Imperial College London

Department of Physics

# Gaussian Resource Theories and Semidefinite Programming Hierarchies for Quantum Information 

Hyejung Jee

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy at Imperial College London, January 2023

I hereby declare that this thesis and the work reported herein was composed by and originated entirely from me. Information derived from the published and unpublished work of others has been acknowledged in the text and references are given in the list of sources.

Hyejung Jee (2023)

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.


#### Abstract

Determining which quantum tasks we can perform with currently available tools and devices is one of the most important goals of quantum information science today. To achieve this requires careful investigation of the capability of current quantum tools as well as development of classical protocols which can assist quantum tasks and amplify their abilities. In this thesis, we approach this problem through two different topics in quantum information theory: Gaussian resource theories and semidefinite programming hierarchies.

In the first part of this thesis, we examine the possibility of implementing quantum information processing tasks in the Gaussian platform through the eyes of quantum resource theories. Gaussian states and operations are primary tools for the study of continuous-variable quantum information processing due to their easy accessibility and concise mathematical descriptions, although it has been discovered that they are subject to a number of limitations for advanced quantum information processing tasks. We explore the capability of the Gaussian platform further in the first part of this thesis. Firstly, we investigate whether introducing convex structure to the Gaussian framework can circumvent the known no-go theorem of Gaussian resource distillation. Surprisingly, we find that resource distillation becomes possible - albeit in a limited fashion - when convexity is introduced. Then, we consider the quantum resource theory of Gaussian thermal operations when catalysts are allowed, and examine the abilities of catalytic Gaussian thermal operations by characterising all possible state transformations under them.

In the second part of this thesis, we address the problem of characterising quantum correlations via semidefinite programming hierarchies. In particular, we focus on characterising quantum correlations of fixed dimension, which is practically relevant to the field of semi-device-independent quantum information processing. Semidefinite programming is a special type of mathematical optimisation, and it is known that some important but difficult problems in quantum information theory admit semidefinite programming relaxations; these include the characterisation of general quantum correlations in the context of non-locality and the distinction of quantum separable states from entangled states. In this second part, we show how to construct a hierarchy of semidefinite programming relaxations for quantum correlations of fixed dimension and derive analytical bounds on the convergence speed of the hierarchy. For the proof, we make a connection to a variant of quantum separability problem and employ multipartite quantum de Finetti theorems with linear constraints.


## Acknowledgements

Above all, I would like to thank Mario, my supervisor, for all the discussion and support he gave me through my PhD and for giving me freedom to explore and follow my interests. All that I have achieved in last three and a half years would not have been possible without his help. I would also like to thank everyone in our group, Francesco, Navneeth, and Samson, for being wonderful colleagues as well as friends and the nice time we spent together. It is sad that we could not spend more time together in-person due to COVID for the last couple of years. I would like to give very special thanks to Carlo for being a great mentor as a researcher and all the support and encouragement he gave me through my PhD.

I would like to thank Florian and Raul for agreeing to be my examiners and taking their time to read my thesis. I also want to thank all my collaborators, Carlo, Omar, Ben, Alessio, Gerardo, Nick, Hyukjoon, David, and Myungshik, for all scientific discussions we had and the results we achieved together. I have learned a lot from all these collaborations, and it was one of the most valuable experiences I obtained during my PhD .

I would also like to thank people in my CDT. First of all, I want to thank Miranda for all support and help she provided during my PhD as well as the patience she showed me when I was not particularly organised. I also want to thank Ben and Steve for their support as the director and the mentor of the CDT. I would like to give special thanks to our cohort, Oli, George, Simon, Cameron, Sam, Sissi, Will, and Alvise. They helped me a lot to adapt to living in London when I first arrived and since then, always have been great friends for me. We had so much fun, especially in our MRes year. In particular, I am grateful to Cameron and Susie for being wonderful regular pub-mates as well as supportive friends.

I want to thank Jiyoon for having been a great friend with me for last thirteen years and a superb flatmate for most of my PhD. She is the one who convinced me to come to the UK, and it was one of the most wonderful choices I made in my life thanks to her.

I would like to thank my family, my mum, dad and brother, for always loving me and supporting me. I cannot imagine what I would have been like without them. I would also like to thank Jane and Paul, who will be soon my mother- and father-in-law, for taking care of me like parents for the last three and a half years. Finally, I would like to give my deepest thanks to Oli, my fiancé, for making my life so happy and being always next to me whenever I needed him. You are definitely the best thing happened to me during my time at Imperial.

## Author Contributions

This thesis is written based on research which is published in the following three papers:

1. Hyejung H Jee, Carlo Sparaciari, and Mario Berta. Resource distillation in convex Gaussian resource theories. Physical Review A, 103(2):022420, 2021.
2. Benjamin Yadin*, Hyejung H Jee*, ${ }^{1}$ Carlo Sparaciari, Gerardo Adesso, and Alessio Serafini. Catalytic Gaussian thermal operations. Journal of Physics A. 55:325301, 2022.
3. Hyejung H Jee, Carlo Sparaciari, Omar Fawzi, and Mario Berta. Quasi-polynomial time algorithms for free quantum games in bounded dimension. In 48th International Colloquium on Automata, Languages, and Programming (ICALP 2021), volume 198, page 82, 2021.

This thesis consists of two parts. In Part I, we discuss the first two papers within the topic of Gaussian resource theories. Chapter 2 and Chapter 3 provide background theory for Part I and do not contain any novel result. Chapter 4 introduces the results in Paper 1. The work in this chapter is completed by myself collaborated with Dr. Carlo Sparaciari, who was a postdoc in our group at the time, under supervision of my supervisor, Dr. Mario Berta. Chapter 5 discusses the work published in Paper 2, which was a product of a collaboration with people outside of our group, including Dr. Benjamin Yadin at Universität Siegen, Prof. Gerardo Adesso at the University of Nottingham, and Prof. Alessio Serafini at University College London . I would particularly like to acknowledge that the work in Section 5.2 and 5.5 was largely derived by Dr. Benjamin Yadin, who gave me a permission to include the work in this thesis. In Part II, we move on to the topic of semidefinite programming hierarchies. Chapter 6 again provides the background theory for the second part of this thesis and does not present any novel result. Chapter 7 contains the work reported in Paper 3, which was a product of the collaboration of our group with Dr. Omar Fawzi at ENS de Lyon and supported by an Imperial European Partners Fund grant.

[^0]The eternal mystery of the world is its comprehensibility.
... The fact that it is comprehensible is a miracle.

- Albert Einstein (1936)


