Quantum Capacities and Entropy Production of Quantum Markov Chains

TECHNISCHE UNIVERSITÄT MÜNCHEN

Alexander Müller-Hermes

Zentrum Mathematik M5

Technische Universität München

A thesis submitted for the degree of

Doctor rerum naturalium

2015 May

TECHNISCHE UNIVERSITÄT MÜNCHEN
ZENTRUM MATHEMATIK

Lehrstuhl für mathematische Physik

Quantum Capacities and Entropy Production of Quantum Markov Chains

Alexander Müller-Hermes

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzende : Univ.-Prof. Dr. Simone Warzel

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Michael M. Wolf
2. Univ.-Prof. Dr. Dr. Holger Boche
3. Prof. Dr. Matthias Christandl (nur schriftliche Beurteilung),
   University of Copenhagen,
   Dänemark

Die Dissertation wurde am 26.05.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 18.08.2015 angenommen.

The road to wisdom? — Well, it's plain
and simple to express:
Err
and err
and err again
but less
and less
and less.

– Piet Hein

Acknowledgements

First and foremost I want to thank Prof. Michael Wolf for supervising my doctoral research study and giving me the opportunity to work in his group. He created the perfect working environment by providing me with interesting ideas to work on and at the same time giving me the freedom to follow my own interests. Thank you.
This work would not have been possible without David Reeb, who always had time and patience to discuss mathematics, physics and academic life in general. Thank you for this.
I am very grateful to my colleagues and collaborators Daniel Stilck França and Oleg Szehr. It was fun working with you.
For their time and effort I would like to thank all the committee members, Prof. Matthias Christandl, Prof. Holger Boche and Prof. Simone Warzel.
For proof-reading the introductory part of this dissertation and for finding many small (and also larger) mistakes I am very thankful to Anna Hashagen.
Furthermore, I would like to thank all the members (and former members) of M5 for turning our set of people into a group.
Finally I am deeply grateful to my girlfriend Antonia Schmalstieg for loving me even when I am consumed by mathematics.

List of contributed articles

I) (*) A. Müller-Hermes, D. Reeb and M.M. Wolf,
Quantum Subdivision Capacities and Continuous-time Quantum Coding,
IEEE Transactions on Information Theory, vol.61, no.1, pp.565-581, Jan. 2015

II) (*) A. Müller-Hermes, D. Reeb and M.M. Wolf,
Positivity of Linear Maps under Tensor Powers,
Journal of Mathematical Physics, 57, 015202 (2016)

III) A. Müller-Hermes, D. Stilck França and M.M. Wolf,
Entropy-production of Doubly-Stochastic Quantum Channels,
arXiv:1505.04678

IV) (*) A. Müller-Hermes, D. Stilck França and M.M. Wolf,
Relative Entropy Convergence for Depolarizing Channels ,
arXiv:1508.07021

V) (*) A. Müller-Hermes and O. Szehr,
Spectral-variation Bounds in Hyperbolic Geometry,
Linear Algebra and its Application, vol.482, pp.131-148, Oct. 2015
doi:10.1016/j.laa.2015.05.020

For the articles marked by (*) I am principal author.

Contents

1 Introduction **1**

1.1 Summary . . . 1

1.2 Outline . . . 5

2 Finite-dimensional quantum mechanics **7**

2.1 Quantum states and measurements . . . 8

2.2 Quantum Markov chains . . . 9

2.2.1 Choi-Jamiolkowski isomorphism and representation theorems . . . 10

2.2.2 Entanglement breaking and completely co-positive channels . . . 11

2.2.3 Continuous-time quantum Markov processes . . . 12

2.3 Distance measures . . . 13

2.3.1 Matrix norms and induced norms for linear maps . . . 13

2.3.2 Fidelity . . . 16

2.3.3 Relative entropy . . . 17

2.4 Information-theoretic quantities . . . 18

2.4.1 Von-Neumann entropy and conditional entropy . . . 18

2.4.2 Conditional min-entropy . . . 19

2.5 Quantum capacities . . . 21

2.5.1 Quantum Shannon-capacity . . . 21

2.5.2 The decoupling approach . . . 23

2.5.3 Quantum capacity assisted by classical communication . . . 25

Bibliography **29**

1

Introduction

Information theory deals with the mathematical description of information processing tasks. Such tasks can often be treated in an abstract fashion such that the concrete physical realizations do not enter the analysis. This is possible as many systems used for communication can be described within the same stochastic formalism. However, with the development of quantum mechanics it became apparent that the fundamental laws of physics were quite different from what had been anticipated. In particular it was realized that no ordinary stochastic theory could explain the experimental data, and a non-commutative stochastic theory was needed. As a consequence also the classical information theory based on the classical probability theory has to be replaced by a non-commutative quantum information theory, when considering information processing tasks involving quantum systems.

This dissertation is about information theory in systems described by quantum Markov processes. The Markov assumption says that the evolution of a system only depends on the current state and time, and not on the particular history that lead to the current state. This assumption is useful both in the classical and in the quantum theory to simplify the description of a system. Most of the articles included in this dissertation are motivated by the transmission of quantum information (represented by quantum states) over such quantum Markov processes.

1.1 Summary

I will begin with a brief summary of the articles included in this dissertation. Often, research does not proceed in a linear fashion and these articles contain results on different topics in quantum information theory. The main line of research concerns capacities and entropy-production of continuous-time quantum Markov processes (articles I, III and IV). Some questions related to discrete-time quantum Markov processes are addressed in article II. Finally article V is not directly related to quantum information theory. It emerged as a side project and addresses an elementary matrix-theoretic problem.

1. *Quantum capacities for continuous-time quantum Markov processes: Article I*

   A basic scenario of quantum Shannon-theory is the transmission of quantum information through many copies of the same quantum channel. The quantum Shannon-capacity quantifies the optimal amount of transmitted information per channel use (i.e. the optimal rate) that can be transmitted when a suitable encoding is chosen before transmission such that the error of decoding after transmission vanishes in the limit of infinitely many channel uses. This framework can also be applied to quantify the optimal rate of storing quantum information in an array of identical quantum memories each affected by the same continuous-time Markovian noise. However, the standard Shannon-theory is too restrictive in this case as we have direct access to the quantum memory and may apply additional coding channels (or even additional continuous-time Markovian processes) during the storage time to counteract the noise. To apply ideas of quantum Shannon-theory in such a scenario we introduce new capacities allowing for this additional freedom. For the quantum subdivision capacity the noise channel may be subdivided into many pieces with coding channels acting in between to protect the information. Depending on which coding channels we allow different quantum subdivision capacities are obtained. When all infinitely divisible quantum channels are allowed, we show that any noise can be removed and perfect storage for arbitrary long times becomes possible. In the case where only unitary quantum channels are allowed we show that the corresponding quantum subdivision capacity is strictly positive for all times, but for depolarizing noise it exponentially decays to zero. Besides quantum subdivision capacities we also introduce the continuous quantum capacity, where an additional continuous-time Markov evolution may be applied to the system in order to protect the information. This idea is similar to continuous-time quantum error correction studied by several authors, e.g. [1, 2, 3]. However, these ideas have to my knowledge never been studied (even in the classical case) in a Shannon-theoretic setting allowing for techniques like randomized coding and decoupling to be applied.

2. *Positivity of linear maps under tensor powers: Article II*

   A linear map acting on a matrix algebra is called positive if the image of any positive semidefinite matrix is again positive semidefinite. In this article we study how positivity behaves under tensor products. In general even the simplest case of the second tensor power, i.e. a tensor product of a positive linear map with itself, will fail to be positive. Nevertheless, there are two trivial classes of positive linear maps (completely positive and completely co-positive maps connected to the transposition) for which any tensor power is again positive. In our work we study the question whether there are non-trivial (i.e. not belonging to the two aforementioned classes) positive maps with this property called *tensor-stable positivity.* Unfortunately, we were not able to show the existence of such a map. However, we prove that their existence would lead to an NPPT bound-entangled state and thereby solve the long-standing NPPT bound entanglement problem [4, 5]. As a direct consequence the only tensor-stable positive maps with input or output dimension two are trivial. The existence of tensor-stable positive maps would have other implications in quantum information theory, e.g. upper bounds (even strong converse bounds) on the quantum Shannon-capacity and proving that any ∞-entanglement annihilating channel has to be entanglement breaking (see [6]). Finally we show that for any fixed number n there is a positive linear map, neither completely positive nor completely co-positive, with positive n-th tensor power. However, it is unclear, whether our construction also yields such a map for all n.

3. *Entropy-production of doubly stochastic quantum channels: Article III*

   Remark: This work was done in a collaboration, where the author contributed but did *not* take the leading role.

   This work originated from article I on quantum capacities for continuous-time Markov-processes, where we used an entropy-production estimate to upper bound the unitary quantum subdivision capacity (i.e. where unitary coding channels may be applied in the intermediate steps) of depolarizing channels. To obtain similar bounds for more general Liouvillians we study logarithmic-Sobolev inequalities. It turns out that for primitive and doubly stochastic quantum Markov-processes the entropy increases exponentially fast to its maximal value with an exponent proportional to the logarithmic-Sobolev constant. It is usually difficult to compute the exact value of logarithmic Sobolev constants even in the classical case. Additionally we need entropy-production estimates scaling well under tensor powers of quantum channels to prove upper bounds on the unitary quantum subdivision capacity. We establish such bounds using in the first step a group-theoretic technique and hypercontractivity to compute a bound on the logarithmic-Sobolev constant for tensor powers of the depolarizing channel. Then by a comparison method we obtain an estimate for the logarithmic-Sobolev constant of any doubly stochastic, primitive continuous-time quantum Markov-process in terms of its spectral gap. This method is new and our result improves on similar ones obtained by more complicated techniques. Using the entropy production estimate derived this way we can prove exponential decay of the unitary quantum subdivision capacity of any doubly stochastic, primitive continuous-time quantum Markov-process.

4. *Relative entropy convergence for depolarizing channels: Article IV*
   While the previous work dealt with logarithmic-Sobolev inequalities of unital continuous-time quantum Markov processes in this article we go back to the case of depolarizing channels now with general fixed points. Our main result is the computation of the logarithmic-Sobolev constant for these kind of channels. We are able to show, that this constant coincides with the logarithmic-Sobolev constant of a classical random walk on a finite set with transition probabilities equal to the eigenvalues of the fixed point. Surprisingly this constant seems to be unknown even in the classical case and by using Lagrange-multipliers we compute it explicitly. As an application we derive an improved concavity inequality for the von-Neumann (or Shannon) entropy of a convex combination of two states with a correction term depending on the relative entropy of the states. This improvement seems to be incomparable to other recent improvements to this inequality, but in some cases gives better bounds than the previously known results (cf. [7]). Further results contained in this paper are an optimal (in terms of spectral data of one of the two states) Pinsker's inequality and a proof of an entropy production bound for tensor powers of a depolarizing channel (stated less general in [8]) using a quantum Shearer's inequality.

5. *Spectral-variation Bounds in Hyperbolic Geometry: Article V*
   This paper is not directly related to quantum information theory and arose from the work of the second author. Given two complex $n \times n$-matrices we consider the problem of bounding the optimal matching distance between their spectra in terms of the norm distance of the two matrices. Despite considerable interest in the question, the optimal constants in these bounds are not known. We introduce an optimal matching distance in the hyperbolic pseudometric on the Poincaré-disc model. To derive spectral-variation

bounds for this distance we use resolvent bounds from the theory of model operators [9] and a Chebyshev-type interpolation result for Blaschke-products [10]. Due to properties of the hyperbolic pseudometric our bound is hard to compare to previous results in the Euclidean metric. However, if the distance between the two matrices is small enough we obtain a Euclidean spectral-variation bound with a constant approaching the conjectured optimal value in the limit of large dimensions.

1.2 Outline

In the following chapter I will introduce the necessary concepts of quantum information theory to understand the articles included in this thesis. As these articles only deal with finite dimensional spaces I will restrict my introduction to this case. Note that by doing so many technical difficulties usually occurring in infinite-dimensional Banach-space theory can be avoided. However, even in finite dimensions there are still many open questions and hard problems mostly due to the non-commutative nature of quantum theory. To make the exposition as simple as possible I only consider matrix algebras and avoid introducing the framework on general C^*-algebras. This does not restrict the theory in finite dimensions.

In section 2.1 I start with a brief introduction of the basic formalism of finite-dimensional quantum mechanics. Section 2.2 contains a brief introduction to quantum channels (quantum Markov chains) and mathematical tools needed to study their properties. Special emphasize is put on continuous quantum Markov-processes towards the end of this section. In section 2.3 distance measures on matrix algebras and state spaces will be introduced. In particular I give a brief introduction to completely bounded p-to-q-norms, which were introduced in operator space theory [26]. In section 2.4 I introduce some information theoretic quantities necessary to understand the articles and the rest of the introduction. Finally in section 2.5 I conclude with a brief introduction to quantum Shannon-theory and the decoupling approach.

Throughout this introduction I will omit references for basic definitions well-known in the field of quantum information theory. These definitions can be found in the standard textbook by Michael A. Nielsen and Isaac L. Chuang [11]. Furthermore, I do not claim any credit for the results summarized in this chapter even if a reference is missing.

2

Finite-dimensional quantum mechanics

Throughout this thesis the standard notation $\mathbb{C}^{d_1 \times d_2}$ for complex $d_1 \times d_2$-matrices is used. If $d_1 = d_2 = d$ we will also write $\mathcal{M}_d := \mathcal{M}\left(\mathbb{C}^d\right) = \mathbb{C}^{d \times d}$ for the algebra of complex $d \times d$-matrices. Note that in finite dimensions we can simply define $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} := \{X \otimes Y \ : \ X \in \mathcal{M}_{d_1}, Y \in \mathcal{M}_{d_2}\}$ where the $\otimes$ (between matrices) denotes the Kronecker product (see [11]). By iterating this definition we obtain $\mathcal{M}\left(\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}\right) := \bigotimes_{i=1}^n \mathcal{M}_{d_i}$. A matrix $X \in \mathcal{M}_d$ is called *positive* (denoted by $X \geq 0$) if and only if (iff) it is Hermitian with $\operatorname{spec}(X) \subset [0, \infty)$. The set of positive matrices forms a cone and is denoted by $\mathcal{M}_d^+ \subset \mathcal{M}_d$. For positive matrices on the tensor product $\mathcal{M}\left(\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}\right)$ we will write $\mathcal{M}^+\left(\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}\right)$.

We will often consider linear maps between matrix algebras. Such maps could in principle be represented by matrices again, but we will rarely do so as the properties we want to consider would not have easy counterparts in these representations. Note that for linear maps $\mathcal{S} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ and $\mathcal{T} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_4}$ we can define the tensor product $\mathcal{S} \otimes \mathcal{T}$ as the unique linear map fulfilling $(\mathcal{S} \otimes \mathcal{T})(X \otimes Y) = \mathcal{S}(X) \otimes \mathcal{T}(Y)$ for all $X \in \mathcal{M}_{d_1}, Y \in \mathcal{M}_{d_2}$.

Formulas quickly get complicated when tensor product spaces are involved. Therefore it is sometimes useful to attach labels to matrices indicating the tensor structure of the space they are contained in. For instance we write $X_{AB} \in \mathcal{M}\left(\mathbb{C}^{d_A} \otimes C^{d_B}\right)$. We might also write $V^{A \to BE} : \mathbb{C}^{d_A} \to \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$ to emphasize that the matrix V (considered as a linear map) maps to a system with a certain tensor structure. Most of the time when only dealing with simple systems where such labels might cause confusion, we will simply omit them for better readability.

2.1 Quantum states and measurements

For $d \in \mathbb{N}$ we denote by $\mathcal{D}_d \subset \mathcal{M}_d$ the set of *quantum states*, i.e. positive matrices with normalized trace, of a d-dimensional quantum system. These are the quantum mechanical analogue of probability distributions. It is easy to see that $\mathcal{D}_d$ is compact, convex and the extremal points of this set, called *pure states*, are the rank-1 projectors $|\psi\rangle\langle\psi|$ for $|\psi\rangle \in \mathbb{C}^d$.

Given $n \in \mathbb{N}$ quantum systems with dimensions $d_1, d_2, \ldots, d_n \in \mathbb{N}$ a composite (multipartite) quantum system can be formed by taking tensor products. The quantum states of this multipartite quantum system are given by $\mathcal{D}\left(\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}\right) \subset \bigotimes_{i=1}^n \mathcal{M}_{d_i}$. A state $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}\right)$ is called *separable* if there are $k \in \mathbb{N}$, a probability distribution $\{p_i\}_{i=1}^k$ and sets of quantum states $\{\sigma_i^m\}_{i=1}^k \subset \mathcal{D}_{d_m}$ for all $m \in \{1, \ldots, n\}$ such that $\rho = \sum_{i=1}^k p_i \sigma_i^1 \otimes \sigma_i^2 \otimes \ldots \otimes \sigma_i^n$. Any state $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}\right)$ that is *not* separable is called *entangled*.

For a multipartite quantum state the *reduced states* on a subsystem is obtained by applying a trace map $\mathrm{tr} : \mathcal{M}_d \to \mathbb{C}$ to the rest of the system. Consider for instance a bipartite quantum state $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$. Tracing out the 'A' part of the state corresponds to applying the *partial trace* $\mathrm{tr}_A = \mathrm{tr} \otimes \mathrm{id}_{d_B}$ to ρ_{AB}. We will denote the quantum state after applying the partial trace by simply omitting the corresponding labels, i.e. in the above case we have $\rho_B := \mathrm{tr}_A(\rho_{AB})$. The reduced state ρ_B corresponds to the quantum information of a local observer having only access to the 'B' part of the system.

Let $\mathbb{1}_d \in \mathcal{M}_d$ denote the identity matrix. We will use the notation $\pi_d := \frac{\mathbb{1}_d}{d} \in \mathcal{D}_d$ (or also $\pi_A \in \mathcal{D}_{d_A}$ if labels are used) for the *maximally mixed state*. Note that this is the analogue of the uniform distribution in classical probability theory. Consider now the composition of two d-dimensional systems. For any choice of orthonormal bases $\{|\psi_i\rangle\}_{i=1}^d$ and $\{|\phi_i\rangle\}_{i=1}^d$ on the two spaces we can define a *maximally entangled state* $\omega_d := |\Omega_d\rangle\langle\Omega_d|$ for $|\Omega_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |\psi_i\rangle \otimes |\phi_i\rangle$. Often all properties we are interested in are preserved under a local change of bases on the two tensor factors. In this case we will consider the so called *computational basis* given by the standard unit vectors $\{|i\rangle\}_{i=1}^d$ with $|i\rangle_j = \delta_{ij}$ and refer to the maximally entangled state with respect to this basis (on each tensor factor) simply as *the* maximally entangled state. When labels are used we will sometimes write $\omega_{RA} \in \mathcal{D}\left(\mathbb{C}^{d_R} \otimes \mathbb{C}^{d_A}\right)$ for $d_A = d_R$ for the maximally entangled state. This might seem strange as $\mathbb{C}^{d_A}$ and $\mathbb{C}^{d_R}$ are mathematically the same space. However, thinking about these spaces as different tensor factors will be helpful when studying coding protocols where they represent different physical systems.

We will conclude this section with a brief description of quantum measurements. An observable on a quantum system is represented by a so called positive operator valued measure. A set $\{E_i\}_{i=1}^k \subset \mathcal{M}_d^+$ is called *positive operator valued measure* (POVM) iff $\sum_{i=1}^k E_i = \mathbb{1}_d$. Here the indices $i \in \{1, \ldots, k\}$ should be thought of as measurement outcomes. Performing the measurement represented by the POVM $\{E_i\}_{i=1}^k \subset \mathcal{M}_d^+$ on a quantum system in a particular state $\rho \in \mathcal{D}_d$ yields outcome $i \in \{1, \ldots, k\}$ with probability $p_i = \mathrm{tr}\,(E_i \rho)$. In the special case where $\mathrm{rk}\,(E_i) = 1$ and $\mathrm{tr}\,(E_i E_j) = \delta_{ij}$, i.e. where $E_i = |\psi_i\rangle\langle\psi_i|$ for mutually orthogonal vectors $|\psi_i\rangle \in \mathbb{C}^d$, the POVM is also called a *von-Neumann measurement*.

2.2 Quantum Markov chains

In the formalism of quantum mechanics physical processes are represented by linear maps mapping quantum states on one physical system to quantum states on another. It is therefore necessary for these maps to preserve both positivity of matrices and the trace. A linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called *positive* iff $\mathcal{T}\,(X) \geq 0$ whenever $X \geq 0$, and it is called *trace-preserving* if it preserves the trace, i.e. $\mathrm{tr}\,[\mathcal{T}\,(X)] = \mathrm{tr}\,[X]$ for any $X \in \mathcal{M}_{d_1}$. An example of such a map is the *identity map* $\mathrm{id}_d : \mathcal{M}_d \to \mathcal{M}_d$ given by $\mathrm{id}_d(X) = X$ for any $X \in \mathcal{M}_d$ (we might also write $\mathrm{id}_A : \mathcal{M}_{d_A} \to \mathcal{M}_{d_A}$ when labels are used).

It turns out that these two conditions are not enough to properly model physical processes in quantum mechanics. A positive map representing a physical process should even preserve positivity when only applied to part of a quantum state describing a larger system (and leaving the rest of the state unchanged). A linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called *completely positive* iff $(\mathrm{id}_n \otimes \mathcal{T}) : \mathcal{M}\left(\mathbb{C}^n \otimes \mathbb{C}^{d_1}\right) \to \mathcal{M}\left(\mathbb{C}^n \otimes \mathbb{C}^{d_2}\right)$ is a positive map for every $n \in \mathbb{N}$ [12]. An example for a positive map that is *not* completely positive is the *matrix transposition* $\vartheta_d : \mathcal{M}_d \to \mathcal{M}_d$ given by $\vartheta_d(X) = X^T$ in any fixed basis. A linear map is called a *quantum channel* iff it is trace-preserving and completely positive. Note that the quantum channel (as defined here) does not depend on the particular history leading to the input state. Therefore, the physical processes described by such channels are Markovian and we will sometimes refer to quantum channels as *quantum Markov processes/chains.*

In the mathematical model of quantum mechanics quantum states evolve by application of quantum channels while observables represented by POVMs stay fixed. This is usually referred to as *"Schrödinger picture"*. An equivalent model can be constructed by letting the observables evolve and the states stay fixed. Note that for a quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$

and any POVM $\{E_i\}_{i=1}^k \subset \mathcal{M}_{d_2}$ we can write the output probabilities of an evolved state as $p_i = \mathrm{tr}\,[E_i \mathcal{T}(\rho)] = \mathrm{tr}\,[\mathcal{T}^*(E_i)\,\rho]$, where $\mathcal{T}^* : \mathcal{M}_{d_2} \to \mathcal{M}_{d_1}$ denotes the dual map of $\mathcal{T}$ with respect to the Hilbert-Schmidt inner product. Thus, without changing the outcome probabilities of measurements (which are the only quantities accessible by experiment) we could take the maps $\mathcal{T}^*$ to represent physical processes. This mathematical model of quantum mechanics is referred to as *"Heisenberg picture"*. It is easy to see that the map $\mathcal{T}^*$ is *unital*, i.e. $\mathcal{T}^*(\mathbb{1}_{d_2}) = \mathbb{1}_{d_1}$, whenever $\mathcal{T}$ is trace-preserving. Thus, the unital completely positive maps are quantum channels in the Heisenberg picture.

2.2.1 Choi-Jamiolkowski isomorphism and representation theorems

In order to check complete positivity and other properties of linear maps between matrix algebras the following theorem is useful.

Theorem 2.2.1 (Choi-Jamiolkowski isomorphism, [13]). *The spaces $\{\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2} \; : \; \mathcal{T}$ linear$\}$ and $\mathcal{M}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\right)$ are isomorphic and an isomorphism is given by*

$$\mathcal{T} \;\longmapsto\; C_{\mathcal{T}} := (id_{d_1} \otimes \mathcal{T})\,(\omega_{d_1})\,. \tag{2.1}$$

The matrix $C_{\mathcal{T}} \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ is called Choi matrix of $\mathcal{T}$.

Some properties of a linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ have useful characterizations in terms of the Choi matrix. For example $\mathcal{T}$ is completely positive iff $C_{\mathcal{T}} \geq 0$ [13]. It is not hard to show (see for instance [13]) that via the Choi-Jamiolkowski isomorphism any rank-1 matrix $|\psi\rangle\langle\psi| \in \mathcal{M}^+\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\right)$ corresponds to a completely positive map $\mathcal{L}_\psi : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ of the form $\mathcal{L}_\psi(X) = AXA^\dagger$ for some fixed matrix $A \in \mathbb{C}^{d_2 \times d_1}$. For any completely positive map $\mathcal{T}$ its positive Choi-matrix $C_{\mathcal{T}}$ admits a spectral decomposition (and thus can be written as a sum of rank-1 matrices). Linearity of the Choi-Jamiolkowski isomorphism then implies:

Theorem 2.2.2 (Kraus decomposition, [13, 14]). *A linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is completely positive iff there are matrices $A_i \in \mathbb{C}^{d_2 \times d_1}$ such that for any $X \in \mathcal{M}_{d_1}$*

$$\mathcal{T}(X) = \sum_{i=1}^{k} A_i X A_i^\dagger.$$

The matrices $\{A_i\}_{i=1}^k$ are called Kraus operators *of $\mathcal{T}$. A completely positive map in the above form is trace-preserving iff $\sum_{i=1}^k A_i^\dagger A_i = \mathbb{1}_{d_1}$.*

The following theorem can be obtained as a consequence of the Kraus decomposition (see [15]), but we will not go into the details of the proof.

Theorem 2.2.3 (Stinespring dilation, [16])**.** *A linear map* $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ *is a quantum channel iff there exists* $d_E \in \mathbb{N}$ *and an isometry* $V : \mathbb{C}^{d_1} \to \mathbb{C}^{d_E} \otimes C^{d_2}$ *such that for any* $X \in \mathcal{M}_{d_1}$

$$\mathcal{T}(X) = tr_E \left[V X V^\dagger \right].$$

The isometry V *is also called the* Stinespring isometry *of the channel* $\mathcal{T}$.

By the Stinespring dilation the output state of a quantum channel corresponds to the reduced state of a larger system in which the initial state is embedded. This illustrates what is called an *open quantum system* instead of a closed one. An observer of an open quantum system can only access part of a quantum system and the rest of the quantum system often called the *environment* might be much larger. This point of view is particularly useful for communication scenarios involving privacy. The information that an adversary (not having access to the output system of the quantum channel) can possibly have about the transmitted state are contained in the environment system after the transmission. Given a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ with Stinespring dilation $\mathcal{T}(X) = \mathrm{tr}_E \left[V X V^\dagger \right]$ the *complementary channel* (see [17]) denoted by $\mathcal{T}^c : \mathcal{M}_{d_A} \to \mathcal{M}_{d_E}$ is defined as $\mathcal{T}^c(X) = \mathrm{tr}_B \left[V X V^\dagger \right]$, where the partial trace acts on the output-system of $\mathcal{T}$. The complementary channel $\mathcal{T}^c$ describes the information flow to the environment, that occurs when the channel $\mathcal{T}$ is applied.

2.2.2 Entanglement breaking and completely co-positive channels

The transmission of quantum entanglement is a fundamental task of quantum information processing. We will introduce two classes of quantum channels that are ill-suited for the transmission of entanglement. A quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called *entanglement breaking* iff $(\mathrm{id}_n \otimes \mathcal{T})(\rho)$ is separable for any $n \in \mathbb{N}$ and any state $\rho \in \mathcal{D}\left(\mathbb{C}^n \otimes \mathbb{C}^d\right)$ [18]. Entanglement breaking channels are not useful for any quantum information processing as they destroy entanglement completely. It can be shown that any entanglement breaking channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ can be written as $\mathcal{T}(\rho) = \sum_{i=1}^k \mathrm{tr}\left[E_i \rho\right] \sigma_i$, i.e. as a measurement with effect operators $\{E_i\}_{i=1}^k \subset \mathcal{M}_{d_1}$ followed by a preparation of new states $\{\sigma_i\}_{i=1}^k \in \mathcal{D}_{d_2}$ depending on the measurement outcome.

We need the following characterization (based on the Hahn-Banach theorem [19]) of separability and some facts about entanglement: A bipartite quantum state $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\right)$ is separable iff $(\mathrm{id}_{d_1} \otimes \mathcal{P})(\rho) \geq 0$ for any positive map $\mathcal{P} : \mathcal{M}_{d_2} \to \mathcal{M}_d$ (with $d \in \mathbb{N}$ arbitrary) [20]. Therefore, we can use positive maps to detect entanglement as any bipartite state with $(\mathrm{id}_{d_1} \otimes \mathcal{P})(\rho) \ngeq 0$ for some positive map has to be entangled. It is clear that only positive maps which are not completely positive yield a useful criterion, because for completely positive maps the expression $(\mathrm{id}_{d_1} \otimes \mathcal{P})(\rho)$ is always positive. One possible choice is the transposition map $\vartheta_d : \mathcal{M}_d \to \mathcal{M}_d$ (in any basis) which is positive but not completely positive. A map of the form $(\mathrm{id}_{d_1} \otimes \vartheta_{d_2}) : \mathcal{M}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\right) \to \mathcal{M}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\right)$ is called a *partial transposition* and states $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\right)$ with $(\mathrm{id}_{d_1} \otimes \vartheta_{d_2})(\rho) \ngeq 0$ are called *NPPT* ("non-positive partial transpose"). Similar ρ is called PPT("positive partial transpose") if $(\mathrm{id}_{d_1} \otimes \vartheta_{d_2})(\rho) \geq 0$. By the previous discussion every NPPT state is entangled and it turns out that a state $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\right)$ with $d_1 d_2 \leq 6$ is separable iff it is PPT [20]. However, this is not true for higher dimensions and already for $d_1 = d_2 = 3$ there are entangled PPT states.

Using the transposition and a completely positive map we can build new positive maps by composition. We call a linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ *completely co-positive* iff $\mathcal{T} = \vartheta_{d_2} \circ \mathcal{S}$ for a completely positive map $\mathcal{S} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ (see for instance [21] where these maps are called copositive). This definition does *not* depend on the basis of the transposition (as positivity is preserved under unitary transformations). Note that a linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is completely co-positive iff $\mathcal{T} = \tilde{\mathcal{S}} \circ \vartheta_{d_1}$ for a completely positive map $\tilde{\mathcal{S}} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$. There are quantum channels which are also completely co-positive. For such a channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ states of the form $(\mathrm{id}_n \otimes \mathcal{T})(\rho)$ are always PPT. It is easy to check that any entanglement-breaking channel is a completely co-positive quantum channel and every completely co-positive quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ for dimensions fulfilling $d_1 d_2 \leq 6$ is entanglement breaking (as the partial transposition detects all entanglement in these dimensions by [20]). In larger dimensions this is no longer true and there are completely co-positive quantum channels that are not entanglement breaking.

2.2.3 Continuous-time quantum Markov processes

A *quantum dynamical semigroup* is a semigroup of quantum channels $\mathcal{T}_t : \mathcal{M}_d \to \mathcal{M}_d$ indexed by a (time-)parameter $t \in \mathbb{R}^+$ that is continuous in t and such that $\mathcal{T}_0 = \mathrm{id}_d$ and $\mathcal{T}_s \circ \mathcal{T}_t = \mathcal{T}_{s+t}$ for all $s, t \in \mathbb{R}^+$. We have the following representation theorem:

Theorem 2.2.4 (Lindblad [22])**.** *For every quantum dynamical semigroup $\mathcal{T}_t : \mathcal{M}_d \to \mathcal{M}_d$ there is a Hermitian matrix $H \in \mathcal{M}_d$ and matrices $\{A_i\}_{i=1}^k \subset \mathcal{M}_d$ such that $\mathcal{T}_t = e^{t\mathcal{L}}$ for the generator $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ of the form*

$$\mathcal{L}(X) = i\,[X, H] + \sum_{i=1}^{k} A_i X A_i^\dagger - \frac{1}{2}\{A_i^\dagger A_i, X\}. \tag{2.2}$$

The generator $\mathcal{L}$ in (2.2) is also called a *Liouvillian*. By differentiation it is easy to check, that for any initial state $\rho \in \mathcal{D}_d$ the time-evolved state $\rho_t := \mathcal{T}_t(\rho)$ fulfills the following differential equation:

$$\frac{d}{dt}\rho_t = \mathcal{L}(\rho).$$

By setting $A_i = 0$ for all $i \in \{1, \ldots, k\}$ we obtain the Schrödinger equation describing the time-evolution of a closed system with a Hamiltonian H. This is solved by $\rho_t = U(t)\rho U(t)^\dagger$ for unitary $U(t) = e^{-iHt}$. Therefore the first part of (2.2) can be interpreted as a unitary time-evolution. The second part can be similarly interpreted as dissipation.

The Liouvillian (2.2) can also be written as

$$\mathcal{L}(X) = \mathcal{S}(X) - KX - XK^\dagger$$

for a completely positive map $\mathcal{S} : \mathcal{M}_d \to \mathcal{M}_d$ and a matrix $K \in \mathcal{M}_d$ such that $\mathcal{S}^*(\mathbb{1}_d) = K + K^\dagger$ (see [23]). By choosing $\mathcal{S} : \mathcal{M}_d \to \mathcal{M}_d$ to be a quantum channel and $K = \frac{1}{2}\mathbb{1}_d$ we see that $\mathcal{L} = \mathcal{S} - \mathrm{id}_d$ is the generator of a quantum dynamical semigroup. In the special case where $\mathcal{S}(X) = \mathrm{tr}[X]\,\sigma$ for a quantum state $\sigma \in \mathcal{D}_d$ this yields the *depolarizing Liouvillian* $\mathcal{L}_\sigma(X) = \mathrm{tr}[X]\,\sigma - X$. This Liouvillian generates the *depolarizing channel* $\mathcal{T}_t(X) = (1 - e^{-t})\mathrm{tr}[X]\,\sigma + e^{-t}X$ depolarizing onto the state σ [11].

A quantum dynamical semigroup $\mathcal{T}_t : \mathcal{M}_d \to \mathcal{M}_d$ is called *primitive* iff there exists a unique full-rank stationary state $\sigma \in \mathcal{D}_d$ such that $\mathcal{T}_t(\rho) \to \sigma$ for any $\rho \in \mathcal{D}_d$ [24]. The depolarizing Liouvillian $\mathcal{L}_\sigma$ generates an example of such a semigroup if $\sigma \in \mathcal{D}_d$ is full-rank. It is clear that generic quantum dynamical semigroups are primitive. For any primitive quantum dynamical semigroup it is easy to show, that there exists a finite time $t_0 \in \mathbb{R}^+$ such that $\mathcal{T}_t$ is entanglement breaking for any $t \geq t_0$.

2.3 Distance measures

We will need some distance measures on $\mathcal{M}_d$ and on the set of linear maps acting on $\mathcal{M}_d$.

2.3.1 Matrix norms and induced norms for linear maps

For $p \in [1, \infty)$ the *p-norm* (also called Schatten-p-norm) of a matrix $X \in \mathcal{M}_d$ is defined as $\|X\|_p := \mathrm{tr}\,[|X|^p]^{1/p}$, where $|X| = \sqrt{X^\dagger X}$ denotes the *absolute value* of X [25]. By the singular value decomposition we can also write $\|X\|_p = \left(\sum_{i=1}^d s_i(X)^p\right)^{1/p}$, where $\{s_i(X)\}_{i=1}^d \subset \mathbb{R}^+$

denote the singular values of X. As for commutative p-norms one consistently defines $\|X\|_\infty = \lim_{p\to\infty}\|X\|_p = \sup_i\{s_i(X)\}$. The Schatten-$p$-norms have the usual properties (Hölder inequality, ordering $\|\cdot\|_p \geq \|\cdot\|_q$ for $p \leq q$, etc.) as their commutative counterparts [25]. The Schatten-1-norm is also called *trace norm*.

When working on tensor products $\mathcal{M}\left(\mathbb{C}^{d_1}\otimes\mathbb{C}^{d_2}\right)$ it will be useful to introduce norms acting like an r-norm on the first tensor factor, but like a p-norm on the second tensor factor. Such norms have been introduced in the context of non-commutative vector-valued L_p-spaces [26]. See also [27] for a more readable presentation of this topic in the special case of finite dimensional matrix algebras. Following their presentation we first define this notion for $r=\infty$. For $X \in \mathcal{M}\left(\mathbb{C}^{d_1}\otimes\mathbb{C}^{d_2}\right)$ we set

$$\|X\|_{(\infty,p)} = \sup_{A,B\in\mathcal{M}_{d_1}} \frac{\|(A\otimes \mathbb{1}_{d_2})X(B\otimes\mathbb{1}_{d_2})\|_p}{\|A\|_{2p}\|B\|_{2p}}.$$

Now one can define the (r,p)*-norm* as

$$\|X\|_{(r,p)} = \inf_{\substack{X=(A\otimes\mathbb{1}_{d_2})Y(B\otimes\mathbb{1}_{d_2})\\ A,B\in\mathcal{M}_{d_1}}} \|A\|_{2r}\|B\|_{2r}\|Y\|_{(\infty,p)}.$$

Note that for a tensor product of matrices this norm fulfills $\|X\otimes Y\|_{(r,p)} = \|X\|_r\|Y\|_p$, which is consistent with the interpretation stated above. Furthermore, we have $\|\cdot\|_{(p,p)} = \|\cdot\|_p$ and $\|\cdot\|_{(r,p)} = \|\cdot\|_p$ on $\mathcal{M}\left(\mathbb{C}\otimes\mathbb{C}^{d_2}\right)\simeq\mathcal{M}_{d_2}$.

For a linear map $\mathcal{T}:\mathcal{M}_{d_1}\to\mathcal{M}_{d_2}$ and $p,q\in[1,\infty]$ the *p-to-q-norm* (see [28]) is given by

$$\|\mathcal{T}\|_{p\to q} = \sup_{X\in\mathcal{M}_{d_1}} \frac{\|\mathcal{T}(X)\|_q}{\|X\|_p}.$$

The p-to-q-norms behave well with respect to composition and it can be shown, that $\|\mathcal{S}\circ\mathcal{T}\|_{p\to q} \leq \|\mathcal{S}\|_{p\to r}\|\mathcal{T}\|_{r\to q}$ for linear maps $\mathcal{T}:\mathcal{M}_{d_1}\to\mathcal{M}_{d_2}$, $\mathcal{S}:\mathcal{M}_{d_2}\to\mathcal{M}_{d_3}$ [28]. In many cases we will deal with the cases $p=q=1$ or $p=q=\infty$. In these cases (which are dual to each other) there is an interesting connection between the above norms and positivity of the linear map $\mathcal{T}$.

Theorem 2.3.1 (Russo-Dye and converse to Russo-Dye, see [25])**.**

(i) For a positive map $\mathcal{T}:\mathcal{M}_{d_1}\to\mathcal{M}_{d_2}$ we have $\|\mathcal{T}\|_{\infty\to\infty} = \|\mathcal{T}(\mathbb{1}_{d_1})\|_\infty$.

(ii) If for a unital linear map $\mathcal{T}:\mathcal{M}_{d_1}\to\mathcal{M}_{d_2}$ we have $\|\mathcal{T}\|_{\infty\to\infty}=1$, then $\mathcal{T}$ is positive.

It is easy to see (by duality of norms [25]), that for a positive map $\mathcal{T}:\mathcal{M}_{d_1}\to\mathcal{M}_{d_2}$ we also have $\|\mathcal{T}\|_{1\to1} = \|\mathcal{T}^*(\mathbb{1}_{d_2})\|_\infty$ and therefore $\|\mathcal{T}\|_{1\to1}=1$ for a trace-preserving positive map $\mathcal{T}$.

Despite all these nice properties the p-to-q norms have some drawbacks when applied to tensor products of linear maps. More specifically we may have $\|\mathcal{S}\otimes\mathcal{T}\|_{p\to q} > \|\mathcal{S}\|_{p\to q}\|\mathcal{T}\|_{p\to q}$ (consider for instance $p = q = 1$ and $\mathcal{S} = \mathrm{id}_2$, $\mathcal{T} = \vartheta_2$) for linear maps $\mathcal{S}$ and $\mathcal{T}$. To resolve this issue one can define *stabilized* versions of the p-to-q-norms. Consider first the case where $p = q$. For $p \in [1,\infty]$ the *completely bounded (CB) p-to-p norm* (see [27]) of $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is given by

$$\|\mathcal{T}\|_{\mathrm{CB},p\to p} = \sup_{n\in\mathbb{N}} \|\mathrm{id}_n \otimes \mathcal{T}\|_{p\to p}. \tag{2.3}$$

In the special case where $p = 1$ the above norm $\|\cdot\|_{\mathrm{CB},1\to 1} = \|\cdot\|_\diamond$ is also called the $\diamond$-*norm*. For $p = \infty$ the norm $\|\cdot\|_{\mathrm{CB},\infty\to\infty} = \|\cdot\|_{\mathrm{CB}}$ is often just called the CB-norm. In finite dimensions the computation of these norms simplifies and for any linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ we have $\|\mathcal{T}\|_\diamond = \|\mathrm{id}_{d_1} \otimes \mathcal{T}\|_{1\to 1}$ (see [12]). By the Russo-Dye theorem we have $\|\mathcal{T}\|_\diamond = 1$ for any quantum channel $\mathcal{T}$. The converse to the Russo-Dye theorem shows that for any trace-preserving linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ that is *not* completely positive we have $\|\mathcal{T}\|_\diamond > 1$.

Note that with the above definition we have $\|\mathcal{S}\otimes\mathcal{T}\|_{\mathrm{CB},p\to p} = \|\mathcal{S}\|_{\mathrm{CB},p\to p}\|\mathcal{T}\|_{\mathrm{CB},p\to p}$ which makes these norms useful for quantum information theory. Also it can be shown that for any completely positive map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ and $p \in [1,\infty]$ we have $\|\mathcal{T}\|_{\mathrm{CB},p\to p} = \|\mathcal{T}\|_{p\to p}$ [27].

With the (p,p)-norms the completely bounded p-to-p can be written as

$$\|\mathcal{T}\|_{CB,p\to p} = \sup_{n\in\mathbb{N}} \|\mathrm{id}_n \otimes \mathcal{T}\|_{(p,p)\to(p,p)}$$

This definition applies a p-to-p norm on the first tensor factor where the n-dimensional identity acts. When $q < p$ we have to be careful as $\|\mathrm{id}_n\|_{p\to q} = n^{\frac{1}{q}-\frac{1}{p}} \to \infty$ as $n \to \infty$. Therefore the naive definition "$\|\mathcal{T}\|_{CB,p\to q} = \sup_{n\in\mathbb{N}} \|\mathrm{id}_n \otimes \mathcal{T}\|_{(p,p)\to(q,q)}$" does not work in general. The *completely bounded p-to-q norm* (see [27]) of a linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is given by

$$\|\mathcal{T}\|_{CB,p\to q} = \sup_{n\in\mathbb{N}} \|\mathrm{id}_n \otimes \mathcal{T}\|_{(s,p)\to(s,q)}$$

with $\|\cdot\|_{(s,p)}$ as defined above and for an arbitrary $s \in [1,\infty]$. It turns out that the choice of $s \in [1,\infty]$ on the first tensor factor is irrelevant and the above norm has the same value in any case [26]. The completely bounded p-to-q-norm has some nice properties. First it is clear that $\|\mathcal{T}\|_{CB,p\to q} \geq \|\mathcal{T}\|_{p\to q}$. The property $\|\mathcal{S}\otimes\mathcal{T}\|_{CB,p\to q} = \|\mathcal{S}\|_{CB,p\to q}\|\mathcal{T}\|_{CB,p\to q}$ holds for completely positive maps $\mathcal{S}$ and $\mathcal{T}$ [27]. For $q \geq p$ and a completely positive map $\mathcal{T}$ we have $\|\mathcal{T}\|_{CB,p\to q} = \|\mathcal{T}\|_{p\to q}$ [27]. Finally as the Riesz-Thorin interpolation theorem holds for the p-to-q-norms (see [29, 30]) it also holds for the completely bounded p-to-q norms [31]:

Theorem 2.3.2 (Riesz-Thorin theorem). *For any linear map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ and $p_0, p_1, q_0, q_1 \in [1, \infty]$ and $\theta \in [0, 1]$ we have*

$$\|\mathcal{T}\|_{CB,p(\theta)\to q(\theta)} \leq \|\mathcal{T}\|^{\theta}_{CB,p_0\to q_0} \|\mathcal{T}\|^{1-\theta}_{CB,p_1\to q_1}$$

for $\frac{1}{p(\theta)} = \frac{\theta}{p_0} + \frac{1-\theta}{p_1}$ *and* $\frac{1}{q(\theta)} = \frac{\theta}{q_0} + \frac{1-\theta}{q_1}$.

2.3.2 Fidelity

The *fidelity* of two quantum states $\rho, \sigma \in \mathcal{D}_d$ is defined as

$$F(\rho, \sigma) = \operatorname{tr}\left[\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\right]. \tag{2.4}$$

Note that $F(\rho, \rho) = 1$ and $F(\rho, \sigma) = 0$ if $\operatorname{supp}(\rho) \cap \operatorname{supp}(\sigma) = \emptyset$. For pure states the above definition simplifies and we have

$$F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\langle\psi|\phi\rangle|$$

for $|\psi\rangle, |\phi\rangle \in \mathbb{C}^d$. We will also write $F(|\psi\rangle, |\phi\rangle)$ for the fidelity of two pure states. The following theorem relates the fidelity with the metric induced by the trace-norm:

Theorem 2.3.3 (Fuchs-van de Graaf-inequalities [32]). *For quantum states $\rho, \sigma \in \mathcal{D}_d$ we have*

$$1 - F(\rho, \sigma) \leq \frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)^2}. \tag{2.5}$$

The fidelity has some properties, that make it a useful quantity in quantum information theory. It can be shown that the fidelity $F(\rho, \sigma)$ can be interpreted in terms of how well ρ and σ can be distinguished by measurements. More specifically for $\rho, \sigma \in \mathcal{D}_d$ we have

$$F(\rho, \sigma) = \min_{\{E_i\}_{i=1}^k} \sum_{i=1}^k \sqrt{\operatorname{tr}[\rho E_i] \operatorname{tr}[\sigma E_i]}$$

where the minimization is over POVMs $\{E_m\}_{i=1}^k \subset \mathcal{M}_d^+$ [11]. It is a simple consequence of this representation (using that adjoints of positive maps are positive) that the application of a positive and trace-preserving map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ two both states $\rho, \sigma \in \mathcal{D}_{d_1}$ increases the fidelity, i.e. $F(\mathcal{P}(\rho), \mathcal{P}(\sigma)) \geq F(\rho, \sigma)$. Furthermore, the fidelity is jointly concave in its two arguments (see [11]), i.e. for quantum states $\rho_i, \sigma_i \in \mathcal{D}_d$ and $p_i \in \mathbb{R}^+$ such that $\sum_{i=1}^k p_i = 1$ we have

$$F(\sum_{i=1}^k p_i\rho_i, \sum_{i=1}^k p_i\sigma_i) \geq \sum_{i=1}^k p_i F(\rho_i, \sigma_i).$$

The fidelity of two quantum states can be expressed as the fidelity of two pure states on a larger system. A *purification* of a quantum state $\rho_A \in \mathcal{D}_{d_A}$ is a pure state $|\psi_{AB}\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ such that $\rho_A = \mathrm{tr}_B\,[|\psi_{AB}\rangle\langle\psi_{AB}|]$ [11].

Theorem 2.3.4 (Uhlmann [33]). *For quantum states $\rho_A, \sigma_A \in \mathcal{D}_{d_A}$ and any $d_B \geq d_A$ we have*

$$F(\rho_A, \sigma_A) = \max_{|\psi_{AB}\rangle, |\phi_{AB}\rangle} F(|\psi_{AB}\rangle, |\phi_{AB}\rangle) = \max_{|\psi_{AB}\rangle, |\phi_{AB}\rangle} |\langle\psi_{AB}|\phi_{AB}\rangle|.$$

where the maxima are over purifications $|\psi_{AB}\rangle, |\phi_{AB}\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ of ρ_A and σ_A respectively.

