Imperial College London
Department of Computing

De Finetti methods in Quantum Information

Francesco Borderi

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Computing.

Alla mia famiglia.

To my family.

Statement of Originality

I declare that the material presented in this thesis is my own work except where otherwise stated and referenced according to the established practices of the field.

Copyright Declaration

Acknowledgements

I want to thank the following people for their support and help during my PhD studies.

- *Mario Berta.* Mario is an exceptional researcher and a great supervisor. I am incredibly grateful for the patience he showed with me and the guidance I have received from him. I have learnt a lot from Mario. Thank you!

- *Omar Fawzi, and Volkher Scholz.* Omar and Volkher are extremely strong researchers, and I am lucky to have them as co-authors. I want to thank them for their comments and suggestions.

- *Dimitrios Myrisiotis.* Dimitrios is a great friend, and I want to thank him for all the helpful discussions we had during our time in London. I admire his unbounded passion for complexity theory, and I wish him good luck!

- *João Ribeiro.* João is a wonderful person and an exceptional researcher. I am grateful for his help and suggestions, and I am sure he will become an outstanding academic.

- *Carlo Sparaciari.* I want to thank Carlo for all the technical and non-technical discussions we had, as well as for his sincere friendship! I am grateful to have met him during my PhD!

- *Hyejung Hailey Jee, Navneeth Ramakrishnan, and Samson Wang.* I want to thank my amazing groupmates for the time we spent together during our PhDs. It has been a lot of fun, and I wish them a bright future!

- *Amani El-Kholy.* Amani is the PhD Programme Manager at the Department of Computing. She is one of the most loved people in our department, and the reason is self-evident: she is always present for the PhD candidates, no matter what! Thank you, Amani!

There are other people I need and want to thank.

First, I want to thank my wonderful mother, Maurizia Zannini, and my amazing brother and sister, Lorenzo Borderi and Maria Borderi. I am incredibly grateful to my father, Marco Borderi, for his help and support. I want to thank my grandparents, who supported me in every imaginable way. Even if you are not with us anymore, this work is also dedicated to you.

I cannot forget to thank my friend Matteo Neri (Mario) for his friendship! Life is always better when we are together!

When I first arrived in London, my radiant friend Ege Savas took the room I wanted. I could not have found a better flatmate to start my London adventures! I also want to thank my previous flatmate Astrid Verstraete for keeping my cumbersome paintings in her room for so long. Finally, I have to recognize the outstanding culinary skills of my current flatmate Mattia Feleppa, who is also constantly keeping me updated on the new music trends, so I can pretend to be younger!

I want to thank my brilliant lawyer Alessandro Morleo, his sister Giulia Morleo, and their beautiful family. In particular, I want to thank zio Elio for his amazing gelato and nonna Angela for the best food in town.

Several friends visited me during my time in London. In particular, my lifelong friend

Edoardo Cappelli, my gorgeous friends and party animals Riccardo Solazzi, Giandomenico Caterino and Federico Armaroli, the bright Lorenzo Duso, my friends Nicola Santullo and Nicolò Zagni, and the ladies Rachele Bagnolini, and Victoria Nikitina.

I want to thank Damarys for our time together; you made me a much better man. Thanks also to Gaia and Veronica, you are special!

I want to thank my brilliant friend Matteo Pontecorvi for sharing his experience and giving honest comments, my friends Andrea Sacco, Filippo Bernardoni and Andrea Bernardoni for their hospitality, Cristian Esposito, Salvatore Guercio for the pittas, Tommaso Mezzetti for helping me and my brother with the wine, Alessio Arancio for our chats, Simone Calzolari for our endless walks, Jacopo Collina, Enrico Pascai, Luca Lovisetto, Nicola Vanzini, and Riccardo Vinerba for our traditional "gran quests" at the Japanese restaurant, Edoardo Venturi, who is always there if I need a taxi ride, and to my sunny friend Jessica Biagini.

I want to thank my previous officemates Dimitrios Letsios, Georgia Kouyialis, Miten Mistry, Simon Olofsson, Francesco Ceccon, Johannes Wiebe, Alexander Thebelt, Juan Campos Salazar, Mehdi Bahri, Jan Kronqvist, and Haoyang Wang. Thank you for adopting me when I started my PhD!

Finally, I want to thank all the professors I worked with at Imperial College London as part of my teaching activity. I have learnt so much from those collaborations. Thank you!

List of Publications

The material presented in this thesis is based on (parts of) the following works.

1. Mario Berta, Francesco Borderi, Omar Fawzi, and Volkher B. Scholz. Quantum coding via semidefinite programming. In *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019.

2. Mario Berta, Francesco Borderi, Omar Fawzi, and Volkher B. Scholz. Semidefinite programming hierarchies for constrained bilinear optimization. *Mathematical Programming*, 2021.

3. Mario Berta, Francesco Borderi, and Omar Fawzi. On de Finetti reductions and representation theorems (*Research Notes*). 2021.

Abstract

The main topic of this thesis is the study of de Finetti methods and their applications in quantum information theory. The primary motivation of a de Finetti representation theorem is to represent, or approximate, a mathematical object symmetric under permutation of its components, into a probabilistic ensemble of elementary independent and identically distributed (i.i.d.) constituents. Approximations are given by finite version of those results, while exact representations are provided by infinite de Finetti representation theorems. One of their most common applications in quantum information theory, is the approximation of the set $\mathrm{Sep}(A:B)$ of separable states. To that purpose, the notion of n-extendibility plays a central role. A quantum state ρ_{AB} is said to be n-extendible if there exists a multipartite extension $\rho_{AB_1^n}$ that is symmetric with respect to A, i.e., invariant under permutation of the B-systems. While every separable state is also n-extendible, there exist n-extendible states that are not separable. Thus, for a fixed n, the set $n\text{-}\mathrm{Ext}(A:B)$ of n-extendible states provides an outer approximation to the set of separable states. Moreover, this approximation is computationally efficient, since it leads to semidefinite programs (SDPs). If we are looking for a better approximation, we can increase n, and, if we take the limit $n \to \infty$ we get an exact representation. In other words, a quantum state that is n-extendible for any n must be separable. In several applications, we are interested in quantum states that are not only separable, but also subject to additional

linear constraints. This observation has been the primary motivation of our research, and our findings include

1. The development of general mathematical techniques that can be used to obtain concrete constrained de Finetti representation theorems for the desired application.

2. The application of those methods to the problem of approximate quantum error correction. In particular, we use our framework to develop asymptotically converging SDP hierarchies that can be used to study the average and worst error cases, as given by the quantum channel fidelity and a channel distance based on the diamond norm, respectively.

De Finetti reductions are another class of techniques that are used to take advantage of permutation symmetries. For example, a quantum de Finetti reduction provides an upper bound to a symmetric quantum state in the form of an integral superposition of product states, weighted by a factor which is polynomial in terms of the number of copies and exponential in terms of the local dimensionality. Our research results in this direction include

1. A new de Finetti reduction in presence of an additional system carrying side information, that can handle various types of linear constraints.

2. The development of entropic techniques that can be used to generate de Finetti representation theorems from a starting de Finetti reduction. In particular, we use those methods to obtain a new proof for finite quantum de Finetti theorems.

List of Figures

Table of Contents

Chapter 1

Introduction

The approximation of the set $\mathrm{Sep}(A:B)$ of separable states is a common but computationally hard problem, which arises in many application in quantum information theory (see, e.g., [6]). A quantum state ρ_{AB} on $AB := A \otimes B$ is said to be *separable*, if it can be written as $\rho_{AB} = \sum_{i\in I} p_i \sigma_A^i \otimes \tau_B^i$, for a probability distribution $\{p_i\}_{i\in I}$, and quantum states $\{\sigma_A^i\}_{i\in I}$, and $\{\tau_B^i\}_{i\in I}$. The elements in $\mathrm{Sep}(A:B)$ describe unentangled states. Thus, being able to characterize $\mathrm{Sep}(A:B)$ is extremely important in order to understand entanglement, which is one of the main features of quantum mechanics. Operationally speaking, the characterization of $\mathrm{Sep}(A:B)$ is connected to the formulation of *separability tests*. A popular approach for the approximation of $\mathrm{Sep}(A:B)$ is via the notion of n-extendibility, where $n > 0$ is a natural number. The state ρ_{AB} is said to be *n-extendible* if there exists a quantum state $\rho_{AB_1^n}$ on $AB_1^n := A \otimes B^{\otimes n}$ satisfying the following two conditions

1. $\mathrm{Tr}_{B_2^n}(\rho_{AB_1^n}) = \rho_{AB}$,
2. $(\mathcal{I}_A \otimes \mathcal{U}_{B^n}^{\pi})(\rho_{AB_1^n}) = \rho_{AB_1^n}$ for every $\pi \in \mathfrak{S}_n$,

with $B_1 := B$, and $\mathfrak{S}_n$ denoting the set of permutations acting on n elements (or letters).

Condition 1. implies that ρ_{AB} is the local state on the system AB (i.e., $\rho_{AB_1^n}$ is an *extension* of ρ_{AB}), while condition 2. requires $\rho_{AB_1^n}$ to be invariant under permutation of the B-systems (i.e., $\rho_{AB_1^n}$ is *symmetric with respect to* A).

While every separable state is also n-extendible, there exist n-extendible states that are not separable [24], [52]. Thus, for a fixed n, the set $n\text{-Ext}(A:B)$ of n-extendible states provides an outer approximation to the set of separable states. In other words, for any given n, $\text{Sep}(A:B)$ is a proper subset of $n\text{-Ext}(A:B)$

$$\text{Sep}(A:B) \subsetneq n\text{-Ext}(A:B). \tag{1.1}$$

Finite *quantum de Finetti theorems* quantify the distance between $n\text{-Ext}(A:B)$ and $\text{Sep}(A:B)$. Moreover, one obtains convergence in the limit $n \to \infty$ [75]. More precisely (see [21, Theorem II.7]), if $\rho_{AB} \in n\text{-Ext}(A:B)$, there exists a probability distribution $\{p_i\}_{i\in I}$ and states ρ_A^i, ρ_B^i, such that

$$\left\| \rho_{AB} - \sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i \right\|_1 \leq \frac{2d_B^2}{n}, \tag{1.2}$$

where d_B denotes the dimension of the Hilbert space B.

This result can be generalized for $k \in \{1, \ldots, n-1\}$ to [21, 57]

$$\left\| \rho_{AB_1^k} - \sum_{i\in I} p_i \rho_A^i \otimes \left(\rho_B^i\right)^{\otimes k} \right\|_1 \leq \frac{2kd_B^2}{n}, \tag{1.3}$$

which is optimal on k and n for a fixed dimension d_B and up to a constant factor[1] (see [21, Theorem II.10]). In other words, if a multipartite state on AB_1^n is symmetric with respect to A, then the reduced state on the first k systems AB_1^k is close to a separable mixture of independent

[1]Moreover, in [21, Lemma III.9] the authors prove that the error term must be at least $\frac{d_B}{2n}\left(1 - \frac{1}{d_B^2}\right)$. In particular, this shows that we cannot obtain a dimension-independent bound for quantum de Finetti theorems, and the dimensional dependence on d_B in (1.2) and (1.3) cannot be exponentially improved.

and identical states for k sufficiently smaller than n. Notice that, in the asymptotic limit $n \to \infty$ and holding k constant, the above inequalities reduce to equalities and the approximations become exact. For our setting, however, we are interested more generally in characterizing bipartite states that are separable, but subject to linear constraints on the quantum states ρ_A^i, ρ_B^i as well[2].

1.1 De Finetti Theorems with Linear Constraints

In particular, we are interested in the study of constrained bilinear optimization problems of the form

$$Q := \max \quad \mathrm{Tr}\left[G_{AB}\left(\sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i\right)\right] \tag{1.4}$$

$$s.t. \quad p_i \geq 0 \quad \forall i \in I, \quad \sum_{i\in I} p_i = 1 \tag{1.5}$$

$$\rho_A^i \succeq 0, \quad \rho_B^i \succeq 0 \quad \forall i \in I \tag{1.6}$$

$$\mathrm{Tr}\left(\rho_A^i\right) = \mathrm{Tr}\left(\rho_B^i\right) = 1 \quad \forall i \in I \tag{1.7}$$

$$\Lambda_{A\to C_A}\left(\rho_A^i\right) = X_{C_A}, \quad \Gamma_{B\to C_B}\left(\rho_B^i\right) = Y_{C_B} \quad \forall i \in I, \tag{1.8}$$

where G_{AB} is a fixed operator, $\Lambda_{A\to C_A}$, and $\Gamma_{B\to C_B}$ are linear maps (also known as *super-operators*), and X_{C_A}, Y_{C_B} are the operators defining the linear constraints in combination with the linear maps. As we see, the optimization is over a subset of $\mathrm{Sep}(A:B)$, determined by the linear constraints

$$\Lambda_{A\to C_A}\left(\rho_A^i\right) = X_{C_A}, \quad \Gamma_{B\to C_B}\left(\rho_B^i\right) = Y_{C_B} \quad \forall i \in I. \tag{1.9}$$

[2] As we will show in Section 4.2, standard de Finetti theorems are not sufficient for our purposes, and new de Finetti representation theorems are indeed needed to capture the additional linear constraints.

Clearly, we are interested in the general case where the linear constraints (1.9) are not trivial, determining a proper subset of $\mathrm{Sep}(A:B)$. In order to outer approximate this subset, a new de Finetti theorem with linear constraints is needed. Thus, we prove the following finite constrained representation result.

Theorem 1.1.1. *Let $\rho_{AB_1^n}$ be a quantum state, $\Lambda_{A\to C_A}, \Gamma_{B\to C_B}$ super-operators, and X_{C_A}, Y_{C_B} operators such that*

$$\mathcal{U}^{\pi}_{B_1^n}(\rho_{AB_1^n}) = \rho_{AB_1^n}\ \forall \pi \in \mathfrak{S}_n \qquad \textit{symmetric with respect to } A \tag{1.10}$$

$$\Lambda_{A\to C_A}(\rho_{AB_1^n}) = X_{C_A}\otimes \rho_{B_1^n} \qquad \textit{linear constraint on } A \tag{1.11}$$

$$\Gamma_{B_n\to C_B}(\rho_{B_1^n}) = \rho_{B_1^{n-1}}\otimes Y_{C_B} \qquad \textit{linear constraint on } B. \tag{1.12}$$

Then, we have that

$$\left\| \rho_{AB} - \sum_{i\in I} p_i \sigma_A^i \otimes \omega_B^i \right\|_1 \leq \min\left\{ f(A,B), f(B|\cdot)\right\} \sqrt{\frac{(2\ln 2)\log\left(d_A\right)}{n}} \tag{1.13}$$

with $\{p_i\}_{i\in I}$ a probability distribution, $\rho_{AB} = \mathrm{Tr}_{B_2^n}\left(\rho_{AB_1^n}\right)$, $\log(\cdot) := \log_2(\cdot)$, and quantum states σ_A^i, ω_B^i such that for every $i \in I$:

$$\Lambda_{A\to C_A}\left(\sigma_A^i\right) = X_{C_A} \quad \textit{and} \quad \Gamma_{B\to C_B}\left(\omega_B^i\right) = Y_{C_B}. \tag{1.14}$$

The quantity $f(A,B)$ is known as minimal distortion for the bipartite system AB, and can be bound as $f(A,B) \leq 18\sqrt{d_A d_B}$ [17, Lemma 14]. The quantity $f(B|\cdot)$ will then be referred as minimal distortion with side information for system B, and can be bound as $f(B|\cdot) \leq 2d_B$ [53, Lemma 8].

We also generalize the above result to $k \in \{1,\ldots,n-1\}$ copies, obtaining the following bound

$$\left\| \rho_{AB_1^k} - \sum_{i\in I} p_i \sigma_A^i \otimes \left(\omega_B^i\right)^{\otimes k} \right\|_1 \leq k f(B|\cdot) \sqrt{(2\ln 2)\frac{\log d_A + (k-1)\log d_B}{n-k+1}}. \tag{1.15}$$

Comparing the bound of (1.15) with (1.3), we see that the room for improvement is fairly limited, i.e., we may be able to improve the square root and the logarithm dependence, but the overall bound cannot be made exponentially better. Using Theorem 1.1.1 we can generate an asymptotically converging hierarchy of semidefinite programs that can be used to approximate Q (1.4). This is formalized by the following theorem.

Theorem 1.1.2. *For the SDPs*

$$\begin{aligned} \mathrm{SDP}_n := \max \quad & \mathrm{Tr}\big[G_{AB}\rho_{AB_1}\big] && (1.16)\\ s.t. \quad & \rho_{AB_1^n} \succeq 0, \mathrm{Tr}(\rho_{AB_1^n}) = 1 && (1.17)\\ & \rho_{AB_1^n} = \mathcal{U}^{\pi}_{B_1^n}\left(\rho_{AB_1^n}\right) \ \forall \pi \in \mathfrak{S}_n && (1.18)\\ & \Lambda_{A\to C_A}\left(\rho_{AB_1^n}\right) = X_{C_A} \otimes \rho_{B_1^n} && (1.19)\\ & \Gamma_{B_n\to C_B}\left(\rho_{B_1^n}\right) = \rho_{B_1^{n-1}} \otimes Y_{C_B}, && (1.20) \end{aligned}$$

we have for $d := \max\{d_A, d_B\}$ *that*

$$0 \leq \mathrm{SDP}_n - Q \leq \frac{\mathrm{poly}(d)}{\sqrt{n}} \quad \textit{implying} \quad Q = \lim_{n\to\infty} \mathrm{SDP}_n. \tag{1.21}$$

It is important to realize that the results of Theorem 1.1.1 and Theorem 1.1.2 contain several degrees of freedom we can choose. Namely, the various underlying Hilbert spaces A, C_A, B, C_B, the operator G_{AB} appearing in the objective function, the two linear maps $\Lambda_{A\to C_A}$ and $\Gamma_{B_n\to C_B}$, and the operators X_{C_A} and Y_{C_B} defining the linear constraints in combination with the linear maps. Thus, the outlined framework can be used to generate the specific de Finetti representation theorem, and associated asymptotically converging SDP hierarchy, needed for the desired application. One application of particular interest is found in the research area of approximate quantum error correction.

1.2 Approximate Quantum Error Correction

Given a noisy classical channel $N_{X\to Y}$, a central quantity of interest in error correction is the *maximum success probability* $p(N, M)$ for transmitting a uniform M-dimensional message under the noise model $N_{X\to Y}$. This is a bilinear maximization problem, which is in general NP-hard to approximate up to a sufficiently small constant factor [8]. Nevertheless, there exists an efficiently computable linear programming relaxation $\mathrm{lp}(N, M)$ (sometimes called *meta-converse* [45, 71]) giving quantifiable upper bounds on $p(N, M)$ [8]. Thus, the gap between $\mathrm{lp}(N, M)$ and $p(N, M)$ is well-understood.

The analogue quantum problem is to determine the *quantum channel fidelity* $F(\mathcal{N}, M)$, which is defined as follows.

Definition 1.2.1. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. The quantum channel fidelity for message dimension M is defined as*

$$F(\mathcal{N}, M) := \max \quad F\Big(\Phi_{\overline{B}R}, \big(\big(\mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}}\big) \otimes \mathcal{I}_R\big)(\Phi_{AR})\Big) \tag{1.22}$$

$$s.t. \quad \mathcal{D}_{B\to\overline{B}}, \mathcal{E}_{A\to\overline{A}} \text{ quantum channels}, \tag{1.23}$$

where $F(\rho, \sigma) := \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2$ denotes the fidelity, Φ_{AR} denotes the maximally entangled state on AR, and we have $M = d_A = d_{\overline{B}} = d_R$.

The optimization is performed over sets of quantum channels (i.e., trace preserving completely positive linear maps between two spaces of quantum states), which is not practical to handle or visualize. Thus, we show that $F(\mathcal{N}, M)$ can be rewritten in a more convenient form, as the following optimization over Choi states

$$F(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)\left(\sum_{i\in I} p_i E^i_{A\overline{A}} \otimes D^i_{B\overline{B}}\right)\right] \tag{1.24}$$

$$s.t. \quad p_i \geq 0 \; \forall i \in I, \; \sum_{i \in I} p_i = 1 \tag{1.25}$$

$$E^i_{A\overline{A}} \succeq 0, \quad D^i_{B\overline{B}} \succeq 0 \tag{1.26}$$

$$E^i_A = \frac{1_A}{d_A}, \quad D^i_B = \frac{1_B}{d_B} \; \forall i \in I, \tag{1.27}$$

where $J^{\mathcal{N}}_{B\overline{A}} := (\mathcal{N}_{\overline{A}\to B} \otimes \mathcal{I}_{\overline{A}})(\Phi_{\overline{A}\,\overline{A}})$ denotes the Choi state of the quantum channel $\mathcal{N}_{\overline{A}\to B}$.

As in the classical case, this is a bilinear optimization problem, only now with operator-valued variables. In order to approximate $F(\mathcal{N}, M)$, an efficiently computable semidefinite programming relaxation $\mathrm{SDP}(\mathcal{N}, M)$ was given in [65]. However, contrary to the classical case, the gap between $\mathrm{SDP}(\mathcal{N}, M)$ and $F(\mathcal{N}, M)$ is not understood. On the other hand, the tools we have developed and outlined in the previous section, can be used to generate a converging hierarchy of efficiently computable semidefinite programming relaxations, allowing us to quantify the gap between these new relaxations and $F(\mathcal{N}, M)$. In fact, we can fix the degrees of freedom available in Theorem 1.1.1 to generate the desired constrained de Finetti representation theorem. By doing so, we have automatically an associated asymptotically converging SDP hierarchy (Theorem 1.1.2), which reads

$$\mathrm{SDP}_n(\mathcal{N}, M) := \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B_1} \otimes \Phi_{A\overline{B}_1}\right) \rho_{A\overline{A}B_1\overline{B}_1}\right] \tag{1.28}$$

$$s.t. \quad \rho_{A\overline{A}(B\overline{B})^n_1} \succeq 0, \; \mathrm{Tr}\left[\rho_{A\overline{A}(B\overline{B})^n_1}\right] = 1 \tag{1.29}$$

$$\rho_{A\overline{A}(B\overline{B})^n_1} = \mathcal{U}^{\pi}_{(B\overline{B})^n_1}\left(\rho_{A\overline{A}(B\overline{B})^n_1}\right) \; \forall \pi \in \mathfrak{S}_n \tag{1.30}$$

$$\rho_{A(B\overline{B})^n_1} = \frac{1_A}{d_A} \otimes \rho_{(B\overline{B})^n_1} \tag{1.31}$$

$$\rho_{A\overline{A}(B\overline{B})^{n-1}_1 B_n} = \rho_{A\overline{A}(B\overline{B})^{n-1}_1} \otimes \frac{1_{B_n}}{d_B}. \tag{1.32}$$

Recalling that the original optimization was over quantum channels, the presented results can be interpreted as a way to approximate permutationally invariant *bipartite*[3] quantum channels

[3]It is important to stress that, in the application of approximate quantum error correction, we have three

by a mixture of product channels, i.e., as de Finetti theorems for bipartite quantum channels. Moreover, we can also state the representation theorem directly in terms of the quantum channels, obtaining an upper bound for the diamond norm distance.

We also study the setting in which we allow for classical forward communication assistance. Thus, we modify Definition 1.2.1 to include the classical channel that can be used to send classical information from one party (say Alice) to the other (say Bob). The corresponding *LOCC(1)-assisted (quantum) channel fidelity* $F^{\mathrm{LOCC}(1)}(\mathcal{N}, M)$ is defined as follows.

Definition 1.2.2. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. The LOCC(1)-assisted channel fidelity for message dimension M is defined as*

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) := \max \quad F\Big(\Phi_{\overline{B}R}, \sum_{i\in I} \big(\big(\mathcal{D}^i_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}^i_{A\to\overline{A}}\big) \otimes \mathcal{I}_R\big)(\Phi_{AR})\Big) \tag{1.33}$$

$$s.t. \quad \sum_{i\in I} \mathcal{E}^i_{A\to\overline{A}} \text{ quantum channel with } \mathcal{E}^i_{A\to\overline{A}} \text{ cp for } i \in I \tag{1.34}$$

$$\mathcal{D}^i_{B\to\overline{B}} \text{ quantum channel } \forall i \in I, \tag{1.35}$$

where Φ_{AR} denotes the maximally entangled state on AR, cp is the abbreviation for completely positive, and we have $M = d_A = d_{\overline{B}} = d_R$.

We then follow the same approach used for the quantum channel fidelity, to rewrite $F^{\mathrm{LOCC}(1)}(\mathcal{N}, M)$ as a bilinear optimization program, and to generate the appropriate constrained de Finetti representation theorem (using Theorem 1.1.1) with associated asymptotically converging hierarchy (using Theorem 1.1.2). Moreover, we show several bounds for the two

types of quantum channels: the fixed noise model $\mathcal{N}_{\overline{A}\to B}$, and the coding schemes given by the various encoder and decoder pairs $(\mathcal{E}_{A\to\overline{A}}, \mathcal{D}_{B\to\overline{B}})$. Each coding scheme then determines a bipartite quantum channel with input system $A \otimes B$ and output system $\overline{A} \otimes \overline{B}$.

fidelity measures, e.g.,

$$F^{\mathrm{LOCC}(1)}(\mathcal{N},M)\geq F(\mathcal{N},M)\geq\left(F^{\mathrm{LOCC}(1)}(\mathcal{N},M)\right)^{2}. \tag{1.36}$$

We analyse our results by performing numerical experiments for the low levels of our hierarchies. The experiments have been done in MATLAB using the QETLAB library [55], CVX [40], MOSEK [1], and SDPT3 [78]. In addition, all the code has been made avilable at the following link: https://github.com/FrancescoBorderi/Quantum-SDPs. While our analysis is limited to the low levels of the SDP hierarchies, due to the size of the optimization programs, we have been able to use to following *rank loop condition* to certify that a certain level of the hierarchy already gives the optimal value.

Lemma 1.2.3. *[68],[51] Let* $\rho_{AB_1^n}=\mathcal{U}^{\pi}_{B_1^n}\left(\rho_{AB_1^n}\right)$ *for all* $\pi\in\mathfrak{S}_n$ *and fixed* $0\leq k\leq n$ *such that* $\rho^{\mathrm{T}_{B_{k+1}^n}}_{AB_1^n}\succeq 0$. *Then,* ρ_{AB_1} *is separable if*

$$\mathrm{rank}(\rho_{AB_1^n})\leq\max\left\{\mathrm{rank}\left(\rho_{AB_1^k}\right),\mathrm{rank}\left(\rho_{B_{k+1}^n}\right)\right\}. \tag{1.37}$$

For most cases the hierarchies of SDPs collapse to the first or second level, without the need to explore the higher levels, which are computationally much more expensive.

The presented fidelity measures, i.e., $F(\mathcal{N},M)$ and $F^{\mathrm{LOCC}(1)}(\mathcal{N},M)$, correspond to the average error case. On the other hand, we can study the worst case error by considering the following channel distance based on the diamond norm.

Definition 1.2.4. *Let* $\mathcal{N}_{\overline{A}\to B}$ *be a quantum channel and* $M\in\mathbb{N}$, *with* $M=d_A=d_{\overline{B}}$. *The channel distance is defined as*

$$\Delta(\mathcal{N},M):=\min\frac{1}{2}\left\|\mathcal{D}_{B\to\overline{B}}\circ\mathcal{N}_{\overline{A}\to B}\circ\mathcal{E}_{A\to\overline{A}}-\mathcal{I}_{A\to\overline{B}}\right\|_{\diamond} \tag{1.38}$$

$$s.t.\ \mathcal{D}_{B\to\overline{B}},\mathcal{E}_{A\to\overline{A}}\ \textit{quantum channels.} \tag{1.39}$$

With some additional manipulation, we show how to write the above optimization program in terms of the Choi states of the quantum channels, in a form suitable for our framework. Finally, we generate an asymptotically converging hierarchy of semidefinite programs, generating lower bounds for $\Delta(\mathcal{N}, M)$.

1.3 De Finetti Reductions with Linear Constraints

In several applications, instead of representation results as given by de Finetti theorems, one may need to establish a generalized order relation between the symmetric mathematical object and the probabilistic ensemble of elementary i.i.d. constituents. De Finetti reductions, previously known as "post-selection techniques" [22] or methods based on "universal states" [46], provide the desired inequality. For example, a *quantum de Finetti reduction* provides an upper bound to a symmetric quantum state in the form of an integral superposition of product states, weighted by a factor which is polynomial in terms of the number of copies and exponential in terms of the local dimensionality

$$\rho_{\mathcal{H}^n} \preceq (n+1)^{d_{\mathcal{H}}^2-1} \int \sigma_{\mathcal{H}}^{\otimes n} \, \mathrm{d}\sigma_{\mathcal{H}}, \tag{1.40}$$

where $\rho_{\mathcal{H}^n}$ is a permutation invariant quantum state, and $\mathrm{d}\sigma_{\mathcal{H}}$ is an appropriate measure over the set of quantum states on $\mathcal{H}$. The generality of expression (1.40) is also its main drawback. On one hand, unlike finite de Finetti representation theorems, (1.40) provides an exact bound, without any parameter controlling the approximation error. On the other hand, all permutation invariant quantum states are upper bounded by the same mixture of tensor product states. Any other information encoded in the permutation invariant state $\rho_{\mathcal{H}^n}$ is lost. There exist in the literature several quantum de Finetti reductions that are able to handle *specific* linear constraints on the permutation invariant state (e.g., [33] and [64]). Those theorems restrict the

support of the measure in order to capture the specific linear constrain on the initial state or introduce a fidelity weight in the integral superposition.

For cryptographic applications and error correction, it is often useful to study the case where a new system, carrying external side information, adds a non-symmetric contribution to the symmetric object.

We prove the following new de Finetti reduction in presence of quantum side information.

Proposition 1.3.1. *Let Q, A and B be Hilbert spaces and $\rho_{QA^nB^n}$ a state symmetric with respect to Q. Moreover, let $\rho_{A^n} = \sigma_A^{\otimes n}$ for a fixed state σ_A. Then, there exist a probability measure $\mathrm{d}\sigma_{AB}$ on the set of extensions σ_{AB} of σ_A and a state ω_Q such that*

$$\rho_{QA^nB^n} \preceq (n+1)^{3d^2} \cdot \omega_Q \otimes \int \sigma_{AB}^{\otimes n} \, \mathrm{d}\sigma_{AB}, \tag{1.41}$$

with $d := d_A d_B^2$.

Our result can be seen as an extension of the constrained de Finetti reduction presented in [33, Corollary 3.2]. Moreover, we show that our de Finetti reduction can handle, in addition to the marginal constraint $\rho_{A^n} = \sigma_A^{\otimes n}$ on the symmetric part, a general linear constraint on the quantum side information. This is the content of the following corollary.

Corollary 1.3.2. *Under the same assumptions of Proposition 1.3.1 with additionally $\Gamma_{Q\to F}$ a linear map and X_F an operator on a Hilbert space F, the state ω_Q can be chosen such that*

$$\Gamma_{Q\to F}(\rho_{QA^n}) = X_F \otimes \sigma_A^{\otimes n} \quad \Longrightarrow \quad \Gamma_{Q\to F}(\omega_Q) = X_F. \tag{1.42}$$

Note that the marginal constraint $\rho_{A^n} = \sigma_A^{\otimes n}$ is a special type of linear constraint, but we do not know if it is possible to extend this to general linear constraints.

So far, no clear or systematic connection between de Finetti reductions and de Finetti representation theorems has been proven in the literature. In this thesis we show how to derive

de Finetti representation theorems from de Finetti reductions. First, we prove the following lemma, showing that de Finetti reductions can be interpreted as relative entropy inequalities.

Lemma 1.3.3. *Let Q and G be Hilbert spaces, and ρ_{QG^n} a state symmetric with respect to Q. Consider a de Finetti reduction of the form*

$$\rho_{QG^n} \preceq \mathrm{poly}(n) \cdot \sigma_Q \otimes \int \sigma_G^{\otimes n} d\sigma_G, \tag{1.43}$$

where $d\sigma_G$ is an appropriate measure over the set of quantum states on G. Then, there exists a discrete random variable X, p_X a probability mass function, and σ_G^x quantum states for every $x \in \mathrm{image}(X)$, such that

$$D\Big(\rho_{QG^n} \Big\| \sigma_Q \otimes \sum_x p_X(x) \left[\sigma_G^x\right]^{\otimes n}\Big) \leq \log \mathrm{poly}(n). \tag{1.44}$$

This finding will be the basis to go from de Finetti reductions to representation theorems. Second, we use a technique based on chain rules for relative entropy to obtain a new proof for the classical de Finetti theorem. Third, we can leverage the obtained result to the quantum setting, giving the following theorem.

Theorem 1.3.4. *Let $k \in \{1, \ldots, n-1\}$, X be a discrete random variable, $G_1 \cdots G_n$ Hilbert spaces with $G_1 \cong \ldots \cong G_n$, ρ_{G^n} and σ_G^x quantum states for every $x \in \mathrm{image}(X)$, p_X a probability mass function, and assume ρ_{G^n} to be symmetric. Whenever we have*

$$D\Big(\rho_{G^n} \Big\| \sum_x p_X(x) [\sigma_G^x]^{\otimes n}\Big) \leq \log \mathrm{poly}(n), \tag{1.45}$$

then there exists a probability mass function q_X such that

$$\left\| \rho_{G^k} - \sum_x q_X(x) [\sigma_G^x]^{\otimes k} \right\|_1 \leq O\left(\sqrt{\frac{k \cdot d_G^{2k}}{n} \cdot \log n}\right). \tag{1.46}$$

Notice that the bound on the approximation error grows exponentially fast with k, which we know that is not optimal, as previously discussed. Whether it is possible to improve that k-dependence *and* maintain the proposed proof technique is still an open question. On the other hand, we already know that the dependence in n is suboptimal.

Finally, we use de Finetti reductions to derive a new de Finetti theorem for quantum channels. In comparison to our results from [13] and [14], which are given for bipartite quantum channels, we show that is possible to drop one constraint and still achieve asymptotic convergence. While we are not able to prove the theoretical minimality of our constraints, the simplification of the existing conditions is definitely a fundamental step in the right direction. Moreover, our new results provide insights on the "power" of the constraints and their effect on the convergence speed.

1.4 Thesis Organization

This thesis is organized as follows.

- In Chapter 2 we present some background material.
- In Chapter 3 we develop new de Finetti theorems with linear constraints and we use them to generate SDP hierarchies for constrained bilinear optimization programs.
- In Chapter 4 we focus on certain optimization problems arising in the context of approximate quantum error correction and we adapt the results of Chapter 3 to the desired setting. Proof of concept numerics are implemented to test the low levels of our hierarchies.
- In Chapter 5 we prove a new constrained de Finetti reduction with side information, and we establish a connection between de Finetti reductions and de Finetti representation

theorems. We use our methods to simplify the SDP hierarchy for quantum channels.

- In Chapter 6 we present some open problems.

Chapter 2

Preliminaries

This chapter provides a concise presentation of the mathematical framework needed to understand the subsequent chapters. While it is mainly based on Watrous's book [87], many other textbooks provide a good introduction to the subject. For an excellent introduction to quantum information theory I recommend the textbook by Nielsen and Chuang [69] and the one by Wilde [89]. For a review of the methods of convex optimization used in this thesis, semidefinite programming in particular, the reader is referred to the textbook by Boyd and Vandenberghe [15].

2.1 Finite Dimensional Hilbert Spaces

In quantum mechanics, Hilbert spaces represent one of the most fundamental mathematical objects. A *complex Hilbert space* $\mathcal{H}$ is a vector space with two defining characteristics

1. $\mathcal{H}$ is equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$,

2. $\mathcal{H}$ is *complete* for the distance induced by the inner product $\langle \cdot, \cdot \rangle$.

In quantum mechanics, unlike several other areas of mathematics, the inner product $\langle \cdot, \cdot \rangle$ is conventionally chosen to be *antilinear in the first argument*, not in the second one[1]. Notice that $\langle \cdot, \cdot \rangle (\mathcal{H} \times \mathcal{H}) \subseteq \mathbb{C}$, clearly showing the presence of the complex field $(\mathbb{C}, +, \cdot)$ in the algebraic structure of the vector space. In this thesis we will implicitly assume all the Hilbert spaces to be complex vector spaces[2].

Using the standard Dirac's bra-ket notation, we will call kets the elements of $\mathcal{H}$, i.e., the vectors. A generic element of $\mathcal{H}$, i.e., a *ket*, will be then denoted by $|\psi\rangle \in \mathcal{H}$. Given the Hilbert space $\mathcal{H}$, we will denote with $\mathcal{H}^*$ its *topological dual space*[3], i.e., the space of all continuous linear functionals from $\mathcal{H}$ to $\mathbb{C}$.

Given a ket $|\psi\rangle \in \mathcal{H}$, we can use the inner product $\langle \cdot, \cdot \rangle$ to create a correspondence between the Hilbert space $\mathcal{H}$ and its topological dual $\mathcal{H}^*$

$$\mathcal{H} \ni |\psi\rangle \to f_{|\psi\rangle} \in \mathcal{H}^*, \tag{2.1}$$

where

$$f_{|\psi\rangle} : \mathcal{H} \to \mathbb{C}, |\phi\rangle \to \langle |\psi\rangle, |\phi\rangle \rangle. \tag{2.2}$$

Riesz representation theorem [27, Theorem 3.7.7] guarantees that (2.2) is a linear bijection, thus realizing an isomorphism between $\mathcal{H}$ and its topological dual $\mathcal{H}^*$. To denote that two

[1] A function $f : V \to W$ between two complex vector spaces V, W is said to be *antilinear* if it is *additive*, i.e., $f(x+y) = f(x) + f(y)$ for every $x, y \in V$, and *conjugate homogeneous*, i.e., $f(\alpha x) = \overline{\alpha} f(x)$ for every $x \in V$ and $\alpha \in \mathbb{C}$.

[2] It is interesting to notice the need for complex numbers in quantum mechanics. For example, in [74] the authors propose a new Bell-type experiment in which the input-output correlations cannot be approximated by a version of quantum mechanics based on real Hilbert spaces.

[3] The topological dual space $\mathcal{H}^*$ is a subset of the *algebraic dual space*, where the linear functionals are not required to be continuous. In this thesis we will work with finite dimensional Hilbert spaces, and the two notions coincide.

spaces are isomorphic we will use the symbol $\cong$. Then, Riesz representation theorem proves that

$$\mathcal{H} \cong \mathcal{H}^*. \tag{2.3}$$

It is immediate to show that two finite-dimensional vector spaces are isomorphic if and only if they have the same dimension. Thus, given a ket $|\psi\rangle \in \mathcal{H}$, we can find a unique element $f_{|\psi\rangle} \in \mathcal{H}^*$ as defined above, and vice versa. Following Dirac's bra-ket notation, we will call the linear functional $f_{|\psi\rangle} \in \mathcal{H}^*$ the *bra* associated with the ket $|\psi\rangle$, and it will be denoted by $\langle\psi| \in \mathcal{H}^*$. Finally, the simplifying notation $\langle\psi|\phi\rangle$ is used in place of $\langle\psi|(|\phi\rangle) = \langle|\psi\rangle, |\phi\rangle\rangle$, i.e.,

$$\langle\psi|\phi\rangle := \langle|\psi\rangle, |\phi\rangle\rangle. \tag{2.4}$$

Here and henceforth we use the symbol $:=$ as *equal by definition*.

The completeness property required in 2. guarantees the convergence of all the Cauchy sequences[4] of points in $\mathcal{H}$ within the space $\mathcal{H}$ itself. As we see, the concept of completeness, which involves Cauchy sequences, requires a metric structure. Given a Hilbert space $\mathcal{H}$, a *norm* is automatically induced by the inner product as

$$\|\cdot\| : \mathcal{H} \to \mathbb{R} : |\psi\rangle \to \sqrt{\langle\psi|\psi\rangle}. \tag{2.5}$$

Thus, every Hilbert space is also a Banach space, i.e., a complete normed vector space. The metric structure is then automatically induced by the norm by defining a *metric/distance* on $\mathcal{H}$ as

$$d(\cdot,\cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{R} : (|\psi\rangle, |\phi\rangle) \to \||\psi\rangle - |\phi\rangle\|. \tag{2.6}$$

[4] A *Cauchy sequence* is a sequence of points such that, for every $\epsilon > 0$, the distance between any two elements of the sequence becomes smaller than ϵ after a certain index (which can depend on ϵ).

Completeness is an important requirement for the infinite dimensional case. However in this thesis we are interested in finite dimensional Hilbert spaces. Given a d-dimensional Hilbert space $\mathcal{H}$, where $d > 0$ is a natural number, the natural isomorphism between $\mathcal{H}$ and $\mathbb{C}^d$ automatically guarantees the completeness of the mathematical structure. More precisely, fixed an orthonormal basis $(|i\rangle)_{i=1,\dots,d}$ for $\mathcal{H}$, we can decompose any ket $|\psi\rangle \in \mathcal{H}$ as

$$|\psi\rangle = \sum_{i=1}^{d} \langle i|\psi\rangle |i\rangle. \tag{2.7}$$

The linear correspondence $\mathcal{H} \to \mathbb{C}^d : |\psi\rangle \to (\langle 1|\psi\rangle, \langle 2|\psi\rangle, \dots, \langle d|\psi\rangle)$ realizes the desired isomorphism between $\mathcal{H}$ and $\mathbb{C}^d$. Moreover, with respect to the canonical basis[5] of $\mathbb{C}^d$, the vector $(\langle 1|\psi\rangle, \langle 2|\psi\rangle, \dots, \langle d|\psi\rangle) \in \mathbb{C}^d$ can be written as the column

$$\begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \vdots \\ \langle d|\psi\rangle \end{pmatrix}, \tag{2.8}$$

which is often identified with the starting ket $|\psi\rangle$. Similarly, a bra can be identified, once a basis is fixed, as a row of d complex numbers.

In this thesis we will implicitly assume all the Hilbert spaces to be finite-dimensional, so we will not have to deal with completeness-related concerns. Given a (finite-dimensional) Hilbert space $\mathcal{H}$, we will denote with $d_{\mathcal{H}}$ its dimension.

Given two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, their *tensor product* $\mathcal{H} \otimes \mathcal{H}'$ is the Hilbert space generated by the linear span[6] of the basis $(|i\rangle \otimes |j'\rangle)_{\substack{i=1,\dots,d_{\mathcal{H}} \\ j=1,\dots,d_{\mathcal{H}'}}}$, where $(|i\rangle)_{i=1,\dots,d_{\mathcal{H}}}$ is an orthonormal basis

[5] The *canonical basis* of $\mathbb{C}^d$ is formed by the following collection of d elements with d entries: $((1,0,\dots,0),(0,1,\dots,0),\dots,(0,0,\dots,1))$.