## Contents

Abstract ..... 3
Acknowledgements ..... 4
Author Contributions ..... 5
1 Introduction ..... 16
1.1 Overview ..... 18
1.2 Notations ..... 22
I Gaussian Resource Theories ..... 24
2 Background theory: quantum continuous variables ..... 25
2.1 Continuous variables quantum systems ..... 26
2.2 Gaussian states ..... 29
2.3 Gaussian operations and measurements ..... 32
2.3.1 Gaussian unitary operations ..... 32
2.3.2 General Gaussian maps ..... 37
2.3.3 Homodyne detection ..... 40
2.4 Normal-mode decomposition of Gaussian states ..... 41
2.5 Examples of pure Gaussian states ..... 46
2.6 Phase-space representation of Gaussian states ..... 48
2.7 Further reading ..... 51
3 Background theory: quantum resource theories ..... 52
3.1 The general structure of quantum resource theories ..... 53
3.2 State convertibility ..... 55
3.2.1 Single-shot case ..... 56
3.2.2 Many-copy case ..... 60
3.3 Quantifying resources: monotones ..... 61
3.4 Multi-resource theories ..... 64
3.5 Further reading ..... 65
4 Convex Gaussian resource theories ..... 66
4.1 Gaussian resource theories ..... 68
4.1.1 No-go theorem for Gaussian resource distillation ..... 69
4.2 Convex Gaussian resource theories ..... 71
4.3 Limitations on resource distillation in convex Gaussian resource theories ..... 73
4.4 Convex distillation protocols ..... 79
4.4.1 Squeezing distillation ..... 79
4.4.2 Entanglement distillation ..... 86
4.5 Summary ..... 90
5 Catalytic Gaussian thermal operations ..... 92
5.1 Preliminaries ..... 95
5.1.1 Gaussian thermal operations ..... 95
5.1.2 Principal mode temperatures and principal mode asymmetries ..... 97
5.2 Gaussian thermal operations without catalysts ..... 100
5.2.1 Multi-mode case ..... 101
5.3 Definition of catalytic Gaussian thermal operations ..... 104
5.4 Single-mode with catalysts: full characterisation ..... 105
5.5 Multi-mode with catalysts ..... 114
5.5.1 Approximate transformations with strong catalysts ..... 114
5.5.2 Necessary majorisation conditions for weak catalysts ..... 116
5.6 Physical implications ..... 117
5.6.1 Concentration of athermal resources ..... 118
5.6.2 General limitations of Gaussian systems ..... 119
5.7 Summary ..... 121
II Semidefinite Programming Hierarchies for Quantum Informa-
tion ..... 123
6 Background theory ..... 124
6.1 Convex sets and convex functions ..... 126
6.2 Semidefinite programming ..... 128
6.3 Semidefinite programming relaxations ..... 133
6.3.1 The Doherty-Parrilo-Spedalieri (DPS) hierarchy ..... 133
6.3.2 The Navascués-Pironio-Acín (NPA) hierarchy ..... 138
6.4 Quantum de Finetti theorems ..... 144
7 Characterising quantum correlations of fixed dimension ..... 146
7.1 Non-local games ..... 150
7.2 Semidefinite programming relaxations for $\omega_{Q(T)}(V, \pi)$ ..... 152
7.2.1 Connection to quantum separability ..... 152
7.2.2 Semidefinite programming relaxations ..... 155
7.3 Convergence of the hierarchy ..... 158
7.3.1 Preliminary: conditional quantum mutual information ..... 159
7.3.2 Tripartite quantum de Finetti theorem with linear constraints ..... 160
7.3.3 Convergence rate of the hierarchy ..... 169
7.4 Combining with the NPA hierarchy ..... 171
7.5 Comparison with previous work ..... 174
7.6 Summary ..... 176
8 Conclusions ..... 178
A Additional information for convex Gaussian resource theories ..... 180
A. 1 Partial selective homodyne measurement ..... 180
A. 2 Monotones for squeezing ..... 182
B Additional information for catalytic Gaussian thermal operations ..... 186
B. 1 Proof of Theorem 12 ..... 186
B. 2 Proof of Theorem 13 ..... 191
B. 3 Proof of Theorem 14 ..... 193
C Additional information for characterising quantum correlations ..... 195
C. 1 General games ..... 195
C. 2 Distortion relative to quantum side information ..... 197
C. 3 Tripartite quantum de Finetti theorem with partial trace constraints ..... 201
C. 4 Constructing NPA matrices with the optimisation variable ..... 207
C. 5 No-signalling value ..... 213
C. 6 Rank-one projective measurements ..... 215
C.6.1 Adapted SDP hierarchies ..... 215
C.6.2 Asymptotic convergence analysis ..... 218