It can be shown that two pure states $|\psi_{AB}\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ and $|\phi_{AB'}\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_{B'}}$ with $d_B \leq d_{B'}$ are purifications of the same state $\rho_A \in \mathcal{D}_{d_A}$ iff they are related by an isometry $V : \mathbb{C}^{d_B} \to \mathbb{C}^{d_{B'}}$, i.e. $|\phi_{AB'}\rangle = (\mathbb{1}_A \otimes V)|\psi_{AB}\rangle$ [11]. Together with Uhlmann's theorem this implies that for pure states $|\psi_{AB}\rangle, |\phi_{AB}\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ and any isometry $V : \mathbb{C}^{d_B} \to \mathbb{C}^{d_{B'}}$ we have $F((\mathbb{1}_A \otimes V)|\psi_{AB}\rangle, (\mathbb{1}_A \otimes V)|\phi_{AB}\rangle) = F(|\psi_{AB}\rangle, |\phi_{AB}\rangle)$. As an immediate consequence we can write the statement of Uhlmann's theorem for fixed purifications $|\psi_{AB}\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ and $|\phi_{AB'}\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_{B'}}$ of $\rho, \sigma \in \mathcal{D}_{d_A}$ and with $d_B \leq d_{B'}$ as

$$F(\rho, \sigma) = \max_V F((\mathbb{1}_A \otimes V)|\psi_{AB}\rangle, |\phi_{AB'}\rangle). \tag{2.6}$$

Here the maximum is over isometries $V : \mathbb{C}^{d_B} \to \mathbb{C}^{d_{B'}}$.

2.3.3 Relative entropy

Throughout this thesis we will write log for the logarithm to base 2. The *relative entropy* (also known as *quantum Kullback-Leibler-divergence*) of quantum states $\rho, \sigma \in \mathcal{D}_d$ is defined as

$$D(\rho\|\sigma) := \begin{cases} \mathrm{tr}[\rho(\log\rho - \log\sigma)], & \text{if } \mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma) \\ +\infty, & \text{otherwise} \end{cases}. \tag{2.7}$$

The relative entropy fulfills a so called *data-processing inequality*, i.e. for any quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ and quantum states $\rho, \sigma \in \mathcal{D}_{d_1}$ we have $D(T(\rho)\|T(\sigma)) \leq D(\rho\|\sigma)$ [34]. Furthermore it is jointly convex in its arguments (see [35]), i.e.

$$D(\sum_{i=1}^k p_i\rho_i \| \sum_{i=1}^k p_i\sigma_i) \leq \sum_{i=1}^k p_i D(\rho_i\|\sigma_i)$$

for quantum states $\rho_i, \sigma_i \in \mathcal{D}_d$ and $p_i \in \mathbb{R}^+$ with $\sum_{i=1}^k p_i = 1$. The relative entropy also has the following property with respect to tensor-powers: For any quantum states $\rho_1, \sigma_1 \in \mathcal{D}_{d_1}$ and $\rho_2, \sigma_2 \in \mathcal{D}_{d_2}$ we have $D(\rho_1 \otimes \rho_2\|\sigma_1 \otimes \sigma_2) = D(\rho_1\|\sigma_1) + D(\rho_2\|\sigma_2)$. We can lower bound the relative entropy in terms of the trace-distance

Theorem 2.3.5 (Pinsker inequality [36]). *For quantum states $\rho, \sigma \in \mathcal{D}_d$ we have*

$$\frac{1}{2}\|\rho - \sigma\|_1 \leq D(\rho, \sigma).$$

2.4 Information-theoretic quantities

In this section we will introduce some quantities commonly used in quantum information theory.

2.4.1 Von-Neumann entropy and conditional entropy

The *von-Neumann entropy* of a quantum state $\rho \in \mathcal{D}_d$ is defined as $S(\rho) = -\mathrm{tr}\left[\rho \log(\rho)\right]$. It can be shown that $S\left(\rho\right) \in [0, \log(d)]$ for any $\rho \in \mathcal{D}_d$ and that $S(\rho) = 0$ iff ρ is a pure state [11]. Furthermore the entropy is a concave function [11]. Note that the relative entropy can be written as $D(\rho\|\sigma) = -S(\rho) - \mathrm{tr}\left[\rho \log(\sigma)\right]$ and many properties of the von-Neumann entropy are related to properties of the relative entropy.

For a bipartite quantum state $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ we have the *subadditivity inequality* $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ [11]. For a tripartite quantum state $\rho_{ABC} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_C}\right)$ we can use the monotonicity of the relative entropy under quantum channels (more specifically the partial trace) to show $D\left(\rho_{AB}\|\rho_A \otimes \rho_B\right) \leq D\left(\rho_{ABC}\|\rho_A \otimes \rho_{BC}\right)$. By rewriting this inequality in terms of the von-Neumann entropy we obtain the so called *strong subadditivity inequality* $S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$ [11].

The von-Neumann entropy is the quantum version of the Shannon-entropy. As its classical counterpart it has an operational interpretation in a source coding (or data compression) protocol. We will need the following lemma:

Lemma 2.4.1 (Typical subspace theorem [11]). *Let $\rho \in \mathcal{D}_d$ be a quantum state and $\epsilon > 0$ fixed. Then for any $\delta > 0$ there exists an $N \in \mathbb{N}$ such that for any $n \geq N$ there is a projector $P(n, \epsilon)$ with:*

1. $tr\left[P(n, \epsilon)\rho^{\otimes n}\right] \geq 1 - \delta$
2. $(1 - \delta)\, 2^{n(S(\rho) - \epsilon)} \leq tr\left[P(n, \epsilon)\right] \leq 2^{n(S(\rho) + \epsilon)}$

The subspace $Im\left(P(n, \epsilon)\right) \subset \mathbb{C}^{d^n}$ is also called ϵ-typical subspace.

As a direct consequence for any fixed $\epsilon > 0$, some $\delta > 0$ and n sufficiently large we can write

$$\rho^{\otimes n} = p_n \rho^n_{\mathrm{typ}} + (1 - p_n)\rho^n_{\mathrm{rest}}$$

with $p_n = \mathrm{tr}\left[P(n, \epsilon)\rho^{\otimes n}\right]$, a state ρ^n_{typ} supported on the typical subspace and a state ρ^n_{rest} with no support on the typical subspace. Note that for $S\left(\rho\right) \ll \log(d)$ the dimension of the typical subspace $d_{\text{typ}} = \operatorname{tr}[P(n,\epsilon)] \leq 2^{n(S(\rho)+\epsilon)} \ll d^n$ will be much smaller than the dimension of the full space. As $\|\rho^{\otimes n} - \rho^n_{\text{typ}}\|_1 \leq 2(1-p_n) \leq 2\delta$ we can (for δ small enough) compress $\rho^{\otimes n}$ to ρ^n_{typ} without introducing much error. This compression can be performed by measuring the 2-outcome POVM given by $\{P(n,\epsilon), (\mathbb{1}_{d^n} - P(n,\epsilon))\}$. This will successfully compress $\rho^{\otimes n}$ with probability $p_n \geq 1-\delta$, which is close to 1 when $\delta > 0$ is chosen small enough. It can also be shown that it is not possible to compress the state further without introducing large errors in the limit $n \to \infty$, i.e. the von-Neumann entropy is the optimal upper bound for the compression rate of a quantum state in the asymptotic limit of many copies [11].

Similar to the classical case where one can consider the Shannon entropy of a conditional probability distribution one can define the conditional entropy in the quantum case. Given a bipartite quantum state $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ the *conditional (von-Neumann) entropy of system A given system B* is defined as $S(A|B)_{\rho_{AB}} := S(\rho_{AB}) - S(\rho_B)$. This quantity can be interpreted as the information that an observer can gain about system A when allowed access to system B. It is a well-known fact, that due to entanglement of ρ_{AB} the conditional entropy $S(A|B)_{\rho_{AB}}$ can be negative unlike its classical analogue [37]. For a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ and a quantum state $\sigma_{RA} \in \mathcal{D}\left(\mathbb{C}^{d_R} \otimes \mathbb{C}^{d_A}\right)$ with $d_R = d_A$ we define $I^{\text{coh}}(\sigma_{RA}, \mathcal{T}) := -S(R|B)_{(\text{id}_R \otimes \mathcal{T})(\sigma_{RA})}$. Then the *coherent information of* $\mathcal{T}$ is given by

$$I^{\text{coh}}(\mathcal{T}) := \max_{\sigma_{RA} \in \mathcal{D}\left(\mathbb{C}^{d_R} \otimes \mathbb{C}^{d_A}\right)} I^{\text{coh}}(\sigma_{RA}, \mathcal{T})$$

The coherent information quantifies how much information can be transmitted through the quantum channel.

2.4.2 Conditional min-entropy

The von-Neumann entropy and the quantities derived from it are useful in scenarios involving asymptotic limits of many copies of the same object (usually quantum states or quantum channels). They seem to be less useful in one-shot scenarios where only one instance of an object is considered. To analyze such scenarios many useful quantities have been introduced (see for instance [38, 39]). We will only use one such quantity namely the conditional min-entropy first introduced in [38].

For a bipartite positive and subnormalized matrix $X_{AB} \in \mathcal{M}^+\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$, i.e. with $\operatorname{tr}[X] \leq 1$, the *conditional min-entropy of A given B* (see [39]) is defined as

$$S_{\min}(A|B)_{X_{AB}} := -\log\left(\min\{\operatorname{tr}[Y_B] \ : \ Y_B \in \mathcal{M}^+_{d_B}, X_{AB} \leq \mathbb{1}_{d_A} \otimes Y_B\}\right).$$

The optimization in the conditional min-entropy is a semidefinite program and hence this quantity can be efficiently computed [39]. For any matrix $X_{AB} \in \mathcal{M}^+\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ with $\mathrm{tr}\left[X_{AB}\right] \leq 1$ we have $X_{AB} \leq \mathbb{1}_{d_A} \otimes \mathbb{1}_{d_B}$, which implies $S_{\min}(A|B)_{X_{AB}} \geq -\log(d_B)$. It can also be shown (see [39]) that for any subnormalized $X_{AB} \in \mathcal{M}^+\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ and any isometries $V : \mathbb{C}^{d_A} \to \mathbb{C}^{d_{A'}}$ and $W : \mathbb{C}^{d_B} \to \mathbb{C}^{d_{B'}}$ we have

$$S_{\min}(A'|B')_{(V\otimes W)X_{AB}(V\otimes W)} = S_{\min}(A|B)_{X_{AB}}.$$

The conditional min-entropy of a quantum state has an operational interpretation as the maximum achievable overlap with a maximally entangled state [40]. More specifically for $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ with $d_A \leq d_B$ we have

$$2^{-S_{\min}(A|B)_{\rho_{AB}}} = d_A \max_{\mathcal{T}} F\left(\left(\mathrm{id}_A \otimes \mathcal{T}\right)\left(\rho_{AB}\right), \omega_{AB}\right)$$

where the maximum is over quantum channels $\mathcal{T} : \mathcal{M}_{d_B} \to \mathcal{M}_{d_B}$.

For $\epsilon > 0$ the *ϵ-smooth conditional min-entropy of A given B* (see [39]) of a bipartite quantum state $\rho \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ is defined as

$$S^{\epsilon}_{\min}(A|B)_{\rho_{AB}} = \max_{X \in B(\rho_{AB},\epsilon)} S_{\min}(A|B)_{X_{AB}}$$

where $B(\rho, \epsilon) = \{X_{AB} \in \mathcal{M}^+\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right) \; : \; \mathrm{tr}\left[X\right] \leq 1, \sqrt{1 - F(\rho, X)^2} \leq \epsilon\}$ denotes an ϵ-ball around ρ in the *purified distance* (where the fidelity $F(\rho, X)$ for subnormalized X can be defined as in (2.4)). By the above lower bound for $S_{\min}(A|B)_{X_{AB}}$ for subnormalized and positive X_{AB} we also have

$$S^{\epsilon}_{\min}(A|B)_{\rho_{AB}} \geq -\log(d_B) \tag{2.8}$$

for any state $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$. The invariance under local isometries mentioned above for $S_{\min}$ also holds for the smooth version, i.e. for any $\epsilon > 0$, any state $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ and any isometries $V : \mathbb{C}^{d_A} \to \mathbb{C}^{d_{A'}}$ and $W : \mathbb{C}^{d_B} \to \mathbb{C}^{d_{B'}}$ we have

$$S^{\epsilon}_{\min}(A'|B')_{(V\otimes W)\rho_{AB}(V\otimes W)} = S^{\epsilon}_{\min}(A|B)_{\rho_{AB}}. \tag{2.9}$$

When the ϵ-smooth conditional min-entropy is applied to tensor-powers of a quantum state the usual (von-Neumann) conditional entropy can be recovered as a limit [39]. More specifically for a quantum state $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ we have for any $\epsilon \in (0, 1)$ that

$$\lim_{n\to\infty} S^{\epsilon}_{\min}(A^n|B^n)_{\rho^{\otimes n}_{AB}} = S(A|B)_{\rho_{AB}}.$$

We will not need this property directly, but only the following bound known as *asymptotic equipartition property*:

Theorem 2.4.1 (Fully quantum AEP, [41]). *For a density operator $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ and $\epsilon > 0$ there exists a sequence $\Delta(n, \rho_{AB}, \epsilon) \to 0$ as $n \to \infty$ and such that for all $n \geq \frac{8}{5}\log\left(\frac{2}{\epsilon^2}\right)$ we have*

$$\frac{1}{n} S^{\epsilon}_{min}(A^n|B^n)_{\rho_{AB}^{\otimes n}} \geq S(A|B)_{\rho_{AB}} - \Delta(n, \rho_{AB}, \epsilon) \ . \tag{2.10}$$

2.5 Quantum capacities

In this section we will introduce two capacities describing the transmission of quantum information over quantum channels. First we will talk about the quantum analogue of the Shannon capacity. In the second part we will introduce the two-way quantum capacity where the transmission is assisted by arbitrary classical communication between the sender and the receiver. Note that there are many other capacities and communication settings not mentioned here describing for instance the transmission of classical information or using different channel models (e.g. scenarios with more than one in- and/or output-systems, see [42] or channels varying in time). Also we will assume that the sender and receiver know the the quantum channel used for transmission exactly. This might not be the case in realistic scenarios.

2.5.1 Quantum Shannon-capacity

Definition 2.5.1 (Quantum capacity, see [43]). *For a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ its* quantum capacity *is defined as*

$$\mathcal{Q}(\mathcal{T}) = \sup\{r \in \mathbb{R}^+ \ : \ r \textit{ achievable rate } \}.$$

A number $r \in \mathbb{R}^+$ is called achievable rate iff there are sequences $(n_\nu)_{\nu\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$ and $(m_\nu)_{\nu\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$ with $n_\nu \to \infty$ as $\nu \to \infty$ and $r = \limsup_{n\to\infty} \frac{n_\nu}{m_\nu}$ such that

$$\inf_{\mathcal{E},\mathcal{D}} \left\| id_2^{\otimes n_\nu} - \mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E} \right\|_\diamond \to 0 \textit{ as } \nu \to \infty \tag{2.11}$$

where the infimum is over encoding quantum channels $\mathcal{E} : \mathcal{M}_2^{\otimes n_\nu} \to \mathcal{M}_{d_A}^{\otimes m_\nu}$ and decoding quantum channels $\mathcal{D} : \mathcal{M}_{d_B}^{\otimes m_\nu} \to \mathcal{M}_2^{\otimes n_\nu}$.

Note that by choosing the identity channel in (2.11) to be two-dimensional we take the qubit as the unit of quantum information. By choosing a different dimension d we obtain essentially the same capacity just scaled by a factor of $1/\log_2(d)$ [43]. It can be shown (see [43]) that a rate $r > 0$ is achievable iff the communication error in (2.11) vanishes for the sequences $n_\nu = r\nu$ and $m_\nu = \nu$.

There are many equivalent definitions for the quantum capacity of a quantum channel using different error-measures. Instead of the $\diamond$-norm it is sometimes useful to consider the so called *channel fidelity* (see for instance [43]) of a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ defined as

$$F_c(\mathcal{T}) := F((\mathrm{id}_R \otimes \mathcal{T})(\omega_{RA}), \omega_{RB})^2 \ . \tag{2.12}$$

This error-measure quantifies how much entanglement is transmitted through $\mathcal{T}$. Requiring $\sup_{\mathcal{E},\mathcal{D}} F_c\left(\mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E}\right) \to 1$ as $\nu \to \infty$ instead of (2.11) to express the vanishing communication error yields an equivalent definition of the capacity. For convenience we will summarize the argument given in [43] to show this fact. The following theorem is a direct consequence of [43, Proposition 4.5].

Theorem 2.5.1 ([43])**.** *Let $\mathcal{T} : \mathcal{M}_d \to \mathcal{M}_d$ and $k \leq d$. Then there exist quantum channels $\mathcal{V}_k : \mathcal{M}_k \to \mathcal{M}_d$ and $\mathcal{S}_k : \mathcal{M}_d \to \mathcal{M}_k$ such that*

$$\left\| id_k - \mathcal{S}_k \circ \mathcal{T} \circ \mathcal{V}_k \right\|_\diamond \leq 8 \left(\frac{1 - F_c\left(\mathcal{T}\right)}{1 - \frac{k}{d}} \right)^{1/4} \ .$$

With this theorem it is possible to see, that the capacity does not change when the $\diamond$-norm is replaced by the channel fidelity to quantify the communication error in (2.11). Assume that there is a rate $r > 0$ and sequences $(n_\nu)_\nu$ and $(m_\nu)_\nu$ with $r = \limsup_{\nu\to\infty} \frac{n_\nu}{m_\nu}$ such that for a sequence of encodings $\mathcal{E}_\nu : \mathcal{M}_{2^{n_\nu}} \to \mathcal{M}_{d^{m_\nu}}$ and decodings $\mathcal{D}_\nu : \mathcal{M}_{d^{m_\nu}} \to \mathcal{M}_{2^{n_\nu}}$ we have $F_c\left(\mathcal{D}_\nu \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E}_\nu\right) \to 1$ as $\nu \to \infty$. Now applying the above theorem for each $k_\nu = 2^{n_\nu - 1}$ yields modified encodings $\tilde{\mathcal{E}}_\nu = \mathcal{E}_\nu \circ \mathcal{V}_\nu$ and modified decodings $\tilde{\mathcal{D}}_\nu = \mathcal{S}_\nu \circ \mathcal{D}_\nu$ such that $\left\| \mathrm{id}_2^{\otimes(n_\nu - 1)} - \tilde{\mathcal{D}}_\nu \circ \mathcal{T}^{\otimes m_\nu} \circ \tilde{\mathcal{E}}_\nu \right\|_\diamond \to 0$ as $\nu \to \infty$. As $\limsup_{\nu\to\infty} \frac{n_\nu - 1}{m_\nu} = r$ this rate is also achievable when using the $\diamond$-norm as in (2.11) to quantify the communication error. The other direction that any rate achievable with respect to the $\diamond$-norm is also achievable with respect to the channel fidelity is straightforward [43].

From the above definition it is still unclear how to compute the quantum capacity for a concrete channel. The following theorem relates the above capacity to the coherent information introduced before:

Theorem 2.5.2 (LSD-theorem [44, 45, 46])**.** *For a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ we have*

$$\mathcal{Q}\left(\mathcal{T}\right) = \lim_{n\to\infty} \frac{1}{n} I^{coh}\left(\mathcal{T}^{\otimes n}\right) .$$

The limit over tensor-powers in the above theorem, also called *regularization*, is essential and for any $n \in \mathbb{N}$ there are examples of quantum channels $\mathcal{T}$ such that $I^{\text{coh}}\left(\mathcal{T}^{\otimes n}\right) = 0$ but $I^{\text{coh}}\left(\mathcal{T}^{\otimes (n+1)}\right) > 0$ [47]. At the time of writing it is *not* known whether there is a formula for the quantum capacity possibly involving other quantities and avoiding a regularization.

Even with Theorem 2.5.2 computing the quantum capacity for a given quantum channel is hard in general, but in some cases the computation simplifies. A quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ with complementary channel $\mathcal{T}^c : \mathcal{M}_{d_A} \to \mathcal{M}_{d_E}$ is called *degradable* (see [17]) iff there is a quantum channel $\mathcal{S} : \mathcal{M}_{d_B} \to \mathcal{M}_{d_E}$ such that $\mathcal{S} \circ \mathcal{T} = \mathcal{T}^c$. For degradable quantum channels the regularization can be removed in Theorem 2.5.2 which simplifies the computation of the capacity. A quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ with complementary channel $\mathcal{T}^c : \mathcal{M}_{d_A} \to \mathcal{M}_{d_E}$ is called *anti-degradable* (see for instance [48]) iff there exists a quantum channel $\mathcal{S} : \mathcal{M}_{d_E} \to \mathcal{M}_{d_B}$ such that $\mathcal{S} \circ \mathcal{T}^c = \mathcal{T}$. It can be shown that anti-degradable channels have vanishing quantum capacity. Note that every entanglement-breaking channel is also anti-degradable [48].

We finish this section by stating the well-known transposition bound on the quantum capacity. Let $\vartheta_d : \mathcal{M}_d \to \mathcal{M}_d$ denote a matrix transposition in any fixed basis.

Theorem 2.5.3 (Transposition bound [49])**.** *For a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ we have*

$$\mathcal{Q}\left(\mathcal{T}\right) \leq \log_2(\|\vartheta_{d_B} \circ \mathcal{T}\|_\diamond) \ .$$

Using the properties of $\|\cdot\|_\diamond$ the transposition bound implies that any completely co-positive quantum channel (i.e. a quantum channel $\mathcal{T}$ for which $\vartheta \circ \mathcal{T}$ is also a quantum channel) has vanishing quantum capacity. At the time of writing the completely co-positive quantum channels and the anti-degradable channels are the only classes of quantum channels for which the quantum capacity is known to vanish.

2.5.2 The decoupling approach

In this section we will illustrate how to show the existence of suitable coding channels in Definition 2.5.1 of the quantum capacity using the decoupling approach. While this method can be used to prove Theorem 2.5.2 in its full generality (see [50, 51]) we will only show that for a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ any rate $r < I_{\text{coh}}\left(\omega_{A'A}, \mathcal{T}\right)$ is achievable. This is sufficient to understand the articles contained in this dissertation and the argument will be slightly easier while equally illuminating. The following discussion is based on results stated in [51].

The idea of the decoupling approach can be described using the following lemma, which is a consequence of Uhlmann's theorem.

Lemma 2.5.1 ([51]). *Let $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ be a quantum channel with Stinespring isometry $V^{A\to BE} : \mathbb{C}^{d_A} \to \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$ and a purification of its Choi-matrix is given by $|\sigma_{RBE}\rangle = (\mathbb{1}_R \otimes V^{A\to BE})|\Omega_{RA}\rangle$. If*

$$\left\| (id_R \otimes \mathcal{T}^c)(\omega_{RA}) - \pi_R \otimes \sigma_E \right\|_1 \leq \epsilon \ ,$$

then there exists a quantum channel $\mathcal{D} : \mathcal{M}_{d_B} \to \mathcal{M}_{d_{A'}}$ for $d_A = d_{A'}$ such that

$$F_c(\mathcal{D} \circ \mathcal{T}) \geq \left(1 - \frac{\epsilon}{2}\right)^2 .$$

Proof. Using the Fuchs-van de Graaf (Theorem 2.3.3) inequalities we have:

$$F((\mathrm{id}_R \otimes T^c)(\omega_{RA}), \pi_R \otimes \sigma_E) \geq 1 - \frac{\epsilon}{2}.$$

Note that a purification of $\pi_R \otimes \sigma_E$ is given by $|\Omega_{RA'}\rangle \otimes |\sigma_{R'B'E}\rangle \in \mathbb{C}^{d_R d_{A'}} \otimes \mathbb{C}^{d_{R'} d_{B'} d_E}$ and also a purification of $(\mathrm{id}_R \otimes T^c)(\omega_{RA})$ is given by $|\sigma_{RBE}\rangle \in \mathbb{C}^{d_R} \otimes \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$. By applying Uhlmann's theorem (in the form (2.6)) there is an isometry $W^{B\to A'R'B'} : \mathbb{C}^{d_B} \to \mathbb{C}^{d_{A'}} \otimes \mathbb{C}^{d_{R'}} \otimes \mathbb{C}^{d_{B'}}$ such that

$$\begin{aligned} 1 - \frac{\epsilon}{2} &\leq F((\mathbb{1}_R \otimes W^{B\to A'R'B'} \otimes \mathbb{1}_E)|\sigma_{RBE}\rangle, |\omega_{RA'}\rangle \otimes |\sigma_{R'B'E}\rangle) \\ &\leq F((\mathrm{id}_{d_R} \otimes \mathcal{D} \circ \mathcal{T})(\omega_{RA}), \omega_{RA'}). \end{aligned}$$

In the last step we used monotonicity of the fidelity under quantum channels (in this case partial traces) and introduced the quantum channel $\mathcal{D} : \mathcal{M}_{d_B} \to \mathcal{M}_{d_{A'}}$ as

$$\mathcal{D}(\rho) = \mathrm{tr}_{R'B'}\left[W^{B\to A'R'B'} \rho \left(W^{B\to A'R'B'} \right)^\dagger \right].$$

□

By the above lemma a decoding channel $\mathcal{D}$ achieving high channel fidelity exists when after information transmission the environment (labeled by 'E') is decoupled from the reference system (labeled by 'R'). Surprisingly, by choosing a random encoding prior to transmission this can be ensured. The following theorem is also called *decoupling theorem*:

Theorem 2.5.4 (Decoupling Theorem [51]). *For a quantum state $\rho_{RA} \in \mathcal{D}\left(\mathbb{C}^{d_R} \otimes \mathbb{C}^{d_A}\right)$ and a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_E}$ with Choi matrix $\sigma_{A'E} = (id_{A'} \otimes \mathcal{T})(\omega_{A'A})$ and an arbitrary $\epsilon > 0$ we have*

$$\int_{\mathfrak{U}(d_A)} \left\| (id_R \otimes \mathcal{T} \circ \mathcal{U})(\rho_{RA}) - \rho_R \otimes \sigma_E \right\|_1 d\mathcal{U} \leq 2^{-\frac{1}{2} S^{\epsilon}_{min}(A'|E)_\sigma - \frac{1}{2} S^{\epsilon}_{min}(A|R)_\rho} + 12\epsilon$$

where the integral is with respect to the Haar measure on the group of unitary matrices $\mathfrak{U}(d_A) \subset \mathcal{M}_{d_A}$.

To apply the decoupling approach consider a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ with Stinespring isometry $V^{A\to BE} : \mathbb{C}^{d_A} \to \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$ and purified Choi-matrix $|\sigma_{A'BE}\rangle = (\mathbb{1}_{A'} \otimes V^{A\to BE})|\Omega_{A'A}\rangle$. For any rate $0 < r < I_{\text{coh}}\left(\omega_{A'A}, \mathcal{T}\right)$ let $\omega_{RR'} \in \mathcal{D}\left(\mathbb{C}^{d_R} \otimes \mathbb{C}^{d_{R'}}\right)$ be the maximally entangled state for $d_R = d_{R'} = 2^{r\nu}$. To show that r is achievable in the sense of Definition 2.5.1 we set $\rho_{RA^\nu} = (\text{id}_R \otimes \mathcal{W}_\nu)\left(\omega_{RR'}\right)$ for isometric channels $\mathcal{W}_\nu(\cdot) = W_\nu \cdot W_\nu^\dagger$ with isometries $W_\nu : \mathbb{C}^{d_{R'}} \to \mathbb{C}^{d_A^\nu}$. Note that by (2.9) and (2.8) we have

$$S^\epsilon_{\min}\left(A^\nu|R\right)_{\rho_{RA^\nu}} = S^\epsilon_{\min}\left(R'|R\right)_{\omega_{RR'}} \geq -\log(d_R).$$

Now consider the quantum channel $(\mathcal{T}^c)^{\otimes\nu}$ and note that the Choi-matrix of this channel is given by $\sigma^{\otimes\nu}_{A'E}$. Using the asymptotic equipartition property (see Theorem 2.4.1) we have

$$S^\epsilon_{\min}\left((A')^\nu|E^\nu\right)_{\sigma^{\otimes\nu}_{A'E}} \geq S(A'|E)_{\sigma_{A'E}} - \Delta(\nu, \sigma_{A'E}, \epsilon) = I_{\text{coh}}\left(\omega_{A'A}, \mathcal{T}\right) - \Delta(\nu, \sigma_{A'E}, \epsilon).$$

Applying Theorem 2.5.4 with $\epsilon = \frac{1}{\nu}$ implies the existence of unitary channels $\mathcal{U}_\nu \in \mathfrak{U}_{d_A^\nu}$ such that

$$\left\| \left(\text{id}_R \otimes (\mathcal{T}^c)^{\otimes\nu} \circ \mathcal{U}_\nu \circ \mathcal{W}_\nu\right)\left(\rho_{RR'}\right) - \rho_R \otimes \sigma_E^{\otimes\nu}\right\|_1 \leq \sqrt{d_R}2^{-\frac{1}{2}I_{\text{coh}}(\omega_{A'A},\mathcal{T})+\frac{1}{2}\Delta(\nu,\sigma_{A'E},\epsilon)} + \frac{12}{\nu}.$$

Finally we can use the Lemma 2.5.1 to obtain quantum channels $\mathcal{D}_\nu : \mathcal{M}_{d_B^\nu} \to \mathcal{M}_{2^{r\nu}}$ such that

$$F_c(\mathcal{D}_\nu \circ \mathcal{T}^{\otimes\nu} \circ \mathcal{U}_\nu \circ \mathcal{W}_\nu) \geq \left(1 - \frac{1}{2}\left(2^{\frac{\nu}{2}(r - I_{\text{coh}}(\omega_{A'A},\mathcal{T}) + \Delta(\nu,\sigma_{A'E},\epsilon))} + \frac{12}{\nu}\right)\right)^2.$$

As $r < I_{\text{coh}}\left(\omega_{A'A}, \mathcal{T}\right)$ we have that $F_c(\mathcal{D}_\nu \circ \mathcal{T}^{\otimes\nu} \circ \mathcal{U}_\nu \circ \mathcal{W}_\nu) \to 1$ in the limit $\nu \to \infty$. By the discussion after Theorem 2.5.1 we can slightly modify the above coding scheme such that the communication error (2.11) in $\diamond$-norm vanishes in the limit $\nu \to \infty$. This proves that any such rate is an achievable rate for the quantum capacity.

2.5.3 Quantum capacity assisted by classical communication

The quantum capacity introduced in the previous section only describes a very basic scenario of information transmission, which often does not capture all resources that the sender and the receiver could use. In this section we introduce another quantum capacity taking into account that the sender and receiver might be able to exchange arbitrary classical information for free. For more information concerning this and similar scenarios see [52]. It should be emphasized that we restrict to the idealized case where the quantum channel used for transmission is known to both communicating parties.

In the following we consider a bipartite system with its two parts labeled 'A' and 'B'. The 'A' part should be thought of as a system controlled by the *sender*, and in the same way the 'B' part is a system controlled by the *receiver*. We introduce essentially classical systems both at the sending and the receiving side where the two parties can store classical information. To distinguish the quantum- and the classical part of our states we will use labels A_q, B_q to refer to the quantum part at A or B and labels A_c, B_c for the classical part.

We will start with a central definition of the special class of quantum channels that can be implemented using local quantum channels and arbitrary classical communication. This set can be found in similar form in [53] where its mathematical properties are discussed.

Definition 2.5.2 (Local operations and classical communication (LOCC) [53])**.** *A quantum channel* $\mathcal{T} : \mathcal{M}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right) \rightarrow \mathcal{M}\left(\mathbb{C}^{d_{A'}} \otimes \mathbb{C}^{d_{B'}}\right)$ *is called an* LOCC-channel with respect to the bipartition $A : B$ *if it can be written as a sequential concatenation of any number of channels* $\mathcal{L} : \mathcal{M}\left(\mathbb{C}^{d_{A_q} d_{A_c}} \otimes \mathbb{C}^{d_{B_q} d_{B_c}}\right) \rightarrow \mathcal{M}\left(\mathbb{C}^{d_{A'_q} d_{A'_c}} \otimes \mathbb{C}^{d_{B'_q} d_{B'_c}}\right)$ *of the following form* $\big($*for* $X_{A_q A_c B_q B_c} \in \mathcal{M}\left(\mathbb{C}^{d_{A_q} d_{A_c}} \otimes \mathbb{C}^{d_{B_q} d_{B_c}}\right)\big)$*:*

$$\mathcal{L}(X_{A_q A_c B_q B_c}) = \sum_{i,j} (K_i^A \otimes K_j^B) X_{A_q A_c B_q B_c} (K_i^A \otimes K_j^B)^\dagger \otimes |j\rangle\langle j|_{A'_c} \otimes |i\rangle\langle i|_{B'_c}, \tag{2.13}$$

where $K_i^A : \mathbb{C}^{d_{A_q} d_{A_c}} \rightarrow \mathbb{C}^{d_{A'_q}}$ *and* $K_j^B : \mathbb{C}^{d_{B_q} d_{B_c}} \rightarrow \mathbb{C}^{d_{B'_q}}$ *are Kraus operators of quantum channels mapping system* $A_q A_c$ *to* A'_q *and system* $B_q B_c$ *to* B'_q *respectively (i.e.* $\sum_i (K_i^A)^\dagger K_i^A = \mathbb{1}_{A_q A_c}$ *and* $\sum_j (K_j^B)^\dagger K_j^B = \mathbb{1}_{B_q B_c}$*), and* $|j\rangle_{A'_c}$ *and* $|i\rangle_{B'_c}$ *are orthonormal bases belonging to (effectively classical) systems* A_c *and* B_c*.*

As the definition of LOCC-channels is complicated it is often helpful to consider the following class of quantum channels, which contains the LOCC-channels as a proper subset [53]:

Definition 2.5.3 (Separable operations [53])**.** *A quantum channel* $\mathcal{T} : \mathcal{M}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right) \rightarrow \mathcal{M}\left(\mathbb{C}^{d_{A'}} \otimes \mathbb{C}^{d_{B'}}\right)$ *is called* separable *iff it can be written as*

$$\mathcal{T}(X) = \sum_i \left(K_i^A \otimes K_i^B\right) X \left(K_i^A \otimes K_i^B\right)^\dagger$$

for matrices $\{K_i^A\}_i \subset \mathbb{C}^{d_{A'} \times d_A}$ *and* $\{K_i^B\} \subset \mathbb{C}^{d_{B'} \times d_B}$.

Separable quantum channels are easier to analyze than LOCC quantum channels. Nevertheless, they still have enough structure for interesting properties to hold; examples of this can be found in the articles contained in this dissertation. With the previous definitions we can define a capacity assisted by unlimited classical two-way communication.

Definition 2.5.4 (Quantum capacity assisted by unlimited 2-way classical communication [52]). *For a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ the* 2-way classical communication assisted capacity *is defined as*

$$\mathcal{Q}_2(\mathcal{T}) = \sup\{r \in \mathbb{R}^+ \; : \; r \text{ achievable rate}\}.$$

A number $r \in \mathbb{R}^+$ is called achievable rate iff there are sequences $(n_\nu)_{\nu\in\mathbb{N}} \in \mathbb{N}^\mathbb{N}$ and $(m_\nu)_{\nu\in\mathbb{N}} \in \mathbb{N}^\mathbb{N}$ with $r = \limsup_{\nu\to\infty} \frac{n_\nu}{m_\nu}$ and such that

$$\inf_{\mathcal{L}_1,\ldots,\mathcal{L}_{m_\nu+1}} \Big\| id_2^{\otimes n_\nu} - \mathcal{L}_{m_\nu+1} \circ (id_{A_c} \otimes \mathcal{T}_{A_1\to B_{m_\nu}} \otimes id_{B_cB_1\cdots B_{m_\nu-1}}) \circ \cdots$$
$$\cdots \circ \mathcal{L}_2 \circ (id_{A_cA_1\cdots A_{m_\nu-1}} \otimes \mathcal{T}_{A_{m_\nu}\to B_1} \otimes id_{B_c}) \circ \mathcal{L}_1 \Big\|_\diamond \to 0$$

as $\nu \to \infty$. In each step the channel $\mathcal{T}$ brings one system of size d_A from the 'A' side to the 'B' side. The infimum is over LOCC-quantum channels

$$\mathcal{L}_i : \mathcal{M}\left(\mathbb{C}^{d_{A_c}d_{A_1}\cdots d_{A_{m_\nu-i+1}}} \otimes \mathbb{C}^{d_{B_c}d_{B_1}\cdots d_{B_{i-1}}}\right) \to \mathcal{M}\left(\mathbb{C}^{d_{A'_c}d_{A_1}\cdots d_{A_{m_\nu-i+1}}} \otimes \mathbb{C}^{d_{B'_c}d_{B_1}\cdots d_{B_{i-1}}}\right)$$

for $1 < i \leq m_\nu$ and

$$\mathcal{L}_1 : \mathcal{M}\left(\mathbb{C}^{2^{n_\nu}}\right) \to \mathcal{M}\left(\mathbb{C}^{d_{A_c}d_{A_1}\cdots d_{A_{m_\nu}}} \otimes \mathbb{C}^{d_{B_c}}\right)$$

and

$$\mathcal{L}_{m_\nu+1} : \mathcal{M}\left(\mathbb{C}^{d_{A_c}} \otimes \mathbb{C}^{d_{B_c}d_{B_1}\cdots d_{B_{m_\nu}}}\right) \to \mathcal{M}\left(\mathbb{C}^{2^{n_\nu}}\right)$$

where the label c indicates classical systems (as explained above) of arbitrary size and $d_{A_k} = d_A$ and $d_{B_l} = d_B$ for all k, l.

The above capacity is connected to the task of entanglement distillation, i.e. the generation of a maximally entangled state from many copies of a given state. We have the following precise definition:

Definition 2.5.5 (Entanglement distillation [4]). *A bipartite quantum state $\rho_{AB} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ is called* distillable *iff*

$$\inf_{\mathcal{L}_n} \left\| \omega_2 - \mathcal{L}\left(\rho_{AB}^{\otimes n}\right) \right\|_1 \to 0$$

with infimum over LOCC-quantum channels $\mathcal{L}_n : \mathcal{M}\left(\mathbb{C}^{d_A^n} \otimes \mathbb{C}^{d_B^n}\right) \to \mathcal{M}\left(\mathbb{C}^2 \otimes \mathbb{C}^2\right)$.

It can be shown that a quantum channel $\mathcal{T} : \mathcal{M}_{d_A} \to \mathcal{M}_{d_B}$ has $\mathcal{Q}_2(\mathcal{T}) > 0$ iff its Choi-matrix $C_\mathcal{T} \in \mathcal{D}\left(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}\right)$ is distillable. By using properties of the transposition map it can be shown that PPT states are not distillable [4]. As there are entangled PPT states this leads to the surprising fact that there are entangled states, which are not distillable. Such states are called *bound entangled states*. As PPT states are not distillable it is natural to ask, whether conversely all NPPT states are distillable. This problem of the existence of an NPPT bound entangled state is often referred to as the *NPPT bound entanglement problem*. It is despite considerable effort [5, 54, 55, 56] still an open problem.

Bibliography

[1] Juan Pablo Paz and Wojciech Hubert Zurek. **Continuous error correction**. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **454**(1969):355–364, 1998. 2

[2] Fernando Pastawski, Lucas Clemente, and J. Ignacio Cirac. **Quantum memories based on engineered dissipation**. *Physical Review A*, **83**(1):012304, 2011. 2

[3] Michael Ben-Or, Daniel Gottesman, and Avinatan Hassidim. **Quantum refrigerator**. *arXiv preprint arXiv:1301.1995*, 2013. 2

[4] Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. **Mixed-state entanglement and distillation: is there a "bound" entanglement in nature?** *Physical Review Letters*, **80**(24):5239, 1998. 3, 27

[5] David P. DiVincenzo, Peter W. Shor, John A. Smolin, Barbara M. Terhal, and Ashish V. Thapliyal. **Evidence for bound entangled states with negative partial transpose**. *Physical Review A*, **61**(6):062312, 2000. 3, 27

[6] Lenka Moravčíková and Mário Ziman. **Entanglement-annihilating and entanglement-breaking channels**. *Journal of Physics A: Mathematical and Theoretical*, **43**(27):275306, 2010. 3

[7] Isaac Kim and Mary Beth Ruskai. **Bounds on the concavity of quantum entropy**. *Journal of Mathematical Physics*, **55**(9):–, 2014. 4

[8] Dorit Aharonov, Michael Ben-Or, Russell Impagliazzo, and Noam Nisan. **Limitations of noisy reversible computation**. *arXiv preprint quant-ph/9611028*, 1996. 4

[9] Nikolai Nikolski. **Condition numbers of large matrices, and analytic capacities**. *St. Petersburg Mathematical Journal*, **17**(4):641–682, 2006. 5

[10] Tuen Wai Ng and Chiu Yin Tsang. **Chebyshev–Blaschke products: Solutions to certain approximation problems and differential equations**. *Journal of Computational and Applied Mathematics*, **277**:106–114, 2015. 5

[11] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010. 5, 7, 13, 16, 17, 18, 19

[12] Vern Paulsen. *Completely bounded maps and operator algebras*, **78**. Cambridge University Press, 2002. 9, 15

[13] Man-Duen Choi. **Completely positive linear maps on complex matrices**. *Linear Algebra and its Applications*, **10**(3):285 – 290, 1975. 10

[14] Karl Kraus. *States, effects and operations*. Springer, 1983. 10

[15] Rajendra Bhatia. *Positive definite matrices*. Princeton University Press, 2009. 10

[16] W. Forrest Stinespring. **Positive functions on C*-algebras**. *Proceedings of the American Mathematical Society*, **6**(2):211–216, 1955. 11

[17] Igor Devetak and Peter W. Shor. **The capacity of a quantum channel for simultaneous transmission of classical and quantum information**. *Communications in Mathematical Physics*, **256**(2):287–303, 2005. 11, 23

[18] Michael Horodecki, Peter W. Shor, and Mary Beth Ruskai. **Entanglement breaking channels**. *Reviews in Mathematical Physics*, **15**(06):629–641, 2003. 11

[19] Walter Rudin. *Functional Analysis*. International series in pure and applied mathematics. McGraw-Hill, 2006. 11

[20] Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. **Separability of mixed states: necessary and sufficient conditions**. *Physics Letters A*, **223**(1):1–8, 1996. 11, 12

[21] Erling Størmer. *Positive linear maps of operator algebras*. Springer Science & Business Media, 2012. 12

[22] Goran Lindblad. **On the generators of quantum dynamical semigroups**. *Communications in Mathematical Physics*, **48**(2):119–130, 1976. 12

[23] Michael M. Wolf. **Quantum Channels and Operations: Lecture notes**. CTAN: `http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf`, 2011. 13

[24] Mikel Sanz, David Pérez-García, Michael M. Wolf, and J. Ignacio Cirac. **A quantum version of Wielandt's inequality**. *Information Theory, IEEE Transactions on*, **56**(9):4668–4673, 2010. 13

[25] Rajendra Bhatia. *Matrix analysis*, **169**. Springer Science & Business Media, 1997. 13, 14

[26] Gilles Pisier. **Non-Commutative Vector Valued L_p-Spaces and Completely p-Summing Maps**. *Astérisque*, **247**, 1998. 5, 14, 15

[27] Igor Devetak, Marius Junge, Christoper King, and Mary Beth Ruskai. **Multiplicativity of completely bounded p-norms implies a new additivity result**. *Communications in mathematical physics*, **266**(1):37–63, 2006. 14, 15

[28] John Watrous. **Notes on super-operator norms induced by Schatten norms**. *arXiv preprint quant-ph/0411077*, 2004. 14

[29] Salman Beigi. **Sandwiched Rényi divergence satisfies data processing inequality**. *Journal of Mathematical Physics*, **54**(12):122202, 2013. 15

[30] Payam Delgosha and Salman Beigi. **Impossibility of local state transformation via hypercontractivity**. *Communications in Mathematical Physics*, **332**(1):449–476, 2014. 15

[31] Gilles Pisier. *Introduction to operator space theory*, **294**. Cambridge University Press, 2003. 15

[32] Christopher A. Fuchs and Jeroen Van De Graaf. **Cryptographic distinguishability measures for quantum-mechanical states**. *Information Theory, IEEE Transactions on*, **45**(4):1216–1227, 1999. 16

[33] Armin Uhlmann. **The "transition probability" in the state space of A*-algebra**. *Reports on Mathematical Physics*, **9**(2):273–279, 1976. 17

[34] Göran Lindblad. **Completely positive maps and entropy inequalities**. *Communications in Mathematical Physics*, **40**(2):147–151, 1975. 17

[35] Armin Uhlmann. **Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory**. *Communications in Mathematical Physics*, **54**(1):21–32, 1977. 17

[36] Masanori Ohya. *Quantum entropy and its use*. Springer Science & Business Media, 2004. 18

[37] Nicolas J. Cerf and Chris Adami. **Negative entropy and information in quantum mechanics**. *Physical Review Letters*, **79**(26):5194, 1997. 19

[38] Renato Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zürich, Nr. 20213, 2005. 19

[39] Marco Tomamichel. *A framework for non-asymptotic quantum information theory*. PhD thesis, ETH Zürich, Nr. 20213, 2012. 19, 20

[40] Robert König, Renato Renner, and Christian Schaffner. **The operational meaning of min-and max-entropy**. *Information Theory, IEEE Transactions on*, **55**(9):4337–4347, 2009. 20

[41] Marco Tomamichel, Roger Colbeck, and Renato Renner. **A fully quantum asymptotic equipartition property**. *Information Theory, IEEE Transactions on*, **55**(12):5840–5847, 2009. 21

[42] Mark M. Wilde. *Quantum information theory*. Cambridge University Press, 2013. 21

[43] Dennis Kretschmann and Reinhard F. Werner. **Tema con variazioni: quantum channel capacity**. *New Journal of Physics*, **6**(1):26, 2004. 21, 22

[44] Seth Lloyd. **Capacity of the noisy quantum channel**. *Physical Review A*, **55**(3):1613, 1997. 22

[45] Peter W. Shor. **The quantum channel capacity and coherent information**. *Lecture Notes, MSRI Workshop on Quant. Comp., Available at http://www.msri.org/realvideo/ln/msri/2002/quantumcrypto/shor/1/*, 2002. 22

[46] Igor Devetak. **The private classical capacity and quantum capacity of a quantum channel**. *Information Theory, IEEE Transactions on*, **51**(1):44–55, 2005. 22

[47] Toby S. Cubitt, David Elkouss, William Matthews, Maris Ozols, David Pérez-García, and Sergii Strelchuk. **Unbounded number of channel uses may be required to detect quantum capacity**. *Nature communications*, **6**, 2015. 23

[48] Toby S. Cubitt, Mary Beth Ruskai, and Graeme Smith. **The structure of degradable quantum channels**. *Journal of Mathematical Physics*, **49**(10):102104, 2008. 23

[49] Alexander S. Holevo and Reinhard F. Werner. **Evaluating capacities of bosonic Gaussian channels**. *Physical Review A*, **63**(3):032312, 2001. 23

[50] Patrick Hayden, Michal Horodecki, Andreas Winter, and Jon Yard. **A decoupling approach to the quantum capacity**. *Open Systems & Information Dynamics*, **15**(01):7–19, 2008. 23

[51] Frédéric Dupuis. *The decoupling approach to quantum information theory*. PhD thesis, Université de Montréal, 2009. 23, 24

[52] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. **Mixed-state entanglement and quantum error correction**. *Physical Review A*, **54**(5):3824, 1996. 25, 27

[53] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. **Everything you always wanted to know about LOCC (but were afraid to ask)**. *Communications in Mathematical Physics*, **328**(1):303–326, 2014. 26

[54] Wolfgang Dür, J. Ignacio Cirac, Maciej Lewenstein, and Dagmar Bruss. **Distillability and partial transposition in bipartite systems**. *Physical Review A*, **61**(6):062313, 2000. 27

[55] Lieven Clarisse. **The distillability problem revisited**. *arXiv preprint quant-ph/0510035*, 2005. 27

[56] Lukasz Pankowski, Marco Piani, Michal Horodecki, and Pawel Horodecki. **A few steps more towards NPT bound entanglement**. *Information Theory, IEEE Transactions on*, **56**(8):4085–4100, 2010. 27

Quantum Subdivision Capacities and Continuous-time Quantum Coding

A. Müller-Hermes, D. Reeb and M. Wolf May 20, 2015

The quantum Shannon capacity quantifies the optimal rate of transmitting quantum information via many copies of a quantum channel. Here we study the optimal storage rates of an array of identical quantum memories each affected by continuous-time Markovian noise in a Shannon theoretic framework. The noise affecting each memory is modeled by a quantum dynamical semigroup generated by a Liouvillian $\mathcal{L}$, i.e. quantum channels $\mathcal{T}_t = e^{t\mathcal{L}}$ [2] describing the accumulated noise up to times t. We introduce new quantum capacities where coding channels can also be applied during the storage time. This possibility is *not* captured by the quantum Shannon capacity. The ideas of our capacities might be also useful in a classical setting where they seem to be new as well.