[6] Given a basis, its *linear span* is defined as the vector space formed by all the linear combinations obtained with the elements of the basis.

for $\mathcal{H}$, and $(|j'\rangle)_{j=1,\dots,d_{\mathcal{H}'}}$ is an orthonormal basis for $\mathcal{H}'$. The scalars used for the linear span belong to the field associated with the two algebraic structures, i.e., $(\mathbb{C}, +, \cdot)$ in this case. It is immediate to see from the definition that the dimension of $\mathcal{H} \otimes \mathcal{H}'$ is the product of the individual dimensions, i.e.,

$$d_{\mathcal{H}\otimes\mathcal{H}'} = d_{\mathcal{H}} \cdot d_{\mathcal{H}'}. \tag{2.9}$$

In order to simplify the notation, we can also omit the $\otimes$ symbol, i.e., we define $\mathcal{H}\mathcal{H}' := \mathcal{H} \otimes \mathcal{H}'$.

In case of n copies of the same Hilbert space $\mathcal{H}$ we will use the notation $\mathcal{H}_1^n$ or also $\mathcal{H}^n$ to indicate $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, where $\mathcal{H}_i$ indicates the ith copy of $\mathcal{H}$, for $i = 1, \dots, n$. Moreover, the notation $\mathcal{H}_1^n$ can be generalized to select a contiguous collection of Hilbert spaces appearing in the tensor product $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. More precisely, for every $i, j = 1, \dots, n$ with $i < j$, the expression $\mathcal{H}_i^j$ stands for $\mathcal{H}_i \otimes \cdots \otimes \mathcal{H}_j$. It is useful to extend the new notation also when $i \geq j$. In particular, when $i = j$, we set $\mathcal{H}_i^i := \mathcal{H}_i$. On the other had, if $i > j$ we consider the expression $\mathcal{H}_i^j$ to be the empty set $\varnothing$.

2.1.1 Operators

Given two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, we can consider the vector space $\mathcal{L}(\mathcal{H}, \mathcal{H}')$ formed by all the linear maps, i.e., *operators*, between $\mathcal{H}$ and $\mathcal{H}'$. When $\mathcal{H} = \mathcal{H}'$, we will write $\mathcal{L}(\mathcal{H})$ in place of $\mathcal{L}(\mathcal{H}, \mathcal{H})$ to denote all the operators from $\mathcal{H}$ onto itself, i.e., the *endomorphisms* of $\mathcal{H}$. As we saw with kets, once we have fixed a basis for $\mathcal{H}$, we can identify any $|\psi\rangle \in \mathcal{H}$ with the corresponding column (2.8). In a similar way, once we have fixed a basis for $\mathcal{H}$ and $\mathcal{H}'$, we can identify any operator $T \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ with a $d_{\mathcal{H}'} \times d_{\mathcal{H}}$ complex matrix[7]. More precisely, if $(|i\rangle)_{i=1,\dots,d_{\mathcal{H}}}$ is an orthonormal basis for $\mathcal{H}$, and $(|j'\rangle)_{j=1,\dots,d_{\mathcal{H}'}}$ is an orthonormal basis for $\mathcal{H}'$,

[7]For this reason, one often finds the term matrix used in place of operator. However, notice that this correspondence requires choosing the bases for the two Hilbert spaces.

$\langle j'|T|i\rangle$ will be the coefficient in the jth row and ith column of the matrix representation for the operator T, for every $i = 1, \ldots, d_{\mathcal{H}}$ and $j = 1, \ldots, d_{\mathcal{H}'}$. This is because T can be decomposed with respect to the two orthonormal bases as

$$T = \sum_{\substack{i=1,\ldots,d_{\mathcal{H}} \\ j=1,\ldots,d_{\mathcal{H}'}}} \langle j'|T|i\rangle |j'\rangle\langle i|. \tag{2.10}$$

Notice that, if $T \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$, we set $T(|\psi\rangle) := T|\psi\rangle$ for every $|\psi\rangle \in \mathcal{H}$. Given an operator $T \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ we can consider its *adjoint operator* $T^\dagger \in \mathcal{L}(\mathcal{H}', \mathcal{H})$, defined by the relation

$$\langle |\phi\rangle, T|\psi\rangle\rangle = \langle T^\dagger|\phi\rangle, |\psi\rangle\rangle, \tag{2.11}$$

which must hold for every $|\psi\rangle \in \mathcal{H}$ and $|\phi\rangle \in \mathcal{H}'$. Using Riesz representation theorem [27, Theorem 3.7.7], it is immediate to show that $T^\dagger$ is unique. If $T \in \mathcal{L}(\mathcal{H})$ and $T = T^\dagger$, then T is said to be *self-adjoint* or *Hermitian*. The set of Hermitian operators acting on the Hilbert space $\mathcal{H}$ will be denoted by $\mathrm{Herm}(\mathcal{H})$. Given a $T \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$, it is important to realize that the notion of adjoint operator $T^\dagger$ is basis-free. On the other hand, the transpose T^{T} is a basis-dependent concept. If $(|i\rangle)_{i=1,\ldots,d_{\mathcal{H}}}$ is an orthonormal basis for $\mathcal{H}$, and $(|j'\rangle)_{j=1,\ldots,d_{\mathcal{H}'}}$ is an orthonormal basis for $\mathcal{H}'$, the *transpose* T^{T} of $T \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ with respect to those two bases is defined by the relation

$$T^{\mathrm{T}} := \sum_{\substack{i=1,\ldots,d_{\mathcal{H}} \\ j=1,\ldots,d_{\mathcal{H}'}}} \langle j'|T|i\rangle |i\rangle\langle j'|. \tag{2.12}$$

Given an operator $T \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$, we define its *kernel* as the vector space formed by all the elements of $\mathcal{H}$ that are mapped by T to the null vector of $\mathcal{H}'$. In order to simplify the notation, we will use the symbol 0 to denote the null vector of a given vector space. The specific context will clearly identify whether 0 is a scalar or a null vector, and, in the second case, the specific

vector space it belongs to. With this convention we can write

$$\ker(T) := \{|\psi\rangle \in \mathcal{H} : T|\psi\rangle = 0\}. \tag{2.13}$$

The *support* of T is the vector space obtained by considering all the vectors in $\mathcal{H}$ that are orthogonal to every element in $\ker(T)$, i.e.,

$$\text{supp}(T) := \{|\psi\rangle \in \mathcal{H} : \forall |\phi\rangle \in \ker(T) \text{ we have } \langle\psi|\phi\rangle = 0\}. \tag{2.14}$$

Finally, the *image* of T is the vector space formed by all the images of the elements of $\mathcal{H}$, obtained via the application of T, i.e.,

$$\text{image}(T) := \{T|\psi\rangle : |\psi\rangle \in \mathcal{H}\}. \tag{2.15}$$

Notice that the notion of image for operators is exactly the same as the one used for ordinary functions. Thus, using the same name does not lead to any inconsistency. The *rank* of T is defined to be the dimension of $\text{image}(T)$, i.e.,

$$\text{rank}(T) := d_{\text{image}(T)}. \tag{2.16}$$

Notice that, while $\ker(T)$ and $\text{supp}(T)$ are subspaces of $\mathcal{H}$, $\text{image}(T)$ is a subspace of $\mathcal{H}'$. Moreover, it is immediate to see that [76]

$$d_{\mathcal{H}} = d_{\ker(T)} + d_{\text{supp}(T)}, \tag{2.17}$$

and,

$$d_{\mathcal{H}} = d_{\ker(T)} + d_{\text{image}(T)}, \tag{2.18}$$

implying $\text{rank}(T) = d_{\text{image}(T)} = d_{\text{supp}(T)}$.

An operator $P \in \mathcal{L}(\mathcal{H})$ acting as the identity on all the elements belonging to its image, is said to be a projector on $\text{image}(P)$. More precisely, given a subspace A of $\mathcal{H}$, a *projector* into A is the unique operator $P \in \mathcal{L}(\mathcal{H})$ satisfying

1. $\text{image}(P) = A$,

2. $P|\psi\rangle = |\psi\rangle$ for all $|\psi\rangle \in A$.

We can use the notation P^A to make explicit the space we project onto. It is also immediate to see that $\text{image}(P) = \text{supp}(P)$. For this reason, it is common to say that P projects onto $\text{supp}(P)$. If $A = \mathcal{H}$, we obtain the identity operator on $\mathcal{H}$, i.e., the unique operator mapping every element to itself. We will use the symbol 1 to denote the *identity operator* on $\mathcal{H}$, i.e., $1 := P^{\mathcal{H}}$. The specific context will clearly identify whether 1 is a scalar or the identity operator, and, in the second case, the specific Hibert space it acts on.

The identity operator allows us to introduce the concept of the *inverse operator*. Given an operator $T \in \mathcal{L}(\mathcal{H})$, we say that is invertible if there exists an operator $T^{-1} \in \mathcal{L}(\mathcal{H})$ such that $T^{-1}T = 1$. If T is invertible, T^{-1} is said to be its inverse. Moreover, when T^{-1} exists, it must be unique and must also satisfy $TT^{-1} = 1$. Recalling that the identity operator is a special kind of projector, the relation $T^{-1}T = 1$ can be written as $T^{-1}T = P^{\mathcal{H}} = P^{\text{supp}(1)}$. This expression is useful to generalize the notion of invertibility for generic operators in $\mathcal{L}(\mathcal{H}, \mathcal{H}')$. In particular, given a $T \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$, we define its *generalized inverse* $T^{-1} \in \mathcal{L}(\mathcal{H}', \mathcal{H})$ as the unique operator satisfying

$$T^{-1}T = P^{\text{supp}(T)}. \tag{2.19}$$

Relation (2.19) shows that the generalized inversion is an inversion on the support of the operator. When the operator is invertible, its generalized inverse clearly coincides with the ordinary inverse. Thus, using the same symbol for both the inverse and the generalized inverse does not lead to any inconsistency.

The concept of inverse operator allows us to define the important class of unitary operators. An operator $U \in \mathcal{L}(\mathcal{H})$ is said to be *unitary* if

1. U is invertible,

2. the inverse of U coincides with its adjoint operator, i.e., $U^{-1} = U^{\dagger}$.

The set of unitary operators acting on the Hilbert space $\mathcal{H}$ will be denoted by $\mathcal{U}(\mathcal{H})$. It is immediate to see that U preserves the induced norm[8] (2.5), i.e., $\|U|\psi\rangle\| = \||\psi\rangle\|$ for every $|\psi\rangle \in \mathcal{H}$. Using the polarization identity [Theorem 4.3.7][27] this fact implies that unitary operators preserve the inner product as well.

Given an operator $T \in \mathcal{L}(\mathcal{H})$, one can compute its trace $\mathrm{Tr}(T)$, which is a complex basis-independent number capturing several important properties of the endomorphism. Formally, the *trace of an operator* in $\mathcal{L}(\mathcal{H})$ can be defined as the linear functional $\mathrm{Tr} : \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ satisfying

1. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ for every $A \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ and $B \in \mathcal{L}(\mathcal{H}', \mathcal{H})$,

2. $\mathrm{Tr}(1) = d_{\mathcal{H}}$.

Property 1. states that the trace is invariant under cyclic permutations. Thus, it is also invariant under unitary conjugation. If $T \in \mathcal{L}(\mathcal{H})$ and $U \in \mathcal{U}(\mathcal{H})$, we can use the unitary operator U to *unitary conjugate* T via the expression $U^{\dagger}TU$. Thus, being invariant under conjugation means that for any $T \in \mathcal{L}(\mathcal{H})$ and $U \in \mathcal{U}(\mathcal{H})$ the relation $\mathrm{Tr}(U^{\dagger}TU) = \mathrm{Tr}(T)$ holds.

We have already pointed out that $\mathcal{L}(\mathcal{H}, \mathcal{H}')$ is a vector space. The trace allows us to define an inner product on such space. In particular, the *Hilbert-Schmidt inner product* is defined as

$$\langle \cdot, \cdot \rangle : \mathcal{L}\left(\mathcal{H}, \mathcal{H}'\right) \times \mathcal{L}\left(\mathcal{H}, \mathcal{H}'\right) \to \mathbb{C}, (A, B) \to \mathrm{Tr}(A^{\dagger}B). \tag{2.20}$$

[8] This means that unitary operators are *isometries*. On the other hand, there exist isometries that are not unitary. From the definition, we immediately see that only surjective isometries correspond to unitary operators.

Thus, $\mathcal{L}(\mathcal{H}, \mathcal{H}')$ is a Hilbert space as well. As described by relation (2.5), we can use the inner product to induce a norm. The *Hilbert-Schmidt norm* is then defined as

$$\|\cdot\| : \mathcal{L}(\mathcal{H}, \mathcal{H}') \to \mathbb{R} : T \to \sqrt{\mathrm{Tr}(T^\dagger T)}. \tag{2.21}$$

Recall that we also use the symbols $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ to indicate the inner product and norm on $\mathcal{H}$. The specific context will clearly identify the meaning of those symbols.

Notice that (2.21) can be written as $\|T\| = \left[\mathrm{Tr}\left(\left[\sqrt{T^\dagger T}\right]^2\right)\right]^{\frac{1}{2}} = \left[\mathrm{Tr}\left(|T|^2\right)\right]^{\frac{1}{2}}$, where we set $|T| := \sqrt{T^\dagger T}$. The generalization of this expression leads to the notion of Schatten norms. For any $p \in [1, \infty)$ we define the *Schatten p-norm* as

$$\|\cdot\|_p : \mathcal{L}(\mathcal{H}, \mathcal{H}') \to \mathbb{R} : T \to \left[\mathrm{Tr}(|T|^p)\right]^{\frac{1}{p}}. \tag{2.22}$$

Notice that the Hilbert-Schmidt norm (2.21) coincides with the Schatten 2-norm. The Schatten 1-norm

$$\|\cdot\|_1 : \mathcal{L}(\mathcal{H}, \mathcal{H}') \to \mathbb{R} : T \to \mathrm{Tr}(|T|) \tag{2.23}$$

is also known as the *trace norm*. One can extend the notion of Schatten norm for $p = \infty$ by considering the limit $p \to \infty$ in (2.22). By doing so, one obtains the definition of the *operator norm*, also known as *infinity norm*, i.e.,

$$\|\cdot\|_\infty : \mathcal{L}(\mathcal{H}, \mathcal{H}') \to \mathbb{R} : T \to \sup_{\substack{|\psi\rangle \in \mathcal{H} \\ \||\psi\rangle\| = 1}} \|T|\psi\rangle\|. \tag{2.24}$$

Schatten norms are *sub-multiplicative*, meaning that, for every $A, B \in \mathcal{L}(\mathcal{H})$

$$\|AB\|_p \leq \|A\|_p \|B\|_p, \tag{2.25}$$

for any $p \in [1, \infty]$.

Recall that $T \in \mathrm{Herm}(\mathcal{H})$ means $T = T^\dagger$. Positive semidefinite operators constitute a special proper subset of the set of Hermitian operators. The set $\mathcal{P}(\mathcal{H}) \subset \mathrm{Herm}(\mathcal{H})$ of *positive semidefinite operators* acting on the Hilbert space $\mathcal{H}$, is formally defined as

$$\mathcal{P}(\mathcal{H}) := \{T \in \mathrm{Herm}(\mathcal{H}) : \langle\psi|T|\psi\rangle \geq 0 \text{ for all } |\psi\rangle \in \mathcal{H}\}. \tag{2.26}$$

If $T \in \mathcal{P}(\mathcal{H})$, we will write $T \succeq 0$. It is immediate to show that $\succeq$ defines a partial order on $\mathrm{Herm}(\mathcal{H})$. In particular, if $A, B \in \mathrm{Herm}(\mathcal{H})$, we say that $A \succeq B$ (or $B \preceq A$) if and only if $A - B \succeq 0$. This partial order relation is known as the *Loewner (partial) order*.

2.1.2 Super-Operators

Given an Hilbert space $\mathcal{H}$, we said that its elements are called kets. Moreover, we have denoted by $\mathcal{L}(\mathcal{H})$ the set of operators from $\mathcal{H}$ onto itself. If $\mathcal{H}'$ is another Hilbert space, we can consider linear maps mapping operators in $\mathcal{L}(\mathcal{H})$ to operators in $\mathcal{L}(\mathcal{H}')$. Those linear maps are called *super-operators*[9] and form the vector space $\mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$. Given a super-operator $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}'))$ we can consider its *adjoint super-operator* $\mathcal{E}^\dagger \in \mathcal{L}(\mathcal{L}(\mathcal{H}'), \mathcal{L}(\mathcal{H}))$, defined by the relation

$$\langle B, \mathcal{E}(A)\rangle = \langle \mathcal{E}^\dagger(B), A\rangle, \tag{2.27}$$

which must hold for every $A \in \mathcal{L}(\mathcal{H})$ and $B \in \mathcal{L}(\mathcal{H}')$. Using Riesz representation theorem [27, Theorem 3.7.7], it is immediate to show that $\mathcal{E}^\dagger$ is unique. We will use the symbol $\circ$ to concatenate super-operators. With $\mathcal{I} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H}))$ we will denote the *identity super-operator* on $\mathcal{L}(\mathcal{H})$.

[9] Notice that super-operators are operators as well. The super- prefix is used to stress the fact that they act on operators.

A super-operator mapping Hermitian operators to Hermitian operators is said to be Hermitian-preserving. In other words, a super-operator $\mathcal{E} \in \mathcal{L}\left(\mathcal{L}\left(\mathcal{H}\right), \mathcal{L}\left(\mathcal{H}'\right)\right)$ is said to be *Hermitian-preserving* if

$$\mathcal{E}(T) \in \mathrm{Herm}(\mathcal{H}'), \tag{2.28}$$

for every Hermitian operator $T \in \mathrm{Herm}(\mathcal{H})$.

Super-operators mapping positive semidefinite operators to positive semidefinite operators are said to be *positive*. However, the concept of positive operator is not robust enough for our purposes. Thus, we need to introduce the notion of completely positive super-operators. A super-operator $\mathcal{E} \in \mathcal{L}\left(\mathcal{L}\left(\mathcal{H}\right), \mathcal{L}\left(\mathcal{H}'\right)\right)$ is said to be *completely positive* (cp) if, for any Hilbert space $\mathcal{H}''$, the super-operator $\mathcal{E} \otimes \mathcal{I} \in \mathcal{L}\left(\mathcal{L}\left(\mathcal{H}\right) \otimes \mathcal{L}\left(\mathcal{H}''\right), \mathcal{L}\left(\mathcal{H}'\right) \otimes \mathcal{L}\left(\mathcal{H}''\right)\right)$ is positive. In other words, if

$$(\mathcal{E} \otimes \mathcal{I})(T) \succeq 0, \tag{2.29}$$

for every $T \in \mathcal{P}(\mathcal{H} \otimes \mathcal{H}'')$. A super-operator $\mathcal{E} \in \mathcal{L}\left(\mathcal{L}\left(\mathcal{H}\right), \mathcal{L}\left(\mathcal{H}'\right)\right)$ is said to be *trace preserving* (tp) if

$$\mathrm{Tr}(\mathcal{E}(T)) = \mathrm{Tr}(T), \tag{2.30}$$

for every $T \in \mathcal{L}\left(\mathcal{H}\right)$. As we will see, *trace preserving completely positive* (tpcp) maps play a fundamental role in quantum information theory. A super-operator $\mathcal{E} \in \mathcal{L}\left(\mathcal{L}\left(\mathcal{H}\right), \mathcal{L}\left(\mathcal{H}'\right)\right)$ is said to be *unital* if it maps the identity operator onto the identity operator, i.e., if

$$\mathcal{E}(1) = 1. \tag{2.31}$$

It is immediate to show the following relation

$$\mathcal{E} \text{ is a tpcp map} \implies \mathcal{E}^{\dagger} \text{ is a cp and unital map.} \tag{2.32}$$

The partial trace is a very common and useful super-operator. Given two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, the *partial trace* $\mathrm{Tr}_{\mathcal{H}'}(\cdot)$ is the tpcp map defined via

$$\mathrm{Tr}_{\mathcal{H}'} : \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}\left(\mathcal{H}'\right) \to \mathcal{L}(\mathcal{H}) : T \to (\mathcal{I} \otimes \mathrm{Tr})(T). \tag{2.33}$$

2.2 Quantum Mechanics in Finite Dimensional Hilbert Spaces

In this section we want to outline how the mathematical concepts that we introduced in the previous section can be specialized and applied to describe the essential ingredients of quantum mechanics. As a remainder, our focus is on the finite-dimensional setting of the theory.

2.2.1 Quantum Systems and Quantum States

In the previous section, we stated that, in quantum mechanics, Hilbert spaces represent one of the most fundamental mathematical objects. The reason is that we model an isolated *quantum system* (or in short *system*) with a Hilbert space $\mathcal{H}$. The *quantum state* of the system (or in short *state*) is described by a unit-trace positive semidefinite operator ρ acting on $\mathcal{H}$, also known as *density operator*. We will denote with $\mathcal{S}(\mathcal{H})$ the *set of quantum states on* $\mathcal{H}$, i.e.,

$$\mathcal{S}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \mathrm{Tr}(\rho) = 1\}. \tag{2.34}$$

If $\mathrm{rank}(\rho) = 1$ the state ρ is *pure*, otherwise ρ is said to be *mixed*. The state $\frac{1}{d_{\mathcal{H}}}$ is known as the *maximally mixed state on* $\mathcal{H}$.

If $\rho \in \mathcal{S}(\mathcal{H})$ is pure, it can be written as $\rho = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is an appropriate ket in $\mathcal{H}$ with unit norm, i.e., $\||\psi\rangle\| = 1$. In such a case, $|\psi\rangle$ is also called a quantum state. Notice that the expression $|\psi\rangle\langle\psi|$ defines a rank-one projector on the ket $|\psi\rangle$.

2.2.2 Quantum Modelling of Classical Systems

Given a sample space Ω and a discrete random variable $Z : \Omega \to \mathcal{Z}$, where $\mathcal{Z}$ is a finite set with cardinality $|\mathcal{Z}| := d_Z$, we can encode a probability mass function $p_Z : \mathcal{Z} \to [0, 1]$ in a density operator ρ. The classical system will be modelled by a d_Z-dimensional Hilbert space Z, also known as *classical register*, and p_Z will be encoded in the quantum state $\rho_Z \in \mathcal{S}(Z)$

$$\rho_Z := \sum_{i=1}^{d_Z} p_Z(i)|i\rangle\langle i|, \tag{2.35}$$

where $(|i\rangle)_{i=1,\ldots,d_Z}$ is an orthonormal basis for the Hilbert space Z. Notice that we used the same letter, i.e., Z, to denote both the random variable and the Hilbert space. This is a common identification, since the two mathematical objects represent the same concept, but it two different mathematical frameworks.

2.2.3 Bipartite Systems

Many interesting properties of quantum mechanics, which make it substantially different from classical physics, arise when considering multiple systems, i.e., *multipartite systems*. For example, *bipartite systems* of the form $\mathcal{H} \otimes \mathcal{H}'$. The quantum state of a multipartite system is said to be a *multipartite state*. For example, the quantum state of a bipartite system is said to be a *bipartite state*. The set $\mathrm{Sep}(\mathcal{H} : \mathcal{H}')$ of separable states is an important subset of $\mathcal{S}(\mathcal{H} \otimes \mathcal{H}')$. A quantum state $\rho \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H}')$ is said to be *separable*, i.e., $\rho \in \mathrm{Sep}(\mathcal{H} : \mathcal{H}')$, if it can be written as

$$\rho = \sum_{i \in I} p_i \sigma^i \otimes \tau^i, \tag{2.36}$$

where I is a finite set, and for every $i \in I$, we have $\sigma^i \in \mathcal{S}(\mathcal{H})$, $\tau^i \in \mathcal{S}(\mathcal{H}')$, $p_i \geq 0$, and $\sum_{i \in I} p_i = 1$. In other words, a separable state is a quantum state that is in the convex hull

of *tensor product states*, i.e., states of the form $\sigma^i \otimes \tau^i$. Notice that the collection $\{p_i\}_{i\in I}$ appearing in (2.36), that we will call a probability distribution, is naturally associated with a probability mass function, i.e., $p : I \to [0,1], i \to p_i$.

Given a quantum state $\rho \in \mathcal{S}(\mathcal{H}\otimes\mathcal{H}')$, we use the partial trace to obtain a local description of the quantum states of the two individual systems. Those states are called *marginal states*. In particular, $\rho_{\mathcal{H}} := \mathrm{Tr}_{\mathcal{H}'}(\rho)$ will denote the marginal of ρ on $\mathcal{H}$, and $\rho_{\mathcal{H}'} := \mathrm{Tr}_{\mathcal{H}}(\rho)$ the marginal of ρ on $\mathcal{H}'$.

Given a quantum state $\rho \in \mathcal{S}(\mathcal{H})$ and a bipartite pure state $|\psi\rangle \in \mathcal{H}\otimes\mathcal{H}'$, $|\psi\rangle$ is said to be *purification* of ρ, if $\rho = \mathrm{Tr}_{\mathcal{H}'}(|\psi\rangle\langle\psi|)$. Using the Schmidt decomposition [69, Theorem 2.7], it is always possible to find such a $|\psi\rangle$ if $\mathcal{H} \cong \mathcal{H}'$. We can generalize this concept by looking for a quantum state $\tilde{\rho} \in \mathcal{S}(\mathcal{H}\otimes\mathcal{H}')$, not necessarily pure, such that $\rho = \mathrm{Tr}_{\mathcal{H}'}(\tilde{\rho})$. In such a case, we will say that $\tilde{\rho}$ is an *extension* of ρ.

We can also consider classical-quantum states, which are bipartite separable states that have a classical and a quantum part. More precisely, a state $\rho \in \mathcal{S}(Z\otimes\mathcal{H})$ is said to be a *classical-quantum state* if it can be written in the form

$$\rho = \sum_{i=1}^{d_Z} p_i |i\rangle\langle i| \otimes \tau^i, \tag{2.37}$$

for a probability distribution $\{p_i\}_{i=1,\ldots,d_Z}$, an orthonormal basis $(|i\rangle)_{i=1,\ldots,d_Z}$ for the Hilbert space Z, and $\tau^i \in \mathcal{S}(\mathcal{H})$ for $i = 1,\ldots,d_Z$. We refer to Z as the classical part, cf., (2.35), of the bipartite classical-quantum system $\mathcal{S}(Z\otimes\mathcal{H})$. We will use the same denomination, i.e., classical-quantum state, to denote also states of the form $\sum_{i=1}^{d_Z} p_i\tau^i \otimes |i\rangle\langle i|$.

Not every state in $\mathcal{S}(\mathcal{H}\otimes\mathcal{H}')$ is a separable state. In other words, $\mathrm{Sep}(\mathcal{H}:\mathcal{H}')$ is a proper subset of $\mathcal{S}(\mathcal{H}\otimes\mathcal{H}')$, i.e.,

$$\mathrm{Sep}(\mathcal{H}:\mathcal{H}') \subsetneq \mathcal{S}(\mathcal{H}\otimes\mathcal{H}'). \tag{2.38}$$

States that are not separable are called *entangled states*. For example, the *maximally entangled state* $\Phi := |\Phi\rangle\langle\Phi|$ on the bipartite system $\mathcal{H} \otimes \mathcal{H}'$, where $\mathcal{H} \cong \mathcal{H}'$, is a pure and not separable state, and is defined via

$$|\Phi\rangle := \frac{1}{\sqrt{d_{\mathcal{H}}}} \sum_{i=1}^{d_{\mathcal{H}}} |i\rangle \otimes |i\rangle, \tag{2.39}$$

where $(|i\rangle)_{i=1,\ldots,d_{\mathcal{H}}}$ is an orthonormal basis for the two Hilbert spaces. Notice that, since $\mathcal{H} \cong \mathcal{H}'$, we can think of $\mathcal{H}'$ as a copy of the Hilbert space $\mathcal{H}$. In fact, one often uses the same letter to label the two systems.

2.2.4 Quantum Channels and the Choi-Jamiołkowski Isomorphism

The evolution of a quantum system is described by the application of a tpcp map on its associated quantum state. Trace preserving completely positive maps are also called *quantum channels* (or in short *channels*). The partial trace (2.33) is an example of a quantum channel. Given a quantum channel $\mathcal{N} : \mathcal{S}(\overline{A}) \to \mathcal{S}(B)$, with $\overline{A}$ and B two Hilbert spaces[10], we can think of $\mathcal{N}$ as a state $J^{\mathcal{N}} \in \mathcal{S}(B \otimes \overline{A}')$, where $\overline{A}'$ is a copy of $\overline{A}$ (thus, in what follows we use the same label $\overline{A}$ for both). This correspondence, known as *channel-state duality*, is realized by the *Choi-Jamiołkowski isomorphism*. The state $J^{\mathcal{N}}$ is known as the *Choi state*, and is given by the expression[11]

$$J^{\mathcal{N}} := (\mathcal{N} \otimes \mathcal{I})(\Phi), \tag{2.40}$$

where $\mathcal{I}$ is the identity channel on $\overline{A}$, and Φ is the maximally entangled state on $\overline{A} \otimes \overline{A}$.

[10] Here we are using the label $\overline{A}$ instead of A, which may seem more logical, to obtain formulas consistent with the notation we will use in the subsequent chapters of this thesis.

[11] Notice that this definition makes sense also for general linear maps. However, in most cases, we use the Choi-Jamiołkowski isomorphism when $\mathcal{N}$ is a quantum channel, i.e., a tpcp map. In such a case, the operator $J^{\mathcal{N}}$ satisfies additional properties that will be discussed in this subsection.

Remark 2.2.1. *In order to simplify the notation, we will use subscripts to keep track of the systems the operators, e.g., quantum states, and super-operators, e.g., quantum channels, act on. For example, if* $|\psi\rangle \in \overline{A} \otimes B$*, and* $\rho \in \mathcal{S}(\overline{A} \otimes B)$*, we will write* $|\psi\rangle_{\overline{A}B}$*, and* $\rho_{\overline{A}B}$*. If* $\mathcal{N} : \mathcal{S}(\overline{A}) \to \mathcal{S}(B)$ *is a quantum channel, we will write* $\mathcal{N}_{\overline{A}\to B}$*. If* $\overline{A} = B$*, we set* $\mathcal{N}_{\overline{A}} := \mathcal{N}_{\overline{A}\to\overline{A}}$*. For example, with this new notation, the defining expression (2.40) becomes*

$$J^{\mathcal{N}}_{B\overline{A}} = \left(\mathcal{N}_{\overline{A}\to B} \otimes \mathcal{I}_{\overline{A}}\right)\left(\Phi_{\overline{A}\,\overline{A}}\right) \tag{2.41}$$

$$= \left(\mathcal{N}_{\overline{A}\to B} \otimes \mathcal{I}_{\overline{A}}\right)\left(\frac{1}{d_{\overline{A}}}\sum_{i,j=1}^{d_{\overline{A}}} |i\rangle\langle j|_{\overline{A}} \otimes |i\rangle\langle j|_{\overline{A}}\right) \tag{2.42}$$

$$= \frac{1}{d_{\overline{A}}}\sum_{i,j=1}^{d_{\overline{A}}} \mathcal{N}_{\overline{A}\to B}\left(|i\rangle\langle j|_{\overline{A}}\right) \otimes |i\rangle\langle j|_{\overline{A}}, \tag{2.43}$$

where $(|i\rangle)_{i=1,\ldots,d_{\overline{A}}}$ *is an orthonormal basis for the two Hilbert spaces. Notice that, for brevity, we can also suppress identity channels and make their presence implicit by the subscripts. For example, we can write* $\mathcal{N}_{\overline{A}\to B}\left(\Phi_{\overline{A}\,\overline{A}}\right)$ *in place of* $\left(\mathcal{N}_{\overline{A}\to B} \otimes \mathcal{I}_{\overline{A}}\right)\left(\Phi_{\overline{A}\,\overline{A}}\right)$*.*

The usage of subscripts is also convenient because it allows us to make implicit any isometry needed to rearrange the underlying Hilbert spaces, e.g., in expressions such as $W_{\overline{A}B}Q_{B\overline{A}}$*, where* $W_{\overline{A}B}$ *is an operator acting on* $\overline{A}\otimes B$*, and* $Q_{B\overline{A}}$ *is an operator acting on* $B\otimes\overline{A}$*. In other words, the expression* $W_{\overline{A}B}Q_{B\overline{A}}$ *must be interpreted as* $W_{\overline{A}B}F_{\overline{A}\leftrightarrow B}Q_{B\overline{A}}$*, where* $F_{\overline{A}\leftrightarrow B} : B\otimes\overline{A} \to \overline{A}\otimes B$ *is the swap operator (or flip operator) exchanging* $\overline{A}$ *with* B*, i.e.,*

$$\left(F_{\overline{A}\leftrightarrow B}\right)\left(W_B \otimes W_{\overline{A}}\right) := W_{\overline{A}} \otimes W_B, \tag{2.44}$$

for every operator $W_{\overline{A}}$ *and* W_B *acting on* $\overline{A}$ *and* B*, respectively. Given the simplicity of (2.44), one often uses the same symbol, i.e.,* $F_{\overline{A}\leftrightarrow B}$*, for both* $F_{\overline{A}\leftrightarrow B}$ *and its inverse operator* $F^{-1}_{\overline{A}\leftrightarrow B} : B\otimes\overline{A} \to \overline{A}\otimes B$*.*

If $\mathcal{N}_{\overline{A}\to B}$ is a quantum channel, it is simple to verify that complete positivity (2.29) and trace preservation (2.30) are translated in the following properties for its Choi state

1. $J^{\mathcal{N}}_{B\overline{A}} \succeq 0$,

2. $\mathrm{Tr}_B\left(J^{\mathcal{N}}_{B\overline{A}}\right) = \frac{1_{\overline{A}}}{d_{\overline{A}}}$.

Tracing out $\overline{A}$ instead of B leads to $\mathrm{Tr}_{\overline{A}}\left(J^{\mathcal{N}}_{B\overline{A}}\right) = \frac{\mathcal{N}_{\overline{A}\to B}\left(1_{\overline{A}}\right)}{d_{\overline{A}}}$. Thus, if $\mathcal{N}_{\overline{A}\to B}$ is unital (2.31), we find

$$\mathrm{Tr}_{\overline{A}}\left(J^{\mathcal{N}}_{B\overline{A}}\right) = \frac{1_B}{d_{\overline{A}}}. \tag{2.45}$$

The inverse of the Choi-Jamiołkowski isomorphism maps Choi states back to quantum channels. Given a quantum state $W_{B\overline{A}}$ with $W_{\overline{A}} = \frac{1_{\overline{A}}}{d_{\overline{A}}}$, its *Choi channel* $\mathcal{N}^W_{\overline{A}\to B}$ is given by the following tpcp map

$$\mathcal{N}^W_{\overline{A}\to B} : \rho_{\overline{A}} \to d_{\overline{A}} \cdot \mathrm{Tr}_{\overline{A}}\left[W_{B\overline{A}}\left(1_B \otimes \rho^{\mathrm{T}}_{\overline{A}}\right)\right], \tag{2.46}$$

where the transpose is taken with respect to the orthonormal basis of the maximally entangled state in (2.41).

Since the Choi state $J^{\mathcal{N}}_{B\overline{A}}$ acts on $B \otimes \overline{A}$, we can multiply $J^{\mathcal{N}}_{B\overline{A}}$ by, for example, tensor product operators of the form $\sigma_B \otimes \tau_{\overline{A}}$. The following lemma is useful to simplify expressions of the form $\mathrm{Tr}\Big[J^{\mathcal{N}}_{B\overline{A}}\left(\sigma_B \otimes \tau_{\overline{A}}\right)\Big]$.

Lemma 2.2.2. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel, $J^{\mathcal{N}}_{B\overline{A}}$ its Choi state, $\sigma_B \in \mathcal{L}(B)$, and $\tau_{\overline{A}} \in \mathcal{L}\left(\overline{A}\right)$, then*

$$\mathrm{Tr}\left[J^{\mathcal{N}}_{B\overline{A}}\left(\sigma_B \otimes \tau_{\overline{A}}\right)\right] = \frac{1}{d_{\overline{A}}}\mathrm{Tr}\left[\sigma_B \mathcal{N}_{\overline{A}\to B}\left(\tau^{\mathrm{T}}_{\overline{A}}\right)\right], \tag{2.47}$$

where the transpose is taken with respect to the orthonormal basis of the maximally entangled state in (2.41).

Proof.

$$\operatorname{Tr}\left[J_{B\overline{A}}^{\mathcal{N}}\left(\sigma_B \otimes \tau_{\overline{A}}\right)\right] = \operatorname{Tr}\left[\left(\frac{1}{d_{\overline{A}}}\sum_{i,j=1}^{d_{\overline{A}}} \mathcal{N}_{\overline{A}\to B}\left(|i\rangle\langle j|_{\overline{A}}\right) \otimes |i\rangle\langle j|_{\overline{A}}\right)\left(\sigma_B \otimes \tau_{\overline{A}}\right)\right] \tag{2.48}$$

$$= \frac{1}{d_{\overline{A}}}\operatorname{Tr}\left[\sum_{i,j=1}^{d_{\overline{A}}} \langle j|\tau_{\overline{A}}|i\rangle \mathcal{N}_{\overline{A}\to B}\left(|i\rangle\langle j|_{\overline{A}}\right)\sigma_B\right] \tag{2.49}$$

$$= \frac{1}{d_{\overline{A}}}\operatorname{Tr}\left[\mathcal{N}_{\overline{A}\to B}\left(\sum_{i,j=1}^{d_{\overline{A}}} \langle j|\tau_{\overline{A}}|i\rangle|i\rangle\langle j|_{\overline{A}}\right)\sigma_B\right] \tag{2.50}$$

$$= \frac{1}{d_{\overline{A}}}\operatorname{Tr}\left[\mathcal{N}_{\overline{A}\to B}\left(\sum_{i,j=1}^{d_{\overline{A}}} \langle i|\tau_{\overline{A}}^{\mathrm{T}}|j\rangle|i\rangle\langle j|_{\overline{A}}\right)\sigma_B\right] \tag{2.51}$$

$$= \frac{1}{d_{\overline{A}}}\operatorname{Tr}\left[\mathcal{N}_{\overline{A}\to B}\left(\tau_{\overline{A}}^{\mathrm{T}}\right)\sigma_B\right] \tag{2.52}$$

$$= \frac{1}{d_{\overline{A}}}\operatorname{Tr}\left[\sigma_B \mathcal{N}_{\overline{A}\to B}\left(\tau_{\overline{A}}^{\mathrm{T}}\right)\right]. \tag{2.53}$$

□

2.2.5 Quantum Measurements

Quantum measurements (or in short *measurements*) are a special case of quantum channels that can be written in the form

$$\mathcal{M}_{A\to Z} : \mathcal{L}(A) \to \mathcal{L}(Z), \rho_A \to \sum_{i=1}^{d_Z} \langle M_A^i, \rho_A\rangle |i\rangle\langle i|_Z \tag{2.54}$$

with an orthonormal basis $(|i\rangle)_{i=1,\ldots,d_Z}$ for the Hilbert space Z, and satisfy the following two properties

1. $M_A^i \succeq 0$ for every $i \in \{1, \ldots, d_Z\}$,
2. $\sum_{i=1}^{d_Z} M_A^i = 1_A$.

Notice that the expression $\sum_{i=1}^{d_Z}\langle M_A^i, \rho_A\rangle|i\rangle\langle i|_Z = \sum_{i=1}^{d_Z} \operatorname{Tr}\left[M_A^i\rho_A\right]|i\rangle\langle i|_Z$ represents the state of a classical system (2.35), described by the probability distribution $\{\langle M_A^i, \rho_A\rangle\}_{i=1,\ldots,d_Z}$. The

type of quantum measurements that we have described is called a *positive operator-valued measure* (POVM), and the operators $\{M_A^i\}_{i=1,\dots,d_Z}$ are known as *POVM elements*.

There exists a special kind of measurement that uniquely determine the state of a quantum system by the measurement statistics they generate. Those measurements are known as *information-complete measurements*.

Definition 2.2.3. *A quantum measurement $\mathcal{M}_{A\to Z} : \mathcal{L}(A) \to \mathcal{L}(Z)$ is said to be informationally complete if its POVM elements $\{M_A^i\}_{i=1,\dots,d_Z}$ span the entire Hilbert space $\mathcal{L}(A)$.*

In other words, $\mathcal{M}_{A\to Z} : \mathcal{L}(A) \to \mathcal{L}(Z)$ is an information-complete measurement if it is an injective map. In such a case, two different quantum states lead to two different classical outcomes probabilities. Informationally complete quantum measurements will play a special role in the proof techniques used in this thesis.

2.2.6 Symmetric States

We use the notation $\mathfrak{S}_n$ to denote the set of permutations acting on n elements (or letters). Its cardinality is $|\mathfrak{S}_n| = n!$. With the permutation composition operation $\circ : \mathfrak{S}_n \to \mathfrak{S}_n$, the algebraic structure $(\mathfrak{S}_n, \circ)$ forms a group, the *symmetric group* of n elements. Clearly the group $(\mathfrak{S}_n, \circ)$ is not abelian, i.e., the composition permutation is a non-commutative operation.

Given a permutation $\pi \in \mathfrak{S}_n$ and a Hilbert space $\mathcal{H}^n$, we indicate with $U_{\mathcal{H}^n}^\pi \in \mathcal{L}(\mathcal{H}^n)$ the *permutation operator* associated with π. In other words, $U_{\mathcal{H}^n}^\pi$ is the operator re-arranging the tensor products of kets according to the rule specified by π. A multipartite ket, e.g., a multipartite pure state, $|\psi\rangle_{\mathcal{H}^n} \in \mathcal{H}^n$ is said to be *permutation invariant*, or *symmetric*, if

$$U_{\mathcal{H}^n}^\pi |\psi\rangle_{\mathcal{H}^n} = |\psi\rangle_{\mathcal{H}^n}, \tag{2.55}$$

for every $\pi \in \mathfrak{S}_n$. Given a multipartite Hilbert space $\mathcal{H}^n$, we indicate with

$$\mathrm{Sym}^n(\mathcal{H}) := \{|\psi\rangle_{\mathcal{H}^n} \in \mathcal{H}^n : U^{\pi}_{\mathcal{H}^n}|\psi\rangle_{\mathcal{H}^n} = |\psi\rangle_{\mathcal{H}^n}, \forall \pi \in \mathfrak{S}_n\}, \tag{2.56}$$

the *symmetric subspace* of $\mathcal{H}^n$. For a review of many quantum information applications of the symmetric subspace see [43].