## List of Figures

4.1 The geometrical illustration of the state spaces. (a) In Gaussian resource theories, the set of free states is the interception of $\mathcal{F}_{f}$, the set of free states for $\mathcal{R}_{f}$, and the set of Gaussian states $\mathcal{G}$ which is the set of free states for $\mathcal{R}_{G} ; \mathcal{F}_{f}^{G}=\mathcal{F}_{f} \cap \mathcal{F}_{G}$. (b) In convex Gaussian resource theories, the new set of free states is the convex hull of $\mathcal{F}_{f}^{G}$, i.e., $\mathcal{F}_{f}^{C}=\operatorname{conv}\left(\mathcal{F}_{f}^{G}\right)=\operatorname{conv}\left(\mathcal{F}_{f} \cap \mathcal{F}_{G}\right)$70
4.2 (a) Schematic of the deterministic one-shot squeezing distillation protocol. The system, described by the state $\rho_{\text {in }}$, is fed into a beam splitter with transmissivity $t=\cos ^{2} \theta$ together with a pointer system in the vacuum state. After the interaction, a homodyne measurement is performed on the pointer system, and depending on the outcome the main system is displaced or not. (b) Schematic of the probabilistic multi-copy squeezing distillation protocol. The pointer system is now replaced by another copy of $\rho_{\text {in }}$, providing better performance due to the fact that it is a mixture of squeezed states (resourceful states). The system and the pointer interact via a $50: 50$ beam splitter, and a homodyne measurement is performed on the pointer. If the outcome falls within a desired range, the protocol is successful, otherwise it is aborted. The protocol can be iterated multiple times, with the output system of one iteration being the input state of the next one.
4.3 Numerical simulation of the deterministic one-shot squeezing distillation protocol, for two different beam splitter settings with transmissivity $t=90 \%$ and $65 \%$. (a) The fidelity between the target state $\rho_{\mathrm{sq}}=|0, r\rangle_{1 \mathrm{SQ}}\left(0,\left.r\right|_{1 \mathrm{SQ}}\right.$ and the final state of the protocol as a function of $d / \sigma$, which is the ratio between the amount of the random displacement noise $d$ and the $\hat{x}$-variance $\sigma=e^{-2 r}$. While the low transmissivity $t=65 \%$ allows for correcting the displacement noise when $d / \sigma$ is small, it also degrades squeezing. Thus, it achieves a lower maximum fidelity for the high $d / \sigma$ region than the one achieved by the higher transmissivity $t=90 \%$. (b) The variance in the $\hat{x}$ quadrature for the final state of the protocol is shown as a function of $d / \sigma$. The protocol is regarded successful when the final state has a variance lower than the best free state, the vacuum state. For the simulation, we set the squeezing parameter $r=0.7$ and the probability of random displacement $p=0.5$.
4.4 Numerical simulation of the multi-copy squeezing distillation protocol for different numbers of copies $N$ of the initial state. (a) The fidelity between the target state $\rho_{\mathrm{sq}}=|0, r\rangle_{1 \mathrm{SQ}}\left\langle 0,\left.r\right|_{1 \mathrm{SQ}}\right.$ and the final state of the protocol as a function of $d / \sigma$, where $\sigma=e^{-2 r}$ is the $\hat{x}$-variance of $|0, r\rangle_{1 S Q}$. The success probability $p_{\text {succ }}$ of the protocol is shown in the inset. In most regions, more copies lead to more successful performance even though the success probability decreases significantly as $N$ increases. However, for smaller $d / \sigma$, using multiple copies of the initial state does not help to improve the performance. (b) The variance of the $\hat{x}$ quadrature for the final state as a function of $d / \sigma$. The protocol is regarded successful when the final state has a smaller variance than the best free state, the vacuum state (dotted grey line). The 'optimal squeezing' (dashed-dotted grey line) is the $\hat{x}$-variance of $|0, r\rangle_{1 S Q}$. For the simulation, we set the squeezing parameter $r=0.7$ and the probability of random displacement to $p=0.5$
4.5 The marginal probability distribution obtained from the Wigner function of the state $\rho_{\text {in }}$. In the $\hat{x}_{1}, \hat{x}_{2}$ space, the distribution is represented by two Gaussian blocks squeezed along the $\hat{x}_{1}+\hat{x}_{2}$ quadrature, one centred in the origin and the other displaced by $d$ along the $\hat{x}_{1}$ quadrature. If Bob was to measure his local quadrature $\hat{x}_{1}$ alone, the two Gaussian blocks would overlap and distinguishing them would be more difficult than measuring the non-local quadrature $\hat{x}_{1}+\hat{x}_{2}$, where we see that two distributions have a smaller overlap.
4.6 Schematic of the probabilistic multi-copy entanglement distillation protocol. Alice and Bob share two copies of the initial state and feed their local modes into 50:50 beam splitters. Then, via classical communication they perform joint homodyne measurement on the quadrature $\hat{x}_{+}=\left(\hat{x}_{1}+\hat{x}_{2}\right) / \sqrt{2}$. If the outcome is in the desired region, the protocol is successful, otherwise it is aborted. The protocol can be iterated multiple times with the output state of one iteration being the input state of the next one.
4.7 Numerical simulation of the multi-copy entanglement distillation protocol for different $N$ copies of the initial state. The graph shows the fidelity between the target state $\rho_{\mathrm{sq}}=|\mathbf{0}, r\rangle_{2 \mathrm{SQ}}\left\langle\mathbf{0},\left.r\right|_{2 \mathrm{SQ}}\right.$ and the final state of the protocol as a function of $d / \sigma$. The success probability $p_{\text {succ }}$ of the protocol is shown in the inset. The protocol is regarded as successful when the final state has a larger fidelity than the initial state as well as the the vacuum state (the dotted grey line), which is the best free state, i.e., closest to $|\mathbf{0}, r\rangle_{2 S Q}[1]$. The results are similar to the ones for the multi-copy squeezing distillation protocol. Using more copies leads to a better performance at the expense of the small success probability for larger $d / \sigma$, but not for small $d / \sigma$. For the simulation, we set the squeezing parameter $r=0.7$, and the probability of random displacement to $p=0.5$.
5.1 Illustration of the parameters in Section 5.1.2 (shown for a single mode). We use a common visual representation of Gaussian states in the phase space, where an ellipse indicates the region with high Wigner probability distribution. Note that the quantities indicated are the squares of the lengths. Dashed circle: thermal state at the background temperature $T$, with variance $\nu$ in all quadrature directions. Solid ellipse: arbitrary Gaussian state (with vanishing first moments). $\mu$ is the principal mode temperature and $\alpha$ is the principal mode asymmetry. . .
5.2 Visualisation of possible state transformations under single-mode strong or weak catalytic GTOs, when the initial state of the system is described by the parameters $\mu_{S}, \alpha_{S}$ or the $\mathrm{CM} \sigma_{S}$. The grey regions represent physical regions in both figures. (a) In the state space described by the principal mode temperature $\mu$ and asymmetry $\alpha$. Each point in the physical region (grey) fully describes a covariance matrix of the single-mode system up to passive transformations, and the origin represents the thermal bath. The red line describes all final states achievable by strong catalytic GTOs; it is same as the non-catalytic single-mode case (thermalisation). The blue region represents all possible final states with weak catalytic GTOs. In this case, we can additionally perform inverse squeezing to the system by coupling it to the catalyst. (b) In the state space described by the eigenvalues of the CM of the system, $\lambda_{1}, \lambda_{2}$ with $\lambda_{1} \geq \lambda_{2}$. $\sigma_{T}$ denotes the CM of the thermal bath. Again, the red line represents all possible state transformations with strong catalytic GTOs, and the blue region describes the same with weak catalytic GTOs. Note that the red line is the same as what we could achieve with non-catalytic single-mode GTOs discussed in Proposition 8.
5.3 Lorenz curves (i.e., partial sums) demonstrating the different kinds of orderings encountered here. If the top curve (solid, black) is $\boldsymbol{x}$, then the lower curves from top to bottom satisfy respectively $\boldsymbol{x}^{\prime} \prec \boldsymbol{x}$ (solid, red), $\boldsymbol{x}^{\prime} \prec_{w} \boldsymbol{x}$ (dashed, blue), and $\boldsymbol{x}^{\prime} \leq \boldsymbol{x}$ (dotted, green).
6.1 The schematic of how each ingredient will be used in the next chapter, where we construct a new hierarchy of semidefinite programming relaxations for quantum correlations with fixed dimension.
6.2 Extendible states. Extendible states form a converging hierarchy of outer approximations for the set of separable states. As $n$ becomes larger, the set of $n$-extendible states becomes tighter and closer to the set of separable states. It actually coincides with the separable set when $n$ goes to infinity.
7.1 Non-local game. The referee gives Alice and Bob questions $q_{1} \in Q_{1}$ and $q_{2} \in Q_{2}$ according to the question probability distribution $\pi\left(q_{1}, q_{2}\right)$, and then Alice and Bob give answers $a_{1} \in A_{1}$ and $a_{2} \in A_{2}$ back to the referee depending on the questions they received. The referee decides whether Alice and Bob win or lose according to the rule function $V: A_{1} \times A_{2} \times Q_{1} \times Q_{2} \rightarrow\{0,1\}$, where 0 denotes losing the game, and 1 denotes winning the game. Alice and Bob cannot communicate with each other during the game, but they can agree on a strategy beforehand. We assume that $\left|Q_{1}\right|=\left|Q_{2}\right|=Q$ and $\left|A_{1}\right|=\left|A_{2}\right|=A$.
B. 1 Illustration of the block partitioning of non-increasingly ordered lists $\boldsymbol{z}=(\boldsymbol{x}, \boldsymbol{y})^{\downarrow}$, $\boldsymbol{z}^{\prime}=\left(\boldsymbol{x}^{\prime}, \boldsymbol{y}\right)^{\downarrow}$. In this case, $B_{1}$ is an increasing block, $B_{2}$ is decreasing, and so on alternately. Therefore $x_{1}^{\prime}>x_{1}, x_{2}^{\prime}>x_{2}, x_{3}^{\prime} \leq x_{3}$ and $x_{4}^{\prime} \leq x_{4}$.

## Chapter 1

## Introduction

Quantum information science is the study of information processing based on the fundamental theory of quantum mechanics. It is a highly interdisciplinary field borrowing ideas from many different subjects, including quantum mechanics, computer science, information theory and cryptography. Each of these fields has experienced a major breakthrough at similar times in the last century: the birth of quantum mechanics in the mid-1920s; the discovery of universal Turing machine in 1930s followed by a series of researches on the Church-Turing thesis in 1960-70s which have become the foundation of computational complexity theory in computer science; and Shannon's remarkable research on the concept of information and communication in 1940s. All of these breakthroughs together contributed to the birth of quantum information science and enabled rich discussions and advances in the field.

The main question that has driven people working in quantum information science is whether there exist quantum information processing tasks which can outperform their classical counterparts. As quantum mechanics often exhibits counter-intuitive features or, more precisely, phenomena which are classically impossible, scientists questioned whether applying quantum theory to information processing could open the door for a new technological development resulting in more powerful information processing devices than the ones based on classical mechanics. This idea was further supported by the discoveries of such as Shor's factoring algorithm [2], Grover's search algorithm [3], and quantum key distribution protocols [4] which all demonstrated some quantum advantage over their classical counterparts. However, due to
the difficulty in building high-fidelity and large-scale quantum devices which are feasible to run these algorithms or protocols, it is still far from being able to apply them to any real-world applications or demonstrate the advantage in real life.

Recently, researchers started concentrating on a more present-focused question: which practical quantum tasks can we perform with currently available tools and quantum devices? This idea was emphasised by the emergence of the noisy intermediate-scale quantum (NISQ) era. As Preskill pointed out in his landmark article [5], answering this question not only involves finding quantum information processing protocols implementable in current quantum devices but also a vast amount of research taking different approaches: such as understanding the limitations of current tools for further improvement, developing classical tools which can assist quantum tasks, re-investigating classical protocols to make a fair comparison with quantum counterparts etc.

In this thesis, we also joined this current of research and present several results investigating the ability of currently available tools for quantum information science. We address this point via two different topics, each of which forms a separate part of this thesis: (i) Gaussian resource theories where we examine the capability of Gaussian platforms for quantum tasks and (ii) semidefinite programming hierarchies where we develop a classical algorithm for characterising an important quantum feature, quantum correlations.