1 Quantum subdivision capacities

For a subset $\mathfrak{C}$ of quantum channels we define the $\mathfrak{C}$-*quantum subdivision capacity* of a noise Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ at time t denoted by $\mathcal{Q}_{\mathfrak{C}}(t\mathcal{L})$ as the supremum over achievable rates $R \in \mathbb{R}^+$. A rate R is called *achievable* iff there exist sequences $(n_\nu)_{\nu=1}^\infty$ and $(m_\nu)_{\nu=1}^\infty$ such that $R = \limsup_{\nu\to\infty} \frac{n_\nu}{m_\nu}$ and such that the approximation error vanishes asymptotically

$$\inf_{k,\mathcal{E},\mathcal{D},\mathcal{C}_1,\ldots,\mathcal{C}_k} \left\| \mathrm{id}_2^{\otimes n_\nu} - \mathcal{D} \circ \prod_{l=1}^{k} \left(\mathcal{C}_l \circ \left(e^{\frac{\mathcal{L}}{k}} \right)^{\otimes m_\nu} \right) \circ \mathcal{E} \right\|_\diamond \to 0 \quad \text{as } \nu \to \infty. \tag{1}$$

The latter infimum is over the number of subdivisions $k \in \mathbb{N}$, arbitrary encoding and decoding quantum channels $\mathcal{E}$ and $\mathcal{D}$ and appropriate intermediate coding channels $\mathcal{C}_l \in \mathfrak{C}$ taken from the chosen subset $\mathfrak{C}$.

In the case where $\mathfrak{C} = \mathrm{ID}$ is the set of infinitely divisible quantum channels, i.e. quantum channels of the form $\prod_{i=1}^{l} e^{\mathcal{L}_i}$ for some $l \in \mathbb{N}$ and Liouvillians $\{\mathcal{L}_i\}_{i=1}^{k}$ we prove:

Theorem 1.1. *For any Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ and any time $t \in \mathbb{R}^+$ we have*

$$\mathcal{Q}_{ID}(t\mathcal{L}) = \log_2(d).$$

Note that $\log_2(d)$ is the capacity of a noiseless channel on a d-dimensional system. The proof constructs a coding scheme by cutting the Markovian noise into pieces each close to the identity channel and hence with capacity close to $\log_2(d)$. By successive de- and encoding in between these pieces we can achieve any rate below $\log_2(d)$ when arbitrary coding channels are allowed. The final part of the proof uses the decoupling approach and a typical subspace argument to show that only a small overhead of channel uses is needed to implement this coding scheme using unitaries and strong depolarizing noise.

When $\mathfrak{C} = \mathrm{U}$ is the set of unitary quantum channels the above techniques cannot be applied anymore. In this case we use Schumacher compression to prove that $\mathcal{Q}_{\mathrm{U}}(t\mathcal{L}) > 0$ for any Liouvillian $\mathcal{L} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ and any time $t \in \mathbb{R}^+$. However, we also prove an upper bound:

Theorem 1.2. *For the depolarizing Liouvillian $\mathcal{L}_{dep} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ with fixed-point $\rho_0 \in \mathcal{D}_d$ defined as $\mathcal{L}_{dep}(\rho) = \mathrm{tr}(\rho)\rho_0 - \rho$ we have*

$$\mathcal{Q}_U\left(t\mathcal{L}_r^{dep}\right) \leq \log_2(d) - \left(1 - e^{-t}\right) S\left(\rho_0\right).$$

The proof uses an estimate for the entropy produced by any tensor-product of the depolarizing channel. The unitary coding maps cannot remove the entropy from the system and as it accumulates the transmission rate decreases.

2 Continuous quantum capacity

We also introduce a more general *continuous quantum capacity* allowing for continuous-time coding Liouvillians to be applied in addition to the noise. These capacities take techniques like continuous error correction [3] and algorithmic cooling [1] into account to quantify the limitations of quantum memories in a setting of quantum Shannon theory. Surprisingly even purely dissipative coding Liouvillians can assist the information storage. As an additional result we give an example of Liouvillians $\mathcal{L}, \mathcal{L}_c$ such that $\mathcal{L}_c$ is purely dissipative and $\mathcal{Q}\left(e^{\mathcal{L}}\right) < \mathcal{Q}\left(e^{\mathcal{L}+\mathcal{L}_c}\right)$ holds for the usual quantum Shannon capacity $\mathcal{Q}$.

3 Legal statement

The project was assigned by Michael Wolf. In all parts of this work I was significantly involved.

References

[1] Michael Ben-Or, Daniel Gottesman, and Avinatan Hassidim. Quantum refrigerator. *arXiv preprint arXiv:1301.1995*, 2013.

[2] Goran Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, 1976.

[3] Juan Pablo Paz and Wojciech Hubert Zurek. Continuous error correction. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):355–364, 1998.

Positivity of Linear Maps under Tensor-powers

A. Müller-Hermes, D. Reeb and M. Wolf May 18, 2015

A linear map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called *positive* iff $\mathcal{P}\left(\mathcal{M}_{d_1}^+\right) \subset \mathcal{M}_{d_2}^+$. We call a linear map $\mathcal{P}$ *n-tensor-stable* positive iff $\mathcal{P}^{\otimes n}$ is positive and *tensor-stable positive* iff it is n-tensor-stable positive for any $n \in \mathbb{N}$. There are two *trivial* classes of tensor-stable positive maps: Completely positive maps and completely co-positive maps (i.e. compositions of completely positive maps and a transposition). It is currently not known, whether there exist non-trivial tensor-stable positive maps. We connect this question to several open problems in quantum information theory (NPPT-bound entanglement [2, 1], capacity bounds, existence of quantum channels with entanglement annihilating properties [3]) and obtain some results on n-tensor-stable positivity.

1 Existence of n-tensor-stable positive maps

A set of product states $\{|\alpha_i\rangle \otimes |\beta_i\rangle\}_{i=1}^k$ is called an *unextendible product set* if there is no product vector in $\left(\mathrm{span}_i\left(|\alpha_i\rangle \otimes |\beta_i\rangle\right)\right)^\perp$.

Theorem 1.1. *For any $n \in \mathbb{N}$ and any $d_1, d_2 \geq 2$ there exists an n-tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$.*

The proof uses unextendible product sets to construct a Choi-matrix staying block-positive under tensor-powers. In the proof we need a quantified version of the tensorization property of unextendible product bases. This is shown using multiplicativity of the minimal output eigenvalue of an entanglement breaking channel.

2 Tensor-stable positivity and NPPT-bound entanglement

We denote by $d_{\mathrm{CP}}(\mathcal{P}) := \frac{1}{2}\left(\|C_\mathcal{P}\|_1 - \mathrm{tr}\left(C_\mathcal{P}\right)\right)$, which measures the distance of $\mathcal{P}$ to the cone of completely positive maps. The following lemma connects tensor-stable positivity to distillation problems:

Lemma 2.1. *Let $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be an n-tensor-stable positive map. Then*

$$\frac{d_{CP}(\mathcal{P})}{\|\mathcal{P}\|_\diamond} \leq \inf_{\mathcal{S}\ sep} \|\omega_{d_1} - \mathcal{S}\left(C_\mathcal{P}^{\otimes(n-1)}\right)\|_1, \tag{1}$$

where the infimum is taken over all separable completely positive maps $\mathcal{S} : \mathcal{M}_{d_1^{n-1}} \otimes \mathcal{M}_{d_2^{n-1}} \to \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_1}$.

Thus, any tensor-stable positive map for which there is a sequence $\mathcal{S}_n$ of separable completely positive maps fulfilling $\|\omega_{d_1} - \mathcal{S}_n\left(C_\mathcal{P}^{\otimes n}\right)\|_1 \to 0$ as $n \to \infty$ has to be completely positive. We show how such a sequence of separable completely positive maps can be constructed by generalizing entanglement distillation protocols (originally only defined for quantum states) to block-positive matrices. By doing so we prove:

Theorem 2.1 (Non-trivial tensor-stable positivity implies NPPT-bound entanglement)**.** *If there exists a non-trivial tensor-stable positive map* $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$*, then there exist NPPT bound-entangled states in* $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_1}$ *as well as in* $\mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}$*.*

As there is no NPPT-bound entanglement [1] in 2-dimensional systems we have:

Theorem 2.2. *Any tensor-stable positive map* $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ *with* $d_1 = 2$ *or* $d_2 = 2$ *is trivial.*

3 Connection to other open problems

We generalize the well-known transposition bound by using any surjective tensor-stable positive map. Furthermore, we show how to obtain strong converse bounds on the quantum capacity using tensor-stable positive maps. For the quantum capacity assisted by 2-way classical communication we show that the transposition bound is actually a strong converse bound.

A quantum channel $\mathcal{T}$ is called ∞-locally entanglement annihilating if for any $n \in \mathbb{N}$ and any input state ρ the output state $\mathcal{T}^{\otimes n}(\rho)$ is fully separable (with respect to the n tensor factors). If any tensor-stable positive map is trivial, then any ∞-locally entanglement annihilating channel would be entanglement breaking. This would answer an open question posed in [3].

4 Legal statement

The project was assigned by Michael Wolf. In all parts of this work, except the strong converse bounds, I was significantly involved. I proved an upper bound on $\mathcal{Q}_2$, which was replaced by the strong converse bound.

References

[1] David P DiVincenzo, Peter W Shor, John A Smolin, Barbara M Terhal, and Ashish V Thapliyal. Evidence for bound entangled states with negative partial transpose. *Physical Review A*, 61(6):062312, 2000.

[2] Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: is there a bound entanglement in nature? *Physical Review Letters*, 80(24):5239, 1998.

[3] Lenka Moravčiková and Mário Ziman. Entanglement-annihilating and entanglement-breaking channels. *Journal of Physics A: Mathematical and Theoretical*, 43(27):275306, 2010.

Positivity of Linear Maps under Tensor Powers

Alexander Müller-Hermes,[1,*] David Reeb,[2,1,†] and Michael M. Wolf[1,‡]

[1]*Zentrum Mathematik, Technische Universität München, 85748 Garching, Germany*

[2]*Institute for Theoretical Physics, Leibniz Universität Hannover, 30167 Hannover, Germany*

We investigate linear maps between matrix algebras that remain positive under tensor powers, i.e., under tensoring with n copies of themselves. Completely positive and completely co-positive maps are trivial examples of this kind. We show that for every $n \in \mathbb{N}$ there exist non-trivial maps with this property and that for two-dimensional Hilbert spaces there is no non-trivial map for which this holds for all n. For higher dimensions we reduce the existence question of such non-trivial "tensor-stable positive maps" to a one-parameter family of maps and show that an affirmative answer would imply the existence of NPPT bound entanglement.

As an application we show that any tensor-stable positive map that is not completely positive yields an upper bound on the quantum channel capacity, which for the transposition map gives the well-known cb-norm bound. We furthermore show that the latter is an upper bound even for the LOCC-assisted quantum capacity, and that moreover it is a strong converse rate for this task.

Contents

I. Introduction and main results 2

II. Notation and preliminaries 4

III. Proof of Theorem 1 4

IV. Applications to quantum information theory 6
- A. Entanglement annihilating channels 6
- B. Upper bounds on the quantum capacity 7
- C. Transposition bound as a strong converse rate for the two-way quantum capacity 9
- D. Strong converse rate from tensor-stable positive maps 13

V. Distillation schemes for tensor-stable positive maps 14
- A. Quantifying the distance from the completely positive maps 14
- B. Proof of Theorem 2 and Theorem 4 17
- C. Proof of Theorem 3 18
- D. Generalization of the reduction criterion 18

VI. Conclusion 21

A. Twirling and families of symmetric matrices 22

B. Minimal output eigenvalue 23

References 24

[*]Electronic address: muellerh@ma.tum.de

[†]Electronic address: reeb.qit@gmail.com

[‡]Electronic address: m.wolf@tum.de

I. INTRODUCTION AND MAIN RESULTS

Within the set $\mathcal{M}_d$ of complex $d \times d$-matrices we denote the cone of positive matrices by $\mathcal{M}_d^+$ (we call "positive semidefinite matrices" simply "positive matrices"). A linear map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called positive if $\mathcal{P}\left(\mathcal{M}_{d_1}^+\right) \subseteq \mathcal{M}_{d_2}^+$, and we then write $\mathcal{P} \geq 0$. We want to study how positivity of a linear map behaves when taking tensor powers. Therefore we consider the following:

Definition 1 (Tensor-stable positivity).

*(i) A linear map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called **n-tensor-stable positive** for some number $n \in \mathbb{N}$ if the map $\mathcal{P}^{\otimes n} : \mathcal{M}_{d_1^n} \to \mathcal{M}_{d_2^n}$ is positive.*

*(ii) A linear map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is called **tensor-stable positive** if the map $\mathcal{P}$ is n-tensor-stable positive for all $n \in \mathbb{N}$.*

Note that every n-tensor-stable positive map is in particular a positive map. The following example displays some maps that are easily seen to be tensor-stable positive. We will call all maps from these classes **trivial tensor-stable positive maps**.

Example I.1 (Trivial tensor-stable positive maps).

1. *All completely positive maps are tensor-stable positive, i.e. all linear maps $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ such that $(\mathrm{id}_d \otimes \mathcal{T}) : \mathcal{M}_d \otimes \mathcal{M}_{d_1} \to \mathcal{M}_d \otimes \mathcal{M}_{d_2}$ is positive for all dimensions $d \in \mathbb{N}$.*
2. *All maps of the form $\vartheta_{d_2} \circ \mathcal{T}$ for a completely positive map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ and the transposition $\vartheta_d : \mathcal{M}_d \to \mathcal{M}_d$ are tensor-stable positive. The maps of this form are called **completely co-positive**.*

We will be concerned with three basic questions:

1. Are there any non-trivial tensor-stable positive maps?
2. How far away can an n-tensor-stable positive map be from the cones of completely positive and completely co-positive maps (i.e. from the two cones of trivial tensor-stable positive maps from Example I.1)?
3. What are the implications of question 1. for quantum information theory?

Our main results are the following. In section III we use (non-orthogonal) unextendible product bases to show:

Theorem 1 (Existence of n-tensor-stable positive maps). *For any $n \in \mathbb{N}$ and any $d_1, d_2 \geq 2$ there exists an n-tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ that is* not *a trivial tensor-stable positive map.*

Our construction used to obtain this theorem does not seem to suffice for constructing a non-trivial tensor-stable positive map (i.e. one for all $n \in \mathbb{N}$), and at the time of writing we do not know whether such a map exists.

In section IV we discuss applications and implications of tensor-stable positive maps for quantum information theory. We show that the existence of an ∞-locally entanglement annihilating channel [11, 12, 21] which is not entanglement breaking [18] implies the existence of non-trivial tensor-stable positive maps. A quantum channel is called ∞-locally entanglement annihilating if any state when sent through arbitrarily many copies of the channel becomes fully separable. It is currently not known whether such channels exist outside the set of entanglement breaking channels [10].

In Section IV B we generalize the well-known transposition bound [14] to show that tensor-stable positive, but not completely positive, maps yield upper bounds on the quantum channel capacity as well as strong converse rates for this task (Section IV D). In Section IV C we show that the transposition bound is an upper bound even on the LOCC-assisted quantum capacity (see also Corollary 2) and constitutes a strong converse rate for this task.

In light of these implications, deciding question 1. would have important consequences for quantum information theory. Whereas we cannot resolve this question in general, in section V we use techniques from the theory of entanglement distillation and a generalization of a technique used in [27] to prove:

Theorem 2 (Only trivial tensor-stable positive maps in $d = 2$)**.** *There are* no *non-trivial tensor-stable positive maps $\mathcal{P} : \mathcal{M}_2 \to \mathcal{M}_d$ or $\mathcal{P} : \mathcal{M}_d \to \mathcal{M}_2$ for any $d \in \mathbb{N}$.*

Furthermore, a non-trivial tensor-stable positive map exists iff one exists within the following one-parameter families based on Werner states [30]:

Theorem 3 (One-parameter family of candidates for non-trivial tensor-stable positivity)**.** *Let $d_1, d_2 \in \mathbb{N}$, $d \in \{d_1, d_2\}$, and for $p \in [-1, 1]$ let*

$$\mathcal{P}_p := \mathcal{W}_p \otimes \left(\vartheta_d \circ \mathcal{W}_p\right) \; : \; \mathcal{M}_d \otimes \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d \,, \tag{1}$$

where we define for $X \in \mathcal{M}_d$:

$$\mathcal{W}_p\left(X\right) := \frac{1}{d^2 - 1}\left(\left(d - p\right)\operatorname{tr}\left(X\right)\mathbb{1}_d - \left(1 - dp\right)X^T\right). \tag{2}$$

(i) If there exists a non-trivial tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$, then there exists $p \in [-1, 0)$ such that the map (1) is tensor-stable positive.

(ii) If for some $p \in [-1, 0)$ the map (1) is tensor-stable positive, then it is non-trivial tensor-stable positive (i.e. it is neither completely positive nor completely co-positive).

The aforementioned connection to the theory of entanglement distillation has the following direct implication:

Theorem 4 (Non-trivial tensor-stable positivity implies NPPT-bound entanglement)**.** *If there exists a non-trivial tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$, then there exist NPPT bound-entangled states [8, 9, 17] in $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_1}$ as well as in $\mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}$.*

After completion of this work, we learned that tensor-stable positive maps have been introduced by M. Hayashi under the name “tensor product positive maps” in [13, chapter 5], where it was furthermore shown that the quantum relative entropy does not increase under the application of any trace-preserving tensor product positive map.

II. NOTATION AND PRELIMINARIES

For every $d \in \mathbb{N}$, we fix an orthonormal basis $\{|i\rangle\}_{i=1}^{d}$ of the Hilbert space $\mathbb{C}^d$, and denote by $\vartheta_d(X) := X^T$ the transposition w.r.t. that basis, the d-dimensional maximally entangled state by $|\Omega_d\rangle := \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|ii\rangle \in \mathcal{M}_{d^2}$ and the corresponding projection by $\omega_d := |\Omega_d\rangle\langle\Omega_d|$. The $d \times d-$identity matrix will be denoted by $\mathbb{1}_d$. The following Lemma collects two frequently used and well-known techniques involving the maximally entangled state and linear maps that can be proved by direct computation.

Lemma 1 (Tricks using the maximally entangled state)**.**

1. *For any $d_2 \times d_1$-matrix X we have* $\left(\mathbb{1}_{d_1} \otimes X\right)|\Omega_{d_1}\rangle = \sqrt{\frac{d_2}{d_1}}\left(X^T \otimes \mathbb{1}_{d_2}\right)|\Omega_{d_2}\rangle$.
2. *For any map $\mathcal{L} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ that is hermiticity-preserving (i.e. maps hermitian matrices to hermitian matrices), we have* $\left(\mathrm{id}_{d_1} \otimes \mathcal{L}\right)(\omega_{d_1}) = \frac{d_2}{d_1}\left(\vartheta_{d_1} \circ \mathcal{L}^* \circ \vartheta_{d_2} \otimes \mathrm{id}_{d_2}\right)(\omega_{d_2})$.

In the above $\mathcal{L}^*$ denotes the adjoint w.r.t. the Hilbert-Schmidt inner product.

We will frequently make use of the Choi-Jamiolkowski isomorphism between linear maps $\mathcal{L} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ and matrices $C \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$. The Choi matrix of such a linear map is defined as $C_{\mathcal{L}} := \left(\mathrm{id}_{d_1} \otimes \mathcal{L}\right)(\omega_{d_1})$. Note that we used the normalized maximally entangled state in this definition. The following implications are well known:

- $\mathcal{L} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is positive iff $C_{\mathcal{L}}$ is block-positive, i.e. $\left(\langle\phi| \otimes \langle\psi|\right) C \left(|\phi\rangle \otimes |\psi\rangle\right) \geq 0$ for all $|\phi\rangle \in \mathbb{C}^{d_1}$, $|\psi\rangle \in \mathbb{C}^{d_2}$.
- $\mathcal{L} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is completely positive iff $C_{\mathcal{L}} \geq 0$.
- $\mathcal{L} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is completely co-positive iff $C_{\mathcal{L}}^{T_2} \geq 0$.

For $C \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ we denote by $C^{T_2} := \left(\mathrm{id}_d \otimes \vartheta_d\right)(C)$ the partial transpose w.r.t. to the second tensor-factor. The paradigm of a block-positive matrix that is not positive is the Choi matrix of the transposition $\omega_d^{T_2} = \frac{1}{d}\mathbb{F}_d$. Here $\mathbb{F}_d : \mathbb{C}^d \otimes \mathbb{C}^d \to \mathbb{C}^d \otimes \mathbb{C}^d$ denotes the flip operator with $\mathbb{F}_d|ij\rangle = |ji\rangle$.

Matrices $C \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ with $C^{T_2} \geq 0$ will be called PPT (positive partial transpose). A matrix is called NPPT (non-positive partial transpose) if it is not PPT. The question of NPPT-bound entanglement [8, 9, 17] concerns the problem of creating a maximally entangled state from many copies of an NPPT-state using only local operations and classical communications (LOCC) [6]. While it is clear that no maximally entangled state can be created from many copies of a PPT-state it is currently unknown whether the same can be true for an NPPT-state.

The $\diamond$-norm [23] is defined as $\|\mathcal{L}\|_\diamond := \sup_{n\in\mathbb{N}} \sup_{\|X\|_1=1} \|\left(\mathrm{id}_n \otimes \mathcal{L}\right)(X)\|_1$ for a linear map $\mathcal{L} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$.

III. PROOF OF THEOREM 1

Our proof of Theorem 1 uses the following quantitative version of the result [7, Lemma 22] about tensor products of generalizations of unextendible product bases [1], whose elements are not necessarily mutually orthogonal. For the following, we call a matrix $P \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ *separable* if it can be written as $P = \sum_{i=1}^{k} A_i \otimes B_i$ for some $k \in \mathbb{N}$ and matrices $A_i \in \mathcal{M}_{d_1}$, $B_i \in \mathcal{M}_{d_2}$ with $A_i \geq 0$ and $B_i \geq 0$.

Lemma 2 (Multiplicativity of minimal overlap with product states)**.** *For a separable matrix* $P \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$*, define*

$$\mu \ := \ \min\{(\langle\psi| \otimes \langle\phi|)\, P\, (|\psi\rangle \otimes |\phi\rangle) : |\psi\rangle \in \mathbb{C}^{d_1}, |\phi\rangle \in \mathbb{C}^{d_2}, \langle\psi|\psi\rangle = \langle\phi|\phi\rangle = 1\}.$$

Then, for all $n \in \mathbb{N}$*, we have*

$$\min\{(\langle\Psi| \otimes \langle\Phi|)\, P^{\otimes n}\, (|\Psi\rangle \otimes |\Phi\rangle) : |\Psi\rangle \in \left(\mathbb{C}^{d_1}\right)^{\otimes n}, |\Phi\rangle \in \left(\mathbb{C}^{d_2}\right)^{\otimes n}, \langle\Psi|\Psi\rangle = \langle\Phi|\Phi\rangle = 1\} \ = \ \mu^n.$$

In particular, if there is no nonzero product vector in the kernel of P*, then there is none in the kernel of* $P^{\otimes n}$*.*

The connection to [7, Lemma 22] becomes clear by noting that any separable matrix $P \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ admits a decomposition of the form $P = \sum_{i=1}^{N} |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i|$ such that $\ker(P) = \left(\mathrm{span}\{|\psi_i\rangle \otimes |\phi_i\rangle\}_{i=1}^{N}\right)^{\perp}$. Hence for $\mu > 0$ the set $\{|\psi_i\rangle \otimes |\phi_i\rangle\}$ forms an unextendible product set.

For the following proof we will need the *minimal output eigenvalue* of a completely positive map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ defined as

$$\lambda^{\min}_{\text{out}}[\mathcal{T}] := \min_{\rho \in \mathcal{D}_{d_1}} \lambda_{\min}(\mathcal{T}(\rho)) \,. \tag{3}$$

Here $\lambda_{\min}(\cdot)$ denotes the minimal eigenvalue and $\mathcal{D}_{d_1}$ is the set of quantum states in $\mathcal{M}_{d_1}$. For any entanglement breaking map $\mathcal{T}$ and any completely positive map $\mathcal{S}$ we prove in Theorem 9 (Appendix B) that

$$\lambda^{\min}_{\text{out}}[\mathcal{T} \otimes \mathcal{S}] = \lambda^{\min}_{\text{out}}[\mathcal{T}]\, \lambda^{\min}_{\text{out}}[\mathcal{S}] \,.$$

Thus, $\lambda^{\min}_{\text{out}}$ is multiplicative for entanglement breaking maps.

proof of Lemma 2. Consider the completely positive map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ such that $P = C_{\mathcal{T}}$. Then we have

$$(\langle\Psi| \otimes \langle\Phi|)\, P^{\otimes k}\, (|\Psi\rangle \otimes |\Phi\rangle) = \frac{1}{d_1^k} \langle\Phi| \mathcal{T}^{\otimes k}\left(\overline{|\Psi\rangle\langle\Psi|}\right) |\Phi\rangle$$

for all $k \in \mathbb{N}$ and all $|\Psi\rangle \in \left(\mathbb{C}^{d_1}\right)^{\otimes k}$, $|\Phi\rangle \in \left(\mathbb{C}^{d_2}\right)^{\otimes k}$. Using the minimal output eigenvalue (3) we have for any $k \in \mathbb{N}$

$$\lambda^{\min}_{\text{out}}(\mathcal{T}^{\otimes k}) = d_1^k \min\left\{ (\langle\Psi|\langle\Phi|)\, P^{\otimes k}\, (|\Psi\rangle|\Phi\rangle) : |\Psi\rangle \in (\mathbb{C}^{d_1})^{\otimes k}, |\Phi\rangle \in (\mathbb{C}^{d_2})^{\otimes k}, \|\Psi\| = \|\Phi\| = 1\right\}.$$

As P is separable the map $\mathcal{T}$ is entanglement breaking [18] and we can apply Theorem 9 from Appendix B. This shows that $\lambda^{\min}_{\text{out}}(\mathcal{T}^{\otimes n}) = \lambda^{\min}_{\text{out}}(\mathcal{T})^n$ and finishes the proof.

□

With this ingredient we prove Theorem 1:

proof of Theorem 1. Choose orthonormal bases $\{|i\rangle\} \subseteq \mathbb{C}^{d_1}$ and $\{|j\rangle\} \subseteq \mathbb{C}^{d_2}$ and define the operator

$$\begin{aligned} P \ := \ & \big(|1\rangle|1\rangle + |2\rangle|2\rangle\big)\big(\langle 1|\langle 1| + \langle 2|\langle 2|\big) \ + \ |1\rangle|2\rangle\langle 1|\langle 2| \ + \ |2\rangle|1\rangle\langle 2|\langle 1| \\ & + \sum_{\substack{(i,j) \\ i>2 \text{ or } j>2}} |i\rangle|j\rangle\langle i|\langle j| \ \in \ \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \,. \end{aligned} \tag{4}$$

It is easy to verify that

$$P = \sum_{k=1}^{3} \frac{1}{3} |\xi_k\rangle\langle\xi_k| \otimes \overline{|\xi_k\rangle\langle\xi_k|} + \sum_{\substack{(i,j)\\ i>2 \text{ or } j>2}} |i\rangle\langle i| \otimes |j\rangle\langle j|$$

for $|\xi_k\rangle = |0\rangle + e^{\frac{2\pi i k}{3}}|1\rangle$. This shows that P is separable as a sum of positive product operators. Now define

$$\mu := \min\{(\langle\psi| \otimes \langle\phi|)\, P\, (|\psi\rangle \otimes |\phi\rangle) : |\psi\rangle \in \mathbb{C}^{d_1}, |\phi\rangle \in \mathbb{C}^{d_2}, \langle\psi|\psi\rangle = \langle\phi|\phi\rangle = 1\}$$

and apply Lemma 2 showing that for any $k \in \mathbb{N}$:

$$\min\{(\langle\Psi| \otimes \langle\Phi|)\, P^{\otimes k}\, (|\Psi\rangle \otimes |\Phi\rangle) : |\Psi\rangle \in \left(\mathbb{C}^{d_1}\right)^{\otimes k}, |\Phi\rangle \in \left(\mathbb{C}^{d_2}\right)^{\otimes k}, \langle\Psi|\Psi\rangle = \langle\Phi|\Phi\rangle = 1\} = \mu^k.$$

As the kernel $\ker(P) = \operatorname{span}\{|1\rangle|1\rangle - |2\rangle|2\rangle\}$ of P in (4) contains no nonzero product vector we have $\mu > 0$. With this we can compute

$$\begin{aligned}\left(\langle\Psi| \otimes \langle\Phi|\right) \left(P - \varepsilon \mathbb{1}_{d_1}\mathbb{1}_{d_2}\right)^{\otimes n} \left(|\Psi\rangle \otimes |\Phi\rangle\right) &\geq \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} \varepsilon^{2k} \mu^{n-2k} - \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n}{2k-1} \varepsilon^{2k-1} \|P\|_\infty^{n-2k+1} \\ &= \frac{(\mu+\varepsilon)^n + (\mu-\varepsilon)^n}{2} - \frac{(\|P\|_\infty + \varepsilon)^n - (\|P\|_\infty - \varepsilon)^n}{2} \\ &\geq \mu^n - (\|P\|_\infty + \varepsilon)^n + \|P\|_\infty^n \geq 0\end{aligned} \tag{5}$$

for any $0 \leq \varepsilon \leq \sqrt[n]{\|P\|_\infty^n + \mu^n} - \|P\|_\infty$. This means that $(P - \varepsilon \mathbb{1}_{d_1 d_2})^{\otimes n} \in (\mathcal{M}_{d_1})^{\otimes n} \otimes (\mathcal{M}_{d_2})^{\otimes n}$ is a block-positive operator for any $\varepsilon \in \left[0, \sqrt[n]{\|P\|_\infty^n + \mu^n} - \|P\|_\infty\right]$, which by the Choi-Jamiolkowski isomorphism (Section II) corresponds to a positive linear map $\mathcal{P}_\varepsilon^{\otimes n} : (\mathcal{M}_{d_1})^{\otimes n} \to (\mathcal{M}_{d_2})^{\otimes n}$. The map $\mathcal{P}_\varepsilon : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ with Choi matrix $(P - \varepsilon \mathbb{1}_{d_1 d_2})$ is thus n-tensor-stable positive. Note that P is rank-deficient and as P^{T_2} equals the expression (4) with the first terms replaced by $(|1\rangle|2\rangle + |2\rangle|1\rangle)(\langle 1|\langle 2| + \langle 2|\langle 1|) + |1\rangle|1\rangle\langle 1|\langle 1| + |2\rangle|2\rangle\langle 2|\langle 2|$ it is rank-deficient as well. Hence, the Choi matrices $(P - \varepsilon \mathbb{1}_{d_1 d_2})$ and $(P^{T_2} - \varepsilon \mathbb{1}_{d_1 d_2})$ of $\mathcal{P}$ respectively $\vartheta_{d_2} \circ \mathcal{P}$ are not positive for $\epsilon > 0$, which finally shows that $\mathcal{P}_\varepsilon$ is not a trivial tensor-stable positive map for any $\varepsilon \in \left(0, \sqrt[n]{\|P\|_\infty^n + \mu^n} - \|P\|_\infty\right]$. □

IV. APPLICATIONS TO QUANTUM INFORMATION THEORY

Deciding the existence of non-trivial tensor-stable positive maps could lead to a solution of other open problems in quantum information theory. Here we will discuss two such connections.

A. Entanglement annihilating channels

In [10–12, 21] the authors study how entanglement in a multipartite setting can be destroyed by dissipative processes. They define the set of k-locally entanglement annihilating channels. These are quantum channels $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ such that $\mathcal{T}^{\otimes k}(\rho)$ is k-partite separable for all input states $\rho \in \mathcal{M}_{d_1^k}$, i.e. for all $\rho \geq 0$ we have $\mathcal{T}^{\otimes k}(\rho) = \sum_{i=1}^{m} p_i \sigma_i^{(1)} \otimes \sigma_i^{(2)} \otimes \cdots \otimes \sigma_i^{(k)}$ for some $m \in \mathbb{N}$, states $\sigma_i^{(j)} \in \mathcal{M}_{d_2}$ and $p_i \in \mathbb{R}^+$ depending on ρ. Furthermore, a channel is called ∞-locally entanglement annihilating if it is k-locally entanglement annihilating for all $k \in \mathbb{N}$.

It is clear that entanglement breaking channels [18] are ∞-locally entanglement annihilating. In [11, 21] examples of 2-locally entanglement annihilating channels are constructed that are not entanglement breaking. However it is not known whether there exist an ∞-locally entanglement annihilating channel, which is not entanglement breaking.

We can prove the following theorem connecting k-locally entanglement annihilating channels to tensor-stable positive maps.

Theorem 5. *If the quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is k-locally entanglement annihilating for some $k \geq 2$, but not entanglement breaking, then there exists a positive map $\mathcal{S} : \mathcal{M}_{d_2} \to \mathcal{M}_{d_1}$ such that $\mathcal{P} : \mathcal{M}_{d_1^2} \to \mathcal{M}_{d_1^2}$ defined as*

$$\mathcal{P} = (\mathcal{S} \circ \mathcal{T}) \otimes (\vartheta \circ \mathcal{S} \circ \mathcal{T}) \tag{6}$$

is a $\lfloor \frac{k}{2} \rfloor$-tensor-stable positive map that is not *a trivial tensor-stable positive map.*

Thus, the existence of a non-entanglement breaking ∞-locally entanglement annihilating channel implies the existence of a non-trivial tensor-stable positive map.

Proof. Assume that $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is a k-locally entanglement annihilating channel. If $\mathcal{T}$ is not entanglement breaking, then there exists a positive map $\mathcal{S} : \mathcal{M}_{d_2} \to \mathcal{M}_{d_1}$ such that $\mathcal{S} \circ \mathcal{T}$ is not completely positive [16]. Now consider the map $\mathcal{P} : \mathcal{M}_{d_1^2} \to \mathcal{M}_{d_1^2}$ defined in (6). As $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is k-entanglement annihilating, $\mathcal{P}$ is $\lfloor \frac{k}{2} \rfloor$-tensor-stable positive. Furthermore it is neither completely positive nor completely co-positive.

□

By our Theorem 4, the existence of a ∞-locally entanglement annihilating but not entanglement breaking channel then implies the existence of NPPT-bound entanglement.

B. Upper bounds on the quantum capacity

The existence of non-trivial tensor-stable positive maps would imply new bounds on the quantum capacity of a quantum channel. By generalizing the proof of the transposition criterion [14, 19] we obtain a quantitative bound on the quantum capacity $\mathcal{Q}(\mathcal{T})$ of a quantum channel. Recall that the quantum capacity is defined as:

Definition 2 (Quantum capacity $\mathcal{Q}$, see [19, 20])**.**

*The **quantum capacity** of a quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is defined as*

$$\mathcal{Q}(\mathcal{T}) := \sup\{R \in \mathbb{R}^+ : \ R \text{ achievable rate}\},$$

where a rate $R \in \mathbb{R}^+$ is called achievable if there exist sequences $(n_\nu)_{\nu=1}^{\infty}, (m_\nu)_{\nu=1}^{\infty}$ such that $R = \limsup_{\nu\to\infty} \frac{n_\nu \log_2(d)}{m_\nu}$ and the approximation error vanishes in the asymptotic limit, i.e.

$$\inf_{\mathcal{E},\mathcal{D}} \frac{1}{2} \left\| \mathrm{id}_d^{\otimes n_\nu} - \mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E} \right\|_\diamond \to 0 \quad \text{as } \nu \to \infty. \tag{7}$$

Here, the infimum runs over all encoding and decoding quantum channels $\mathcal{E} : \mathcal{M}_d^{\otimes n_\nu} \to \mathcal{M}_{d_1}^{\otimes m_\nu}$ and $\mathcal{D} : \mathcal{M}_{d_2}^{\otimes m_\nu} \to \mathcal{M}_d^{\otimes n_\nu}$, and $d \geq 2$ is any fixed integer (note, the value of $\mathcal{Q}(\mathcal{T})$ does not depend on the choice of d [19]).

Currently all channels known to have zero quantum capacity come from two classes [26]. These are the classes of anti-degradable channels [2, 5] and of completely co-positive quantum channels. The latter can be shown using the quantitative transposition bound [14]

$$\mathcal{Q}(\mathcal{T}) \leq \log_2\left(\|\vartheta_{d_2} \circ \mathcal{T}\|_\diamond\right) \tag{8}$$

on the quantum capacity of any quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$. We will now prove a generalization of this bound using any surjective, unital and tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$ that is not completely positive. Note that any surjective linear map $\mathcal{P} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$ has a linear right-inverse $\mathcal{P}^{-1} : \mathcal{M}_{d_2} \to \mathcal{M}_{d_3}$ (generally not unique) satisfying $\mathcal{P} \circ \mathcal{P}^{-1} = \mathrm{id}_{d_2}$.

Theorem 6. *Let $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be a quantum channel and $\mathcal{P} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$ be a surjective, unital and tensor-stable positive map that is not completely positive, and let $\mathcal{P}^{-1}$ be any right-inverse of $\mathcal{P}$. Then we have*

$$\mathcal{Q}(\mathcal{T}) \leq \frac{\log_2\left(\|\mathcal{P}^{-1} \circ \mathcal{T}\|_\diamond \|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty\right)\log_2(d_2)}{\log_2\left(\|\mathcal{P}^*\|_\diamond\right)}$$

Note that the transposition bound (8) *is retrieved for $\mathcal{P} = \vartheta_{d_2}$.*

Proof. As $\mathcal{P}^*$ is trace-preserving but not completely positive, we have $\|\mathcal{P}^*\|_\diamond > 1$ [23], and from the definition of the diamond norm it is furthermore $\|\mathcal{P}^*\|_\diamond = \|\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2}\|_\diamond$. Now we can do the following calculation, which generalizes the proof of the transposition bound [14, 19]. Let $\mathcal{E} : \mathcal{M}_{d_2}^{\otimes n_\nu} \to \mathcal{M}_{d_1}^{\otimes m_\nu}$ and $\mathcal{D} : \mathcal{M}_{d_2}^{\otimes m_\nu} \to \mathcal{M}_{d_2}^{\otimes n_\nu}$ denote arbitrary quantum channels. Then:

$$\begin{aligned}
\|\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2}\|_\diamond^{n_\nu} &= \|(\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2})^{\otimes n_\nu} \circ \left(\mathrm{id}_{d_2}^{\otimes n_\nu} - \mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E} + \mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E}\right)\|_\diamond \\
&\leq \|\,(\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2})^{\otimes n_\nu} \circ (\mathrm{id}_{d_2}^{\otimes n_\nu} - \mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E})\|_\diamond + \|\,(\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2})^{\otimes n_\nu} \circ \mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E}\|_\diamond \\
&\leq 2\epsilon_\nu \|\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2}\|_\diamond^{n_\nu} + \|\,(\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2})^{\otimes n_\nu} \circ \mathcal{D} \circ \mathcal{P}^{\otimes m_\nu}\|_\diamond \|\mathcal{P}^{-1} \circ \mathcal{T}\|_\diamond^{m_\nu},
\end{aligned}$$

where we abbreviated the approximation error by $\epsilon_\nu := \|\mathrm{id}_{d_2}^{\otimes n_\nu} - \mathcal{D} \circ \mathcal{T}^{\otimes m_\nu} \circ \mathcal{E}\|_\diamond/2$ and used well-known properties [23] of the $\diamond$-norm. Note that by Lemma 1 we have

$$\begin{aligned}
&\left(\mathrm{id}_{d_3^{m_\nu}} \otimes \left[(\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2})^{\otimes n_\nu} \circ \mathcal{D} \circ \mathcal{P}^{\otimes m_\nu}\right]\right)\left(\omega_{d_3}^{\otimes m_\nu}\right) \\
&\quad = \left(\frac{d_2}{d_3}\right)^{m_\nu} \vartheta_{d_3^{\otimes(m_\nu+n_\nu)}} \circ (\mathcal{P}^*)^{\otimes(m_\nu+n_\nu)} \circ \vartheta_{d_2^{\otimes(m_\nu+n_\nu)}} \circ \left(\mathrm{id}_{d_2^{m_\nu}} \otimes \mathcal{D}\right)\left(\omega_{d_2}^{\otimes m_\nu}\right) \;\geq\; 0,
\end{aligned}$$

since $\mathcal{P}^*$ is also tensor-stable positive. Thus, the map $(\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2})^{\otimes n_\nu} \circ \mathcal{D} \circ \mathcal{P}^{\otimes m_\nu}$ is completely positive, and by unitality of $\mathcal{P}$ we have

$$\|\,(\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2})^{\otimes n_\nu} \circ \mathcal{D} \circ \mathcal{P}^{\otimes m_\nu}\|_\diamond = \|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty^{m_\nu}$$

for all quantum channels $\mathcal{D}$, as any positive map attains its operator norm at the identity matrix [23, Corollary 2.9]. Inserting this into the above calculation we have

$$(1 - 2\epsilon_\nu)\|\mathcal{P}^*\|_\diamond^{n_\nu} = (1 - 2\epsilon_\nu)\|\vartheta_{d_3} \circ \mathcal{P}^* \circ \vartheta_{d_2}\|_\diamond^{n_\nu} \leq \|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty^{m_\nu}\|\mathcal{P}^{-1} \circ \mathcal{T}\|_\diamond^{m_\nu}.$$

Applying the logarithm and taking the limit $\nu \to \infty$ we obtain

$$R = \limsup_{\nu\to\infty} \frac{n_\nu \log_2(d_2)}{m_\nu} \leq \frac{\log_2\left(\|\mathcal{P}^{-1} \circ \mathcal{T}\|_\diamond \|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty\right)\log_2(d_2)}{\log_2\left(\|\mathcal{P}^*\|_\diamond\right)}$$

for any achievable rate R (see Definition 2) and corresponding coding schemes $\mathcal{E}, \mathcal{D}$ with $\epsilon_\nu \to 0$. □

To apply Theorem 6 it is enough to have a surjective and tensor-stable positive map $\mathcal{R} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$ which is not completely positive. Note that as $\mathcal{R}$ is surjective, it is easy to see that the operator $\mathcal{R}(\mathbb{1}_{d_3})$ is strictly positive, and thus the map $\mathcal{P} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$ defined by $\mathcal{P}(X) := \mathcal{R}(\mathbb{1}_{d_3})^{-1/2}\mathcal{R}(X)\mathcal{R}(\mathbb{1}_{d_3})^{-1/2}$ is unital, surjective and tensor-stable positive. Furthermore, $\mathcal{P}$ is completely (co-)positive if and only if $\mathcal{R}$ was completely (co-)positive. Thus, we constructed a map $\mathcal{P}$ as needed for Theorem 6.

Note that for completely co-positive maps $\mathcal{P}$ the capacity bound from Theorem 6 is worse than the transposition bound given by (8). To prove this let $\mathcal{P} = \vartheta_{d_2} \circ \mathcal{S}$ for a surjective, unital and completely positive map $\mathcal{S} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$. Then, due to the invertibility of ϑ_{d_2}, any right-inverse $\mathcal{P}^{-1}$ of $\mathcal{P}$ can be written as $\mathcal{P}^{-1} = \mathcal{S}^{-1} \circ \vartheta_{d_2}$ with a right-inverse $\mathcal{S}^{-1} : \mathcal{M}_{d_2} \to \mathcal{M}_{d_3}$ of $\mathcal{S}$. By unitality of $\mathcal{P}$ and basic properties of the $\diamond$-norm (see for instance [23, Exercise 3.11 and Corollary 2.9]) we have $\|\mathcal{P}^*\|_\diamond \leq d_2\|\mathcal{P}^*\|_{1\to1} = d_2$, and furthermore $\|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty = \|\mathcal{S}^*(\mathbb{1}_{d_2})\|_\infty = \|\mathcal{S}\|_\diamond$ since $\mathcal{S}$ is completely positive. Thus, for any quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ we have:

$$\frac{\log_2\left(\|\mathcal{P}^{-1}\circ\mathcal{T}\|_\diamond \|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty\right)\log_2 d_2}{\log_2\left(\|\mathcal{P}^*\|_\diamond\right)} \geq \log_2\left(\|\mathcal{S}^{-1}\circ\vartheta_{d_2}\circ\mathcal{T}\|_\diamond\|\mathcal{S}\|_\diamond\right) \geq \log_2\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond \geq \mathcal{Q}(\mathcal{T}).$$

Therefore, to obtain a capacity bound stronger than the transposition bound (8), one would need a non-trivial tensor-stable positive map $\mathcal{P}$.

Similarly, if $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_3}$ is a trace-preserving and tensor-stable positive map that is not completely positive and that has a left-inverse $\mathcal{P}^{-1} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_1}$, then the following bound holds for any quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$:

$$\mathcal{Q}(\mathcal{T}) \leq \frac{\log_2\left(\|\mathcal{T}\circ\mathcal{P}^{-1}\|_\diamond\right)\log_2(d_1)}{\log_2\left(\|\mathcal{P}^*\|_\diamond/\|\mathcal{P}(\mathbb{1}_{d_1})\|_\infty\right)}. \tag{9}$$

The proof works in the same way as the proof of Theorem 6, and again, this bound reduces to the transposition bound (8) for $\mathcal{P} = \vartheta_{d_1}$.

C. Transposition bound as a strong converse rate for the two-way quantum capacity

We now prove that the transposition bound (8) is even an upper bound on the capacity $\mathcal{Q}_2(\mathcal{T}) \geq \mathcal{Q}(\mathcal{T})$ of any channel $\mathcal{T}$ for forward communication of quantum information assisted by unrestricted two-way classical side communication between both parties and arbitrary local quantum operations (LOCC).

For this, we first define an *LOCC channel* (w.r.t. bipartitions $A : B$ and $A' : B'$ of the input and output systems, respectively) to be any quantum channel $\mathcal{L}_{A:B\to A':B'} : \mathcal{M}_{d_A} \otimes \mathcal{M}_{d_B} \to \mathcal{M}_{d_{A'}} \otimes \mathcal{M}_{d_{B'}}$ that can be written as a sequential concatenation of any number of channels $\mathcal{L}_{A_q:B_q\to A'_qA'_c:B'_qB'_c}$ of the following form ($X_{A_qB_q} \in \mathcal{M}_{d_{A_q}} \otimes \mathcal{M}_{d_{B_q}}$):

$$\mathcal{L}_{A_q:B_q\to A'_qA'_c:B'_qB'_c}(X_{A_qB_q}) = \sum_{i,j}(K_i^A\otimes K_j^B)X_{A_qB_q}(K_i^A\otimes K_j^B)^\dagger \otimes |j\rangle\langle j|_{A'_c}\otimes|i\rangle\langle i|_{B'_c}, \tag{10}$$

where $K_i^A : \mathbb{C}^{|A_q|} \to \mathbb{C}^{|A'_q|}$ and $K_j^B : \mathbb{C}^{|B_q|} \to \mathbb{C}^{|B'_q|}$ ($i \in I, j \in J$) are Kraus operators of quantum channels mapping system A to A'_q and system B to B'_q respectively (i.e. $\sum_i (K_i^A)^\dagger K_i^A = \mathbb{1}_{A_q}$ and $\sum_j (K_j^B)^\dagger K_j^B = \mathbb{1}_{B_q}$), and $|j\rangle_{A'_c}$ and $|i\rangle_{B'_c}$ are orthonormal bases belonging to (effectively classical) systems A_c and B_c of dimension $|J|$ and $|I|$ (see [6] for more details). When one of the systems, such as B, is trivial (i.e. one-dimensional), we also speak of a LOCC channel $\mathcal{L}_{A\to A':B'}$, omitting the indices of the trivial subsystems. From the definition it is clear that any LOCC channel $\mathcal{L}_{A:B\to A':B'} : \mathcal{M}_{d_A} \otimes \mathcal{M}_{d_B} \to \mathcal{M}_{d_{A'}} \otimes \mathcal{M}_{d_{B'}}$ is *PPT preserving* (w.r.t. bipartitions $A : B$ and $A' : B'$), meaning that the map $(\mathrm{id}_{A'} \otimes \vartheta_{B'})\mathcal{L}_{A:B\to A':B'}(\mathrm{id}_A \otimes \vartheta_B)$ is completely positive and therefore a quantum channel, whose $\diamond$-norm equals 1. We can now define the two-way quantum capacity.