Given a multipartite Hilbert space $\mathcal{H}^n$ and a permutation operator $U^{\pi}_{\mathcal{H}^n}$, where $\pi \in \mathfrak{S}_n$, we can build the associated *permutation channel* as

$$\mathcal{U}^{\pi}_{\mathcal{H}^n}(\cdot) := U^{\pi}_{\mathcal{H}^n}(\cdot)U^{\pi^{-1}}_{\mathcal{H}^n}, \tag{2.57}$$

where π^{-1} indicates the inverse of the permutation π with respect to the composition operation $\circ$. It is now easy to extend the concept of permutation invariance also to operators on $\mathcal{H}^n$, e.g., mixed states. In fact, a multipartite operator $\rho_{\mathcal{H}^n} \in \mathcal{L}(\mathcal{H}^n)$ is said to be *permutationally invariant*, or *symmetric*, if

$$\mathcal{U}^{\pi}_{\mathcal{H}^n}(\rho_{\mathcal{H}^n}) = \rho_{\mathcal{H}^n}, \tag{2.58}$$

for every $\pi \in \mathfrak{S}_n$.

Finally, given the Hilbert spaces Q and $\mathcal{H}$, a multipartite operator $\rho_{Q\mathcal{H}^n} \in \mathcal{L}(Q \otimes \mathcal{H}^n)$, is said to be *symmetric with respect to* Q, if it is invariant under permutation of the $\mathcal{H}$-systems keeping Q fixed, i.e., if

$$(\mathcal{I}_Q \otimes \mathcal{U}^{\pi}_{\mathcal{H}^n})(\rho_{Q\mathcal{H}^n}) = \rho_{Q\mathcal{H}^n}, \tag{2.59}$$

for every $\pi \in \mathfrak{S}_n$. A bipartite state $\rho_{Q\mathcal{H}} \in \mathcal{S}(Q \otimes \mathcal{H})$ is said to be *n-extendible* if there exists a multipartite extension $\rho_{Q\mathcal{H}^n} \in \mathcal{S}(Q \otimes \mathcal{H}^n)$, i.e., $\mathrm{Tr}_{\mathcal{H}_2^n}(\rho_{Q\mathcal{H}^n}) = \rho_{Q\mathcal{H}}$, that is symmetric with respect to Q. The notion of n-extendibility will be extremely important for this thesis.

2.2.7 Diamond Norm

In order to quantify the distance between quantum states, we can use the metric induced by one of the Schatten p-norms (2.22), (2.24). Those norms can be used also to define norms for quantum channels. For example, given a quantum channel $\mathcal{N}_{\overline{A}\to B}$, the trace norm $\|\cdot\|_1$ is used to define the *diamond norm*

$$\|\mathcal{N}_{\overline{A}\to B}\|_\diamond := \sup_{\substack{X\in\overline{A}\otimes\overline{A}\\ \|X\|_1\leq 1}} \|(\mathcal{N}_{\overline{A}\to B}\otimes\mathcal{I}_{\overline{A}})(X_{\overline{A}\,\overline{A}})\|_1, \tag{2.60}$$

which is a popular norm used to quantify the distance between quantum channels via its induced metric.

The following lemma relates the trace norm of Choi states to the diamond norm of their isomorphically associated quantum channels.

Lemma 2.2.4. *[82, Lemma 7] Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel, $J^{\mathcal{N}}_{B\overline{A}}$ its Choi state, then*

$$\|J^{\mathcal{N}}_{B\overline{A}}\|_1 \leq \|\mathcal{N}_{\overline{A}\to B}\|_\diamond \leq d_{\overline{A}}\|J^{\mathcal{N}}_{B\overline{A}}\|_1. \tag{2.61}$$

Notice that, even if we have stated the above lemma for quantum channels, it holds even if $\mathcal{N}_{\overline{A}\to B}$ is a generic Hermitian-preserving super-operator (2.28).

2.3 Semidefinite Programming

In this section we introduce the essential concepts from semidefinite programming that are needed for this thesis.

There are three ingredients that are used to specify a *semidefinite program* (SDP)

1. A Hermitian-preserving super-operator $\Phi_{A\to B} : \mathcal{L}(A) \to \mathcal{L}(B)$,

2. a "primal" Hermitian operator $P_A \in \mathrm{Herm}(A)$,

3. a "dual" Hermitian operator $D_B \in \mathrm{Herm}(B)$.

Thus, a semidefinite program can be seen as the triple $(\Phi_{A\to B}, P_A, D_B)$. Notice that, in practice, semidefinite programs are stated in a simplified, less formal way. However, it is always possible to reformulate any semidefinite program in the form $(\Phi_{A\to B}, P_A, D_B)$ to fit within the above framework.

One then defines the *primal problem* as

$$\alpha := \sup \quad \langle P_A, X_A \rangle \tag{2.62}$$

$$s.t. \quad \Phi_{A\to B}(X_A) = D_B, \tag{2.63}$$

$$X_A \succeq 0. \tag{2.64}$$

Every primal problem has an associated *dual problem*, defined as

$$\beta := \inf \quad \langle D_B, Y_B \rangle \tag{2.65}$$

$$s.t. \quad \Phi^{\dagger}_{B\to A}(Y_B) \succeq P_A, \tag{2.66}$$

$$Y_B \in \mathrm{Herm}(B). \tag{2.67}$$

Operators satisfying the constraints of an optimization program are said to be *feasible operators*. Thus, an operator $X_A \in \mathcal{L}(A)$ satisfying $\Phi_{A\to B}(X_A) = D_B$, and $X_A \succeq 0$, is said to be *primal feasible*. The set formed by all the primal feasible operators is the *primal feasible set* $\mathcal{A}$

$$\mathcal{A} := \{X_A \in \mathcal{L}(A) : \Phi_{A\to B}(X_A) = D_B, X_A \succeq 0\}. \tag{2.68}$$

As we see, the *primal optimum value* α is the supremum of the *primal objective function* $X_A \to \langle P_A, X_A \rangle$ over the primal feasible set $\mathcal{A}$. If there are no primal feasible operators, i.e., if $\mathcal{A}$ is the empty set $\varnothing$, then we define $\alpha := -\infty$.

Similarly, an operator $Y_B \in \mathcal{L}(B)$ satisfying $\Phi^\dagger_{B\to A}(Y_B) \succeq P_A$, and $Y_B \in \mathrm{Herm}(B)$, is said to be *dual feasible*. The set formed by all the dual feasible operators is the *dual feasible set* $\mathcal{B}$

$$\mathcal{B} := \{Y_B \in \mathcal{L}(B) : \Phi^\dagger_{B\to A}(Y_B) \succeq P_A, Y_B \in \mathrm{Herm}(B)\}. \tag{2.69}$$

The *dual optimum value* β is the infimum of the *dual objective function* $Y_B \to \langle D_B, Y_B\rangle$ over the dual feasible set $\mathcal{B}$. If there are no dual feasible operators, i.e., if $\mathcal{B}$ is the empty set $\varnothing$, then we define $\beta := +\infty$.

Duality relations establish connections between the primal optimum value α and the dual optimum value β. The first duality relation, which always hold, states that the primal optimum value is always less than or equal to the dual optimum value.

Theorem 2.3.1. (Weak Duality) *If* $(\Phi_{A\to B}, P_A, D_B)$ *is a SDP, then*

$$\alpha \leq \beta. \tag{2.70}$$

The equality condition $\alpha = \beta$, which does not necessarily always hold, is known as *strong duality*. Strong duality is not a rare condition in common practical applications of semidefinite programming. Nevertheless, the following theorem provides a set of sufficient conditions for strong duality

Theorem 2.3.2. (Slater's theorem) *If* $(\Phi_{A\to B}, P_A, D_B)$ *is a SDP, then* $\alpha = \beta$ *if one of the following conditions holds*

1. α *is finite and there exists a dual feasible operator* $Y_B \in \mathcal{B}$ *such that* $\Phi^\dagger_{B\to A}(Y_B) \succ P_A$,

2. β *is finite and there exists a primal feasible operator* $X_A \in \mathcal{A}$ *such that* $X_A \succ 0$.

If 1. holds, then there exists a primal feasible operator $X_A \in \mathcal{A}$ *achieving the primal optimum value, i.e.,* $\langle P_A, X_A\rangle = \alpha$. *If 2. holds, then there exists a dual feasible operator* $Y_B \in \mathcal{B}$ *achieving the dual optimum value, i.e.,* $\langle D_B, Y_B\rangle = \beta$.

Chapter 3

De Finetti Theorems with Linear Constraints

The primary motivation of de Finetti theorems is to represent, or approximate, mathematical objects symmetric under permutations of their components into a probabilistic ensemble of elementary independent and identically distributed (i.i.d.) constituents. In the classical case, the mathematical objects are probability mass functions, and the related theorems are known as finite classical de Finetti theorems [28]. In the quantum case, the mathematical objects are quantum states, and the related theorems are known as finite quantum de Finetti theorems [59]. Infinite version of those theorems are known in the literature, and they give exact alternative representations for the desired mathematical object [19]. The infinite de Finetti representation theorems can be found as limits of the finite versions, which are typically stated in the form of upper bounds to the approximation error. In this work, we consider finite versions of de Finetti theorems, while their generalization to the infinite case is easily obtained by taking the asymptotic limit.

3.1 Classical De Finetti theorem

De Finetti [26] first proved the classical version of the theorem in the asymptotic limit $n \to \infty$. The classical de Finetti theorem is based on the notion of exchangeability for random variables (or their joint probability distribution).

Definition 3.1.1. *A collection of n discrete random variables $X_1, \ldots, X_n$ is said to be symmetric, or finitely exchangeable, if the joint probability mass function $p_{X_1,\ldots,X_n}$ is invariant under permutation of its arguments, i.e., if*

$$p_{X_1,\ldots,X_n}(x_{\pi(1)}, \ldots, x_{\pi(n)}) = p_{X_1,\ldots,X_n}(x_1, \ldots, x_n), \tag{3.1}$$

for every $\pi \in \mathfrak{S}_n$, and $x_i \in \text{image}(X)$ for $i = 1, \ldots, n$, where $\text{image}(X) := \text{image}(X_1)$.

If the collection $X_1, \ldots, X_n$ can be seen as arising from an infinite sequence of symmetric random variable, it is said to be *exchangeable* (or *infinitely exchangeable*). More formally, the joint probability mass function $p_{X_1,\ldots,X_n}$ must satisfy the following two conditions

1. it must be invariant under permutation of its arguments,
2. it can be seen as the marginal probability mass function of a symmetric probability mass function of arbitrarily many random variables.

Condition 2. means that there exists a probability mass function $p_{X_1,\ldots,X_n,X_{n+1},\ldots,X_{n+m}}$ that is symmetric, and such that

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \sum_{x_{n+1},\ldots,x_{n+m}} p_{X_1,\ldots,X_n,X_{n+1},\ldots,X_{n+m}}(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m}), \tag{3.2}$$

for every $m > 0$, and the sum is over $[\text{image}(X)]^{\times m}$.

I.i.d. implies exchangeability, as shown by the following lemma.

Lemma 3.1.2. *If $X_1, \ldots, X_n$ is a collection of n i.i.d. discrete random variables, then $X_1, \ldots, X_n$ is finitely exchangeable.*

Proof.

$$p_{X_1,\ldots,X_n}(x_{\pi(1)}, \ldots, x_{\pi(n)}) = \prod_{i=1}^{n} p_{X_i}(x_{\pi(i)}) \tag{3.3}$$

$$= \prod_{i=1}^{n} p_{X_i}(x_i) \tag{3.4}$$

$$= p_{X_1,\ldots,X_n}(x_1, \ldots, x_n). \tag{3.5}$$

□

Example 3.1.3. (Coin Tossing) *Consider a classical experiment where a coin, not necessarily fair, is tossed repeatedly n times, and let X_i denote the outcome of the ith toss, for $i = 1, \ldots, n$. Clearly, the collection $X_1, \ldots, X_n$ is i.i.d., so it is finitely exchangeable.*

We have proved that exchangeability is implied by i.i.d. The opposite is not true. In other words, exchangeability is weaker than the concept of i.i.d. This can be shown with the following example which makes use of an urn model.

Example 3.1.4. (Pólya's Urn Scheme [35, Chapter V]) *Consider an urn containing $N_B > 0$ black and $N_R > 0$ red balls. A balls is drawn at random and is replaced. Moreover, $n_+ > 0$ extra balls of the color drawn are also added to the urn. Another ball is drawn, and the process continues as described above for a total of n drafts. Let X_i denote the outcome of the ith draft. The probability of drawing n_B black and then n_R red balls, where $n_B + n_R = n$, is given by*

$$\frac{N_B(N_B + n_+) \cdots (N_B + (n_B - 1) \cdot n_+) N_R(N_R + n_+) \cdots (N_R + (n_R - 1) \cdot n_+)}{(N_B + N_R)(N_B + N_R + n_+) \cdots (N_B + N_R + (n-1) \cdot n_+)}. \tag{3.6}$$

The key point is that the above is also the probability of any sequence with n_B black and n_R red balls, independently on the order in which they have been drawn. In fact, the probability of any

such sequence will clearly have the same denominator as in (3.6), and the factors appearing in the numerators will also be the same ones as in (3.6), but arranged in a different order. Thus, the collection $X_1, \ldots, X_n$ *is finitely exchangeable. On the other hand, it is self-evident that* $X_1, \ldots, X_n$ *are not i.i.d. random variables, since the composition of the urn changes after every draft. For example,* $\Pr(X_2 = \text{red}|X_1 = \text{black}) = \frac{N_R}{N_B+N_R+n_+} \neq \frac{N_R}{N_B+N_R} = \Pr(X_2 = \text{red})$*, since* $n_+ > 0$*.*

We have seen that exchangeability and i.i.d. are different concepts. In particular, i.i.d. random variables are exchangeable but exchangeable random variables do not need to be i.i.d. However, the classical de Finetti theorem [26] shows that, in some sense, a notion of independence is applicable also in the context of exchangeability.

Theorem 3.1.5. *A collection* $X_1, \ldots, X_n$ *of discrete random variables is exchangeable if and only if there exists a parameter* θ *and a measure* P *on it, such that the joint probability mass function* $p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ *can be written as*

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \int \prod_{i=1}^{n} p(x_i|\theta) P(d\theta). \tag{3.7}$$

The classical de Finetti theorem states that a collection $X_1, \ldots, X_n$ of discrete random variables is exchangeable if and only its joint probability mass function can be represented as an "integral mixture" of likelihoods that are conditionally independent with respect to a parameter θ, having P as prior. In other words, if the observations of an experiment are exchangeable, they can be seen as a random sample from some model, which is determined by a parameter θ, and there must exist a prior probability distribution over θ.

3.2 Quantum De Finetti theorem

Similar to the previous section, a quantum state $\rho_{\mathcal{H}^n} \in \mathcal{S}(\mathcal{H}^n)$ is said to be *exchangeable* (or *infinitely exchangeable*) if it satisfies the following two conditions $\rho_{\mathcal{H}^n} \in \mathcal{S}(\mathcal{H}^n)$

1. it must be symmetric (2.58),
2. it can be seen as the marginal density operator of a symmetric quantum state acting on arbitrarily many quantum systems.

Condition 2. means that $\rho_{\mathcal{H}^n}$ admits a symmetric extension $\rho_{\mathcal{H}^{n+m}} \in \mathcal{S}(\mathcal{H}^{n+m})$ for every $m > 0$. The following is the quantum version [19, Section III] of the classical de Finetti theorem (3.1.5).

Theorem 3.2.1. *A quantum state $\rho_{\mathcal{H}^n} \in \mathcal{S}(\mathcal{H}^n)$ is exchangeable if and only if it can be written in the form*

$$\rho_{\mathcal{H}^n} = \int_{\mathcal{S}(\mathcal{H})} P(\rho)\rho^{\otimes n} d\rho, \tag{3.8}$$

where P is a unique probability density function over $\mathcal{S}(\mathcal{H})$ and $d\rho$ is a measure on that set.

In analogy with its classical version, we see that the quantum de Finetti theorem states that a quantum state $\rho_{\mathcal{H}^n} \in \mathcal{S}(\mathcal{H}^n)$ is exchangeable if and only if it can be represented as an "integral mixture" of separable i.i.d. states.

3.3 Approximating Separable States with PPT States

While the quantum de Finetti Theorem 3.2.1 is an interesting result, it is not in the form we need for practical applications. In particular, approximating the set $\mathrm{Sep}(A : B)$ of separable

states is a ubiquitous but computationally hard problem in quantum information theory (see, e.g., [6, Section 9.1] for hardness results with respect to approximations). The set $\mathrm{Sep}(A:B)$ has been formally defined in Subsection 2.2.3. In particular, we said that $\rho_{AB} \in \mathrm{Sep}(A:B)$ if it can be written as $\rho_{AB} = \sum_{i\in I} p_i \sigma_A^i \otimes \tau_B^i$, for a probability distribution $\{p_i\}_{i\in I}$, and quantum states $\{\sigma_A^i\}_{i\in I}$, and $\{\tau_B^i\}_{i\in I}$. Since elements in $\mathrm{Sep}(A:B)$ describe unentangled states, being able to characterize $\mathrm{Sep}(A:B)$ is extremely important in order to understand entanglement, which is one of the main features of quantum mechanics. Operationally speaking, the characterization of $\mathrm{Sep}(A:B)$ is connected to the formulation of *separability tests*. I.e., criteria that can be used to check whether a given quantum state ρ_{AB} is separable or not. One of the most famous separability tests is based on the following result by Horodecki [50].

Theorem 3.3.1. (Horodecki criterion) *A quantum state $\rho_{AB} \in \mathcal{S}(A\otimes B)$ is separable, i.e., $\rho_{AB} \in Sep(A:B)$, if and only if*

$$(\Phi_{A\to C} \otimes \mathcal{I}_B)(\rho_{AB}) \succeq 0, \tag{3.9}$$

for every Hilbert space C, and positive super-operator $\Phi_{A\to C} : \mathcal{L}(A) \to \mathcal{L}(C)$.

Notice that in the above theorem we can choose $C = A$. The Horodecki criterion can be directly used to test separability. If we find a positive super-operator $\Phi_{A\to A}$ such that $(\Phi_{A\to A} \otimes \mathcal{I}_B)(\rho_{AB}) \not\succeq 0$, then we conclude that ρ_{AB} is not separable. Vice versa, it is clear that we cannot try all the possible positive super-operators to prove that ρ_{AB} is separable. Thus, in practice, one restricts the analysis to well-known positive super-operators that are easy to compute. The most common positive super-operator used to test separability in the

context of the Horodecki criterion is the transpose map[1]

$$\text{Transpose} : X \to X^{\mathrm{T}}, \tag{3.10}$$

where the transpose is taken with respect to an orthonormal basis of the Hilbert space. Moreover, we define

$$(\cdot)^{\mathrm{T}_A} := (\text{Transpose}_A \otimes \mathcal{I}_B)(\cdot), \tag{3.11}$$

$$(\cdot)^{\mathrm{T}_B} := (\mathcal{I}_A \otimes \text{Transpose}_B)(\cdot), \tag{3.12}$$

which are known as *partial transposes*. Given the importance of partial transposition to test separability, the following definition is natural.

Definition 3.3.2. *A quantum state $\rho_{AB} \in \mathcal{S}(A \otimes B)$ is said to be a positive partial transpose (PPT) state if*

$$\rho_{AB}^{\mathrm{T}_A} \succeq 0. \tag{3.13}$$

We will denote with $\mathrm{PPT}(A:B)$ the set of PPT states[2], i.e.,

$$\mathrm{PPT}(A:B) := \{\rho_{AB} \in \mathcal{S}(A \otimes B) : \rho_{AB}^{\mathrm{T}_A} \succeq 0\}. \tag{3.14}$$

The following corollary is a direct consequence of the Horodecki criterion.

Corollary 3.3.3. *Let $\rho_{AB} \in \mathcal{S}(A \otimes B)$ be a quantum state, then*

$$\rho_{AB} \in \mathrm{Sep}(A:B) \implies \rho_{AB} \in \mathrm{PPT}(A:B). \tag{3.15}$$

[1]The transpose map is clearly a positive super-operator since the spectrum of an operator is invariant under transposition. In other words, X and X^{T} share the same eigenvalues.

[2]Notice that this definition is independent on the choice of the system to be transposed. In fact, $\rho_{AB}^{\mathrm{T}_A}$ and $\left(\rho_{AB}^{\mathrm{T}_A}\right)^{\mathrm{T}} = \rho_{AB}^{\mathrm{T}_B}$ share the same spectrum. Thus, $\rho_{AB}^{\mathrm{T}_A} \succeq 0$ if and only if $\rho_{AB}^{\mathrm{T}_B} \succeq 0$.

The above corollary shows that $\mathrm{Sep}(A:B)$ is a subset of $\mathrm{PPT}(A:B)$, i.e., every separable state is also PPT. On the other hand, one can prove that there exist PPT states that are not separable, i.e., entangled PPT states[3]. In other words, $\mathrm{Sep}(A:B)$ is a proper subset of $\mathrm{PPT}(A:B)$

$$\mathrm{Sep}(A:B) \subsetneq \mathrm{PPT}(A:B). \tag{3.16}$$

While entangled PPT states in general do exist, their entanglement is highly constrained by the PPT condition. In some sense, they are the "most classical" of the entangled states, and they exhibit similar properties to separable states. For example, PPT states cannot be "too entangled", in the sense that their overlap, as measured by the inner product, with the maximally entangled state is small ([87, Proposition 6.42]).

3.4 Approximating Separable States with n-extendible States

Another approach for the approximation of $\mathrm{Sep}(A:B)$ is via the notion of n-extendibility introduced in Subsection 2.2.6. Recall that a quantum state $\rho_{AB} \in \mathcal{S}(A \otimes B)$ is said to be n-extendible if there exists a multipartite extension $\rho_{AB_1^n} \in \mathcal{S}(A \otimes B^n)$ that is symmetric with respect to A. In other words, if there exists a quantum state $\rho_{AB_1^n} \in \mathcal{S}(A \otimes B^n)$ satisfying the following two conditions

1. $\mathrm{Tr}_{B_2^n}(\rho_{AB_1^n}) = \rho_{AB}$,

2. $(\mathcal{I}_A \otimes \mathcal{U}^{\pi}_{B^n})(\rho_{AB_1^n}) = \rho_{AB_1^n}$ for every $\pi \in \mathfrak{S}_n$,

with $B_1 := B$. In such a case, we will say that ρ_{AB} is an element of the set of n-extendible

[3]It is interesting to note that entangled PPT states do *not* exist for 2×2, 2×3, and 3×2 bipartite systems. For those low-dimensional quantum systems, every PPT state is also separable [50, Theorem 3].

states, which will be denoted by n-Ext$(A:B)$. It is easy to show that, given a natural number $n > 0$, every separable state is n-extendible. This is formalized by the following proposition.

Proposition 3.4.1. *Let $\rho_{AB} \in \mathcal{S}(A \otimes B)$ be a quantum state, then*

$$\rho_{AB} \in \text{Sep}(A:B) \implies \rho_{AB} \in n\text{-Ext}(A:B), \tag{3.17}$$

for every natural $n > 0$.

Proof. If ρ_{AB} is separable, it can be written as $\rho_{AB} = \sum_{i\in I} p_i \sigma_A^i \otimes \tau_B^i$, for a probability distribution $\{p_i\}_{i\in I}$, and quantum states $\{\sigma_A^i\}_{i\in I}$, and $\{\tau_B^i\}_{i\in I}$. Let's consider the following state

$$\rho_{AB_1^n} := \sum_{i\in I} p_i \sigma_A^i \otimes (\tau_B^i)^{\otimes n}. \tag{3.18}$$

The state $\rho_{AB_1^n}$ is clearly an extension of ρ_{AB}, in fact

$$\begin{aligned}
\text{Tr}_{B_2^n}(\rho_{AB_1^n}) &= \sum_{i\in I} p_i \sigma_A^i \otimes \text{Tr}_{B_2^n}\left[(\tau_B^i)^{\otimes n}\right] && (3.19)\\
&= \sum_{i\in I} p_i \sigma_A^i \otimes \tau_B^i \otimes \text{Tr}(\tau_B^i)^{n-1} && (3.20)\\
&= \sum_{i\in I} p_i \sigma_A^i \otimes \tau_B^i && (3.21)\\
&= \rho_{AB}, && (3.22)
\end{aligned}$$

where we used the normalization condition $\text{Tr}(\tau_B^i) = 1$ for every $i \in I$. Moreover, it is clear that $\rho_{AB_1^n}$ is symmetric with respect to A. This is because $(\tau_B^i)^{\otimes n}$ is formed by n copies of the same quantum state. Permuting those density operators will not change the overall quantum state $\rho_{AB_1^n}$. □

The above proposition shows that, given a $n > 0$, Sep$(A:B)$ is a subset of n-Ext$(A:B)$, i.e., every separable state is also n-extendible. On the other hand, one can prove that there

exist n-extendible states that are not separable [24], [52]. In other words, for any given $n > 0$, $\mathrm{Sep}(A:B)$ is a proper subset of $n\text{-Ext}(A:B)$

$$\mathrm{Sep}(A:B) \subsetneq n\text{-Ext}(A:B). \tag{3.23}$$

Comparing (3.23) with (3.16), we see that both $\mathrm{PPT}(A:B)$ and $n\text{-Ext}(A:B)$ provide outer approximations to the set $\mathrm{Sep}(A:B)$ of separable states. The great advantage in considering n-extendibility is that, while there exist n-extendible states that are not separable, a quantum state that is n-extendible for any n must be separable. This statement is formalized by the following theorem.

Theorem 3.4.2. *[30, Theorem 1] Let $\rho_{AB} \in \mathcal{S}(A \otimes B)$ be a quantum state, then*

$$\rho_{AB} \in \mathrm{Sep}(A:B) \iff \rho_{AB} \in n\text{-Ext}(A:B), \tag{3.24}$$

for every natural $n > 0$.

Notice that the direction $\implies$ in Theorem 3.4.2 is proven by Proposition 3.4.1.

On the other hand, the proof of the opposite direction $\impliedby$ is not trivial. Theorem 3.4.2 naturally leads to a test for separability based on n-extendibility. However, we find ourselves in the same practical problem we encountered with the Horodecki criterion in the previous section. To prove that a quantum state ρ_{AB} is separable, we have to show its n-extendibility for every $n > 0$. In practice, this is impossible. On the other hand, we can stop at a certain n and quantify the "error" we commit by working with n-extendible states in place of separable states. Crucially, n-extendibility has a *semidefinite representation* and this then immediately gives efficient semidefinite approximations of the set $\mathrm{Sep}(A:B)$ for any fixed n. Finite *finite quantum de Finetti theorems* quantify, with upper bounds, the distance of n-extendible states to separable states [21], with convergence in the limit $n \to \infty$ [75]. More precisely, [21, Theorem

II.7] gives that for states ρ_{AB} n-extendible to $\rho_{AB_1^n}$, there exists a probability distribution $\{p_i\}_{i\in I}$ and states ρ_A^i, ρ_B^i on A and B, respectively, such that

$$\left\| \rho_{AB} - \sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i \right\|_1 \leq \frac{2d_B^2}{n}. \tag{3.25}$$

As pointed out, in the limit $n \to \infty$ the distance between separable states and n-extendible states shrinks to zero. In other words, separable states and ∞-extendible states are the same thing. Moreover, inequality (3.25) can be generalized to [21, 57]

$$\left\| \rho_{AB_1^k} - \sum_{i\in I} p_i \rho_A^i \otimes \left(\rho_B^i\right)^{\otimes k} \right\|_1 \leq \frac{2kd_B^2}{n}, \tag{3.26}$$

for $k \in \{1, \cdots, n-1\}$, which represents the state-of-the-art bound. Inequalities of the form (3.26) will be referred as *generalized finite quantum de Finetti theorems*. Those results state that if a multipartite state on AB_1^n is symmetric with respect to A, then the reduced state on the first k systems AB_1^k is close to a separable mixture of independent and identical states for k sufficiently smaller than n. Notice that the dependence of the approximation error on k is linear, meaning that it does not slows down the convergence "too much". Again, in the asymptotic limit $n \to \infty$ and holding k constant, the inequality reduces to an equality and the approximation becomes exact. The special case $k = 1$ exactly recovers (3.25), which characterizes $\mathrm{Sep}(A : B)$. For our setting, however, we are interested more generally in characterizing bipartite states that are separable, but subject to linear constraints on the ρ_A^i, ρ_B^i as well.

3.5 Constrained Bilinear Optimization

As pointed out at the end of the previous section, we are interested in the study of constrained bilinear optimization problems of the form

$$Q = \max \quad \mathrm{Tr}\big[H(D \otimes E)\big] \tag{3.27}$$

$$s.t. \quad D \in \mathcal{P}_D = \Pi_{A\to D}\left(\mathcal{S}(\mathcal{A}) \cap \mathrm{Aff}_A\right) \tag{3.28}$$

$$E \in \mathcal{P}_E = \Pi_{B\to E}\left(\mathcal{S}(\mathcal{B}) \cap \mathrm{Aff}_B\right), \tag{3.29}$$

where $H \in \mathcal{L}(\mathcal{H}_D \otimes \mathcal{H}_E)$ is a fixed operator, and $\mathcal{P}_D$ and $\mathcal{P}_E$ are positive semidefinite representable sets such that

- $\Pi_{A\to D} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_D)$ and $\Pi_{B\to E} : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_E)$ are super-operators,
- Aff_A and Aff_B are affine subspaces of $\mathcal{L}(\mathcal{H}_A)$ and $\mathcal{L}(\mathcal{H}_B)$, respectively.

As we see, the optimization is performed over the set of operators of the form $D \otimes E$, where D and E must represent proper quantum states, subject to additional linear constraints implicitely specified by the affine subspaces Aff_A and Aff_B. In this thesis we will use the generic expression "linear constraints" to include affine constraints as well.

Our main motivation to study problems of the form (3.27) comes from quantum information theory, or more specifically from the problem of *approximate quantum error correction*. We present this application and its motivation in detail in Chapter 4, but continue here with the general mathematical setting.

To discuss our approach, we first rewrite (3.27) by defining $G_{AB} := (\Pi^\dagger_{D\to A} \otimes \Pi^\dagger_{E\to B})(H)$. This leads to the form

$$Q = \max \quad \mathrm{Tr}\big[G_{AB}(\rho_A \otimes \rho_B)\big] \tag{3.30}$$

$$s.t. \quad \rho_A \succeq 0, \quad \rho_B \succeq 0 \tag{3.31}$$

$$\mathrm{Tr}(\rho_A) = \mathrm{Tr}(\rho_B) = 1 \tag{3.32}$$

$$\Lambda_{A\to C_A}(\rho_A) = X_{C_A}, \quad \Gamma_{B\to C_B}(\rho_B) = Y_{C_B}, \tag{3.33}$$

where $G_{AB} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a fixed operator, $\Lambda_{A\to C_A} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{C_A})$, and $\Gamma_{B\to C_B} : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{C_B})$ are super-operators, and $X_{C_A} \in \mathcal{L}(\mathcal{H}_{C_A})$ and $Y_{C_B} \in \mathcal{L}(\mathcal{H}_{C_B})$ are the operators defining Aff_A and Aff_B as the affine subspaces associated with the kernels of the linear maps $\Lambda_{A\to C_A}$ and $\Gamma_{B\to C_B}$, respectively.

Now, by the linearity of the objective function we can equivalently optimise over the convex hull of feasible points

$$Q = \max \quad \mathrm{Tr}\left[G_{AB}\left(\sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i\right)\right] \tag{3.34}$$

$$s.t. \quad p_i \geq 0 \quad \forall i \in I, \quad \sum_{i\in I} p_i = 1 \tag{3.35}$$

$$\rho_A^i \succeq 0, \quad \rho_B^i \succeq 0 \quad \forall i \in I \tag{3.36}$$

$$\mathrm{Tr}\left(\rho_A^i\right) = \mathrm{Tr}\left(\rho_B^i\right) = 1 \quad \forall i \in I \tag{3.37}$$

$$\Lambda_{A\to C_A}\left(\rho_A^i\right) = X_{C_A}, \quad \Gamma_{B\to C_B}\left(\rho_B^i\right) = Y_{C_B} \quad \forall i \in I. \tag{3.38}$$

In fact, making the constraints of the optimization programs implicit, on one hand it is clear that

$$Q = \max \mathrm{Tr}\left[G_{AB}(\rho_A \otimes \rho_B)\right] \leq \max \mathrm{Tr}\left[G_{AB}\left(\sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i\right)\right]. \tag{3.39}$$

On the other hand, for every quantum state ρ_A^i and ρ_B^i, we have

$$\mathrm{Tr}\left[G_{AB}\left(\sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i\right)\right] = \sum_{i\in I} p_i \mathrm{Tr}\left[G_{AB}\left(\rho_A^i \otimes \rho_B^i\right)\right] \tag{3.40}$$

$$\leq \sum_{i\in I} p_i \max \mathrm{Tr}\big[G_{AB}(\rho_A^i \otimes \rho_B^i)\big] \tag{3.41}$$

$$= \sum_{i\in I} p_i Q \tag{3.42}$$

$$= Q. \tag{3.43}$$

Taking the maximum, we find

$$\max \mathrm{Tr}\left[G_{AB}\left(\sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i\right)\right] \leq Q. \tag{3.44}$$

Thus, combining (3.39) with (3.44),

$$Q \leq \max \mathrm{Tr}\left[G_{AB}\left(\sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i\right)\right] \leq Q \tag{3.45}$$

$$\implies \max \mathrm{Tr}\left[G_{AB}\left(\sum_{i\in I} p_i \rho_A^i \otimes \rho_B^i\right)\right] = Q, \tag{3.46}$$

which proves the equivalence of the two optimization programs (3.30) and (3.34).

As we see, the transformed program (3.34) requires an optimization over a subset of $\mathrm{Sep}(A:B)$, specified by the additional linear constraints imposed via the linear maps $\Lambda_{A\to C_A}$, $\Gamma_{B\to C_B}$, and the operators X_{C_A} and Y_{C_B}.

3.6 Quantum De Finetti theorems with Linear Constraints

In the following, we start by providing a brief sketch of the main ideas behind the proof. For simplicity we restrict to $k = 1$, which is the relevant case for (3.34). Namely, we start with a multipartite state $\rho_{AB_1^n}$ symmetric with respect to A, and the goal is to identify constraints such that $\rho_{AB} := \rho_{AB_1}$ is approximated by a mixture of states of the form

$$\rho_A^i \otimes \rho_B^i, \tag{3.47}$$

$$\text{with } \Lambda_{A\to C_A}\left(\rho_A^i\right) = X_{C_A}, \tag{3.48}$$

$$\text{and } \Gamma_{B\to C_B}\left(\rho_B^i\right) = Y_{C_B}. \tag{3.49}$$

3.6.1 Proof methods

The standard approach for proving de Finetti theorems [21] proceeds by measuring the B systems with the uniform measurement on the symmetric subspace given by $\left\{|\psi\rangle\langle\psi|_B^{\otimes n}\right\}_\psi$. In this case, the candidate mixture of product states is given by the expression

$$\int p\left(\psi\right) d|\psi\rangle\rho_{A|\psi} \otimes |\psi\rangle\langle\psi|_B\,, \tag{3.50}$$

where the integral is computed with respect to the Haar measure [87, Definition 7.18], $p(\psi)d|\psi\rangle$ denotes the probability of outcome ψ, and $\rho_{A|\psi}$ is the quantum the state on A conditioned on obtaining outcome ψ in the measurement. However, by doing so, the states $|\psi\rangle\langle\psi|_B$ appearing in the integral (3.50), will not satisfy, in general, the desired condition $\Gamma_{B\to C_B}\left(|\psi\rangle\langle\psi|_B\right) = Y_{C_B}$. More precisely, the measurement $\left\{|\psi\rangle\langle\psi|_B^{\otimes n}\right\}_\psi$ does not guarantee that the set $\{|\psi\rangle_B : \Gamma_{B\to C_B}\left(|\psi\rangle\langle\psi|_B\right) \neq Y_{C_B}\}$ has zero measure.

In principle, one could try to modify the measurement so that we only get $|\psi\rangle\langle\psi|_B$ satisfying the desired constraint. However, this approach seems difficult. Instead, we use an alternative approach, where the candidate mixture of product states is chosen differently [58, 17]. Namely, starting from $\rho_{AB_1^n}$, a well-chosen measurement on the systems B_2^n with measurement outcomes z_2^n leads to the candidate mixture of product states

$$\mathop{\mathbf{E}}_{z_2^n}\left\{\rho_{A|z_2^n} \otimes \rho_{B|z_2^n}\right\}. \tag{3.51}$$

The advantage of this candidate state is that, by enforcing the right constraints on the global state $\rho_{AB_1^n}$, we can ensure that $\Lambda_{A\to C_A}(\rho_{A|z_2^n}) = X_{C_A}$ and $\Gamma_{B\to C_B}(\rho_{B|z_2^n}) = Y_{C_B}$. Note that, in

order for this strategy to work properly, we need the chosen measurement to be informationally complete (2.2.3), i.e., allowing to estimate the expectation value of arbitrary operators [2], and have a small distortion in the sense that the loss in distinguishability resulting from applying the measurement is small. This concept will be made more precise in the next subsection.

3.6.2 Information-theoretic tools

The starting point for our proof technique is the use of the chain rule of the conditional mutual information, first used in this context in [16] and further exploited in [17]. More precisely, we will use the *quantum relative entropy* defined as

$$D(\rho\|\sigma) := \begin{cases} \operatorname{Tr}(\rho\log\rho) - \operatorname{Tr}(\rho\log\sigma) & \text{if } \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma) \\ \infty & \text{otherwise} \end{cases}, \tag{3.52}$$

where ρ and σ are quantum states and the logarithm is taken with respect to the basis two, i.e., $\log(\cdot) := \log_2(\cdot)$. The following theorem [87, Theorem 5.38] relates the quantum relative entropy $D(\rho\|\sigma)$ to the trace distance $\|\rho-\sigma\|_1$.

Theorem 3.6.1. (Quantum Pinsker's inequality) *Let $\rho,\sigma \in \mathcal{S}(\mathcal{H})$, then*

$$D(\rho\|\sigma) \geq \frac{1}{2\ln 2}\|\rho-\sigma\|_1^2 . \tag{3.53}$$

Via the quantum relative entropy we define the *quantum mutual information* as

$$I(A:B)_\rho := D(\rho_{AB}\|\rho_A\otimes\rho_B). \tag{3.54}$$

If $B := Z$ is a classical system, we have the following upper bound (see, e.g., [69, Chapter 11])

$$I(A:Z) \leq \log d_A. \tag{3.55}$$

In other words, one can bound the quantum mutual information of a classical-quantum state (2.37) with the quantum entropy of the maximally mixed state on the quantum system.

The following lemma, which can be found in [17], says that if some classical systems Z_1^n are symmetric with respect to A, then conditioning on Z_1^m for some value of m breaks the correlations between A and Z_{m+1}. Before stating the lemma, we introduce notation that will be used throughout the section. For a state ρ_{AZ} with a classical Z-system, we write

$$\rho_{A|z} := \frac{\mathrm{Tr}_Z\Big[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\Big]}{\mathrm{Tr}\Big[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\Big]}, \tag{3.56}$$

to denote the quantum state on A if we have obtained z as the outcome of the Z-system measurement. Notice that the term $\mathrm{Tr}\Big[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\Big]$ at the denominator is just a normalization factor whose purpose is to guarantee that $\rho_{A|z}$ is a properly defined quantum state

$$\mathrm{Tr}(\rho_{A|z}) = \mathrm{Tr}_A(\rho_{A|z}) = \frac{\mathrm{Tr}_{AZ}\Big[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\Big]}{\mathrm{Tr}\Big[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\Big]} = \frac{\mathrm{Tr}\Big[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\Big]}{\mathrm{Tr}\Big[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\Big]} = 1. \tag{3.57}$$

We simply write $\underset{z_1^m}{\mathbf{E}}\{\cdot\}$ for the expectation over the choices of z_1^m and the probability distribution will be clear from the context.

Lemma 3.6.2. *[17] Let $\rho_{AZ_1^n}$ be a classical-quantum state with the Z_1^n-systems classical and $\mathcal{U}^\pi_{Z^n}(\rho_{AZ_1^n}) = \rho_{AZ_1^n}$ for all $\pi \in \mathfrak{S}_n$. Then, there exists $0 \leq m < n$ such that*

$$\underset{z_1^m}{\mathbf{E}}\left\{D(\rho_{AZ_{m+1}|z_1^m} \| \rho_{A|z_1^m} \otimes \rho_{Z_{m+1}|z_1^m})\right\} \leq \frac{\log d_A}{n} \tag{3.58}$$

as well as

$$\underset{z_1^m}{\mathbf{E}}\left\{\|\rho_{AZ_{m+1}|z_1^m} - \rho_{A|z_1^m} \otimes \rho_{Z_{m+1}|z_1^m}\|_1^2\right\} \leq \frac{(2\ln 2)\log d_A}{n}, \tag{3.59}$$

where $\ln(\cdot)$ denotes the natural logarithm.

Proof. Since the Z_1^n-systems are classical, we can use the upper bound (3.55) to obtain the following inequality that is independent on the dimension of the classical systems and on their number n, i.e.,

$$I\left(A : Z_1^n\right)_\rho \leq \log d_A. \tag{3.60}$$

The quantum mutual information satisfies the following *chain rule* (see, e.g., [69, Chapter 11])

$$I\left(A : Z_1^n\right)_\rho = \sum_{m=0}^{n-1} I(A : Z_{m+1}|Z_1^m)_\rho, \tag{3.61}$$

where we used the *quantum conditional mutual information*

$$I(A : Z_{m+1}|Z_1^m)_\rho := I(A : Z_1^{m+1})_\rho - I(A : Z_1^m)_\rho. \tag{3.62}$$

Since $\sum_{m=0}^{n-1} I(A : Z_{m+1}|Z_1^m)_\rho \leq \log d_A$, and the sum is formed by n terms, it is not possible to have $I(A : Z_{m+1}|Z_1^m)_\rho > \frac{\log d_A}{n}$ for each $m = 0, \ldots, n-1$. As a result, there exists an $m \in \{0, \ldots, n-1\}$ such that $I(A : Z_{m+1}|Z_1^m)_\rho \leq \frac{\log d_A}{n}$, which implies

$$\mathop{\mathbf{E}}_{z_1^m} \left\{ I(A : Z_{m+1})_{\rho_{AZ_{m+1}|z_1^m}} \right\} \leq \frac{\log d_A}{n}, \tag{3.63}$$

where we used

$$I(A : Z_{m+1}|Z_1^m)_\rho = \mathop{\mathbf{E}}_{z_1^m} \left\{ I(A : Z_{m+1})_{\rho_{AZ_{m+1}|z_1^m}} \right\}, \tag{3.64}$$

which holds since the conditioning systems are classical [33].