In Part I, we explore the possibility of implementing quantum information processing tasks with Gaussian states and Gaussian operations from the perspective of quantum resource theories. Gaussian states and operations are often regarded as easily accessible resources in many experimental set-ups. Indeed, there are a large number of proposed physical platforms for quantum information processing in the continuous-variable regime which can be described by the Gaussian framework; including photonics, optomechanics, atomic ensembles, and trapped ions. The Gaussian framework is also a favourable tool for a theoretical research due to the existence of finite-dimensional representations of the Gaussian dynamics. Despite being a manageable and useful tool for quantum information science, it is known that Gaussian platforms are subject to several limitations when implementing advanced quantum information processing tasks, such as universal quantum computing [6-8], error correction [9] and resource distillations [10].

Nevertheless, non-Gaussian elements are difficult to employ and tend to be more vulnerable to errors/noise, and thus it is beneficial to have more studies on the capabilities of Gaussian platforms for quantum tasks or finding solutions to circumvent those limitations. In this regard, we address in this thesis the tasks of (i) resource distillations and (ii) thermodynamical processes with catalysts using Gaussian platforms.

In Part II, we proceed to the next topic and consider the problem of characterising quantum correlations via semidefinite programming hierarchies. Quantum correlations are an important concept in quantum information science which does not have any counterpart in the classical world. For example, it is known that some quantum correlation can violate locality constraints (known as Bell inequalities) which must be satisfied by any local (classical) correlations, and it was one of the earliest demonstrations of quantum advantage. This feature has been widely exploited to construct useful quantum information processing tasks, notably in quantum communication and cryptography [11]. Such attempts were highlighted by the field of device- or semi-deviceindependent quantum information processing [12], which has significant practical applications within the NISQ era; in device-independent or semi-device-independent quantum information processing protocols, the violation of such constraints is used to certify a desired quantum feature of the device. To construct protocols exploiting these quantum correlations, it is essential to be able to characterise quantum correlations and understand their properties. Characterising quantum correlations is a difficult task, but nevertheless it has been pointed out that borrowing tools from the mathematical optimisation theory in classical computer science could make this particular task more manageable. In such an approach, the set of quantum correlations are approximated via a hierarchy of constraints for a special type of optimisation problems, called semidefinite programming. We follow this approach in this thesis and explore the characterisation of quantum correlations with fixed dimension via semidefinite programming hierarchy.

### 1.1 Overview

Here, we explain the structure of this thesis and provide a short summary for each chapter.

## Part I

The main theme of Part I is Gaussian resource theories. In the first two chapters, Chapter 2 and 3, we provide preliminary information for Part I. In Chapter 2, we revise the background theory of continuous-variable quantum systems. We give the definition of continuous-variable quantum systems and explain why they necessarily admit an infinite-dimensional Hilbert space as their state space. Then, we introduce the Gaussian states, operations and measurements and consider how they make studying continuous-variable quantum systems more tractable by admitting finite-dimensional representations. We also provide useful mathematical tools for studying the Gaussian framework, such as normal-mode decomposition and phase-space representation. In Chapter 3, we introduce another important mathematical framework for Part I, called quantum resource theories. We first look at the definition and the general structure of quantum resource theories. Then, we introduce the state-convertibility problem, which is one of the basic quantum information processing tasks that can be studied within the framework of quantum resource theories. We explain that the state-convertibility problem can provide some insights into how we can quantify the amount of resource present in an arbitrary quantum state. This idea leads to the introduction of resource monotones, key resource measures in quantum resource theories. Finally, we discuss how to generalise the aforementioned framework to the case when there exists more than one relevant physical resource.

In Chapter 4, which is the first chapter containing novel results, we explore the task of resource distillation using Gaussian states and operations. Gaussian resource theories, the main theme of Part I, are quantum resource theories with the additional physical restriction that one is only allowed to use Gaussian states and operations. It is known that resource distillation is in general impossible in such Gaussian resource theories [10]. To overcome this limitation, we examine whether the non-convex structure of the Gaussian framework contributes to this impossibility by investigating convex Gaussian resource theories, which are Gaussian resource theories equipped with classical randomness and conditional Gaussian operations, i.e., Gaussian resource theories that have an additional convex structure. This approach is well-motivated from the practical point of view, since classical randomness and conditional Gaussian operations are more accessible in a laboratory than non-Gaussian elements. Surprisingly, we find that
resource distillation becomes possible for convex Gaussian resource theories albeit in a limited fashion; the impossibility of resource distillation is replaced by a limitation on the amount of distillable resources in convex Gaussian resource theories. We fully characterise this limitation by studying a new resource measure for convex Gaussian states. We also provide explicit examples of resource distillation protocols for squeezing and entanglement, showing that the derived upper bound on the amount of distillable resource can be saturated in a special case - which implies that the derived upper bound is tight.

In Chapter 5, we continue to look at another instance of Gaussian resource theories: the Gaussian resource theory of quantum thermodynamics. Quantum thermodynamics in continuousvariable quantum systems through the framework of quantum resource theories is not a wellstudied topic despite a vast amount of results for the discrete-variable case. Especially, the role played by catalysts, which have a significant influence in discrete-variable quantum thermodynamics, remains unclear in continuous-variable quantum thermodynamics. Our main goal in this chapter is to characterise the capability of catalytic Gaussian thermal operations, which are defined as all energy-preserving Gaussian operations in the presence of catalysts. We carry out this goal by characterising all possible state transformations for Gaussian states under Gaussian thermal operations when catalysts are allowed. We introduce two different models for catalysts: strong catalysts, where final correlations with the system are forbidden, and weak catalysts, where final correlations with the system are allowed. We show that strong catalysts are useless for the single-mode case in the sense that they do not expand the set of achievable states for a given initial state via Gaussian thermal operations. Then, we prove that weak catalysts are instead capable of reaching a larger region of final states and derive exact conditions for state transformation for the single-mode case. We also derive necessary conditions for the general multipartite case for strong catalysts and approximate transformation as well as for weak catalysts. As the notion of catalysts generalises any thermal machine that operates in a cycle, we discuss the implications of our results for thermal devices operating with Gaussian elements.

## Part II

In Part II, we address the topic of semidefinite programming hierarchies for quantum information science. Chapter 6, the first chapter of Part II, contains necessary background materials for Part II without any novel results. In Chapter 6, we explain two main concepts used in Part II, semidefinite programming relaxations and quantum de Finetti theorems. We start with studying some mathematical background including convex sets and convex functions and move on to the topic of semidefinite programming, a special sub-class of mathematical optimisation problems. We show that one can construct semidefinite programming relaxations for some general optimisation problem that is difficult to solve and provide two widely-used hierarchies of semidefinite programming relaxations in quantum information theory: the Doherty-ParriloSpedalieri (DPS) and Navascués-Pironio-Acín (NPA) hierarchy. Then, we continue to the second topic, quantum de Finetti theorems, and give a short introduction including the connection to the DPS hierarchy.

Chapter 7 contains our novel results on characterising quantum correlations of fixed dimension via a hierarchy of semidefinite programming relaxations. As we discussed before, characterising quantum correlations of fixed dimension is related to the field of semi-device-independent quantum information processing where the dimensionality of quantum devices plays an important role. In this chapter, we show how to construct a converging semidefinite programming hierarchy of outer approximations for the set of quantum correlations of fixed dimension. More specifically, we employ non-local games, which can be seen as a generalisation of the Bell scenario, for formulating correlations and derive a hierarchy of constraints in the form of semidefinite programming, which can approximate quantum correlations of fixed dimension. We derive analytical bounds on the convergence speed of the hierarchy using improved multipartite quantum de Finetti theorems with additional linear constraints, which we also derive in the same chapter. As a way to improve the relaxations, we provide a procedure to combine our semidefinite programming relaxations with the dimension-agnostic NPA hierarchy. We compare the performance of our semidefinite programming relaxations with previous work in the literature before summarising the chapter.