Definition 3 (Two-way quantum capacity $\mathcal{Q}_2$)**.**

Given a quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$, *we define an* (N, m, ε)-scheme for quantum communication with two-way classical communication *to be any set of LOCC channels* $\mathcal{L}_{A_i:B_i^tB_i\to A_{i+1}^tA_{i+1}:B_{i+1}}$ *for* $i = 0, \ldots, m$, *where the initial A-system and final B-system are of the same dimension* $N = |A_0| = |B_{m+1}|$ *and are identified with each other,* $A_0 = B_{m+1}$, *the initial B-system and final A-system are trivial,* $|B_0^t| = |B_0| = |A_{m+1}^t| = |A_{m+1}| = 1$, *and the subsystems used for quantum transmission (hence the superscript "t") are of dimensions* $|A_i^t| = d_1$ *and* $|B_i^t| = d_2$ *for* $i = 1, \ldots, m$, *and* ε *is the* $\diamond$*-norm error of the scheme,*

$$\begin{aligned}\varepsilon = \big\|\mathrm{id}_{A_0\to B_{m+1}} - \mathcal{L}_{A_m:B_m^tB_m\to B_{m+1}} \circ \mathcal{T}_{A_m^t\to B_m^t} \circ \mathcal{L}_{A_{m-1}:B_{m-1}^tB_{m-1}\to A_mA_m^tB_m} \circ \mathcal{T}_{A_{m-1}^t\to B_{m-1}^t} \circ \ldots \\ \ldots \circ \mathcal{T}_{A_2^t\to B_2^t} \circ \mathcal{L}_{A_1:B_1^tB_1\to A_2^tA_2:B_2} \circ \mathcal{T}_{A_1^t\to B_1^t} \circ \mathcal{L}_{A_0\to A_1^tA_1:B_1}\big\|_\diamond / 2\,, \qquad (11)\end{aligned}$$

omitting for brevity the action of the identity channel on some subsystems, e.g. in $\mathcal{T}_{A_i^t\to B_i^t} \equiv (\mathcal{T}_{A_i^t\to B_i^t} \otimes \mathrm{id}_{A_i} \otimes \mathrm{id}_{B_i})$.

We call $R \in \mathbb{R}^+$ *an* achievable rate *for quantum communication over the channel* $\mathcal{T}$ *assisted by two-way classical communication if there exists for each* $\nu \in \mathbb{N}$ *a* $(N_\nu, m_\nu, \varepsilon_\nu)$*-scheme as just defined in such a way that* $R = \limsup_{\nu\to\infty} \frac{\log_2(N_\nu)}{m_\nu}$ *and* $\lim_{\nu\to\infty} \varepsilon_\nu = 0$. *The* two-way quantum capacity $\mathcal{Q}_2(\mathcal{T})$ *is defined to be the supremum of all such achievable rates.*

To prove the following statements about $\mathcal{Q}_2$ we need only the PPT preserving property of the LOCC channels in the above coding scheme. The statements hold therefore more generally for quantum communication assisted by any PPT preserving channels.

Lemma 3 (Error of two-way coding schemes)**.** *Let* $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ *be a quantum channel and suppose there exists a* (N, m, ε)*-scheme for quantum communication with two-way classical side communication. Then:*

$$\varepsilon \geq 1 - \frac{\|\vartheta_{d_2} \circ \mathcal{T}\|_\diamond^m}{N}\,.$$

Proof. The following proof generalizes ideas from the examples in [22, Section III]. We follow through the m steps of the given (N, m, ε)-scheme (cf. Definition 3) and examine how the partially transposed communication channel between the two parties evolves. For this, let $\mathcal{S}^{(1)}_{A_0\to A_1^tA_1B_1} := \mathcal{L}_{A_0\to A_1^tA_1B_1}$ and for $i = 1, \ldots, m$,

$$\mathcal{S}^{(i+1)}_{A_0\to A_{i+1}^tA_{i+1}B_{i+1}} := (\mathcal{L}_{A_i:B_i^tB_i\to A_{i+1}^tA_{i+1}:B_{i+1}}) \circ (\mathcal{T}_{A_i^t\to B_i^t} \otimes \mathrm{id}_{A_i} \otimes \mathrm{id}_{B_i}) \circ \mathcal{S}^{(i)}_{A_0\to A_i^tA_iB_i}\,.$$

As each LOCC map in the communication scheme is PPT preserving and using that the transposition is an involution, i.e. $\vartheta_{B_i^tB_i} \circ (\vartheta_{B_i^t} \otimes \vartheta_{B_i}) = \mathrm{id}_{B_i^tB_i}$ we have:

$$
\begin{aligned}
&\big\|(\mathrm{id}_{A^t_{i+1}A_{i+1}}\otimes\vartheta_{B_{i+1}})\circ\mathcal{S}^{(i+1)}_{A_0\to A^t_{i+1}A_{i+1}B_{i+1}}\big\|_\diamond \\
&\qquad\qquad = \big\|(\mathrm{id}_{A^t_{i+1}A_{i+1}}\otimes\vartheta_{B_{i+1}})\circ(\mathcal{L}_{A_i:B^t_iB_i\to A^t_{i+1}A_{i+1}:B_{i+1}})\circ(\mathrm{id}_{A_i}\otimes\vartheta_{B^t_iB_i})\circ \\
&\qquad\qquad\qquad \circ(\mathrm{id}_{A_i}\otimes\vartheta_{B^t_i}\otimes\vartheta_{B_i})(\mathcal{T}_{A^t_i\to B^t_i}\otimes\mathrm{id}_{A_i}\otimes\mathrm{id}_{B_i})\circ\mathcal{S}^{(i)}_{A_0\to A^t_iA_iB_i}\big\|_\diamond \\
&\qquad\qquad \le \big\|(\mathrm{id}_{A^t_{i+1}A_{i+1}}\otimes\vartheta_{B_{i+1}})\circ(\mathcal{L}_{A_i:B^t_iB_i\to A^t_{i+1}A_{i+1}:B_{i+1}})\circ(\mathrm{id}_{A_i}\otimes\vartheta_{B^t_iB_i})\big\|_\diamond \\
&\qquad\qquad\qquad \cdot\big\|\vartheta_{B^t_i}\circ\mathcal{T}_{A^t_i\to B^t_i}\big\|_\diamond\cdot\big\|(\mathrm{id}_{A^t_iA_i}\otimes\vartheta_{B_i})\circ\mathcal{S}^{(i)}_{A_0\to A^t_iA_iB_i}\big\|_\diamond \\
&\qquad\qquad = \big\|\vartheta_{d_2}\circ\mathcal{T}\big\|_\diamond\cdot\big\|(\mathrm{id}_{A^t_iA_i}\otimes\vartheta_{B_i})\circ\mathcal{S}^{(i)}_{A_0\to A^t_iA_iB_i}\big\|_\diamond
\end{aligned}
$$

for $i=1,\dots,m$, and $\|(\mathrm{id}_{A^t_1A_1}\otimes\vartheta_{B_1})\circ\mathcal{S}^{(1)}_{A_0\to A^t_1A_1B_1}\|_\diamond=\|(\mathrm{id}_{A^t_1A_1}\otimes\vartheta_{B_1})\circ\mathcal{L}_{A_0\to A^t_1A_1B_1}\|_\diamond=1$. From these relations we obtain inductively, recalling that A^t_{m+1} and A_{m+1} are trivial one-dimensional systems whereas $A_0=B_{m+1}$ are N-dimensional and abbreviating $\mathcal{S}:=\mathcal{S}^{(m+1)}_{A_0\to B_{m+1}}:\mathcal{M}_N\to\mathcal{M}_N$:

$$
\big\|\vartheta_N\circ\mathcal{S}\big\|_\diamond = \big\|\vartheta_{B_{m+1}}\circ\mathcal{S}^{(m+1)}_{A_0\to B_{m+1}}\big\|_\diamond \le \big\|\vartheta_{d_2}\circ\mathcal{T}\big\|^m_\diamond\,. \tag{12}
$$

Next, we bound the $\diamond$-norm error ε of the communication scheme (see Definition 3) from below by evaluating at the N-dimensional maximally entangled state $\omega_N=\omega_{A_0R}$ between the two N-dimensional systems A_0 and R and twirling over a representation of the unitary group $\mathcal{U}(N)$. For this we note that the twirled state is

$$
\int_{\mathcal{U}(N)} dU\,(U\otimes\overline{U})\,(\mathcal{S}\otimes\mathrm{id}_N)(\omega_N)\,(U\otimes\overline{U})^\dagger = p\omega_N+(1-p)(\mathbb{1}_{N^2}-\omega_N)/(N^2-1)
$$

with $p:=\mathrm{tr}\,(\omega_N\,(\mathcal{S}\otimes\mathrm{id}_N)(\omega_N))$ by Appendix A.

$$
\begin{aligned}
\varepsilon &= \frac12\big\|\mathrm{id}_N-\mathcal{S}\big\|_\diamond \ge \frac12\big\|((\mathrm{id}_N-\mathcal{S})\otimes\mathrm{id}_N)(\omega_N)\big\|_1 \\
&= \frac12\int_{\mathcal{U}(N)} dU\,\big\|(U\otimes\overline{U})(\omega_N-(\mathcal{S}\otimes\mathrm{id}_N)(\omega_N))(U\otimes\overline{U})^\dagger\big\|_1 \\
&\ge \frac12\Big\|\omega_N-\int_{\mathcal{U}(N)} dU\,(U\otimes\overline{U})\,(\mathcal{S}\otimes\mathrm{id}_N)(\omega_N)\,(U\otimes\overline{U})^\dagger\Big\|_1 \\
&= \frac12\big\|(1-p)\omega_N-(1-p)\big(\mathbb{1}_{N^2}-\omega_N\big)/(N^2-1)\big\|_1 = 1-p\,.
\end{aligned}
$$

We now derive an upper bound on p, by using similar steps starting from (12) and noting that $N(\vartheta_N\otimes\mathrm{id}_N)(\omega_N)=\mathbb{F}_N$ is the flip operator:

$$\begin{aligned}
\left\|\vartheta_{d_2}\circ\mathcal{T}\right\|_\diamond^m &\geq \left\|\vartheta_N\circ\mathcal{S}\right\|_\diamond \geq \left\|\big((\vartheta_N\circ\mathcal{S})\otimes\mathrm{id}_N\big)(\omega_N)\right\|_1\\
&= \int_{\mathcal{U}(N)} dU\,\left\|(\overline{U}\otimes\overline{U})\,\big((\vartheta_N\circ\mathcal{S})\otimes\mathrm{id}_N\big)(\omega_N)\,(\overline{U}^\dagger\otimes\overline{U}^\dagger)\right\|_1\\
&\geq \left\|\big(\vartheta_N\otimes\mathrm{id}_N\big)\left(\int_{\mathcal{U}(N)} dU\,(U\otimes\overline{U})\,(\mathcal{S}\otimes\mathrm{id}_N)(\omega_N)\,(U^\dagger\otimes\overline{U}^\dagger)\right)\right\|_1\\
&= \left\|\big(\vartheta_N\otimes\mathrm{id}_N\big)\left(p\omega_N+\frac{1-p}{N^2-1}(\mathbb{1}_{N^2}-\omega_N)\right)\right\|_1\\
&= \left\|\frac{Np+1}{N(N+1)}\,\frac{\mathbb{1}_{N^2}+\mathbb{F}_N}{2}+\frac{Np-1}{N(N-1)}\,\frac{\mathbb{1}_{N^2}-\mathbb{F}_N}{2}\right\|_1\\
&= |Np+1|/2+|Np-1|/2 \geq Np\,.
\end{aligned}$$

Combining this bound with the above relation between p and ε yields the claim. □

We can now state our capacity bound:

Theorem 7 (Strong converse upper bound on the two-way capacity $\mathcal{Q}_2(\mathcal{T})$)**.** *Let $\mathcal{T}:\mathcal{M}_{d_1}\to\mathcal{M}_{d_2}$ be a quantum channel. Then:*

$$\mathcal{Q}_2(\mathcal{T}) \leq \log_2\left(\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond\right)\,.$$

Moreover, let for each $\nu\in\mathbb{N}$ an $(N_\nu,m_\nu,\varepsilon_\nu)$-scheme for quantum communication over $\mathcal{T}$ assisted by two-way classical communication be given in such a way that $\lim_{\nu\to\infty}m_\nu=\infty$, and define the lower code rate *$R_{\inf}:=\liminf_{\nu\to\infty}\frac{\log_2(N_\nu)}{m_\nu}$. If $R_{\inf}>\log_2\left(\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond\right)$, then the $\diamond$-norm error ε_ν of the sequence converges to 1 (exponentially fast in m_ν).*

Proof. To prove the first statement, suppose that a rate $R=\limsup_{\nu\to\infty}\frac{\log_2(N_\nu)}{m_\nu}>\log_2(\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond)$ is achievable by schemes with parameters $(N_\nu,m_\nu,\varepsilon_\nu)$ (cf. Definition 3). Then, for any $\chi\in\mathbb{R}$ with $\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond<\chi<2^R$, we have $N_\nu\geq\chi^{m_\nu}$ for infinitely many values of $\nu\in\mathbb{N}$. Thus, by Lemma 3,

$$\limsup_{\nu\to\infty}\varepsilon_\nu \geq 1-\liminf_{\nu\to\infty}\frac{\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond^{m_\nu}}{N_\nu} \geq 1-\liminf_{\nu\to\infty}\left(\frac{\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond}{\chi}\right)^{m_\nu} > 0\,.$$

which contradicts the requirement $\lim_{\nu\to\infty}\varepsilon_\nu=0$.

The second statement follows similarly by noting that for any $\chi<2^{R_{\inf}}$, one has $N_\nu\geq\chi^{m_\nu}$ for almost all $\nu\in\mathbb{N}$. □

The second part of Theorem 7 means that $\log_2(\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond)$ is not only an upper bound on the two-way capacity $\mathcal{Q}_2(\mathcal{T})$, but even a *strong converse rate* for quantum communication over $\mathcal{T}$ assisted by free two-way classical communication. This generalizes the examples in [22, Section III], which are obtained for completely co-positive channels $\mathcal{T}$, where $\mathcal{Q}_2(\mathcal{T})=\log_2(\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond)=0$, and for the identity channel $\mathcal{T}=\mathrm{id}_d$, where $\log_2(\|\vartheta_d\circ\mathcal{T}\|_\diamond)=\log_2(d)=\mathcal{Q}_2(\mathcal{T})$. The *entanglement cost* $E_C(\mathcal{T})$ has been established as a strong converse rate for $\mathcal{Q}_2$ [3], although it can be larger than our bound. In recent work [29] is has been shown that the upper bound $\log_2(\|\vartheta_{d_2}\circ\mathcal{T}\|_\diamond)$ from Eq. (8), and improvements thereof, are strong converse rates for the usual quantum capacity $\mathcal{Q}$ from Definition 2, even when allowing for arbitrary LOCC operations at the beginning and the end of the protocol. The case of free LOCC communication during the protocol as in Definition 3 has however not been resolved in ref. [29].

Even the capacity bound on $\mathcal{Q}_2(\mathcal{T})$ from the first part of Theorem 7 seems to be new. In particular, an upper bound on $\mathcal{Q}_2(\mathcal{T})$ for pure-loss bosonic channels was derived in [28, Section 6] based on the *squashed entanglement of* $\mathcal{T}$. And while this was noted for pure-loss channels $\mathcal{T}$ to agree with the transposition bound (8) on $\mathcal{Q}(\mathcal{T})$, the question was left open whether the transposition bound is a general upper bound on two-way capacity $\mathcal{Q}_2(\mathcal{T})$.

D. Strong converse rate from tensor-stable positive maps

With ideas from the proofs of Lemma 3 and Theorem 6, we can use any surjective, unital and tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$ that is not completely positive to derive a strong converse rate for the usual quantum capacity $\mathcal{Q}(\mathcal{T})$ of any quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ (see Definition 2). The strong converse rate we obtain is

$$\frac{\log_2\left(\|\mathcal{P}^{-1}\circ\mathcal{T}\|_\diamond \|\mathcal{P}^*\left(\mathbb{1}_{d_2}\right)\|_\infty\right)\log_2(d_2)}{\log_2\left(\|(\mathcal{P}^*\otimes \mathrm{id}_{d_2})(\omega_{d_2})\|_1\right)}\,, \tag{13}$$

which is always at least as big as our upper bound on $\mathcal{Q}(\mathcal{T})$ from Theorem 6, due to $\|\mathcal{P}^*\|_\diamond \geq \|(\mathcal{P}^*\otimes \mathrm{id}_{d_2})(\omega_{d_2})\|_1$. The proof that (13) is a strong converse rate for the desired task follows from the following Lemma in the same way as Theorem 7 follows from Lemma 3.

Lemma 4. *Let $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be a quantum channel and $\mathcal{P} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_2}$ be a surjective, unital and tensor-stable positive map that is not completely positive, and let $\mathcal{P}^{-1}$ be any right-inverse of $\mathcal{P}$. Let $n, m \in \mathbb{N}$. Then:*

$$\inf_{\mathcal{E},\mathcal{D}} \frac{1}{2}\left\|\mathrm{id}_{d_2}^{\otimes n} - \mathcal{D}\circ\mathcal{T}^{\otimes m}\circ\mathcal{E}\right\|_\diamond \geq 1 - \frac{1}{d_2^{2n}} - \frac{\left(\|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty \|\mathcal{P}^{-1}\circ\mathcal{T}\|_\diamond\right)^m + 2}{\|(\mathcal{P}^*\otimes\mathrm{id}_{d_2})(\omega_{d_2})\|_1^n}\,,$$

where the infimum is over all quantum channels $\mathcal{E} : \mathcal{M}_{d_2}^{\otimes n} \to \mathcal{M}_{d_1}^{\otimes m}$ and $\mathcal{D} : \mathcal{M}_{d_2}^{\otimes m} \to \mathcal{M}_{d_2}^{\otimes n}$.

Proof. Fix $\mathcal{E}$, $\mathcal{D}$. As in the proof of Lemma 3, we bound the $\diamond$-norm with the maximally entangled state $\omega_{d_2^n}$ of dimension d_2^n and then twirl:

$$\frac{1}{2}\left\|\mathrm{id}_{d_2}^{\otimes n} - \mathcal{D}\circ\mathcal{T}^{\otimes m}\circ\mathcal{E}\right\|_\diamond \geq \frac{1}{2}\left\|\omega_{d_2^n} - \int_{\mathcal{U}(d_2^n)} dU\,(U\otimes\overline{U})\,\rho\,(U\otimes\overline{U})^\dagger\right\|_1 = 1-p\,,$$

where we denoted $\rho := ((\mathcal{D}\circ\mathcal{T}^{\otimes m}\circ\mathcal{E})\otimes\mathrm{id}_{d_2^n})(\omega_{d_2^n})$, and used $\int dU\,(U\otimes\overline{U})\,\rho\,(U\otimes\overline{U})^\dagger = p\omega_{d_2^n} + (1-p)(\mathbb{1}_{d_2^{2n}} - \omega_{d_2^n})/(d_2^{2n}-1)$ for $p := \mathrm{tr}\,(\omega_N\,\rho)$. For any unitary $U\in\mathcal{U}(d_2^n)$ we now define the unital quantum channel $\mathcal{C}_U : \mathcal{M}_{d_2}^{\otimes n} \to \mathcal{M}_{d_2}^{\otimes n}$ by $\mathcal{C}_U(X) := UXU^\dagger$ for all $X\in\mathcal{M}_{d_2}^{\otimes n}$ and reuse some arguments from the proof of Theorem 6:

$$\begin{aligned}
\|\mathcal{P}^*(\mathbb{1}_{d_2})\|_\infty^m \|\mathcal{P}^{-1}\circ\mathcal{T}\|_\diamond^m &\geq \int_{\mathcal{U}(d_2^n)} dU\,\left\|(\vartheta_{d_3}\circ\mathcal{P}^*\circ\vartheta_{d_2})^{\otimes n}\circ\mathcal{C}_U\circ\mathcal{D}\circ\mathcal{P}^{\otimes m}\right\|_\diamond \left\|(\mathcal{P}^{-1})^{\otimes m}\circ\mathcal{T}^{\otimes m}\circ\mathcal{E}\right\|_\diamond \\
&\geq \int_{\mathcal{U}(d_2^n)} dU\,\left\|\left(((\vartheta_{d_3}\circ\mathcal{P}^*\circ\vartheta_{d_2})^{\otimes n}\circ\mathcal{C}_U\circ\mathcal{D}\circ\mathcal{T}^{\otimes m}\circ\mathcal{E})\otimes\mathcal{C}_{\overline{U}}\right)(\omega_{d_2^n})\right\|_1 \\
&\geq \left\|\left((\vartheta_{d_3}\circ\mathcal{P}^*\circ\vartheta_{d_2})^{\otimes n}\otimes\mathrm{id}_{d_2^n}\right)\left(p\omega_{d_2^n} + \frac{1-p}{d_2^{2n}-1}(\mathbb{1}_{d_2^{2n}} - \omega_{d_2^n})\right)\right\|_1 \\
&\geq \frac{pd_2^{2n}-1}{d_2^{2n}-1}\left\|\left((\vartheta_{d_3}\circ\mathcal{P}^*\circ\vartheta_{d_2})^{\otimes n}\otimes\mathrm{id}_{d_2^n}\right)\left(\omega_{d_2^n}\right)\right\|_1 - \frac{1-p}{d_2^{2n}-1}\left\|\left((\vartheta_{d_3}\circ\mathcal{P}^*\circ\vartheta_{d_2})^{\otimes n}\otimes\mathrm{id}_{d_2^n}\right)\left(\mathbb{1}_{d_2^{2n}}\right)\right\|_1 \\
&\geq \left(p - \frac{1}{d_2^{2n}}\right)\|(\mathcal{P}^*\otimes\mathrm{id}_{d_2})\,(\omega_{d_2})\|_1^n - 2\,.
\end{aligned}$$

Converting this to an upper bound on p and combining with the above relation, we obtain the claim. $\square$

V. DISTILLATION SCHEMES FOR TENSOR-STABLE POSITIVE MAPS

A. Quantifying the distance from the completely positive maps

For a given hermiticity-preserving map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ we define a distance from the set of completely positive maps as

$$d_{\mathrm{CP}}\left(\mathcal{P}\right) := \frac{1}{2}\left(\|C_{\mathcal{P}}\|_1 - \mathrm{tr}\left(C_{\mathcal{P}}\right)\right). \tag{14}$$

By the Choi-Jamiolkowski isomorphism, $\mathcal{P}$ is completely positive iff $C_{\mathcal{P}} \geq 0$, i.e. iff $d_{\mathrm{CP}}(\mathcal{P}) = 0$, whereas $d_{\mathrm{CP}}(\mathcal{P}) > 0$ otherwise. The distance $d_{\mathrm{CP}}(\mathcal{P})$ is just the absolute value of the sum of negative eigenvalues of the Choi matrix $C_{\mathcal{P}}$ of $\mathcal{P}$. The following lemma gives a useful upper bound on d_{CP}:

Lemma 5. *Let $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be a positive map. If there exists a linear map $\mathcal{R} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_1}$ such that $\mathcal{R} \otimes \mathcal{P}$ is a positive map, then*

$$d_{CP}(\mathcal{P}) \leq \|\mathrm{id}_{d_1} - \mathcal{R}\|_\diamond \|\mathcal{P}\|_\diamond. \tag{15}$$

Proof. By elementary properties of the $\diamond$-norm and using positivity of $\mathcal{R} \otimes \mathcal{P}$ we have

$$\begin{aligned}\|\mathrm{id}_{d_1} - \mathcal{R}\|_\diamond \|\mathcal{P}\|_\diamond &\geq \|\omega_{d_1} - \left(\mathcal{R} \otimes \mathrm{id}_{d_1}\right)\left(\omega_{d_1}\right)\|_1 \|\mathcal{P}\|_\diamond \geq \|(\mathrm{id}_{d_1} \otimes \mathcal{P})(\omega_{d_1}) - (\mathcal{R} \otimes \mathcal{P})(\omega_{d_1})\|_1 \\ &\geq \inf_{X \geq 0} \|C_{\mathcal{P}} - X\|_1.\end{aligned}$$

And for any hermitian matrix H we have $\inf_{X\geq 0} \|H - X\|_1 = \frac{1}{2}\left(\|H\|_1 - \mathrm{tr}\left(H\right)\right)$ by Weyl's inequalities [4, Corollary III.2.2]. $\square$

Note that in the case of $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ being completely positive we can use $\mathcal{R} = \mathrm{id}_{d_1}$ in order to verify that $d_{\mathrm{CP}}(\mathcal{P}) = 0$ using Lemma 5.

To apply Lemma 5 for an n-tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ we have to find a suitable map $\mathcal{R} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_1}$ such that $\mathcal{R} \otimes \mathcal{P}$ is positive and $\mathcal{R}$ is close to the identity map. A convenient way to construct such an $\mathcal{R}$ is by considering generalized "coding schemes" of the form

$$\mathcal{R} = \sum_{i=1}^{m} \mathcal{D}_i \circ \mathcal{P}^{\otimes(n-1)} \circ \mathcal{E}_i \tag{16}$$

with completely positive maps $\mathcal{E}_i : \mathcal{M}_{d_1} \to \mathcal{M}_{d_1^{n-1}}$ and $\mathcal{D}_i : \mathcal{M}_{d_2^{n-1}} \to \mathcal{M}_{d_1}$. Indeed, as $\mathcal{P}^{\otimes n} \geq 0$ we have

$$\mathcal{R} \otimes \mathcal{P} = \sum_{i=1}^{m} \left(\mathcal{D}_i \otimes \mathrm{id}_{d_2}\right) \circ \left(\mathcal{P}^{\otimes n}\right) \circ \left(\mathcal{E}_i \otimes \mathrm{id}_{d_1}\right) \geq 0.$$

As $\mathcal{R} \otimes \mathcal{P}$ is positive for all choices of $\mathcal{E}_i$ and $\mathcal{D}_i$ in (16) we can optimize over these completely positive maps trying to make $\|\mathrm{id}_{d_1} - \mathcal{R}\|_\diamond$ as small as possible. This proves:

Corollary 1. *Let $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be an n-tensor-stable positive map, $\mathcal{P} \neq 0$. Then*

$$\frac{d_{CP}(\mathcal{P})}{\|\mathcal{P}\|_\diamond} \leq \inf_{m,\mathcal{E}_i,\mathcal{D}_i} \|\mathrm{id}_{d_1} - \sum_{i=1}^{m} \mathcal{D}_i \circ \mathcal{P}^{\otimes(n-1)} \circ \mathcal{E}_i\|_\diamond, \tag{17}$$

where the infimum is taken over $m \in \mathbb{N}$ and completely positive maps $\mathcal{E}_i, \mathcal{D}_i$.

The map $\mathcal{R}$ in (16) can be interpreted as a coding scheme where quantum information is encoded by the completely positive maps $\mathcal{E}_i$, sent through $(n-1)$ uses of the map $\mathcal{P}$, and decoded using the maps $\mathcal{D}_i$. The indices i can be seen as classical information which is communicated from the sender to the receiver for free and without noise. A special case of this technique for $m = 1$ and projectors $\mathcal{E}_1, \mathcal{D}_1$ has been used in ref. [27].

If $\mathcal{P}$ is tensor-stable positive we can take the limit $n \to \infty$ of the approximation error on the right-hand-side of (17). As the left-hand-side in (17) does not depend on n, the approximation error cannot vanish in the limit $n \to \infty$ unless $\mathcal{P}$ is completely positive.

As a first application of this idea we derive sufficient criteria for a quantum channel $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ to have $\mathcal{Q}_2(\mathcal{T}) = 0$ (see Definition 3). For this, note that the alternating application of the LOCC maps and the m channel uses in Eq. (11) can be written as

$$\mathcal{L}_{A_m:B^t_m B_m \to B_{m+1}} \circ \ldots \circ \mathcal{T}_{A^t_1 \to B^t_1} \circ \mathcal{L}_{A_0 \to A^t_1 A_1 : B_1}(\rho) \;=\; \sum_k K^B_k\, T^{\otimes m}\left(K^A_k\, \rho\, (K^A_k)^\dagger\right) (K^B_k)^\dagger\,,$$

where here the K^A_k (with a multi-index k) are simply all the time-ordered products of Kraus operators $(K^A_i \otimes |j\rangle)$ on the sender's side from (10) occurring in the LOCC maps in (11); similarly for K^B_k on the receiver's side. Thus, by defining completely positive maps $\mathcal{E}_k : \mathcal{M}_{A_0} \to \mathcal{M}^{\otimes m}_{d_1}$ and $\mathcal{D}_k : \mathcal{M}^{\otimes m}_{d_2} \to \mathcal{M}_{B_{m+1}}$ by $\mathcal{E}_k(X) := K^A_k\, X\, (K^A_k)^\dagger$ and $\mathcal{D}_k(Y) := K^B_k\, Y\, (K^B_k)^\dagger$, we have shown the existence of completely positive maps $\mathcal{E}_k$, $\mathcal{D}_k$ such that

$$\|\mathrm{id}_N - \sum_k \mathcal{D}_k \circ \mathcal{T}^{\otimes m} \circ \mathcal{E}_k\|_\diamond \;=\; 2\varepsilon\,,$$

whenever there exists a (N, m, ε)-scheme for LOCC-assisted quantum communication according to Definition 3. If the quantum channel $\mathcal{T}$ has positive two-way capacity $\mathcal{Q}_2(\mathcal{T}) > 0$, then for any fixed N, one can certainly transmit an N-dimensional quantum system with arbitrarily low error ($\varepsilon \to 0$) in the limit of arbitrarily many channel uses ($m \to \infty$).

Now consider the case where the quantum channel $\mathcal{T}$ with $\mathcal{Q}_2(\mathcal{T}) > 0$ is of the form $\mathcal{T} = \sum_j \mathcal{V}_j \circ \mathcal{P} \circ \mathcal{W}_j$ for a tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_4}$ that is *not* completely positive and for completely positive maps $\mathcal{V}_j : \mathcal{M}_{d_1} \to \mathcal{M}_{d_3}$ and $\mathcal{W}_j : \mathcal{M}_{d_4} \to \mathcal{M}_{d_2}$. Then, by setting $N := d_3$ in the previous paragraph, we have

$$\begin{aligned} 0 \;&=\; \lim_{m\to\infty} \inf_{\mathcal{D}_k,\mathcal{E}_k} \|\mathrm{id}_{d_3} - \sum_k \mathcal{D}_k \circ \mathcal{T}^{\otimes m} \circ \mathcal{E}_k\|_\diamond \\ &=\; \lim_{m\to\infty} \inf_{\mathcal{D}_k,\mathcal{E}_k} \left\|\mathrm{id}_{d_3} - \sum_{k,j_1,\ldots,j_m} \left(\mathcal{D}_k \circ \bigotimes_{\ell=1}^{m} \mathcal{V}_{j_\ell}\right) \circ \mathcal{P}^{\otimes m} \circ \left(\bigotimes_{\ell=1}^{m} \mathcal{W}_{j_\ell} \circ \mathcal{E}_k\right)\right\|_\diamond . \end{aligned}$$

But this leads to a contradiction since, by interpreting $\widetilde{\mathcal{D}}_i := \mathcal{D}_k \circ (\bigotimes_{\ell=1}^m \mathcal{V}_{j_\ell})$ and $\widetilde{\mathcal{E}}_i := (\bigotimes_{\ell=1}^m \mathcal{W}_{j_\ell}) \circ \mathcal{E}_k$ with the multi-index $i \equiv (k, j_1, \ldots, j_m)$ as the encoding and decoding maps for the map $\mathcal{P}$, Corollary 1 would imply that $d_{\mathrm{CP}}(\mathcal{P}) = 0$, meaning that $\mathcal{P}$ would be completely positive contrary to assumption. This proves the following:

Corollary 2. *Let $\mathcal{T}$ be a quantum channel of the form $\mathcal{T} = \sum_j \mathcal{V}_j \circ \mathcal{P} \circ \mathcal{W}_j$ for a tensor-stable positive map $\mathcal{P}$ that is not completely positive and for completely positive maps $\mathcal{V}_j$, $\mathcal{W}_j$. Then:*

$$\mathcal{Q}_2(\mathcal{T}) = 0.$$

The special case $\mathcal{T} = \mathcal{V} \circ \vartheta$ of this theorem, i.e. where $\mathcal{T}$ is completely co-positive, was already established in [24]. It however appears that Corollary 2 could give new channels $\mathcal{T} = \mathcal{P} \circ \mathcal{W}$ or $\mathcal{T} = \mathcal{V} \circ \mathcal{P}$ with $\mathcal{Q}(\mathcal{T}) = 0$ beyond Theorem 6 and Eq. (9), at least when $\mathcal{P}$ does not possess a right- or left-inverse.

Using Corollary 2 one can show that any *non-trivial* tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ will immediately yield new channels $\mathcal{T} : \mathcal{M}_d \to \mathcal{M}_d$ with $\mathcal{Q}_2(\mathcal{T}) = 0$ (for both $d = d_1$ and $d = d_2$). To see this, note that by writing the separable map $\mathcal{S}$ from Lemma 6 into single Kraus operators as in Section V C, we can construct completely positive maps $\mathcal{V}_j$, $\mathcal{W}_j$ such that $\sum_j \mathcal{V}_j \circ \mathcal{P} \circ \mathcal{W}_j = \mathcal{W}_{\widetilde{p}}$, where $\mathcal{W}_{\widetilde{p}} : \mathcal{M}_d \to \mathcal{M}_d$ with $\widetilde{p} \in [-1, 0)$ is a quantum channel from the family (2) whose Choi matrix is an entangled Werner state. Thus, by Corollary 2 and the depolarizing idea from Section V C, all the channels $\mathcal{W}_p$ with $p \in [\widetilde{p}, 0)$ have vanishing two-way capacity $\mathcal{Q}_2(\mathcal{W}_p) = 0$, although these channels are not detected as such by the existing criterion from [24] (or by Theorem 7) as they are not completely co-positive. The channels constructed in this way are however already known to have vanishing one-way quantum capacity $\mathcal{Q}(\mathcal{W}_p) = 0$, since they possess a symmetric extension (note, the case $d \leq 2$ does not occur here due to Theorem 2) and are thus anti-degradable [2, 5, 26].

In the following chapters we will use another way of thinking about coding schemes of the form (16). Recall that a completely positive map $\mathcal{S} : \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_3} \to \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_4}$ is called separable if its Kraus operators are product operators $\{A_i \otimes B_i\}_{i=1}^m$, i.e. $\mathcal{S}(X) = \sum_{i=1}^m (A_i \otimes B_i) X (A_i \otimes B_i)^\dagger$ for all $X \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_3}$.

The application of a separable map $\mathcal{S} : \mathcal{M}_{d_1^{n-1}} \otimes \mathcal{M}_{d_2^{n-1}} \to \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_1}$ to $(n-1)$ copies of the Choi-matrix $C_\mathcal{P}$ of some linear map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ corresponds via the Choi-Jamiolkowski isomorphism to a map $\mathcal{R}$ as

$$C_\mathcal{R} = \mathcal{S}\left(C_\mathcal{P}^{\otimes(n-1)}\right),$$

with $\mathcal{R}(X) = \sum_{i=1}^m B_i \mathcal{P}^{\otimes(n-1)}\left(A_i^T X \overline{A_i}\right) B_i^\dagger$. The map $\mathcal{R}$ is of the form (16), which by slightly modifying the proof of Lemma 5 implies:

Corollary 3. *Let $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be an n-tensor-stable positive map. Then*

$$\frac{d_{CP}(\mathcal{P})}{\|\mathcal{P}\|_\diamond} \leq \inf_{\mathcal{S}\ sep} \|\omega_{d_1} - \mathcal{S}\left(C_\mathcal{P}^{\otimes(n-1)}\right)\|_1, \tag{18}$$

where the infimum is taken over all separable completely positive maps $\mathcal{S} : \mathcal{M}_{d_1^{n-1}} \otimes \mathcal{M}_{d_2^{n-1}} \to \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_1}$.

If the Choi-matrix $C_\mathcal{P}$ is a quantum state, then the problem of finding separable maps $\mathcal{S}$ to minimize the error on the right-hand-side of (18) is well-studied in quantum information theory: A state $C_\mathcal{P} \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ is distillable iff there exists a sequence of LOCC-maps $\mathcal{S}_n$ such that $\mathcal{S}_n\left(C_\mathcal{P}^{\otimes n}\right) \to \omega_{d_1}$. As LOCC-maps are in particular separable this sequence leads to a vanishing (in the limit $n \to \infty$) right-hand-side in (18).

Note that any positive map that is not completely co-positive has an NPPT, but not necessarily positive, Choi-matrix. We generalize distillation schemes from quantum states to arbitrary block-positive matrices to show (using Corollary 3) that tensor-stable positivity implies complete positivity for certain classes of non-completely co-positive maps.

B. Proof of Theorem 2 and Theorem 4

To prove Theorem 2 and Theorem 4 we will use the theory of entanglement distillation. For convenience we collect some basic definitions and results in Appendix A. The central result we will need is Lemma 8, which shows that applying the twirl [30] to a block-positive and NPPT matrix yields (up to normalization) a Werner state, i.e. it yields in particular a positive matrix. This allows us to extend the theory of entanglement distillation to block-positive matrices. We will start with a basic lemma:

Lemma 6 (Werner states from positive maps)**.** *Let $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be a positive map and $d \in \{d_1, d_2\}$. If $\mathcal{P}$ is not completely co-positive, then there exists a separable completely positive map $\mathcal{S} : \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \to \mathcal{M}_d \otimes \mathcal{M}_d$ such that $\mathcal{S}(C_{\mathcal{P}})$ is an entangled d-dimensional Werner state (see Appendix A).*

Proof. This proof works similar to the protocol introduced in [15] for states. Consider $d = d_2$ now, and we will treat the case $d = d_1$ later. As $\mathcal{P}$ is not completely co-positive, there exists a normalized vector $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ with $\langle\psi|C_{\mathcal{P}}^{T_2}|\psi\rangle < 0$. Express this vector as $|\psi\rangle = (A \otimes \mathbb{1}_{d_2})|\Omega_{d_2}\rangle$ for some $d_1 \times d_2$ matrix A and the maximally entangled state $|\Omega_{d_2}\rangle$. Now define a new linear map $\mathcal{P}'$ via the Choi-Jamiolkowski isomorphism by applying a local filtering operation

$$C_{\mathcal{P}'} := (A^\dagger \otimes \mathbb{1}_{d_2})C_{\mathcal{P}}(A \otimes \mathbb{1}_{d_2}) \in \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2},$$

i.e. $\mathcal{P}'(X) := \mathcal{P}\left(\overline{A} X A^T\right)$ for all $X \in \mathcal{M}_{d_2}$.

The matrix $C_{\mathcal{P}'}$ is block-positive and fulfills $\mathrm{tr}\,(C_{\mathcal{P}'}\mathbb{F}_{d_2}) = d_2\langle\psi|C_{\mathcal{P}}^{T_2}|\psi\rangle < 0$. Therefore we can use Lemma 8 and conclude that applying the UU-twirl leads to a positive matrix. After normalization we obtain a Werner state

$$\rho_W = \frac{1}{\mathrm{tr}\,(C_{\mathcal{P}'})} \int_{U \in \mathcal{U}(d_2)} (U \otimes U)\, C_{\mathcal{P}'}\, (U \otimes U)^\dagger \,\mathrm{dU} \in \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}.$$

Due to $\mathrm{tr}\,(C_{\mathcal{P}'}\mathbb{F}_{d_2}) < 0$, this state is entangled. Finally, the composition of the twirl (which is separable, see Appendix A) with the filtering map is a separable completely positive map.

If one chooses $d = d_1$, then write $|\psi\rangle = (\mathbb{1}_{d_1} \otimes B)|\Omega_{d_1}\rangle$ with a $d_2 \times d_1$-matrix B, and define $C_{\mathcal{P}'} := (\mathbb{1}_{d_1} \otimes B^T)C_{\mathcal{P}}(\mathbb{1}_{d_1} \otimes \overline{B})$. The proof goes then through similarly.

□

From this Lemma we get:

proof of Theorem 4. For $d \in \{d_1, d_2\}$ the Choi-matrix $C_{\mathcal{P}}$ of every non-trivial tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ yields an entangled Werner state by the application of a separable completely positive map $\mathcal{S}$ according to Lemma 6. If this Werner state is distillable, there exists a sequence of separable (even LOCC) completely positive maps $(\mathcal{S}_n)_{n\in\mathbb{N}}$ such that $\|\omega_d - \mathcal{S}_n \circ \mathcal{S}^{\otimes n}\left(C_{\mathcal{P}}^{\otimes n}\right)\|_1 \to 0$ as $n \to \infty$. But then Corollary 3 implies that $\mathcal{P}$ is completely positive contradicting the assumptions.

□

proof of Theorem 2. As all entangled Werner states on $\mathcal{M}_2 \otimes \mathcal{M}_2$ are distillable [15, 17] there is no non-trivial tensor-stable positive map $\mathcal{P} : \mathcal{M}_2 \to \mathcal{M}_d$ or $\mathcal{P} : \mathcal{M}_d \to \mathcal{M}_2$ for $d \in \mathbb{N}$ according to Theorem 4.

□

C. Proof of Theorem 3

Using the techniques from section V B we can define one-parameter families of non-trivial positive maps such that there exists a non-trivial tensor-stable positive map iff it exists within this family.

proof of Theorem 3. ad (ii): For $p \in [-1, 0)$ the Werner state $\rho_W^{(p)} \in \mathcal{M}_{d^2}$ is NPPT. Therefore the map $\mathcal{P}_p : \mathcal{M}_{d^2} \to \mathcal{M}_{d^2}$ is neither completely positive nor completely co-positive, as its Choi-matrix is $\rho_W^{(p)} \otimes \left(\rho_W^{(p)}\right)^{T_2}$.

ad (i): For a non-trivial tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$, neither $\mathcal{P}$ nor $\vartheta_{d_2} \circ \mathcal{P}$ are completely co-positive. According to Lemma 6 there exist $p_1, p_2 \in [-1, 0)$ and separable completely positive maps $\mathcal{S}_1, \mathcal{S}_2 : \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \to \mathcal{M}_d \otimes \mathcal{M}_d$ such that

$$\begin{aligned}
\rho_W^{(p_1)} &= \mathcal{S}_1(C_{\mathcal{P}}) \,, \\
\rho_W^{(p_2)} &= \mathcal{S}_2(C_{\vartheta_{d_2} \circ \mathcal{P}}) = \mathcal{S}_2 \circ (\mathrm{id}_{d_1} \otimes \vartheta_{d_2})(C_{\mathcal{P}}).
\end{aligned}$$

It is obvious that for the separable completely positive map $\mathcal{S}_2(X) = \sum_i (A_i \otimes B_i) X (A_i \otimes B_i)^\dagger$ the map $\tilde{\mathcal{S}}_2 = (\mathrm{id}_d \otimes \vartheta_d) \circ \mathcal{S}_2 \circ (\mathrm{id}_{d_1} \otimes \vartheta_{d_2})$ is again separable completely positive. Thus, the separable map $\mathcal{S}_1 \otimes \tilde{\mathcal{S}}_2$ applied to two tensor copies of $C_{\mathcal{P}}$ gives:

$$\left(\mathcal{S}_1 \otimes \tilde{\mathcal{S}}_2\right)(C_{\mathcal{P}} \otimes C_{\mathcal{P}}) = \rho_W^{(p_1)} \otimes \left(\rho_W^{(p_2)}\right)^{T_2} .$$

By applying a depolarizing channel $\mathcal{D}_\alpha : \mathcal{M}_d \to \mathcal{M}_d$ of the form $\mathcal{D}_\alpha(X) = (1-\alpha)\mathrm{tr}\,(X)\, \frac{\mathbb{1}_d}{d} + \alpha X$ (with α chosen appropriately) to one half of either $\rho_W^{(p_1)}$ (if $p_1 < p_2$) or $\left(\rho_W^{(p_2)}\right)^{T_2}$ (if $p_1 > p_2$) we can increase the corresponding parameter to obtain the desired state $\rho_W^{(p)} \otimes \left(\rho_W^{(p)}\right)^{T_2}$ with $p = \max(p_1, p_2) < 0$. Thus, there exists a separable and completely positive map $\mathcal{R} : (\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2})^{\otimes 2} \to (\mathcal{M}_d \otimes \mathcal{M}_d)^{\otimes 2}$, given by the composition of $\mathcal{S}_1 \otimes \tilde{\mathcal{S}}_2$ with $(\mathcal{D}_\alpha \otimes \mathrm{id}_d) \otimes \mathrm{id}_{d^2}$ or $\mathrm{id}_{d^2} \otimes (\mathcal{D}_\alpha \otimes \mathrm{id}_d)$, such that $C_{\mathcal{P}_p} = \rho_W^{(p)} \otimes \left(\rho_W^{(p)}\right)^{T_2} = \mathcal{R}(C_{\mathcal{P}^{\otimes 2}}) = \sum_i (C_i \otimes D_i) C_{\mathcal{P}^{\otimes 2}} (C_i \otimes D_i)^\dagger$ where $\mathcal{P}_p$ was defined in (1). By the Choi-Jamiolkowski isomorphism we can thus write $\mathcal{P}_p(X) = \sum_i D_i \mathcal{P}^{\otimes 2}(C_i^T X \overline{C}_i) D_i^\dagger$, which shows that $\mathcal{P}_p$ is tensor-stable positive as $\mathcal{P}$ was. □

Note that the construction from the proof of Theorem 3 also works for an n-tensor-stable positive map $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$. The positive map $\mathcal{P}_p$ of the from (1) obtained this way is $\lfloor \frac{n}{2} \rfloor$-tensor-stable positive.

D. Generalization of the reduction criterion

In this section we will generalize the reduction criterion and use the well-known recurrence protocol [15] to prove bounds on $d_{\mathrm{CP}}(\mathcal{P})$ for an n-tensor-stable positive map $\mathcal{P}$.

We will need the following lemma (an analogue of Lemma 6):

Lemma 7 (Reduction criterion)**.** *Let $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be a positive map. Let $\Gamma_d : \mathcal{M}_d \to \mathcal{M}_d$ denote the reduction map $\Gamma(X) := \mathrm{tr}\,(X)\, \mathbb{1}_d - X$. Then we have:*

1. *If $\Gamma_{d_2} \circ \mathcal{P}$ is not completely positive there exists a separable completely positive map $\mathcal{S} : \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \to \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}$ s.th. $\mathcal{S}\,(C_{\mathcal{P}})$ is an entangled isotropic state (see Appendix A).*

2. *If $\mathcal{P} \circ \Gamma_{d_1}$ is not completely positive there exists a separable completely positive map $\mathcal{S}$: $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} \to \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_1}$ s.th. $\mathcal{S}(C_\mathcal{P})$ is an entangled isotropic state (see Appendix A).*

Proof. Again the proof works similar to the protocol introduced in [15] for states. We will start with the first case.

As $\Gamma_{d_2} \circ \mathcal{P}$ is not completely positive the Choi-matrix $C_{\Gamma_{d_2} \circ \mathcal{P}} = \mathcal{P}^*(\mathbb{1}_{d_2})^T \otimes \mathbb{1}_{d_2} - C_\mathcal{P}$, derived using Lemma 1, is not positive. Thus, there exists a normalized vector $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ with

$$\frac{1}{d_1}\langle\psi|\mathcal{P}^*(\mathbb{1}_{d_2})^T \otimes \mathbb{1}_{d_2}|\psi\rangle < \langle\psi|C_\mathcal{P}|\psi\rangle.$$

Express this vector as $|\psi\rangle = (A \otimes \mathbb{1}_{d_2})|\Omega_{d_2}\rangle$ for some $d_1 \times d_2$-matrix A and define a new linear map $\mathcal{P}'$ with Choi matrix

$$C_{\mathcal{P}'} = (A^\dagger \otimes \mathbb{1}_{d_2})C_\mathcal{P}(A \otimes \mathbb{1}_{d_2}) \in \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}.$$

Note that by construction $\mathcal{P}'$ is a tensor-stable positive map obtained from $\mathcal{P}$ via a separable (even local) completely positive map. Furthermore we have using Lemma 1

$$\begin{aligned}\mathrm{tr}\,(C_{\mathcal{P}'}) =& \langle\Omega_{d_1}|AA^\dagger \otimes \mathcal{P}^*(\mathbb{1}_{d_2})|\Omega_{d_1}\rangle \\ =& \langle\Omega_{d_2}|A^\dagger\mathcal{P}^*(\mathbb{1}_{d_2})^T A \otimes \mathbb{1}_{d_2}|\Omega_{d_2}\rangle \cdot \frac{d_2}{d_1} \\ =& \langle\psi|\mathcal{P}^*(\mathbb{1}_{d_2})^T \otimes \mathbb{1}_{d_2}|\psi\rangle \cdot \frac{d_2}{d_1} < d_2\langle\psi|C_\mathcal{P}|\psi\rangle.\end{aligned}$$

Therefore we have $\mathrm{tr}\,(C_{\mathcal{P}'}\omega_{d_2}) = \langle\psi|C_\mathcal{P}|\psi\rangle > \frac{\mathrm{tr}(C_{\mathcal{P}'})}{d_2} > 0$. Note that $\mathrm{tr}\,(C_{\mathcal{P}'}) > 0$ as $C_{\mathcal{P}'}$ is block-positive and $C_{\mathcal{P}'} \neq 0$ as $\Gamma_{d_2} \circ \mathcal{P}$ is not completely positive. By applying Lemma 8 we conclude that

$$\rho_I^{(p)} = \frac{1}{\mathrm{tr}\,(C_{\mathcal{P}'})}\int_{U \in \mathcal{U}(d_2)} \left(U \otimes \overline{U}\right) C_{\mathcal{P}'} \left(U \otimes \overline{U}\right)^\dagger \mathrm{dU} \in \mathcal{M}_{d_2} \otimes \mathcal{M}_{d_2}.$$

is an isotropic state, with $p = \frac{\mathrm{tr}\left(C_{\mathcal{P}'}\omega_{d_2}\right)}{\mathrm{tr}(C_{\mathcal{P}'})} > \frac{1}{d_2}$. Thus this state is entangled and the composition of the twirl (which is a separable completely positive map, see Appendix A) with the filtering map is separable and completely positive.