The second statement then follows directly from Pinsker's inequality, i.e., Theorem (3.6.1).

□

The next lemma can be seen as a generalization of the law of total probability for classical-quantum states.

Lemma 3.6.3. *Let ρ_{AZ} be a classical-quantum state with the Z-system classical. Then,*

$$\mathop{\mathbf{E}}_{z}\left\{\rho_{A|z}\right\} = \rho_A. \tag{3.65}$$

Proof. Since ρ_{AZ} is a classical-quantum state, it can be written in the form (Subsection 2.2.3)

$$\rho_{AZ} = \sum_z p_z \tau_A^z \otimes |z\rangle\langle z|, \tag{3.66}$$

for a probability distribution $\{p_z\}_z$, quantum states $\{\tau_A^z\}_z$ and an orthonormal basis $(|z\rangle)_z$ for the Hilbert space Z. On one hand, it is clear that

$$\rho_A = \sum_z p_z \tau_A^z. \tag{3.67}$$

On the other, we have

$$\mathop{\mathbf{E}}_{z}\left\{\rho_{A|z}\right\} = \mathop{\mathbf{E}}_{z}\left\{\frac{\mathrm{Tr}_Z\left[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\right]}{\mathrm{Tr}\left[\rho_{AZ}\left(1_A \otimes |z\rangle\langle z|\right)\right]}\right\} \tag{3.68}$$

$$= \mathop{\mathbf{E}}_{z}\left\{\frac{\mathrm{Tr}_Z\left[p_z \tau_A^z \otimes |z\rangle\langle z|\right]}{\mathrm{Tr}\left[p_z \tau_A^z \otimes |z\rangle\langle z|\right]}\right\} \tag{3.69}$$

$$= \mathop{\mathbf{E}}_{z}\left\{\frac{p_z \tau_A^z}{p_z}\right\} \tag{3.70}$$

$$= \mathop{\mathbf{E}}_{z}\left\{\tau_A^z\right\} \tag{3.71}$$

$$= \sum_z p_z \tau_A^z. \tag{3.72}$$

Thus, $\mathop{\mathbf{E}}_{z}\left\{\rho_{A|z}\right\} = \rho_A$. □

To prove the de Finetti theorem, we will crucially make use of so-called informationally complete measurements (2.2.3) for which the loss in distinguishability, or *distortion*, can be bounded.

Lemma 3.6.4. *[17, Lemma 14] There exists a informationally complete product measurement $\mathcal{M}_A \otimes \mathcal{M}_B$ with finitely many outcomes such that, for any Hermitian and traceless operator ξ_{AB} on $A \otimes B$, we have*

$$\|(\mathcal{M}_A \otimes \mathcal{M}_B)(\xi_{AB})\|_1 \geq \frac{1}{18\sqrt{d_A d_B}} \|\xi_{AB}\|_1. \tag{3.73}$$

The above lemma follows from the methods in [63], and we stated the version for bipartite quantum systems. More generally, we define the *minimal distortion* for the bipartite system $A \otimes B$ as

$$f(A,B) := \inf_{\mathcal{M}_A, \mathcal{M}_B} \max_{\substack{\xi_{AB}^\dagger = \xi_{AB} \\ \xi_A = 0, \xi_B = 0}} \frac{\|\xi_{AB}\|_1}{\|(\mathcal{M}_A \otimes \mathcal{M}_B)(\xi_{AB})\|_1}, \tag{3.74}$$

where the infimum is over all product measurements on $A \otimes B$. In this notation, Lemma 3.6.4 shows that

$$f(A,B) \leq 18\sqrt{d_A d_B}. \tag{3.75}$$

Note that in the definition of $f(A,B)$ we restricted the maximization to operators satisfying $\xi_A = 0$ and $\xi_B = 0$ because this is sufficient for our purposes. Notice that operators satisfying $\xi_A = 0$ and $\xi_B = 0$, are also traceless. Thus, it is possible to use them in Lemma 3.6.4.

A drawback of Lemma 3.6.4 is that the distortion depends on the dimension d_A. This is not surprising since, in that lemma, we are measuring both systems. On the other hand, in certain applications we may want to measure B only. In such a case, it is interesting to investigate whether one can remove the A-system dimensional dependence. First, we define the *minimal distortion with side information* for a system B as

$$f(B|\cdot) := \inf_{\mathcal{M}_B} \sup_{\substack{\xi_{AB}^\dagger = \xi_{AB} \\ \xi_A = 0, \xi_B = 0}} \frac{\|\xi_{AB}\|_1}{\|(\mathcal{I}_A \otimes \mathcal{M}_B)(\xi_{AB})\|_1}, \tag{3.76}$$

where the infimum is over all measurements on B and the supremum is over all finite-dimensional systems A. In [14, Lemma D.1] we give an elementary proof that

$$f(B|\cdot) \leq d_B^2(d_B + 1) \tag{3.77}$$

using state two-designs and properties of weighted non-commutative L_p-spaces. With methods from operator space theory [18, Equation 66] gave the stronger bound

$$f(B|\cdot) \leq \sqrt{18 d_B^3}. \tag{3.78}$$

Finally, the following optimal bound has been shown in [53, Lemma 8]

$$f(B|\cdot) \leq 2 d_B, \tag{3.79}$$

which is linear in the dimension of the system being measured.

3.6.3 Main Theorem

Combining the tools from the previous subsection we find the following de Finetti theorem with linear constraints.

Theorem 3.6.5. *Let $\rho_{AB_1^n}$ be a quantum state, $\Lambda_{A\to C_A}, \Gamma_{B\to C_B}$ super-operators, and X_{C_A}, Y_{C_B} operators such that*

$$\mathcal{U}^{\pi}_{B_1^n}(\rho_{AB_1^n}) = \rho_{AB_1^n}\ \forall \pi \in \mathfrak{S}_n \qquad \textit{symmetric with respect to } A \tag{3.80}$$

$$\Lambda_{A\to C_A}(\rho_{AB_1^n}) = X_{C_A} \otimes \rho_{B_1^n} \qquad \textit{linear constraint on } A \tag{3.81}$$

$$\Gamma_{B_n\to C_B}(\rho_{B_1^n}) = \rho_{B_1^{n-1}} \otimes Y_{C_B} \qquad \textit{linear constraint on } B. \tag{3.82}$$

Then, we have that

$$\left\| \rho_{AB} - \sum_{i\in I} p_i \sigma_A^i \otimes \omega_B^i \right\|_1 \leq \min\left\{ f(A,B), f(B|\cdot) \right\} \sqrt{\frac{(2\ln 2)\log(d_A)}{n}} \tag{3.83}$$

with $\{p_i\}_{i\in I}$ *a probability distribution,* $\rho_{AB} := \mathrm{Tr}_{B_2^n}\left(\rho_{AB_1^n}\right)$*, and quantum states* σ_A^i, ω_B^i *such that for every* $i \in I$*:*

$$\Lambda_{A\to C_A}\left(\sigma_A^i\right) = X_{C_A} \quad \textit{and} \quad \Gamma_{B\to C_B}\left(\omega_B^i\right) = Y_{C_B}. \tag{3.84}$$

As stated in Section 3.6.2, we can, e.g., take $f(A,B) \leq 18\sqrt{d_A d_B}$ *or* $f(B|\cdot) \leq 2d_B$*.*

Proof. Let $\mathcal{M}_B$ be a measurement of the B system and call the outcome system Z. Consider the state $\rho_{AZ_1^n}$ obtained by measuring all the B systems with the same quantum measurement $\mathcal{M}_B$. This is a classical-quantum state symmetric with respect to A and so we can apply Lemma 3.6.2. We find that there exists an $m \in \{0,\ldots,n-1\}$ such that

$$\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AZ_{m+1}|z_1^m} - \rho_{A|z_1^m}\otimes\rho_{Z_{m+1}|z_1^m}\|_1^2\right\} \leq \frac{(2\ln 2)\log d_A}{n}. \tag{3.85}$$

For any collection z_1^m of measurement outcomes, we can rewrite the quantum states $\rho_{AZ_{m+1}|z_1^m}$ and $\rho_{Z_{m+1}|z_1^m}$ appearing in (3.85) as

$$\rho_{AZ_{m+1}|z_1^m} = (\mathcal{I}_A\otimes\mathcal{M}_B)(\rho_{AB_{m+1}|z_1^m}), \tag{3.86}$$

and

$$\rho_{Z_{m+1}|z_1^m} = \mathcal{M}_B(\rho_{B_{m+1}|z_1^m}). \tag{3.87}$$

Now, we choose the measurement $\mathcal{M}_B$ achieving $f(B|\cdot)$ in (3.76), and we get that $\|\xi_{AB}\|_1^2 \leq f(B|\cdot)^2\|(\mathcal{I}_A\otimes\mathcal{M}_B)(\xi_{AB})\|_1^2$, where we set $\xi_{AB} := \rho_{AB_{m+1}|z_1^m} - \rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}$. Notice that ξ_{AB}, being a difference of two Hermitian operators, is Hermitian, and it satisfies

$$\xi_A = \rho_{A|z_1^m} - \rho_{A|z_1^m} = 0, \tag{3.88}$$

and

$$\xi_B = \rho_{B_{m+1}|z_1^m} - \rho_{B_{m+1}|z_1^m} = 0. \tag{3.89}$$

As a result, we have

$$\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1^2\right\}\leq f(B|\cdot)^2\frac{(2\ln 2)\log d_A}{n}. \tag{3.90}$$

Now, using the convexity of the square function, we get

$$\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1\right\}\leq\sqrt{\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1^2\right\}} \tag{3.91}$$

$$\leq f(B|\cdot)\sqrt{\frac{(2\ln 2)\log d_A}{n}}. \tag{3.92}$$

To arrive to the above inequality, we measured only the B systems. On the other hand, we can also choose measurements $\mathcal{M}_A$ and $\mathcal{M}_B$ achieving $f(A,B)$ in (3.74). In this case,

$$\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1^2 \tag{3.93}$$

$$\leq f(A,B)^2\|(\mathcal{M}_A\otimes\mathcal{M}_B)(\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m})\|_1^2 \tag{3.94}$$

$$\leq f(A,B)^2\|(\mathcal{M}_A\otimes\mathcal{I}_B)[(\mathcal{I}_A\otimes\mathcal{M}_B)(\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m})]\|_1^2 \tag{3.95}$$

$$\leq f(A,B)^2\|(\mathcal{I}_A\otimes\mathcal{M}_B)(\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m})\|_1^2 \tag{3.96}$$

$$= f(A,B)^2\|\rho_{AZ_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{Z_{m+1}|z_1^m}\|_1^2, \tag{3.97}$$

where we used the fact that the trace norm cannot increase when applying the quantum channel $\mathcal{M}_A$ [87, Theorem 3.39]. As a result, we get

$$\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1^2\right\}\leq f(A,B)^2\frac{(2\ln 2)\log d_A}{n}. \tag{3.98}$$

Again, using the convexity of the square function, we get

$$\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1\right\}\leq f(A,B)\sqrt{\frac{(2\ln 2)\log d_A}{n}}. \tag{3.99}$$

Thus, we can bound the expectation $\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1\right\}$ in two different ways. The best upper bound will be given by the minimum of the two ones, i.e.,

$$\mathop{\mathbf{E}}_{z_1^m}\left\{\|\rho_{AB_{m+1}|z_1^m}-\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\|_1\right\}\leq\min\left\{f(A,B),f(B|\cdot)\right\}\sqrt{\frac{(2\ln 2)\log d_A}{n}}. \tag{3.100}$$

Then, using the convexity of the norm (which is true for every norm due to the triangle inequality), and Lemma 3.6.3

$$\mathop{\mathbf{E}}_{z_1^m}\left\{\rho_{AB_{m+1}|z_1^m}\right\} = \rho_{AB_{m+1}}, \tag{3.101}$$

we obtain

$$\left\|\rho_{AB_{m+1}} - \mathop{\mathbf{E}}_{z_1^m}\left\{\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\right\}\right\|_1 \leq \min\left\{f(A,B), f(B|\cdot)\right\}\sqrt{\frac{(2\ln 2)\log d_A}{n}}. \tag{3.102}$$

The state $\mathop{\mathbf{E}}_{z_1^m}\left\{\rho_{A|z_1^m}\otimes\rho_{B_{m+1}|z_1^m}\right\}$ corresponds to our candidate mixture of product states.

It now remains to show that all the states in the mixture satisfy the linear constraints. Indeed we have for any collection of measurement outcomes z_1^m, writing M_B^z for the POVM elements of the measurement $\mathcal{M}_B$,

$$\Lambda_{A\to C_A}(\rho_{A|z_1^m}) = \frac{\mathrm{Tr}_{B_1^m}\left[(1_A\otimes M_B^{z_1}\otimes\cdots\otimes M_B^{z_m})\Lambda_{A\to C_A}(\rho_{AB_1^m})\right]}{\mathrm{Tr}\left[(1_A\otimes M_B^{z_1}\otimes\cdots\otimes M_B^{z_m})\rho_{AB_1^m}\right]} \tag{3.103}$$

$$= \frac{\mathrm{Tr}_{B_1^m}\left[(1_A\otimes M_B^{z_1}\otimes\cdots\otimes M_B^{z_m})(X_{C_A}\otimes\rho_{B_1^m})\right]}{\mathrm{Tr}\left[(1_A\otimes M_B^{z_1}\otimes\cdots\otimes M_B^{z_m})\rho_{AB_1^m}\right]} \tag{3.104}$$

$$= X_{C_A}, \tag{3.105}$$

and similarly

$$\Gamma_{B\to C_B}(\rho_{B_{m+1}|z_1^m}) = \frac{\mathrm{Tr}_{B_1^m}\left[(M_B^{z_1}\otimes\cdots\otimes M_B^{z_m}\otimes 1_{C_B})\Gamma_{B_{m+1}\to C_B}(\rho_{B_1^{m+1}})\right]}{\mathrm{Tr}\left[(M_B^{z_1}\otimes\cdots\otimes M_B^{z_m}\otimes 1_{B_{m+1}})\rho_{B_1^{m+1}}\right]} \tag{3.106}$$

$$= \frac{\mathrm{Tr}_{B_1^m}\left[(M_B^{z_1}\otimes\cdots\otimes M_B^{z_m}\otimes 1_{B_{m+1}})(\rho_{B_1\cdots B_m}\otimes Y_{C_B})\right]}{\mathrm{Tr}\left[(M_B^{z_1}\otimes\cdots\otimes M_B^{z_1}\otimes 1_{B_{m+1}})\rho_{B_1^{m+1}}\right]} \tag{3.107}$$

$$= Y_{C_B}. \tag{3.108}$$

□

Theorem 3.6.5 allows us to approximate the set of separable states subject to linear constraints on A and B, with a proper subset of n-Ext$(A:B)$, formed by n-extendible states satisfying two appropriate linear conditions. Comparing the bound of Theorem 3.6.5 with (3.25), we see that the room for improvement is fairly limited, i.e., we may be able to improve the square root and the logarithm dependence, but the overall bound cannot be made exponentially better.

In the next subsection we will generalize such result, by keeping a generic number $0 < k < n$ of B-systems, instead of $k = 1$ (which will correspond to the setting of Theorem 3.6.5).

3.6.4 Generalizing the Main Theorem to k Copies

As pointed out at the end of the previous subsection, Theorem 3.6.5 can be extended to a generalized finite quantum de Finetti theorem for any reduced state $\rho_{AB_1^k}$ with $0 < k < n$.

Theorem 3.6.6. *For the same setting as in Theorem 3.6.5, we have for $0 < k < n$ that*

$$\left\| \rho_{AB_1^k} - \sum_{i\in I} p_i \sigma_A^i \otimes \left(\omega_B^i\right)^{\otimes k} \right\|_1 \leq k f(B|\cdot) \sqrt{(2\ln 2)\frac{\log d_A + (k-1)\log d_B}{n-k+1}}. \tag{3.109}$$

Proof. Note that the for the state $\rho_{AB_1^{k-1}B_k^n}$, the systems B_k^n are symmetric with respect to AB_1^{k-1}. As such, we can apply the same argument used in the proof of Lemma 3.6.2 and Theorem 3.6.5, but this time starting from the following chain rule

$$I\left(AB_1^{k-1} : Z_k^n\right)_\rho = \sum_{m=k-1}^{n-1} I(AB_1^{k-1} : Z_{m+1}|Z_k^m)_\rho. \tag{3.110}$$

Notice that the sum in the chain rule contains $(n-1)-(k-1)+1 = n-k+1$ terms. Thus, we find that there exists a $m \in \{k, \ldots, n\}$ such that

$$\mathop{\mathbf{E}}_{z_{k+1}^m} \left\{ \|\rho_{AB_1^k|z_{k+1}^m} - \rho_{AB_1^{k-1}|z_{k+1}^m} \otimes \rho_{B_k|z_{k+1}^m}\|_1 \right\} \leq f(B|\cdot) \sqrt{\frac{(2\ln 2)\log\left(d_A d_B^{k-1}\right)}{n-k+1}}. \tag{3.111}$$

Summing the $n-k+1$ inequalities labelled by m, we find

$$\frac{1}{n-k+1}\sum_{m=k}^{n}\underset{z_{k+1}^m}{\mathbf{E}}\left\{\|\rho_{AB_1^k|z_{k+1}^m}-\rho_{AB_1^{k-1}|z_{k+1}^m}\otimes\rho_{B_k|z_{k+1}^m}\|_1\right\} \tag{3.112}$$

$$\leq f(B|\cdot)\sqrt{\frac{(2\ln 2)\log\left(d_A d_B^{k-1}\right)}{n-k+1}}. \tag{3.113}$$

By symmetry of the B-systems, for any $i\in\{1,\ldots,k\}$, we also have

$$\frac{1}{n-k+1}\sum_{m=k}^{n}\underset{z_{k+1}^m}{\mathbf{E}}\left\{\|\rho_{AB_1^i|z_{k+1}^m}-\rho_{AB_1^{i-1}|z_{k+1}^m}\otimes\rho_{B_i|z_{k+1}^m}\|_1\right\} \tag{3.114}$$

$$\leq f(B|\cdot)\sqrt{\frac{(2\ln 2)\log\left(d_A d_B^{i-1}\right)}{n-k+1}}. \tag{3.115}$$

Now, using the triangle inequality $k-1$ times, we get for any $m\in\{k,\ldots,n\}$ and collection z_{k+1}^m of measurement outcomes, that

$$\left\|\rho_{AB_1^k|z_{k+1}^m}-\rho_{A|z_{k+1}^m}\otimes\rho_{B_1|z_{k+1}^m}\otimes\cdots\otimes\rho_{B_k|z_{k+1}^m}\right\|_1 \tag{3.116}$$

$$\leq\sum_{i=1}^{k}\Big\|\rho_{AB_1^i|z_{k+1}^m}\otimes\rho_{B_{i+1}|z_{k+1}^m}\otimes\cdots\otimes\rho_{B_k|z_{k+1}^m} \tag{3.117}$$

$$-\rho_{AB_1^{i-1}|z_{k+1}^m}\otimes\rho_{B_i|z_{k+1}^m}\otimes\rho_{B_{i+1}|z_{k+1}^m}\otimes\cdots\otimes\rho_{B_k|z_{k+1}^m}\Big\|_1 \tag{3.118}$$

$$=\sum_{i=1}^{k}\left\|\rho_{AB_1^i|z_{k+1}^m}-\rho_{AB_1^{i-1}|z_{k+1}^m}\otimes\rho_{B_i|z_{k+1}^m}\right\|_1\left\|\rho_{B_{i+1}|z_{k+1}^m}\otimes\cdots\otimes\rho_{B_k|z_{k+1}^m}\right\|_1 \tag{3.119}$$

$$=\sum_{i=1}^{k}\left\|\rho_{AB_1^i|z_{k+1}^m}-\rho_{AB_1^{i-1}|z_{k+1}^m}\otimes\rho_{B_i|z_{k+1}^m}\right\|_1, \tag{3.120}$$

where we used

$$\left\|\rho_{B_{i+1}|z_{k+1}^m}\otimes\cdots\otimes\rho_{B_k|z_{k+1}^m}\right\|_1=\mathrm{Tr}(\rho_{B_{i+1}|z_{k+1}^m}\otimes\cdots\otimes\rho_{B_k|z_{k+1}^m})=1. \tag{3.121}$$

Taking the average over m and z_{k+1}^m and using (3.115), we get

$$\frac{1}{n-k+1}\sum_{m=k}^{n}\underset{z_{k+1}^m}{\mathbf{E}}\left\{\left\|\rho_{AB_1^k|z_{k+1}^m}-\rho_{A|z_{k+1}^m}\otimes\rho_{B_1|z_{k+1}^m}\otimes\cdots\otimes\rho_{B_k|z_{k+1}^m}\right\|_1\right\} \tag{3.122}$$

$$\leq \sum_{i=1}^{k} f(B|\cdot)\sqrt{\frac{(2\ln 2)\log\left(d_A d_B^{i-1}\right)}{n-k+1}} \tag{3.123}$$

$$\leq \sum_{i=1}^{k} f(B|\cdot)\sqrt{\frac{(2\ln 2)\log\left(d_A d_B^{k-1}\right)}{n-k+1}} \tag{3.124}$$

$$= kf(B|\cdot)\sqrt{\frac{(2\ln 2)\log\left(d_A d_B^{k-1}\right)}{n-k+1}}. \tag{3.125}$$

As a result, there is an m such that the previous inequality holds. Then, as before, we use the convexity of the norm to put the expectation inside, getting the existence of an m such that

$$\left\| \rho_{AB_1^k} - \mathop{\mathbf{E}}_{z_{k+1}^m} \left\{ \rho_{A|z_{k+1}^m} \otimes \rho_{B_1|z_{k+1}^m} \otimes \cdots \otimes \rho_{B_k|z_{k+1}^m} \right\} \right\|_1 \tag{3.126}$$

$$\leq kf(B|\cdot)\sqrt{(2\ln 2)\frac{\log d_A + (k-1)\log d_B}{n-k+1}}. \tag{3.127}$$

To conclude, it suffices to observe that, by symmetry, $\rho_{B_i|z_{k+1}^m} = \rho_{B_1|z_{k+1}^m}$ for all $i \in \{1,\ldots,k\}$ and the linear constraints are satisfied by the same calculation as in the proof of Theorem 3.6.5.

□

The same comment we made on the dimensional dependence for the bound provided by Theorem 3.6.5 does apply to its generalization as given by Theorem 3.6.6. I.e., the error term cannot be exponentially improved.

3.7 Application to Constrained Bilinear Optimization

As stated in Section 3.5, the constrained bilinear optimization problem we are interested in, takes the form

$$Q = \max \quad \mathrm{Tr}\left[G_{AB}(\rho_A \otimes \rho_B)\right] \tag{3.128}$$

$$s.t. \quad \rho_A \succeq 0, \quad \rho_B \succeq 0 \tag{3.129}$$

$$\mathrm{Tr}(\rho_A) = \mathrm{Tr}(\rho_B) = 1 \tag{3.130}$$

$$\Lambda_{A\to C_A}(\rho_A) = X_{C_A}, \quad \Gamma_{B\to C_B}(\rho_B) = Y_{C_B}. \tag{3.131}$$

Lower bounds on the optimal value can, e.g., be derived by means of seesaw methods [60] (see [88] for an example in quantum information theory). Those methods often converge in practice and sometimes even provably reach a local maxima. What was missing, however, is a general method to give an approximation guarantee to the global maximum.

The de Finetti theorem with linear constraints (Theorem 3.6.5) gives an SDP hierarchy of outer bounds, that exactly provides such a criterion.

Theorem 3.7.1. *For the SDPs*

$$\mathrm{SDP}_n := \max \quad \mathrm{Tr}\big[G_{AB}\rho_{AB_1}\big] \tag{3.132}$$

$$s.t. \quad \rho_{AB_1^n} \succeq 0, \mathrm{Tr}(\rho_{AB_1^n}) = 1 \tag{3.133}$$

$$\rho_{AB_1^n} = \mathcal{U}^{\pi}_{B_1^n}\left(\rho_{AB_1^n}\right) \ \forall \pi \in \mathfrak{S}_n \tag{3.134}$$

$$\Lambda_{A\to C_A}\left(\rho_{AB_1^n}\right) = X_{C_A} \otimes \rho_{B_1^n} \tag{3.135}$$

$$\Gamma_{B_n\to C_B}\left(\rho_{B_1^n}\right) = \rho_{B_1^{n-1}} \otimes Y_{C_B} \tag{3.136}$$

and Q defined as above, we have for $d := \max\{d_A, d_B\}$ that

$$0 \leq \mathrm{SDP}_n - Q \leq \frac{\mathrm{poly}(d)}{\sqrt{n}} \quad \textit{implying} \quad Q = \lim_{n\to\infty} \mathrm{SDP}_n. \tag{3.137}$$

Proof. We have by construction $0 \leq \mathrm{SDP}_n - Q$ and the remaining inequality arises from

$$\mathrm{Tr}\left[G_{AB}\rho_{AB_1}\right] = \mathrm{Tr}\left[G_{AB}(\rho_A \otimes \rho_B)\right] + \mathrm{Tr}\left[G_{AB}\left(\rho_{AB_1} - \rho_A \otimes \rho_B\right)\right] \tag{3.138}$$

$$\leq \mathrm{Tr}\left[G_{AB}(\rho_A \otimes \rho_B)\right] + \|G_{AB}\|_\infty \cdot \|\rho_{AB_1} - \rho_A \otimes \rho_B\|_1 \tag{3.139}$$

$$\leq \mathrm{Tr}\left[G_{AB}(\rho_A \otimes \rho_B)\right] + \frac{\mathrm{poly}(d)}{\sqrt{n}}, \tag{3.140}$$

where we used the linearity of the objective function as described in Section 3.5, the $(\infty, 1)$ Hölder's inequality [87, Inequality 1.174]

$$\mathrm{Tr}\left[G_{AB}\left(\rho_{AB_1} - \rho_A \otimes \rho_B\right)\right] \leq \|G_{AB}\|_\infty \cdot \|\rho_{AB_1} - \rho_A \otimes \rho_B\|_1, \tag{3.141}$$

and the de Finetti argument as in Theorem 3.6.5. □

The bounds from Theorem 3.6.5 give worst case convergence guarantees that are "fairly slow", as to ensure that the approximation error is small we need at least the level $n = \mathrm{poly}(d)$. However, note that constrained bilinear optimization contains as a special case the best separable state problem and so we cannot expect much better bounds on the convergence speed in general. We refer to [44] and the references therein for a detailed discussion about the computational complexity of the best separable state problem.

We can add positive partial transpose (PPT) constraints

$$\rho_{AB_1^n}^{\mathrm{T}_A} \succeq 0,\ \rho_{AB_1^n}^{\mathrm{T}_{B_1}} \succeq 0,\ \rho_{AB_1^n}^{\mathrm{T}_{B_1^2}} \succeq 0, \ldots,\ \rho_{AB_1^n}^{\mathrm{T}_{B_1^{n-1}}} \succeq 0 \tag{3.142}$$

to SDP_n and we denote the resulting relaxations by $\mathrm{SDP}_{n,\mathrm{PPT}}$. It is important to point out that any separable state is also a PPT state (Corollary 3.3.3), and hence we still have a valid relaxation to the problem (3.30). It is an interesting question to study if these constraints can lead to a faster convergence speed, cf. the discussions in [68, 32]. Based on the PPT constraints, we can give a sufficient condition when already

$$\mathrm{SDP}_{n,\mathrm{PPT}} = Q \text{ for some finite } n. \tag{3.143}$$

The condition — known as *rank loop condition* — is based on [68], which in turn builds on [51].

Lemma 3.7.2. *[68],[51] Let $\rho_{AB_1^n} = \mathcal{U}_{B_1^n}^{\pi}\left(\rho_{AB_1^n}\right)$ for all $\pi \in \mathfrak{S}_n$ and fixed $0 \leq k \leq n$ such that $\rho_{AB_1^n}^{\mathrm{T}_{B_{k+1}^n}} \succeq 0$. Then, ρ_{AB_1} is separable if*

$$\mathrm{rank}(\rho_{AB_1^n}) \leq \max\left\{\mathrm{rank}\left(\rho_{AB_1^k}\right), \mathrm{rank}\left(\rho_{B_{k+1}^n}\right)\right\}. \tag{3.144}$$

Proof. The proof is based on [51], where they prove the following implication

$$\rho_{AB} \in \mathrm{PPT}(A:B) \text{ and } \mathrm{rank}(\rho_{AB}) \leq \mathrm{rank}(\rho_A) \implies \rho_{AB} \in \mathrm{SEP}(A:B). \tag{3.145}$$

Applying the above result to the quantum state $\rho_{AB_1^n}$ with respect to the bipartite system $AB_1^k \otimes B_{k+1}^n$, we find

$$\rho_{AB_1^n} \in \mathrm{PPT}(AB_1^k : B_{k+1}^n) \text{ and } \mathrm{rank}(\rho_{AB_1^n}) \leq \mathrm{rank}(\rho_{AB_1^k}) \implies \rho_{AB_1^n} \in \mathrm{SEP}(AB_1^k : B_{k+1}^n) \tag{3.146}$$

and,

$$\rho_{AB_1^n} \in \mathrm{PPT}(AB_1^k : B_{k+1}^n) \text{ and } \mathrm{rank}(\rho_{AB_1^n}) \leq \mathrm{rank}(\rho_{B_{k+1}^n}) \implies \rho_{AB_1^n} \in \mathrm{SEP}(AB_1^k : B_{k+1}^n). \tag{3.147}$$

Thus, $\rho_{AB_1^n} \in \mathrm{SEP}(AB_1^k : B_{k+1}^n)$ if

$$\mathrm{rank}(\rho_{AB_1^n}) \leq \max\left\{\mathrm{rank}\left(\rho_{AB_1^k}\right), \mathrm{rank}\left(\rho_{B_{k+1}^n}\right)\right\}. \tag{3.148}$$

Finally, since $\rho_{AB_1^n}$ is symmetric with respect to A, we have

$$\rho_{AB_1^n} \in \mathrm{SEP}(AB_1^k : B_{k+1}^n) \implies \rho_{AB_1} \in \mathrm{SEP}(A:B_1). \tag{3.149}$$

□

Note that instead of extending the B-systems we could equally well extend the A-systems to get another, possibly non-equivalent, hierarchy.

In the next chapter we will use the methods developed in this chapter to study the problem of approximate quantum error correction. Moreover, our methods can be readily applied to the quantum marginal problem[4] and to an entire class of problems expressed via rank-constrained SDPs, as subsequently studied in [91, 90].

[4] In the *quantum marginal problem* the question of interest is whether a given collection of quantum states can be seen as the marginals of a, not necessarily unique, global quantum state.

Chapter 4

Approximate Quantum Error Correction

In this chapter we apply the results of the previous one to the problem of approximate quantum error correction. First we introduce the problem and its relevance and applications in quantum information theory. We will then use a specialized version of Theorem 3.7.1 to obtain convergent hierarchies to the desired problem. Corresponding numerical tests can be found in Section 4.4.

Given a noisy classical channel $N_{X\to Y}$, a central quantity of interest in error correction is the *maximum success probability* $p(N, M)$ for transmitting a uniform M-dimensional message under the noise model $N_{X\to Y}$. This is a bilinear maximization problem, which is in general NP-hard[1] to approximate up to a sufficiently small constant factor [8]. Nevertheless, there are

[1]We can think of NP-hard problems as the class of problems that are at least as hard as NP-complete problems, the latter being the hardest problems in the NP class. The *halting problem* is a classic example of a NP-hard decision problem. Notice that the halting problem is not NP-complete, since it is not decidable in a finite number of steps. Given a quantum state ρ_{AB}, it is NP-hard to decide whether $\rho_{AB} \in \mathrm{Sep}(A : B)$ [41, Theorem 6.7]. Finally, the class of NP-hard problems is not limited to decision problems, such as the halting

efficient methods for constructing feasible coding schemes approximating $p(N, M)$ from below as well as an efficiently computable linear programming relaxation $\mathrm{lp}(N, M)$ (sometimes called *meta-converse* [45, 71]) giving upper bounds on $p(N, M)$.[2] In fact, it was shown in [8] that $p(N, M)$ and $\mathrm{lp}(N, M)$ cannot be very far from each other

$$p(N, M) \leq \mathrm{lp}(N, M) \leq \frac{1}{1 - \frac{1}{e}} \cdot p(N, M). \tag{4.1}$$

Furthermore, the meta-converse has many appealing analytic properties, such as, e.g., the ability to evaluate it efficiently in the limit of many independent repetitions $N_{X \to Y}^{\times n}$, leading to very precise asymptotic bounds, i.e., considering the limit $n \to \infty$, on the capacity[3] of noisy classical channels [8].

The analogue quantum problem is to determine the *quantum channel fidelity* $F(\mathcal{N}, M)$, a quantity that will be formally defined later (Definition 4.1.1), for transmitting one part of a maximally entangled state of dimension M over a noisy quantum channel $\mathcal{N}_{A \to B}$. As in the classical case, this is a bilinear optimization problem, only now with matrix-valued variables. In order to approximate $F(\mathcal{N}, M)$, an efficiently computable semidefinite programming relaxation $\mathrm{SDP}(\mathcal{N}, M)$ was given in [65].[4] However, contrary to the classical case the gap between

problem, but it also includes other kind of problems, e.g., optimization problems (see [5]).

[2]Operationally, $\mathrm{lp}(N, M)$ corresponds to the *non-signalling assisted maximum success probability* discussed in [67]. In other words, the two parties of the protocol are allowed to share additional resources that are not useful for communication by themselves, e.g., shared randomness. Such kind of resources, which do not allow the two parties to send information to each other, are known in the literature as *non-signalling boxes*.

[3]The *capacity* of a channel is defined as the maximum rate at which a sender can send information to a receiver through the channel [25].

[4]Operationally, $\mathrm{SDP}(\mathcal{N}, M)$ corresponds to the *PPT-preserving, non-signalling assisted maximum fidelity*. *PPT-preserving channels* map bipartite PPT states into bipartite PPT states, and include all unassisted and forward-classical-assisted communication. However, not all entanglement-assisted communication protocols can be represented with PPT-preserving channels [65].

$\mathrm{SDP}(\mathcal{N}, M)$ and $F(\mathcal{N}, M)$ is not understood. On the other hand, the tools introduced in the previous chapter will exactly be used to generate a converging hierarchy of efficiently computable semidefinite programming relaxations, allowing us to quantify the gap between these new relaxations and $F(\mathcal{N}, M)$.

Moreover, the relaxation $\mathrm{SDP}(\mathcal{N}, M)$ is lacking most of the analytic properties of its classical analogue $\mathrm{lp}(N, M)$. In fact, in quantum communication theory so-called non-additivity problems[5] caused by quantum correlations make it notoriously hard to compute asymptotic limits in the first place [29]. Hence, we propose to use methods from optimization theory to directly study the maximum fidelity $F(\mathcal{N}, M)$ in order to quantify the ability of a quantum channel to transmit quantum information. The goal is then to identify a quantum version of the meta-converse for approximating $F(\mathcal{N}, M)$, having similar properties as the classical meta-converse $\mathrm{lp}(N, M)$ for approximating $p(N, M)$. This approach can also be justified by the fact that most of the quantum devices that will be available in the near future are likely to be noisy and small in size. As such, efficient algorithms approximating $F(\mathcal{N}, M)$ for reasonable error models $\mathcal{N}$ and dimension M are more relevant in such settings than computing the asymptotic limit of the rate achievable for multiple copies of a given noise model.

Numerical lower bound methods for $F(\mathcal{N}, M)$ are available through iterative seesaw methods[6] that lead to efficiently computable semidefinite programs [73, 72, 37, 36, 61, 77, 54]. These algorithms often converge in practice and sometimes even provably reach a *local maximum*. What was previously missing, however, is a general method to give an approximation guarantee

[5] For example, while the classical channel capacity is additive over independent channel repetitions, this is not true for the quantum channel capacity and related quantities (see [42]).

[6] An *iterative seesaw method* tackles a joint optimization by alternating the optimization over a subset of variables, with the others kept fixed. In this case, a constrained bilinear optimization would lead to a sequence of SDP optimizations.

to the *global maximum*. Here, the techniques as developed in Section 3.7 exactly lead to a converging hierarchy of efficiently computable semidefinite programming relaxations on the maximum fidelity $F(\mathcal{N}, M)$. As such, this can be seen as a tool for benchmarking existing quantum error correction codes and to understand in what direction to look for improved codes.

We note that references [80, 83, 84, 56] gave refined relaxations on the size of a maximally entangled state that can be sent over a noisy quantum channel for fixed fidelity $1-\varepsilon$. These approaches are complementary[7] to our work and contrary to our findings they do not lead to a converging hierarchy of efficiently computable bounds.

4.1 Setting

The mathematical setting of approximate quantum error correction we study is as follows. First, we define the main quantity of interest, i.e., the *quantum channel fidelity* (or in short *channel fidelity*).

Definition 4.1.1. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. The channel fidelity for message dimension M is defined as*

$$F(\mathcal{N}, M) := \max \quad F\Big(\Phi_{\overline{B}R}, \big(\big(\mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}}\big) \otimes \mathcal{I}_R\big)(\Phi_{AR})\Big) \tag{4.2}$$

$$s.t. \quad \mathcal{D}_{B\to\overline{B}}, \mathcal{E}_{A\to\overline{A}} \text{ quantum channels}, \tag{4.3}$$

where $F(\rho,\sigma) := \big\|\sqrt{\rho}\sqrt{\sigma}\big\|_1^2$ denotes the fidelity, Φ_{AR} denotes the maximally entangled state on AR, and we have $M = d_A = d_{\overline{B}} = d_R$.

[7] Notice that in this thesis we fix the size M of the message, while the fidelity measure of the protocol is not fixed. Thus, we are interested in obtaining bounds on the fidelity of the protocol, not on the message size.

Operationally, one creates a reference copy R of the input system A (hence $d_A = d_R$) and sends the maximally entangled state on AR through the quantum channel

$$\left(\mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}}\right) \otimes \mathcal{I}_R. \tag{4.4}$$

Then, the channel fidelity $F(\mathcal{N}, M)$ is defined as the maximum fidelity between the output of this protocol and the maximally entangled state on $\overline{B}R$ (for consistency, we then need $d_A = d_{\overline{B}}$). In information-theoretic language, the channel fidelity corresponds to an *average error criterion* for preserving uniformly distributed information. Alternatively, we might also aim for a *worst error criterion*. To do so, we need to move away from the quantum channel fidelity and use another metric (the diamond norm (2.60)). This approach, and the related analysis, will be discussed in Section 4.5.

We will use the following well-known result, which allows us to simplify the fidelity when one of the two quantum states is pure.

Lemma 4.1.2. *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, if ρ is a pure quantum state, then*

$$F(\rho, \sigma) = \mathrm{Tr}(\rho\sigma). \tag{4.5}$$

Proof. If $\rho \in \mathcal{S}(\mathcal{H})$ is pure, it can be written as $\rho = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is an appropriate ket in $\mathcal{H}$ with unit norm. Then,

$$\begin{aligned}
F(\rho, \sigma) &= \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2 = \left(\mathrm{Tr}\left|\sqrt{\rho}\sqrt{\sigma}\right|\right)^2 && (4.6)\\
&= \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 = \left(\mathrm{Tr}\sqrt{|\psi\rangle\langle\psi|\sigma|\psi\rangle\langle\psi|}\right)^2 && (4.7)\\
&= \langle\psi|\sigma|\psi\rangle \left(\mathrm{Tr}\sqrt{|\psi\rangle\langle\psi|}\right)^2 = \langle\psi|\sigma|\psi\rangle && (4.8)\\
&= \mathrm{Tr}(\langle\psi|\sigma|\psi\rangle) = \mathrm{Tr}(|\psi\rangle\langle\psi|\sigma) && (4.9)\\
&= \mathrm{Tr}(\rho\sigma). && (4.10)
\end{aligned}$$

□

By the Choi-Jamiołkowski isomorphism (2.41) the channel fidelity is conveniently rewritten as a bilinear optimization.

Lemma 4.1.3. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. Then, the channel fidelity can be written as*

$$F(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)\left(E_{A\overline{A}} \otimes D_{B\overline{B}}\right)\right] \tag{4.11}$$

$$s.t. \quad E_{A\overline{A}} \succeq 0, \quad D_{B\overline{B}} \succeq 0 \tag{4.12}$$

$$E_A = \frac{1_A}{d_A}, \quad D_B = \frac{1_B}{d_B}, \tag{4.13}$$

where $J^{\mathcal{N}}_{B\overline{A}} := (\mathcal{N}_{\overline{A}\to B} \otimes \mathcal{I}_{\overline{A}})(\Phi_{\overline{A}\,\overline{A}})$ denotes the Choi state of $\mathcal{N}_{\overline{A}\to B}$ (see 2.41).

The advantage of this notation is that all A-systems, i.e., A and $\overline{A}$, are with the sender (termed Alice) and all B-systems, i.e., B and $\overline{B}$, are with the receiver (termed Bob), which is consistent with the conventions used in [65].