### 1.2 Notations

In this section, we establish some basic notations used throughout this thesis.
A mathematical set is denoted with bold symbol or capital letter, and when there are a finite number of elements in the set, says $n$ elements, we often denote it for instance by $\left\{x_{i}\right\}_{i=1}^{n}$, or if the number of elements is clear in the context we simply write $\left\{x_{i}\right\}_{i}$. The cardinality of a given set $X$ is denoted by $|X|$. We use the following notations for different ranges of real numbers: $\mathbb{R}_{\geq 0}$ for all non-negative real numbers, $\mathbb{R}_{\leq 0}$ for all non-positive real numbers, and likewise for similar cases. A vector is also denoted by bold symbol; $\mathbf{x}=\left(x_{1}, \ldots, x_{n}\right)^{\mathrm{T}}$.

Let $\mathcal{H}$ be a either finite- or infinite-dimensional Hilbert space over $\mathbb{C}$ with inner product $\langle\cdot, \cdot\rangle$. For a linear operator $A$ on $\mathcal{H}$, the complex conjugate $A^{*}$ of $A$ is the operator whose matrix representation has entries that are complex conjugates of the ones for $A$. The transpose $A^{\mathrm{T}}$ of $A$ is the operator whose matrix representation is obtained by transposing the matrix representation of $A$. Then, the adjoint $A^{\dagger}$ of $A$ is the operator obtained by taking the complex conjugate and transpose of $A$. A self-adjoint (for an infinite-dimensional $\mathcal{H}$ ) or hermitian (for a finite-dimensional $\mathcal{H})$ operator is a linear operator $A$ such that $\langle x, A y\rangle=\langle A x, y\rangle$ for all $x, y \in \mathcal{H}$. We denote the set of all bounded linear operators acting on $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$. A self-joint or hermitian operator $A$ is called positive semidefinite and denoted by $A \geq 0$ (or $A \succeq 0$ ) if $\operatorname{Tr}[A|\psi\rangle\langle\psi|] \geq 0$ for all $\psi \in \mathcal{H}$, and called positive definite and denoted by $A>0($ or $A \succ 0)$ if $\operatorname{Tr}[A|\psi\rangle\langle\psi|]>0$ for all $\psi \in \mathcal{H}$. A density operator $\rho$ is a positive semidefinite operator with trace one. We denote the set of all density operators on $\mathcal{H}$ by $\mathcal{D}(\mathcal{H})$. We denote the identity operator acting on $d$-dimensional $\mathcal{H}$ by $\mathbb{1}_{d}$. When it is obvious in the context, we often omit the dimension subscription and use the notation $\mathbb{1}$.

A map $\Gamma: \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ is called positive if $\Gamma(A) \geq 0$ for all $A \geq 0$ and completely-positive $(\mathrm{CP})$ if $\left(\mathcal{I}_{k} \otimes \Gamma\right)(A) \geq 0$ for any positive semidefinite operator $A$ on $\mathbb{C}^{k} \otimes \mathcal{H}$ and any $k \in \mathbb{N}$, where $\mathcal{I}_{k}$ denotes the identity map on $\mathbb{C}^{k}$. We denote the set of all CP maps from $\mathcal{B}\left(\mathcal{H}_{\text {in }}\right)$ to $\mathcal{B}\left(\mathcal{H}_{\text {out }}\right)$ by $\mathcal{C}\left(\mathcal{H}_{\text {in }}, \mathcal{H}_{\text {out }}\right)$. A map $\Gamma: \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ is called trace-preserving (TP) when it preserves the trace of the argument. Then, a quantum map (also called quantum channel or quantum operation) is a completely-positive and trace-preserving (CPTP) map acting on $\mathcal{D}(\mathcal{H})$. The
most general quantum measurements are described by positive operator valued measurements (POVMs). A POVM with $k$ measurement outcomes is defined with a set of positive semidefinite operators $\left\{M_{i}\right\}_{i=1}^{k}$ such that $\sum_{i=1}^{k} M_{i}=\mathbb{1}$.

Given a vector $\psi$ in $d$-dimensional $\mathcal{H}$, we define the $\ell_{p}$ norm of the vector $\psi$ for $1 \leq p<\infty$ as $\|\psi\|_{p}=\left(\sum_{i=1}^{d}\left|\psi_{i}\right|^{p}\right)^{1 / p}$, where $\left\{\psi_{i}\right\}_{i}$ are the components of the vector $\psi$. We can also define a similar concept, the Schattern p-norms, for a linear operator $A$ on $\mathcal{H}$ with any real number $p \geq 1$ as $\|A\|_{p}=\left(\operatorname{Tr}\left[\left(A^{\dagger} A\right)^{p / 2}\right]\right)^{1 / p}$. We particularly call the case with $p=1$ the trace norm. The fidelity between two positive semidefinite operators $A$ and $B$ is defined as $F(A, B)=\operatorname{Tr}[\sqrt{\sqrt{A} B \sqrt{A}}]$.

We say a matrix $U$ acts by congruence on $A$ when $U A U^{\mathrm{T}}$. When we consider continuousvariable quantum systems, we denote operators acting on the infinite-dimensional Hilbert space with hat, for example $\hat{x}$, with the exception of density operators $\rho$, and denote operators acting on a finite-dimensional space such as the phase space without hat, for example $x$, following the notation in [13].

For a tensor-product of Hilbert spaces involving multiple copies of identical quantum registers, for example $\bigotimes_{i=1}^{k} \mathcal{H}_{A}$, we use the concise notation $\mathcal{H}_{A^{\otimes k}}$ or $\mathcal{H}_{A^{k}}$. We do likewise for the subscription in quantum states; for example, $\rho_{A^{\otimes k}}$ or $\rho_{A^{k}}$.

## Part I

## Gaussian Resource Theories

## Chapter 2

## Background theory: quantum

## continuous variables

Continuous variable quantum systems, along with discrete variable quantum systems (e.g., qubits), are one of the two forms for realising quantum information processing tasks and play an important role in quantum information science. A vast amount of physical platforms proposed for quantum information processing are described by continuous variable quantum systems; some famous examples are quantised modes of bosonic systems including electromagnetic fields, motional degrees of freedom in trapped ions and atomic ensembles, and various platforms within the field of optomechanics. In particular, the theory of quantum continuous variables is highly related to quantum communication and cryptography which have propagating photons as their natural platform [14-18]. However, the study of dynamics in continuous variable quantum systems is often restricted to Gaussian states and operations since infinite-dimensional representations of continuous variable quantum systems in general make their dynamics intractable. The Gaussian resources, namely Gaussian states, operations, and measurements, are thus a primary tool for continuous variable quantum information science.

In this section, we revise the fundamental theory of continuous variable quantum systems focusing on the Gaussian framework. The main purpose of this section is to understand how the Gaussian framework resolves the intractability problem of working with continuous variable quantum systems. In Section 2.1, we define continuous variable quantum systems together
with the important canonical commutation relation and canonical observables. We explain why any continuous variable quantum system admits an infinite dimensional representation. Then, we start introducing the Gaussian framework with the definition of Gaussian states in Section 2.2, where we show that any Gaussian state can be fully characterised by a finite number of elements. We continue to introduce Gaussian operations and measurements in Section 2.3, where we provide their concise descriptions in terms of statistical moments and list some notable examples. In Section 2.4, we present a useful mathematical tool in the Gaussian framework called the normal-mode decomposition of Gaussian states, which is largely exploited in later chapters (especially in Chapter 5). We present some famous examples of pure Gaussian states in Section 2.5 and introduce the phase-space representation of Gaussian states in Section 2.6. This chapter is written based on the following textbook and review papers on quantum continuous variables: [13, 19, 20].