The second part works similar to the first part. Note that $\mathcal{P} \circ \Gamma_{d_1}$ not being completely positive is equivalent to the existence of a normalized vector $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ with

$$\frac{1}{d_1}\langle\psi|\mathbb{1}_{d_1} \otimes \mathcal{P}(\mathbb{1}_{d_1})|\psi\rangle < \langle\psi|C_\mathcal{P}|\psi\rangle.$$

Express this vector as $|\psi\rangle = (\mathbb{1}_{d_1} \otimes B)|\Omega_{d_1}\rangle$ for some $d_2 \times d_1$-matrix B and define a new linear map $\mathcal{P}'$ with Choi matrix

$$C_{\mathcal{P}'} = (\mathbb{1}_{d_1} \otimes B^\dagger)C_\mathcal{P}(\mathbb{1}_{d_1} \otimes B) \in \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_1}.$$

Now by a similar calculation as before we have $\mathrm{tr}\,(C_{\mathcal{P}'}) = \langle\psi|\mathbb{1}_{d_1} \otimes \mathcal{P}(\mathbb{1}_{d_1})|\psi\rangle < d_1\langle\psi|C_\mathcal{P}|\psi\rangle$. The rest of the proof works the same as for the first case.

□

Lemma 7 shows how to obtain an entangled isotropic state from the Choi-matrix of a positive map violating the reduction criterion. It is well-known that these states are distillable by the recurrence protocol [15]. More precisely there exists a separable completely positive map $\mathcal{S} : \mathcal{M}_{d^2} \to \mathcal{M}_d$ with

$$T_{U\overline{U}} \circ \mathcal{S}\left((\rho_I^{(p)})^{\otimes 2} \right) = \rho_I^{(r(p))}, \tag{19}$$

where $T_{U\overline{U}}$ denotes the $U\overline{U}$-twirl and where

$$r(p) = \frac{1 + p\left(pd(d^2 + d - 1) - 2\right)}{p^2 d^3 - 2pd + d^2 + d - 1}.$$

It can be easily seen that for $p > \frac{1}{d}$ we have $r^{(m)}(p) \to 1$ as $m \to \infty$, where the notation $r^{(m)}(p)$ means that we concatenate m applications of the function r, i.e. $r^{(m)}(p) := r(r(\ldots r(p)))$. Therefore iterating the protocol using up many copies of the input state $\rho_I^{(p)}$ leads to isotropic states close to the maximally entangled state ω_d. In the following we use this protocol and Corollary 3 to upper-bound the distance of an n-tensor-stable positive map violating the reduction criterion to the cone of completely positive maps.

Note that the original protocol [15] has a sufficiently small but non-zero probability of failure. As the separable completely positive maps $\mathcal{S}$ in Corollary 3 do not have to be trace-preserving we can avoid the possibility of failure by choosing only the Kraus operators corresponding to a successful measurement for $\mathcal{S}$.

Theorem 8 (Bound from the recurrence protocol)**.** *Let $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ be a positive map and such that $\mathcal{P} \circ \Gamma_{d_1}$ is not completely positive, i.e.*

$$p := \sup_{|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}} \frac{\langle\psi| C_{\mathcal{P}} |\psi\rangle}{\langle\psi| \mathbb{1}_{d_1} \otimes \mathcal{P}\left(\mathbb{1}_{d_1}\right) |\psi\rangle} \tag{20}$$

$$= \lambda_{max}\left[\left(\mathbb{1}_{d_1} \otimes \mathcal{P}\left(\mathbb{1}_{d_1}\right)\right)^{-1/2} C_{\mathcal{P}} \left(\mathbb{1}_{d_1} \otimes \mathcal{P}\left(\mathbb{1}_{d_1}\right)\right)^{-1/2} \right] \in (1/d_1, 1], \tag{21}$$

using generalized inverses and denoting by $\lambda_{max}[\,\cdot\,]$ the maximum eigenvalue. If $\mathcal{P}$ is n-tensor-stable positive, then

$$d_{CP}(\mathcal{P}) \leq 2(1-p)\left(g_{d_1}(p)\right)^{\lfloor \log_2(n-1) \rfloor} \tag{22}$$

where $g_d(p) := \frac{d(d+1) - 2 - p(pd(d-1) + 2(d-1))}{(1-p)(p(pd^3 - 2d) + d^2 + d - 1)}$. Note that $g_d(p) \in [0, 1)$ for $p \in (\frac{1}{d}, 1]$.

Proof. By Lemma 7 (and its proof) there is a separable completely positive map $\mathcal{S}_1 : \mathcal{M}_{(d_1 d_2)^{(n-1)}} \to \mathcal{M}_{(d_2^1)^{(n-1)}}$ with $\mathcal{S}_1(C_{\mathcal{P}}^{\otimes(n-1)}) = (\rho_I^{(p)})^{\otimes(n-1)}$. By (19) we can apply the recurrence protocol for $\lfloor \log_2(n-1) \rfloor$ levels yielding $\rho_I^{(p')} \in \mathcal{M}_{d_1^2}$ with $p' = r^{(\lfloor \log_2(n-1) \rfloor)}(p)$. Composing these two protocols gives a separable completely positive map $\mathcal{S} : \mathcal{M}_{(d_1 d_2)^{(n-1)}} \to \mathcal{M}_{(d_1^2)}$ with

$$\|\omega_{d_1} - \mathcal{S}\left(C_{\mathcal{P}}^{\otimes(n-1)}\right)\|_1 = 2(1-p').$$

A simple calculation gives

$$1 - p' = 1 - r^{(\lfloor \log_2(n-1) \rfloor)}(p) \leq \left(g_{d_1}(p)\right)^{\lfloor \log_2(n-1) \rfloor}(1-p),$$

since $g_{d_1}(p) = \frac{1-r(p)}{1-p}$ is strictly monotonously decreasing for $p \in (\frac{1}{d_1}, 1)$ and is equal to the expression above.

By Corollary 3 we finally have

$$\frac{d_{\mathrm{CP}}(T)}{\|\mathcal{P}\|_\diamond} \leq 2(1-p)\left(g_{d_1}(p)\right)^{\lfloor \log_2(n-1) \rfloor}.$$

□

For dimension $d_1 = 2$ we have $\Gamma_2 = \mathcal{U} \circ \vartheta_2$ for some unitary conjugation $\mathcal{U} : \mathcal{M}_2 \to \mathcal{M}_2$. Therefore the positive maps $\mathcal{P} : \mathcal{M}_d \to \mathcal{M}_2$ such that $\Gamma_2 \circ \mathcal{P}$ is completely positive are precisely the completely co-positive maps. For general maps $\mathcal{P} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$, if $\Gamma_{d_2} \circ \mathcal{P}$ is not completely positive, then $\vartheta_{d_2} \circ \mathcal{P}$ is not completely positive, i.e. $\mathcal{P}$ is not completely co-positive.

VI. CONCLUSION

We have introduced the notions of *n-tensor-stable positive* and *tensor-stable positive maps*, and have investigated whether such maps exist outside of the cones of completely positive or completely co-positive maps. We showed that tensor-stable positive maps outside these families would provide novel bounds on the quantum capacity of quantum channels. Our main technique was to apply coding schemes from distillation theory to block-positive operators rather than to density matrices. Thereby and by the Choi correspondence between block-positive operators and positive maps, we related the existence of tensor-stable positive maps to the existence of NPPT bound entanglement. We also showed that the cb-norm bound coming from the transposition map yields a strong converse rate for the two-way quantum capacity $\mathcal{Q}_2$, and established strong converse rates on the usual quantum capacity $\mathcal{Q}$ coming from other tensor-stable positive maps.

The main question left open by our work is whether non-trivial tensor-stable positive maps exist at all, i.e. maps outside of the above cones that are n-tensor-stable positive for all $n \in \mathbb{N}$. We have reduced this existence question to certain one-parameter families of candidate maps (Theorem 3). But can this reduction be used to decide the existence, or at least to prove the non-existence result of Theorem 2 directly?

Furthermore, the converse of Theorem 4 is open: Does the existence of NPPT bound entanglement imply the existence of non-trivial tensor-stable positive maps? Note that such an equivalence would be rather different from the equivalence result of [8], linking NPPT bound entanglement to completely co-positive maps which are not completely positive but such that all their tensor powers are 2-positive. The map of interest in the latter scenario lies within the completely co-positive cone, which is among the trivial cases for our work.

Our existence result of an n-tensor stable positive map for every $n \in \mathbb{N}$ (Theorem 1) is analogous to the result in the theory of entanglement distillation which guarantees for every n the existence of NPPT states that are not n-copy distillable [8, 9]. Our Lemma 2 however appears to be to weak to show the existence of a map that is n-tensor stable positive for all n, see Eq. (5).

Finally, note that one may relate the existence of a tensor-stable positive map to the stability of operator norms under tensor products [23]: a positive unital map $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ is n-tensor stable positive if and only if the induced operator norm $\|\mathcal{T}^{\otimes n}\|_{\infty\to\infty} = 1$.

Acknowledgements. We thank Sergey Filippov for useful discussions and the Isaac Newton Institute for the Mathematical Sciences for its hospitality and support during the programme "Quantum Control Engineering" (M.M.W.). We acknowledge financial support from the CHIST-ERA/BMBF project CQC (A.M.H. and M.M.W.), the Marie Curie Intra-European Fellowship QUINTYL (D.R.), and the John Templeton Foundation (D.R. and M.M.W.). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

Appendix A: Twirling and families of symmetric matrices

The main ingredient in the distillation protocols we will apply is the UU**-twirl** operation $T_{UU} : \mathcal{M}_d \otimes \mathcal{M}_d \to \mathcal{M}_d \otimes \mathcal{M}_d$ [30], defined as

$$T_{UU}(X) := \int_{U\in\mathcal{U}(d)} (U \otimes U)\, X\, (U \otimes U)^\dagger \,\mathrm{d}U.$$

An application of the Schur-Weyl duality gives [30] (for $d \geq 2$)

$$\int_{U\in\mathcal{U}(d)} (U \otimes U)\, X\, (U \otimes U)^\dagger \,\mathrm{dU} = \left[\frac{\mathrm{tr}\,(X)}{d^2-1} - \frac{\mathrm{tr}\,(X\mathbb{F}_d)}{d(d^2-1)}\right] (\mathbb{1}_d \otimes \mathbb{1}_d) - \left[\frac{\mathrm{tr}\,(X)}{d(d^2-1)} - \frac{\mathrm{tr}\,(X\mathbb{F}_d)}{d^2-1}\right] \mathbb{F}_d. \tag{A1}$$

It is easy to verify that this matrix is positive iff $\mathrm{tr}\,(X\mathbb{F}_d) \in [-\mathrm{tr}\,(X)\,, \mathrm{tr}\,(X)]$ (and $\mathrm{tr}\,(X) \geq 0$); the twirled matrix has positive partial transpose iff $\mathrm{tr}\,(X\mathbb{F}_d) \in [0, d\mathrm{tr}\,(X)]$ (and $\mathrm{tr}\,(X) \geq 0$).

Using unitary 2-designs [25] it is well-known that the twirl is a separable completely positive map, i.e. there exists a finite set of unitary product matrices $\{U_i \otimes U_i\}_{i=1}^m$ such that $T_{UU}(X) = \frac{1}{m}\sum_{i=1}^m (U_i \otimes U_i)\, X\, (U_i \otimes U_i)^\dagger$.

States of the form (A1) are clearly invariant under the UU-twirl operation and are called **Werner states** [30]. We denote these states by $\rho_W^{(p)}$, parametrized by $p := \mathrm{tr}\left(\rho_W^{(p)}\mathbb{F}_d\right) \in [-1, 1]$ and satisfying $\mathrm{tr}\left(\rho_W^{(p)}\right) = 1$. It is well-known that these states are entangled (and NPPT) for $p \in [-1, 0)$ and separable for $p \in [0, 1]$ (thus, PPT). Furthermore for $d = 2$ all entangled Werner states are distillable [15, 17]. But for $d > 2$ it is *not* known whether all entangled Werner states are distillable.

By partially transposing the matrices of form (A1) we obtain matrices invariant under the $U\overline{U}$-twirl operation, i.e. invariant under the operation

$$X \mapsto \int_{U\in\mathcal{U}(d)} \left(U \otimes \overline{U}\right) X \left(U \otimes \overline{U}\right)^\dagger \mathrm{d}U.$$

We will denote the states obtained in this way by $\rho_I^{(p)}$, which are the **isotropic states** [15] parametrized by $p := \mathrm{tr}\left(\rho_I^{(p)}\omega\right) = \frac{1}{d}\mathrm{tr}\left((\rho_I^{(p)})^{T_2}\mathbb{F}_d\right) \in [0, 1]$ and normalized to $\mathrm{tr}\left(\rho_I^{(p)}\right) = 1$. These states are entangled (and NPPT) for $p \in (\frac{1}{d}, 1]$, and separable for $p \in [0, \frac{1}{d}]$ (thus, PPT). It is well-known that all entangled isotropic states are distillable [15].

To obtain distillation schemes as needed in our proofs we will apply suitable twirling operations to the Choi-matrix $C_\mathcal{P}$ of a positive map $\mathcal{P} : \mathcal{M}_d \to \mathcal{M}_d$. The matrix $C_\mathcal{P}$ is in general not positive, but the next lemma proves that under certain conditions the UU-twirl leads to positive matrix, which can then be distilled using the existing theory.

Lemma 8 (Twirl of block-positive matrices)**.** *If* $C \in \mathcal{M}_d \otimes \mathcal{M}_d$ *is block-positive and such that* $\mathrm{tr}\left(C\mathbb{F}_d\right) \leq 0$*, then*

$$\int_{U \in \mathcal{U}(d)} (U \otimes U)\, C\, (U \otimes U)^{\dagger}\, dU \geq 0. \tag{A2}$$

Similarly, if $C \in \mathcal{M}_{d^2}$ *is block-positive and such that* $\mathrm{tr}\left(C\omega_d\right) \geq 0$*, then*

$$\int_{U \in \mathcal{U}(d)} \left(U \otimes \overline{U}\right) C \left(U \otimes \overline{U}\right)^{\dagger} dU \geq 0. \tag{A3}$$

Proof. By block-positivity we have $\mathrm{tr}\left(C\right) \geq 0$ and

$$\mathrm{tr}\left(C\left(\mathbb{1}_d \otimes \mathbb{1}_d + \mathbb{F}_d\right)\right) \geq 0$$

because the Werner state $\rho_W^{(1)} = (\mathbb{1}_d \otimes \mathbb{1}_d + \mathbb{F}_d)/(d(d+1))$ is separable. Together with the assumption $\mathrm{tr}\left(C\mathbb{F}_d\right) \leq 0$ this implies $\mathrm{tr}\left(C\mathbb{F}_d\right) \in \left[-\mathrm{tr}\left(C\right), 0\right]$ which shows the first statement (A2).

Secondly, as the Werner state $\rho_W^{(0)}$ is also separable we get by block-positivity

$$\mathrm{tr}\left(C^{T_2}\left(\mathbb{1}_d \otimes \mathbb{1}_d - \frac{1}{d}\mathbb{F}_d\right)\right) \geq 0,$$

which implies $d\mathrm{tr}\left(C^{T_2}\right) \geq \mathrm{tr}\left(C^{T_2}\mathbb{F}_d\right)$. Together with the assumption $\mathrm{tr}\left(C^{T_2}\mathbb{F}_d\right) = d\mathrm{tr}\left(C\omega_d\right) \geq 0$ this implies that the UU-twirl of C^{T_2} has positive partial transpose. As $\left[(U \otimes U)\, X^{T_2}\, (U \otimes U)^{\dagger}\right]^{T_2} = \left(U \otimes \overline{U}\right) X \left(U \otimes \overline{U}\right)^{\dagger}$ this finishes the proof.

□

Appendix B: Minimal output eigenvalue

Here we prove the multiplicativity of the minimal output eigenvalue (3) for entanglement breaking completely positive maps.

Theorem 9. *Let* $\mathcal{T} : \mathcal{M}_{d_1} \to \mathcal{M}_{d_2}$ *be entanglement breaking, i.e.* $\left(\mathrm{id}_n \otimes \mathcal{T}\right)(\rho)$ *is separable for all* $n \in \mathbb{N}$ *and positive* $\rho \in \mathcal{M}_n \otimes \mathcal{M}_{d_1}$*, and* $\mathcal{S} : \mathcal{M}_{d_3} \to \mathcal{M}_{d_4}$ *be completely positive. Then we have*

$$\lambda_{out}^{min}\left[\mathcal{T} \otimes \mathcal{S}\right] = \lambda_{out}^{min}\left[\mathcal{T}\right] \lambda_{out}^{min}\left[\mathcal{S}\right].$$

Proof. By inserting product states it is clear that $\lambda_{\mathrm{out}}^{\mathrm{min}}\left[\mathcal{T} \otimes \mathcal{S}\right] \leq \lambda_{\mathrm{out}}^{\mathrm{min}}\left[\mathcal{T}\right] \lambda_{\mathrm{out}}^{\mathrm{min}}\left[\mathcal{S}\right]$.

For the other direction, let the minimum in (3) for the computation of $\lambda_{\mathrm{out}}^{\mathrm{min}}\left[\mathcal{T} \otimes \mathcal{S}\right]$ be attained at $\rho = \tau$. Then there exists a pure state $|\phi\rangle$ such that

$$\begin{aligned}\lambda_{\mathrm{out}}^{\mathrm{min}}\left[\mathcal{T} \otimes \mathcal{S}\right] &= \langle\phi|(\mathcal{T} \otimes \mathcal{S})(\tau)|\phi\rangle \\ &= \langle\phi| \sum_{i=1}^{k} \left[\sigma_i \otimes \mathcal{S}(\rho_i)\right] |\phi\rangle\end{aligned}$$

using that there exist non-zero $\sigma_i \geq 0$ and $\rho_i \geq 0$ such that $(\mathcal{T} \otimes \mathrm{id}_{d_3})(\tau) = \sum_{i=1}^k \sigma_i \otimes \rho_i$ as $\mathcal{T}$ is entanglement breaking. Note that $\mathcal{T}(\mathrm{tr}_2(\tau)) = \sum_{i=1}^k \mathrm{tr}(\rho_i)\sigma_i$. Thus:

$$\begin{aligned}
\lambda_{\mathrm{out}}^{\min}[\mathcal{T} \otimes \mathcal{S}] &= \langle\phi| \sum_{i=1}^k \left[\mathrm{tr}(\rho_i)\sigma_i \otimes \mathcal{S}(\frac{\rho_i}{\mathrm{tr}(\rho_i)})\right] |\phi\rangle \\
&\geq \lambda_{\mathrm{out}}^{\min}[\mathcal{S}] \langle\phi| \sum_{i=1}^k \mathrm{tr}(\rho_i)\sigma_i \otimes \mathbb{1}_{d_4} |\phi\rangle \\
&= \lambda_{\mathrm{out}}^{\min}[\mathcal{S}] \,\mathrm{tr}\left[\mathcal{T}(\mathrm{tr}_2(\tau))\,\mathrm{tr}_2(|\phi\rangle\langle\phi|)\right] \\
&\geq \lambda_{\mathrm{out}}^{\min}[\mathcal{S}]\, \lambda_{\min}\big(\mathcal{T}(\mathrm{tr}_2(\tau))\big) \\
&\geq \lambda_{\mathrm{out}}^{\min}[\mathcal{S}]\, \lambda_{\mathrm{out}}^{\min}[\mathcal{T}].
\end{aligned}$$

□

---

[1] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal. Unextendible product bases and bound entanglement. *Phys. Rev. Lett.*, 82:5385–5388, Jun 1999.

[2] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin. Capacities of quantum erasure channels. *Physical Review Letters*, 78(16):3217, 1997.

[3] M. Berta, M. Christandl, F. G. Brandao, and S. Wehner. Entanglement cost of quantum channels. In *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pages 900–904. IEEE, 2012.

[4] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer New York, 1997.

[5] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin. Optimal universal and state-dependent quantum cloning. *Physical Review A*, 57(4):2368, 1998.

[6] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter. Everything you always wanted to know about locc (but were afraid to ask). *Communications in Mathematical Physics*, 328(1):303–326, 2014.

[7] T. Cubitt, J. Chen, and A. Harrow. Superactivation of the asymptotic zero-error classical capacity of a quantum channel. *Information Theory, IEEE Transactions on*, 57(12):8114–8126, Dec 2011.

[8] D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and A. V. Thapliyal. Evidence for bound entangled states with negative partial transpose. *Phys. Rev. A*, 61:062312, May 2000.

[9] W. Dür, J. I. Cirac, M. Lewenstein, and D. Bruß. Distillability and partial transposition in bipartite systems. *Phys. Rev. A*, 61:062313, May 2000.

[10] S. N. Filippov, A. A. Melnikov, and M. Ziman. Dissociation and annihilation of multipartite entanglement structure in dissipative quantum dynamics. *Physical Review A*, 88(6):062328, 2013.

[11] S. N. Filippov, T. Rybár, and M. Ziman. Local two-qubit entanglement-annihilating channels. *Physical Review A*, 85(1):012303, 2012.

[12] S. N. Filippov and M. Ziman. Bipartite entanglement-annihilating maps: Necessary and sufficient conditions. *Physical Review A*, 88(3):032316, 2013.

[13] M. Hayashi. *Quantum Information – An Introduction*. Springer Berlin Heidelberg, 2006.

[14] A. S. Holevo and R. F. Werner. Evaluating capacities of bosonic gaussian channels. *Physical Review A*, 63(3):032312, 2001.

[15] M. Horodecki and P. Horodecki. Reduction criterion of separability and limits for a class of distillation protocols. *Phys. Rev. A*, 59:4206–4216, Jun 1999.

[16] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1):1–8, 1996.

[17] M. Horodecki, P. Horodecki, and R. Horodecki. Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature? *Phys. Rev. Lett.*, 80:5239–5242, Jun 1998.

[18] M. Horodecki, P. W. Shor, and M. B. Ruskai. Entanglement breaking channels. *Reviews in Mathematical Physics*, 15(06):629–641, 2003.

[19] D. Kretschmann and R. F. Werner. Tema con variazioni: quantum channel capacity. *New Journal of Physics*, 6(1):26, 2004.

[20] S. Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55:1613–1622, Mar 1997.

[21] L. Moravčíková and M. Ziman. Entanglement-annihilating and entanglement-breaking channels. *Journal of Physics A: Mathematical and Theoretical*, 43(27):275306, 2010.

[22] C. Morgan and A. Winter. "Pretty strong" converse for the quantum capacity of degradable channels. *Information Theory, IEEE Transactions on*, 60(1):317–333, Jan 2014.

[23] V. Paulsen. *Completely bounded maps and operator algebras*, volume 78. Cambridge University Press, 2002.

[24] E. M. Rains. A semidefinite program for distillable entanglement. *Information Theory, IEEE Transactions on*, 47(7):2921–2933, 2001.

[25] A. Roy and A. J. Scott. Unitary designs and codes. *Designs, codes and cryptography*, 53(1):13–31, 2009.

[26] G. Smith and J. A. Smolin. Detecting incapacity of a quantum channel. *Phys. Rev. Lett.*, 108:230507, Jun 2012.

[27] E. Størmer. Tensor powers of 2-positive maps. *J. of Math. Phys*, 51(10), 2010.

[28] M. Takeoka, S. Guha, and M. Wilde. The squashed entanglement of a quantum channel. *Information Theory, IEEE Transactions on*, 60(8):4987–4998, Aug 2014.

[29] M. Tomamichel, M. M. Wilde, and A. Winter. Strong converse rates for quantum communication. *arXiv:1406.2946*, 2014.

[30] R. F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277, 1989.

Entropy Production of Doubly Stochastic Quantum Channels

Alexander Müller-Hermes*, Daniel Stilck França†, and Michael M. Wolf‡

Department of Mathematics, Technische Universität München, 85748 Garching, Germany

May 19, 2015

We study the entropy increase of quantum systems evolving under primitive, doubly stochastic Markovian noise and thus converging to the maximally mixed state. This entropy increase can be quantified by a logarithmic-Sobolev constant of the Liouvillian generating the noise. We prove a universal lower bound on this constant that stays invariant under taking tensor-powers. Our methods involve a new comparison method to relate logarithmic-Sobolev constants of different Liouvillians and a technique to compute logarithmic-Sobolev inequalities of Liouvillians with eigenvectors forming a projective representation of a finite abelian group. Our bounds improve upon similar results established before and as an application we prove an upper bound on continuous-time quantum capacities. In the last part of this work we study entropy production estimates of discrete-time doubly-stochastic quantum channels by extending the framework of discrete-time logarithmic-Sobolev inequalities to the quantum case.

Contents

1. Introduction **2**

2. Notations and preliminaries **3**
2.1. The LS-1 Constant 4
2.2. The LS-2 Constant 5
2.3. The spectral gap and relations between the LS-constants 5

3. Continuous LS inequalities for doubly stochastic Liouvillians **6**
3.1. Tensor-stable LS-inequalities 6
3.2. Qubit Liouvillians of the form $\mathcal{L} = T - \mathrm{id}$ 10

4. Discrete LS inequalities for doubly stochastic channels **13**

5. Applications **17**
5.1. Upper bounds on unitary quantum subdivision capacities 17
5.2. Entropy production for random Pauli channels 18

6. Conclusion **19**

*muellerh@ma.tum.de
†dsfranca@mytum.de
‡wolf@ma.tum.de

1. Introduction

Consider a quantum system affected by Markovian noise driving every initial state to the maximally mixed state. For such a, so-called doubly stochastic and primitive, noise the von-Neumann entropy S will steadily increase in time. Here we want to quantify how much entropy is produced by such a noise channel. More specifically let the Markovian noise channel be modeled by a quantum dynamical semigroup $T_t = e^{t\mathcal{L}}$ and recall that the von-Neumann entropy of a state ρ is given by $S(\rho) = -\mathrm{tr}(\rho \log(\rho))$. We want to establish bounds of the form

$$S(T_t(\rho)) - S(\rho) \geq C_t(log(d) - S(\rho)) \tag{1}$$

for some time-dependent $C_t < 1$ independent of the state ρ. To find good C_t for a given noise channel T_t in the above bound we apply the framework of logarithmic Sobolev (LS) inequalities [OZ99, KT13]. For special channels bounds of the form of (1) have been considered in [ABIN96]. Similar bounds have also been studied in terms of contractive properties of the channel with respect to different norms in [Str85, Rag02, cf. Remark 4.1]. However, in these cases the lower bound is given in terms of the trace or Hilbert-Schmidt distance between the state ρ and the maximally mixed state, which are upper bounded by a constant independent of the dimension, while the left-hand side of (1) is of order $\log(d)$. The entropy production of quantum dynamical semigroups has also been investigated in a more general setting in [Spo78].

For many applications in quantum information theory it will be important to quantify the entropy production of tensor-powers $T_t^{\otimes n}$ of the noise channel. In our main result we obtain bounds of this form, where C_t does not depend on n. We improve on recent results by Temme et al. [TPK14] who established such bounds using spectral theory. The invariance under taking tensor-powers makes these bounds important for many applications including the study of stability in Markovian systems [CLMP13], mixing-time bounds [TPK14] and quantifying the storage time in quantum memories [MHRW15].

In the second part of this paper we consider entropy production estimates of the form (1) for discrete-time doubly stochastic quantum channels. We introduce the framework of discrete LS inequalities, which allows us to generalize results from classical Markov chains, where there already is a vast literature on the subject [DSC96, DSC93, BT06, Mic97].

This paper is organized as follows:

- In section 2 we introduce our notation, definitions and show how the framework of logarithmic Sobolev inequalities relates to the entropy production of a doubly stochastic Markovian time-evolution.

- In section 3.1 we prove our main result. This is an improved lower bound on the LS constant of tensor powers of doubly stochastic semigroups (Theorem 3.3), which directly implies an entropy production estimate (Corollary 3.3) for tensor-products of Markovian time-evolutions. Previous approaches to this problem focused on spectral and interpolation techniques [BZ00, TPK14]. Here we obtain better bounds with simpler proofs using group theoretic techniques similar to the ones developed in [JPPP15] and comparison inequalities.

- In section 3.2 we consider Liouvillians of the form $\mathcal{L} = T - \mathrm{id}$, where T is a quantum channel. We show how to use LS constants of classical Markov chains to analyze the entropy production of such semigroups. As an application of our techniques we compute the LS constant of all doubly stochastic qubit Liouvillians of this form.

- In section 4 we extend techniques from LS inequalities to analyze discrete-time quantum channels. Here we not only get bounds on the entropy production (Theorem 4.2) but also on the hypercontractivity (Theorem 4.5) of these channels. However, the obtained bounds are in general weaker and become trivial as we increase the number of copies of the channel. These results are mostly a generalization of [Mic97].

- In section 5 we apply the results from section 2 to unitary quantum subdivision capacities introduced by some of the authors [MHRW15]. We show (Theorem 5.1) that the unitary quantum subdivision capacity of any doubly stochastic and primitive Liouvillian has to decay exponentially in time. Our bound improves similar results found in [ABIN96, BOGH13]. In the second part of the section we compute entropy production estimates for random Pauli channels.

2. Notations and preliminaries

Throughout this paper $\mathcal{M}_d$ will denote the set of complex $d \times d$-matrices and $\mathcal{M}_d^+ \subset \mathcal{M}_d$ the cone of strictly positive matrices. The set of d-dimensional density matrices or states, i.e. positive matrices in $\mathcal{M}_d$ with trace 1, will be denoted by $\mathcal{D}_d$ and the set of strictly positive states will be denoted by $\mathcal{D}_d^+$. The $d \times d$ identity matrix will be denoted by $\mathbb{1}_d$.

We will call a completely positive, trace preserving linear map $T : \mathcal{M}_d \to \mathcal{M}_d$ a quantum channel and will denote its adjoint with respect to the Hilbert-Schmidt scalar product by T^*. A quantum channel T is said to be doubly stochastic if

$$T\left(\mathbb{1}_d\right) = T^*\left(\mathbb{1}_d\right) = \mathbb{1}_d.$$

A family of quantum channels $\{T_t\}_{t \in \mathbb{R}_+}$ parametrized by a nonnegative parameter will be called a quantum dynamical semigroup if $T_0 = \mathrm{id}_d$ (the identity map in d dimensions), $T_{s+t} = T_s T_t$ for any $s, t \in \mathbb{R}_+$ and T_t depends continuously on t. Physically a quantum dynamical semigroup describes a Markovian evolution in continuous time. It is well known [Lin76, GKS76] that any quantum dynamical semigroup is generated by a Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ of the form

$$\mathcal{L}(X) = \Phi(X) - \kappa X - X\kappa^\dagger,$$

for $\kappa \in \mathcal{M}_d$ and $\Phi : \mathcal{M}_d \to \mathcal{M}_d$ completely positive such that $\Phi^*(\mathbb{1}_d) = \kappa + \kappa^\dagger$.

A quantum dynamical semigroup generated by a Liouvillian $\mathcal{L}$ is said to have a fixed point $\sigma \in \mathcal{D}_d$ if $\mathcal{L}(\sigma) = 0$. Then all quantum channel in $\{e^{t\mathcal{L}}\}_{t \in \mathbb{R}_+}$ have σ as a fixed point. In the special case where the fixed point is $\sigma = \frac{\mathbb{1}_d}{d}$ we call the quantum dynamical semigroup and its Liouvillian doubly stochastic. If the generator is hermitian with respect to the Hilbert-Schmidt scalar product, i.e. $\mathcal{L} = \mathcal{L}^*$, we will call it reversible. We will be interested in the asymptotic behavior of quantum dynamical semigroups. A quantum dynamical semigroup $T_t : \mathcal{M}_d \to \mathcal{M}_d$ with a full-rank fixed point $\sigma \in \mathcal{D}_d$ is called primitive if $\lim_{t\to\infty} T_t(\rho) = \sigma$ for all states $\rho \in \mathcal{D}_d$. In the following we will be interested in particular in tensor-products of quantum dynamical semigroups. Given a Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ we denote by

$$\mathcal{L}^{(n)} := \sum_{i=1}^{n} \mathrm{id}_d^{\otimes i-1} \otimes \mathcal{L} \otimes \mathrm{id}_d^{\otimes(n-i)} \tag{2}$$

the generator of the quantum dynamical semigroup $(e^{t\mathcal{L}})^{\otimes n}$.

We will need distance measures on the set $\mathcal{M}_d$. Recall the family of Schatten p-norms for $p \in [1, \infty)$ defined as

$$\|X\|_p := \left(\sum_{i=1}^{d} s_i(X)^p\right)^{1/p}$$

where $s_i(X)$ denotes the i-th singular value of $X \in \mathcal{M}_d$ and $s(X) \in \mathbb{R}^d$ is the vector containing the ordered singular values of X as entries. Note that we can consistently define $\|X\|_\infty := \sup_{i\in\{1,\dots,d\}} s_i(X)$ for any $X \in \mathcal{M}_d$.

Another distance measure (although not a metric) on the set $\mathcal{D}_d$ of states is the relative entropy (also known as Kullback-Leibler divergence):

$$D(\rho\|\sigma) = \begin{cases} \mathrm{tr}[\rho(\log\rho - \log\sigma)], & \text{if } \mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma) \\ +\infty, & \text{otherwise} \end{cases}$$

for $\rho, \sigma \in \mathcal{D}_d$. Recall Pinsker's inequality:

$$D(\rho\|\sigma) \geq \frac{1}{2}\|\rho - \sigma\|_1^2$$

for any $\rho, \sigma \in \mathcal{D}_d$. This inequality implies in particular that $D(\rho\|\sigma) = 0$ iff $\rho = \sigma$.

2.1. The LS-1 Constant

Consider an inequality of the form:

$$D\left(T_t(\rho)\middle\|\frac{\mathbb{1}_d}{d}\right) \leq e^{-2\alpha t} D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) \tag{3}$$

for some doubly stochastic Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$, and where $T_t = e^{t\mathcal{L}}$, $\rho \in \mathcal{D}_d$ and $\alpha \in \mathbb{R}_+$ is a constant independent of ρ. Using $D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) = \log(d) - S(\rho)$ the above inequality is clearly equivalent to (1) for $C_t = (1 - e^{-2\alpha t})$. The framework of logarithmic Sobolev-1-inequalities (LS-1 inequality) allows us to determine the optimal α such that (3) holds. To this end define the function $f(t) = D\left(T_t(\rho)\middle\|\frac{\mathbb{1}_d}{d}\right)$. If we can show

$$\frac{df}{dt} \leq -2\alpha f \tag{4}$$

for some $\alpha \in \mathbb{R}_+$ it follows that $f(t) \leq e^{-2\alpha t} f(0)$. The time derivative of the relative entropy at $t = 0$, also called the entropy production[Spo78], is given by:

$$\begin{aligned}\frac{d}{dt} D\left(T_t(\rho)\middle\|\frac{\mathbb{1}_d}{d}\right)\Bigg|_{t=0} &= \mathrm{tr}[\mathcal{L}(\rho)(\log(d) + \log(\rho))] \\ &= \mathrm{tr}[\mathcal{L}(\rho)\log(\rho)]\end{aligned}$$

as $\mathrm{tr}(\mathcal{L}(\rho)) = 0$ for any $\rho \in \mathcal{D}_d$. This motivates the definition of the LS-1 constant:

Definition 2.1 (LS-1 constant)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic Liouvillian. We define its **LS-1 constant** as*

$$\alpha_1(\mathcal{L}) := \inf_{\rho \in \mathcal{D}_d^+} -\frac{1}{2}\frac{tr[\mathcal{L}(\rho)\log(\rho)]}{D(\rho\|\frac{\mathbb{1}_d}{d})} \tag{5}$$

By the above discussion (3) is valid for $\alpha = \alpha_1(\mathcal{L})$ as it is true for small times and the time can be extended by iterating the bound. Also by definition it is the optimal constant such that (3) holds independent of ρ. Note that we may only consider states with full rank in the optimization of Definition 2.1 due to the continuity of the relative entropy and entropy production. Also note that if the Liouvillian $\mathcal{L}$ has another fixed point different from $\frac{\mathbb{1}_d}{d}$, then $\alpha_1(\mathcal{L}) = 0$ and (3) reduces to the data processing inequality. In the following we will always consider primitive Liouvillians and thereby avoid this issue.

Using $D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) = \log(d) - S(\rho)$ the following theorem follows from (3):

Corollary 2.1 (Entropy increase)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic Liouvillian and $\alpha \leq \alpha_1(\mathcal{L})$ then*

$$S(T_t(\rho)) - S(\rho) \geq (1 - e^{-2\alpha t})(log(d) - S(\rho))$$

for any $\rho \in \mathcal{D}_d$ and where $T_t = e^{\mathcal{L}t}$ denotes the semigroup generated by $\mathcal{L}$.

2.2. The LS-2 Constant

The non-linearity of the optimization problem in (5) makes it hard to compute the LS-1 constant analytically. In fact there are only few examples even for classical Markov chains [BT06] where the LS-1 constant is known. For many applications, however, it will be enough to have good lower bounds on α_1. Here these lower bounds will be in terms of the so-called LS-2 constant:

Definition 2.2 (LS-2 constant). *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic Liouvillian. We define its **LS-2 constant** as*

$$\alpha_2(\mathcal{L}) := \inf_{X \in \mathcal{M}_d^+} \frac{\mathcal{E}_{\mathcal{L}}^2(X)}{Ent_2(X)}.$$

*Here we used the so-called **2-Dirichlet form***

$$\mathcal{E}_{\mathcal{L}}^2(X) := -\frac{1}{d} tr[\mathcal{L}(X)X]$$

*and the so-called **2-relative entropy** given by*

$$Ent_2(X) := \frac{1}{2d} tr\left[X^2 \left(\log\left(\frac{X^2}{tr(X^2)}\right) + \log(d)\right)\right]$$

It is well known that a Liouvillian is primitive iff we have a unique strictly positive density matrix in the kernel of $\mathcal{L}$. From this it is easy to see that $\alpha_2(\mathcal{L}) = 0$ if the Liouvillian is not primitive, as in this case there exists $X \in \mathcal{M}_d^+$ s.t. $X \notin \text{span}\{\mathbb{1}\}$ and $\mathcal{E}_2^{\mathcal{L}}(X) = 0$. We will later see that the LS-2 constant is strictly positive if the Liouvillian is primitive.

As the 2-Dirichlet form is bilinear, α_2 is easier to compute than α_1, where a logarithm occurs in the numerator (see (5)). Also by $\alpha_2(\mathcal{L}) = \alpha_2(\frac{\mathcal{L}+\mathcal{L}^*}{2})$ we may always suppose that the Liouvillian is reversible when computing α_2. Another advantage of the LS-2 constant in comparison to the LS-1 constant is the following hypercontractive characterization. This characterization allows the use of tools from other areas of mathematics, such as interpolation theory, to compute the LS-2 constant.

Theorem 2.1 (Hypercontractive picture [OZ99]). *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive doubly stochastic Liouvillian and $T_t = e^{t\mathcal{L}}$ its associated semigroup. Then:*

1. *If there exists $\alpha > 0$ such that $\|T_t(X)\|_{p(t),\frac{\mathbb{1}_d}{d}} \leq \|X\|_{2,\frac{\mathbb{1}_d}{d}}$ for all $X \in \mathcal{M}_d^+$, where $p(t) = 1 + e^{2\alpha t}$, it follows that $\alpha_2(\mathcal{L}) \geq \alpha$.*
2. *For $\alpha_2(\mathcal{L}) > 0$, we have $\|T_t(X)\|_{p(t),\frac{\mathbb{1}_d}{d}} \leq \|X\|_{2,\frac{\mathbb{1}_d}{d}}$ for all $X \in \mathcal{M}_d^+$, with $p(t) = 1 + e^{2\alpha_2 t}$ for $\mathcal{L}$ reversible and $p(t) = 1 + e^{\alpha_2 t}$ if not.*

*Here we used the $\frac{\mathbb{1}_d}{d}$-**weighted** l_p-**norm** on $\mathcal{M}_d$ given by:*

$$\|Y\|_{p,\frac{\mathbb{1}_d}{d}} := \frac{1}{d^{\frac{1}{p}}} (tr[|Y|^p])^{\frac{1}{p}}$$

We will state the connection between the LS-2 and the LS-1 constants in Theorem 2.2 below.

2.3. The spectral gap and relations between the LS-constants

Another important constant for studying the convergence properties of quantum dynamical semigroups is the spectral gap. Usually a unital Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ is said to have a spectral gap λ' if the 0 eigenvalue corresponding to $\frac{\mathbb{1}_d}{d}$ is the only eigenvalue with real part 0 whereas $|\text{Re}\lambda_i| \geq \lambda'$ for all other eigenvalues of $\mathcal{L}$. In the context of LS-inequalities the following definition is used:

Definition 2.3 (Spectral Gap)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic Liouvillian. The **spectral gap** is defined as:*

$$\lambda(\mathcal{L}) := \inf_{X \in \mathcal{M}_d : X = X^\dagger} \frac{\mathcal{E}^2_{\mathcal{L}}(X)}{Var(X)}.$$

*Here we used the **variance** with respect to the maximally mixed state defined as*

$$Var_{\frac{\mathbb{1}_d}{d}}(Y) := \|Y - tr(Y)\frac{\mathbb{1}_d}{d}\|^2_{2,\frac{\mathbb{1}_d}{d}} .$$

It agrees with the usual definition for reversible Liouvillians and is the spectral gap of the additive symmetrization $\frac{\mathcal{L}+\mathcal{L}^*}{2}$. Indeed, for reversible Liouvillians the spectrum is nonpositive and we may assume w.l.o.g. that the eigenvector X_i corresponding to an eigenvalue λ_i is hermitian. By the orthogonality of eigenvectors, we have that $\mathrm{tr}[\mathbb{1}_d X_i] = 0$, thus all eigenvectors that do not correspond to the eigenvalue 0 are also traceless and are invariant under the transformation $X \mapsto X - tr[X]\frac{\mathbb{1}}{d}$. By these considerations,

$$\inf_{X \in \mathcal{M}_d : X = X^\dagger} \frac{\mathcal{E}^2_{\mathcal{L}}(X)}{\mathrm{Var}(X)} \tag{6}$$

is the second largest eigenvalue of $-\mathcal{L}$ if $\mathcal{L}$ is reversible or the of $-\frac{\mathcal{L}+\mathcal{L}^*}{2}$ if not, coinciding with the usual definition.

Finally we can state how the LS-constants and the spectral gap relate to each other.

Theorem 2.2 ([KT13])**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic Liouvillian. If $\mathcal{L}$ is reversible, the spectral gap, LS-2 and LS-1 constant satisfy:*

$$\lambda(\mathcal{L})\frac{2(1-\frac{2}{d})}{\log(d-1)} \leq \alpha_2(\mathcal{L}) \leq \alpha_1(\mathcal{L}) \leq \lambda(\mathcal{L})$$

If $\mathcal{L}$ is not reversible they satisfy:

$$\lambda(\mathcal{L})\frac{(1-\frac{2}{d})}{\log(d-1)} \leq \frac{\alpha_2(\mathcal{L})}{2} \leq \alpha_1(\mathcal{L}) \leq \lambda(\mathcal{L})$$

For $d = 2$ the function $d \mapsto \frac{2(1-\frac{2}{d})}{\log(d-1)}$ can be extended continuously by 1.

The previous theorem establishes a connection between the LS-1 and LS-2 constants. We will use this connection as it is usually easier to derive bounds on the LS-2 constant than on the LS-1 constant directly. Note that by combining Theorem 2.2 with Theorem 2.1 we immediately obtain the bound

$$S(T_t(\rho)) - S(\rho) \geq (1 - e^{-\alpha_2(\mathcal{L})t})(\log(d) - S(\rho)) \tag{7}$$

for any $\rho \in \mathcal{D}_d$ and where $T_t = e^{\mathcal{L}t}$ denotes the semigroup generated by $\mathcal{L}$.

3. Continuous LS inequalities for doubly stochastic Liouvillians

3.1. Tensor-stable LS-inequalities

For doubly stochastic and primitive Liouvillians $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ we consider the generator $\mathcal{L}^{(n)}$ (see (2)) of the tensor-product semigroup $\left(e^{t\mathcal{L}}\right)^{\otimes n}$. We will prove a lower bound on $\alpha_2\left(\mathcal{L}^{(n)}\right)$ that does not depend on n. By (7) such bounds directly lead to entropy production inequalities for tensor-products of quantum dynamical semigroups (see Corollary 3.3). These inequalities turn out to be useful for the analysis of quantum memories (see section 5.1).

For our bounds on the LS-2 constant we will first compute a lower bound on $\alpha_2\left(\mathcal{L}^{(n)}_{\text{dep}}\right)$, where $\mathcal{L}_{\text{dep}}:\mathcal{M}_d\to\mathcal{M}_d$ denotes the depolarizing Liouvillian given by

$$\mathcal{L}_{\text{dep}}(X)=\text{tr}(X)\frac{\mathbb{1}_d}{d}-X\ . \tag{8}$$

Then we use a comparison technique to derive the desired bound for general doubly stochastic and primitive Liouvillians.

We will need the following theorem proved in [TPK14] showing that it suffices to show hypercontractivity for one fixed time to lower bound $\alpha_2(\mathcal{L})$:

Theorem 3.1. *[TPK14, Theorem 5] Let $\mathcal{L}:\mathcal{M}_d\to\mathcal{M}_d$ be a reversible, doubly stochastic Liouvillian with spectral gap λ. Suppose that for some $t_0\in\mathbb{R}_+$ we have $\|T_{t_0}\|_{2\to4,\frac{\mathbb{1}_d}{d}}\leq 1$. Then:*

$$\alpha_2(\mathcal{L})\geq\frac{\lambda}{4\lambda t_0+2}$$

Using the group theoretic techniques and definitions introduced in the appendix we prove the following bound on the $2\to4$ norm of tensor powers of the depolarizing channel. Similar bounds have been developed in [JPPP15].

Theorem 3.2. *Let $T_t:\mathcal{M}_d\to\mathcal{M}_d$ denote the semigroup $T_t=e^{t\mathcal{L}_{dep}}$ generated by $\mathcal{L}_{dep}:\mathcal{M}_d\to\mathcal{M}_d$ as defined in* (8). *Then we have*

$$\|T_{t_0}^{\otimes n}\|_{2\to4,\frac{1}{d^n}}\leq 1 \tag{9}$$

for $t_0=\frac{\log(3)\log(d^2-1)}{4(1-2d^{-2})}$.

Proof. Note that the Weyl system (52) forms an almost commuting unitary eigenbasis (see Definition A.1) for the depolarizing Liouvillian $\mathcal{L}_{\text{dep}}$. The unitaries of the Weyl system can be associated to characters on $\mathbb{Z}_d\times\mathbb{Z}_d$. As explained in the appendix we can associate a classical semigroup P_t (see (54)) acting on the space $V(\mathbb{Z}_d\times\mathbb{Z}_d)$ of complex functions on $\mathbb{Z}_d\times\mathbb{Z}_d$. It is easy to verify that the generator L of this classical semigroup coincides with the generator of the random walk on the complete graph with d^2 vertices and uniform distribution. In [DSC96, Theorem A.1] it was shown that:

$$\alpha_2(L)=\frac{2\left(1-2d^{-2}\right)}{\log(d^2-1)} \tag{10}$$

Also it is known that for classical semigroups [DSC96, Lemma 3.2]

$$\alpha_2(L_1\otimes\text{id}+\text{id}\otimes L_2)=\min\{\alpha_2(L_1),\alpha_2(L_2)\}. \tag{11}$$

Thus, by the hypercontractive characterization of the LS-2 constant[DSC96, Theorem 3.5] we have

$$\|P_t^{\otimes n}\|_{2\to p(t)}\leq 1,$$

for any $n\in\mathbb{N}$ where $p(t)=1+e^{2\alpha_2(L)t}$. With $t_0=\frac{\log(3)}{2\alpha_2(L)}$ we have $p(t_0)=4$ and the the claim follows if we apply Theorem A.1 inductively. $\square$

As the spectral gap of $\mathcal{L}_{\text{dep}}$ is 1 we obtain the following corollary by applying Theorem 3.2 and Theorem 3.1.