Proof. By using the adjoint map in Hilbert-Schmidt inner product (2.27) and multiple times the Choi-Jamiołkowski isomorphism as given in (2.41), and noting that the pure state $\Phi_{\overline{B}R}$ allows us to use the simplified expression for the fidelity when one of the two arguments is pure (Lemma 4.1.2), we can write the objective function from Definition 4.1.1 as

$$F\Big(\Phi_{\overline{B}R}, \big(\left(\mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}}\right) \otimes \mathcal{I}_R\big)(\Phi_{AR})\Big) \tag{4.14}$$

$$= \mathrm{Tr}\left[\Phi_{\overline{B}R}\big(\left(\mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}}\right) \otimes \mathcal{I}_R\big)(\Phi_{AR})\right] \tag{4.15}$$

$$= \mathrm{Tr}\left[J^{\mathcal{D}^\dagger}_{BR}\left(\mathcal{N}_{\overline{A}\to B} \otimes \mathcal{I}_R\right)\left(J^{\mathcal{E}}_{\overline{A}R}\right)\right]. \tag{4.16}$$

Taking advantage of $d_A = d_{\overline{B}} = d_R$, we relabel the systems and we proceed as follows

$$F\Big(\Phi_{\overline{B}R}, \big(\left(\mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}}\right) \otimes \mathcal{I}_R\big)(\Phi_{AR})\Big) \tag{4.17}$$

$$= \operatorname{Tr}\left[J^{\mathcal{D}^\dagger}_{BR}\left(\mathcal{N}_{\overline{A}\to B}\otimes\mathcal{I}_R\right)\left(J^{\mathcal{E}}_{\overline{A}R}\right)\right] \tag{4.18}$$

$$= \operatorname{Tr}\left[J^{\mathcal{D}^\dagger}_{B\overline{B}}\left(\mathcal{N}_{\overline{A}\to B}\otimes\mathcal{I}_{A\to\overline{B}}\right)\left(J^{\mathcal{E}}_{\overline{A}A}\right)\right] \tag{4.19}$$

$$= d_A d_{\overline{A}}\cdot\operatorname{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B}\otimes\Phi_{A\overline{B}}\right)\left(\left(J^{\mathcal{E}}_{A\overline{A}}\right)^{\mathrm{T}}\otimes J^{\mathcal{D}^\dagger}_{B\overline{B}}\right)\right] \tag{4.20}$$

$$= d_{\overline{A}} d_B\cdot\operatorname{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B}\otimes\Phi_{A\overline{B}}\right)\left(\left(J^{\mathcal{E}}_{A\overline{A}}\right)^{\mathrm{T}}\otimes\frac{d_A}{d_B}\cdot J^{\mathcal{D}^\dagger}_{B\overline{B}}\right)\right], \tag{4.21}$$

where the transpose is taken with respect to the orthonormal basis of the maximally entangled state, and the dimensional factors come from Lemma 2.2.2. Due to the basic proprieties of the Choi-Jamiołkowski isomorphism discussed in Subsection 2.2.4, it is immediate to see that $(J^{\mathcal{E}}_{A\overline{A}})^{\mathrm{T}}$ can be identified with the $E_{A\overline{A}}$ of Lemma 4.1.3. In addition, we have $\frac{d_A}{d_B}\cdot J^{\mathcal{D}^\dagger}_{B\overline{B}}\succeq 0$, and tracing out the $\overline{B}$ system as well as using $d_A = d_{\overline{B}}$, we get

$$\frac{d_A}{d_B}\cdot J^{\mathcal{D}^\dagger}_{B} = \frac{d_A}{d_B}\cdot\mathcal{D}^\dagger\left(\frac{1_{\overline{B}}}{d_{\overline{B}}}\right) \tag{4.22}$$

$$= \frac{d_A}{d_B}\cdot\frac{1}{d_{\overline{B}}}\cdot 1_B \tag{4.23}$$

$$= \frac{1_B}{d_B}. \tag{4.24}$$

Thus, we can identify $\frac{d_A}{d_B}\cdot J^{\mathcal{D}^\dagger}_{B\overline{B}}$ with the $D_{B\overline{B}}$ of Lemma 4.1.3. □

The following simple dimension bounds hold for the channel fidelity.

Lemma 4.1.4. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M\in\mathbb{N}$. Then, we have*

$$0\leq F(\mathcal{N},M)\leq\min\left\{1,\left(\frac{d_{\overline{A}}}{M}\right)^2,\frac{d_B}{M}\right\}. \tag{4.25}$$

Proof. The lower bound is trivial and the upper bounds follow directly from the more general statements about the optimal fidelity under additional classical communication assistance as given in Lemma 4.3.4. □

By the linearity of the objective function, we can repeat the same approach followed in Section 3.5, and rewrite the channel fidelity as

$$F(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \operatorname{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)\left(\sum_{i\in I} p_i E^i_{A\overline{A}} \otimes D^i_{B\overline{B}}\right)\right] \tag{4.26}$$

$$s.t. \quad p_i \geq 0 \ \forall i \in I, \ \sum_{i\in I} p_i = 1 \tag{4.27}$$

$$E^i_{A\overline{A}} \succeq 0, \quad D^i_{B\overline{B}} \succeq 0 \tag{4.28}$$

$$E^i_A = \frac{1_A}{d_A}, \quad D^i_B = \frac{1_B}{d_B} \ \forall i \in I. \tag{4.29}$$

4.2 De Finetti theorems for quantum channels

Recall from Subsection 2.2.4 that a quantum channel is just a trace preserving completely positive (tpcp) map between two spaces of quantum states. Here, we establish a sufficient criterion under which permutation invariance of a quantum channel implies that it can be well approximated by a mixture of product quantum channels.

Theorem 4.2.1. *Let $\rho_{A\overline{A}(B\overline{B})^n_1}$ be a quantum state with*

$$\rho_{A\overline{A}(B\overline{B})^n_1} = \mathcal{U}^{\pi}_{(B\overline{B})^n_1}(\rho_{A\overline{A}(B\overline{B})^n_1}) \ \forall \pi \in \mathfrak{S}_n \tag{4.30}$$

$$\rho_{A(B\overline{B})^n_1} = \frac{1_A}{d_A} \otimes \rho_{(B\overline{B})^n_1} \tag{4.31}$$

$$\rho_{(B\overline{B})^{n-1}_1 B_n} = \rho_{(B\overline{B})^{n-1}_1} \otimes \frac{1_{B_n}}{d_B}. \tag{4.32}$$

Then, we have for $0 < k < n$ that

$$\left\| \rho_{A\overline{A}(B\overline{B})^k_1} - \sum_{i\in I} p_i \sigma^i_{A\overline{A}} \otimes \left(\omega^i_{B\overline{B}}\right)^{\otimes k} \right\|_1 \tag{4.33}$$

$$\leq k f(B\overline{B}|\cdot) \sqrt{(2\ln 2)\frac{\log\left(d_A d_{\overline{A}}\right) + (k-1)\log\left(d_B d_{\overline{B}}\right)}{n-k+1}} \tag{4.34}$$

with $\{p_i\}_{i\in I}$ *a probability distribution, and* $\sigma^i_{A\overline{A}}, \omega^i_{B\overline{B}} \succeq 0$ *such that* $\sigma^i_A = \frac{1_A}{d_A}$ *and* $\omega^i_B = \frac{1_B}{d_B}$ *for* $i \in I$.

Proof. The proof is a straightforward application of Theorem 3.6.6. In particular, referring to the notation of Theorem 3.6.6, we need to make the following choices for the systems $A := A\overline{A}$, $B := B\overline{B}$, $C_A := A$, $C_B := B$, for the operators $X_{C_A} := \frac{1_A}{d_A}$, $Y_{C_B} := \frac{1_B}{d_B}$, and for the linear maps $\Lambda_{A\overline{A}\to A} := \mathrm{Tr}_{\overline{A}}$, and $\Gamma_{B\overline{B}\to B} := \mathrm{Tr}_{\overline{B}}$. □

We emphasize that, according to the representation we obtain in this theorem, $\rho_{A\overline{A}(B\overline{B})^k_1}$ is close to a mixture of products of Choi states of completely positive *and* trace-preserving maps. We note that applying standard de Finetti theorems for quantum states would only show that $\rho_{A\overline{A}(B\overline{B})^k_1}$ is close to a mixture of products states, or in other words Choi states of completely positive maps that are in general not even trace-non-increasing. This is not sufficient for our applications, and having the constraints (4.31) and (4.32) are needed in our proofs to achieve this stronger statement. We discuss this in more detail by means of the following examples.

Example 4.2.2. *For* $\overline{A}$, $\overline{B}$ *trivial and* $k = 1$ *Theorem 4.2.1 says that* ρ_{AB} *is close to the product state* $\frac{1_{AB}}{d_A d_B}$, *as this is the only valid state satisfying the linear constraints. However, having only the permutation invariance condition* (4.30) *without the other two conditions* (4.32) *and* (4.31), *this conclusion does not hold. In fact, choose* $\rho_{AB^n_1}$ *to be maximally classically correlated between all the* $n+1$ *systems* $A, B_1, \ldots, B_n$

$$\rho_{AB^n_1} = \frac{1}{d}\sum_i |i\rangle\langle i|^{\otimes n+1}, \tag{4.35}$$

where $d := d_A = d_B$. *Clearly, the systems* B^n_1 *are symmetric with respect to* A, *i.e.,* $(\mathcal{I}_A \otimes \mathcal{U}^\pi_{B^n_1})(\rho_{AB^n_1}) = \rho_{AB^n_1}$ *for every* $\pi \in \mathfrak{S}_n$. *Thus, the permutation invariance condition* (4.30) *is satisfied. On the other hand, it is clear that the other two conditions* (4.31) *and* (4.32) *are not*

satisfied. For example, $\rho_{AB_1^n} \neq \frac{1_A}{d_A} \otimes \rho_{B_1^n}$. *Finally, we also notice that the conclusion of the theorem does not hold, since* ρ_{AB_1} *is not close to the state* $\frac{1_{AB_1}}{d_A d_B}$.

Example 4.2.3. *We now want to show that imposing the constraint* $\rho_{AB_1} = \frac{1_{AB_1}}{d_A d_B}$ *is not enough either. Let* $A, \overline{A}, B, \overline{B}$ *all be of dimension* $d \geq 2$. *Then, define for any* $n \geq 1$

$$\rho_{A\overline{A}(B\overline{B})_1^n} = \frac{1}{d^2} \sum_{i,j} |j\rangle\langle j|_A \otimes |i\rangle\langle i|_{\overline{A}} \otimes |i\rangle\langle i|_B^{\otimes n} \otimes |i\rangle\langle i|_{\overline{B}}^{\otimes n}. \tag{4.36}$$

Then, the state is invariant under permutations of the $B\overline{B}$ *systems and*

$$\begin{aligned} \rho_{AB_1} &= \frac{1}{d^2} \sum_{i,j} |j\rangle\langle j|_A \otimes |i\rangle\langle i|_{B_1} && (4.37)\\ &= \frac{1}{d^2} \sum_{j} |j\rangle\langle j|_A \otimes \sum_{i} |i\rangle\langle i|_{B_1} && (4.38)\\ &= \frac{1}{d^2} 1_A \otimes 1_{B_1} && (4.39)\\ &= \frac{1_{AB_1}}{d^2}. && (4.40) \end{aligned}$$

However, the reduced state $\rho_{A\overline{A}B_1\overline{B}_1}$ *is not close to states of the form*

$$\sum_{\ell} p_\ell \sigma^\ell_{A\overline{A}} \otimes \omega^\ell_{B_1\overline{B}_1} \quad \textit{with} \quad \sigma^\ell_A = \frac{1_A}{d}, \omega^\ell_{B_1} = \frac{1_{B_1}}{d}, \tag{4.41}$$

i.e., convex combinations of tensor products of Choi states. To see this, consider the projector $\Pi_{\overline{A}B_1} := \sum_i 1_A \otimes |i\rangle\langle i|_{\overline{A}} \otimes |i\rangle\langle i|_{B_1} \otimes 1_{\overline{B}_1}$. *Then, we get*

$$\mathrm{Tr}(\Pi_{\overline{A}B_1} \rho_{A\overline{A}B_1\overline{B}_1}) = \mathrm{Tr}(\rho_{A\overline{A}B_1\overline{B}_1}) = 1, \tag{4.42}$$

but

$$\begin{aligned} \mathrm{Tr}\left(\Pi_{\overline{A}B_1} \sum_{\ell} p_\ell \sigma^\ell_{A\overline{A}} \otimes \omega^\ell_{B_1\overline{B}_1}\right) &= \sum_{\ell} p_\ell \mathrm{Tr}(\Pi_{\overline{A}B_1} \sigma^\ell_{\overline{A}} \otimes \omega^\ell_{B_1}) && (4.43)\\ &= \sum_{\ell} p_\ell \mathrm{Tr}\left(\Pi_{\overline{A}B_1} \sigma^\ell_{\overline{A}} \otimes \frac{1_{B_1}}{d}\right) && (4.44) \end{aligned}$$

$$= \sum_\ell p_\ell \mathrm{Tr}(\sigma^\ell_{\overline{A}}) \frac{1}{d} = \sum_\ell p_\ell \frac{1}{d} = \frac{1}{d}. \tag{4.45}$$

Finally, using the relation between the trace distance and projectors (see [69, Section 9.2])

$$\left\| \rho_{A\overline{A}B_1\overline{B}_1} - \sum_\ell p_\ell \sigma^\ell_{A\overline{A}} \otimes \omega^\ell_{B_1\overline{B}_1} \right\|_1 \tag{4.46}$$

$$= 2 \max_{P_{A\overline{A}B_1\overline{B}_1}} \mathrm{Tr}\left[P_{A\overline{A}B_1\overline{B}_1} \left(\rho_{A\overline{A}B_1\overline{B}_1} - \sum_\ell p_\ell \sigma^\ell_{A\overline{A}} \otimes \omega^\ell_{B_1\overline{B}_1} \right) \right] \tag{4.47}$$

$$\geq \mathrm{Tr}\left[\Pi_{\overline{A}B_1} \left(\rho_{A\overline{A}B_1\overline{B}_1} - \sum_\ell p_\ell \sigma^\ell_{A\overline{A}} \otimes \omega^\ell_{B_1\overline{B}_1} \right) \right] \tag{4.48}$$

$$= 1 - \frac{1}{d}, \tag{4.49}$$

where the maximization is taken over all projectors $P_{A\overline{A}B_1\overline{B}_1}$.

By the Choi-Jamiołkowski isomorphism and relating the trace norm of Choi states to the diamond norm of the quantum channels (Lemma 2.2.4), we can alternatively state the bounds from Theorem 4.2.1 directly in terms of the quantum channels.

Corollary 4.2.4. *Let $\mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n}$ be a quantum channel such that*

$$\mathcal{U}^\pi_{\overline{B}_1^n} \left(\mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n}(\cdot) \right) = \mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n} \left(\mathcal{U}^\pi_{B_1^n}(\cdot) \right) \ \forall \pi \in \mathfrak{S}_n \tag{4.50}$$

$$\mathrm{Tr}_{\overline{B}_n} \left[\mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n}(\cdot) \right] = \mathrm{Tr}_{\overline{B}_n} \left[\mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n} \left(\mathrm{Tr}_{B_n}[\cdot] \otimes \frac{1_{B_n}}{d_B} \right) \right] \tag{4.51}$$

$$\mathrm{Tr}_{\overline{A}} \left[\mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n}(\cdot) \right] = \mathrm{Tr}_{\overline{A}} \left[\mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n} \left(\frac{1_A}{d_A} \otimes \mathrm{Tr}_A[\cdot] \right) \right]. \tag{4.52}$$

Then, we have for $0 < k < n$ with

$$\mathcal{N}_{AB_1^k \to \overline{A}\,\overline{B}_1^k}(X_{AB_1^k}) := \mathrm{Tr}_{\overline{B}_{k+1}^n} \left[\mathcal{N}_{AB_1^n \to \overline{A}\,\overline{B}_1^n} \left(X_{AB_1^k} \otimes \frac{1_{B_{k+1}^n}}{d_B^{n-k}} \right) \right] \tag{4.53}$$

that

$$\left\| \mathcal{N}_{AB_1^k \to \overline{A}\,\overline{B}_1^k} - \sum_{i \in I} p_i \mathcal{E}^i_{A \to \overline{A}} \otimes \left(\mathcal{D}^i_{B \to \overline{B}} \right)^{\otimes k} \right\|_\diamond \tag{4.54}$$

$$\leq \; d_A d_B^k \cdot k f(B\overline{B}|\cdot)\sqrt{(2\ln 2)\frac{\log\left(d_A d_{\overline{A}}\right) + (k-1)\log\left(d_B d_{\overline{B}}\right)}{n-k+1}} \tag{4.55}$$

with $\{p_i\}_{i\in I}$ *a probability distribution and* $\mathcal{D}^i_{B\to\overline{B}}, \mathcal{E}^i_{A\to\overline{A}}$ *quantum channels for* $i \in I$.

In (4.53) we chose a specific extension of $X_{AB_1^k}$ to define $\mathcal{N}_{AB_1^k\to\overline{A}\,\overline{B}_1^k}(X_{AB_1^k})$, namely $X_{AB_1^k} \otimes \frac{1_{B_{k+1}^n}}{d_B^{n-k}}$. This is still well-defined as the conditions (4.50) and (4.51) we require of $\mathcal{N}_{AB_1^n\to\overline{A}\,\overline{B}_1^n}$ actually say that the choice of extension does not matter. That is, we have for any $X_{AB_1^n}$ that

$$\begin{aligned}
\mathrm{Tr}_{\overline{B}_{k+1}^n}\left[\mathcal{N}_{AB_1^n\to\overline{A}\,\overline{B}_1^n}(X_{AB_1^n})\right] &= \mathrm{Tr}_{\overline{B}_{k+1}^{n-1}}\left[\mathrm{Tr}_{\overline{B}_n}\left[\mathcal{N}_{AB_1^n\to\overline{A}\,\overline{B}_1^n}\left(X_{AB_1^{n-1}}\otimes\frac{1_{B_n}}{d_B}\right)\right]\right] && (4.56)\\
&= \mathrm{Tr}_{\overline{B}_{k+1}^n}\left[\mathcal{N}_{AB_1^n\to\overline{A}\,\overline{B}_1^n}\left(X_{AB_1^k}\otimes\frac{1_{B_{k+1}^n}}{d_B^{n-k}}\right)\right] && (4.57)\\
&= \mathcal{N}_{AB_1^k\to\overline{A}\,\overline{B}_1^k}\left(X_{AB_1^k}\right), && (4.58)
\end{aligned}$$

where we used (4.51) for the first equality as well as (4.50) and (4.51) multiple times for the second equality. Thus, we can in fact choose

$$X_{AB_1^n} := X_{AB_1^k}\otimes\frac{1_{B_{k+1}^n}}{d_B^{n-k}}. \tag{4.59}$$

In the following we state several comments about de Finetti theorems for quantum channels:

- In contrast to the bound for Choi states (Theorem 4.2.1), the diamond norm bound in Corollary 4.2.4 does not have a polynomial dependence in d_B and k. We leave it as an open question to give a de Finetti theorem for quantum channels in terms of the diamond norm distance with a dimension dependence polynomial in d_B and k. (For our purposes we only need the $k = 1$ bound, in terms of the Choi states.)

- In the case $k = 1$, the conditions of the above theorem can be seen as approximations for the convex hull of product quantum channels, just as extendible states provide an

approximation for the set of separable states.[8] We note that in SDP hierarchies for the quantum separability problem the permutation invariance can be replaced by the stronger *Bose symmetric* condition [68]. That is, the state in question is supported on the symmetric subspace. The reason is that every separable state can without loss of generality be decomposed in a convex combination of *pure* product states. However, in our setting, we cannot assume that we have a mixture of a product of pure channels[9], and so we keep the more general notion of permutation invariance.

- In the following, we never directly make use of Corollary 4.2.4 but rather state it for connecting to the previous literature. In particular, when choosing $A\overline{A}$ trivial as a special case we find a finite version of the asymptotic de Finetti theorem for quantum channels from [39, 38].[10] We emphasize that our derived conditions then become a finite version of the notion of *exchangeable sequences of quantum channels* of [39] defined as a sequence of channels $\{\mathcal{N}_{B_1^n\to\overline{B}_1^n}\}$ satisfying for all n that

$$\mathcal{U}^{\pi}_{\overline{B}_1^n}\left(\mathcal{N}_{B_1^n\to\overline{B}_1^n}(\cdot)\right) = \mathcal{N}_{B_1^n\to\overline{B}_1^n}\left(\mathcal{U}^{\pi}_{B_1^n}(\cdot)\right) \ \forall \pi \in \mathfrak{S}_n \tag{4.60}$$

$$\mathcal{N}_{B_1^{n-1}\to\overline{B}_1^{n-1}}\left(\mathrm{Tr}_{B_n}[\cdot]\right) = \mathrm{Tr}_{\overline{B}_n}\left[\mathcal{N}_{B_1^n\to\overline{B}_1^n}(\cdot)\right]. \tag{4.61}$$

They show that under these conditions, for any k, the channel $\mathcal{N}_{B_1^k\to\overline{B}_1^k}$ is in the convex hull of tensor power channels. In Corollary 4.2.4, we start with a channel[11] $\mathcal{N}_{B_1^n\to\overline{B}_1^n}$ and quantify the closeness of such $\mathcal{N}_{B_1^k\to\overline{B}_1^k}$ to convex combinations of tensor product

[8] The class of channels we consider here is more restricted than general separable channels, which usually refers to a mixture of product completely positive and not necessarily trace-preserving maps [87].

[9] A *pure channel* is a quantum channel having associated a pure Choi state.

[10] We also refer to [70] for previous related work and [23] for a classical version. Moreover, following [56], conditions related to our (4.50) – (4.52) give rise to extendible channels in the resource theory of unextendibility.

[11] This is equivalent to being given a finite sequence $\mathcal{N}_{B_1^k\to\overline{B}_1^k}$ for $k \in \{1,\ldots,n\}$ satisfying the exchangeability condition, as the reduced channels are then completely determined by $\mathcal{N}_{B_1^n\to\overline{B}_1^n}$ (see [87])

channels $\sum_i p_i \left(\mathcal{D}^i_{B\to\overline{B}}\right)^{\otimes k}$.

Channels that are written as mixtures of product channels, i.e., channels of the form $\mathcal{E}_{A\to\overline{A}} \otimes \mathcal{D}_{B\to\overline{B}}$ where $\mathcal{E}_{A\to\overline{A}}$ and $\mathcal{D}_{B\to\overline{B}}$ are quantum channels, correspond to communication protocols in which the two parties have access to shared randomness but no communication [87, Section 6.1]. There is a natural relaxation to this set of channels, often called LOCC(1) channels [20], corresponding to channels that can be implemented with additional classical communication from A to B. Mathematically, a *LOCC(1) channel* is a tpcp map that can be written in the form

$$\sum_{i\in I} \mathcal{E}^i_{A\to\overline{A}} \otimes \mathcal{D}^i_{B\to\overline{B}}, \tag{4.62}$$

where $\mathcal{D}^i_{B\to\overline{B}}$ are channels, and $\mathcal{E}^i_{A\to\overline{A}}$ are completely positive but not necessarily trace-preserving maps, summing to a channel. I.e., $\sum_{i\in I} \mathcal{E}^i_{A\to\overline{A}}$ is a quantum channel. Channels of the form (4.62) are also known as *one-way right LOCC channels*, where Alice is the sender and is assumed to be on the left, while Bob, who is the receiver, is assumed to be on the right. We discuss this variation of approximate quantum error correction in Section 4.3.

4.2.1 Hierarchy of outer bounds

Following the de Finetti theorem for quantum channels as given in Theorem 4.2.1 for $k = 1$, the n-th level of the SDP hierarchy for the quantum channel fidelity becomes

$$\begin{aligned}
\mathrm{SDP}_n(\mathcal{N}, M) := \max \quad & d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B_1} \otimes \Phi_{A\overline{B}_1}\right) \rho_{A\overline{A}B_1\overline{B}_1}\right] && (4.63)\\
s.t. \quad & \rho_{A\overline{A}(B\overline{B})^n_1} \succeq 0,\ \mathrm{Tr}\left[\rho_{A\overline{A}(B\overline{B})^n_1}\right] = 1 && (4.64)\\
& \rho_{A\overline{A}(B\overline{B})^n_1} = \mathcal{U}^{\pi}_{(B\overline{B})^n_1}\left(\rho_{A\overline{A}(B\overline{B})^n_1}\right)\ \forall \pi \in \mathfrak{S}_n && (4.65)\\
& \rho_{A(B\overline{B})^n_1} = \frac{1_A}{d_A} \otimes \rho_{(B\overline{B})^n_1} && (4.66)
\end{aligned}$$

$$\rho_{A\overline{A}(B\overline{B})_1^{n-1}B_n} = \rho_{A\overline{A}(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}. \tag{4.67}$$

Here, we identified $B_1 := B$ and hence the n-th level of the hierarchy then corresponds to taking $n-1$ extensions. Note that instead of stating the last condition for the final block B_n we could have equivalently stated it for any block B_j with $j = 1, \ldots, n$ (by the permutation invariance). Iteratively, the condition then also holds on all pairs of blocks of size two, and so on. Moreover, we slightly strengthened the last condition by including the A-systems compared to the minimal condition needed for Theorem 4.2.1, i.e.,

$$\rho_{(B\overline{B})_1^{n-1}B_n} = \rho_{(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}. \tag{4.68}$$

We then immediately have asymptotic convergence.

Theorem 4.2.5. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $n, M \in \mathbb{N}$. Then, we have*

$$0 \leq \mathrm{SDP}_n(\mathcal{N}, M) - F(\mathcal{N}, M) \leq \frac{\mathrm{poly}(d)}{\sqrt{n}} \tag{4.69}$$

implying

$$F(\mathcal{N}, M) = \lim_{n\to\infty} \mathrm{SDP}_n(\mathcal{N}, M), \tag{4.70}$$

where $d := \max\{d_A, d_{\overline{A}}, d_B, d_{\overline{B}}\}$.

Proof. By construction $0 \leq \mathrm{SDP}_n(\mathcal{N}, M) - F(\mathcal{N}, M)$, and the remaining inequality arises from

$$\begin{aligned}
&d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right) \rho_{A\overline{A}B\overline{B}}\right] && (4.71)\\
&\quad = d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right) \left(E_{A\overline{A}} \otimes D_{B\overline{B}}\right)\right] && (4.72)\\
&\quad + d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right) \left(\rho_{A\overline{A}B\overline{B}} - E_{A\overline{A}} \otimes D_{B\overline{B}}\right)\right] && (4.73)\\
&\quad \leq d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right) \left(E_{A\overline{A}} \otimes D_{B\overline{B}}\right)\right] && (4.74)
\end{aligned}$$

$$+ d_{\overline{A}} d_B \cdot \|J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\|_\infty \cdot \|\rho_{A\overline{A}B\overline{B}} - E_{A\overline{A}} \otimes D_{B\overline{B}}\|_1 \tag{4.75}$$

$$\leq d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)\left(E_{A\overline{A}} \otimes D_{B\overline{B}}\right)\right] + \frac{\mathrm{poly}(d)}{\sqrt{n}}, \tag{4.76}$$

where we used the linearity of the objective function as described in Section 3.5, the $(\infty, 1)$ Hölder's inequality and Theorem 4.2.1 with $k = 1$. □

We note that the worst case convergence guarantee is "fairly slow", as to ensure that the approximation error becomes small, we need at least the level $n = \mathrm{poly}(d)$. As already pointed out in Section 3.7, this slow convergence in the worst case is as expected from the quantum separability problem. However, in practice the convergence speed may be much better. We will numerically analyse in detail this aspect in Section 4.4.

Remark 4.2.6. *Instead of extending the B-systems we could alternatively extend the A-systems, which leads to the non-equivalent asymptotically converging hierarchy*

$$\overline{\mathrm{SDP}}_n(\mathcal{N}, M) := \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}_1 B} \otimes \Phi_{A_1\overline{B}}\right)\rho_{A_1\overline{A}_1 B\overline{B}}\right] \tag{4.77}$$

$$s.t. \quad \rho_{(A\overline{A})_1^n B\overline{B}} \succeq 0,\ \mathrm{Tr}\left[\rho_{(A\overline{A})_1^n B\overline{B}}\right] = 1 \tag{4.78}$$

$$\rho_{(A\overline{A})_1^n B\overline{B}} = \mathcal{U}^\pi_{(A\overline{A})_1^n}\left(\rho_{(A\overline{A})_1^n B\overline{B}}\right)\ \forall \pi \in \mathfrak{S}_n \tag{4.79}$$

$$\rho_{(A\overline{A})_1^n B} = \rho_{(A\overline{A})_1^n} \otimes \frac{1_B}{d_B} \tag{4.80}$$

$$\rho_{(A\overline{A})_1^{n-1} A_n B\overline{B}} = \frac{1_{A_n}}{d_A} \otimes \rho_{(A\overline{A})_1^{n-1} B\overline{B}}. \tag{4.81}$$

For the first level we have

$$\overline{\mathrm{SDP}}_1(\mathcal{N}, M) = \mathrm{SDP}_1(\mathcal{N}, M), \tag{4.82}$$

by inspection. However, for the higher levels it depends on the input-output dimensions $d_{\overline{A}}, d_B$ which hierarchy is potentially more powerful, i.e., faster to converge.

The relaxations $\mathrm{SDP}_n(\mathcal{N}, M)$ behave naturally with respect to the first two bounds of Lemma 4.1.4.

Lemma 4.2.7. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $n, M \geq 1$. Then, we have*

$$0 \leq \mathrm{SDP}_n(\mathcal{N}, M) \leq \min\left\{1, \left(\frac{d_{\overline{A}}}{M}\right)^2\right\}. \tag{4.83}$$

Proof. The lower bound is trivial. By the monotonicity in n (Theorem 4.2.5), it is enough to restrict to $n = 1$ for the upper bounds. Alternatively, the upper bound of one can directly be deduced operationally from [65, Theorem 3], where $\mathrm{SDP}_1(\mathcal{N}, M)$ was identified as the non-signalling assisted channel fidelity. We use that for any bipartite quantum state ρ_{XY} we have[12] [79, Lemma A.2]

$$d_X \cdot 1_X \otimes \rho_Y \succeq \rho_{XY}. \tag{4.84}$$

For the first upper bound we find $\frac{d_{\overline{B}}}{d_B} \cdot \rho_{A\overline{A}} \otimes 1_{B_1\overline{B}_1} \succeq \rho_{A\overline{A}B_1\overline{B}_1}$, which gives for the objective function

$$\begin{aligned}
\mathrm{SDP}_1(\mathcal{N}, M) &\leq d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B_1} \otimes \Phi_{A\overline{B}_1}\right)\left(\frac{d_{\overline{B}}}{d_B} \cdot \rho_{A\overline{A}} \otimes 1_{B_1\overline{B}_1}\right)\right] && (4.85)\\
&= d_{\overline{A}} d_{\overline{B}} \cdot \mathrm{Tr}\left[\left(\frac{1_A}{d_A} \otimes \frac{1_{\overline{A}}}{d_{\overline{A}}}\right)\rho_{A\overline{A}}\right] && (4.86)\\
&= \mathrm{Tr}\left(\rho_{A\overline{A}}\right) = 1. && (4.87)
\end{aligned}$$

For the second upper bound we find similarly as for the first upper bound $\frac{d_{\overline{A}}}{d_A} \cdot 1_{A\overline{A}} \otimes \rho_{B_1\overline{B}_1} \succeq \rho_{A\overline{A}B_1\overline{B}_1}$, which then leads to the claim by the same argument as for the second upper bound in Lemma 4.1.4. □

[12] It is interesting to note that, while $d_X \cdot 1_X \otimes \rho_Y \succeq \rho_{XY}$ is always valid, removing d_X leads to the so-called *reduction criterion*. The reduction criterion, which does not always hold, is connected to the separability problem for low-dimensional quantum systems [49].

We can again add all the PPT constraints and denote the resulting relaxations by $\mathrm{SDP}_{n,\mathrm{PPT}}(\mathcal{N}, M)$. In the following we study more closely these levels $\mathrm{SDP}_{n,\mathrm{PPT}}(\mathcal{N}, M)$, which are our tightest outer bound relaxations on the channel fidelity.

4.2.2 Low level relaxations

For $n = 1$, we find the first-level relaxation

$$\mathrm{SDP}_{1,\mathrm{PPT}}(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{AB} \otimes \Phi_{A\overline{B}}\right) \rho_{A\overline{A}B\overline{B}}\right] \tag{4.88}$$

$$s.t. \quad \rho_{A\overline{A}B\overline{B}} \succeq 0, \ \mathrm{Tr}\left(\rho_{A\overline{A}B\overline{B}}\right) = 1 \tag{4.89}$$

$$\rho_{AB\overline{B}} = \frac{1_A}{d_A} \otimes \rho_{B\overline{B}} \tag{4.90}$$

$$\rho_{A\overline{A}B} = \rho_{A\overline{A}} \otimes \frac{1_B}{d_B} \tag{4.91}$$

$$\rho^{\mathrm{T}_{B\overline{B}}}_{A\overline{A}B\overline{B}} \succeq 0, \tag{4.92}$$

which is the SDP outer bound[13] found in [65, Section IV], up to their a priori stronger condition

$$\rho_{AB} = \frac{1_{AB}}{d_A d_B} \text{ instead of our } \mathrm{Tr}\left(\rho_{A\overline{A}B\overline{B}}\right) = 1. \tag{4.93}$$

However, as implicitly shown in [65, Theorem 3] these two conditions actually become equivalent because of the structure of the objective function. Operationally $\mathrm{SDP}_1(\mathcal{N}, M)$ corresponds to the non-signalling assisted channel fidelity, whereas $\mathrm{SDP}_{1,\mathrm{PPT}}(\mathcal{N}, M)$ adds the PPT-preserving constraint — as discussed in [65, Corollary 4]. Moreover, in the objective function the symmetry[14]

$$\int \left(\overline{U}_A \otimes U_{\overline{B}}\right) (\cdot) \left(\overline{U}_A \otimes U_{\overline{B}}\right)^\dagger \, \mathrm{d}U \tag{4.94}$$

[13] In the introduction we referred to this semidefinite programming relation with the notation $\mathrm{SDP}(\mathcal{N}, M)$.

[14] Here, $\overline{U}_A$ denotes the complex conjugate of U_A with respect to some standard basis. The super-operator (4.94) is commonly known as the *isotropic twirling channel* (see [87, Example 7.25]).

can be used to achieve a dimension reduction[15] of M^2 leading to [65, Theorem 3]

$$\begin{aligned}
\mathrm{SDP}_{1,\mathrm{PPT}}(\mathcal{N}, M) = \max \quad & d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[J^{\mathcal{N}}_{\overline{A}B} Y_{\overline{A}B}\right] && (4.95)\\
s.t. \quad & \rho_{\overline{A}} \otimes \frac{1_B}{d_B} \succeq Y_{\overline{A}B} \succeq 0,\ \mathrm{Tr}(\rho_{\overline{A}}) = 1 && (4.96)\\
& M^2 \cdot Y_B = \frac{1_B}{d_B} && (4.97)\\
& \rho_{\overline{A}} \otimes \frac{1_B}{d_B} \succeq M \cdot Y^{\mathrm{T}_B}_{\overline{A}B} \succeq -\rho_{\overline{A}} \otimes \frac{1_B}{d_B}. && (4.98)
\end{aligned}$$

The level $n = 2$ reads as

$$\begin{aligned}
\mathrm{SDP}_{2,\mathrm{PPT}}(\mathcal{N}, M) = \max \quad & d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B_1} \otimes \Phi_{A\overline{B}_1}\right) \rho_{A\overline{A}B_1\overline{B}_1}\right] && (4.99)\\
s.t. \quad & \rho_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2} \succeq 0,\ \mathrm{Tr}\left(\rho_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2}\right) = 1 && (4.100)\\
& \mathcal{U}^{\pi}_{B_1B_2\overline{B}_1\overline{B}_2}\left(\rho_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2}\right) = \rho_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2}\ \forall \pi \in \mathfrak{S}_2 && (4.101)\\
& \rho_{AB_1B_2\overline{B}_1\overline{B}_2} = \frac{1_A}{d_A} \otimes \rho_{B_1B_2\overline{B}_1\overline{B}_2} && (4.102)\\
& \rho_{A\overline{A}B_1B_2\overline{B}_1} = \rho_{A\overline{A}B_1\overline{B}_1} \otimes \frac{1_{B_2}}{d_B} && (4.103)\\
& \rho^{\mathrm{T}_{A\overline{A}}}_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2} \succeq 0,\ \rho^{\mathrm{T}_{B_2\overline{B}_2}}_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2} \succeq 0. && (4.104)
\end{aligned}$$

Numerical evaluations of (4.95) and (4.99) can be found in Section 4.4.

4.3 Classically-assisted approximate quantum error correction

4.3.1 Setting

It is often useful to add classical forward communication assistance to the problem of quantum error correction. The corresponding assisted channel fidelity is defined as follows.

[15] The use of the isotropic twirling channel allows us to remove the quantum systems A and $\overline{B}$. Since $M = d_A = d_{\overline{B}}$, the achieved dimension reduction is of M^2.

Definition 4.3.1. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. The LOCC(1)-assisted channel fidelity for message dimension M is defined as*

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) := \max \quad F\Big(\Phi_{\overline{B}R}, \sum_{i\in I} \big(\big(\mathcal{D}^i_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}^i_{A\to\overline{A}}\big) \otimes \mathcal{I}_R\big)(\Phi_{AR})\Big) \tag{4.105}$$

$$s.t. \quad \sum_{i\in I} \mathcal{E}^i_{A\to\overline{A}} \text{ quantum channel with } \mathcal{E}^i_{A\to\overline{A}} \text{ cp for } i \in I \tag{4.106}$$

$$\mathcal{D}^i_{B\to\overline{B}} \text{ quantum channel } \forall i \in I, \tag{4.107}$$

where Φ_{AR} denotes the maximally entangled state on AR, cp is the abbreviation for completely positive, and we have $M = d_A = d_{\overline{B}} = d_R$.

By the Choi-Jamiołkowski isomorphism this can again be rewritten as a bilinear optimization.

Lemma 4.3.2. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. Then, the LOCC(1)-assisted channel fidelity can be written as*

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right) \left(\sum_{i\in I} E^i_{A\overline{A}} \otimes D^i_{B\overline{B}} \right) \right] \tag{4.108}$$

$$s.t. \quad E^i_{A\overline{A}} \succeq 0, \; D^i_{B\overline{B}} \succeq 0 \quad \forall i \in I \tag{4.109}$$

$$\sum_{i\in I} E^i_A = \frac{1_A}{d_A} \tag{4.110}$$

$$D^i_B = \frac{1_B}{d_B} \quad \forall i \in I. \tag{4.111}$$

The proof follows the same steps as in Lemma 4.1.3 about plain quantum error correction, and is based on the manipulation of the objective function

$$F\Big(\Phi_{\overline{B}R}, \sum_{i\in I} \big(\big(\mathcal{D}^i_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}^i_{A\to\overline{A}}\big) \otimes \mathcal{I}_R\big)(\Phi_{AR})\Big) \tag{4.112}$$

by using the Choi-Jamiołkowski isomorphism. As we show in the following lemma, we have that $F^{\mathrm{LOCC}(1)}(\mathcal{N}, M)$ is closely connected to the channel fidelity $F(\mathcal{N}, M)$.

Lemma 4.3.3. *Let $\mathcal{N}$ be a quantum channel and $M \in \mathbb{N}$. Then, we have*

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \geq F(\mathcal{N}, M) \geq \left(F^{\mathrm{LOCC}(1)}(\mathcal{N}, M)\right)^2. \tag{4.113}$$

Asymptotically this corresponds to the well-known statement that forward classical communication assistance does not increase the capacity [9].

Proof. The first inequality is trivial because the addition of a forward classical communication channel cannot decrease the channel fidelity.

The fact that $\left(F^{\mathrm{LOCC}(1)}(\mathcal{N}, M)\right)^2$ gives a lower bound on $F(\mathcal{N}, M)$ can be seen from [62, Proposition 4.5]. Consider an arbitrary coding scheme for the quantum channel $\mathcal{N}$ assisted with a forward classical communication channel and call $\mathcal{F}_{\mathrm{LOCC}(1)}$ the channel fidelity obtained using that scheme. We then want to show that it is always possible to find a coding scheme for the quantum channel $\mathcal{N}$ alone allowing us to achieve a channel fidelity $\mathcal{F} \geq \mathcal{F}^2_{\mathrm{LOCC}(1)}$. Say we are able to send, through the forward classical communication channel, a symbol in the set $\{1, \ldots, S\}$ with $S \in \mathbb{N}$. An arbitrary coding scheme for the assisted quantum channel can be modelled by a collection of instruments $\{\mathcal{E}^s_{A\to\overline{A}}\}_{s\in\{1,\ldots,S\}}$, i.e., trace-nonincreasing cp maps summing up to a channel, and channels $\{\mathcal{D}^s_{B\to\overline{B}}\}_{s\in\{1,\ldots,S\}}$. It is then easy to show that there must exist a symbol $\tilde{s}$ such that the fidelity of the map $\mathcal{D}^{\tilde{s}} \circ \mathcal{N} \circ \frac{\mathcal{E}^{\tilde{s}}}{e^{\tilde{s}}}$ is lower bounded by $\mathcal{F}_{\mathrm{LOCC}(1)}$, where the factor $e^{\tilde{s}}$ is chosen such that the completely positive map $\frac{\mathcal{E}^{\tilde{s}}}{e^{\tilde{s}}}$ becomes trace preserving with respect to the maximally mixed state $\frac{1_A}{d_A}$, as done in [62, Proposition 5.1]. Using the polar decomposition it is possible to find an isometric encoder[16] $\mathcal{V}^{\tilde{s}}$ such that

[16] An *isometric channel* $\mathcal{V} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ is a quantum channel that can be expressed as $\mathcal{V}(\cdot) = V(\cdot)V^\dagger$, where $V : \mathcal{H} \to \mathcal{H}'$ is an isometry, i.e., a norm-preserving operator. In practice, an isometric channel is a generalization of a *unitary channel*, where the dimensions of the two Hilbert spaces do not need to be equal. More precisely, we only need $d_{\mathcal{H}} \leq d_{\mathcal{H}'}$.

the channel fidelity $\mathcal{F}$ obtained using the coding scheme with encoder $\mathcal{V}^{\tilde{s}}$ and decoder $\mathcal{D}^{\tilde{s}}$ is lower bounded by the squared fidelity of the map $\mathcal{D}^{\tilde{s}} \circ \mathcal{N} \circ \frac{\mathcal{E}^{\tilde{s}}}{e^{\tilde{s}}}$. This implies $\mathcal{F} \geq \mathcal{F}^2_{\mathrm{LOCC}(1)}$. □

We have the dimension bounds for the LOCC(1)-assisted setting. Notice that the following result readily implies Lemma 4.1.4.