## Notations

Unless stated otherwise, we always assume the natural unit $\hbar=1$, where $\hbar$ is the reduced Plank constant. In this section, we mainly consider a quantum system associated with an infinite dimensional Hilbert space $\mathcal{H}$. We emphasise again that operators on the Hilbert space are denoted with hat, $\hat{x}$, with exception of quantum states (e.g., $\rho$ ), and operators on finitedimensional spaces such as the phase space are denoted without hat, e.g., $x$. Unless we indicate with $\rho_{G}$, most definitions, formulas, and concepts presented in this chapter are valid for general continuous variable quantum states.

### 2.1 Continuous variables quantum systems

Continuous variable (CV) quantum systems naturally arise when we consider canonical quantisation procedure, one way of quantising a classical theory. Consider a pair of canonical conjugate quantities, for example position and momentum. We can define corresponding quantum mechanical operators $\hat{x}$ and $\hat{p}$ for these variables, so-called quadrature operators, satisfying the
following canonical commutation relation (CCR)

$$
\begin{equation*}
[\hat{x}, \hat{p}]=i \hbar \hat{\mathbb{1}}, \tag{2.1}
\end{equation*}
$$

which replaces Poisson brackets in classical Hamiltonian dynamics. CV quantum systems are quantum systems which obey this CCR. It is not difficult to observe that the CCR cannot be satisfied for operators $\hat{x}$ and $\hat{p}$ acting on finite-dimensional spaces; taking the trace on both sides of Eq. (2.1) always leads to a contradiction in finite dimensions unless $\hbar$ vanishes, since the left hand side (LHS) becomes zero due to the linearity and the cyclic property of the trace. Thus, quantum CV systems inevitably have infinite-dimensional Hilbert spaces as their state spaces.

It is well known that there exist infinite-dimensional representations for CV quantum systems. Let us consider the space of square-integrable functions on the real space, $L^{2}(\mathbb{R})$, which forms a Hilbert space. Then, the canonical operators $\hat{x}$ and $\hat{p}$ can be defined as operators acting on $L^{2}(\mathbb{R})$ in the following way:

$$
\hat{x}|\psi\rangle=x|\psi\rangle, \quad \hat{p}|\psi\rangle=-i \hbar \frac{\mathrm{~d}}{\mathrm{~d} x}|\psi\rangle \quad \forall|\psi\rangle \in L^{2}(\mathbb{R}) .
$$

We can check that these two operators satisfy the CCR.

$$
[\hat{x}, \hat{p}]|\psi\rangle=\hat{x} \hat{p}|\psi\rangle-\hat{p} \hat{x}|\psi\rangle=-i \hbar \hat{x} \frac{\mathrm{~d}|\psi\rangle}{\mathrm{d} x}-\hat{p} x|\psi\rangle=-i \hbar x \frac{\mathrm{~d}|\psi\rangle}{\mathrm{d} x}+i \hbar|\psi\rangle+i \hbar x \frac{\mathrm{~d}|\psi\rangle}{\mathrm{d} x}=i \hbar|\psi\rangle .
$$

CV quantum systems necessarily adopt this infinite-dimensional Hilbert space $L^{2}(\mathbb{R})$ as their state spaces. Any element in $L^{2}(\mathbb{R})$, which now represents a state of CV quantum systems, can be decomposed into the eigenstates of $\hat{x}$ or $\hat{p} .^{1}$ Their eigenvalues form continuous spectra covering the whole real numbers, and the terminology 'quantum continuous variables' came from this factor.

We can find another expedient basis set for a CV quantum system. We can define the ladder

[^1]operators in terms of the canonical operators as
$$
\hat{a}=\frac{\hat{x}+i \hat{p}}{\sqrt{2}}, \quad \hat{a}^{\dagger}=\frac{\hat{x}-i \hat{p}}{\sqrt{2}},
$$
where natural unit is adopted, i.e., $\hbar=1$. In the context of quantum harmonic oscillators, $\hat{a}$ and $\hat{a}^{\dagger}$ corresponds to the annihilation and creation of a quantum of energy respectively, and in bosonic quantum field theory, they describe the creation and destruction of particles. The CCR in Eq. (2.1) can be also expressed with the ladder operators;
$$
\left[\hat{a}, \hat{a}^{\dagger}\right]=\mathbb{1}
$$

We now define a Hermitian operator $\hat{n}:=\hat{a}^{\dagger} \hat{a}$ in terms of the ladder operators and call it the number operator. The state space of a CV quantum system with single degree of freedom, which we denote $\mathcal{H}$, is then spanned by the eigenstates of this number operator, $\{|n\rangle\}_{n=0}^{\infty}$ with $\hat{n}|n\rangle=n|n\rangle$, and this countable basis is called the Fock or number state basis. Thus, the Hilbert space $\mathcal{H}$ is infinite-dimensional and separable ${ }^{2}$ while it admits continuous representations with the operators $\hat{x}$ or $\hat{p}$.

When $N$ degrees of freedom are considered, we have $N$ pairs of canonical operators $\left\{\hat{x}_{i}\right\}_{i=1}^{N}$ and $\left\{\hat{p}_{i}\right\}_{i=1}^{N}$ satisfying $\left[\hat{x}_{i}, \hat{p}_{j}\right]=i \delta_{i j} \mathbb{1} \forall i, j$, and each pair is connected to the ladder operators of the corresponding degree of freedom via $\hat{a}_{i}=\left(\hat{x}_{i}+i \hat{p}_{i}\right) / \sqrt{2}$ and $\hat{a}_{i}^{\dagger}=\left(\hat{x}_{i}-i \hat{p}_{i}\right) / \sqrt{2}$. In CV quantum systems, we often call each degree of freedom a mode borrowing the name from the area of quantum optics. The total state space of an $N$-mode CV quantum system is then a tensor-product Hilbert space $\bigotimes_{i=1}^{N} \mathcal{H}_{i}$ where each $\mathcal{H}_{i}$ is the single-mode Hilbert space. We define a vector of canonical operators as

$$
\hat{\mathbf{r}}=\left(\hat{x}_{1}, \hat{p}_{1}, \ldots, \hat{x}_{n}, \hat{p}_{n}\right)^{\mathrm{T}}
$$

[^2]and the symplectic form matrix
\[

\Omega=\bigoplus_{i=1}^{N} \Omega_{1} \quad with \quad \Omega_{1}=\left($$
\begin{array}{cc}
0 & 1  \tag{2.2}\\
-1 & 0
\end{array}
$$\right)
\]

Then, the CCRs for an $N$-mode CV quantum system can be written as $\left[\hat{r}_{i}, \hat{r}_{j}\right]=i \Omega_{i j}$ where we now omitted the identity operator on the right-hand side (RHS) and adapted the ' $c$-numbers'3 notation. More concisely, we can express the CCRs as

$$
\left[\hat{\mathbf{r}}, \hat{\mathbf{r}}^{\mathrm{T}}\right]=i \Omega
$$

where the commutator between row and column vectors is applied in the same way as an outer product, and the relation is satisfied element-wise with the omitted identity operator on the RHS. Note that the symplectic form matrix is a real orthogonal matrix, i.e., $\Omega^{\mathrm{T}} \Omega=\mathbb{1}$, and also satisfies $\Omega^{\mathrm{T}}=-\Omega$ and $\Omega^{2}=1$. We will sometimes denote by $\Omega_{n}$ the $2 n \times 2 n$-dimensional symplectic form matrix for $n$ modes. Without subscription, it denotes the symplectic form matrix with right dimension in the context.

In summary, even with finite degrees of freedom, a CV quantum system admits an infinitedimensional Hilbert space, and it requires an infinite number of elements to represent its states. This makes the analysis of their dynamics intractable, and in general, it is very cumbersome to work with CV quantum systems analytically. In the next section, we will discuss how to circumvent this problem.

