).

[88] P. Shor, *Additivity of the classical capacity of entanglement-breaking quantum channels*, J. Math. Phys., **43** 4334–4340, (2003).

[89] D. DiVincenzo, P. Shor, J. Smolin, *Quantum-channel capacity of very noisy channels*, Phys. Rev. A, **57** 830–839 (1998).

[90] A. Holevo, *The additivity problem in quantum information theory*, Russian Mathematical Surveys, **61** 301–339 (2006).

[91] R. Alicki, K. Lendi, *Quantum dynamical semigroups and applications*, Springer-Verlag, Berlin (1987).

[92] E. B. Davies, *Markovian master equations*, Commun. Math. Phys., **39** 91–110 (1974).

[93] V. Gorrini, A. Kossakowski and E. Sudarshan, *Completely positive dynamical semigroups of n-level systems*, J. Math. Phys., **17** 821–825 (1976).

[94] T. Cubitt, J. Eisert, M. Wolf, *Deciding whether a Quantum Channel is Markovian is NP-hard*, arXiv/math-phys:0908.2128v1 (2009).

[95] M. Wolf, J. Eisert, T. S. Cubitt and J. I. Cirac, *Assessing non-Markovian dynamics*, Phys. Rev. Lett., **101** 150402 (4 pp) (2008).

[96] G. S. Agarwal, *Open quantum Markovian systems and the microreversibility*, Z. Phys., **258** 409–422 (1973).

[97] G. Kimura, *Restriction on relaxation times derived from the Lindblad-type master equations for two-level systems*, Phys. Rev. A, **66** 062113 (4 pp) (2002).

[98] S. Daffer, K. Wódkiewicz and J. K. McIver: *Quantum Markov channels for qubits*, Phys. Rev. A, **67** 062312 (13 pp) (2003).

[99] C. King, N. Koldan *New multiplicativity results for qubit maps*, J. Math. Phys., **47** 042106 (9 pp) (2006).

[100] J. M. Ortega, *Matrix Theory, A Second Course*, Plenum Press, New Yourk (1987).

[101] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge (1985).

[102] P. Coles, *Non-negative discord strengthens the subadditivity of quantum entropy functions*, arXiv:1101.1717 (2011).

[103] H. Ollivier, W. Zurek, *Quantum Discord: A Measure of the Quantumness of Correlations*, Phys. Rev. Lett., **88** 017901 (4 pp) (2002).

[104] P. Coles, L. Yu, V. Gheorghiu, R. Griffiths, *Information-theoretic treatment of tripartite systems and quantum channels*, Phys. Rev. A, **83** 062338 (2011).

[105] P. W. Lamberti, M. Portesi, J. Sparacino, *Natural metric for quantum information theory*, International Journal of Quantum Information, **7** 1009–1019 (2009).

[106] V. Belavkin, *Contravariant densities, complete distances and relative fidelities for quantum channels*, Rep. Math. Phys., **55** 61–77 (2005).

[107] R. Alicki, M. Fannes, *Note on Multiple Additivity of Minimal Rényi Entropy Output of the Werner-Holevo Channels*, Open Sys. & Information Dyn., **11** 339–342 (2004).