Lemma 4.3.4. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. Then, we have*

$$0 \leq F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq \min\left\{1, \left(\frac{d_{\overline{A}}}{M}\right)^2, \frac{d_B}{M}\right\}. \tag{4.114}$$

Proof. The lower bound is trivial. For the upper bounds, we use that for any quantum state ρ_{XY} we have [79, Lemma A.2]

$$d_X \cdot 1_X \otimes \rho_Y \succeq \rho_{XY}. \tag{4.115}$$

Now, for the first upper bound note that $\frac{d_{\overline{B}}}{d_B} \cdot 1_{B\overline{B}} = d_{\overline{B}} \cdot 1_{\overline{B}} \otimes D^i_B \succeq D^i_{B\overline{B}}$ for all $i \in I$, and hence we get for the objective function (with $d_A = d_{\overline{B}} = M$)

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq d_{\overline{A}} d_{\overline{B}} \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)^{1/2} \left(\sum_{i\in I} E^i_{A\overline{A}} \otimes 1_{B\overline{B}}\right) \left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)^{1/2}\right] \tag{4.116}$$

$$= d_{\overline{A}} d_{\overline{B}} \cdot \mathrm{Tr}\left[\left(\frac{1_{\overline{A}}}{d_{\overline{A}}} \otimes \frac{1_A}{d_A}\right) \sum_{i\in I} E^i_{A\overline{A}}\right] \tag{4.117}$$

$$= \mathrm{Tr}\left[\sum_{i\in I} E^i_{A\overline{A}}\right] \tag{4.118}$$

$$= 1. \tag{4.119}$$

For the second upper bound, note that from $E^i_{A\overline{A}} \succeq 0,\ D^i_{B\overline{B}} \succeq 0$ we get

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq \max\ d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right) \left(\sum_{i\in I} E^i_{A\overline{A}} \otimes \sum_{j\in I} D^j_{B\overline{B}}\right)\right]. \tag{4.120}$$

Now, we employ that $d_{\overline{A}} \cdot E^i_A \otimes 1_{\overline{A}} \succeq E^i_{A\overline{A}}$ giving $\frac{d_{\overline{A}}}{d_A} \cdot 1_{A\overline{A}} \succeq \sum_{i\in I} E^i_{A\overline{A}}$, which in turn leads to

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq \frac{d_{\overline{A}}^2 d_B}{d_A} \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)\left(1_{A\overline{A}} \otimes \sum_{j\in I} D^j_{B\overline{B}}\right)\right] \tag{4.121}$$

$$= \frac{d_{\overline{A}}^2 d_B}{d_A} \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_B \otimes \frac{1_{\overline{B}}}{d_{\overline{B}}}\right)\sum_{j\in I} D^j_{B\overline{B}}\right] \tag{4.122}$$

$$= \frac{d_{\overline{A}}^2 d_B}{d_A^2 d_{\overline{B}}} \cdot \mathrm{Tr}\left[J^{\mathcal{N}}_B \sum_{j\in I} D^j_B\right] \tag{4.123}$$

$$= \frac{d_{\overline{A}}^2 d_B}{d_A^2 d_{\overline{B}}} \cdot \mathrm{Tr}\left[J^{\mathcal{N}}_B d_A \frac{1_B}{d_B}\right] \tag{4.124}$$

$$= \frac{d_{\overline{A}}^2}{d_A d_{\overline{B}}}. \tag{4.125}$$

For the third upper bound, note that $1_{B\overline{B}} \succeq D^i_{B\overline{B}}$ and thus

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right)\left(\sum_{i\in I} E^i_{A\overline{A}} \otimes 1_{B\overline{B}}\right)\right] \tag{4.126}$$

$$= d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(\frac{1_{\overline{A}}}{d_{\overline{A}}} \otimes \frac{1_A}{d_A}\right)\sum_{i\in I} E^i_{A\overline{A}}\right] \tag{4.127}$$

$$= \frac{d_B}{d_A} \cdot \mathrm{Tr}\left[\sum_{i\in I} E^i_{A\overline{A}}\right] \tag{4.128}$$

$$= \frac{d_B}{d_A}. \tag{4.129}$$

□

4.3.2 Hierarchy of outer bounds

Following what we did in Theorem 4.2.1, we get the following approximation for the set of LOCC(1) channels, stated in terms of the corresponding Choi states.

Proposition 4.3.5. *Let $\rho_{A\overline{A}(B\overline{B})^n_1}$ be a quantum state with*

$$\rho_{A\overline{A}(B\overline{B})^n_1} = \mathcal{U}^{\pi}_{(B\overline{B})^n_1}(\rho_{A\overline{A}(B\overline{B})^n_1}) \; \forall \pi \in \mathfrak{S}_n \tag{4.130}$$

$$\rho_A = \frac{1_A}{d_A} \tag{4.131}$$

$$\rho_{(B\overline{B})_1^{n-1}B_n} = \rho_{(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}. \tag{4.132}$$

Then, we have for $0 < k < n$ *that* $\left\| \rho_{A\overline{A}(B\overline{B})_1^k} - \sum_{i\in I} \sigma^i_{A\overline{A}} \otimes \left(\omega^i_{B\overline{B}}\right)^{\otimes k} \right\|_1$ *is upper bounded by the same term as in Theorem 4.2.1, where* $\omega^i_{B\overline{B}} \succeq 0$ *with* $\omega^i_B = \frac{1_B}{d_B}$ *and* $\sigma^i_{A\overline{A}} \succeq 0$ *with* $\sum_{i\in I} \sigma^i_A = \frac{1_A}{d_A}$.

Comparing the above proposition with Theorem 4.2.1, we see that the condition $\rho_A = \frac{1_A}{d_A}$ replaces the previous $\rho_{A(B\overline{B})_1^n} = \frac{1_A}{d_A} \otimes \rho_{(B\overline{B})_1^n}$. This is because the constraint we have now to reproduce on the states is $\sum_{i\in I} \sigma^i_A = \frac{1_A}{d_A}$, while before we had $\sigma^i_A = \frac{1_A}{d_A}$ for every $i \in I$.

The n-th level of the SDP hierarchy then becomes

$$\mathrm{SDP}_n^{\mathrm{LOCC}(1)}(\mathcal{N}, M) := \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B_1} \otimes \Phi_{A\overline{B}_1}\right) \rho_{A\overline{A}B_1\overline{B}_1}\right] \tag{4.133}$$

$$s.t. \quad \rho_{A\overline{A}(B\overline{B})_1^n} \succeq 0 \tag{4.134}$$

$$\mathcal{U}^{\pi}_{(B\overline{B})_1^n}\left(\rho_{A\overline{A}(B\overline{B})_1^n}\right) = \rho_{A\overline{A}(B\overline{B})_1^n} \quad \forall \pi \in \mathfrak{S}_n \tag{4.135}$$

$$\rho_{AB_1^n} = \frac{1_{AB_1^n}}{d_A d_B^n} \tag{4.136}$$

$$\rho_{A\overline{A}(B\overline{B})_1^{n-1}B_n} = \rho_{A\overline{A}(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}. \tag{4.137}$$

By inspection, the only difference between $\mathrm{SDP}_n(\mathcal{N}, M)$ and $\mathrm{SDP}_n^{\mathrm{LOCC}(1)}(\mathcal{N}, M)$ is the weakened second to last condition. The asymptotic convergence follows immediately from Proposition 4.3.5.

Theorem 4.3.6. *Let* $\mathcal{N}$ *be a quantum channel and* $n, M \in \mathbb{N}$. *Then, we have*

$$\mathrm{SDP}_{n+1}^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq \mathrm{SDP}_n^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \tag{4.138}$$

and,

$$F^{\mathrm{LOCC}(1)}(\mathcal{N}, M) = \lim_{n\to\infty} \mathrm{SDP}_n^{\mathrm{LOCC}(1)}(\mathcal{N}, M). \tag{4.139}$$

Note that for $\mathrm{SDP}_n^{\mathrm{LOCC}(1)}(\mathcal{N}, M)$ we slightly strengthened the last two conditions by including some more A- and B-systems in the conditions compared to the minimal conditions

$$\rho_A = \frac{1_A}{d_A} \quad \text{and} \quad \rho_{(B\overline{B})_1^{n-1}B_n} = \rho_{(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B} \tag{4.140}$$

needed for Proposition 4.3.5. By an iterative argument the last condition implies in particular that

$$\rho_{A\overline{A}B_1^n\overline{B}_1} = \rho_{A\overline{A}B_1\overline{B}_1} \otimes \frac{1_{B_2^n}}{d_B^n}, \tag{4.141}$$

which together with the other three conditions in $\mathrm{SDP}_n^{\mathrm{LOCC}(1)}(\mathcal{N}, M)$ then corresponds to the notion of *extendible channels* from [56, Definition 5] (also see [31] for similar conditions). We note, however, that when relaxing the conditions to n-extendible channels our proofs for the asymptotic convergence of the resulting outer bounds do not apply.

The SDP relaxations again behave naturally in the sense that they are upper bounded by one.

Lemma 4.3.7. *Let $\mathcal{N}$ be a quantum channel and $n, M \in \mathbb{N}$. Then, we have*

$$0 \leq \mathrm{SDP}_n^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq 1. \tag{4.142}$$

Proof. The lower bound is trivial. For the upper bound, by the monotonicity in n (Theorem 4.3.6) it is enough to restrict to $n = 1$. As in the proof of Lemma 4.3.4, we make use of $\frac{d_{\overline{B}}}{d_B} \cdot \rho_{A\overline{A}} \otimes 1_{B_1\overline{B}_1} \succeq \rho_{A\overline{A}B_1\overline{B}_1}$. This again gives

$$\mathrm{SDP}_1^{\mathrm{LOCC}(1)}(\mathcal{N}, M) \leq d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J_{\overline{A}B}^{\mathcal{N}} \otimes \Phi_{A\overline{B}}\right) \frac{d_{\overline{B}}}{d_B} \cdot \rho_{A\overline{A}} \otimes 1_{B_1\overline{B}_1}\right] \tag{4.143}$$

$$= 1. \tag{4.144}$$

□

As before, we can again add PPT constraints and we denote the resulting relaxations by $\mathrm{SDP}^{\mathrm{LOCC}(1)}_{n,\mathrm{PPT}}(\mathcal{N}, M)$. In the following we study more closely these levels $\mathrm{SDP}^{\mathrm{LOCC}(1)}_{n,\mathrm{PPT}}(\mathcal{N}, M)$, which are our tightest outer bound relaxations on the LOCC(1)-assisted channel fidelity. We find

$$\mathrm{SDP}^{\mathrm{LOCC}(1)}_{1,\mathrm{PPT}}(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B} \otimes \Phi_{A\overline{B}}\right) \rho_{A\overline{A}B\overline{B}}\right] \tag{4.145}$$

$$s.t. \quad \rho_{A\overline{A}B\overline{B}} \succeq 0 \tag{4.146}$$

$$\rho_{AB} = \frac{1_{AB}}{d_A d_B} \tag{4.147}$$

$$\rho_{A\overline{A}B} = \rho_{A\overline{A}} \otimes \frac{1_B}{d_B} \tag{4.148}$$

$$\rho^{\mathrm{T}_{B\overline{B}}}_{A\overline{A}B\overline{B}} \succeq 0 \tag{4.149}$$

This is exactly the SDP outer bound found in [65, Section IV], which simplifies to

$$\mathrm{SDP}^{\mathrm{LOCC}(1)}_{1,\mathrm{PPT}}(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[J^{\mathcal{N}}_{\overline{A}B} X_{\overline{A}B}\right] \tag{4.150}$$

$$s.t. \quad \rho_{\overline{A}} \otimes \frac{1_B}{d_B} \succeq X_{\overline{A}B} \succeq 0, \ \mathrm{Tr}[\rho_{\overline{A}}] = 1 \tag{4.151}$$

$$\rho_{\overline{A}} \otimes \frac{1_B}{d_B} \succeq M \cdot X^{\mathrm{T}_B}_{\overline{A}B} \succeq -\rho_{\overline{A}} \otimes \frac{1_B}{d_B}. \tag{4.152}$$

By inspection, this corresponds to $\mathrm{SDP}_{1,\mathrm{PPT}}(\mathcal{N}, M)$ but with one missing constraint, namely $M^2 X_B = \frac{1_B}{d_B}$. For $n = 2$ we get

$$\mathrm{SDP}^{\mathrm{LOCC}(1)}_{2,\mathrm{PPT}}(\mathcal{N}, M) = \max \quad d_{\overline{A}} d_B \cdot \mathrm{Tr}\left[\left(J^{\mathcal{N}}_{\overline{A}B_1} \otimes \Phi_{A\overline{B}_1}\right) \rho_{A\overline{A}B_1\overline{B}_1}\right] \tag{4.153}$$

$$s.t. \quad \rho_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2} \succeq 0 \tag{4.154}$$

$$\mathcal{U}^{\pi}_{B_1B_2\overline{B}_1\overline{B}_2}\left(\rho_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2}\right) = \rho_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2} \quad \forall \pi \in \mathfrak{S}_2 \tag{4.155}$$

$$\rho_{AB_1B_2} = \frac{1_{AB_1B_2}}{d_A d_B^2} \tag{4.156}$$

$$\rho_{A\overline{A}B_1B_2\overline{B}_1} = \rho_{A\overline{A}B_1\overline{B}_1} \otimes \frac{1_{B_2}}{d_B} \tag{4.157}$$

$$\rho^{\mathrm{T}_{A\overline{A}}}_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2} \succeq 0,\ \rho^{\mathrm{T}_{B_2\overline{B}_2}}_{A\overline{A}B_1B_2\overline{B}_1\overline{B}_2} \succeq 0, \tag{4.158}$$

and we recover the exact same conditions as for the notion of extendible channels [56, Definition 5].

4.4 Numerical examples

4.4.1 Methods

In the following we present the proof of concept numerics we implemented to test the low levels of our hierarchy for the application of approximate quantum error correction. Moreover, given the size of the programs, our focus is limited to qubit and qutrit channels. In order to explore more complex quantum channels, or to increase the number of channel repetitions, one needs to simplify further the optimization programs, by taking advantage of the potential symmetries of the particular noise model. We do that for the qubit depolarizing channel in Subsection 4.4.4.

The experiments have been done in MATLAB using the QETLAB library [55], CVX [40], MOSEK [1], and SDPT3 [78].[17]

Remark 4.4.1. *In Remark 2.2.1 we introduced subscripts to keep track of the systems the operators act on. Moreover, the usage of subscripts has allowed us to make implicit any isometry needed to rearrange the underlying Hilbert spaces. For example, the expression* $W_{\overline{A}B}Q_{B\overline{A}}$ *must be interpreted as* $W_{\overline{A}B}F_{\overline{A}\leftrightarrow B}Q_{B\overline{A}}$*, where* $F_{\overline{A}\leftrightarrow B} : B \otimes \overline{A} \to \overline{A} \otimes B$ *is the swap operator exchanging* $\overline{A}$ *with* B*. This is a standard convention, which has been used through this thesis. However, one needs to be careful when implementing the optimization programs for numerical purposes. In fact, while expressions such as* $W_{\overline{A}B}Q_{B\overline{A}}$ *make perfectly sense according to our conventions,*

[17] All the code is available at https://github.com/FrancescoBorderi/Quantum-SDPs.

numerical solvers cannot understand the implicit arrangement of the underlying Hilbert spaces. Thus, we need to implement explicitly the isometries needed to obtain the correct order. To do that, we used the `PermuteSystems` *function available from the* QETLAB *library.*

As discussed in Lemma 3.7.2, the authors of [68] gave a *rank loop condition* to certify that a certain level of the hierarchy already gives the optimal value. We restate the condition here in the exact form needed for approximate quantum error correction.

Lemma 4.4.2. *Let* $\rho_{A\overline{A}(B\overline{B})_1^n} = \mathcal{U}^{\pi}_{(B\overline{B})_1^n}\left(\rho_{A\overline{A}(B\overline{B})_1^n}\right)$ *for all* $\pi \in \mathfrak{S}_n$ *and fixed* $0 \leq k \leq n$ *such that* $\rho^{\mathrm{T}_{(B\overline{B})_{k+1}^n}}_{A\overline{A}(B\overline{B})_1^n} \succeq 0$. *If we have*

$$\operatorname{rank}\left(\rho_{A\overline{A}(B\overline{B})_1^n}\right) \leq \max\left\{\operatorname{rank}\left(\rho_{A\overline{A}(B\overline{B})_1^k}\right), \operatorname{rank}\left(\rho_{(B\overline{B})_{k+1}^n}\right)\right\}, \tag{4.159}$$

then $\rho_{A\overline{A}B\overline{B}}$ *is separable with respect to the bipartite system* $A\overline{A} \otimes B\overline{B}$, *i.e.,* $\rho_{A\overline{A}B\overline{B}} \in \operatorname{Sep}(A\overline{A} : B\overline{B})$.

Using Lemma 4.4.2 it is in principle possible to, e.g., certify the optimality of the first level using the second level of our hierarchy. Moreover, if the criterion is fulfilled it can also allow us to extract the actual encoder and decoder of the optimal quantum error correction code. However, in order to facilitate the search for solutions having rank loops we need to look for low rank solutions $\rho_{A\overline{A}(B\overline{B})_1^n}$. It is not possible to directly write a rank condition into our semidefinite programs because rank constraints are not convex, as shown in the following remark.

Remark 4.4.3. *It is easy to show, with an explicit example, that rank constraints are not convex. In other words, it is not true that for every* $t \in [0,1]$ *and operators* X, Y *on* $\mathcal{H}$, *we have the inequality* $\operatorname{rank}(tX + (1-t)Y) \leq t\operatorname{rank}(X) + (1-t)\operatorname{rank}(Y)$. *For example, consider a two-dimensional Hilbert space* $\mathcal{H}$ *spanned by the orthonormal basis* $(|0\rangle, |1\rangle)$, *and*

choose $t := \frac{1}{2}$, $X := |0\rangle\langle 0|$ *and* $Y := |1\rangle\langle 1|$, *implying* $\mathrm{rank}(X) = \mathrm{rank}(Y) = 1$. *Thus,* $\mathrm{rank}(tX + (1-t)Y) = \mathrm{rank}\left(\frac{1}{2}\left[|0\rangle\langle 0| + |1\rangle\langle 1|\right]\right) = 2$. *On the other hand, we have* $t\mathrm{rank}(X) + (1-t)\mathrm{rank}(Y) = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 1 = 1$. *Since* $2 \not\leq 1$, *the rank is not a convex functional.*

In addition, SDP solvers typically give high rank solutions since they tend to look for solutions at the interior of the convex set. For our optimization programs, we noticed that SDPT3 compared to MOSEK gives results having in general lower rank. Nevertheless, a possible strategy is to find a solution $\rho_{A\overline{A}(B\overline{B})_1^n}$ and then employ a heuristic to minimize the rank while keeping the hierarchy constraints. The heuristic we found the most effective for our purposes was the *log-det method* described in [34]. The idea is to minimize the first-order Taylor series expansion of

$$\log\det\left(\rho_{A\overline{A}(B\overline{B})_1^n} + \delta\cdot 1_{A\overline{A}(B\overline{B})_1^n}\right), \tag{4.160}$$

which is used as a smooth surrogate[18] for $\mathrm{rank}\left(\rho_{A\overline{A}(B\overline{B})_1^n}\right)$ and $\delta > 0$ is a small regularization constant introduced to ensure the invertibility of the operators involved in the various iterations. The procedure is iterative, meaning that we start from $\rho_0 = 1_{A\overline{A}(B\overline{B})_1^n}$, then compute ρ_1 minimizing the log-det objective function, and so on. In particular, the choice $\rho_0 = 1_{A\overline{A}(B\overline{B})_1^n}$ connects the method to the *trace heuristic* [34], which is known to be an effective heuristic for rank reduction. More precisely, the log-det method can be seen as a sequence of weighted trace minimization problems. The (arbitrary) choice $\rho_0 = 1_{A\overline{A}(B\overline{B})_1^n}$ implies that ρ_1 is the outcome of the trace heuristic, which is an ordinary trace minimization. Thus, the further iterations of

[18] A *surrogate* is a function used to approximate another function. A surrogate function should be easy to evaluate. In this way, one can evaluate many times the surrogate to find the best approximation to the optimum value of the original objective function. The function $\log\det(\cdot)$ is a popular surrogate for the rank, because its global minimization leads to non-invertible operators, and hence to rank minimization (recall that a non-invertible operator cannot have full rank).

the log-det method can be seen as an improvement of the result given by the trace heuristic. We stop after a certain number i of iterations and then we find a solution ρ_i having, hopefully, lower rank than the original $\mathrm{rank}\big(\rho_{A\overline{A}(B\overline{B})_1^n}\big)$.

4.4.2 Qubit Channels

We computed SDP relaxations in the plain coding setting for the most common qubit channels: depolarizing, amplitude damping, bit flip, phase flip, bit-phase flip, Werner-Holevo and generalized Werner-Holevo channel. We found the upper bounds

$$\mathrm{SDP}_{1,\mathrm{PPT}}(\mathcal{N}_2, 2) = \mathrm{SDP}_{2,\mathrm{PPT}}(\mathcal{N}_2, 2) \tag{4.161}$$

$$= \mathrm{SDP}_{3,\mathrm{PPT}}(\mathcal{N}_2, 2) \tag{4.162}$$

$$= \mathrm{SDP}_1(\mathcal{N}_2, 2) \tag{4.163}$$

$$= \mathrm{SDP}_2(\mathcal{N}_2, 2) \tag{4.164}$$

$$= \mathrm{SDP}_3(\mathcal{N}_2, 2), \tag{4.165}$$

where the subscript in $\mathcal{N}_2$ refers to the two-dimensional input and output of the channel. These identities also remain true for random qubit channels[19] and one might then conjecture that for qubit channels indeed already $\mathrm{SDP}_1(\mathcal{N}_2, 2)$ captures $F(\mathcal{N}, 2)$.

For the qubit depolarizing channel the trivial coding scheme is known to be optimal[20] and

[19] To sample random channels we used the `RandomSuperoperator` function available from the QETLAB library. The rationale behind the usage of random channels in the numerics, is to move away from the highly symmetric settings provided by the most popular quantum channels (which may be the cause of the observed identities).

[20] In particular, for less than 5 repetitions of the depolarizing channel, the trivial coding scheme turns out to be the optimal error correction strategy [73]. With *trivial coding scheme* we mean to do no error correction at all. This result implies that, for the depolarizing channel, error correction becomes interesting in presence of at least 5 channel repetitions. Thus, we will study this setting in Subsection 4.4.4.

we retrieve this result using the rank loop condition of the second level based on the log-det method.

Similarly, for the qubit bit flip channel with parameter $p = 0.1$ we find a rank-one state solution of the second level using again the log-det method, implying that the rank loop condition holds. In this case the solution is not just the state associated with the trivial coding scheme via the Choi isomorphism but the resulting encoder/decoder pair with optimal fidelity 0.9 is given by the unitary channels with Kraus operators[21] $U_E = -|1\rangle\langle 0| + |0\rangle\langle 1|$ and $U_D = |0\rangle\langle 0| - |1\rangle\langle 1|$, respectively. Note that the trivial coding scheme is largely suboptimal for a qubit bit flip channel with $p = 0.1$, as the corresponding fidelity is 0.1.

4.4.3 Qutrit Channels

We computed SDP relaxations in the plain coding setting for the following qutrit channels: depolarizing, Werner-Holevo and generalized Werner-Holevo channel. We found the upper bounds

$$\mathrm{SDP}_{1,\mathrm{PPT}}(\mathcal{N}_3, 2) = \mathrm{SDP}_{2,\mathrm{PPT}}(\mathcal{N}_3, 2) \tag{4.166}$$

and this identity also remains true for random qutrit channels. Removing the PPT conditions, however, we found qutrit channels $\mathcal{N}_3$ such that

$$\mathrm{SDP}_2(\mathcal{N}_3, 2) < \mathrm{SDP}_1(\mathcal{N}_3, 2). \tag{4.167}$$

[21] A quantum channel $\mathcal{N} : \mathcal{S}(A) \to \mathcal{S}(B)$, can be represented by the finite sum $\mathcal{N}(\cdot) = \sum_k E_k(\cdot)E_k^\dagger$, for E_k linear maps between A and B satisfying the property $\sum_k E_k^\dagger E_k = 1$. Those linear maps are known as the *Kraus operators* of the channel.

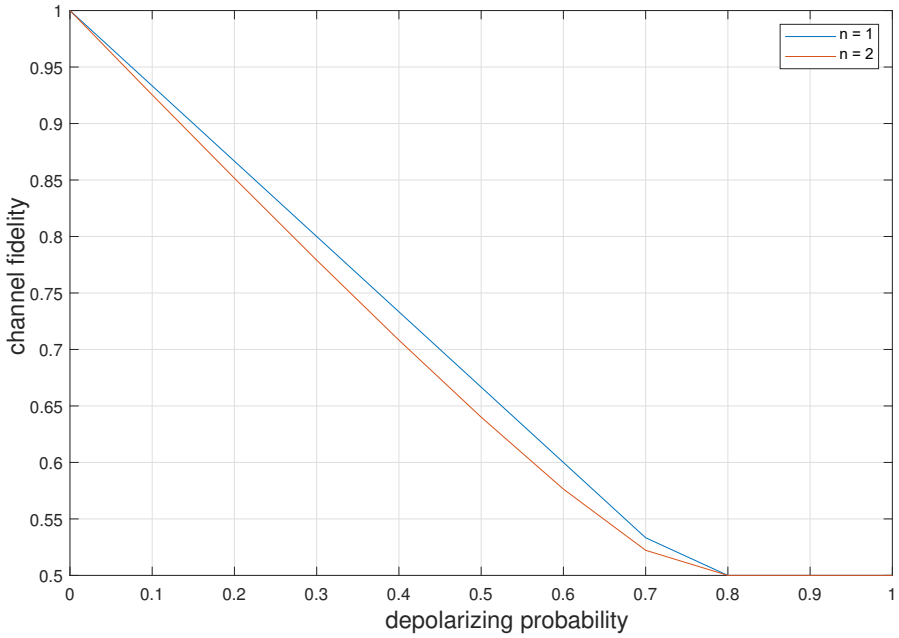

Figure 4.1: Comparison of the SDP upper bounds $n = 1, 2$ on the channel fidelity of the 3-dimensional depolarizing channel for LOCC(1)-assisted coding (see Section 4.3). We see an improvement for the second level for $p \in (0, 0.8)$.

4.4.4 Depolarizing channel

The depolarizing channel for $p \in [0, 4/3]$ is given as

$$Dep_d : \rho_{\overline{A}} \to p \cdot \mathrm{Tr}(\rho_{\overline{A}}) \frac{1_B}{d_B} + (1 - p) \cdot \rho_B, \tag{4.168}$$

where $d := d_{\overline{A}} = d_B$ denotes the dimension of the input and output. Notice that even though often the channel is only studied for $p \in [0, 1]$ where we can interpret p as a depolarizing probability, the above expression also represents a channel for $p \in (1, 4/3]$ (as, e.g., discussed in [73, Chapter 3]).

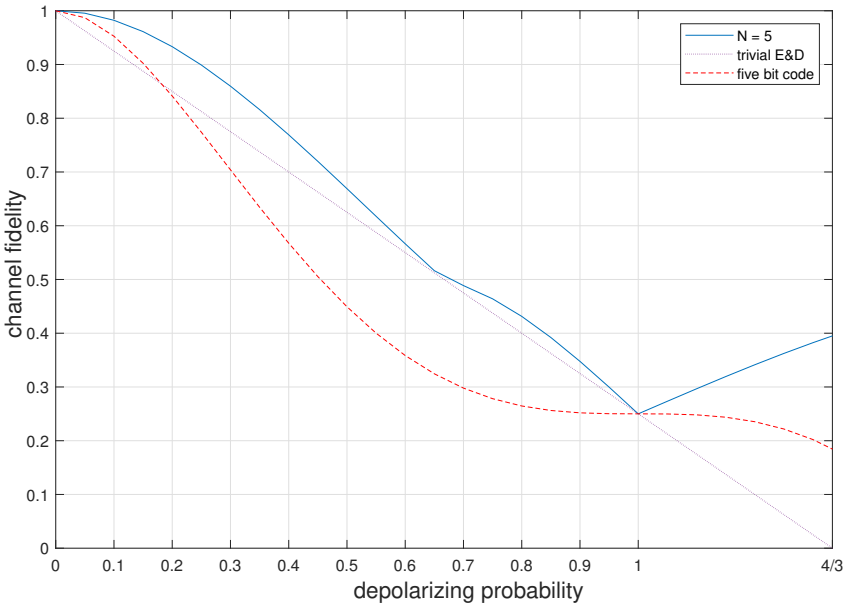

Figure 4.2: Comparison of the SDP upper bound $n = 1$ on the channel fidelity for five repetitions of the qubit depolarizing channel in the plain coding setting, with the trivial coding scheme and the 5 qubit code from [10]. Notice the intersection of the 5 qubit code and the trivial scheme in the region $p \in (0.1, 0.2)$ and the singular behaviour of the first level in the region $p \in (0.6, 0.7)$. In addition, for $p \in [1, 4/3]$ the behaviour of the first level seems to match exactly with the lower bound obtained with an iterative seesaw algorithm reported in Figure 3.7 of [73, Chapter 3].

We find that

$$\mathrm{SDP}_{1,\mathrm{PPT}}(Dep_2, 2) = \mathrm{SDP}_{2,\mathrm{PPT}}(Dep_2, 2) \tag{4.169}$$

$$= \mathrm{SDP}_{1,\mathrm{PPT}}(Dep_3, 2) \tag{4.170}$$

$$= \mathrm{SDP}_{2,\mathrm{PPT}}(Dep_3, 2). \tag{4.171}$$

However, in Section 4.4.3 we found that in general removing the PPT conditions allows us to see a difference for the first two levels. This behaviour is not shown by the qutrit depolarizing

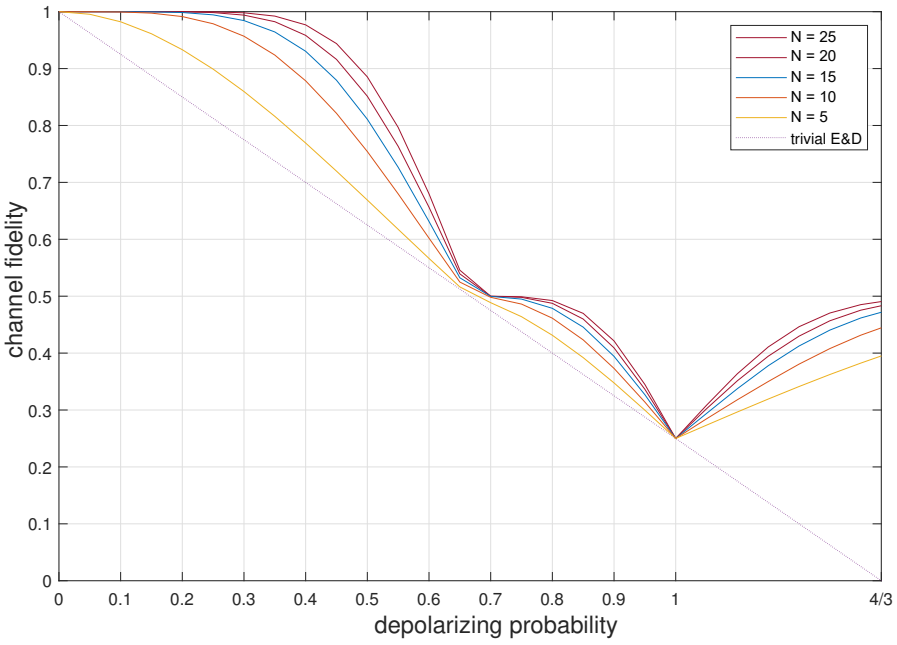

Figure 4.3: Comparison of the SDP upper bound $n = 1$ on the channel fidelity for 5, 10, 15, 20, 25 repetitions of the 2-dimensional depolarizing channel in the plain coding setting. Notice that the singular behaviour of the first level in the region $p \in (0.6, 0.7)$ is even more accentuated with the increase of the number of repetitions.

channel, probably due to its highly symmetrical structure. We computed the upper bound for LOCC(1) coding and found for $p \in (0, 0.8)$ that

$$\mathrm{SDP}_{2,\mathrm{PPT}}^{\mathrm{LOCC}(1)}(Dep_2, 2) = \mathrm{SDP}_{1,\mathrm{PPT}}^{\mathrm{LOCC}(1)}(Dep_2, 2) \tag{4.172}$$

while,

$$\mathrm{SDP}_{2,\mathrm{PPT}}^{\mathrm{LOCC}(1)}(Dep_3, 2) < \mathrm{SDP}_{1,\mathrm{PPT}}^{\mathrm{LOCC}(1)}(Dep_3, 2). \tag{4.173}$$

We compared, for the plain coding setting, the $n = 1$ level for five repetitions of the qubit depolarizing channel with the fidelity of the trivial coding scheme, as well as the 5 qubit code

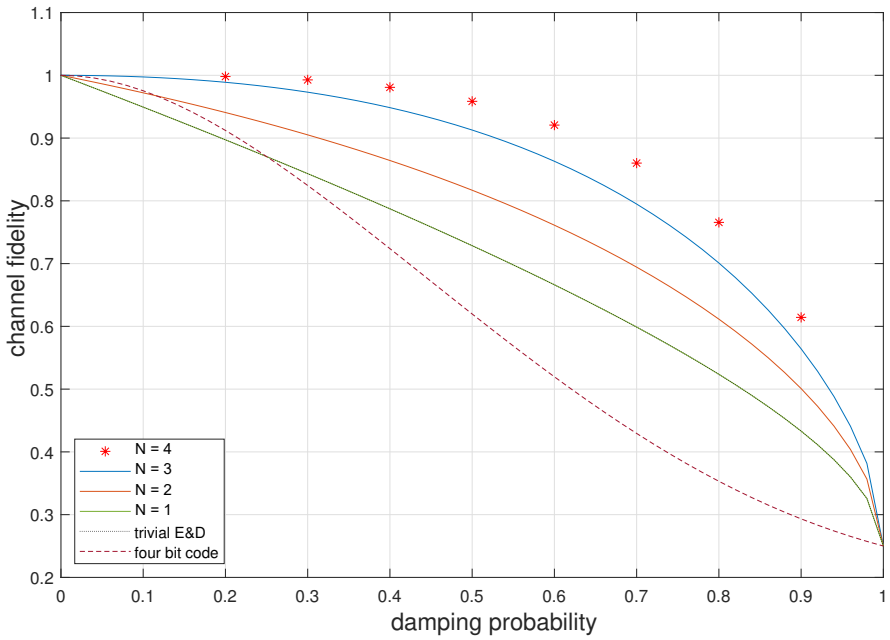

Figure 4.4: Comparison of the SDP upper bound $n = 1$ on the channel fidelity of the qubit amplitude damping channel for 1, 2, 3 and 4 repetitions in the plain coding setting, as well as the trivial encoder and decoder and the 4 qubit code from [66].

from [10]. In particular, following [84] we exploited the symmetries of the qubit depolarizing channel to get the linear program

$$\mathrm{SDP}_{1,\mathrm{PPT}}\left(Dep_2^{\otimes N}, 2\right) = \max \quad \sum_{i=0}^{N} \binom{N}{i} \left(1 - \frac{3p}{4}\right)^i \left(\frac{3p}{4}\right)^{N-i} m_i \tag{4.174}$$

$$s.t. \quad 0 \leq m_i \leq 1 \quad i \in \{0, \ldots, N\} \tag{4.175}$$

$$-\frac{1}{2} \leq \sum_{i=0}^{N} x_{i,k} m_i \leq \frac{1}{2} \quad k \in \{0, \ldots, N\} \tag{4.176}$$

$$\sum_{i=0}^{N} \binom{N}{i} 3^{N-i} m_i = 2^{2N-2}. \tag{4.177}$$

where $x_{i,k} := \frac{1}{d^N} \sum_{r=\max\{0,i+k-N\}}^{\min\{i,k\}} \binom{k}{r} \binom{N-k}{i-r} (-1)^{i-r} (d-1)^{k-r} (d+1)^{N-k+r-i}$ for every $i, k \in$

$\{0, \ldots, N\}$. Notice that the number of variables grows linearly with N. The results are reported in Figure 4.2. Comparing these with Figure 3.7 in [73, Chapter 3], it seems that the first level of the hierarchy matches their lower bounds in the region $p \in [1, 4/3]$. Notice the intersection of the five qubit code and the trivial coding scheme in the region $p \in (0.1, 0.2)$ and the singular behaviour in the region $p \in (0.6, 0.7)$. We have also examined five, ten, fifteen, twenty and twenty five repetitions of the qubit depolarizing channel, again using the above linear program. The results are shown in Figure 4.3. Notice that the singular behaviour noted in Figure 4.2 is now even more accentuated when increasing the number of repetitions.

4.4.5 Amplitude damping channel

The qubit amplitude damping channel with damping probability $\gamma \in [0, 1]$ is given as

$$Amp_\gamma : \rho_{\overline{A}} \to E_B^0 \rho_B {E_B^0}^\dagger + E_B^1 \rho_B {E_B^1}^\dagger \tag{4.178}$$

where

$$E_B^0 := |0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|, \tag{4.179}$$

$$E_B^1 := \sqrt{\gamma}|0\rangle\langle 1|. \tag{4.180}$$

We compared the results given by one, two, three, and four repetitions of the channel for the level $n = 1$. The bounds are shown in Figure 4.4, compared with the fidelity of the trivial coding scheme, and the 4 qubit code from [66]. Notice the overlap between the first level of the hierarchy and the trivial coding scheme for the one-shot setting, i.e., with a single repetition of the channel. Comparing these results with Figure 3.12 in [73, Chapter 3] we see that there is gap between their lower bounds (that significantly improves on the trivial coding scheme) and our upper bounds.

4.5 Worst case error criterion

4.5.1 Setting

So far we have used the channel fidelity from Definition 4.1.1 as the measure to study approximate quantum error correction, which corresponds to the *average error case*. In this section, we consider the diamond norm (2.60) to study the *worst case error*[22] and we find a program for which the hierarchy can be used to generate, in this case, lower bounds[23]. We prove the sequence of semidefinite relaxations do in fact converge to the exact value of the original optimization program.

Definition 4.5.1. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$, with $M = d_A = d_{\overline{B}}$. The channel distance is defined as*

$$\Delta(\mathcal{N}, M) := \min \frac{1}{2} \left\| \mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}} - \mathcal{I}_{A\to\overline{B}} \right\|_\diamond \tag{4.181}$$

$$s.t.\ \mathcal{D}_{B\to\overline{B}}, \mathcal{E}_{A\to\overline{A}}\ \textit{quantum channels.} \tag{4.182}$$

The following lemma writes the channel distance as given in Definition 4.5.1 in terms of the Choi states of the encoder $\mathcal{E}_{A\to\overline{A}}$ and decoder $\mathcal{D}_{B\to\overline{B}}$, respectively.

Lemma 4.5.2. *Let $\mathcal{N}_{\overline{A}\to B}$ be a quantum channel and $M \in \mathbb{N}$. Then, we have that*

$$\Delta(\mathcal{N}, M) = \min \lambda \tag{4.183}$$

$$\text{s.t. } E_{A\overline{A}} \succeq 0,\ E_A = \frac{1_A}{d_A} \tag{4.184}$$

[22] The diamond norm is used to measure the worst case error because of its direct connection with the worst case probability of failing to distinguish the outputs of two quantum channels, given any common input state [11]. See also [87, Theorem 3.52].

[23] In this setting our hierarchy generates lower bounds because we are looking at outer approximations for a minimization problem.

$$D_{B\overline{B}} \succeq 0,\ D_B = \frac{1_B}{d_B} \tag{4.185}$$

$$Z_{A\overline{B}} \succeq 0,\ \frac{\lambda}{d_A} \cdot 1_A \succeq Z_A \tag{4.186}$$

$$Z_{A\overline{B}} + \Phi_{A\overline{B}} \succeq d_{\overline{A}} d_B \cdot \mathrm{Tr}_{\overline{A}B}\left[(1_A \otimes J^{\mathcal{N}}_{\overline{A}B} \otimes 1_{\overline{B}})(E_{A\overline{A}} \otimes D_{B\overline{B}})\right], \tag{4.187}$$

where $J^{\mathcal{N}}_{\overline{A}B}$ denotes the Choi state of $\mathcal{N}_{\overline{A}\to B}$ (see 2.41).

Proof. Following [86], the channel distance $\Delta(\mathcal{N}, M)$ can be written as

$$\Delta(\mathcal{N}, M) = \min\ \|Z_A\|_\infty \tag{4.188}$$

$$\text{s.t. } \mathcal{D}_{B\to\overline{B}}, \mathcal{E}_{A\to\overline{A}} \text{ quantum channels} \tag{4.189}$$

$$Z_{A\overline{B}} \succeq 0,\ Z_{A\overline{B}} \succeq d_A \cdot J^{\mathcal{D}\circ\mathcal{N}\circ\mathcal{E}-\mathcal{I}}_{A\overline{B}}. \tag{4.190}$$

We simplify

$$J^{\mathcal{D}\circ\mathcal{N}\circ\mathcal{E}-\mathcal{I}}_{A\overline{B}} = J^{\mathcal{D}\circ\mathcal{N}\circ\mathcal{E}}_{A\overline{B}} - J^{\mathcal{I}}_{A\overline{B}} \tag{4.191}$$

$$= J^{\mathcal{D}\circ\mathcal{N}\circ\mathcal{E}}_{A\overline{B}} - \Phi_{A\overline{B}}, \tag{4.192}$$

write for the infinity norm $\|Z_A\|_\infty = \min\{\lambda \in \mathbb{R} : \lambda \cdot 1_A \succeq Z_A\}$ [81], and relabel $\frac{Z_{A\overline{B}}}{d_A}$ as $Z_{A\overline{B}}$, leading to

$$\Delta(\mathcal{N}, M) = \min \lambda \tag{4.193}$$

$$\text{s.t. } \mathcal{D}_{B\to\overline{B}}, \mathcal{E}_{A\to\overline{A}} \text{ quantum channels} \tag{4.194}$$

$$Z_{A\overline{B}} \succeq 0, \frac{\lambda}{d_A} \cdot 1_A \succeq Z_A \tag{4.195}$$

$$Z_{A\overline{B}} + \Phi_{A\overline{B}} \succeq J^{\mathcal{D}\circ\mathcal{N}\circ\mathcal{E}}_{A\overline{B}}. \tag{4.196}$$

Following [65] and in particular [85, Equation 7], we have the Choi state

$$J^{\mathcal{D}\circ\mathcal{N}\circ\mathcal{E}}_{A\overline{B}} = d_{\overline{A}} d_B \cdot \mathrm{Tr}_{\overline{A}B}\left[(1_A \otimes J^{\mathcal{N}}_{\overline{A}B} \otimes 1_{\overline{B}})\left(J^{\mathcal{E}}_{A\overline{A}} \otimes J^{\mathcal{D}}_{B\overline{B}}\right)\right] \tag{4.197}$$

and writing $J^{\mathcal{E}}_{A\overline{A}} = E_{A\overline{A}}$ as well as $J^{\mathcal{D}}_{B\overline{B}} = D_{B\overline{B}}$ concludes the proof. □

4.5.2 Hierarchy of lower bounds

Similarly as in Section 4.2.1, we define a hierarchy of semidefinite programs labelled by an index n. Our framework directly applies as the structure of the optimization problem derived in Lemma 4.5.2 involves the tensor product $E_{A\overline{A}} \otimes D_{B\overline{B}}$. The n-th level of the SDP hierarchy then generates the lower bounds $\mathrm{SDP}_n^{\Delta}(\mathcal{N}, M)$ for the distance $\Delta(\mathcal{N}, M)$ as

$$\mathrm{SDP}_n^{\Delta}(\mathcal{N}, M) := \min \lambda \tag{4.198}$$

$$\text{s.t. } \rho_{A\overline{A}(B\overline{B})_1^n} \succeq 0,\ \mathrm{Tr}\left[\rho_{A\overline{A}(B\overline{B})_1^n}\right] = 1 \tag{4.199}$$

$$\rho_{A\overline{A}(B\overline{B})_1^n} = \mathcal{U}^{\pi}_{(B\overline{B})_1^n}\left(\rho_{A\overline{A}(B\overline{B})_1^n}\right)\ \forall \pi \in \mathfrak{S}_n \tag{4.200}$$

$$\rho_{A(B\overline{B})_1^n} = \frac{1_A}{d_A} \otimes \rho_{(B\overline{B})_1^n} \tag{4.201}$$

$$\rho_{A\overline{A}(B\overline{B})_1^{n-1}B_n} = \rho_{A\overline{A}(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B} \tag{4.202}$$

$$Z_{A\overline{B}} \succeq 0,\ \frac{\lambda}{d_A} \cdot 1_A \succeq Z_A \tag{4.203}$$

$$Z_{A\overline{B}} + \Phi_{A\overline{B}} \succeq d_{\overline{A}} d_B \cdot \mathrm{Tr}_{\overline{A}B}\left[\left(1_A \otimes J^{\mathcal{N}}_{\overline{A}B} \otimes 1_{\overline{B}}\right) \rho_{A\overline{A}B\overline{B}}\right]. \tag{4.204}$$

We can also add PPT constraints and denote the resulting relaxations by $\mathrm{SDP}^{\Delta}_{n,\mathrm{PPT}}(\mathcal{N}, M)$. The following theorem states the convergence of the hierarchy.