### 2.2 Gaussian states

One way to make CV quantum systems more tractable is to be restricted to Gaussian states. In this section, we will revise the definition of Gaussian states and how they can avoid the problem of infinite-dimensional representations. We will largely follow the approach described in [13] and introduce Gaussian states as ground and thermal states of second-order Hamiltonians, which

[^3]provides a plenty of intuitions proven to be useful later in this part of the thesis.
Second-order Hamiltonians are the Hamiltonians which include canonical operators at most to the power of two, i.e., which can be written as a polynomial of order two in terms of the canonical operators. In terms of the vector of the canonical operators $\hat{\mathbf{r}}$, a generic second-order Hamiltonian $\hat{H}$ of an $N$-mode CV quantum system can be parametrised as
\[

$$
\begin{equation*}
\hat{H}=\frac{1}{2} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}+\hat{\mathbf{r}}^{\mathrm{T}} \overline{\mathbf{r}} \tag{2.3}
\end{equation*}
$$

\]

with a $2 N \times 2 N$ symmetric matrix $H$ and a $2 N$-dimensional real vector $\overline{\mathbf{r}} . H$ is sometimes called the Hamiltonian matrix and should not be confused with the Hamiltonian operator $\hat{H}$ itself; $H$ does not act on the state space $L^{2}\left(\mathbb{R}^{N}\right)$. Eq. (2.3) includes all strictly quadratic Hamiltonians as well as all linear Hamiltonian. Such second-order Hamiltonians are common among experimental set-ups in physics, e.g., quantum optics and optomechanics, and emerge whenever we have negligible higher-order terms. Henceforth, we will only consider the case when $H>0$ for our convenience. Positive definite Hamiltonian matrices imply that Hamiltonian operators are bounded from below, which is necessary for stable systems.

We now define Gaussian states as all ground and thermal states of second-order Hamiltonians with $H>0$.

Definition 1. (Gaussian states) Any Gaussian state $\rho_{G}$ can be written as either

$$
\begin{equation*}
\rho_{G}=\frac{e^{-\beta \hat{H}}}{\operatorname{Tr}\left[e^{-\beta \hat{H}}\right]} \tag{2.4}
\end{equation*}
$$

or

$$
\begin{equation*}
\rho_{G}=\lim _{\beta \rightarrow \infty} \frac{e^{-\beta \hat{H}}}{\operatorname{Tr}\left[e^{-\beta \hat{H}}\right]} \tag{2.5}
\end{equation*}
$$

with $\beta \in \mathbb{R}^{+}$and some second-order Hamiltonian $\hat{H}$ as defined in Eq. (2.3).

Note that Gaussian states of the form in Eq. (2.4) are necessarily mixed states, and any pure Gaussian state should have the form in Eq. (2.5), which describes the ground state of some
second-order Hamiltonian $\hat{H}$.
Definition 1 shows that Gaussian states can be parametrised by the Hamiltonian $\hat{H}$ and the (inverse-temperature like) parameter $\beta$. Following the parametrisation of $\hat{H}$ in Eq. (2.3), Gaussian states are thus fully characterised by the Hamiltonian matrix $H$ and the vector $\overline{\mathbf{r}}$ describing $\hat{H}$, and the parameter $\beta$. This is a crucial point worth stressing as the Hamiltonian matrix $H$ and the vector $\overline{\mathbf{r}}$ are finite-dimensional while $\hat{H}$ as well as the state $\rho_{G}$ are operators on the infinite-dimensional Hilbert space. This property of Gaussian states makes the study of their dynamics tractable and manifests the benefit of being restricted to Gaussian states.

The fact that we can represent Gaussian states with finite number of elements is reminiscent of classical Gaussian distributions, which are fully characterised by their first and second statistical moments. Let us strengthen this connection further. For an $N$-mode Gaussian state $\rho_{G}$, we define a vector of first moments of the canonical operators as

$$
\overline{\mathbf{r}}:=\langle\hat{\mathbf{r}}\rangle_{\rho_{G}}=\left(\left\langle\hat{x}_{1}\right\rangle_{\rho_{G}},\left\langle\hat{p}_{1}\right\rangle_{\rho_{G}}, \ldots,\left\langle\hat{x}_{N}\right\rangle_{\rho_{G}},\left\langle\hat{p}_{N}\right\rangle_{\rho_{G}}\right)^{\mathrm{T}}
$$

where $\langle\hat{A}\rangle_{\rho}=\operatorname{Tr}[\hat{A} \rho]$ is the expectation value of the operator $\hat{A}$ for the state $\rho$. We also define the symmetrised version of the second moments, so called the covariance matrix (CM), as

$$
\begin{equation*}
\sigma=\operatorname{Tr}\left[\left\{(\hat{\mathbf{r}}-\overline{\mathbf{r}}),(\hat{\mathbf{r}}-\overline{\mathbf{r}})^{\mathrm{T}}\right\} \rho_{G}\right], \tag{2.6}
\end{equation*}
$$

where the anti-commutator between a column vector and a row vector acts like an outer product. More specifically, Eq. (2.6) is equivalent to $\sigma_{i j}=\left\langle\left\{\left(\hat{\mathbf{r}}_{i}-\overline{\mathbf{r}}_{i}\right),\left(\hat{\mathbf{r}}_{j}-\overline{\mathbf{r}}_{j}\right)\right\}\right\rangle_{\rho_{G}} \forall i, j$. It is not a coincidence that the vector of first moments is denoted by the same symbol as the parameter $\overline{\mathbf{r}}$ of the Hamiltonian operator $\hat{H}$. It can be shown that the vector of first moments of a Gaussian state $\rho_{G}$ always coincides with the parameter $\overline{\mathbf{r}}$ of the Hamiltonian operator $\hat{H}$ which describes the state $\rho_{G}$ via Definition 1. We will later see that the CM of a Gaussian state $\rho_{G}$ also contains all information about the Hamiltonian matrix $H$ of the operator $\hat{H}$ and the parameter $\beta$ describing $\rho_{G}$ (see Section 2.4). As a result, any Gaussian state can be fully characterised by a pair of its $\overline{\mathbf{r}}$ and $\sigma$. This fact will become more clear when we talk about the phase-space representation of Gaussian states in Section 2.6.

By definition, all covariance matrices of an $N$-mode CV quantum system are $2 N \times 2 N$, real and symmetric matrices. However, not all real symmetric matrices are CMs of quantum states. The CM of any quantum state (not only Gaussian states) must satisfy the following uncertainty principle:

$$
\begin{equation*}
\sigma+i \Omega \geq 0 \tag{2.7}
\end{equation*}
$$

where 0 denotes the null matrix with appropriate dimension. Note that taking the transpose of the LHS of Eq. (2.7) should not change the positivity of the matrix so that $\sigma-i \Omega \geq 0$. One can show that the relation in Eq. (2.7) implies that $\sigma>0$, i.e., all CMs must be positive definite.

The important point to remember from this section is that, despite the underlying infinitedimensional Hilbert space, any Gaussian state can be characterised by a finite number of elements, which provides a remarkable platform for exploring CV quantum systems. We will look at some notable examples of Gaussian states later in this chapter.

### 2.3 Gaussian operations and measurements

Gaussian operations in general refer to completely positive (CP) channels which map Gaussian states into Gaussian states. There are two types of Gaussian operations, Gaussian unitary operations and (more general) Gaussian non-unitary maps. In this chapter, we look at these two different categories of Gaussian operations as well as some important Gaussian measurement.

### 2.3.1 Gaussian unitary operations

All Gaussian unitary operations are the ones generated by a second-order Hamiltonian as described in Eq. (2.3); $e^{i \hat{H} t}=e^{i\left(\frac{1}{2} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}+\hat{\mathbf{r}}^{\mathrm{T}} \mathbf{r}\right) t}$ for some time duration $t$. As all Gaussian states can be characterised by a finite-dimensional representation - by the first and second statistical moments of $\hat{\mathbf{r}}$ - it is convenient to find out the effect of the unitary operation $\left.e^{i\left(\frac{1}{2} \mathbf{r}^{\mathbf{r}} H \hat{\mathbf{r}}\right.}+\hat{\mathbf{r}}^{\mathrm{T}}\right) t t$ on this finite-dimensional representation.