Corollary 3.1 (Lower bound on LS-2 for tensor powers of the depolarizing channel)**.** *Let $\mathcal{L}_{dep}:\mathcal{M}_d\to\mathcal{M}_d$ be the depolarizing Liouvillian* (8)*. Then:*

$$\alpha_2(\mathcal{L}^{(n)}_{dep})\geq\frac{\left(1-2d^{-2}\right)}{\log(3)\log(d^2-1)+2\left(1-2d^{-2}\right)}.$$

Note that any doubly stochastic Liouvillian $\mathcal{L}$ commutes with the depolarizing Liouvillian $\mathcal{L}_{\text{dep}}$. We can use this simple observation to prove the following comparison theorem, which will lead to our main result.

Theorem 3.3 (Comparison with the depolarizing Liouvillian)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic and primitive Liouvillian with spectral gap λ and $\|\frac{\mathcal{L}+\mathcal{L}^*}{2}\|$ the operator norm of its additive symmetrization. For any $n \in \mathbb{N}$ we have*

$$\left\|\frac{\mathcal{L}+\mathcal{L}^*}{2}\right\|\alpha_2(\mathcal{L}_{dep}^{(n)}) \geq \alpha_2(\mathcal{L}^{(n)}) \geq \lambda\alpha_2(\mathcal{L}_{dep}^{(n)})$$

where $\mathcal{L}^{(n)}$ is defined as in (2)*.*

Proof. When working with the LS-2 constant (see Definition 2.2 and the discussion following this definition) we may consider the additive symmetrization $\frac{\mathcal{L}+\mathcal{L}*}{2}$ instead of $\mathcal{L}$. Therefore, we may assume that $\mathcal{L}$ is reversible without loss of generality.

Using $\text{tr}(\mathcal{L}(Y)) = 0$ for all $Y \in \mathcal{M}_d$ and $\mathcal{L}(\mathbb{1}_d) = 0$ it is easily seen that $[\mathcal{L}_{\text{dep}}, \mathcal{L}] = 0$. This shows that $\mathcal{L}$ and $\mathcal{L}_{\text{dep}}$ can be simultaneously diagonalized. The same also holds for $\mathcal{L}^{(n)}$ and $\mathcal{L}_{\text{dep}}^{(n)}$.

Let $\{Y_i\}_{0\leq i\leq d^2-1}$ be an orthonormal basis for $\mathcal{M}_d$ with respect to the normalized Hilbert-Schmidt scalar product $\langle\cdot|\cdot\rangle_{\frac{\mathbb{1}_d}{d}} = \frac{1}{d}\langle\cdot|\cdot\rangle_{\text{HS}}$ consisting of eigenvectors of $\mathcal{L}$ and $\mathcal{L}_{\text{dep}}$ with $Y_0 = \mathbb{1}_d$. Let λ_i denote the corresponding eigenvalues of $\mathcal{L}$, i.e. such that $\mathcal{L}(Y_i) = \lambda_i Y_i$. We will show that

$$\frac{1}{\|\mathcal{L}\|}\mathcal{E}^2_{\mathcal{L}^{(n)}}(X) \leq \mathcal{E}^2_{\mathcal{L}^{(n)}_{\text{dep}}}(X) \leq \frac{1}{\lambda}\mathcal{E}^2_{\mathcal{L}^{(n)}}(X)$$

for $X \in \mathcal{M}^+_{d^n}$.

Given a multi-index $\nu \in [d^2-1]^n$, where $[d^2-1] = \{0, ..., d^2-1\}$, we define

$$Y_\nu := \bigotimes_{i=1}^n Y_{\nu(j)}.$$

Clearly $\{Y_\nu\}_{\nu\in[d^2-1]^n}$ forms an orthonormal basis of eigenvectors of $\mathcal{L}^{(n)}$ and $\mathcal{L}^{(n)}_{\text{dep}}$. For a multi-index $\nu \in [d^2-1]^n$ we define $\text{supp}\,(\nu) := \{i \in \{1, 2, \ldots, n\} : \nu(i) \neq 0\}$. As $\mathcal{L}(Y_0) = \mathcal{L}_{\text{dep}}(Y_0) = 0$ we have

$$\mathcal{L}^{(n)}(Y_\nu) = \left(\sum_{i\in\text{supp}(\nu)} \lambda_{\nu(i)}\right) Y_\nu$$

and

$$\mathcal{L}^{(n)}_{\text{dep}}(Y_\nu) = -|\text{supp}\,(\nu)\,|Y_\nu$$

for any multi-index $\nu \in [d^2-1]^n$. With

$$\lambda_\nu := \sum_{i\in\text{supp}(\nu)} \lambda_{\nu(i)}$$

we can express the 2-Dirichlet forms (see Definition 2.2) as

$$\mathcal{E}^2_{\mathcal{L}^{(n)}}(X) = -\sum_{\nu\in[d^2-1]^n} \lambda_\nu \left|\langle X|Y_\nu\rangle_{\frac{\mathbb{1}_{d^n}}{d^n}}\right|^2 \tag{12}$$

and

$$\mathcal{E}^2_{\mathcal{L}^{(n)}_{\text{dep}}}(X) = \sum_{\nu\in[d^2-1]^n} |\text{supp}\,(\nu)\,| \left|\langle X|Y_\nu\rangle_{\frac{\mathbb{1}_{d^n}}{d^n}}\right|^2. \tag{13}$$

We know that the spectral gap is given by $\lambda = \min_{i\neq 0}\{-\lambda_i\}$ and $\|\mathcal{L}\| = \max_i\{-\lambda_i\}$. As a consequence we have

$$-\lambda_\nu \geq |\mathrm{supp}\,(\nu)\,|\lambda, \tag{14}$$

$$-\lambda_\nu \leq |\mathrm{supp}\,(\nu)\,|\|\mathcal{L}\|. \tag{15}$$

Combining (12) and (14) leads to

$$\mathcal{E}^2_{\mathcal{L}^{(n)}}(X) \geq \lambda \mathcal{E}^2_{\mathcal{L}^{(n)}_{\mathrm{dep}}}.$$

By Definition 2.2 of the LS-2 constant this shows $\alpha_2(\mathcal{L}^{(n)}) \geq \lambda\alpha_2(\mathcal{L}^{(n)}_{\mathrm{dep}})$.

In the same way combining (15) and (12) implies $\alpha_2(\mathcal{L}^{(n)}) \leq \|\mathcal{L}\|\alpha_2(\mathcal{L}^{(n)}_{\mathrm{dep}})$.

□

The previous result tells us that, considering Liouvillians with fixed operator norms, the depolarizing channel is the most hypercontractive one, as it has the largest LS-2 constant in this class.

One can also introduce LS constants for Liouvillians $\mathcal{L}$ having a stationary state $\sigma \in \mathcal{D}^+_d$ that is not necessarily maximally mixed[KT13, OZ99]. When $\mathcal{L}$ is primitive and reversible the same proof as for Theorem 3.3 yields

$$\|\mathcal{L}\|\alpha_2(\mathcal{L}^{(n)}_{\mathrm{dep},\sigma}) \geq \alpha_2(\mathcal{L}^{(n)}) \geq \lambda\alpha_2(\mathcal{L}^{(n)}_{\mathrm{dep},\sigma})$$

for the generalized depolarizing Liouvillian $\mathcal{L}_{\mathrm{dep},\sigma}(X) := \mathrm{tr}(X)\sigma - X$.

By combining Theorem 3.3 with Corollary 3.1 we can finally establish an explicit lower bound on $\alpha_2(\mathcal{L}^{(n)})$. For an explicit upper bound we can also apply Corollary 3.1 to

$$\alpha_2(\mathcal{L}^{(n)}_{\mathrm{dep}}) \leq \alpha_2(\mathcal{L}_{\mathrm{dep}}) = \frac{2(1-\frac{2}{d})}{\log(d-1)}$$

where we used that the LS-2 constant can only decrease when taking tensor-powers (see Definition 2.2). We also used the explicit formula for the LS-2 constant of a single depolarizing Liouvillian $\mathcal{L}_{\mathrm{dep}} : \mathcal{M}_d \to \mathcal{M}_d$ calculated in [KT13]. Summarizing these observations we obtain the following corollary:

Corollary 3.2. *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic and primitive Liouvillian with spectral gap λ and $\|\frac{\mathcal{L}+\mathcal{L}^*}{2}\|$ the operator norm of its additive symmetrization. Then:*

$$\left\|\frac{\mathcal{L}+\mathcal{L}^*}{2}\right\|\frac{2(1-\frac{2}{d})}{\log(d-1)} \geq \alpha_2(\mathcal{L}^{(n)}) \geq \frac{\lambda\left(1-2d^{-2}\right)}{\log(3)\log(d^2-1)+2\left(1-2d^{-2}\right)}.$$

Note that the lower bound in Corollary 3.2 is slightly better than the one that follows from the results in [BZ00, TPK14] given by

$$\alpha_2(\mathcal{L}^{(n)}) \geq \frac{\lambda}{5\log(d)+11}\ .$$

Using (7) we obtain the following corollary on the entropy production of a tensor-product semigroup:

Corollary 3.3 (Entropy production of tensor-product semigroups)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic and primitive Liouvillian with spectral gap λ. Then we have*

$$S(T_t^{\otimes n}(\rho)) - S(\rho) \geq (1-e^{-2t\alpha(d)})\left(\log(d^n) - S(\rho)\right)$$

with $\alpha(d) = \frac{\lambda\left(1-2d^{-2}\right)}{\log(3)\log(d^2-1)+2(1-2d^{-2})}$ for the quantum dynamical semigroup $T_t = e^{t\mathcal{L}}$ generated by $\mathcal{L}$.

In [Kin14][Theorem 3] the multiplicativity of the $2 \to q$ norm for $q \geq 2$ has been proven for doubly stochastic and reversible quantum channels $T : \mathcal{M}_2 \to \mathcal{M}_2$. That is, for a completely positive map $\Omega : \mathcal{M}_{d'} \to \mathcal{M}_{d'}$ we have

$$\|\Omega \otimes T\|_{2\to q} = \|\Omega\|_{2\to q}\|T\|_{2\to q} \ .$$

It then follows from the hypercontractive characterization (Theorem 2.1) of the LS-2 that $\alpha_2(\mathcal{L}^{(n)}) = \alpha_2(\mathcal{L})$ for any doubly stochastic, primitive and reversible qubit Liouvillian $\mathcal{L} : \mathcal{M}_2 \to \mathcal{M}_2$, but for non-reversible Liouvillians it only follows that $\alpha_2\left(\mathcal{L}^{(n)}\right) \geq \frac{\lambda}{2}$.

Here we give a proof of this lower-bound without needing reversibility:

Corollary 3.4. *Let $\mathcal{L} : \mathcal{M}_2 \to \mathcal{M}_2$ be a doubly stochastic and primitive Liouvillian with spectral gap λ. Then:*

$$\alpha_2(\mathcal{L}^{(n)}) = \alpha_2(\mathcal{L}) = \lambda$$

Proof. In [KT13, Lemma 25], it was proven that $\alpha_2(\mathcal{L}^{(n)}_{\text{dep}}) = \alpha_2(\mathcal{L}_{\text{dep}}) = 1$. Theorem 3.3 then implies $\alpha_2(\mathcal{L}^{(n)}) \geq \lambda$. But, by Theorem 2.2, $\alpha_2(\mathcal{L}^{(n)}) \leq \lambda$, which proves the claim. □

Note that if we are interested in the hypercontractivity of $\mathcal{L}^{(n)} : \mathcal{M}_{2^n} \to \mathcal{M}_{2^n}$, this result implies

$$\|e^{t\mathcal{L}^{(n)}}\|_{2\to p(t),\frac{\mathbb{1}_{2^n}}{2^n}} \leq 1$$

with $p(t) = 1 + e^{2\lambda t}$ if $\mathcal{L}$ is reversible and with $p(t) = 1 + e^{\lambda t}$ if $\mathcal{L}$ is not reversible.

3.2. Qubit Liouvillians of the form $\mathcal{L} = T - \mathrm{id}$

In this section we consider quantum dynamical semigroups on a qubit system generated by Liouvillians of the form $\mathcal{L} = T - \mathrm{id}_2$ for a doubly stochastic quantum channel $T : \mathcal{M}_2 \to \mathcal{M}_2$. For these Liouvillians we will compute the LS-1 and LS-2 constants using the corresponding constants for classical Markov chains [DSC96]. By the general theory of LS-inequalities (see Section 2) this leads to entropy production estimates for this kind of Liouvillians.

We start with the LS-2 constant: to establish the connection with classical Markov chains note that we can split the infimum in Definition 2.2 and optimize over spectrum and basis separately. Let $\mathcal{U}_d$ denote the set of unitary $d \times d$-matrices and denote the diagonal matrix with entries $s \in \mathbb{R}^d$ by $\mathrm{diag}(s)$. For a doubly stochastic quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ and the Liouvillian $\mathcal{L} = T - \mathrm{id}_d$ the definitions of $\mathcal{E}^2_{\mathcal{L}}$ and Ent_2 lead to:

$$\alpha_2(\mathcal{L}) = \inf_{U\in\mathcal{U}_d} \inf_{s\in\mathbb{R}^d_+} \frac{\mathcal{E}^2_{\mathcal{L}}(U\mathrm{diag}(s)U^\dagger)}{\mathrm{Ent}_2(U\mathrm{diag}(s)U^\dagger)} \tag{16}$$

$$= \inf_{U\in\mathcal{U}_d} \inf_{s\in\mathbb{R}^d_+} \frac{-2\langle s| (M_U - \mathbb{1}_d) |s\rangle}{\sum_i s_i^2(\log(\frac{s_i}{\|s\|_2}) + \log(d))}. \tag{17}$$

Here $M_U \in \mathcal{M}_d$ is a doubly stochastic matrix depending on $U \in \mathcal{U}_d$ defined as

$$(M_U)_{ij} = \langle j| U^\dagger T(U|i\rangle\langle i|U^\dagger)U |j\rangle \tag{18}$$

for some fixed orthonormal basis $\{|i\rangle\} \subset \mathbb{C}^d$. Note that M_U is doubly stochastic for any $U \in \mathcal{U}_d$ and thus the matrix $M_U - \mathbb{1}_d$ defines a classical Markov kernel on a d-point set. Finally let $\alpha_2^{(c)}(M_U - \mathbb{1}_d)$ denote the classical LS-2 constant [DSC96, equation (3.1)] of the Markov kernel $M_U - \mathbb{1}_d$. By direct comparison we have

$$\alpha_2(\mathcal{L}) = \inf_{U\in\mathcal{U}_d} \alpha_2^{(c)}(M_U - \mathbb{1}_d). \tag{19}$$

The LS-1 constant of the Liouvillian $\mathcal{L} = T - \mathrm{id}_d$ can be treated in the same way as the LS-2 constant above. A similar reasoning then leads to

$$\alpha_1(\mathcal{L}) = \inf_{U \in \mathcal{U}_d} \alpha_1^{(c)}(M_U - \mathbb{1}_d) \tag{20}$$

where $\alpha_1^{(c)}(M_U - \mathbb{1}_d)$ denotes the classical LS-1 constant [BT06, Equation 1.5] of the Markov kernel $M_U - \mathbb{1}_d$, see (18).

The above technique works for every dimension $d \geq 2$. We will now restrict to $d = 2$, i.e. $T : \mathcal{M}_2 \to \mathcal{M}_2$ is a doubly stochastic qubit quantum channel. Such a quantum channel can be represented as an affine transformation on $\mathbb{R}^3$, the so-called Bloch sphere representation [NC00]. In this representation quantum states $\rho \in \mathcal{D}_2$ are identified with vectors $x \in \mathbb{R}^3$ by

$$\rho = \frac{\mathbb{1}_2 + \sum_{i=1}^3 x_i \sigma_i}{2}$$

where we used the Pauli-matrices, i.e.

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

In this representation a doubly stochastic quantum channel $T : \mathcal{M}_2 \to \mathcal{M}_2$ corresponds to a matrix $\hat{T} \in \mathcal{M}_3$ acting on the corresponding vectors in $\mathbb{R}^3$. Note that the adjoint channel T^* with respect to the Hilbert-Schmidt scalar product corresponds to the transposed matrix $\widehat{(T^*)} = (\hat{T})^T$.

We now show how to express the LS constants of a doubly stochastic qubit Liouvillian of the form $\mathcal{L} = T - \mathrm{id}_2$ in terms of the matrix $\hat{T}$ representing the action of the quantum channel T.

Theorem 3.4. *Let $\mathcal{L} : \mathcal{M}_2 \to \mathcal{M}_2$ be a doubly stochastic and primitive Liouvillian of the form $\mathcal{L} = T - id_2$, where $T : \mathcal{M}_2 \to \mathcal{M}_2$ is a doubly stochastic quantum channel. Then we have*

$$\alpha_2(\mathcal{L}) = 1 - \left\| \frac{\hat{T} + \hat{T}^T}{2} \right\|$$

where $\hat{T} \in \mathcal{M}_3$ is the matrix representing T on the Bloch sphere.

Proof. Note that we may replace $\mathcal{L}$ by the symmetrized Liouvillian $\frac{\mathcal{L}+\mathcal{L}^*}{2}$ as we are computing the LS-2 constant. Thus, we may consider $\frac{T+T^*}{2}$, which is represented by a symmetric matrix, instead of T.

Diaconis and Saloff-Coste showed that the LS-2 constant of a Markov kernel $M - \mathbb{1}_d$ on a two-dimensional state space is given by $2M_{1,2}$ [DSC96, Corollary A.3][1]. For $U \in \mathcal{U}_d$ consider M_U, see (18). Using (19) and the aforementioned result we get:

$$\alpha_2(\mathcal{L}) = \inf_{U \in \mathcal{U}_d} \alpha_2^{(c)}(M_U - \mathbb{1}_d) = 2 \inf_{U \in \mathcal{U}_d} \langle 1| U^\dagger \left(\frac{T+T^*}{2} \right) \left(U|0\rangle\langle 0|U^\dagger \right) U |1\rangle \tag{21}$$

Let x be the vector corresponding to $U|0\rangle$ on the Bloch sphere. Then $-x$ is the vector corresponding to $U|1\rangle$. By changing to the Bloch sphere representation we get:

$$\langle 1| U^\dagger \left(\frac{T+T^*}{2} \right) \left(U|0\rangle\langle 0|U^\dagger \right) U |1\rangle = \frac{1}{2} - \frac{\langle x| \left(\frac{\hat{T}+\hat{T}^T}{2} \right) x\rangle}{2}. \tag{22}$$

Taking the infimum over $\mathcal{U}_d$ in (21) corresponds to taking the infimum over all unit vectors in $\mathbb{R}^3$ in (22). This gives

$$\alpha_2(\mathcal{L}) = 1 - \sup_{x \in \mathbb{R}^3 : \|x\|=1} \langle x| \left(\frac{\hat{T}+\hat{T}^T}{2} \right) x\rangle = 1 - \left\| \left(\frac{\hat{T}+\hat{T}^T}{2} \right) \right\|.$$

$\square$

[1] Note that our definition of the 2-entropy in Definition 2.2 is half of the corresponding function in [DSC96]. This explains the difference by a factor of 2 between our results.

Similarly we can also compute the LS-1 constant:

Theorem 3.5. *Let $\mathcal{L} : \mathcal{M}_2 \to \mathcal{M}_2$ be a doubly stochastic and primitive Liouvillian of the form $\mathcal{L} = T - id$, where $T : \mathcal{M}_2 \to \mathcal{M}_2$ is a quantum channel. Then,*

$$\alpha_1(\mathcal{L}) = 1 - \sup_{x \in \mathbb{R}^3 : \|x\|=1} |\langle x|\hat{T}x\rangle| \ ,$$

where $\hat{T} \in \mathcal{M}_3$ is the matrix representing the action of T on the Bloch sphere.

Proof. For fixed $U \in \mathcal{U}_2$ we define $p := (M_U)_{1,2}$, see (18). As M_U is doubly stochastic for a doubly stochastic quantum channel T we may write the Markov kernel as:

$$M_U - \mathbb{1}_2 = \begin{pmatrix} -p & p \\ p & -p \end{pmatrix} =: 2pQ, \tag{23}$$

where $Q := \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$ denotes the kernel of a random walk on the complete graph with two vertices and uniform distribution. In [DSC96] the classical LS-1 constant is defined as

$$\alpha_1^{(c)}(M_U - \mathbb{1}_2) = \inf_{s \in \mathbb{R}^2_+} -\frac{1}{2}\frac{\langle 2pQ \log(s)\,|s\rangle}{\mathrm{Ent}_1(X)} = 2p\alpha_1^{(c)}(Q)$$

where we used (23). It has been shown in [BT06] that $\alpha_1^{(c)}(Q) = 1$. By (20) we can now compute

$$\alpha_1(\mathcal{L}) = \inf_{U \in \mathcal{U}_d} \alpha_1^{(c)}(M_U - \mathbb{1}_d) = \inf_{U \in \mathcal{U}_d} 2\,\langle 1|\, U^\dagger T\left(U|0\rangle\langle 0|U^\dagger\right) U\,|1\rangle \,.$$

Changing to the Bloch sphere representation as in the proof of Theorem 3.4 finishes the proof. □

Using Theorem 2.1 we obtain the following entropy production estimate from Theorem 3.5:

Corollary 3.5 (Entropy production of qubit Liouvillians of the form $\mathcal{T} - \mathrm{id}$)**.**
Let $\mathcal{L} : \mathcal{M}_2 \to \mathcal{M}_2$ be a doubly stochastic and primitive Liouvillian of the form $\mathcal{L} = T - id$ for a quantum channel represented by $\hat{T} \in \mathcal{M}_3$ in the Bloch sphere picture with $r = \sup_{x \in \mathbb{R}^3 : \|x\|=1} |\langle x|\hat{T}x\rangle|$. Then

$$S(T_t(\rho)) - S(\rho) \geq (1 - e^{-2(1-r)t})(log(d) - S(\rho))$$

for any $\rho \in \mathcal{D}_d$ and where $T_t = e^{\mathcal{L}t}$ denotes the semigroup generated by $\mathcal{L}$.

The results from the previous section show that for doubly stochastic, primitive and reversible qubit Liouvillians of the form $\mathcal{L} = T - \mathrm{id}$ we have $\alpha_1(\mathcal{L}) = \alpha_2(\mathcal{L})$. However, in general we might have $\alpha_1(\mathcal{L}) \gg \alpha_2(\mathcal{L})$ even for reversible Liouvillians of this type. This is demonstrated for instance by the depolarizing Liouvillian $\mathcal{L}(X) = \mathrm{tr}(X)\frac{\mathbb{1}}{d} - X$ where

$$\alpha_2(\mathcal{L}) = \frac{2(1 - \frac{2}{d})}{\log(d-1)} \to 0 \text{ as } d \to \infty \quad , \text{ but } \quad \alpha_1(\mathcal{L}) \geq \frac{1}{2}.$$

Therefore, methods based on hypercontractivity may lead to entropy production estimates that are far from optimal, as the optimal constant is described by the LS-1 constant and the separation between LS-2 and LS-1 can be arbitrarily large. For the depolarizing channels in any dimension the authors succeeded in computing the exact LS-1 constant and thus the optimal entropy production in [MHSFW15].

4. Discrete LS inequalities for doubly stochastic channels

In this section we show how LS inequalities may be used to derive an entropy production estimate for doubly stochastic quantum channels in discrete time. We build upon and generalize some results of [Mic97]. As for continuous time-evolutions, we say that a doubly stochastic quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ is primitive iff $\lim_{n\to\infty} T^n(\rho) = \frac{\mathbb{1}_d}{d}$ for any $\rho \in \mathcal{D}_d$. We will need the following characterization of primitive channels:

Theorem 4.1 ([SPGWC10])**.** *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic quantum channel. The following are equivalent:*

1. *T is primitive.*
2. *There exists $n \in \mathbb{N}$ such that we have $T^n(\rho) > 0$ for any $\rho \in \mathcal{D}_d$.*
3. *T has only one eigenvalue of magnitude 1 counting multiplicities.*

We define the discrete LS constant of a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ by the usual LS-2 constant, see Definition 2.2, of a continuous semigroup associated to T.

Definition 4.1 (Discrete LS constant)**.** *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic quantum channel such that T^*T is primitive. Define the discrete LS constant of T as*

$$\alpha_D(T) := \frac{1}{2}\alpha_2(T^*T - id_d) \tag{24}$$

This definition is motivated by the following entropy production estimate:

Theorem 4.2. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic quantum channel such that T^*T is primitive. Then for any $\rho \in \mathcal{D}_d$ we have:*

$$D\left(T(\rho)\middle\|\frac{\mathbb{1}_d}{d}\right) \leq (1 - \alpha_D(T))\, D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) \tag{25}$$

Note that this is equivalent to the entropy production estimate:

$$S(T(\rho)) - S(\rho) \geq \alpha_D(T)(\log(d) - S(\rho))$$

Proof. Suppose first that $\rho \in \mathcal{D}_d^+$ and define $X = d\rho$. Our goal is to show the following inequality:

$$D\left(T(\rho)\middle\|\frac{\mathbb{1}_d}{d}\right) \leq D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) - \mathcal{E}^2_{T^*T-\mathrm{id}}\left(X^{\frac{1}{2}}\right). \tag{26}$$

For that we will use the theory of discrete LS inequalities for classical Markov chains introduced in [Mic97]. Let $\{|a_i\rangle\}_{1\leq i\leq d}$ and $\{|b_j\rangle\}_{1\leq j\leq d}$ be orthonormal bases of $\mathbb{C}^d$ consisting of eigenvectors of X and $T(X)$, respectively. As T is a doubly stochastic quantum channel the matrix $P \in \mathcal{M}_d$ defined as $(P)_{i,j} := \langle b_i| T(|a_j\rangle\langle a_j|)|b_i\rangle$ is doubly-stochastic. In the following denote by $s(X) \in \mathbb{R}^d_+$ the vector of eigenvalues of X decreasingly ordered.

Using that as $\mathbb{1}_d$ commutes with any $d \times d$-matrix we have $D(\rho\|\frac{\mathbb{1}_d}{d}) = D^{(c)}(s(\rho)\|\pi_d)$ and $D(T(\rho)\|\frac{\mathbb{1}_d}{d}) = D^{(c)}(Ps(\rho)\|\pi_d)$ where $D^{(c)}(\cdot\|\cdot)$ denotes the classical relative entropy and π_d the d-dimensional uniform distribution. Then with [Mic97, Equation 8] [2] one can easily show:

$$D\left(T(\rho)\middle\|\frac{\mathbb{1}_d}{d}\right) \leq D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) - \frac{1}{d}\sum_{i,j=1}^{d}\sqrt{s(X)_i}\left(\sqrt{s(X)_i} - \sqrt{s(X)_j}\right)\left(P^TP\right)_{i,j}. \tag{27}$$

[2] Note that in [Mic97] the evolution of a probability distribution is given by a left multiplication with the transition matrix P, while we work with right multiplication. This explains why the order of P and P^T is reversed.

Let $V \in \mathcal{U}_d$ be a unitary operator such that $\left[VT(X)V^\dagger, T\left(X^{\frac{1}{2}}\right)\right] = 0$, i.e. both operators have the same eigenvectors $\{|c_k\rangle\}_{1\leq k\leq d}$. Now define the doubly-stochastic matrix $(Q)_{i,k} = \langle c_k| T(|a_i\rangle\langle a_i|) |c_k\rangle$. As Q is doubly-stochastic so is $Q^T Q$ and we have:

$$\sum_{i,k=1}^{d} s(X)_i \left(Q^T Q\right)_{i,k} = \mathrm{tr}(X). \tag{28}$$

By construction, $Qs\left(X^{\frac{1}{2}}\right) = s\left(T\left(X^{\frac{1}{2}}\right)\right)$ and so:

$$\sum_{i,k=1}^{d} s\left(X^{\frac{1}{2}}\right)_i s\left(X^{\frac{1}{2}}\right)_k \left(Q^T Q\right)_{i,k} = \langle Qs\left(X^{\frac{1}{2}}\right)|Qs\left(X^{\frac{1}{2}}\right)\rangle = \mathrm{tr}\left[T\left(X^{\frac{1}{2}}\right)^2\right] \tag{29}$$

Using unitary invariance of the relative entropy, (27) and equations (28) and (29) we have that:

$$\begin{aligned}
D\left(T(\rho)\,\middle\|\,\frac{\mathbb{1}_d}{d}\right) &= D\left(VT(\rho)V^\dagger\,\middle\|\,\frac{\mathbb{1}_d}{d}\right) &(30)\\
&\leq D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) - \frac{1}{d}\sum_{i,k=1}^{d}\sqrt{s(X)_i}\left(\sqrt{s(X)_i} - \sqrt{s(X)_k}\right)\left(Q^T Q\right)_{i,k} &(31)\\
&= D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) - \frac{1}{d}\mathrm{tr}\left(X - T\left(X^{\frac{1}{2}}\right)^2\right) &(32)\\
&= D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) - \mathcal{E}^2_{T^*T-\mathrm{id}}\left(X^{\frac{1}{2}}\right). &(33)
\end{aligned}$$

By Definition 4.1 of the discrete LS-2 constant and by definition of X we have

$$D\left(T(\rho)\,\middle\|\,\frac{\mathbb{1}_d}{d}\right) \leq D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right) - \mathcal{E}^2_{T^*T-\mathrm{id}}\left((d\rho)^{\frac{1}{2}}\right) \leq \left(1 - \alpha_D(T)\right) D\left(\rho\middle\|\frac{\mathbb{1}_d}{d}\right)$$

for full rank ρ. The inequality follows for all states by a continuity argument.

□

Note that if T^*T is primitive, then $\alpha_D(T)$ is strictly positive by the lower bound given in Theorem 2.2 as primitivity implies the existence of a positive spectral gap by Theorem 4.1.

If the channel T is normal, i.e. $T^*T = TT^*$, it was shown in [BCG+13, Proposition 13] that T^*T being primitive is also a necessary condition for T being primitive. However, as being a strict contraction w.r.t. the relative entropy is not a necessary condition for primitivity, it can't hold that $\alpha_D(T) > 0$ for all primitive channels T. We now show that the assumption of T being primitive is sufficient to ensure that $\alpha_D(T^n) > 0$ for some $n \in \mathbb{N}$ and also derive another characterization of primitive, doubly stochastic channels:

Theorem 4.3. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic quantum channel. The following are equivalent:*

1. *T is primitive*
2. *$\exists n \in \mathbb{N}$ such that $(T^*)^n T^n$ is primitive*

Proof. If T is primitive, then by Theorem 4.1 there exists $n \in \mathbb{N}$ such that $T^n(\rho) > 0$ for all $\rho \in \mathcal{D}_d$. As T is doubly stochastic, also $(T^*)^n T^n(\rho) > 0$ holds, which implies that $(T^*)^n T^n$ is primitive. On the other hand, if $(T^*)^n T^n$ is primitive for some $n \in \mathbb{N}$, then $\alpha_D(T^n) > 0$. By Theorem 4.2 we have $\lim_{k\to\infty} T^{kn}(\rho) = \frac{\mathbb{1}_d}{d}$. As $\frac{\mathbb{1}_d}{d}$ is a fixed point the convergence for a subsequence implies the convergence of the sequence. □

We were not able to determine if inequality (25) is tight, i.e. whether there is a primitive doubly stochastic channel $T : \mathcal{M}_d \to \mathcal{M}_d$ with

$$\sup_{\rho \in \mathcal{D}_d, \rho \neq \frac{\mathbb{1}}{d}} \frac{D\left(T(\rho) \| \frac{\mathbb{1}}{d}\right)}{D\left(\rho \| \frac{\mathbb{1}}{d}\right)} = 1 - \alpha_D(T).$$

However, it is clear that the best constant in the equation (25) is not always given by the discrete LS constant. Consider for instance the completely depolarizing channel $T(\rho) := \operatorname{tr}(\rho) \frac{\mathbb{1}_d}{d}$. Then, $\alpha_D(T^*T - \mathrm{id}) = \frac{\left(1 - \frac{2}{d}\right)}{\log(d-1)} < 1$ [KT13, Corollary 27].

Now we want to determine a bound on the discrete LS constant. For this we need:

Theorem 4.4. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic, primitive quantum channel. Then the function $k \mapsto \alpha_2\left((T^*)^k T^k - id_d\right)$ is monotone increasing.*

Proof. As T is doubly stochastic, the operator Schwarz inequality[Pau02, Proposition 3.3] implies that for $X \in \mathcal{M}_d^+$

$$T^{k+1}(X)^2 \leq T\left(T^k(X)^2\right). \tag{34}$$

As T is trace-preserving we get

$$\|T^{k+1}(X)\|^2_{2, \frac{\mathbb{1}_d}{d}} = \frac{1}{d}\operatorname{tr}\left(T^{k+1}(X)^2\right) \leq \frac{1}{d}\operatorname{tr}\left(T^k(X)^2\right) = \|T^k(X)\|^2_{2, \frac{\mathbb{1}_d}{d}} \tag{35}$$

where again $\|X\|^2_{2,\sigma} = \frac{1}{d}\operatorname{tr}\left(X^2\right)$. Observe that

$$\mathcal{E}^2_{(T^*)^k T^k - \mathrm{id}_d}(X) = -\frac{1}{d}\operatorname{tr}\left(\left((T^*)^k T^k\right)(X) X - X^2\right) = \|T^k(X)\|^2_{2, \frac{1}{d}} - \|X\|^2_{2, \frac{1}{d}}.$$

Then we have by (35) that $\mathcal{E}^2_{(T^*)^{k+1} T^{k+1} - \mathrm{id}_d} \geq \mathcal{E}^2_{((T^*)^k T^k) - \mathrm{id}_d}$, which implies the claim by the variational definition of the LS-2 constant.

□

The next corollary shows that the discrete LS constant becomes less useful if the dimension d of the system grows large.

Corollary 4.1. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic quantum channel s.t. T^*T is primitive. Then*

$$\min\left\{\frac{\lambda}{2}, \frac{\left(1 - \frac{2}{d}\right)}{\log(d-1)}\right\} \geq \alpha_D(T) \geq \lambda \frac{\left(1 - \frac{2}{d}\right)}{\log(d-1)}, \tag{36}$$

where λ is the spectral gap of $T^*T - id$. Again we have $\frac{2\left(1 - \frac{2}{d}\right)}{\log(d-1)} = 1$ for $d = 2$ by continuity.

Proof. By Theorem 4.4, for $k \in \mathbb{N}$ we have $\alpha_D(T) \leq \alpha_D\left(T^k\right)$ and so

$$\alpha_D(T) \leq \liminf_{k \to \infty} \alpha_D\left(T^k\right) \tag{37}$$

By Theorem 3.3:

$$\liminf_{k \to \infty} \alpha_D\left(T^k\right) \leq \lim_{k \to \infty} \|(T^*)^k T^k - \mathrm{id}_d\| \frac{\left(1 - \frac{2}{d}\right)}{\log(d-1)}$$

For primitive channels $\lim_{k \to \infty} T^k = T_\infty$, where $T_\infty(X) = \operatorname{tr}(X) \frac{\mathbb{1}_d}{d}$. By the continuity of norms, multiplication and conjugation of linear operators and $\|\mathrm{id}_d - T_\infty\| = 1$:

$$\alpha_D(T) \leq \frac{\left(1 - \frac{2}{d}\right)}{\log(d-1)}$$

The lower bound follows from [KT13, Corollary 27] and the upper bound in terms of the spectral gap from Theorem 2.2. □

With Theorem 4.2 we immediately get the following entropy production estimate for a doubly stochastic quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$

$$S(T(\rho)) - S(\rho) \geq \lambda \frac{\left(1 - \frac{2}{d}\right)}{\log(d-1)} (\log(d) - S(\rho)) \tag{38}$$

for any $\rho \in \mathcal{D}_d$, where λ denotes the spectral gap of $T^*T - \mathrm{id}$.

Remark 4.1. The bound given in Equation (38) is similar to the one in [Str85, Lemma 2.1], given by

$$S(T(\rho)) - S(\rho) \geq \frac{\lambda}{2} \|\rho - \frac{\mathbb{1}}{d}\|_2^2,$$

where λ is again the spectral gap of $T^*T - \mathrm{id}$. However, if $\alpha_D(T)$ and λ are of the same order of magnitude, then (25) gives an improvement of order $\log(d)$. Similar holds for the bound in [Rag02].

Given the usefulness of the hypercontractive characterization of the LS-2 constant in continuous time, it is natural to ask whether we have a similar characterization of the discrete LS constant. We now show that the discrete LS constant implies hypercontractivity of the channel, but first we will prove some technical lemmas. The following lemma is the generalization of [Mic97, Lemma 3] to the quantum case.

Lemma 4.1. *For $X \in \mathcal{M}_d^+$ and $q \geq 2$:*

$$\|X\|_{q,\frac{\mathbb{1}_d}{d}} - \|X\|_{2,\frac{\mathbb{1}_d}{d}} \leq \frac{q-2}{q} \|X\|_{q,\frac{\mathbb{1}_d}{d}}^{1-q} \operatorname{Ent}_2\left(X^{\frac{q}{2}}\right) \tag{39}$$

Proof. Working in the eigenbasis of X, the proof is identical to [Mic97, Lemma 3]. □

Lemma 4.2. *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic quantum channel, $X \in \mathcal{M}_d^+$ and $q \geq 2$. Then:*

$$\|T(X)\|_{q,\frac{\mathbb{1}_d}{d}}^q - \|X\|_{q,\frac{\mathbb{1}_d}{d}}^q \leq -\mathcal{E}_{T^*T-id}^2\left(X^{\frac{q}{2}}\right) \tag{40}$$

Proof. As the r.h.s. of 40 is equal to

$$\frac{1}{d}\left[\operatorname{tr}\left(T\left(X^{\frac{q}{2}}\right)^2\right) - \operatorname{tr}\left(X^q\right)\right] \tag{41}$$

and the l.h.s. is equal to

$$\frac{1}{d}\left[\operatorname{tr}\left(T(X)^q\right) - \operatorname{tr}\left(X^q\right)\right], \tag{42}$$

the claim is equivalent to $\operatorname{tr}\left(T(X)^q\right) \leq \operatorname{tr}\left(T\left(X^{\frac{q}{2}}\right)^2\right)$. As T is a doubly stochastic channel and $q \geq 2$ [Dav57]:

$$T(X) \leq T\left(X^{\frac{q}{2}}\right)^{\frac{2}{q}} \tag{43}$$

Both $T(X)$ and $T\left(X^{\frac{q}{2}}\right)^{\frac{2}{q}}$ are positive operators, which implies by Weyl's monotonicity theorem [Bha97, Corollary III.2.3]:

$$s\left(T(X)\right)_i \leq s\left(T\left(X^{\frac{q}{2}}\right)^{\frac{2}{q}}\right)_i \tag{44}$$

As $\operatorname{tr}\left(T(X)^q\right) = \sum_{i=1}^{d} s\left(T(X)\right)_i^q$, we have:

$$\operatorname{tr}\left(T(X)^q\right) \leq \sum_{i=1}^{d} s\left(T\left(X^{\frac{q}{2}}\right)\right)_i^2 = \operatorname{tr}\left(T\left(X^{\frac{q}{2}}\right)^2\right) \tag{45}$$

and the claim follows. □

Theorem 4.5 (Discrete hypercontractivity)**.** *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a doubly stochastic quantum channel s.t. T^*T is primitive. Then $\|T\|_{2\to q,\frac{\mathbb{1}_d}{d}} \leq 1$ for $q \leq 2 + 2\alpha_D(T)$*

Proof. For $X \in \mathcal{M}_d^+$, we have by Lemma (4.1):

$$\|X\|_{q,\frac{\mathbb{1}_d}{d}} - \|X\|_{2,\frac{\mathbb{1}_d}{d}} \leq \frac{q-2}{q}\|X\|^{1-q}_{q,\frac{\mathbb{1}_d}{d}} \mathrm{Ent}_2\left(X^{\frac{q}{2}}\right) \tag{46}$$

As the function $g(x) = x^{\frac{1}{q}}$ is concave on $\mathbb{R}_+$, we have the following inequality for $a, b \in \mathbb{R}_+$:

$$a^{\frac{1}{q}} - b^{\frac{1}{q}} \leq \frac{1}{q}\left(b^{\frac{1}{q}-1}(a-b)\right) \tag{47}$$

Plugging in $a = \|T(X)\|^q_{q,\frac{\mathbb{1}_d}{d}}$ and $b = \|X\|^q_{q,\frac{\mathbb{1}_d}{d}}$ we get:

$$\|T(X)\|_{q,\frac{\mathbb{1}_d}{d}} - \|X\|_{q,\frac{\mathbb{1}_d}{d}} \leq \frac{1}{q}\|X\|^{1-q}_{q,\frac{\mathbb{1}_d}{d}}\left(\|T(X)\|^q_{q,\frac{\mathbb{1}_d}{d}} - \|X\|^q_{q,\frac{\mathbb{1}_d}{d}}\right) \tag{48}$$

Summing inequalities (46) and (48) and applying Lemma (4.2), we finally get:

$$\|T(X)\|_{q,\frac{\mathbb{1}_d}{d}} - \|X\|_{2,\frac{\mathbb{1}_d}{d}} \leq \frac{1}{q}\|X\|^{1-q}_{q,\frac{\mathbb{1}_d}{d}}\left((q-2)\mathrm{Ent}_2\left(X^{\frac{q}{2}}\right) - \mathcal{E}^2\left(X^{\frac{q}{2}}\right)\right) \tag{49}$$

By the definition of the discrete LS inequality, if $q \leq 2 + 2\alpha_D(T)$ the right-hand side in (4.2) is negative, which implies for $X \in \mathcal{M}_d^+$:

$$\frac{\|T(X)\|_{q,\frac{\mathbb{1}_d}{d}}}{\|X\|_{2,\frac{\mathbb{1}_d}{d}}} \leq 1 \tag{50}$$

As shown in [Aud09, Wat05], we may restrict ourselves to positive semi-definite operators when considering the $2 \to q$ norm of completely positive maps. By continuity, inequality (50) is also valid for all positive semi-definite operators, implying $\|T\|_{2\to q,\frac{\mathbb{1}_d}{d}} \leq 1$.

□

It is not to be expected that the other direction holds, at least in a naive way. That is, that a bound of the form:

$$\|T\|_{2\to q,\frac{\mathbb{1}_d}{d}} \leq 1$$

for some $q > 2$ and a doubly stochastic quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ gives a lower-bound on $\alpha_D(T)$ that is independent of the dimension d, as we have for continuous time.

To see why this is the case, consider a doubly stochastic qubit channel $T : \mathcal{M}_2 \to \mathcal{M}_2$. As we have mentioned before, we have [Kin14]:

$$\|T^{\otimes n}\|_{2\to q,\frac{\mathbb{1}_{d^n}}{d^n}} = \|T\|^n_{2\to q,\frac{\mathbb{1}_d}{d}}$$

If hypercontractivity of the channel would imply a lower-bound on $\alpha_D(T)$ that is independent of the dimension, we would then obtain a lower-bound strictly positive for $T^{\otimes n}$ for all $n \in \mathbb{N}$. But it follows from Corollary 4.1 that $\lim\limits_{n\to\infty} \alpha_D(T^{\otimes n}) = 0$.

5. Applications

5.1. Upper bounds on unitary quantum subdivision capacities

In [MHRW15] some of the authors introduced quantum capacities for continuous Markovian time-evolutions. These capacities are similar to the usual quantum capacity, but in addition to applying encoding and decoding operations in the beginning and end of the protocol additional operations may be applied in intermediate steps. Here we will only consider the case where these additional operations are unitary quantum channels. The precise definition of this unitary quantum subdivision capacity is:

Definition 5.1 (Unitary quantum subdivision capacity $\mathcal{Q}_{\mathfrak{C}}$[MHRW15])**.**
The $\mathfrak{U}$*-**quantum subdivision capacity** of a Liouvillian* $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ *at a time* $t \in \mathbb{R}_+$ *is defined as*

$$\mathcal{Q}_{\mathfrak{U}}(t\mathcal{L}) := \sup\{R \in \mathbb{R}_+ : \ R \text{ achievable rate}\}$$

where a rate $R \in \mathbb{R}_+$ *is called achievable if there exist sequences* $(n_\nu)_{\nu=1}^{\infty}, (m_\nu)_{\nu=1}^{\infty}$ *such that* $R = \limsup_{\nu\to\infty} \frac{n_\nu}{m_\nu}$ *and we have*

$$\inf_{k,E,D,U_1,\ldots,U_k} \left\| id_2^{\otimes n_\nu} - D \circ \prod_{l=1}^{k} \left(U_l \circ T_{\frac{t}{k}}^{\otimes m_\nu} \right) \circ E \right\|_{\diamond} \to 0 \quad \text{as } \nu \to \infty. \tag{51}$$

The latter infimum is over the number of subdivisions $k \in \mathbb{N}$ *for which the channels* $T_{\frac{t}{k}} := e^{\frac{t}{k}\mathcal{L}}$ *are defined, arbitrary encoding and decoding quantum channels* $E : \mathcal{M}_2^{\otimes n_\nu} \to \mathcal{M}_d^{\otimes m_\nu}$ *and* $D : \mathcal{M}_d^{\otimes m_\nu} \to \mathcal{M}_2^{\otimes n_\nu}$ *and unitary channels* $U_l \in \mathfrak{U}$ *from the chosen subset.*

The unitary quantum subdivision capacity quantifies the highest possible rate of information storage in a quantum memory, when unitary gates may be applied to protect the information against the noise. Using our Corollary 3.2 we obtain the following upper bound on $Q_{\mathfrak{U}}$:

Theorem 5.1. *(Upper bound for doubly stochastic, primitive Liouvillians)*
Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ *be a doubly stochastic and primitive Liouvillian with spectral gap* λ*. Then*

$$\mathcal{Q}_{\mathfrak{U}}(t\mathcal{L}) \leq e^{-2t\alpha(d)} \log(d) \quad .$$

with $\alpha(d) = \frac{\lambda(1-2d^{-2})}{\log(3)\log(d^2-1)+2(1-2d^{-2})}$.

Proof. Using Corollary 3.3 gives

$$\begin{aligned} S(T_t^{\otimes n}(\rho)) &\geq (1 - e^{-2t\alpha(d)}) \log(d^n) + e^{-2t\alpha(d)} S(\rho) \\ &\geq (1 - e^{-2t\alpha(d)}) \log(d^n) \end{aligned}$$

where $T_t = e^{t\mathcal{L}}$ and $n \in \mathbb{N}$. The rest of the proof follows the lines of the proof in [MHRW15, Theorem 5.1].

□

The above theorem shows that in a quantum memory affected by a doubly stochastic, primitive and self-adjoint noise Liouvillian the storage rate is exponentially small in time, when only unitary correction operations are allowed. This result is similar in flavor to results by Ben-Or et al. [BOGH13, ABIN96].

5.2. Entropy production for random Pauli channels

As an application of the discrete LS inequalities from Section 4, we derive an entropy production estimate for random Pauli channels.

Definition 5.2 (Random Pauli channel)**.** *A channel* $T : \mathcal{M}_2 \to \mathcal{M}_2$ *is said to be a random Pauli channel if it can be written as*

$$T(\rho) = p_1\sigma_1\rho\sigma_1 + p_2\sigma_2\rho\sigma_2 + p_3\sigma_3\rho\sigma_3 + (1 - p_1 - p_2 - p_3)\rho,$$

for a probability distribution $(p_1, p_2, p_3, 1 - p_1 - p_2 - p_3)$*. Here* $\sigma_1, \sigma_2, \sigma_3$ *are the Pauli matrices.*

First we use the results from Section 3.2 to prove:

Theorem 5.2. *Let* $T : \mathcal{M}_2 \to \mathcal{M}_2$ *be a random Pauli channel and* $\mathcal{L} : \mathcal{M}_2 \to \mathcal{M}_2$ *given by* $\mathcal{L} = T - id_2$*. Then the LS-2 constant (Definition 2.2) can be computed as*

$$\alpha_2(\mathcal{L}) = 2\min\{p_1 + p_2, p_2 + p_3, p_3 + p_1\}.$$

Proof. Note that we have $\alpha_2(\mathcal{L}) = 0$ if $\mathcal{L}$ is not primitive. Therefore, we only have to show the claim for $\mathcal{L}$ primitive. As T is reversible, Corollary 3.4 implies $\alpha_1(\mathcal{L}) = \alpha_2(\mathcal{L}) = \lambda(\mathcal{L})$. One can check that the spectrum of a random Pauli channel is given by:

$$\{1, 1-2(p_1+p_2), 1-2(p_3+p_2), 1-2(p_1+p_3)\}$$

and thus the spectral gap of $\mathcal{L}$ is given by $2\min\{p_1+p_2, p_2+p_3, p_3+p_1\}$. □

Theorem 5.3. *Let $T : \mathcal{M}_2 \to \mathcal{M}_2$ be a random Pauli channel. Define*

$$p = 2\min\{p_1p_2 + p_1p_3, p_2p_1 + p_2p_3, p_3p_1 + p_3p_2\}.$$

Then $S(T(\rho)) - S(\rho) \geq p(\log(d) - S(\rho))$.