Theorem 4.5.3. *Let $\mathcal{N}$ be a quantum channel and $n, M \in \mathbb{N}$. Then, we have*

$$0 \leq \Delta(\mathcal{N}, M) - \mathrm{SDP}_n^{\Delta}(\mathcal{N}, M) \leq \frac{\mathrm{poly}(d)}{\sqrt{n}} \tag{4.205}$$

implying

$$\Delta(\mathcal{N}, M) = \lim_{n\to\infty} \mathrm{SDP}_n^{\Delta}(\mathcal{N}, M), \tag{4.206}$$

where $d := \max\{d_A, d_{\overline{A}}, d_B, d_{\overline{B}}\}$.

Proof. The bound $0 \leq \Delta(\mathcal{N}, M) - \mathrm{SDP}_n^{\Delta}(\mathcal{N}, M)$ holds by construction and thus we consider the upper bound. First, note that again applying (4.193) we can write

$$\mathrm{SDP}_n^{\Delta}(\mathcal{N}, M) = \min \frac{1}{2} \left\| \mathcal{W}(\mathcal{N})_{A\to\overline{B}} - \mathcal{I}_{A\overline{B}} \right\|_{\diamond} \tag{4.207}$$

$$\text{s.t. } \rho_{A\overline{A}(B\overline{B})_1^n} \succeq 0,\ \mathrm{Tr}\left[\rho_{A\overline{A}(B\overline{B})_1^n}\right] = 1 \tag{4.208}$$

$$\rho_{A\overline{A}(B\overline{B})_1^n} = \mathcal{U}^{\pi}_{(B\overline{B})_1^n}\left(\rho_{A\overline{A}(B\overline{B})_1^n}\right)\ \forall \pi \in \mathfrak{S}_n \tag{4.209}$$

$$\rho_{A(B\overline{B})_1^n} = \frac{1_A}{d_A} \otimes \rho_{(B\overline{B})_1^n} \tag{4.210}$$

$$\rho_{A\overline{A}(B\overline{B})_1^{n-1}B_n} = \rho_{A\overline{A}(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}, \tag{4.211}$$

with the quantum channel $\mathcal{W}(\mathcal{N})_{A\to\overline{B}}$ defined via its Choi state

$$J^{\mathcal{W}(\mathcal{N})}_{A\overline{B}} := d_{\overline{A}} d_B \cdot \mathrm{Tr}_{\overline{A}B}\left[\left(1_A \otimes J^{\mathcal{N}}_{\overline{A}B} \otimes 1_{\overline{B}}\right) \rho_{A\overline{A}B\overline{B}}\right]. \tag{4.212}$$

Second, using the de Finetti Theorem 3.6.5 we get that for every feasible Choi state $\rho_{A\overline{A}(B\overline{B})_1^n}$ in $\mathrm{SDP}_n^{\Delta}(\mathcal{N}, M)$, there exists a feasible Choi state $E_{A\overline{A}} \otimes D_{B\overline{B}}$ in $\Delta(\mathcal{N}, M)$ from Lemma 4.5.2, such that

$$\left\| E_{A\overline{A}} \otimes D_{B\overline{B}} - \rho_{A\overline{A}B\overline{B}} \right\|_1 \leq \frac{\mathrm{poly}(d)}{\sqrt{n}}. \tag{4.213}$$

Third, employing the triangle inequality for the diamond norm we have

$$\left\| \mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}} - \mathcal{I}_{A\to\overline{B}} \right\|_{\diamond} - \left\| \mathcal{W}(\mathcal{N})_{A\to\overline{B}} - \mathcal{I}_{A\to\overline{B}} \right\|_{\diamond} \tag{4.214}$$

$$\leq \left| \left\| \mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}} - \mathcal{I}_{A\to\overline{B}} \right\|_{\diamond} - \left\| \mathcal{W}(\mathcal{N})_{A\to\overline{B}} - \mathcal{I}_{A\to\overline{B}} \right\|_{\diamond} \right| \tag{4.215}$$

$$\leq \left\| \mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}} - \mathcal{W}(\mathcal{N})_{A\to\overline{B}} \right\|_{\diamond}. \tag{4.216}$$

Forth, relating the trace norm distance of Choi states to the diamond norm distance of quantum channels (Lemma 2.2.4), we have

$$\left\| \mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}} - \mathcal{W}(\mathcal{N})_{A\to\overline{B}} \right\|_{\diamond} \leq d_A \cdot \left\| J^{\mathcal{D}\circ\mathcal{N}\circ\mathcal{E}}_{A\overline{B}} - J^{\mathcal{W}(\mathcal{N})}_{A\overline{B}} \right\|_1 \tag{4.217}$$

and thanks to the monotonicity under partial trace [87, Theorem 3.39], the sub-multiplicativity of the partial trace (2.25), and the $(\infty, 1)$ Hölder's inequality, this bounds (4.216) as

$$\left\|\mathcal{D}_{B\to\overline{B}} \circ \mathcal{N}_{\overline{A}\to B} \circ \mathcal{E}_{A\to\overline{A}} - \mathcal{W}(\mathcal{N})_{A\to\overline{B}}\right\|_{\diamond} \tag{4.218}$$

$$\leq d_A d_{\overline{A}} d_B \cdot \left\|\mathrm{Tr}_{\overline{A}B}\left[\left(1_A \otimes J^{\mathcal{N}}_{\overline{A}B} \otimes 1_{\overline{B}}\right)\left(E_{A\overline{A}} \otimes D_{B\overline{B}} - \rho_{A\overline{A}B\overline{B}}\right)\right]\right\|_1 \tag{4.219}$$

$$\leq d_A d_{\overline{A}} d_B \cdot \left\|\left(1_A \otimes J^{\mathcal{N}}_{\overline{A}B} \otimes 1_{\overline{B}}\right)\left(E_{A\overline{A}} \otimes D_{B\overline{B}} - \rho_{A\overline{A}B\overline{B}}\right)\right\|_1 \tag{4.220}$$

$$\leq d_A d_{\overline{A}} d_B \cdot \left\|1_A \otimes J^{\mathcal{N}}_{\overline{A}B} \otimes 1_{\overline{B}}\right\|_{\infty} \left\|E_{A\overline{A}} \otimes D_{B\overline{B}} - \rho_{A\overline{A}B\overline{B}}\right\|_1 \tag{4.221}$$

$$\leq \frac{\mathrm{poly}(d)}{\sqrt{n}}, \tag{4.222}$$

with $d := \max\{d_A, d_{\overline{A}}, d_B, d_{\overline{B}}\}$.

Finally, optimising in (4.216) over all feasible Choi states $\rho_{A\overline{A}(B\overline{B})^n_1}$ and then optimising over all feasible Choi states $E_{A\overline{A}} \otimes D_{B\overline{B}}$, we get the claimed upper bound

$$\Delta(\mathcal{N}, M) - \mathrm{SDP}^{\Delta}_n(\mathcal{N}, M) \leq \frac{\mathrm{poly}(d)}{\sqrt{n}}. \tag{4.223}$$

□

Numerically, we have found that for the qubit depolarizing channel the first level of our hierarchy already gives the exact optimal value

$$\Delta(Dep_2, 2) = \mathrm{SDP}^{\Delta}_{1,\mathrm{PPT}}(Dep_2, 2), \tag{4.224}$$

which coincides with $1 - F(Dep_2, 2)$. That is, for the qubit depolarizing channel the average and worst case error criteria become the same.

Chapter 5

De Finetti Reductions with Linear Constraints

The previous chapters of this thesis investigated de Finetti theorems. Those results allow to represent, or approximate, mathematical objects symmetric under permutations of their components into a probabilistic ensemble of elementary independent and identically distributed (i.i.d.) constituents. In particular, we have shown how to develop a family of such representation theorems in presence of additional linear constraints. In several applications, instead of representation results as given by de Finetti theorems, one may need to establish a generalized order relation between the symmetric mathematical object and the probabilistic ensemble of elementary i.i.d. constituents. De Finetti reductions, previously known as "post-selection techniques" [22] or methods based on "universal states" [46], provide the desired inequality. For example, a quantum de Finetti reduction provides an upper bound to a symmetric quantum state in the form of an integral superposition of product states, weighted by a factor which is

polynomial in terms of the number of copies and exponential in terms of the local dimensionality

$$\rho_{\mathcal{H}^n} \preceq (n+1)^{d_{\mathcal{H}}^2-1} \int \sigma_{\mathcal{H}}^{\otimes n} \, \mathrm{d}\sigma_{\mathcal{H}}, \tag{5.1}$$

where $\rho_{\mathcal{H}^n}$ is a permutation invariant quantum state, and $\mathrm{d}\sigma_{\mathcal{H}}$ is an appropriate measure over the set of quantum states on $\mathcal{H}$. The generality of expression (5.1) is also its main drawback. On one hand, unlike finite de Finetti representation theorems, (5.1) provides an exact bound, without any parameter controlling the approximation error. On the other hand, all permutation invariant quantum states are upper bounded by the same mixture of tensor product states. Any other information encoded in the permutation invariant state $\rho_{\mathcal{H}^n}$ is lost. A way to obtain a state-dependent upper-bound is via a so-called "flexible" de Finetti reduction [64]. In a flexible de Finetti reduction, each tensor product state appearing in the integral superposition is weighted by its fidelity with the symmetric state, although an affine adjustment to the local dimensionality is required

$$\rho_{\mathcal{H}^n} \preceq (n+1)^{3d_{\mathcal{H}}^2-1} \int F(\rho_{\mathcal{H}^n}, \sigma_{\mathcal{H}}^{\otimes n}) \sigma_{\mathcal{H}}^{\otimes n} \, \mathrm{d}\sigma_{\mathcal{H}}, \tag{5.2}$$

where $\rho_{\mathcal{H}^n}$ is a permutation invariant quantum state, $F(\rho, \sigma) := \left\| \sqrt{\rho}\sqrt{\sigma} \right\|_1^2$ denotes the fidelity[1], and $\mathrm{d}\sigma_{\mathcal{H}}$ is an appropriate measure over the set of quantum states on $\mathcal{H}$. As we see from (5.2), only the tensor product states $\sigma_{\mathcal{H}}^{\otimes n}$ that are close (in fidelity) to $\rho_{\mathcal{H}^n}$ should bring a relevant contribution to the integral superposition. This effectively allows us to obtain a state-dependent expression upper-bounding the permutation invariant state $\rho_{\mathcal{H}^n}$. Flexible de Finetti reductions can be applied to the study of the optimal winning strategy for certain types of multi-player games [4]. Those flexible versions of de Finetti reductions can be seen as a complementary approach to our work, but our techniques are different, and we do not use the fidelity.

[1] In order to avoid confusion, it is important to keep in mind that some authors use the name fidelity for the square root of this quantity. In other words, they define the fidelity as $\sqrt{F(\rho,\sigma)} = \left\| \sqrt{\rho}\sqrt{\sigma} \right\|_1$. This is the case of [64].

On the other hand, one can study both de Finetti representation theorems and de Finetti reductions in the presence of additional constraints on the symmetric state (see Chapter 3, [33], [18], and [64]). For cryptographic applications and error correction (see Chapter 4), it is often useful to study the case where a new system, carrying external side information, adds a non-symmetric contribution to the symmetric object. So far, no clear or systematic connection between de Finetti reductions and de Finetti representation theorems has been proven in the literature. Thus, in this chapter we have three main interests

1. How to extend de Finetti reductions to include the case with side information,

2. how to incorporate various types of constraints in the de Finetti reduction,

3. how to derive de Finetti representation theorems from de Finetti reductions.

The content of this chapter is largely based on our research notes [12] .

5.1 Classical Relative Entropy and Chain Rules

The starting point for our proof techniques is the use of various forms of chain rules for the classical relative entropy. The *classical relative entropy*, also known as the *Kullback-Leibler divergence*, is the classical version of (3.52), and is defined as[2]

$$D_{KL}(p_X\|q_X) := \begin{cases} \sum_{\substack{x\in \text{image}(X) \\ p_X(x)>0}} p_X(x)\log\left(\frac{p_X(x)}{q_X(x)}\right) & \text{if } \text{supp}(p_X)\subseteq \text{supp}(q_X) \\ \infty & \text{otherwise} \end{cases}, \tag{5.3}$$

[2]The support of a probability mass function is formed by the set of points where the function is greater than zero. Thus, $\text{supp}(p_X)\subseteq\text{supp}(q_X)$ implies that for every $x\in\text{image}(X)$ such that $p_X(x)>0$, then $q_X(x)>0$ as well.

where X is a discrete random variable, p_X and s_X are two probability mass functions, and the logarithm is taken with respect to the basis two, i.e., $\log(\cdot) := \log_2(\cdot)$. For simplicity, we will typically write $D_{KL}(p_X\|q_X) = \sum_x p_X(x) \log\left(\frac{p_X(x)}{q_X(x)}\right)$ in place of the complete notation of (5.3).

In Section 3.6.2 we have presented Quantum Pinsker's inequality (Theorem 3.6.1), which relates the quantum relative entropy to the trace distance. The following is the classical version of Pinsker's inequality, which can be seen as a special case of Theorem 3.6.1.

Theorem 5.1.1. (Classical Pinsker's inequality) *Let X a discrete random variable and p_X, q_X probability mass functions, then*

$$D_{KL}(p_X\|q_X) \geq \frac{1}{2\ln 2} \|p_X - q_X\|_1^2, \tag{5.4}$$

where $\|p_X - q_X\| := \sum_{x\in\text{image}(X)} \|p_X(x) - q_X(x)\|$.

The following lemma provides a well-known chain rule for the Kullback–Leibler divergence [25, Theorem 2.5.3].

Lemma 5.1.2. *Let X, Y discrete random variables, p_{XY}, q_{XY} probability mass functions, then*

$$D_{KL}(p_{XY}\|q_{XY}) = D_{KL}(p_X\|q_X) + \mathop{\mathbf{E}}_X \left\{ D_{KL}(p_{Y|X}\|q_{Y|X}) \right\}, \tag{5.5}$$

where the expectation is computed with respect to p_X.

Proof. The proof is obtained by a direct computation

$$\begin{aligned} D_{KL}(p_{XY}\|q_{XY}) &= \sum_{x,y} p_{XY}(x,y) \log\left(\frac{p_{XY}(x,y)}{q_{XY}(x,y)}\right) && (5.6)\\ &= \sum_{x,y} p_X(x)p(y|x) \log\left(\frac{p_X(x)p(y|x)}{q_X(x)q(y|x)}\right) && (5.7)\\ &= \sum_{x,y} p_X(x)p(y|x) \log\left(\frac{p_X(x)}{q_X(x)}\right) + \sum_{x,y} p_X(x)p(y|x) \log\left(\frac{p(y|x)}{q(y|x)}\right) && (5.8) \end{aligned}$$

$$= \sum_x p_X(x) \left[\sum_y p(y|x) \right] \log \left(\frac{p_X(x)}{q_X(x)} \right) \tag{5.9}$$

$$+ \sum_x p_X(x) \left[\sum_y p(y|x) \log \left(\frac{p(y|x)}{q(y|x)} \right) \right] \tag{5.10}$$

$$= \sum_x p_X(x) \log \left(\frac{p_X(x)}{q_X(x)} \right) + \sum_x p_X(x) D_{KL}(p_{Y|X=x} \| q_{Y|X=x}) \tag{5.11}$$

$$= D_{KL}(p_X \| q_X) + \mathop{\mathbf{E}}_X \left\{ D_{KL}(p_{Y|X} \| q_{Y|X}) \right\}. \tag{5.12}$$

□

The above lemma can be easily generalized for conditional distributions. For example, if we add an additional discrete random variable C with respect to which we do the conditioning, we find the following chain rule

$$D_{KL}(p_{XY|C} \| q_{XY|C}) = D_{KL}(p_{X|C} \| q_{X|C}) + \mathop{\mathbf{E}}_{X|C} \left\{ D_{KL}(p_{Y|X,C} \| q_{Y|X,C}) \right\}, \tag{5.13}$$

where now the expectation is computed with respect to $p_{X|C} := \frac{p_{XC}}{p_C}$.

It is possible to transform convex combinations of Kullback–Leibler divergences into a quantum relative entropy. This result is the content of the following lemma.

Lemma 5.1.3. *Let Q, Y discrete random variables, r_Q and p_Y^q, s_Y^q probability mass functions for every $q \in \text{image}(Q)$, then*

$$\sum_q r_Q(q) D_{KL}(p_Y^q \| s_Y^q) \tag{5.14}$$

$$= D\left(\sum_q r_Q(q) |q\rangle\langle q| \otimes \sum_y p_Y^q(y) |y\rangle\langle y| \middle\| \sum_q r_Q(q) |q\rangle\langle q| \otimes \sum_y s_Y^q(y) |y\rangle\langle y| \right), \tag{5.15}$$

where the quantum states form an orthonormal basis for the associated Hilbert spaces.

Proof. The proof is obtained by direct computation of the quantum relative entropy

$$D\left(\sum_q r_Q(q) |q\rangle\langle q| \otimes \sum_y p_Y^q(y) |y\rangle\langle y| \middle\| \sum_q r_Q(q) |q\rangle\langle q| \otimes \sum_y s_Y^q(y) |y\rangle\langle y| \right) \tag{5.16}$$

$$= \operatorname{Tr}\left[\sum_q r_Q(q)|q\rangle\langle q| \otimes \sum_y p_Y^q(y)|y\rangle\langle y| \log\left(\sum_{q'} r_Q(q')|q'\rangle\langle q'| \otimes \sum_{y'} p_Y^{q'}(y')|y'\rangle\langle y'|\right)\right] \tag{5.17}$$

$$- \operatorname{Tr}\left[\sum_q r_Q(q)|q\rangle\langle q| \otimes \sum_y p_Y^q(y)|y\rangle\langle y| \log\left(\sum_{q'} r_Q(q')|q'\rangle\langle q'| \otimes \sum_{y'} s_Y^{q'}(y')|y'\rangle\langle y'|\right)\right] \tag{5.18}$$

$$= \operatorname{Tr}\left[\sum_q r_Q(q)|q\rangle\langle q| \otimes \sum_y p_Y^q(y)|y\rangle\langle y| \sum_{q'}\sum_{y'} \log\left(\frac{r_Q(q')p_Y^{q'}(y')}{r_Q(q')s_Y^{q'}(y')}\right)|q'\rangle\langle q'| \otimes |y'\rangle\langle y'|\right] \tag{5.19}$$

$$= \operatorname{Tr}\left[\sum_q \sum_y r_Q(q)p_Y^q(y) \log\left(\frac{p_Y^q(y)}{s_Y^q(y)}\right)|q\rangle\langle q| \otimes |y\rangle\langle y|\right] \tag{5.20}$$

$$= \sum_q r_Q(q) \sum_y p_Y^q(y) \log\left(\frac{p_Y^q(y)}{s_Y^q(y)}\right) \tag{5.21}$$

$$= \sum_q r_Q(q) D_{KL}(p_Y^q \| s_Y^q). \tag{5.22}$$

□

The following lemma is the classical version of the above result.

Lemma 5.1.4. *Let Q, Y discrete random variables, r_Q and p_Y^q, s_Y^q probability mass functions for every $q \in \operatorname{image}(Q)$, then*

$$\sum_q r_Q(q) D_{KL}(p_Y^q \| s_Y^q) = D_{KL}(\tilde{p}_{QY} \| \tilde{s}_{QY}), \tag{5.23}$$

where $\tilde{p}_{QY}$ and $\tilde{s}_{QY}$ are probability mass functions defined by $\tilde{p}_{QY} := r_Q(q)p_Y^q(y)$ and $\tilde{s}_{QY} := r_Q(q)s_Y^q(y)$ for every $q \in \operatorname{image}(Q)$ and $y \in \operatorname{image}(Y)$.

Proof. The proof is obtained by a simple manipulation

$$\sum_q r_Q(q) D_{KL}(p_Y^q \| s_Y^q) = \sum_q r_Q(q) \sum_y p_Y^q(y) \log\left(\frac{p_Y^q(y)}{s_Y^q(y)}\right) \tag{5.24}$$

$$= \sum_{q,y} r_Q(q)p_Y^q(y) \log\left(\frac{r_Q(q)p_Y^q(y)}{r_Q(q)s_Y^q(y)}\right) \tag{5.25}$$

$$= \sum_{q,y} \tilde{p}_{QY} \log \left(\frac{\tilde{p}_{QY}}{\tilde{s}_{QY}} \right) \tag{5.26}$$

$$= D_{KL}(\tilde{p}_{QY} \| \tilde{s}_{QY}). \tag{5.27}$$

□

5.2 Constrained de Finetti Reductions with Side Information

De Finetti reductions are useful inequalities that are commonly used to simplify the computation of certain bounds for the action of functionals on permutation invariant states. In this thesis we are mainly interested in quantum de Finetti reductions, which are stated for permutation invariant quantum states. However, de Finetti reductions do exist also in the classical setting. For example, in [3] the authors prove various classical de Finetti reductions for permutation invariant conditional probability distributions. Their approach is based, mainly, on combinatorial arguments and can be generalized to handle additional types of symmetries (e.g., the CHSH-type symmetry defined in [3, Definition 5]). In [48], the authors prove a classical de Finetti reduction for permutation invariant probability distributions by using the method of types in the context of composite hypothesis testing and its connection to Rényi information measures. The method of types is also used in [7]. Moreover, the authors introduce new proof techniques and derive classical flexible versions of de Finetti reductions.

As pointed out, there exist in the literature several quantum de Finetti reductions that are able to handle specific linear constraints on the permutation invariant state (e.g., [33] and [64]). Those theorems restrict the support of the measure in order to capture the specific linear constrain on the initial state or introduce a fidelity weight in the integral superposition. In this section, we introduce a new ingredient: the quantum side information. We prove a new de

Finetti reduction in presence of quantum side information. Moreover, we show that our result can handle, at the same time, two different types of constraints, a marginal constraint on the symmetric part, and a general linear constraint on the quantum side information. Our result can be seen as an extension of the constrained de Finetti reduction presented in [33, Corollary 3.2].

Proposition 5.2.1. *Let Q, A and B be Hilbert spaces and $\rho_{QA^nB^n}$ a state symmetric with respect to Q. Moreover, let $\rho_{A^n} = \sigma_A^{\otimes n}$ for a fixed state σ_A. Then, there exist a probability measure $\mathrm{d}\sigma_{AB}$ on the set of extensions σ_{AB} of σ_A and a state ω_Q such that*

$$\rho_{QA^nB^n} \preceq (n+1)^{3d^2} \cdot \omega_Q \otimes \int \sigma_{AB}^{\otimes n} \, \mathrm{d}\sigma_{AB}, \tag{5.28}$$

with $d := d_A d_B^2$.

Notice that in general we have $\omega_Q \neq \rho_Q$, but as pointed out at the beginning of this section, Proposition 5.2.1 can be extended to handle linear constraints on the system carrying the quantum side information Q.

Corollary 5.2.2. *Under the same assumptions of Proposition 5.2.1 with additionally $\Gamma_{Q\to F}$ a linear map and X_F an operator on a Hilbert space F, the state ω_Q can be chosen such that*

$$\Gamma_{Q\to F}(\rho_{QA^n}) = X_F \otimes \sigma_A^{\otimes n} \quad \Longrightarrow \quad \Gamma_{Q\to F}(\omega_Q) = X_F. \tag{5.29}$$

Note that the marginal constraint $\rho_{A^n} = \sigma_A^{\otimes n}$ is a special type of linear constraint, but we do not know if it is possible to extend this to general linear constraints. Moreover, with the introduction of the quantum side information Q, the above results fit the framework of our approximate quantum error correction example, which has been studied in Chapter 4. The proofs of Proposition 5.2.1 and Corollary 5.2.2 are based on the extended Schur-Weyl duality framework laid out in [33, Appendix C] and are given in Section 5.5.

5.3 From de Finetti Reductions to de Finetti Theorems

5.3.1 From de Finetti Reductions to Relative Entropy Inequalities

Noteworthy, de Finetti reductions directly allow to bound the relative entropy distance[3] of symmetric quantum states to convex combinations of tensor product states.

Lemma 5.3.1. *Let Q and G be Hilbert spaces, and ρ_{QG^n} a state symmetric with respect to Q. Consider a de Finetti reduction of the form*

$$\rho_{QG^n} \preceq \mathrm{poly}(n) \cdot \sigma_Q \otimes \int \sigma_G^{\otimes n} d\sigma_G, \tag{5.30}$$

where $d\sigma_G$ is an appropriate measure over the set of quantum states on G. Then, there exists a discrete random variable X, p_X a probability mass function, and σ_G^x quantum states for every $x \in \mathrm{image}(X)$, such that

$$D\Big(\rho_{QG^n} \Big\| \sigma_Q \otimes \sum_x p_X(x) \left[\sigma_G^x\right]^{\otimes n}\Big) \leq \log \mathrm{poly}(n). \tag{5.31}$$

This finding will be the basis to go from de Finetti reductions to representation theorems.

Proof. Thanks to Carathéodory's theorem[4] (see [87, Theorem 1.9]), we can find a discrete random variable X, p_X a probability mass function, σ_G^x quantum states for every $x \in \mathrm{image}(X)$, such that

$$\int \sigma_G^{\otimes n} \, d\sigma_G = \sum_x p_X(x) \left[\sigma_G^x\right]^{\otimes n}. \tag{5.32}$$

[3] Here the word "distance" must be read as "statistical distance". The term *statistical distance* is a general expression used to denote a functional quantifying the similarity between two statistical objects. A statistical distance does not need to be a proper distance, in the metric sense. For example, the relative entropy is not even symmetric, thus it is not a metric.

[4] Carathéodory's theorem states that any point belonging to the convex hull of a set P, subset of a D-dimensional real vector space, can be represented as a convex combination of at most $D+1$ points in P. For complex vector spaces, one can identify $\mathbb{C}$ with $\mathbb{R}^2$.

The following sequence of monotonic operations concludes the proof

$$\rho_{QG^n} \preceq \mathrm{poly}(n) \cdot \sigma_Q \otimes \int \sigma_G^{\otimes n} \, d\sigma_G \tag{5.33}$$

$$\rho_{QG^n} \preceq \mathrm{poly}(n) \cdot \sigma_Q \otimes \sum_x p_X(x) \, [\sigma_G^x]^{\otimes n} \tag{5.34}$$

$$\log(\rho_{QG^n}) - \log\Big(\sigma_Q \otimes \sum_x p_X(x) \, [\sigma_G^x]^{\otimes n}\Big) \preceq \log \mathrm{poly}(n) \cdot 1_G^{\otimes n} \tag{5.35}$$

$$\rho_{QG^n}^{1/2}\Big(\log(\rho_{QG^n}) - \log\Big(\sigma_Q \otimes \sum_x p_X(x) \, [\sigma_G^x]^{\otimes n}\Big)\Big)\rho_{QG^n}^{1/2} \preceq \log \mathrm{poly}(n) \cdot \rho_{QG^n} \tag{5.36}$$

$$\mathrm{Tr}\left[\rho_{QG^n}\Big(\log(\rho_{QG^n}) - \log\Big(\sigma_Q \otimes \sum_x p_X(x) \, [\sigma_G^x]^{\otimes n}\Big)\Big)\right] \leq \log \mathrm{poly}(n) \tag{5.37}$$

$$D\Big(\rho_{QG^n}\Big\|\sigma_Q \otimes \sum_x p_X(x) \, [\sigma_G^x]^{\otimes n}\Big) \leq \log \mathrm{poly}(n), \tag{5.38}$$

where we employed the operator monotonicity of the logarithm as well as of positive maps. □

Applying the de Finetti reduction from Proposition 5.2.1, Lemma 5.3.1 immediately leads to the following bound.

Corollary 5.3.2. *Under the same assumptions of Proposition 5.2.1, there exist a discrete random variable X, p_X a probability mass function, a state ω_Q on Q, and σ_{AB}^x extensions of σ_A for every $x \in \mathrm{image}(X)$, such that*

$$D\Big(\rho_{QA^nB^n}\Big\|\omega_Q \otimes \sum_x p_X(x)[\sigma_{AB}^x]^{\otimes n}\Big) \leq 3d^2 \cdot \log(n+1), \tag{5.39}$$

with $d := d_A d_B^2$.

Notice that the right-hand side in Corollary 5.3.2 is not small for any non-trivial dimension. In other words, if $d \neq 0$, the right-hand side diverges when taking the asymptotic limit $n \to \infty$. However, in the next subsections we show how this corollary can be employed to derive de Finetti representation theorems.

5.3.2 Classical case

The following gives a proof for the classical de Finetti theorem based on entropy inequalities. The basic idea is to condition iteratively on pairs of random variables by using the chain rule for the Kullback–Leibler divergence (Lemma 5.1.2). This is reminiscent of the information-theoretic proof strategy from [18], based on the chain rule of the conditional mutual information, which was also employed in our papers [13] and [14], and in Section 3.6.

Proposition 5.3.3. *Let $k \in \{1, \dots, n-1\}$, X and $G_1 \cdots G_n$ discrete random variables, $r_{G_1 \cdots G_n}$, p_X and $\prod_{i=1}^n s^x_{G_i}$ probability mass functions for every $x \in \mathrm{image}(X)$, and assume $r_{G_1,\dots,G_n}$ to be symmetric. Whenever we have*

$$D_{KL}\Big(r_{G_1 \cdots G_n} \Big\| \sum_x p_X(x) \prod_{i=1}^n s^x_{G_i}\Big) \leq \log \mathrm{poly}(n), \tag{5.40}$$

then there exists a probability mass function q_X such that

$$\Big\| r_{G_1 \cdots G_k} - \sum_x q_X(x) \prod_{i=1}^k s^x_{G_i} \Big\|_1 \leq O\left(\sqrt{\frac{k}{n} \cdot \log n}\right). \tag{5.41}$$

Note that, in general, $q_X \neq p_X$, and the bound $O\left(\sqrt{\frac{k}{n} \cdot \log n}\right)$ is known to be suboptimal in n (see, e.g., [48, Lemma 1]). Nevertheless, our strategy provides a novel proof technique that can be used to systematically generate classical de Finetti theorems from classical de Finetti reductions.

Proof. In what follows we assume, for simplicity, $k = 2$ and n even. The generalization to an arbitrary $k \in \{1, \dots, n-1\}$ is obtained by grouping variables in groups of k in the subsequent proof. Moreover, we prove a slightly more general statement, where we do not assume the probability mass function $r_{G_1 \cdots G_n}$ to be permutation invariant. More precisely, without assuming permutation invariance, we show that there exist a $m \in \{0, \dots, n/2 - 1\}$

and a probability mass function q_X such that

$$\left\| r_{G_{2m+1}G_{2m+2}} - \sum_x q_X(x) s^x_{G_{2m+1}} s^x_{G_{2m+2}} \right\|_1 \leq \sqrt{2\ln(2)} \cdot \sqrt{\frac{2\log \mathrm{poly}(n)}{n}} \ . \tag{5.42}$$

If we then assume, as in the statement of Proposition 5.3.3, $r_{G_1,\ldots,G_n}$ to be symmetric, the existential quantification can be replaced, by symmetry, with a universal quantification.

We abbreviate $s^x_{G_1\cdots G_h} := \prod_{i=1}^h s^x_{G_i}$ for $h = 1, \ldots, n$, and using the chain rule for the Kullback–Leibler divergence (Lemma 5.1.2), we obtain

$$D_{KL}\Big(r_{G_1\cdots G_n} \Big\| \sum_x p_X(x) s^x_{G_1\cdots G_n}\Big) = D_{KL}\Big(r_{G_1G_2} \Big\| \sum_x p_X(x) s^x_{G_1} s^x_{G_2}\Big) \tag{5.43}$$

$$+ \mathop{\mathbf{E}}_{G_1G_2} \left\{ D_{KL}\Big(r_{G_3G_4|G_1G_2} \Big\| \frac{\sum_x p_X(x) s^x_{G_1\cdots G_4}}{\sum_x p_X(x) s^x_{G_1G_2}}\Big)\right\} + \ldots \tag{5.44}$$

with the sum formed of $n/2$ terms and the expectation value taken with respect to $r_{G_1G_2}$, and $r_{G_3G_4|G_1G_2} := \frac{r_{G_1G_2G_3G_4}}{r_{G_1G_2}}$. Defining $p(x|g_1g_2) := \frac{p_X(x) s^x_{G_1G_2}(g_1g_2)}{\sum_x p_X(x) s^x_{G_1G_2}(g_1g_2)}$ for every $x \in \mathrm{image}(X)$, $g_1 \in \mathrm{image}(G_1)$ and $g_2 \in \mathrm{image}(G_2)$, we simplify the above sum as

$$D_{KL}\Big(r_{G_1\cdots G_n} \Big\| \sum_x p_X(x) s^x_{G_1\cdots G_n}\Big) = D_{KL}\Big(r_{G_1G_2} \Big\| \sum_x p_X(x) s^x_{G_1} s^x_{G_2}\Big) \tag{5.45}$$

$$+ \mathop{\mathbf{E}}_{G_1G_2} \left\{ D_{KL}\Big(r_{G_3G_4|G_1G_2} \Big\| \sum_x p(x|G_1G_2) s^x_{G_3} s^x_{G_4}\Big)\right\} + \ldots \tag{5.46}$$

with similar definitions and simplifications for the other addends. Because each term in the sum is non-negative and their sum is, by assumption, smaller than or equal to $\log \mathrm{poly}(n)$, there must be at least a term in the sum smaller than or equal to $\frac{\log \mathrm{poly}(n)}{n/2}$. In other words, there exists a $m \in \{0, \ldots, n/2 - 1\}$ such that

$$\mathop{\mathbf{E}}_{G_1\cdots G_{2m}} \left\{ D_{KL}\Big(r_{G_{2m+1}G_{2m+2}|G_1\cdots G_{2m}} \Big\| \sum_x p(x|G_1\cdots G_{2m}) s^x_{G_{2m+1}} s^x_{G_{2m+2}}\Big)\right\} \tag{5.47}$$

$$\leq \frac{\log \operatorname{poly}(n)}{n/2} \tag{5.48}$$

and thanks to the joint convexity of Kullback–Leibler divergence [87, Corollary 5.12] we obtain

$$D_{KL}\Big(\mathop{\mathbf{E}}_{G_1\cdots G_{2m}} \left\{r_{G_{2m+1}G_{2m+2}|G_1\cdots G_{2m}}\right\} \Big\| \sum_x \mathop{\mathbf{E}}_{G_1\ldots G_{2m}} \left\{p(x|G_1\cdots G_{2m})\right\} s^x_{G_{2m+1}} s^x_{G_{2m+2}}\Big) \tag{5.49}$$

$$\leq \frac{\log \operatorname{poly}(n)}{n/2}. \tag{5.50}$$

Defining $q_X(x) := \mathop{\mathbf{E}}_{G_1\cdots G_{2m}} \{p(x|G_1\cdots G_{2m})\}$ for every $x \in \operatorname{image}(X)$, and using the law of total probability $\mathop{\mathbf{E}}_{G_1\cdots G_{2m}} \left\{r_{G_{2m+1}G_{2m+2}|G_1\cdots G_{2m}}\right\} = r_{G_{2m+1}G_{2m+2}}$, we find

$$D_{KL}\Big(r_{G_{2m+1}G_{2m+2}} \Big\| \sum_x q_X(x) s^x_{G_{2m+1}} s^x_{G_{2m+2}}\Big) \leq \frac{\log \operatorname{poly}(n)}{n/2}. \tag{5.51}$$

classical Pinsker's inequality (Theorem 5.1.1) then concludes the proof. □

5.3.3 Quantum case

Using informationally complete measurements (Definition 2.2.3), we can leverage the previous classical result to the quantum setting, thus obtaining a new proof for finite quantum de Finetti theorems employing de Finetti reductions.

Theorem 5.3.4. *Let $k \in \{1, \ldots, n-1\}$, X be a discrete random variable, $G_1 \cdots G_n$ Hilbert spaces with $G_1 \cong \ldots \cong G_n$, ρ_{G^n} and σ^x_G quantum states for every $x \in \operatorname{image}(X)$, p_X a probability mass function, and assume ρ_{G^n} to be symmetric. Whenever we have*

$$D\Big(\rho_{G^n} \Big\| \sum_x p_X(x) [\sigma^x_G]^{\otimes n}\Big) \leq \log \operatorname{poly}(n), \tag{5.52}$$

then there exists a probability mass function q_X such that

$$\Big\|\rho_{G^k} - \sum_x q_X(x) [\sigma^x_G]^{\otimes k}\Big\|_1 \leq O\left(\sqrt{\frac{k \cdot d_G^k}{n} \cdot \log n}\right). \tag{5.53}$$

Notice that the bound on the approximation error grows exponentially fast with k. Whether it is possible to improve that k-dependence and maintain the proposed proof technique is still an open question. On the other hand, we already know that the dependence in n is suboptimal. This is not surprising following our comment on Proposition 5.3.3, which is used to prove Theorem 5.3.4.

Proof. In what follows we assume, for simplicity, $k = 2$. The generalization to an arbitrary $k \in \{1, \ldots, n-1\}$ is obtained using Proposition 5.3.3 in its full generality. We start by measuring ρ_{G^n} and $[\sigma_G^x]^{\otimes n}$ with the same product measurement, i.e., we choose $\mu_{G_1 \cdots G_n} := \otimes_{i=1}^n \mu_{G_i}$. Thanks to the monotonicity of the quantum relative entropy under positive maps, we have

$$D_{KL}\Big(\mu_{G_1\cdots G_n}(\rho_{G^n})\Big\|\sum_x p_X(x)\mu_{G_1\cdots G_n}([\sigma_G^x]^{\otimes n})\Big) \leq D\Big(\rho_{G^n}\Big\|\sum_x p_X(x)[\sigma_G^x]^{\otimes n}\Big) \tag{5.54}$$

$$\leq \log \mathrm{poly}(n) \tag{5.55}$$

and using Proposition 5.3.3 for the post-measurement probability distributions we find

$$\Big\|(\mu_{G_1}\otimes\mu_{G_2})(\rho_{G_1G_2} - \sum_x q_X(x)\sigma_{G_1}^x \otimes \sigma_{G_2}^x)\Big\|_1 \leq \sqrt{2\ln(2)}\cdot\sqrt{\frac{\log \mathrm{poly}(n)}{n}} \tag{5.56}$$

for some probability mass function q_X. Moreover, [17, Lemma 14] shows that we can choose the product measurement $\mu_{G_1} \otimes \mu_{G_2}$ such that

$$\Big\|\rho_{G_1G_2} - \sum_x q_X(x)\sigma_{G_1}^x \otimes \sigma_{G_2}^x\Big\|_1 \leq 18 d_G \cdot \Big\|(\mu_{G_1}\otimes\mu_{G_2})(\rho_{G_1G_2} - \sum_x q_X(x)\sigma_{G_1}^x \otimes \sigma_{G_2}^x)\Big\|_1. \tag{5.57}$$

This concludes the proof. □

5.4 De Finetti theorems for quantum channels: simplifying the constraints

In this section we use de Finetti reductions to derive a new de Finetti theorem for quantum channels. In comparison to our results from Section 4.2, which were given for bipartite quantum channels, here we look at generic single-input/single-ouput quantum channels as in [39], and we show that is possible to drop one constraint and still achieve asymptotic convergence. While we are not able to prove the theoretical minimality of our constraints, the simplification of the existing conditions is definitely a fundamental step in the right direction. Moreover, our new results provide insights on the "power" of the constraints and their effect on the convergence speed.

For the purposes of this section we do not need side information, which would instead be required for bipartite quantum channels. If we remove the system Q carrying the quantum side information, Proposition 5.2.1 reduces to the quantum de Finetti reduction of [33, Corollary 3.2] with a slightly worse dimensional dependence in the scalar prefactor. Thus, in what follows we will directly use their quantum de Finetti reduction.

Corollary 5.4.1. *[33, Corollary 3.2] Let B and $\overline{B}$ be Hilbert spaces and $\rho_{B^n\overline{B}^n}$ a state invariant under permutation of the $B\overline{B}$-systems. Moreover, let $\rho_{B^n} = \sigma_B^{\otimes n}$ for a fixed state σ_B. Then, there exists a probability measure $\mathrm{d}\sigma_{B\overline{B}}$ on the set of extensions $\sigma_{B\overline{B}}$ of σ_B such that*

$$\rho_{B^n\overline{B}^n} \preceq (n+1)^{d^2-1} \int \sigma_{B\overline{B}}^{\otimes n} \mathrm{d}\sigma_{B\overline{B}}, \tag{5.58}$$

with $d := d_B d_{\overline{B}}^2$.

Applied to the above de Finetti reduction, Lemma 5.3.1 gives the following relative entropy

bound

$$D\left(\rho_{B^n\overline{B}^n}\middle\|\sum_x p_X(x)[\sigma^x_{B\overline{B}}]^{\otimes n}\right)\leq (d^2-1)\cdot\log(n+1), \tag{5.59}$$

where X is a discrete random variable, p_X a probability mass function, and $\sigma^x_{B\overline{B}}$ extensions of σ_B for every $x\in\text{image}(X)$.

Using Theorem 5.3.4 with $k=2$, we then conclude the validity of the following de Finetti representation theorem

$$\left\|\rho_{(B\overline{B})^2_1}-\sum_x q_X(x)[\sigma^x_{B\overline{B}}]^{\otimes 2}\right\|_1\leq 18\cdot\sqrt{2}\cdot d_{B\overline{B}}\cdot\sqrt{2\ln(2)}\cdot\sqrt{\frac{(d^2-1)\cdot\log(n+1)}{n}}, \tag{5.60}$$

where q_X is an appropriate probability mass function. We formalize this result in the following theorem, which is stated for an arbitrary $k\in\{1,\ldots,n-1\}$.

Theorem 5.4.2. *Let $k\in\{1,\ldots,n-1\}$, B and $\overline{B}$ Hilbert spaces, and $\rho_{(B\overline{B})^n_1}$ a state invariant under permutation of the $B\overline{B}$-systems. Moreover, let $\rho_{B^n}=\sigma_B^{\otimes n}$ for a fixed state σ_B. Then, there exist a discrete random variable X, a probability mass function q_X, and $\sigma^x_{B\overline{B}}$ extensions of σ_B for every $x\in\text{image}(X)$, such that*

$$\left\|\rho_{(B\overline{B})^k_1}-\sum_x q_X(x)[\sigma^x_{B\overline{B}}]^{\otimes k}\right\|_1\leq 18\cdot\sqrt{k}\cdot d_{B\overline{B}}^{k/2}\cdot\sqrt{2\ln(2)}\cdot\sqrt{\frac{(d_Bd_{\overline{B}}^2-1)\cdot\log(n+1)}{n}}. \tag{5.61}$$

If we set $\sigma_B=\frac{1_B}{d_B}$, we obtain the following corollary, valid for quantum channels and expressed via the Choi states (as in Theorem 4.2.1).

Corollary 5.4.3. *Let $k\in\{1,\ldots,n-1\}$, B and $\overline{B}$ Hilbert spaces, and $\rho_{(B\overline{B})^n_1}$ a state invariant under permutation of the $B\overline{B}$-systems. Moreover, let $\rho_{B^n}=\frac{1_B^{\otimes n}}{d_B^n}$. Then, there exist a discrete random variable X, a probability mass function q_X, and $\sigma^x_{B\overline{B}}$ quantum states satisfying $\sigma^x_B=\frac{1_B}{d_B}$ for every $x\in\text{image}(X)$, such that*

$$\left\|\rho_{(B\overline{B})^k_1}-\sum_x q_X(x)[\sigma^x_{B\overline{B}}]^{\otimes k}\right\|_1\leq 18\cdot\sqrt{k}\cdot d_{B\overline{B}}^{k/2}\cdot\sqrt{2\ln(2)}\cdot\sqrt{\frac{(d_Bd_{\overline{B}}^2-1)\cdot\log(n+1)}{n}}. \tag{5.62}$$

We can compare the above corollary with the de Finetti theorems for quantum channels from Section 4.2. In particular, consider a quantum state $\rho_{(B\overline{B})_1^n}$ satisfying the following two conditions

1. $\rho_{(B\overline{B})_1^n} = \mathcal{U}^{\pi}_{(B\overline{B})_1^n}(\rho_{(B\overline{B})_1^n})\ \forall \pi \in \mathfrak{S}_n$
2. $\rho_{(B\overline{B})_1^{n-1}B_n} = \rho_{(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}$

then Theorem 4.2.1 guarantees the bound

$$\left\| \rho_{(B\overline{B})_1^k} - \sum_x p_X(x) [\sigma^x_{B\overline{B}}]^{\otimes k} \right\|_1 \leq k \cdot 2 d_{B\overline{B}} \cdot \sqrt{2\ln(2)} \cdot \sqrt{\frac{(k-1)\log\left(d_{B\overline{B}}\right)}{n-k+1}}, \tag{5.63}$$

with $k \in \{1, \ldots, n-1\}$, for an appropriate probability mass function p_X. In comparison, Corollary 5.4.3 replaces the marginal constraint $\rho_{(B\overline{B})_1^{n-1}B_n} = \rho_{(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}$ with a much simpler one, i.e., $\rho_{B^n} = \frac{1_B^{\otimes n}}{d_B^n}$. However, this simplification comes at a price. In fact in Corollary 5.4.3 we find a new factor exponential in k, i.e., $d_{B\overline{B}}^{k/2}$. We wonder if this prefactor is just an artefact of our derivation or if it is really necessary. In the latter case, we would be able to characterize the power of the non-trivial constraint $\rho_{(B\overline{B})_1^{n-1}B_n} = \rho_{(B\overline{B})_1^{n-1}} \otimes \frac{1_{B_n}}{d_B}$, with respect to the more basic $\rho_{B^n} = \frac{1_B^{\otimes n}}{d_B^n}$. We leave it as an open question.

5.5 Proof of Proposition 5.2.1

First, we prove the following lemma, which is a version of Proposition 5.2.1 for pure states. Second, we generalize the statement to mixed states by employing Lemma 5.5.2 on symmetric purifications of permutation invariant states.

Lemma 5.5.1. *Let Q, A and B be Hilbert spaces, and let $\rho_{QA^nB^n}$ a pure quantum state symmetric with respect to Q. Moreover, assume $\rho_{QA^nB^n}$ satisfies the marginal constraint*

$\rho_{A^n} = \sigma_A^{\otimes n}$, for a given quantum state σ_A on A. Then, there exist a probability measure $\mathrm{d}\phi$ on the set of purifications $|\phi\rangle\langle\phi|_{AB}$ of σ_A and a quantum state ω_Q on Q, such that

$$\rho_{QA^nB^n} \preceq (n+1)^{3d^2} \cdot \omega_Q \otimes \int |\phi\rangle\langle\phi|_{AB}^{\otimes n} \, \mathrm{d}\phi, \tag{5.64}$$

with $d := \max\{d_A, d_B\}$.

Proof. The idea behind the proof is based on [33, Lemma 3.1]. In particular, we assume without loss of generality that $d = d_A = d_B$, and that σ_A is invertible on A. In fact, it is always possible to embed the smaller space into a larger system of dimension d and replace σ_A by $\sigma_A + \epsilon \, 1_A$, for $\epsilon > 0$. The claim is then obtained by taking the limit $\epsilon \to 0$. We define the non-normalized maximally entangled state (cf. (2.39))

$$|\theta\rangle_{AB} := \sum_i |d_i\rangle_A \otimes |e_i\rangle_B, \tag{5.65}$$

where $\{|d_i\rangle_A\}_i$ and $\{|e_i\rangle_B\}_i$ are orthonormal bases of A and B, respectively. Let now

$$T_{A^nB^n} := \int (1_{A^n} \otimes U_B^{\otimes n}) |\theta\rangle\langle\theta|_{AB}^{\otimes n} (1_{A^n} \otimes U_B^{\otimes n})^\dagger \mathrm{d}U, \tag{5.66}$$

where $\mathrm{d}U$ is the Haar measure on the group of unitaries on B. As $(\sigma_A^{1/2} \otimes U_B)|\theta\rangle\langle\theta|(\sigma_A^{1/2} \otimes U_B^\dagger)$ is a purification of σ_A for any unitary U_B [69, Subsection 9.2.2], we can write

$$\begin{aligned} \tau_{A^nB^n} &:= (\sigma_A^{\otimes n} \otimes 1_{B^n})^{1/2} T_{A^nB^n} (\sigma_A^{\otimes n} \otimes 1_{B^n})^{1/2} && (5.67) \\ &= \int \left[(\sigma_A^{1/2} \otimes U_B) |\theta\rangle\langle\theta|_{AB} (\sigma_A^{1/2} \otimes U_B^\dagger) \right]^{\otimes n} \mathrm{d}U && (5.68) \\ &= \int |\phi\rangle\langle\phi|_{AB}^{\otimes n} \, \mathrm{d}\phi && (5.69) \end{aligned}$$

for some measure $\mathrm{d}\phi$ on the set of purifications $|\phi\rangle\langle\phi|_{AB}$ of σ_A.

We proceed by analysing the structure of $T_{A^nB^n}$. For this purpose, we employ the well-known

Schur-Weyl duality[5], which equips the product space $(A \otimes B)^{\otimes n}$ with the structure

$$A^{\otimes n} \cong \bigoplus_{\lambda} U_{A,\lambda} \otimes V_{A,\lambda} \tag{5.70}$$

and

$$B^{\otimes n} \cong \bigoplus_{\lambda} U_{B,\lambda} \otimes V_{B,\lambda} \tag{5.71}$$

where λ indexes the Young diagrams. We then define the operator (cf. [33, Lemma 3.1])

$$S_{A^nB^n} := (\kappa_{A^n} \otimes 1_{B^n})^{-1/2} T_{A^nB^n} (\kappa_{A^n} \otimes 1_{B^n})^{-1/2} \tag{5.72}$$

with $\kappa_{A^n} := \sum_\lambda \frac{d_{V_\lambda}}{d_{U_\lambda}} \cdot 1_{U_{A,\lambda}} \otimes 1_{V_{A,\lambda}}$, featuring $T_{A^nB^n}$ and by properties of Schur-Weyl duality we have [33, Lemma C.1]

$$S_{A^nB^n} = 1_{\mathrm{Sym}^n(A \otimes B)}. \tag{5.73}$$

Consider now the operator

$$R_{QA^nB^n}(\omega) := \left(\omega_Q^{-1/2} \otimes \kappa_{A^n}^{-1/2} \left(\sigma_A^{\otimes n}\right)^{-1/2} \otimes 1_{B^n}\right) \rho_{QA^nB^n} \left(\omega_Q^{-1/2} \otimes \left(\sigma_A^{\otimes n}\right)^{-1/2} \kappa_{A^n}^{-1/2} \otimes 1_{B^n}\right) \tag{5.74}$$

parametrized by an arbitrary state ω_Q on Q, where ω_Q^{-1} denotes the generalized inverse of ω_Q (see (2.19)). Since $\rho_{QA^nB^n}$ is pure and symmetric with respect to Q, its support is contained in $Q \otimes \mathrm{Sym}^n(A \otimes B)$ and we have

$$\mathrm{supp}(\rho_{QA^nB^n}) \subseteq Q \otimes \mathrm{Sym}^n(A \otimes B) \Longrightarrow \mathrm{supp}(R_{QA^nB^n}(\omega)) \subseteq Q \otimes \mathrm{Sym}^n(A \otimes B). \tag{5.75}$$

[5] *Schur-Weyl duality* asserts that one can isomorphically decompose the n-fold tensor product space $\left(\mathbb{C}^d\right)^{\otimes n}$ into a direct sum of tensor products $U_\lambda \otimes V_\lambda$ of irreducible representations of the unitary group and the symmetric group, for the various Young diagrams λ of size n with at most d rows. I.e., $\left(\mathbb{C}^d\right)^{\otimes n} \cong \bigoplus_\lambda U_\lambda \otimes V_\lambda$.

Using this fact, we find [81]

$$R_{QA^nB^n}(\omega) \preceq \|R_{QA^nB^n}(\omega)\|_\infty \cdot 1_Q \otimes 1_{\mathrm{Sym}^n(A\otimes B)} \tag{5.76}$$

$$= \|R_{QA^nB^n}(\omega)\|_\infty \cdot 1_Q \otimes S_{A^nB^n}, \tag{5.77}$$

which is rewritten as

$$\rho_{QA^nB^n} \preceq \|R_{QA^nB^n}(\omega)\|_\infty \cdot \omega_Q \otimes (\sigma_A^{\otimes n} \otimes 1_{B^n})^{1/2} T_{A^nB^n} (\sigma_A^{\otimes n} \otimes 1_{B^n})^{1/2} \tag{5.78}$$

$$= \|R_{QA^nB^n}(\omega)\|_\infty \cdot \omega_Q \otimes \tau_{A^nB^n}. \tag{5.79}$$

We now make an appropriate choice for the state ω_Q such that we have $\|R_{QA^nB^n}(\omega)\|_\infty \leq (n+1)^{3d^2}$, and hence together with the above what we set out to prove.

For that, we choose ω_Q as the reduced state of

$$\omega_{QA^nB^n} := \frac{1}{\mathrm{Tr}\left(W_{QA^nB^n}\right)} \cdot W_{QA^nB^n}, \tag{5.80}$$

with

$$W_{QA^nB^n} := \left(1_Q \otimes \kappa_{A^n}^{-1/2} (\sigma_A^{\otimes n})^{-1/2} \otimes 1_B^n\right) \rho_{QA^nB^n} \left(1_Q \otimes (\sigma_A^{\otimes n})^{-1/2} \kappa_{A^n}^{-1/2} \otimes 1_B^n\right) \tag{5.81}$$

Because of the structure of the operator $W_{QA^nB^n}$, its support is a subset of $Q \otimes \mathrm{Sym}^n(A \otimes B)$ and we bound the denominator as

$$\mathrm{Tr}\left(W_{QA^nB^n}\right) = \mathrm{Tr}\left[\kappa_{A^n}^{-1/2} (\sigma_A^{\otimes n})^{-1/2} \rho_{A^n} (\sigma_A^{\otimes n})^{-1/2} \kappa_{A^n}^{-1/2}\right] \tag{5.82}$$

$$= \mathrm{Tr}\left[\kappa_{A^n}^{-1/2} (\sigma_A^{\otimes n})^{-1/2} \sigma_A^{\otimes n} (\sigma_A^{\otimes n})^{-1/2} \kappa_{A^n}^{-1/2}\right] \tag{5.83}$$

$$= \mathrm{Tr}\left(\kappa_{A^n}^{-1}\right) \tag{5.84}$$

$$= \sum_\lambda d_{U_\lambda}^2 \tag{5.85}$$

$$\leq (n+1)^{2d} \cdot \sum_\lambda 1 \tag{5.86}$$

$$\leq (n+1)^{3d}, \tag{5.87}$$

where we used [47, Lemma 10] to bound $d_{U_\lambda} \leq (n+1)^d$ as well as $\sum_\lambda 1 \leq (n+1)^d$. We then obtain the desired bound for $\|R_{QA^nB^n}\|_\infty$ as

$$\begin{aligned}
\|R_{QA^nB^n}(\omega)\|_\infty &= \left\|(\omega_Q^{-1/2} \otimes 1_{A^nB^n}) W_{QA^nB^n} (\omega_Q^{-1/2} \otimes 1_{A^nB^n})\right\|_\infty && (5.88)\\
&= \mathrm{Tr}\left(W_{QA^nB^n}\right) \cdot \left\|(\omega_Q^{-1/2} \otimes 1_{A^nB^n}) \omega_{QA^nB^n} (\omega_Q^{-1/2} \otimes 1_{A^nB^n})\right\|_\infty && (5.89)\\
&\leq (n+1)^{3d} \cdot d_{\mathrm{Sym}^n(A\otimes B)} && (5.90)\\
&\leq (n+1)^{d^2+3d-1} && (5.91)\\
&\leq (n+1)^{3d^2}, && (5.92)
\end{aligned}$$

where the first inequality in

$$\begin{aligned}
\left\|\left(\omega_Q^{-1/2} \otimes 1_{A^nB^n}\right) \omega_{QA^nB^n} \left(\omega_Q^{-1/2} \otimes 1_{A^nB^n}\right)\right\|_\infty &\leq d_{\mathrm{Sym}^n(A\otimes B)} && (5.93)\\
&\leq (n+1)^{d^2-1} && (5.94)
\end{aligned}$$

is the application of [79, Proposition 4.3] to the state $\omega_{QA^nB^n}$, and the second one is a standard upper bound on the dimension of the symmetric subspace [87, Corollary 7.3]. This finishes the proof. □

It is well known that permutation invariant states can be purified by symmetric (pure) states [21, Lemma II.5]. The following lemma shows that this is still true in the presence of an additional quantum system, i.e., with side information. This result and its proof can be seen as a generalization of the methods presented in [21].

Lemma 5.5.2. *Let C, H be Hilbert spaces, and ρ_{CH^n} a state symmetric with respect to C. Then, there exists a pure state $|\psi_\rho\rangle \in C \otimes R_C \otimes \mathrm{Sym}^n(H \otimes R_H)$ with $R_C \cong C$ and $R_H \cong H$,*

such that

$$\mathrm{Tr}_{R_C R_H^n}\left(|\psi_\rho\rangle\langle\psi_\rho|\right) = \rho_{CH^n}. \tag{5.95}$$

Proof. Let $\{|j\rangle_C\}_{j=1,\ldots,d_C}$,$\{|\tilde{j}\rangle_{R_C}\}_{j=1,\ldots,d_C}$, $\{|i\rangle_H\}_{i=1,\ldots,d_H}$,$\{|\tilde{i}\rangle_{R_H}\}_{i=1,\ldots,d_H}$, orthonormal bases for C, R_C, H and R_H respectively. We show that the following choice for $|\psi_\rho\rangle$ satisfies the requirements of the lemma

$$|\psi_\rho\rangle := \left(\sqrt{\rho_{CH^n}} \otimes 1_{R_C R_H^n}\right) \cdot \left[\left(\sum_j |j\rangle_C|\tilde{j}\rangle_{R_C}\right) \otimes \left(\sum_i |i\rangle_H|\tilde{i}\rangle_{R_H}\right)^{\otimes n}\right]. \tag{5.96}$$

Since ρ_{CH^n} is symmetric with respect to C, the same will hold for $\sqrt{\rho_{CH^n}}$. In fact, let $U^\pi_{H^n}$ an arbitrary permutation operator on H^n, then

$$\left[(1_C \otimes U^\pi_{H^n})\sqrt{\rho_{CH^n}}(1_C \otimes (U^\pi_{H^n})^\dagger)\right]^2 = (1_C \otimes U^\pi_{H^n})\sqrt{\rho_{CH^n}}\sqrt{\rho_{CH^n}}(1_C \otimes (U^\pi_{H^n})^\dagger) \tag{5.97}$$

$$= \rho_{CH^n}, \tag{5.98}$$

which implies $(1_C \otimes U^\pi_{H^n})\sqrt{\rho_{CH^n}}(1_C \otimes (U^\pi_{H^n})^\dagger) = \sqrt{\rho_{CH^n}}$.

Now, to show that $|\psi_\rho\rangle \in C \otimes R_C \otimes \mathrm{Sym}^n(H \otimes R_H)$, let $U^\pi_{H^n} \otimes U^\pi_{R_H^n}$ be an arbitrary permutation on $(H \otimes R_H)^{\otimes n}$. We have

$$(1_C \otimes 1_{R_C} \otimes U^\pi_{H^n} \otimes U^\pi_{R_H^n})|\psi_\rho\rangle \tag{5.99}$$

$$= \left((1_C \otimes U^\pi_{H^n})\sqrt{\rho_{CH^n}} \otimes 1_{R_C} \otimes U^\pi_{R_H^n}\right) \cdot \left[\left(\sum_j |j\rangle_C|\tilde{j}\rangle_{R_C}\right) \otimes \left(\sum_i |i\rangle_H|\tilde{i}\rangle_{R_H}\right)^{\otimes n}\right] \tag{5.100}$$

$$= \left((1_C \otimes U^\pi_{H^n})\sqrt{\rho_{CH^n}} \otimes 1_{R_C} \otimes U^\pi_{R_H^n}\right) \tag{5.101}$$

$$\cdot \left[\left(\sum_j |j\rangle_C|\tilde{j}\rangle_{R_C}\right) \otimes ((U^\pi_{H^n})^\dagger \otimes (U^\pi_{R_H^n})^\dagger)\left(\sum_i |i\rangle_H|\tilde{i}\rangle_{R_H}\right)^{\otimes n}\right] \tag{5.102}$$

$$= \left((1_C \otimes U^\pi_{H^n})\sqrt{\rho_{CH^n}}(1_C \otimes (U^\pi_{H^n})^\dagger) \otimes 1_{R_C R_H^n}\right) \tag{5.103}$$

$$\cdot \left[\left(\sum_j |j\rangle_C|\tilde{j}\rangle_{R_C}\right) \otimes \left(\sum_i |i\rangle_H|\tilde{i}\rangle_{R_H}\right)^{\otimes n}\right] \tag{5.104}$$

$$= \left(\sqrt{\rho_{CH^n}} \otimes 1_{R_C R_H^n}\right) \cdot \left[\left(\sum_j |j\rangle_C |\tilde{j}\rangle_{R_C}\right) \otimes \left(\sum_i |i\rangle_H |\tilde{i}\rangle_{R_H}\right)^{\otimes n}\right] \tag{5.105}$$

$$= |\psi_\rho\rangle, \tag{5.106}$$

where we used that $\left(\sum_i |i\rangle_H |\tilde{i}\rangle_{R_H}\right)^{\otimes n} \in \mathrm{Sym}^n(H \otimes R_H)$. Finally, we compute the partial trace

$$\mathrm{Tr}_{R_C R_H^n}\left(|\psi_\rho\rangle\langle\psi_\rho|\right) \tag{5.107}$$

$$= \mathrm{Tr}_{R_H^n}\left[\left(\sqrt{\rho_{CH^n}} \otimes 1_{R_H^n}\right)\left[1_C \otimes \left(\sum_{i,i'} |i\rangle_H \langle i'|_H \otimes |\tilde{i}\rangle_{R_H}\langle\tilde{i'}|_{R_H}\right)^{\otimes n}\right]\left(\sqrt{\rho_{CH^n}} \otimes 1_{R_H^n}\right)\right] \tag{5.108}$$

$$= \sqrt{\rho_{CH^n}}\left(1_C \otimes 1_{H^n}\right)\sqrt{\rho_{CH^n}} \tag{5.109}$$

$$= \rho_{CH^n}. \tag{5.110}$$

□

With the above lemmas, we can now prove Proposition 5.2.1 and Corollary 5.2.2. The proofs are done through straightforward extensions of the purification technique [33, Corollary 3.2].

Proof of Proposition 5.2.1. : Using Lemma 5.5.2 we see that $\rho_{QA^nB^n}$ has a symmetric purification $\rho_{QR_QA^nB^nR_{AB}^n}$ with purifying system $R_Q \otimes R_{AB}^{\otimes n}$, where the local dimensions are $d_{R_Q} = d_Q$ and $d_{R_{AB}} = d_A d_B$. Lemma 5.5.1 with Q replaced by $Q \otimes R_Q$ and B replaced by $B \otimes R_{AB}$, applied to $\rho_{QR_QA^nB^nR_{AB}^n}$, yields

$$\rho_{QR_QA^nB^nR_{AB}^n} \preceq (n+1)^{3d^2} \omega_{QR_Q} \otimes \int |\phi\rangle\langle\phi|_{ABR_{AB}}^{\otimes n} \mathrm{d}\phi \ , \tag{5.111}$$

where $\mathrm{d}\phi$ is a measure on the purifications $|\phi\rangle\langle\phi|_{ABR_{AB}}$ of σ_A, and the dimensional factor is $d = \max\{d_A, d_{(B\otimes R_{AB})}\} = d_A d_B^2$. Taking the partial trace over $R_Q \otimes R_{AB}^{\otimes n}$ on both sides gives

$$\rho_{QA^nB^n} \preceq (n+1)^{3d^2} \omega_Q \otimes \int \left(\mathrm{Tr}_{R_{AB}} \left[|\phi\rangle\langle\phi|_{ABR_{AB}} \right] \right)^{\otimes n} \mathrm{d}\phi \ . \tag{5.112}$$

The claim follows because the measure $\mathrm{d}\phi$ on the pure states $|\phi\rangle\langle\phi|_{ABR_{AB}}$ can be replaced by the induced measure $\mathrm{d}\sigma_{AB}$ on the marginal states $\sigma_{AB} = \mathrm{Tr}_{R_{AB}}[|\phi\rangle\langle\phi|_{ABR_{AB}}]$. This concludes the proof. □

Proof of Corollary 5.2.2. : Keeping in mind the proof of Proposition 5.2.1, we now show by direct evaluation that $\Gamma_{Q\to F}(\rho_{QA^n}) = X_F \otimes \sigma_A^{\otimes n}$ implies $\Gamma_{Q\to F}(\omega_Q) = X_F$. In particular, from the proof of Lemma 5.5.1, we have the structure

$$\omega_{QR_Q} = \mathrm{Tr}_{A^nB^nR_{AB}^n} \left[\frac{W_{Q_1QR_QA^nB^nR_{AB}^n}}{\mathrm{Tr}\left[W_{QR_QA^nB^nR_{AB}^n}\right]} \right], \tag{5.113}$$

featuring the operator

$$W_{QR_QA^nB^nR_{AB}^n} \tag{5.114}$$

$$= (1_{QR_Q} \otimes \kappa_{A^n}^{-1/2} (\sigma_A^{\otimes n})^{-1/2} \otimes 1_{B^nR_{AB}^n}) \rho_{QR_QA^nB^nR_{AB}^n} (1_{QR_Q} \otimes (\sigma_A^{\otimes n})^{-1/2} \kappa_{A^n}^{-1/2} \otimes 1_{B^nR_{AB}^n}). \tag{5.115}$$

Hence, we have

$$\Gamma_{Q\to F}(\omega_Q) = \mathrm{Tr}_{R_Q}[\Gamma_{Q\to F}(\omega_{QR_Q})] \tag{5.116}$$

$$= \mathrm{Tr}_{R_Q} \left[\Gamma_{Q\to F} \left(\mathrm{Tr}_{A^nB^nR_{AB}^n} \left[\frac{W_{QR_QA^nB^nR_{AB}^n}}{\mathrm{Tr}\left[W_{QR_QA^nB^nR_{AB}^n}\right]} \right] \right) \right] \tag{5.117}$$

$$= \Gamma_{Q\to F} \left(\mathrm{Tr}_{A^nB^n} \left[\frac{W_{QA^nB^n}}{\mathrm{Tr}[W_{QA^nB^n}]} \right] \right) \tag{5.118}$$

$$= \mathrm{Tr}_{A^nB^n} \left[\frac{\Gamma_{Q\to F}(W_{QA^nB^n})}{\mathrm{Tr}[W_{QA^nB^n}]} \right] \tag{5.119}$$

$$= \mathrm{Tr}_{A^n} \left[\frac{\Gamma_{Q\to F}(W_{QA^n})}{\mathrm{Tr}\left[\kappa_{A^n}^{-1}\right]} \right] \tag{5.120}$$

$$= \frac{X_F \cdot \mathrm{Tr}\left[\kappa_{A^n}^{-1/2}(\sigma_A^{\otimes n})^{-1/2}\sigma_A^{\otimes n}(\sigma_A^{\otimes n})^{-1/2}\kappa_{A^n}^{-1/2}\right]}{\mathrm{Tr}\left[\kappa_{A^n}^{-1}\right]} \tag{5.121}$$

$$= X_F. \tag{5.122}$$

This concludes the proof. □

Chapter 6

Discussion

Our work establishes many new results involving de Finetti methods, i.e., de Finetti representation theorems and de Finetti reductions, and their application in quantum information. In particular, we have developed a class of finite constrained de Finetti representation theorems that can be used to generate asymptotically converging hierarchies of semidefinite programs. This is done by fixing the several degrees of freedom, e.g., the underlying Hilbert spaces, operators and linear maps, that parametrize the various representations. For a suitable choice of those parameters, we have generated multiple SDP hierarchies and used them to approximate constrained bilinear optimization programs arising in the context of approximate quantum error correction. We performed numerical simulations to explore the low levels of our hierarchies, analyzing the actual convergence speed of the generated approximations. With the rank loop condition, we have been able to certify the optimality of the low levels for many low-dimensional channels. We have derived a new constrained de Finetti reduction with side information, and we have established a connection between de Finetti reductions and de Finetti representation theorems. In particular, we have shown how to derive de Finetti representation theorems from

de Finetti reductions, and we used our novel technique to obtain a new de Finetti representation theorem for quantum channels with simplified constraints.

6.1 Open Problems

We believe that this line of work leaves some interesting open problems, for future work, as follows.

1. Comparing the bound of our finite constrained de Finetti representation theorem (Theorem 3.6.6) with inequality (3.26), we see that the room for improvement is fairly limited. However, it would be interesting to see if one can improve the square root and the logarithm dependence. Moreover, finding the actual minimal conditions that still guarantee the asymptotic convergence of the SDP hierarchies is still an open question.

2. Given the generality of our framework, one can adapt our techniques to approximate other quantities of interest, generating the desired asymptotically converging SDP hierarchy by fixing the various degrees of freedom in our theorems.

3. On the numerical side, one can explore more complex quantum channels or increase the number of channel repetitions. For our low-dimensional examples, we certified the optimality of the low levels of our hierarchy using the rank loop condition. It would be interesting to see if this behaviour is also observed for higher-dimensional cases and to explore the role of the PPT conditions in the collapse. In order to study more complex settings, one needs to simplify further the optimization programs by taking advantage of the potential symmetries of the particular noise model, as we did for the qubit depolarizing channel in Subsection 4.4.4.

4. Our techniques to generate a de Finetti representation theorem from a starting de Finetti reduction lead to the bound of Theorem 5.3.4, which is suboptimal in n, and grows exponentially fast with k. Moreover, our approach does not seem to be directly applicable in the presence of quantum side information. It is an interesting open question whether it is possible to adapt our methods to improve the dimensional dependence and to handle side information.

Bibliography

[1] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 8.1.*, 2017.

[2] G. M. D. Ariano, P. Perinotti, and M. F. Sacchi. Informationally complete measurements and group representation. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(6):S487–S491, may 2004.

[3] Rotem Arnon-Friedman and Renato Renner. De Finetti reductions for correlations. *Journal of Mathematical Physics*, 56(5):052203, 2015.

[4] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. Non-signaling parallel repetition using de Finetti reductions. *IEEE Transactions on Information Theory*, 62(3):1440–1457, 2016.

[5] S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2006.

[6] Boaz Barak, Fernando G. S. L. Brandao, Aram W. Harrow, Jonathan Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications.

In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, page 307, 2012.

[7] Ivan Bardet, Cécilia Lancien, and Ion Nechita. de finetti reductions for partially exchangeable probability distributions. *arXiv preprint arXiv:1801.05240*, 2018.

[8] Siddharth Barman and Omar Fawzi. Algorithmic aspects of optimal channel coding. *IEEE Transactions on Information Theory*, 64(2):1038, 2018.

[9] H. Barnum, E. Knill, and Michael A. Nielsen. On Quantum Fidelities and Channel Capacities. *IEEE Transactions on Information Theory*, 46:1317–1329, jul 2000.

[10] 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.

[11] Cédric Bény. Conditions for the approximate correction of algebras. In Andrew Childs and Michele Mosca, editors, *Theory of Quantum Computation, Communication, and Cryptography*, pages 66–75, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[12] Mario Berta, Francesco Borderi, and Omar Fawzi. On de Finetti reductions and representation theorems (*Research Notes*). 2021.

[13] Mario Berta, Francesco Borderi, Omar Fawzi, and Volkher B. Scholz. Quantum coding via semidefinite programming. In *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019.

[14] Mario Berta, Francesco Borderi, Omar Fawzi, and Volkher B. Scholz. Semidefinite programming hierarchies for constrained bilinear optimization. *Mathematical Programming*, 2021.

[15] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[16] Fernando G. S. L. Brandao, Matthias Christandl, and Jon Yard. Faithful squashed entanglement. *Communications in Mathematical Physics*, 306(3):80, 2011.

[17] Fernando G. S. L. Brandao and Aram W. Harrow. Product-state approximations to quantum ground states. *Communications in Mathematical Physics*, 342(1):47, 2016.

[18] Fernando G. S. L. Brandao and Aram W Harrow. Quantum de Finetti theorems under local measurements with applications. *Communications in Mathematical Physics*, 353(2):469, 2017.

[19] Carlton M. Caves, Christopher A. Fuchs, and Rüdiger Schack. Unknown quantum states: The quantum de Finetti representation. *Journal of Mathematical Physics*, 43(9):4537–4559, 2002.

[20] 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.

[21] Matthias Christandl, Robert König, Graeme Mitchison, and Renato Renner. One-and-a-half quantum de Finetti theorems. *Communications in Mathematical Physics*, 273(2):473, 2007.

[22] Matthias Christandl, Robert König, and Renato Renner. Postselection technique for quantum channels with applications to quantum cryptography. *Physical Review Letters*, 102(2):020504, 2009.

[23] Matthias Christandl and Ben Toner. Finite de Finetti theorem for conditional probability distributions describing physical theories. *Journal of Mathematical Physics*, 50(4):042104, 2009.

[24] Valerie Coffman, Joydip Kundu, and William K. Wootters. Distributed entanglement. *Phys. Rev. A*, 61:052306, apr 2000.

[25] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, 1991.

[26] Bruno de Finetti. La prévision : ses lois logiques, ses sources subjectives. *Annales de l'institut Henri Poincaré*, 7(1):1, 1937.

[27] Lokenath Debnath. *Introduction to Hilbert spaces with applications.* Elsevier, 2000.

[28] Persi Diaconis. Finite forms of de Finetti's theorem on exchangeability. *Synthese*, 36(2):271–281, 1977.

[29] David DiVincenzo, Peter Shor, and John Smolin. Quantum-channel capacity of very noisy channels. *Physical Review A*, 57(2):830, 1998.

[30] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Complete family of separability criteria. *Physical Review A*, 69(2):022308, 2004.

[31] Runyao Duan and Andreas Winter. No-signalling-assisted zero-error capacity of quantum channels and an information theoretic interpretation of the lovász number. *IEEE Transactions on Information Theory*, 62(2):891, 2016.

[32] Kun Fang and Hamza Fawzi. The sum-of-squares hierarchy on the sphere and applications in quantum information theory. *Mathematical Programming*, 2020.

[33] Omar Fawzi and Renato Renner. Quantum conditional mutual information and approximate Markov chains. *Communications in Mathematical Physics*, 340(2):575, 2015.

[34] M. Fazel, H. Hindi, and S. P. Boyd. Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices. In *Proceedings of the 2003 American Control Conference, 2003.*, volume 3, pages 2156–2162 vol.3, 2003.

[35] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, jan 1968.

[36] Andrew S. Fletcher. *Channel-Adapted Quantum Error Correction*. PhD thesis, Massachusetts Institue of Technology, 2007.

[37] Andrew S. Fletcher, Peter W. Shor, and Moe Z. Win. Optimum quantum error recovery using semidefinite programming. *Physical Review A*, 75(1):012338, 2007.

[38] Christopher A. Fuchs and Rüdiger Schack. *Unknown Quantum States and Operations, a Bayesian View*, page 147. Springer Berlin Heidelberg, 2004.

[39] Christopher A. Fuchs, Rüdiger Schack, and Petra F. Scudo. De Finetti representation theorem for quantum-process tomography. *Physical Review A*, 69(6):062305, 2004.

[40] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, 2008.

[41] Leonid Gurvits. Classical deterministic complexity of edmonds' problem and quantum entanglement. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '03, page 10–19, New York, NY, USA, 2003. Association for Computing Machinery.

[42] Laszlo Gyongyosi, Sandor Imre, and Hung Viet Nguyen. A survey on quantum channel capacities. *IEEE Communications Surveys Tutorials*, 20(2):1149–1205, 2018.

[43] Aram W. Harrow. The Church of the Symmetric Subspace. *arXiv e-prints*, page arXiv:1308.6595, 2013.

[44] Aram W. Harrow, Anand Natarajan, and Xiaodi Wu. Limitations of semidefinite programs for separable states and entangled games. *arXiv preprint arXiv:1612.09306*, 2016.

[45] Masahito Hayashi. Information spectrum approach to second-order coding rate in channel coding. *IEEE Transactions on Information Theory*, 55(11):4947, 2009.

[46] Masahito Hayashi. Universal Coding for Classical-Quantum Channel. *Communications in Mathematical Physics*, 289(3):1087–1098, may 2009.

[47] Masahito Hayashi and Keiji Matsumoto. Quantum universal variable-length source coding. *Physical Review A*, 66(2), aug 2002.

[48] Masahito Hayashi and Marco Tomamichel. Correlation detection and an operational interpretation of the Rényi mutual information. *Journal of Mathematical Physics*, 57(10):102201, 2016.

[49] Tohya Hiroshima. Majorization criterion for distillability of a bipartite quantum state. *Phys. Rev. Lett.*, 91:057902, aug 2003.

[50] M. Horodecki, P. Horodecki, and R. Horodeck. Separability of Mixed States: Necessary and Sufficient Conditions. *Physics Letters A*, 223(1-2):1–8, nov 1996.

[51] Paweł Horodecki, Maciej Lewenstein, Guifré Vidal, and Ignacio Cirac. Operational criterion and constructive checks for the separability of low-rank density matrices. *Physical Review A*, 62(3):032310, 2000.

[52] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, jun 2009.

[53] Hyejung H. Jee, Carlo Sparaciari, Omar Fawzi, and Mario Berta. Quasi-Polynomial Time Algorithms for Free Quantum Games in Bounded Dimension. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*, volume 198 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 82:1–82:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[54] Peter D. Johnson, Jonathan Romero, Jonathan Olson, Yudong Cao, and Alán Aspuru-Guzik. QVECTOR: an algorithm for device-tailored quantum error correction. *arXiv:1711.02249*, 2017.

[55] Nathaniel Johnston. Qetlab: A matlab toolbox for quantum entanglement, version 0.9, 2016.

[56] Eneet Kaur, Siddhartha Das, Mark M. Wilde, and Andreas Winter. Extendibility limits the performance of quantum processors. *Physical Review Letters*, 123:070502, 2019.

[57] Robert Koenig and Graeme Mitchison. A most compendious and facile quantum de Finetti theorem. *Journal of Mathematical Physics*, 50(1):012105, 2009.

[58] Robert Koenig and Renato Renner. A de Finetti representation for finite symmetric quantum states. *Journal of Mathematical Physics*, 46(12):122108, 2005.

[59] Robert König and Renato Renner. A de Finetti representation for finite symmetric quantum states. *Journal of Mathematical Physics*, 46(12):122108, 2005.

[60] Hiroshi Konno. A cutting plane algorithm for solving bilinear programs. *Mathematical Programming*, 11(1):14, 1976.

[61] Robert L. Kosut and Daniel A. Lidar. Quantum error correction via convex optimization. *Quantum Information Processing*, 8(5):443, 2009.

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

[63] Cécilia Lancien and Andreas Winter. Distinguishing multi-partite states by local measurements. *Communications in Mathematical Physics*, 323(2):555, 2013.

[64] Cécilia Lancien and Andreas Winter. Flexible constrained de Finetti reductions and applications. *Journal of Mathematical Physics*, 58(9):092203, 2017.

[65] Debbie Leung and William Matthews. On the power of PPT-preserving and non-signalling codes. *IEEE Transactions on Information Theory*, 61(8):4486, 2015.

[66] Debbie W. Leung, M. A. Nielsen, Isaac L. Chuang, and Yoshihisa Yamamoto. Approximate quantum error correction can lead to better codes. *Physical Review A*, 56(4):2567, 1997.

[67] William Matthews. A linear program for the finite block length converse of Polyanskiy-Poor-Verdú via nonsignaling codes. *IEEE Transactions on Information Theory*, 58(12):7036, 2012.

[68] Miguel Navascués, Masaki Owari, and Martin B. Plenio. Power of symmetric extensions for entanglement detection. *Physical Review A*, 80(5):052306, 2009.

[69] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[70] Lukasz Pankowski, Fernando G. S. L. Brandao, Michael Horodecki, and Graeme Smith. Entanglement distillation by extendible maps. *Quantum Information and Computation*, 13(9–10):751–770, 2013.

[71] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdú. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307, 2010.

[72] Michael Reimpell. *Quantum Information and Convex Optimization*. PhD thesis, TU Braunschweig, 2008.

[73] Michael Reimpell and Reinhard F. Werner. Iterative optimization of quantum error correcting codes. *Physical Review Letters*, 94(8):080501, 2005.

[74] Marc-Olivier Renou, David Trillo, Mirjam Weilenmann, Le Phuc Thinh, Armin Tavakoli, Nicolas Gisin, Antonio Acin, and Miguel Navascues. Quantum physics needs complex numbers. *arXiv preprint arXiv:2101.10873*, 2021.

[75] Erling Størmer. Symmetric states of infinite tensor products of C*-algebras. *Journal of Functional Analysis*, 3(1):48, 1969.

[76] Gilbert Strang. The fundamental theorem of linear algebra. *The American Mathematical Monthly*, 100(9):848–855, 1993.

[77] Soraya Taghavi, Robert L. Kosut, and Daniel A. Lidar. Channel-optimized quantum error correction. *IEEE Transactions on Information Theory*, 56(3):1461, 2010.

[78] Kim-Chuan Toh, Michael J. Todd, and Reha H. Tütüncü. *On the Implementation and Usage of SDPT3 – A Matlab Software Package for Semidefinite-Quadratic-Linear Programming, Version 4.0*, pages 715–754. Springer US, Boston, MA, 2012.

[79] Marco Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory*. PhD thesis, ETH Zurich, mar 2012.

[80] Marco Tomamichel, Mario Berta, and Joseph M. Renes. Quantum coding with finite resources. *Nature Communications*, 7:11419, 2016.

[81] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[82] Joel J. Wallman and Steven T Flammia. Randomized benchmarking with confidence. *New Journal of Physics*, 16(10):103032, 2014.

[83] Xin Wang and Runyao Duan. A semidefinite programming upper bound of quantum capacity. In *Proceedings IEEE ISIT 2016*, page 1690, 2016.

[84] Xin Wang, Kun Fang, and Runyao Duan. Semidefinite programming converse bounds for quantum communication. *IEEE Transactions on Information Theory*, 65(4):2581–2592, 2018.

[85] Xin Wang, Wei Xie, and Runyao Duan. Semidefinite programming strong converse bounds for classical capacity. *IEEE Transactions on Information Theory*, 64:640–653, 2017.

[86] John Watrous. Semidefinite programs for completely bounded norms. *Theory of Computing*, 5(11):217–238, 2009.

[87] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, USA, 1st edition, 2018.

[88] Reinhard F. Werner and Michael M. Wolf. Bell inequalities and entanglement. *Quantum Information and Computation*, 1(3):1, 2001.

[89] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, jun 2013.

[90] Xiao-Dong Yu, Timo Simnacher, H. Chau Nguyen, and Otfried Gühne. Quantum-inspired hierarchy for rank-constrained optimization. *arXiv preprint arXiv:2012.00554*, 2020.

[91] Xiao-Dong Yu, Timo Simnacher, Nikolai Wyderka, H. Chau Nguyen, and Otfried Gühne. A complete hierarchy for the pure state marginal problem in quantum mechanics. *Nature Communications*, 12(1):1012, Feb 2021.