Proof. It is easy to check that if T is a random Pauli channel,

$$T^*T(\rho) = q_1\sigma_1\rho\sigma_1 + q_2\sigma_2\rho\sigma_2 + q_3\sigma_3\rho\sigma_3 + (1-q_1-q_2-q_3)\rho$$

is a random Pauli channel as well with $q_1 = 2p_2p_3, q_2 = 2q_1q_3$ and $q_3 = 2q_1q_2$. By Theorem 5.2, $\alpha_D(T) = p$ and the claim follows from Theorem 4.2.

□

6. Conclusion

We have extended the use of group theoretic techniques to study LS inequalities for doubly stochastic, primitive Markovian time-evolutions. These bounds lead to entropy-production estimates for tensor-powers of this kind of semigroups, which are independent of the number of tensor-powers. We applied these estimates to derive upper bounds on quantum subdivision capacities. For discrete doubly stochastic quantum channels we generalized discrete LS inequalities to the quantum case.

There are some directions of possible future research which should be emphasized. It would be desirable to try our group theoretic approach with other relevant semigroups and generalize it to semigroups that are not doubly stochastic. Another concrete open question is, whether the LS-2 constant can actually decrease under taking tensor powers.

Concerning LS inequalities for discrete channels there are similar open problems. Again proving analogous statements for channels with arbitrary fixed points and determine the discrete LS constant for more channels would be interesting. New techniques that would yield bounds stable under tensor powers of discrete channels would also be of great interest.

We believe that our proofs illustrate how comparison inequality techniques can be useful and finding systematic methods to establish them, as there are for classical Markov chains, would be interesting.

Acknowledgements

A. Müller-Hermes and M. Wolf acknowledge support from the CHIST-ERA/BMBF project CQC. D. Stilck França acknowledges support from the graduate program TopMath of the Elite Network of Bavaria, the TopMath Graduate Center of TUM Graduate School at Technische Universität München and the Deutscher Akademischer Austauschdienst(DAAD). M. Wolf is also supported by the Alfried Krupp von Bohlen und Halbach-Stiftung and by grant #48322 from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

A. Hypercontractivity via group theory

Here we will consider reversible quantum dynamical semigroups with an eigenbasis consisting of unitaries commuting up to a phase. By relating such quantum semigroups to classical semigroups defined on finite abelian groups we can use the classical theory to prove hypercontractivity. We explore and extend ideas similar to [JPPP15]. In particular we get a bound on the $2 \to 4$ norm of tensor products of depolarizing channels. We start by reviewing some basic facts about Fourier analysis on abelian groups, the proofs of which can be found in [SS11, Chapter 7].

Given a finite abelian group G of order $|G|$ we will denote by $V(G)$ the vector space of all functions $f : G \to \mathbb{C}$. For $f \in V(G)$ we define its l_p-norm by

$$\|f\|_p^p = \frac{1}{|G|} \sum_{g \in G} |f(g)|^p$$

and for linear operators $A : V(G) \to V(G)$ we define

$$\|A\|_{p \to q} = \sup_{f \in V(G)} \frac{\|Af\|_q}{\|f\|_p}.$$

The characters of G are functions $\chi : G \to \{z \in \mathbb{C} : |z| = 1\}$ such that $\chi(g_1 g_2) = \chi(g_1)\chi(g_2)$ holds for any $g_1, g_2 \in G$. We will denote by $\hat{G}$ the set of all characters of G, which is again a group isomorphic to G under multiplication. The characters form an orthonormal basis for the space of functions $V(G)$ under the scalar product

$$\langle f|h \rangle = \frac{1}{|G|} \sum_{g \in G} f(g)^* h(g).$$

We have $f = \sum_{\chi_i \in \hat{G}} \hat{f}(i)\chi_i$ for $\hat{f}(i) = \langle f|\chi_i \rangle$. We also define $G \times G'$ to be the direct product of two groups G, G' and we have

$$\widehat{G \times G'} \cong \hat{G} \times \hat{G}',$$

i.e. all characters of $G \times G'$ are of the form $\chi\chi'$ for $\chi \in \hat{G}$ and $\chi' \in \hat{G}'$

Definition A.1 (Almost commuting unitary Basis). *A set of unitaries $\{U_i\}_{0 \le i \le d^2-1} \subset \mathcal{M}_d$ with $U_0 = \mathbb{1}_d$ is called an almost commuting unitary basis (associated to an abelian group G) if:*

1. *$tr[U_i^\dagger U_j] = d\delta_{i,j}$ for all $0 \le i, j \le d^2 - 1$.*
2. *The $\{U_i\}_{0 \le i \le d^2-1}$ are a projective representation of G, i.e. for all $0 \le i, j \le d^2 - 1$ we have $U_i U_j = \phi(i,j) U_j U_i$ and $U_i U_j = \phi'(i,j) U_{i+j}$ for some $\phi'(i,j), \phi(i,j) \in \mathbb{C}$ with $|\phi'(i,j)| = |\phi(i,j)| = 1$, where in the index we mean addition in the group G.*

We can then associate each unitary to a character in $\hat{G}$.

A prominent example of an almost commuting unitary basis is the discrete Weyl system of unitaries $\{U_{k,l}\}_{0 \le k,l \le d} \subset \mathcal{M}_d$ given by:

$$U_{k,l} = \sum_{r=0}^{d-1} \nu^{rl} |k + r\rangle \langle r| , \quad \nu = e^{\frac{2i\pi}{d}}. \tag{52}$$

It is easy to check the properties:

1. $\operatorname{tr}\left(U_{i,j}^\dagger U_{k,l}\right) = d\delta_{i,k}\delta_{j,l}$ for any $i, j, k, l \in \{0, \ldots, d-1\}$.
2. $U_{i,j} U_{k,l} = \nu^{jk} U_{i+k,j+l}$
3. $U_{k,l}^{-1} = \nu^{kl} U_{-k,-l}$ for any $k, l \in \{0, \ldots, d-1\}$.

These unitaries are a projective representation of $\mathbb{Z}_d \times \mathbb{Z}_d$ in $PU(d)$. This basis has been explored in [JPPP15] to derive similar results on hypercontractivity. However, it should be noted that there are other examples of almost commuting bases and that tensoring leads to further examples.

By associating an almost commuting basis on $\mathcal{M}_d$ to the orthonormal basis of characters on $V(G)$ we can also relate norms on $\mathcal{M}_d$ to corresponding norms on $V(G)$. For any $X \in \mathcal{M}_d$ we define

$$f_X := \sum_{i=0}^{d^2-1} \hat{f}_X(i)\chi_i \tag{53}$$

with $\hat{f}_X(i) = \langle U_i|X\rangle_{\frac{1}{d}}$.

Lemma A.1. *Let $\{U_i\}_{0\leq i\leq d^2-1}$ be an almost commuting unitary basis and G the group associated to it. For any $X \in \mathcal{M}_d$ and f_X as in* (53) *we have*

$$\|X\|^2_{2,\frac{1}{d}} = \|f_X\|^2_2.$$

Proof. It follows immediately from the definition of an almost commuting unitary basis that it is an orthonormal basis w.r.t. $\langle\cdot|\cdot\rangle_{\frac{1}{d}}$ and so we have:

$$\|X\|^2_{2,\frac{1}{d}d} = \langle X|X\rangle_{\frac{1}{d}} = \sum_{i=0}^{d^2-1} |\langle U_i|X\rangle_{\frac{1}{d}}|^2 = \sum_{i=0}^{d^2-1} |\hat{f}(i)|^2 = \|f\|^2$$

as by Plancherel's identity $\|f\|_2^2 = \sum_{i=0}^{d^2-1} |\hat{f}(i)|^2$.

□

In the following $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ will denote a unital, reversible Liouvillian with spectrum $\{\lambda_i\}_{0\leq i\leq d^2-1} \subset \mathbb{R}$ and with unitary eigenvectors $\{U_i\}_{0\leq i\leq d^2-1}$ forming an almost commuting unitary basis. To such a Liouvillian we associate a classical semigroup $P_t : V(G) \to V(G)$ defined as:

$$P_t f := \sum_{i=0}^{d^2-1} e^{\lambda_i t}\hat{f}(i)\chi_i. \tag{54}$$

With this definition we can state the following theorem:

Theorem A.1. *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a unital, reversible Liouvillian with an almost commuting unitary eigenbasis and $P_t : V(g) \to V(g)$ the associated classical semigroup as in* (54)*. Then:*

$$\|e^{t\mathcal{L}}\|_{2\to 4,\frac{1}{d}} \leq \|P_t\|_{2\to 4}$$

Proof. Let $\{\lambda_i\}_{0\leq i\leq d^2-1}$ denote the spectrum of $\mathcal{L}$ and $\{U_i\}_{0\leq i\leq d^2-1}$ the almost commuting unitary eigenbasis. Any $X \in \mathcal{M}_d$ can be written as $X = \sum_{i=0}^{d^2-1} \hat{f}(i)U_i$ with $\hat{f}_X(i) = \langle U_i|X\rangle_{\frac{1}{d}}$ and we can also define f_X as in (53). Then we have:

$$\|X\|^4_{4,\frac{1}{d}} = \frac{1}{d}\mathrm{tr}[X^\dagger X X^\dagger X] = \frac{1}{d}\sum_{i_1,i_2,i_3,i_4} \hat{f}_X(i_1)^* \hat{f}_X(i_2) \hat{f}_X(i_3)^* \hat{f}_X(i_4)\, \mathrm{tr}[U^\dagger_{i_1} U_{i_2} U^\dagger_{i_3} U_{i_4}]$$

The unitaries $\{U_i\}_{i=0}^{d^2-1}$ commute and form a group up to a phase. Also by the orthogonality condition we have $\mathrm{tr}(U_i) = 0$ for any $i \neq 0$, which implies

$$\frac{1}{d}|\mathrm{tr}[U^\dagger_{i_1} U_{i_2} U^\dagger_{i_3} U_{i_4}]| = \delta_{i_2+i_4-i_3-i_1,0}.$$

By the triangle inequality we have:

$$\|X\|_{4,\frac{1}{d}}^4 \leq \sum_{i_2+i_4-i_3-i_1=0} |\hat{f}_X(i_1)\,\hat{f}_X(i_2)\,\hat{f}_X(i_3)\,\hat{f}_X(i_4)\,| \tag{55}$$

Now define $f'_X \in V(G)$ as $f'_X = \sum_{i=0}^{d^2-1} |\hat{f}_X(i)|\chi_i$ and note that

$$\begin{aligned}\|f_X\|_4^4 = \|f'_X\|_4^4 &= \frac{1}{d^2}\sum_{g\in G}\left|\sum_{i=0}^{d^2-1}\left|\hat{f}_X(i)\right|\chi_i(g)\right|^4 \\ &= \frac{1}{d^2}\sum_{g\in G}\sum_{i_1,i_2,i_3,i_4}\left|\hat{f}_X(i_1)\,\hat{f}_X(i_2)\,\hat{f}_X(i_3)\,\hat{f}_X(i_4)\right|\chi_{i_2+i_4-i_3-i_1}(g).\end{aligned}$$

where we have used the identities $\chi_i(g)^* = \chi_{-i}(g)$ and $\chi_{i_1}(g)\chi_{i_2}(g) = \chi_{i_1+i_2}(g)$ for any $g \in G$. By the orthogonality of characters we have $\sum_{g\in G}\chi_a(g) = d^2\delta_{a0}$ and therefore

$$\|f_X\|_4^4 = \sum_{i_2+i_4-i_3-i_1=0} |\hat{f}_X(i_1)\,\hat{f}_X(i_2)\,\hat{f}_X(i_3)\,\hat{f}_X(i_4)\,|.$$

It then follows from (55) that $\|X\|_{4,\frac{1}{d}}^4 \leq \|f_X\|_4^4$.

The semigroup $e^{t\mathcal{L}}$ acts as a multiplication operator in the almost commuting unitary eigenbasis $\{U_i\}_{i=0}^{d^2}$. By construction the classical semigroup P_t from (54) associated to $e^{t\mathcal{L}}$ acts in the same way in the basis of characters. This implies:

$$\|e^{t\mathcal{L}}X\|_{4,\frac{1}{d}} \leq \|P_t f_X\|_4 \leq \|P_t\|_{2\to4}\|f_X\|_2 = \|P_t\|_{2\to4}\|X\|_{2,\frac{1}{d}} \tag{56}$$

using Lemma A.1 for the last equality.

□

Note that the proof of the previous theorem can be used for any $2 \to q$ norm with q an even integer [JPPP15].

Consider two unital, reversible Liouvillians $\mathcal{L}_1 : \mathcal{M}_{d_1} \to \mathcal{M}_{d_1}$ and $\mathcal{L}_2 : \mathcal{M}_{d_2} \to \mathcal{M}_{d_2}$ with spectra $\{\lambda_i\}_{i=0}^{d_1^2-1}$ and $\{\mu_j\}_{j=0}^{d_2^2-1}$ and almost commuting unitary eigenbases $\{U_i^1\}_{0\leq i\leq d_1^2-1}$ and $\{U_j^2\}_{0\leq j\leq d_2^2-1}$ associated to abelian groups G_1 and G_2. Now we can apply the above theorem to the tensor product semigroup $e^{t\mathcal{L}} = e^{t\mathcal{L}_1} \otimes e^{t\mathcal{L}_2}$ generated by $\mathcal{L} = \mathcal{L}_1 \otimes \mathrm{id}_{d_2} + \mathrm{id}_{d_1} \otimes \mathcal{L}_2$. It can be verified easily that $\{U_i^1 \otimes U_j^2\}_{0\leq i\leq d_1^2-1, 0\leq j\leq d_2^2-1}$ is an almost commuting unitary eigenbasis for $\mathcal{L}$ associated to the abelian group $G_1 \times G_2$. Let Q_t denote the classical semigroup acting on $V(G_1)\otimes V(G_2) \cong V(G_1 \times G_2)$ associated to $e^{t\mathcal{L}}$ as in (54). Also let P_t^1 and P_t^2 denote the classical semigroups associated in the same way to $e^{t\mathcal{L}_1}$ and $e^{t\mathcal{L}_2}$ respectively. Note that for any $\chi_{i,j} \in \widehat{G_1 \times G_2}$ we have

$$Q_t\chi_{i,j} = Q_t\chi_i\chi_j = e^{\lambda_i t}e^{\mu_j t}\chi_i\chi_j = P_t^1 \otimes P_t^2\chi_{i,j}$$

and hence $Q_t = P_t^1 \otimes P_t^2$ as the characters (in $\widehat{G_1 \times G_2}$) form a basis of $V(G_1 \times G_2)$. This proves the following corollary:

Corollary A.1. *Let $\mathcal{L}_1 : \mathcal{M}_{d_1} \to \mathcal{M}_{d_1}$ and $\mathcal{L}_2 : \mathcal{M}_{d_2} \to \mathcal{M}_{d_2}$ be unital, reversible Liouvillians with almost commuting unitary eigenbases associated to abelian groups G_1 and G_2. Furthermore, let P_t^1 and P_t^2 be the associated classical semigroups as in* (54) *acting on $V(G_1)$ and $V(G_2)$. Then:*

$$\|e^{t\mathcal{L}_1} \otimes e^{t\mathcal{L}_2}\|_{2\to4,\frac{1}{d_1 d_2}} \leq \|P_t^1 \otimes P_t^2\|_{2\to4}$$

References

[ABIN96] Dorit Aharonov, Michael Ben-Or, Russell Impagliazzo, and Noam Nisan. Limitations of Noisy Reversible Computation. *eprint arXiv:quant-ph/9611028*, November 1996.

[Aud09] Koenraad M.R. Audenaert. A note on the p→q norms of 2-positive maps. *Linear Algebra and Its Applications*, 430(4):1436–1440, 2009.

[BCG+13] Daniel Burgarth, Giulio Chiribella, Vittorio Giovannetti, Paolo Perinotti, and Kazuya Yuasa. Ergodic and mixing quantum channels in finite dimensions. *New Journal of Physics*, 15(7):073045, July 2013.

[Bha97] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer, 1997.

[BOGH13] Michael Ben-Or, Daniel Gottesman, and Avinatan Hassidim. Quantum refrigerator. *arXiv preprint arXiv:1301.1995*, 2013.

[BT06] Sergey G. Bobkov and Prasad Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *Journal of Theoretical Probability*, 19(2):289–336, 2006.

[BZ00] Thierry Bodineau and Boguslaw Zegarlinski. Hypercontractivity via spectral theory. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 03(01):15–31, 2000.

[CLMP13] Toby S. Cubitt, Angelo Lucia, Spyridon Michalakis, and David Perez-Garcia. Stability of local quantum dissipative systems. *ArXiv e-prints*, March 2013.

[Dav57] Chandler Davis. A Schwarz inequality for convex operator functions. *Proceedings of the American Mathematical Society*, 8(1):pp. 42–44, 1957.

[DSC93] Persi Diaconis and Laurent Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.*, 3(3):696–730, 08 1993.

[DSC96] Persi Diaconis and Laurent Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695–750, 08 1996.

[GKS76] Vittorio Gorini, Andrzej Kossakowski, and E. C. George Sudarshan. Completely positive dynamical semigroups of N-level systems. *Journal of Mathematical Physics*, 17:821–825, May 1976.

[JPPP15] Marius Junge, Carlos Palazuelos, Javier Parcet, and Mathilde Perrin. Hypercontractivity in finite-dimensional matrix algebras. *Journal of Mathematical Physics*, 56(2):–, 2015.

[Kin14] Christopher King. Hypercontractivity for semigroups of unital qubit channels. *Communications in Mathematical Physics*, 328(1):285–301, 2014.

[KT13] Michael J. Kastoryano and Kristan Temme. Quantum logarithmic Sobolev inequalities and rapid mixing. *Journal of Mathematical Physics*, 54(5):052202, May 2013.

[Lin76] Göran Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, 1976.

[MHRW15] Alexander Müller-Hermes, David Reeb, and Michael M. Wolf. Quantum subdivision capacities and continuous-time quantum coding. *Information Theory, IEEE Transactions on*, 61(1):565–581, Jan 2015.

[MHSFW15] Alexander Müller-Hermes, Daniel Stilck França, and Michael M. Wolf. Relative entropy convergence for depolarizing channels. *In preparation*, 2015.

[Mic97] Laurent Miclo. Remarques sur l'hypercontractivité et l'évolution de l'entropie pour des chaînes de Markov finies. In *Séminaire de Probabilités XXXI*, volume 1655 of *Lecture Notes in Mathematics*, pages 136–167. Springer Berlin Heidelberg, 1997.

[NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.

[OZ99] Robert Olkiewicz and Boguslaw Zegarlinski. Hypercontractivity in noncommutative lp spaces. *Journal of Functional Analysis*, 161(1):246 – 285, 1999.

[Pau02] Vern Paulsen. *Completely Bounded Maps and Operator Algebras*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2002.

[Rag02] Maxim Raginsky. Entropy production rates of bistochastic strictly contractive quantum channels on a matrix algebra. *Journal of Physics A: Mathematical and General*, 35(41):L585–L590, 2002.

[SPGWC10] Mikel Sanz, David Perez-Garcia, Michael M. Wolf, and Juan I. Cirac. A quantum version of Wielandt's inequality. *Information Theory, IEEE Transactions on*, 56(9):4668–4673, Sept 2010.

[Spo78] Herbert Spohn. Entropy production for quantum dynamical semigroups. *Journal of Mathematical Physics*, 19(5):1227–1230, 1978.

[SS11] Elias M. Stein and Rami Shakarchi. *Fourier analysis: an introduction*, volume 1. Princeton University Press, 2011.

[Str85] Ray F. Streater. Convergence of the quantum Boltzmann map. *Communications in Mathematical Physics*, 98(2):177–185, 1985.

[TPK14] Kristan Temme, Fernando Pastawski, and Michael J. Kastoryano. Hypercontractivity of quasi-free quantum semigroups. *ArXiv e-prints*, March 2014.

[Wat05] John Watrous. Notes on super-operator norms induced by Schatten norms. *Quantum Info. Comput.*, 5(1):58–68, January 2005.

Relative Entropy Convergence for Depolarizing Channels

A. Müller-Hermes, D. Stilck França and M. Wolf May 19, 2015

We study quantum Markovian time-evolutions generated by general depolarizing Liouvillians, i.e. $\mathcal{L}_\sigma : \mathcal{M}_d \to \mathcal{M}_d$ given by $\mathcal{L}_\sigma(\rho) = \mathrm{tr}(\rho)\sigma - \rho$ for some quantum state $\sigma \in \mathcal{D}_d$. Here we restrict to the case where σ is full-rank (otherwise the logarithmic-Sobolev framework is not well-defined). The logarithmic-Sobolev constant $\alpha_1(\mathcal{L}_\sigma)$ can be seen as the optimal exponent $\alpha \in \mathbb{R}$ such that

$$D(T_t^\sigma(\rho)\|\sigma) \leq e^{-2\alpha t} D(\rho\|\sigma) \tag{1}$$

holds for any $\rho \in \mathcal{D}_d$ and any $t \in \mathbb{R}^+$.

1 Main result

Our main result is the computation of $\alpha_1(\mathcal{L}_\sigma)$. As a direct consequence of (1) we have to solve the following optimization problem:

$$\alpha_1(\mathcal{L}_\sigma) = \inf_{\rho \in \mathcal{D}_d^+} \frac{1}{2}\Big(1 + \frac{D(\sigma\|\rho)}{D(\rho\|\sigma)}\Big).$$

The quotient of relative entropies appearing in the optimization is a quasi-linear function in the entries of the doubly stochastic matrix $P_{ij} = |\langle v_i | w_j \rangle|^2$ depending on the eigenbases $\{|v_i\rangle\}$ and $\{|w_j\rangle\}$ of ρ and σ. Then Birkhoff's theorem implies that the optimal ρ has to commute with σ. By applying Lagrange-multipliers we obtain:

Theorem 1.1. *Let $\mathcal{L}_\sigma : \mathcal{M}_d \to \mathcal{M}_d$ be the depolarizing Liouvillian with full-rank fixed point $\sigma \in \mathcal{D}_d$. Then we have*

$$\alpha_1(\mathcal{L}_\sigma) = \min_{x \in [0,1]} \frac{1}{2}\left(1 + \frac{D_2(s_{min}(\sigma)\|x)}{D_2(x\|s_{min}(\sigma))}\right),$$

where $s_{min}(\sigma)$ denotes the minimal eigenvalue of σ.

Our proof shows, that the above constant is also the logarithmic-Sobolev constant for a classical random walk on $\{1, \ldots, d\}$ with transition probabilities equal to the eigenvalues of σ. To the best of our knowledge this constant has not been computed before.

2 Concavity of von-Neumann entropy

As an application of Theorem 1.1 we improve the concavity inequality of the von-Neumann entropy:

Theorem 2.1 (Improved Concavity of the von-Neumann Entropy)**.** *Let $\rho \in \mathcal{D}_d$ and $\sigma \in \mathcal{D}_d^+$ with minimal eigenvalue $s_{\min}(\sigma)$. Then for $q \in [0,1]$ we have*

$$S((1-q)\sigma + q\rho) \geq (1-q)S(\sigma) + qS(\rho) + q(1-q^{c(\sigma)})D(\rho\|\sigma),$$

with

$$c(\sigma) = \min_{x\in[0,1]} \frac{D_2(s_{\min}(\sigma)\|x)}{D_2(x\|s_{\min}(\sigma))}$$

The proof is a straightforward rewriting of (1) with $\alpha = \alpha_1(\mathcal{L}_\sigma)$. Our result seems to be incomparable to similar improvements [2] using different quantities for the correction terms. While in some cases our bound performs better (by numerical value), there are other cases where it performs worse. Furthermore, our proof gives a similar result for the Shannon entropy.

3 Tensor-powers

Let $\mathcal{L}_\sigma^{(n)}$ denote the generator of the semigroup $(T_t^\sigma)^{\otimes n}$. In the special case where $\sigma = \frac{\mathbb{1}_d}{d}$ is the maximally mixed state we can prove the lower bound

$$\alpha_1\left(\mathcal{L}^{(n)}_{\frac{\mathbb{1}_d}{d}}\right) \geq \frac{1}{2}$$

for any $n \in \mathbb{N}$ and any $d \geq 2$. This bound is a direct consequence of the following entropy-production estimate:

Theorem 3.1. *For any $\sigma, \rho \in \mathcal{D}_d$ (not necessarily full rank) we have*

$$S((T_t^\sigma)^{\otimes n}(\rho)) \geq e^{-t}S(\rho) + (1-e^{-t})S(\sigma^{\otimes n}).$$

This theorem was first considered (though with wrong proof) in [1] for the special case $\sigma = \frac{\mathbb{1}_2}{2}$. We prove the above theorem using a quantum version of Shearer's inequality for entropies.

4 Legal statement

In all parts of this work, except the section on the improved Pinsker's inequality, I was significantly involved.

References

[1] Dorit Aharonov, Michael Ben-Or, Russell Impagliazzo, and Noam Nisan. Limitations of noisy reversible computation. *arXiv preprint quant-ph/9611028*, 1996.

[2] Isaac Kim and Mary Beth Ruskai. Bounds on the concavity of quantum entropy. *Journal of Mathematical Physics*, 55(9), 2014.

Relative Entropy Convergence for Depolarizing Channels

Alexander Müller-Hermes*, Daniel Stilck França†, and Michael M. Wolf‡

Department of Mathematics, Technische Universität München, 85748 Garching, Germany

August 27, 2015

We study the convergence of states under continuous-time depolarizing channels with full rank fixed points in terms of the relative entropy. The optimal exponent of an upper bound on the relative entropy in this case is given by the log-Sobolev-1 constant. Our main result is the computation of this constant. As an application we use the log-Sobolev-1 constant of the depolarizing channels to improve the concavity inequality of the von-Neumann entropy. This result is compared to similar bounds obtained recently by Kim et al. and we show a version of Pinsker's inequality, which is optimal and tight if we fix the second argument of the relative entropy. Finally, we consider the log-Sobolev-1 constant of tensor-powers of the completely depolarizing channel and use a quantum version of Shearer's inequality to prove a uniform lower bound.

1. Introduction

Let $\mathcal{M}_d$ denote the set of complex $d \times d$-matrices, $\mathcal{D}_d \subset \mathcal{M}_d$ the set of quantum states, i.e. positive matrices with trace equal to 1, and $\mathcal{D}_d^+$ the set of strictly positive states. The relative entropy (also called quantum Kullback-Leibler divergence) of $\rho, \sigma \in \mathcal{D}_d$ is defined as

$$D(\rho\|\sigma) := \begin{cases} \operatorname{tr}[\rho(\log\rho - \log\sigma)], & \text{if } \operatorname{supp}(\rho) \subset \operatorname{supp}(\sigma) \\ +\infty, & \text{otherwise} \end{cases}. \tag{1}$$

*muellerh@ma.tum.de
†dsfranca@mytum.de
‡wolf@ma.tum.de

The relative entropy defines a natural distance measure to study the convergence of Markovian time-evolutions. For some state $\sigma \in \mathcal{D}_d$ consider the generalized depolarizing Liouvillian $\mathcal{L}_\sigma : \mathcal{M}_d \to \mathcal{M}_d$ defined as

$$\mathcal{L}_\sigma(\rho) := \mathrm{tr}[\rho]\,\sigma - \rho. \tag{2}$$

This Liouvillian generates the generalized depolarizing channel $T^\sigma_t : \mathcal{M}_d \to \mathcal{M}_d$ with $T^\sigma_t(\rho) := e^{t\mathcal{L}_\sigma}(\rho) = (1-e^{-t})\mathrm{tr}[\rho]\,\sigma + e^{-t}\rho$, where $t \in \mathbb{R}^+$ denotes a time parameter. As $T^\sigma_t(\rho) \to \sigma$ for $t \to \infty$ we can study the convergence speed of the depolarizing channel with a full rank fixed point $\sigma \in \mathcal{D}_d$ by determining the largest constant $\alpha \in \mathbb{R}^+$ such that

$$D(T^\sigma_t(\rho)\|\sigma) \leq e^{-2\alpha t} D(\rho\|\sigma) \tag{3}$$

holds for any $\rho \in \mathcal{D}_d$ and any $t \in \mathbb{R}^+$. This constant is known as the logarithmic Sobolev-1 constant [1, 2] of $\mathcal{L}_\sigma$, denoted by $\alpha_1(\mathcal{L}_\sigma)$. In the following we will compute this constant and then use it to derive an improvement on the concavity of von-Neumann entropy.

2. Preliminaries and notation

Consider a primitive[1] Liouvillian $\mathcal{L}$ with full rank fixed point $\sigma \in \mathcal{D}_d$ and denote by $T_t := e^{t\mathcal{L}}$ the quantum dynamical semigroup generated by $\mathcal{L}$. Consider the function $f(t) := D\left(T_t(\rho)\middle\|\sigma\right)$ for some initial state $\rho \in \mathcal{D}_d$ and note that if

$$\frac{df}{dt} \leq -2\alpha f$$

holds for some $\alpha \in \mathbb{R}_+$, then it follows that $f(t) \leq e^{-2\alpha t} f(0)$. The time derivative of the relative entropy at $t = 0$, also called the entropy production [3], is given by:

$$\frac{d}{dt} D\left(T_t(\rho)\middle\|\sigma\right)\bigg|_{t=0} = -\mathrm{tr}[\mathcal{L}(\rho)(\log(\sigma) - \log(\rho))] \tag{4}$$

as $\mathrm{tr}(\mathcal{L}(\rho)) = 0$ for any $\rho \in \mathcal{D}_d$. This motivates the following definition:

Definition 2.1 (log-Sobolev-1 constant, [1, 2])**.** For a primitive Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ with full rank fixed point $\sigma \in \mathcal{D}_d$ we define its **log-Sobolev-1 constant** as

$$\alpha_1(\mathcal{L}) := \sup\left\{\alpha \in \mathbb{R} : \mathrm{tr}[\mathcal{L}(\rho)(\log(\sigma) - \log(\rho))] \geq 2\alpha D(\rho\|\sigma), \forall \rho \in \mathcal{D}^+_d\right\} \tag{5}$$

[1] A Liouvillian is primitive if, and only if, it has a unique full rank fixed point σ and for any $\rho \in \mathcal{D}_d$ we have $e^{t\mathcal{L}}(\rho) \to \sigma$ as $t \to \infty$.

For a primitive Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ the preceding discussion shows that (3) holds for any $\alpha \leq \alpha_1(\mathcal{L})$. Furthermore, $\alpha_1(\mathcal{L})$ is the optimal constant for which this inequality holds independent of $\rho \in \mathcal{D}_d$ (for states ρ not of full rank this follows from a simple continuity argument).

In the following we will need some functions defined as continuous extensions of quotients of relative entropies. We denote by $Q_\sigma : \mathcal{D}_d^+ \to \mathbb{R}$ the continuous extension of the function $\rho \mapsto \frac{D(\sigma\|\rho)}{D(\rho\|\sigma)}$ (see Appendix B) given by

$$Q_\sigma(\rho) := \begin{cases} \frac{D(\sigma\|\rho)}{D(\rho\|\sigma)}, & \rho \neq \sigma \\ 1, & \rho = \sigma \end{cases}. \tag{6}$$

Note that for $x \in [0,1]$ and $y \in (0,1)$ the binary relative entropy is defined as

$$D_2(x\|y) := x \log\left(\frac{x}{y}\right) + (1-x)\log\left(\frac{1-x}{1-y}\right). \tag{7}$$

This is the classical relative entropy of the probability distributions $(x, 1-x)$ and $(y, 1-y)$. For $y \in (0,1)$ we denote by $q_y : (0,1) \to \mathbb{R}$ the continuous extension of $x \mapsto \frac{D_2(y\|x)}{D_2(x\|y)}$ given by

$$q_y(x) := \begin{cases} \frac{D_2(y\|x)}{D_2(x\|y)}, & x \neq y \\ 1, & x = y \end{cases}. \tag{8}$$

3. Log-Sobolev-1 constant for the depolarizing Liouvillian

Note that for the depolarizing Liouvillian $\mathcal{L}_\sigma$ with $\sigma \in \mathcal{D}_d^+$ as defined in (2) we have

$$\mathrm{tr}[\mathcal{L}_\sigma(\rho)(\log(\sigma) - \log(\rho))] = D(\rho\|\sigma) + D(\sigma\|\rho).$$

Inserting this into Definition 2.1 we can write

$$\alpha_1(\mathcal{L}_\sigma) = \inf_{\rho \in \mathcal{D}_d^+} \frac{1}{2}\Big(1 + Q_\sigma(\rho)\Big). \tag{9}$$

Our main result is the following theorem:

Theorem 3.1. *Let $\mathcal{L}_\sigma : \mathcal{M}_d \to \mathcal{M}_d$ be the depolarizing Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d$ as defined in* (2)*. Then we have*

$$\alpha_1(\mathcal{L}_\sigma) = \min_{x \in [0,1]} \frac{1}{2}\left(1 + q_{s_{min}(\sigma)}(x)\right),$$

where $s_{min}(\sigma)$ denotes the minimal eigenvalue of σ.

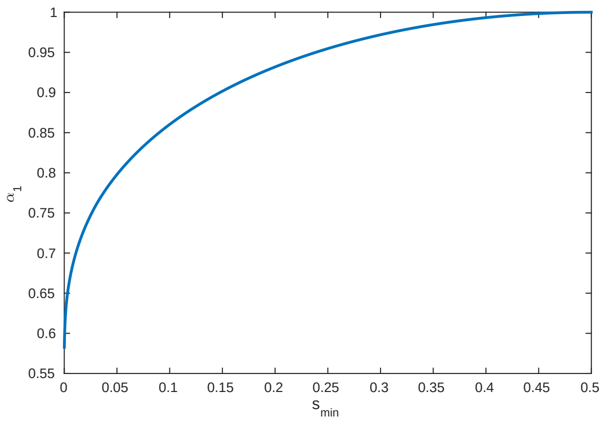


Figure 1: $\alpha_1(\mathcal{L}_\sigma)$ for $s_{\min}(\sigma) \in [0,1]$.

In Figure 1 the values of $\alpha_1(\mathcal{L}_\sigma)$ depending on $s_{\min}(\sigma) \in [0,1]$ are plotted. Note that by Theorem 3.1 we have $\alpha_1(\mathcal{L}_\sigma) \to 1/2$ in the limit $s_{\min}(\sigma) \to 0$ (as $D_2(s_{\min}(\sigma)\|x) \to 0$ and $D_2(x\|s_{\min}(\sigma)) \to \infty$ in this case).

Before we state the proof of Theorem 3.1 we need to make a technical comment. By (6) we have $Q_\sigma(\rho) \to +\infty$ as $\rho \to \partial\mathcal{D}_d$, i.e. as ρ converges to a rank-deficient state. Therefore, the infimum in (9) will be attained in a full rank state $\tilde{\rho} \in \mathcal{D}_d^+$ and we can restrict the optimization to the compact set $K_\sigma \subset \mathcal{D}_d$ (depending on σ) defined as

$$K_\sigma = \{\rho \in \mathcal{D}_d^+ : s_{\min}(\rho) \geq s_{\min}(\tilde{\rho}) - \epsilon\} \tag{10}$$

for some fixed $\epsilon \in (0, s_{\min}(\tilde{\rho}))$ and where $s_{\min}(\cdot)$ denotes the minimal eigenvalue. Note that the minimizing state $\tilde{\rho}$ is contained in the interior of K_σ. Now we have to solve the following optimization problem for fixed $\sigma \in \mathcal{D}_d^+$:

$$\inf_{\rho \in \mathcal{D}_d^+} Q_\sigma(\rho) = \inf_{\rho \in K_\sigma} Q_\sigma(\rho). \tag{11}$$

To prove Theorem 3.1 we will need the following lemma showing that the infimum in (11) is attained at states $\rho \in \mathcal{D}_d$ commuting with the fixed point σ.

Lemma 3.1. *For any $\sigma \in \mathcal{D}_d^+$ we have*

$$\inf_{\rho \in K_\sigma} Q_\sigma(\rho) = \inf_{\rho \in K_\sigma, [\rho,\sigma]=0} Q_\sigma(\rho)$$

where $Q_\sigma : D_d^+ \to \mathbb{R}$ denotes the continuous extension of $\rho \mapsto \frac{D(\sigma\|\rho)}{D(\rho\|\sigma)}$ (see (6)).

Proof. Consider the spectral decomposition $\sigma = \sum_{i=1}^{d} s_i|v_i\rangle\langle v_i|$ for $s \in \mathbb{R}_+^d$ and fix a vector $r \in \mathbb{R}_+^d$ which is not a permutation of s and which fulfills $\min_i(r_i) \geq s_{\min}(\tilde{\rho}) - \epsilon$ (see (10)) and $\sum_{j=1}^{d} r_j = 1$. For some fixed orthonormal basis $\{|w_j\rangle\}_j$ consider $\rho := \sum_{j=1}^{d} r_j|w_j\rangle\langle w_j| \in K_\sigma$. Inserting ρ into Q_σ gives:

$$Q_\sigma(\rho) = \frac{D(\sigma\|\rho)}{D(\rho\|\sigma)} = \frac{-S(\sigma) - \operatorname{tr}[\sigma\log(\rho)]}{-S(\rho) - \operatorname{tr}[\rho\log(\sigma)]} = \frac{-S(\sigma) - \langle s, P\log(r)\rangle}{-S(\rho) - \langle\log(s), Pr\rangle} =: F(P) \qquad (12)$$

where we introduced $P \in \mathcal{M}_d$ given by $P_{ij} = |\langle v_i|w_j\rangle|^2$ and $\log(s), \log(r) \in \mathbb{R}^d$ are defined as $(\log(s))_i = \log(s_i)$ and $(\log(r))_j = \log(r_j)$. Note that P is a unistochastic matrix, i.e. a doubly stochastic matrix whose entries are squares of absolute values of the entries of a unitary matrix. We will show that the minimum of F over unistochastic matrices P is attained at a permutation matrix. By definition of P this shows that there exists a state $\rho' \in K_\sigma$ with spectrum r and commuting with σ, which fulfills $Q_\sigma(\rho') \leq Q_\sigma(\rho)$.

As the set of unistochastic matrices is in general not convex [4], we want to consider the set of doubly stochastic matrices instead. By Birkhoff's theorem [5, Theorem II.2.3] we can write any doubly stochastic $D \in \mathcal{M}_d$ as $D = \sum_{i=1}^{k} \lambda_i P_i$ for some $k \in \mathbb{N}$, numbers $\lambda_i \in [0,1]$ with $\sum_{i=1}^{k} \lambda_i = 1$ and permutation matrices P_i. Now we can write the denominator of $F(D)$ as

$$-S(\rho) - \langle\log(s), Dr\rangle = \sum_{i=1}^{k} \lambda_i \left(-S(\rho) - \langle\log(s), P_i r\rangle\right) = \sum_{i=1}^{k} \lambda_i D(\rho_i\|\sigma) > 0,$$

where ρ_i is the state obtained by permuting the eigenvectors of ρ with P_i. In the last step we used Klein's inequality [6, p. 511] together with the fact that $\rho_i \neq \sigma$ for any $1 \leq i \leq k$ as their spectra are different. The previous estimate shows that F is also well-defined on doubly stochastic matrices.

Any unistochastic matrix is also doubly stochastic and we have

$$\inf\Big\{F(P) : P \in \mathcal{M}_d \text{ doubly stochastic}\Big\} \leq \inf\Big\{F(P) : P \in \mathcal{M}_d \text{ unistochastic}\Big\}.$$

Note that $S(\sigma)$ and $S(\rho)$ in (12) only depend on $s \in \mathbb{R}_+^d$ and $r \in \mathbb{R}_+^d$ and thus the numerator and the denominator of F are positive affine functions in P. This shows that F is a quasi-linear function [7, p. 91] on the set of doubly stochastic matrices. It can be shown (see [7]) that the minimum of such a function over a compact and convex set is always attained in an extremal point of the set. By Birkhoff's theorem [5, Theorem II.2.3] the extremal points of the compact and convex set of doubly stochastic matrices are the permutation matrices. As these are also unistochastic matrices we have

$$\inf\Big\{F(P) : P \in \mathcal{M}_d \text{ unistochastic}\Big\} = \inf\Big\{F(P) : P \in \mathcal{M}_d \text{ permutation matrix}\Big\}.$$

This finishes the first part.

To prove the lemma note that we have

$$\inf_{\rho \in K_\sigma} Q_\sigma(\rho) = Q_\sigma(\tilde{\rho})$$

for some minimizing full rank state $\tilde{\rho} \in \mathcal{D}_d^+$. Now consider some sequence $(\rho_n)_{n\in\mathbb{N}} \in K_\sigma^{\mathbb{N}}$ with $\rho_n \to \tilde{\rho}$ as $n \to \infty$ and such that the spectra of the ρ_n are no permutations of the spectrum of σ. By the first part of the proof we find a sequence $(\rho'_n)_{n\in\mathbb{N}} \in K_\sigma^{\mathbb{N}}$ commuting with σ, such that

$$Q_\sigma(\tilde{\rho}) \leq Q_\sigma(\rho'_n) \leq Q_\sigma(\rho_n) \to Q_\sigma(\tilde{\rho})$$

as $n \to \infty$.Thus $Q_\sigma(\rho'_n) \to Q_\sigma(\tilde{\rho})$ as $n \to \infty$. On the compact set K_σ the sequence $(\rho'_n)_n$ has a converging subsequence $\left(\rho'_{n_k}\right)_{k\in\mathbb{N}}$ with $\rho'_{n_k} \to \rho' \in K_\sigma$ as $k \to \infty$. By continuity of Q_σ we have $Q_\sigma(\rho') = Q_\sigma(\tilde{\rho}) = \inf_{\rho\in K_\sigma} Q_\sigma(\rho)$ and by continuity of the commutator $\rho \mapsto [\rho, \sigma]$ we have $[\rho', \sigma] = 0$.

□

With this lemma we can prove our main result:

Proof of Theorem 3.1. By Lemma 3.1 we may restrict the optimization in (11) to states which commute with σ. Thus, we can repeat the construction of the compact set K_σ (see 10) for a minimizer $\tilde{\rho} \in \mathcal{D}_d^+$ with $[\tilde{\rho}, \sigma] = 0$. By construction $\tilde{\rho}$ lies in the interior of K_σ, which will be important for the following argument involving Lagrange-multipliers.

To find necessary conditions on the minimizers of (11) we abbreviate $C := \inf_{\rho\in K_\sigma} Q_\sigma(\rho)$ and note that $C > 0$. To see this, note that we may extend $Q_\sigma(\rho)$ continuously to 1 at σ, so there exists $\delta > 0$ s.t. for $\|\rho - \sigma\|_1 \leq \delta$ we have $Q_\sigma(\rho) \geq \frac{1}{2}$ and for ρ s.t. $\|\rho - \sigma\|_1 > \delta$ we have $Q_\sigma(\rho) \geq \frac{\delta^2}{2\log(s_{\min}(\sigma^{-1}))}$ Using Pinsker's inequality and $D(\rho\|\sigma) \leq \log(s_{\min}(\sigma))$. For any $\rho \in K_\sigma$ with $[\rho, \sigma] = 0$ and $\rho \neq \sigma$ have

$$\frac{D(\sigma\|\rho)}{D(\rho\|\sigma)} \geq C$$

which is equivalent to

$$S(\sigma) \leq CS(\rho) + C\sum_{i=1}^{d} r_i \log(s_i) - \sum_{i=1}^{d} s_i \log(r_i). \tag{13}$$

Here $\{r_i\}_{i=1}^d$ denote the eigenvalues of $\rho \in K_\sigma$ (see 10) fulfilling $[\rho, \sigma] = 0$ and $\{s_i\}_{i=1}^d$ the eigenvalues of σ. As $\tilde{\rho}$ is a minimizer of (11) and commutes with σ its spectrum is a minimizer of the right-hand-side of (13) minimized over the set $\mathcal{S} := \{r \in \mathbb{R}^d : \min_i(r_i) \geq s_{\min}(\tilde{\rho}) - \epsilon\} \subset \mathbb{R}^d$ with ϵ chosen in the construction of K_σ (see (10)). We will now compute necessary conditions on the spectrum of $\tilde{\rho}$ using the formalism of Lagrange-multipliers (note that by construction the spectrum of $\tilde{\rho}$ lies in the interior of $\mathcal{S}$).

Consider the Lagrange function $F : \mathcal{S} \times \mathbb{R} \to \mathbb{R}$ given by

$$F(r_1, \dots, r_d, \lambda) = CS(\rho) + C\sum_{i=1}^{d} r_i \log(s_i) - \sum_{i=1}^{d} s_i \log(r_i) + \lambda\left(\sum_{i=1}^{d} r_i - 1\right).$$

The gradient of F is given by:

$$[\nabla F(r_1, \dots, r_d, \lambda)]_j = \begin{cases} C(-\log(r_j) - 1 + \log(s_j)) - \frac{s_j}{r_j} + \lambda & 1 \leq j \leq d \\ \sum_{i=1}^{d} r_i - 1 & j = d+1 \end{cases} \tag{14}$$

By the formalism of Lagrange-multipliers any minimizer $r = (r_1, \dots, r_d)$ of the right-hand-side of (13) in the interior of $\mathcal{S}$ has to fulfill $\nabla F(r_1, \dots, r_d, \lambda) = 0$ for some $\lambda \in \mathbb{R}$. Summing up the first d of these equations (where the jth equation is multiplied with r_j) implies

$$\lambda = 1 + C(1 + D(\rho\|\sigma)).$$

Inserting this back into the equations $[\nabla F(r_1, \dots, r_d, \lambda)]_j = 0$ and using $u_j = \frac{r_j}{s_j}$ we obtain

$$u_j(1 + CD(\rho\|\sigma)) - 1 = Cu_j \log(u_j) \tag{15}$$

for $1 \leq j \leq d$. For fixed $D(\rho\|\sigma)$ there are only two values for u_j solving the equations (15), as an affine functions (the left-hand-side) can only intersect a strictly convex function (the right-hand-side) in at most two points. Thus, for a minimizer $\{r_i\}_{i=1}^d$ of the right-hand-side of (13) in the interior of $\mathcal{S}$ there are constants $c_1, c_2 \in \mathbb{R}^+$ such that for each $i \in \{1, \dots, d\}$ either $r_i = c_1 s_i$ or $r_i = c_2 s_i$ holds.

We have obtained the following conditions on the spectrum of the minimizer $\tilde{\rho} \in K_\sigma$ (fulfilling $[\tilde{\rho}, \sigma] = 0$) of (11): There exist constants $c_1, c_2 \in \mathbb{R}^+$ a permutation $\nu \in S_d$ (where S_d denotes the group of permutations on $\{1, \dots, d\}$) and some $0 \leq n \leq d$ such that the spectrum $r \in \mathbb{R}_d^+$ of $\tilde{\rho}$ fulfills $r_i = c_1 s_i$ for any $0 \leq i \leq n$ and $r_i = c_2 s_i$ for any $n+1 \leq i \leq d$. Note that the cases $c_1 = c_2 = 1$, $n = 0$ and $n = d$ all correspond to the case $\rho = \sigma$ where we have $Q_\sigma(\sigma) = 1$. Thus, we can exclude the cases $n = 0$ and $n = d$ as long as we optimize over $c_1 = c_2 = 1$. Furthermore, note that we can use the normalization of $\tilde{\rho}$, i.e. $c_1 \sum_{i=1}^n s_i + c_2 \sum_{i=n+1}^d s_i = 1$ to eliminate c_2. Given a permutation $\nu \in S_d$ and $n \in \{1, \dots, d\}$, we define $p(\nu, n) = \sum_{i=1}^n s_{\nu(i)}$. Inserting the above conditions into (11) and setting $c_1 = x$ and $0 < n < d$ yields

$$\inf_{\rho \in K_\sigma} Q_\sigma(\rho) = \inf_{\nu \in S_d} \inf_{1 \leq n < d} \inf_{x \in [0, p(\nu,n)^{-1}]} q_{p(\nu,n)}\left(xp(\nu, n)\right) \tag{16}$$

$$= \inf_{\nu \in S_d} \inf_{1 \leq n < d} \inf_{x \in [0,1]} q_{p(\nu,n)}(x) \tag{17}$$

where $q_y : [0,1] \to \mathbb{R}$ denotes the continuous extension of $x \mapsto \frac{D_2(y\|x)}{D_2(x\|y)}$ (see (8)). By Lemma A.1 in the appendix the function $y \mapsto q_y(x)$ is continuous and quasi-concave and hence the minimum over any convex and compact set is attained at the boundary. Thus, we have

$$q_{s_{\min}(\sigma)}(x) \geq \inf_{\nu \in S_d} \inf_{1 \leq n \leq d} q_{p(\nu,n)}(x) \geq \inf_{y \in [s_{\min}(\sigma), 1 - s_{\min}(\sigma)]} q_y(x) = q_{s_{\min}(\sigma)}(x)$$

using $q_{1-s_{\min}(\sigma)}(x) = q_{s_{\min}(\sigma)}(x)$ for any $x \in [0,1]$. Inserting this into (17) leads to

$$\inf_{\rho \in K_\sigma} Q_\sigma(\rho) = \inf_{x\in[0,1]} q_{s_{\min}(\sigma)}(x).$$

□

Lemma 3.1 implies that the log-Sobolev-1 constant of the depolarizing channels coincides with the classical one of the random walk on the complete graph with d vertices and distribution given by the spectrum of σ. This constant has been shown to imply other inequalities, such as in [8, Proposition 3.13]. Using this result, Theorem 3.1 implies a refined transportation inequality on graphs.

Using the correspondence with the classical log-Sobolev-1 constant of a random walk on the complete graph, we may apply [9, Example 3.10], which proves:

Corollary 3.1. *Let $\mathcal{L}_\sigma : \mathcal{M}_d \to \mathcal{M}_d$ be the depolarizing Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d$ as defined in* (2)*. Then we have*

$$\alpha_1(\mathcal{L}_\sigma) \geq \frac{1}{2} + \sqrt{s_{\min}(\sigma)\,(1 - s_{\min}(\sigma))}$$

with equality iff $s_{\min}(\sigma) = \frac{1}{2}$. Again $s_{min}(\sigma)$ denotes the minimal eigenvalue of σ.

4. Application: Improved concavity of von-Neumann entropy

It is a well-known fact that the von-Neumann entropy $S(\rho) = -\mathrm{tr}\,[\rho\log(\rho)]$ is concave in ρ. Using Theorem 3.1 we can improve the concavity inequality:

Theorem 4.1 (Improved concavity of the von-Neumann entropy)**.** *For $\rho, \sigma \in \mathcal{D}_d$ and $q \in [0,1]$ we have*

$$S((1-q)\sigma + q\rho) - (1-q)S(\sigma) - qS(\rho) \geq \max\begin{cases} q(1-q^{c(\sigma)})D(\rho\|\sigma) \\ (1-q)(1-(1-q)^{c(\rho)})D(\sigma\|\rho) \end{cases},$$

with

$$c(\sigma) = \min_{x\in[0,1]} \frac{D_2(s_{\min}(\sigma)\|x)}{D_2(x\|s_{\min}(\sigma))}$$

and $c(\rho)$ defined in the same way.

Note that this bound becomes trivial if both σ and ρ are not of full rank (as we have $c(\rho) = c(\sigma) = 0$ in this case). However, as long as $D(\rho\|\sigma)$ or $D(\sigma\|\rho) < \infty$, we may still get a bound by restricting both density matrices to the support of σ or ρ, respectively.

Proof. Note that for the Liouvillian $\mathcal{L} := -\log(q)\mathcal{L}_\sigma$ we have:

$$e^{\mathcal{L}}(\rho) = q\rho + (1-q)\sigma.$$

By Theorem 3.1 and (3) we have

$$D\left(e^{\mathcal{L}}(\rho)\|\sigma\right) \leq e^{(1+c(\sigma))\log(q)} D\left(\rho\|\sigma\right) \tag{18}$$

Rearranging and expanding the terms in (18) we get

$$\begin{aligned} S(q\rho + (1-q)\sigma) &\geq (1-q)S(\sigma) - q\mathrm{tr}\left[\rho\log(\sigma)\right] + q^{1+c(\sigma)}D(\rho\|\sigma) \\ &= (1-q)S(\sigma) + qS(\rho) + q(1-q^{c(\sigma)})D(\rho\|\sigma). \end{aligned}$$

Interchanging the roles of ρ and σ in the above proof gives the second case under the maximum.

□

In [10] another improvement on the concavity of the von-Neumann entropy is shown:

$$S((1-q)\sigma + q\rho) - (1-q)S(\sigma) - qS(\rho) \geq \frac{q(1-q)}{(1-2q)^2}\max\begin{cases} D(\rho_{\text{avg}}\|\rho_{\text{rev}}) \\ D(\rho_{\text{rev}}\|\rho_{\text{avg}}) \end{cases} \tag{19}$$

$$\geq \frac{1}{2}q(1-q)\|\rho-\sigma\|_1^2 \tag{20}$$

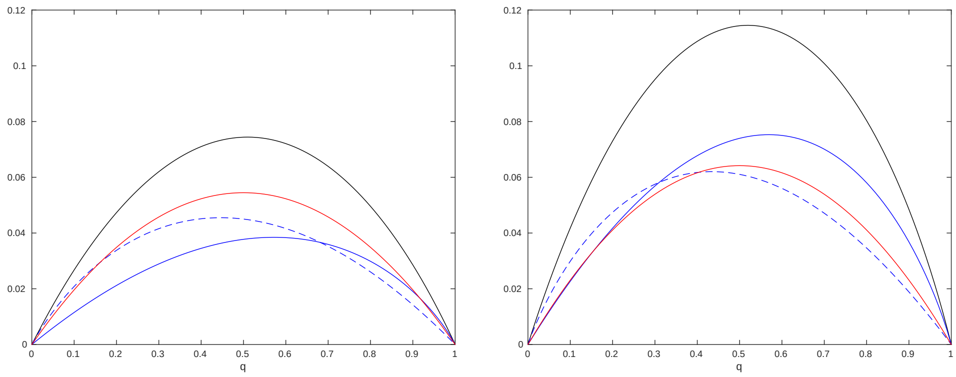
where $\rho_{\text{avg}} = (1-q)\sigma + q\rho$ and $\rho_{\text{rev}} = (1-q)\rho + q\sigma$. Note that this bound is valid for all states $\rho, \sigma \in \mathcal{D}_d$ while our bound in Theorem 4.1 becomes trivial unless the support ρ is contained in the support of σ or the other way around. We will therefore consider only full rank states in the following analysis.

By simple numerical experiments our bound from Theorem 4.1 seems to be worse than (19). However, one can argue that (19) is not much simpler than the left-hand-side itself. In particular the dependence on ρ and σ is only implicit via the relative entropy between ρ_{avg} and ρ_{rev}. Our bound from Theorem 4.1 depends on some spectral data (in terms of the smallest eigenvalues of ρ or σ), but whenever this is given, we have a bound for any $q \in [0,1]$ in terms of the relative entropies of ρ and σ.

Again we can do simple numerical experiments to compare the bounds (20) and Theorem 4.1. Recall that our bound is given in terms of the relative entropy and (20) in terms of the trace norm. In Figure 2 the bounds are compared for randomly generated quantum states in dimension $d = 10$. These plots show that the bounds are *not* comparable and depending of the choice of the states the bound from Theorem 4.1 will perform better than (20) or vice versa. Note that for q close to 0 or 1 our bound seems to perform better in both Figures. This is to be expected as $\alpha_1(\mathcal{L}_\sigma)$ is defined as the optimal constant α bounding the entropy production (4) (in $t = 0$) by $-2\alpha D(\rho\|\sigma)$. Therefore, Theorem 4.1 should be the optimal bound (in terms of relative entropy) for q near 0 or 1.

Note that by applying Pinsker's inequality:

$$D\left(\rho\|\sigma\right) \geq \frac{1}{2}\|\rho-\sigma\|_1^2 \tag{21}$$



(a) (b)

Figure 2: Comparison of bound (20) (red) and the bound from Theorem 4.1 (blue, where both choices of the ordering of ρ and σ are plotted) and the exact value $S((1-q)\sigma + q\rho) - (1-q)S(\sigma) - qS(\rho)$ (black) two pairs of randomly generated 10×10 quantum states and $q \in [0,1]$.

for states $\rho, \sigma \in \mathcal{D}_d$ to our bound from Theorem 4.1 we can obtain an improvement on the concavity inequality in terms of the trace-distance similar to (20). Unfortunately a simple computation shows that the resulting trace-norm bound is always worse than (20). In the next section we will show that Pinsker's inequality can be improved in the case where the second argument in the relative entropy is fixed (which is the case in the bound from Theorem 4.1). This will lead to an additional improvement of the trace-norm bound obtained from Theorem 4.1, such that in some (but only very few) cases the bound becomes better than (20).

5. State-Dependent Optimal Pinsker's Inequality

Pinsker's inequality (21) can be applied to the bound in Theorem 4.1 to get an improvement of the concavity in terms of the trace distance of the two density matrices. It can also be applied to (3) to get a mixing time bound [2] for the depolarizing channel. Note that in both of these cases the second argument of the relative entropy is fixed. Other improvements have been considered in the literature [11], but here we will improve Pinsker's inequality in terms of the second argument of the relative entropy. More specifically we compute the optimal constant $C(\sigma)$ (depending on σ) such that $D(\rho\|\sigma) \geq C(\sigma)\|\rho - \sigma\|_1^2$ holds when σ has full rank.

We will follow a strategy similar to the one pursued in [12] in proving this, where the analogous problem was considered for classical probability distributions. For a state $\rho \in \mathcal{D}_d$ let $s(\rho) = (s_1(\rho), \ldots, s_d(\rho))$ denote its vector of eigenvalues decreasingly ordered.

Lemma 5.1. *Let $\sigma \in \mathcal{D}_d^+$ and for $A \subseteq \{1,\ldots,d\}$ define $P_\sigma(A) = \sum_{i\in A} s_i(\sigma)$. Then we have for $\epsilon > 0$:*

$$\min_{\rho:\|\rho-\sigma\|_1\geq\epsilon} D(\rho\|\sigma) = \min_{A\subseteq\{1,\ldots,d\}} D_2(P_\sigma(A)+\epsilon\|P_\sigma(A))$$

Proof. Let $\rho \in \mathcal{D}_d$ be such that $\|\rho-\sigma\|_1 = \delta$, with $\delta \geq \epsilon$. By Lidskii's theorem [5, Corollary III.4.2], we have:

$$s(\rho-\sigma) = s(\sigma) - Ls(\rho),$$

where L is a doubly stochastic matrix. Define ρ' to be the state which has eigenvalues $Ls(\rho)$ and commutes with σ. Then we have:

$$\|\rho-\sigma\|_1 = \|\rho'-\sigma\|_1$$

By the operational interpretation for the $1-$norm [6, Theorem 9.1] there exist hermitian projections $Q, Q' \in \mathcal{M}_n$ such that

$$2\mathrm{tr}[Q(\rho-\sigma)] = \|\rho-\sigma\|_1 = \|\rho'-\sigma\|_1 = 2\mathrm{tr}[Q'(\rho'-\sigma)]. \tag{22}$$

Now define the quantum channel $T : \mathcal{M}_d \to \mathcal{M}_2$ given by:

$$T(\rho) = \mathrm{tr}[Q\rho]|0\rangle\langle 0| + \mathrm{tr}[(\mathbb{1}-Q)\rho]|1\rangle\langle 1|.$$

where $|0\rangle, |1\rangle$ is an orthonormal basis of $\mathbb{C}^2$. By the data processing inequality we have:

$$D(\rho\|\sigma) \geq D(T(\rho)\|T(\sigma)) \tag{23}$$

It is easy to see that the image of Q' must be spanned by eigenvectors of σ. Thus, we may associate a subset $A \subseteq \{1,\ldots,d\}$ to the projector Q' indicating the eigenvectors of σ spanning this subspace. Using (22) and the assumption that $\|\rho-\sigma\|_1 = \delta$ we have:

$$\mathrm{tr}[Q'\rho'] = P_\sigma(A) + \frac{\delta}{2}.$$

Also observe that

$$D(T(\rho)\|T(\sigma)) = D_2\left(P_\sigma(A) + \frac{\delta}{2}\middle\|P_\sigma(A)\right) \geq D_2\left(P_\sigma(A) + \frac{\epsilon}{2}\middle\|P_\sigma(A)\right)$$

as the binary relative entropy is convex and $\delta \geq \epsilon$ was assumed. With (23) we have:

$$\min_{\rho:\|\rho-\sigma\|_1\geq\epsilon} D(\rho\|\sigma) \geq \min_{A\subseteq\{1,\ldots,d\}} D_2\left(P_\sigma(A) + \frac{\epsilon}{2}\middle\|P_\sigma(A)\right) \tag{24}$$

Now given any $A \subseteq \{1,\ldots,d\}$ such that $P_\sigma(A)+\frac{\epsilon}{2} < 1$ (otherwise $D_2(P_\sigma(A)+\epsilon\|P_\sigma(A)) = +\infty$) , define a state $\tau \in \mathcal{D}_d$ which commutes with σ and has spectrum:

$$s_i(\tau) = \begin{cases} \frac{(P_\sigma(A)+\epsilon/2)s_i(\sigma)}{P_\sigma(A)} & \text{for } i \in A \\ \frac{(1-P_\sigma(A)-\epsilon/2)s_i(\sigma)}{1-P_\sigma(A)} & \text{else.} \end{cases}$$

Note that $\|\sigma - \tau\|_1 = \epsilon$ and $D(\tau\|\sigma) = D_2\left(P_\sigma(A) + \frac{\epsilon}{2}\|P_\sigma(A)\right)$, i.e. the lower bound in (24) is attained.

$\square$

We define the function $\phi : [0, \frac{1}{2}] \to \mathbb{R}$ as

$$\phi(p) = \frac{1}{1-2p}\log\left(\frac{1-p}{p}\right) \tag{25}$$

extended continuously by $\phi\left(\frac{1}{2}\right) = 2$. Furthermore for any $\sigma \in \mathcal{D}_d$ we define

$$\pi(\sigma) = \max_{A\subseteq\{1,\ldots,d\}} \min\left\{\frac{1}{2}, \sum_{i\in A} s_i(\sigma)\right\}. \tag{26}$$

With essentially the same proof as given in [12] for the classical case we obtain the following improvement on Pinsker's inequality:

Theorem 5.1 (State-dependent Pinsker's Inequality)**.** *For $\sigma, \rho \in \mathcal{D}_d$ we have:*

$$D(\rho\|\sigma) \geq \frac{\phi(\pi(\sigma))}{4}\|\rho - \sigma\|_1^2 \tag{27}$$

with ϕ as in (25) and $\pi(\sigma)$ as in (26). Moreover, this inequality is tight.

Proof. For convenience set $\|\rho - \sigma\|_1 = \delta$. Then we have

$$D(\rho\|\sigma) \geq \min_{\rho' : \|\rho' - \sigma\|_1 \geq \delta} D(\rho'\|\sigma) = \min_{A\subseteq\{1,\ldots,d\}} D_2\left(P_\sigma(A) + \frac{\delta}{2}\|P_\sigma(A)\right) \tag{28}$$

using Theorem 5.1. By [12, Proposition 2.2] for $p \in [0, \frac{1}{2}]$ and $\epsilon \geq 0$ we have

$$D_2(p + \epsilon\|p) \leq D_2(1 - p + \epsilon\|1 - p)$$

so we may assume $P_\sigma(A) \leq \frac{1}{2}$ in (28). In [13, Theorem 1] it is shown that for $p \in [0, \frac{1}{2}]$ we have

$$\inf_{\epsilon\in(0,1-p]} \frac{D_2(p + \epsilon\|p)}{\epsilon^2} = \phi(p) \tag{29}$$

which implies:

$$\min_{A\subseteq\{1,\ldots,d\}} D_2\left(P_\sigma(A) + \frac{\delta}{2}\middle\|P_\sigma(A)\right) \geq \min_{A\subseteq\{1,\ldots,d\}} \frac{\phi(P_\sigma(A))}{4}\|\rho - \sigma\|_1^2.$$

By [12, Proposition 2.4] the function ϕ is strictly decreasing. Thus, we have

$$\min_{A\subseteq\{1,\ldots,d\}} \frac{\phi(P_\sigma(A))}{4} = \frac{\phi(\pi(\sigma))}{4}$$

which, after combining the previous inequalities, finishes the proof of (27). To show that the inequality is tight, we may again follow the proof of [12, Proposition 2.1]. Let $B \subseteq \{1, \ldots, d\}$ be the subset such that $\pi(\sigma) = P_\sigma(B) =: p$. Define a minimizing sequence $\{\epsilon_i\}_{i \in \mathbb{N}}$ with $\epsilon_i > 0$ for the infimum (with respect to p) in (29), i.e. such that

$$\lim_{i \to \infty} \frac{D_2(p + \epsilon_i \| p)}{\epsilon_i^2} = \phi(p).$$

Next define a sequence of states ρ_i that commute with σ and have spectrum:

$$s_j(\rho_i) = \begin{cases} \frac{(p+\epsilon_i) s_i(\sigma)}{p} & \text{for } j \in B \\ \frac{(1-p-\epsilon_i) s_i(\sigma)}{1-p} & \text{else.} \end{cases}$$

One can check that $\|\rho_i - \sigma\|_1 = 2\epsilon_i$ and $D(\rho_i \| \sigma) = D_2(p + \epsilon_i \| p)$, from which we get:

$$\lim_{i \to \infty} \frac{D(\rho_i \| \sigma)}{\|\rho_i - \sigma\|^2} = \frac{\phi(\pi(\sigma))}{4}$$

□

In some cases the bound can be made more explicit, as illustrated in the next corollary:

Corollary 5.1. *Let $\sigma, \rho \in \mathcal{D}_d$ be such that $\|\sigma\|_\infty \geq \frac{1}{2}$. Then:*

$$D(\rho \| \sigma) \geq \frac{\phi(1 - \|\sigma\|_\infty)}{4} \|\rho - \sigma\|_1^2 \tag{30}$$

Proof. In this case it is clear that $\pi(\sigma) = 1 - \|\sigma\|_\infty$. □

Note that we have $\phi(x) \to +\infty$ for $x \to 0$. Thus, there might be an arbitrary large improvement of (27) compared to the usual Pinsker's inequality (21). This happens for instance in Corollary 5.1 when $\|\sigma\|_\infty \to 1$, i.e. when σ converges to a pure state.

By applying the improved inequality (27) to Theorem 4.1 we obtain for quantum states $\rho, \sigma \in \mathcal{D}_d$ and $q \in [0, 1]$

$$S((1-q)\sigma + q\rho) - (1-q)S(\sigma) - qS(\rho) \geq \max \begin{cases} q(1 - q^{c(\sigma)}) \frac{\phi(\pi(\sigma))}{4} ||\rho - \sigma||^2 \\ (1-q)(1 - (1-q)^{c(\rho)}) \frac{\phi(\pi(\rho))}{4} ||\rho - \sigma||^2 \end{cases}$$

with ϕ as in (25) and $\pi(\sigma)$ as in (26).

Even using this refinement of Pinsker's inequality, some numerical experiments indicate that (20) is stronger for randomly generated states. From Corollary 5.1 we can expect our bound to perform well if σ has a large eigenvalue and the smallest eigenvalue is as large as possible. Such states have spectrum of the form $\left(p, \frac{1-p}{d-1}, \ldots, \frac{1-p}{d-1}\right)$. Indeed for $\sigma \in \mathcal{D}_5$ with spectrum $(0.99, 0.0025, 0.0025, 0.0025, 0.0025)$ and $q < 0.2$ our bound performs better than (19) for randomly generated ρ. However, even in this case the improvement is *not* significant.

Still we can expect that Theorem 5.1 will find more applications, for instance improving the mixing time bounds. Such bounds have been derived from log-Sobolev inequalities in [2]. The next theorem can be used to improve these results:

Theorem 5.2. *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive Liouvillian with fixed point σ that satisfies*

$$D\left(e^{t\mathcal{L}}\rho\|\sigma\right) \leq e^{-2\alpha t} D\left(\rho\|\sigma\right) \tag{31}$$

for some $\alpha > 0$ and for all $\rho \in \mathcal{D}_d$ and $t \in \mathbb{R}^+$. Then we have

$$\|e^{t\mathcal{L}}(\rho) - \sigma\|_1 \leq 2e^{-\alpha t}\sqrt{\frac{\log\left(s_{\min}(\sigma)\right)}{\phi\left(\pi\left(\sigma\right)\right)}} \tag{32}$$

with ϕ as in (25)*, $\pi(\sigma)$ as in* (26) *and where $s_{\min}(\sigma)$ is the smallest eigenvalue of σ.*

Proof. This is a direct consequence of (3) and (27). □

6. Tensor products of depolarizing channels

For a Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ generating the channel $\mathcal{T}_t = e^{t\mathcal{L}}$ and any $n \in \mathbb{N}$ we denote by $\mathcal{L}^{(n)} : \mathcal{M}_{d^n} \to \mathcal{M}_{d^n}$ the generator of the tensor-product semigroup $(T_t)^{\otimes n}$, i.e. $\mathcal{L}^{(n)} := \sum_{i=1}^{n} \mathrm{id}_d^{\otimes i-1} \otimes \mathcal{L} \otimes \mathrm{id}_d^{\otimes(n-i)}$.

Here we study $\alpha_1\left(\mathcal{L}_\sigma^n\right)$ in the special case where $\sigma = \frac{\mathbb{1}_d}{d}$. For simplicity we denote the depolarizing Liouvillian onto $\sigma = \frac{\mathbb{1}_d}{d}$ by $\mathcal{L}_d := \mathcal{L}_{\frac{\mathbb{1}_d}{d}}$ and by $T_t^d = e^{t\mathcal{L}^d}$ the generated semigroup. In the case $d = 2$ it is known [2] that $\alpha_1\left(\mathcal{L}_2^{(n)}\right) = 1$ for any $n \in \mathbb{N}$. It is, however, an open problem to determine this constant for any $d > 2$ and any $n \geq 2$. We will now show the inequality $\alpha_1\left(\mathcal{L}_d^{(n)}\right) \geq \frac{1}{2}$ for any $d \geq 2$ and $n \geq 1$, which is the best possible lower bound that is independent of the local dimension. Note that for $\sigma = \frac{\mathbb{1}_d}{d}$ inequality (3) for the channel $\left(\mathcal{T}_t^d\right)^{\otimes n}$ can be rewritten as the entropy production inequality:

$$S((T_t^d)^{\otimes n}(\rho)) \geq (1 - e^{-t})n\log(d) + e^{-t}S(\rho).$$

This inequality has been studied in [14] for the case where $d = 2$, for wich, however, an incorrect proof was given. We will provide a proof of a more general statement, from which the claim $\alpha_1\left(\mathcal{L}_{\mathrm{dep}}^{(n)}\right) \geq \frac{1}{2}$ readily follows by the previous discussion.

Theorem 6.1. *For any $\sigma, \rho \in \mathcal{D}_d$ (not necessarily full rank) we have*

$$S((T_t^\sigma)^{\otimes n}(\rho)) \geq e^{-t}S(\rho) + (1 - e^{-t})S(\sigma^{\otimes n}).$$

For the proof we will need a special case of the quantum Shearer's inequality. We will denote by $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n}\right)$ a multipartite density matrix (where the d_i are the local dimensions of each tensor factor). Furthermore we write $S(i_1, i_2, \ldots, i_k)_\rho$

for the entropy of the reduced density matrix ρ on the tensor factors specified by the indices $i_1, i_2, \dots, i_k$. Similarly we write

$$S(i_1, \dots, i_k | j_1, \dots, j_l)_\rho = S(i_1, \dots, i_k, j_1, \dots, j_l)_\rho - S(j_1, \dots, j_l)_\rho$$

for a conditional entropy. The proof of the quantum version of Shearer's inequality is essentially the same as the proof given by Radhakrishnan and Llewellyn for the classical version (see [15]). For convenience we provide the full proof:

Lemma 6.1 (Quantum Shearer's inequality)**.** *Consider $t \in \mathbb{N}$ and a family $\mathcal{F} \subset 2^{\{1,\dots,n\}}$ of subsets of $\{1, \dots, n\}$ such that each $i \in \{1, \dots, n\}$ is included in exactly t elements of $\mathcal{F}$. Then for any $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n}\right)$ we have*

$$S(1, 2 \dots, n)_\rho \leq \frac{1}{t} \sum_{F \in \mathcal{F}} S(F)_\rho \ . \tag{33}$$

Proof. For $F \subset \{1, \dots, n\}$ denote its elements by $(i_1, \dots, i_k)$, increasingly ordered. For any $\rho \in \mathcal{D}\left(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n}\right)$ we have

$$\begin{aligned}\sum_{j=1}^{|F|} S(i_j | i_1, \dots, i_{j-1})_\rho &= S(i_1)_\rho + S(i_2 | i_1)_\rho + \cdots + S(i_{|F|} | i_1, i_2, \dots, i_{|F|-1})_\rho \\ &= S(i_1, i_2, \dots, i_{|F|})_\rho = S(F)_\rho\end{aligned}$$

where we used a telescopic sum trick. By strong subadditivity [16] conditioning decreases the entropy. This implies

$$\sum_{j=1}^{|F|} S(i_j | 1, 2, \dots, i_j - 1)_\rho \leq \sum_{j=1}^{|F|} S(i_j | i_1, \dots, i_{j-1})_\rho = S(F)_\rho. \tag{34}$$

Now consider a family $\mathcal{F} \subset 2^{\{1,\dots,n\}}$ with the properties stated in the assumptions. Using (34) for the first inequality gives:

$$\sum_{F \in \mathcal{F}} S(F)_\rho \geq \sum_{F \in \mathcal{F}} \sum_{j=1}^{|F|} S(i_j | 1, 2, \dots, i_j - 1)_\rho \tag{35}$$

$$= t \sum_{i=1}^{n} S(i | 1, 2, \dots, i - 1)_\rho = tS(1, 2, \dots, n)_\rho. \tag{36}$$

Here we used the assumption that each $i \in \{1, \dots, n\}$ is contained in exactly t elements of $\mathcal{F}$ and (34) in the special case of $F = \{1, \dots, n\}$ for the final equality. □

Note that in the classical case Shearer's inequality is true under the weaker assumption that any $i \in \{1, \dots, d\}$ is contained in *at least* t elements of $\mathcal{F}$. However, as the quantum conditional entropy might be negative [17] we have to use the stronger assumption to get the equality between (35) and (36) where an $\geq$ would be enough.

In the special case where $\mathcal{F} = \mathcal{F}_k := \{F \subseteq \{1,\ldots,n\} : |F| = k\}$ denotes the family of k-element subsets of $\{1,\ldots,n\}$ (i.e. every $i \in \{1,\ldots,d\}$ is contained in exactly $\binom{n-1}{k-1} = \frac{k}{n}\binom{n}{k}$ elements of $\mathcal{F}_k$) the quantum Shearer inequality gives

$$\frac{k}{n} S(1,\ldots,n) \leq \frac{1}{\binom{n}{k}} \sum_{F \in \mathcal{F}_k} S(F). \tag{37}$$

This inequality was also proved in [18], but in a more complicated way and without mentioning the more general quantum Shearer's inequality. It is also used as a lemma (with wrong proof) in [14], where the rest of the proof of their entropy production estimate is correct. The proof of Theorem 6.1 follows the same lines. For completeness we will include the full proof here:

Proof of Theorem 6.1. In the following we will abbreviate $p := e^{-t}$. For a subset $F \subset \{1,\ldots,n\}$ we denote by $\rho|_F$ the reduced density matrix on the tensor factors specified by F. Using this notation we can write

$$(T_t^\sigma)^{\otimes n}(\rho) = \sum_{k=0}^{n} \sum_{F \in \mathcal{F}_k} (1-p)^k p^{n-k} \left(\bigotimes_{l \in F} \sigma \otimes \rho|_{F^c} \right)$$

where $F^c = \{1,\ldots,n\} \setminus F$. Concavity of the von-Neumann entropy implies

$$\begin{aligned} S\left((T_t^\sigma)^{\otimes n}(\rho)\right) &\geq \sum_{k=0}^{n} \sum_{F \in \mathcal{F}_k} (1-p)^k p^{n-k} \left(kS(\sigma) + S(F^c)_\rho\right) \\ &\geq (1-p)nS(\sigma) + \sum_{k=0}^{n} \binom{n}{n-k} \frac{n-k}{n} (1-p)^k p^{n-k} S(\rho) \\ &= (1-p)S(\sigma^{\otimes n}) + pS(\rho). \end{aligned}$$

Here we used the elementary identity $\sum_{k=0}^{n} \binom{n}{k}(1-p)^k p^{n-k} k = (1-p)n$ and (37) for the $(n-k)$-element subsets F^c. $\square$

Acknowledgments

We thank Ion Nechita for pointing out Shearer's inequality to us and David Reeb for interesting and helpful discussions. D.S.F acknowledges support from the graduate program TopMath of the Elite Network of Bavaria, the TopMath Graduate Center of TUM Graduate School at Technische Universität München and the Deutscher Akademischer Austauschdienst(DAAD). We also acknowledge financial support from the CHIST-ERA/BMBF project CQC (A.M.H. and M.M.W.).

A. Quasi-concavity of a quotient of relative entropies

In this appendix we will prove the quasi-concavity of the function $y \mapsto q_y(x)$ for any $x \in (0,1)$. As defined in (8) the function $q_y : (0,1) \to \mathbb{R}$ denotes the continuous extension of $x \mapsto \frac{D_2(y\|x)}{D_2(x\|y)}$. In the following we consider $f_x : [0,1] \to \mathbb{R}$ defined as $f_x(y) = q_y(x)$ for any $y \in (0,1)$ and with $f_x(0) = f_x(1) = 1$. It can be checked easily that f_x is continuous for any $x \in (0,1)$. We have the following Lemma:

Lemma A.1. *For any $x \in (0,1)$ the function $f_x : [0,1] \to \mathbb{R}$ given by $f_x(y) = \frac{D_2(y\|x)}{D_2(x\|y)}$ for $y \notin \{0,x,1\}$ and extended continuously by $f_x(x) = 1$ and $f_x(0) = f_x(1) = 0$ is quasi-concave.*

Proof. Note that without loss of generality we can assume $x \geq \frac{1}{2}$, as $f_x(y) = f_{1-x}(1-y)$. By continuity it is clear that there exists an $m_f \in (0,1)$ (we can exclude the boundary points since $f_x(x) > f_x(0) = f_x(1)$) such that $f_x(m_f)$ is the global maximum. By [7][p. 99] it is sufficient to show that f_x is unimodal, i.e. that f_x is monotonically increasing on $[0, m_f)$ and monotonically decreasing on $(m_f, 1]$. We will use the method of L'Hospital type rules for monotonicity developed in [19, 20].

For any $x \in (0,1)$ and $y \in (0,1)$ with $x \neq y$ we compute

$$\partial_y D_2\left(y\|x\right) = \log\left(\frac{y(1-x)}{x(1-y)}\right) \qquad \partial_y D_2\left(x\|y\right) = \frac{y-x}{y(1-y)}$$
$$\partial_y \log\left(\frac{y(1-x)}{x(1-y)}\right) = \frac{1}{y(1-y)} \qquad \partial_y \frac{y-x}{y(1-y)} = \frac{y^2+x-2yx}{(1-y)^2y^2}$$

and define

$$g_x(y) = \frac{\partial_y D_2\left(y\|x\right)}{\partial_y D_2\left(x\|y\right)} = \frac{\log\left(\frac{x(1-y)}{y(1-x)}\right) y(1-y)}{x-y} \tag{38}$$
$$h_x(y) = \frac{\partial_y \log\left(\frac{x(1-y)}{y(1-x)}\right)}{\partial_y \frac{x-y}{y(1-y)}} = \frac{y(1-y)}{y^2+x-2yx} \tag{39}$$

where again g_x is extended continuously by $g_x(0) = g_x(1) = 0$ and $g_x(x) = 1$. As $y \mapsto y^2+x-2yx$ has no real zeros for $x \in (0,1)$ the rational function h_x is continuously differentiable on $(0,1)$. A straightforward calculation reveals that for $x \geq \frac{1}{2}$ and on $(0,1)$ the derivative h'_x only vanishes in

$$m_h = \begin{cases} \frac{x-\sqrt{x(1-x)}}{2x-1} & \text{for } x > \frac{1}{2} \\ \frac{1}{2} & \text{for } x = \frac{1}{2} \end{cases}.$$

which has to be a maximum as $h_x(0) = h_x(1) = 0$. By the lack of further points with vanishing derivative we have $h'_x(y) < 0$ for any $y < m_h$ and also $h'_x(y) > 0$ for any $y > m_h$. Note that $m_h \leq x$ for any $x \geq \frac{1}{2}$.

Consider first the interval $(x,1) \subset (0,1)$. For $y \to x$ we have $\log\left(\frac{x(1-y)}{y(1-x)}\right) \to 0$ and $\frac{x-y}{y(1-y)} \to 0$. Also it is clear that $y \mapsto \frac{x-y}{y(1-y)}$ does not change sign on the interval $(x,1)$. Therefore and by (39) we see that the pair g_x and h_x satisfy the assumptions of [19, Proposition 1.1.] and as h_x is decreasing we have that $g'_x(y) < 0$ for any $y \in (x,1)$.

We can use the same argument for the (possibly empty) interval (m_h, x) where h_x is decreasing as well and obtain $g_x'(y) < 0$ for any $y \in (m_h, x)$. By continuity of g_x in x we see that g_x is decreasing on $(m_h, 1)$.

Note that in the case where $x = \frac{1}{2}$ we can directly apply [19, Proposition 1.1.] to the remaining interval $(0, \frac{1}{2})$ where $h_{1/2}$ is increasing. This proves $g_{1/2}'(y) > 0$ for any $y \in (0, \frac{1}{2})$. By continuity $m_g = \frac{1}{2}$ is the maximum point of $g_{1/2}$. For $x \neq \frac{1}{2}$, where the remaining interval is $(0, m_h)$ we apply the more general Proposition 2.1. in [20]. It can be checked easily that the assumptions of this proposition are fulfilled for the pair g_x and h_x. As for $y \in (0, m_h)$ we have $\frac{y-x}{y(1-y)} \frac{y^2+x-2yx}{(1-y)^2y^2} < 0$ and as h_x is increasing the proposition shows that $g_x'(y) > 0$ for any $y \in (0, m_g)$ and $g_x'(y) < 0$ for any $y \in (m_g, m_h)$. Here $m_g \in (0, m_h)$ denotes the maximum point of g_x (note that a maximum m_g has to exist due to continuity and $g_x(0) = g_x(1) = 0$).

The previous argument shows that for any $x \geq \frac{1}{2}$ there exists a point $m_g \in (0, m_h] \subset (0, x]$ (we have $m_g = m_h = \frac{1}{2}$ for $x = \frac{1}{2}$) such that $g_x'(y) > 0$ for $y \in (0, m_g)$ and $g_x'(y) < 0$ for $y \in (m_g, 1) \setminus \{x\}$. We can now repeat the above argument for the pair f_x and g_x. This gives the existence of a point $m_f \in (0, m_g]$ such that $f_x'(y) > 0$ for any $y \in (0, m_f)$ and $f_x'(y) < 0$ for any $y \in (m_f, 1) \setminus \{x\}$. By continuity in x this shows that the function f_x is unimodal and therefore quasi-concave.

□

B. Continuous extension of a quotient of relative entropies

In this section we show that the function $Q_\sigma : \mathcal{D}_d^+ \to \mathbb{R}$ as defined in (6) is indeed continuous. As Q_σ is clearly continuous in any point $\rho \neq \sigma$ we have to prove the following:

Lemma B.1. *For $\sigma \in \mathcal{D}_d^+$ and $X \in \mathcal{M}_d$ with $X = X^\dagger$, $\mathrm{tr}[X] = 0$ and $X \neq 0$ we have*

$$\lim_{\epsilon \to 0} \frac{D(\sigma || \sigma + \epsilon X)}{D(\sigma + \epsilon X || \sigma)} = 1.$$

Proof. To show the claim we will expand the relative entropy in terms of ϵ up to second order. Observe that for $\rho \in \mathcal{D}_d$ we have

$$D(\rho \| \sigma) = \int_0^\infty \mathrm{tr}\left[\rho\left((\rho + t)^{-1} - (\sigma + t)^{-1}\right)\right] dt. \tag{40}$$

In the following we assume $\epsilon > 0$ to be small enough such that $\sigma + \epsilon X \in \mathcal{D}_d^+$. To simplify the notation, we introduce $A(t) := (\sigma + t)^{-1}$ and $B(t) := (\sigma + \epsilon X + t)^{-1}$. Applying the recursive relation

$$B(t) = -\epsilon B(t) X A(t) + A(t),$$

twice leads to

$$\begin{aligned}B(t)-A(t)&=-\epsilon B(t)XA(t)=\epsilon^2B(t)XA(t)XA(t)-\epsilon A(t)XA(t)\\&=\epsilon^2A(t)XA(t)XA(t)-\epsilon A(t)XA(t)+\mathcal{O}\left(\epsilon^3\right).\end{aligned}$$

Inserting this into (40) gives

$$D\left(\sigma\|\sigma+\epsilon X\right)=\int_0^\infty \operatorname{tr}\left[\epsilon\sigma A(t)XA(t)-\epsilon^2\sigma A(t)XA(t)XA(t)+\mathcal{O}(\epsilon^3)\right]dt \tag{41}$$

and

$$D\left(\sigma+\epsilon X\|\sigma\right)=\int_0^\infty \operatorname{tr}\left[-\epsilon\sigma A(t)XA(t)+\epsilon^2\sigma A(t)XA(t)XA(t)-\epsilon^2XA(t)XA(t)+\mathcal{O}(\epsilon^3)\right]dt. \tag{42}$$

As $[A(t),\sigma]=0$ we can diagonalize these operators in the same orthonormal basis $\{|i\rangle\}\subset\mathbb{C}^d$, which leads to

$$\int_0^\infty \operatorname{tr}\left[\sigma A(t)XA(t)\right]dt=\sum_{i=1}^d\langle i|X|i\rangle\int_0^\infty\frac{s_i}{(s_i+t)^2}dt=\sum_{i=1}^d\langle i|X|i\rangle=0 \tag{43}$$

where $\{s_i\}_{i=1}^d$ denotes the spectrum of σ. Note that again by diagonalizing σ and $A(t)$ in the same basis we have

$$\int_0^\infty \operatorname{tr}\left[\left(2\sigma A(t)-\mathbb{1}\right)\left(XA(t)XA(t)\right)\right]dt \tag{44}$$

$$\begin{aligned}&=\sum_{i,j=1}^d|\langle i|X|j\rangle|^2\int_0^\infty\frac{2s_i}{(s_i+t)^2(s_j+t)}-\frac{1}{(s_i+t)(s_j+t)}dt\\&=\sum_{i,j=1}^d\frac{|\langle i|X|j\rangle|^2}{(s_i-s_j)^2}\left(2(s_i-s_j)-(s_i+s_j)\log\left(\frac{s_i}{s_j}\right)\right)\qquad(45)\\&=0.\end{aligned}$$

The last equality follows from the fact that the expression in (45) clearly changes its sign when s_i and s_j are exchanged. This is only possible if the value of the integral (44) vanishes. Rearranging the integral (44) gives:

$$\int_0^\infty \operatorname{tr}\left[\sigma A(t)XA(t)XA(t)\right]dt=\int_0^\infty \operatorname{tr}\left[-\sigma A(t)XA(t)XA(t)+XA(t)XA(t)\right]dt. \tag{46}$$

Finally applying (43),(46) to the formulas for the relative entropies (41) and (42) gives:

$$\frac{D(\sigma\|\sigma+\epsilon X)}{D(\sigma+\epsilon X\|\sigma)}=\frac{c+\mathcal{O}(\epsilon)}{c+\mathcal{O}(\epsilon)}\to 1$$

as $\epsilon \to 0$. Here $c := \int_0^\infty \operatorname{tr}\left[\sigma A(t)XA(t)XA(t)\right]dt > 0$ as $\sigma, A(t) > 0$ for any $t \in [0,\infty)$ and $X \neq 0$ is Hermitian.

□

References

[1] R. Olkiewicz and B. Zegarlinski, “Hypercontractivity in noncommutative lpspaces,” *Journal of Functional Analysis*, vol. 161, no. 1, pp. 246 – 285, 1999.

[2] M. J. Kastoryano and K. Temme, “Quantum logarithmic Sobolev inequalities and rapid mixing,” *Journal of Mathematical Physics*, vol. 54, no. 5, p. 052202, May 2013.

[3] H. Spohn, “Entropy production for quantum dynamical semigroups,” *Journal of Mathematical Physics*, vol. 19, no. 5, pp. 1227–1230, 1978.

[4] I. Bengtsson, A. Ericsson, M. Kuś, W. Tadej, and K. Życzkowski, “Birkhoff’s Polytope and Unistochastic Matrices, N = 3 and N = 4,” *Communications in Mathematical Physics*, vol. 259, pp. 307–324, Oct. 2005.

[5] R. Bhatia, *Matrix analysis.* Springer, 1997, vol. 169.

[6] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, ser. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.

[7] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[8] M. D. Sammer, “Aspects of mass transportation in discrete concentration inequalities,” *PhD-thesis, Georgia Institute of Technology*, 2005.

[9] S. G. Bobkov and P. Tetali, “Modified logarithmic sobolev inequalities in discrete settings,” *Journal of Theoretical Probability*, vol. 19, no. 2, pp. 289–336, 2006.

[10] I. Kim and M. B. Ruskai, “Bounds on the concavity of quantum entropy,” *Journal of Mathematical Physics*, vol. 55, no. 9, 2014.

[11] K. M. Audenaert and J. Eisert, “Continuity bounds on the quantum relative entropy,” *Journal of mathematical physics*, vol. 46, no. 10, p. 102104, 2005.

[12] E. Ordentlich and M. Weinberger, "A distribution dependent refinement of pinsker's inequality," *Information Theory, IEEE Transactions on*, vol. 51, no. 5, pp. 1836–1840, May 2005.

[13] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American statistical association*, vol. 58, no. 301, pp. 13–30, 1963.

[14] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, "Limitations of Noisy Reversible Computation," *eprint arXiv:quant-ph/9611028*, Nov. 1996.

[15] J. Radhakrishnan, "Entropy and counting," *Chapter in: Computational mathematics, modelling and algorithms (Editor: J. C. Misra)*, pp. 146–168, 2003.

[16] E. H. Lieb and M. B. Ruskai, "Proof of the strong subadditivity of quantummechanical entropy," *Journal of Mathematical Physics*, vol. 14, no. 12, pp. 1938–1941, 1973.

[17] N. J. Cerf and C. Adami, "Negative entropy and information in quantum mechanics," *Phys. Rev. Lett.*, vol. 79, pp. 5194–5197, Dec 1997.

[18] M. Junge and C. Palazuelos, "Cb-norm estimates for maps between noncommutative l_p-spaces and quantum channel theory," *arXiv preprint arXiv:1407.7684*, 2014.

[19] I. Pinelis, "L'Hôspital type rules for monotonicity, with applications," *J. Ineq. Pure & Appl. Math*, vol. 3, no. 1, 2002.

[20] ——, "Non-strict L'Hôspital-type rules for monotonicity: Intervals of constancy," *J. Ineq. Pure & Appl. Math*, vol. 8, no. 1, pp. 1–8, 2007.

Spectral Variation Bounds in Hyperbolic Geometry

A. Müller-Hermes and O. Szehr May 18, 2015

The (Euclidean) optimal matching distance of two sets $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ is defined as

$$d_E(\{a_i\},\{b_i\}) = \min_{\sigma\in S_n}\max_{1\leq i\leq n}\left|a_i - b_{\sigma(i)}\right|$$

where S_n denotes the group of permutations of $\{1,\dots,n\}$. For two complex $n\times n$-matrices $A,B\in\mathcal{M}_n$ with spectra $\sigma(A),\sigma(B)\subset\mathbb{C}$ spectral variation bounds of the form

$$d_E(\sigma(A),\sigma(B)) \leq C_n\left(\|A\|+\|B\|\right)^{1-\frac{1}{n}}\|A-B\|^{\frac{1}{n}} \tag{1}$$

have been studied for C_n independent of A and B (see for instance [1] and the references therein). Currently the best known constant in the above bound is $C_n = \frac{16}{3\sqrt{3}}$ [2] and it is conjectured that $C_n = 2$ [1] is the optimal constant. In our work we apply recent bounds (proved in [3]) on resolvents to obtain spectral variation bounds with respect to a hyperbolic pseudometric. These lead to an improvement of (1) in certain cases.

1 Main result

For two sets $\{a_i\}_{i=1}^n,\{b_i\}_{i=1}^n\subset\overline{\mathbb{D}}$ define the hyperbolic optimal matching distance as

$$d_H(\{a_i\},\{b_i\}) = \min_{\sigma\in S_n}\max_{1\leq i\leq n}\left|\frac{a_i - b_{\sigma(i)}}{1-\overline{a}_i b_{\sigma(i)}}\right|.$$

With this we can prove:

Theorem 1.1. *For $A,B\in\mathcal{M}_n$ with $\|A\|,\|B\|<1$ let $|m|$ denote the degree of the minimal polynomial of A and let $\rho(B)\leq\|B\|$ denote the spectral radius of B. Then*

$$d_H(\sigma(A),\sigma(B)) \leq \frac{2^{2-\frac{1}{|m|}}}{\left(1-\rho(B)\|A\|\right)^{\frac{1}{|m|}}}\|A-B\|^{\frac{1}{|m|}}.$$

Using a Chebyshev-type interpolation theorem for finite Blaschke products we even prove a slightly stronger statement than the above theorem in terms of elliptic functions. However, as it is difficult to work with this stronger bound, we have to rely on the above statement to improve (1).

2 Improved Euclidean spectral variation bound

Note that a "hyperbolic disk" with radius $r \in \mathbb{R}^+$ and center $a \in \mathbb{D}$ coincides with a Euclidean disc with possibly different center and radius. More specifically we have

$$\{z \in \mathbb{D} \mid \left|\frac{a-z}{1-\bar{a}z}\right| \leq r\} = \{z \in \mathbb{D} \mid |C - z| < R\}$$

with center $C \in \mathbb{D}$ and radius $R \in [0, 1]$ given by

$$C = \frac{1-r^2}{1-r^2|a|^2}a \qquad \text{and} \qquad R = \frac{1-|a|^2}{1-r^2|a|^2}r$$

As $R \to 0$ for $|a| \to 1$ the hyperbolic bound is stronger for eigenvalues close to the boundary of $\mathbb{D}$, because then the corresponding discs are smaller. This effect of the hyperbolic geometry can be exploited by choosing a good scaling for the two matrices.

Corollary 2.1. *For $A, B \in \mathcal{M}_n$ with $M_2 := \max\{\|A\|, \|B\|\}$ and distance*

$$\|A - B\| \leq \left(\frac{1}{2M_2}\right)^{n-1} \left(\frac{n+1}{n-1}\right)^n \alpha_n^n \min_{a \in \sigma(A)\setminus\{0\}} |a|^n \tag{2}$$

we get

$$d_E(\sigma(A), \sigma(B)) \leq \frac{1}{\alpha_n} (2M_2)^{1-\frac{1}{n}} \|A - B\|^{\frac{1}{n}}$$

where

$$\alpha_n := \frac{1}{2}\left(\frac{2}{\sqrt{n^2-1}}\right)^{\frac{1}{n}} \sqrt{\frac{n-1}{n+1}}.$$

Note that the condition (2) is needed to get a uniform bound independent of the concrete eigenvalue pair under consideration. As $\alpha_n \to 2$ as $n \to \infty$ this bound converges to the conjectured optimal spectral variation bound (1).

3 Legal statement

The idea to use the techniques from [3] and the proof of Theorem 3 (in the paper) are due to Oleg Szehr. In all other parts of this work I was significantly involved.

References

[1] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 1997.

[2] Gerd M Krause. Bounds for the variation of matrix eigenvalues and polynomial roots. *Linear algebra and its applications*, 208:73–82, 1994.

[3] Nikolai Nikolski. Condition numbers of large matrices, and analytic capacities. *St. Petersburg Mathematical Journal*, 17(4):641–682, 2006.