Let us consider an $N$-mode CV quantum system. We firstly introduce an important operator
called a Weyl operator defined as

$$
\hat{D}_{\mathbf{r}}:=e^{i \mathbf{r}^{\mathrm{T}} \Omega \hat{\mathbf{r}}}
$$

for some vector $\mathbf{r} \in \mathbb{R}^{2 N}$. Note that $\hat{D}_{\mathbf{r}}^{\dagger}=\hat{D}_{-\mathbf{r}}$. The action of Weyl operators on the vector of canonical operators $\hat{\mathbf{r}}$ is

$$
\begin{equation*}
\hat{D}_{-\mathbf{r}} \hat{\mathbf{r}} \hat{D}_{\mathbf{r}}=\hat{\mathbf{r}}-\mathbf{r}, \tag{2.8}
\end{equation*}
$$

where the Weyl operator acts on each element of the vector $\hat{\mathbf{r}}^{4}$ Thus, Weyl operators simply shift the expectation values of the first statistical moments of the canonical operators. Due to this property, Weyl operators are also known as displacement operators in the quantum optics context. Using Weyl operators, we can derive an alternative form of a generic second-order Hamiltonian as

$$
\begin{equation*}
\hat{H}=\frac{1}{2}(\hat{\mathbf{r}}-\mathbf{r})^{\mathrm{T}} H(\hat{\mathbf{r}}-\mathbf{r})=\frac{1}{2} \hat{D}_{-\mathbf{r}} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}} \hat{D}_{\mathbf{r}} \tag{2.9}
\end{equation*}
$$

This new form is equivalent to Eq. (2.3) up to redefining $\mathbf{r}$ and an irrelevant constant. Eq. (2.9) shows that any linear term in a second-order Hamiltonian shifts the vector of the canonical operators $\hat{\mathbf{r}}$ by the constant vector $\mathbf{r}$ and can be represented by an action of displacement operators on $\hat{\mathbf{r}}$.

Now, let us consider effects of the purely quadratic term $\frac{1}{2} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}$ in a second-order Hamiltonian upon the vector of the canonical operators $\hat{\mathbf{r}}$. By definition, any valid unitary transformation must preserve the CCRs when applied on the canonical operators. It can be shown that, in the finite-dimensional picture, this condition is equivalent to preserve the symplectic form matrix $\Omega$. The group of such transformations is well known from classical Hamiltonian mechanics; the real symplectic group in dimension $2 N, S p(2 N)$, defined as the set of transformations such that for

[^4]all $S \in S p(2 N)$
$$
S \Omega S^{\mathrm{T}}=\Omega
$$

Then, using the Heisenberg evolution of $\hat{\mathbf{r}}$ under the purely quadratic Hamiltonian $\frac{1}{2} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}$, we can derive that

$$
\begin{equation*}
\hat{U} \hat{\mathbf{r}} \hat{U}^{\dagger}=S_{H} \hat{\mathbf{r}}, \tag{2.10}
\end{equation*}
$$

where $\hat{U}=e^{i \frac{1}{\mathbf{r}} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}}$ and $S_{H}=e^{\Omega H} \in S p(2 N)$ [13, Section 3.2.2]. Here, we include the time $t$ in the Hamiltonian matrix $H$ for simplifying the notation. Note that on the LHS, the unitary operation $\hat{U}$ acts on each entry of $\hat{\mathbf{r}}$, but on the RHS, the $2 N \times 2 N$ symplectic matrix $S_{H}$ acts on the $2 N$-dimensional vector $\hat{\mathbf{r}}$. Eq. (2.10) shows that applying the unitary operations induced by $\frac{1}{2} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}$ to the vector of the canonical operators $\hat{\mathbf{r}}$ has the same effect as applying the symplectic transformation $S_{H}=e^{\Omega H}$ to $\hat{\mathbf{r}}$ in the finite-dimensional picture. This manifests that not only Gaussian states but also Gaussian unitary dynamics admit finite-dimensional representations.

The above observations allow us to derive the full description of Gaussian unitary dynamics in the level of the statistical moments of $\hat{\mathbf{r}}$. A linear term in the second-order Hamiltonian generates a displacement $\hat{D}_{\mathbf{r}}$ for some $\mathbf{r} \in \mathbb{R}^{2 N}$ and maps the first moments $\overline{\mathbf{r}}$ and the CM $\sigma$ of the initial state $\rho$ into

$$
\begin{aligned}
& \overline{\mathbf{r}} \mapsto \overline{\mathbf{r}}+\mathbf{r}, \\
& \sigma \mapsto \sigma .
\end{aligned}
$$

Similarly, a purely quadratic term $\frac{1}{2} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}$ in the second-order Hamiltonian generates a unitary operation $e^{\frac{i}{2} \hat{\mathbf{r}}^{\mathrm{T}} H \hat{\mathbf{r}}}$ and maps $\overline{\mathbf{r}}$ and $\sigma$ of the initial state $\rho$ into

$$
\begin{aligned}
& \overline{\mathbf{r}} \mapsto S_{H} \overline{\mathbf{r}}, \\
& \sigma \mapsto S_{H} \sigma S_{H}^{\mathrm{T}},
\end{aligned}
$$

where $S_{H}=e^{\Omega H} \in \operatorname{Sp}(2 N)$.
Now, let us look into more details about the symplectic transformations generated by purely quadratic Hamiltonians and what they can do. Symplectic transformations admit many useful decompositions, and we will introduce the most relevant one here.

Proposition 1. (Singular value decomposition of symplectic transformations) Any symplectic matrix $S \in S p(2 N)$ admits the following decomposition:

$$
S=O_{1} Z O_{2}
$$

where $O_{1}$ and $O_{2}$ are orthogonal symplectic matrices, i.e., $O_{1}, O_{2} \in S p(2 N) \cap O(2 N)^{5}$, and

$$
Z=\bigoplus_{i=1}^{N}\left(\begin{array}{cc}
z_{i} & 0 \\
0 & z_{i}^{-1}
\end{array}\right)
$$

The proof of Proposition 1 is outside of the scope of this thesis, and we refer to [13, Appendix B.2] for the proof. We look into this decomposition more closely in the next two sections.

## Passive transformations: phase shifters and beam splitters

Let us consider the free Hamiltonian $\hat{H}_{\text {free }}$ which is an equally weighted sum of squares of all canonical operators. It is not difficult to figure out that the Hamiltonian matrix $H$ of $\hat{H}_{\text {free }}$ is the $2 N \times 2 N$ identity matrix. The transformations $O_{1}$ and $O_{2}$ in Proposition 1 preserve this identity matrix as they are orthogonal, i.e., $O O^{\mathrm{T}}=O^{\mathrm{T}} O=\mathbb{1}$. Physically, this means that the transformations $O_{1}, O_{2}$ preserve the energy of the system. Thus, such transformations are often called passive or energy-preserving transformations. There are two types of passive transformations: phase shifters and beam splitters, and any passive symplectic transformation can be decomposed into these two types of transformations.

A phase shifter is defined as the following symplectic transformation on the vector of $\left(\hat{x}_{i}, \hat{p}_{i}\right)^{\mathrm{T}}$

[^5]of the corresponding mode $i$ :
\[

S_{\mathrm{PS}}(\phi):=\left($$
\begin{array}{cc}
\cos \phi & \sin \phi \\
-\sin \phi & \cos \phi
\end{array}
$$\right)
\]

for some angle $\phi$. If we consider the ladder operator $\hat{a}_{i}=\left(\hat{x}_{i}+i \hat{p}_{i}\right) / \sqrt{2}$ and its expectation value $\alpha_{i}=\left\langle\hat{a}_{i}\right\rangle$, a phase shifter introduces an additional phase $e^{i \phi}$ to $\alpha_{i}$, which corresponds to a rotation in the single-mode phase space associated with $\hat{x}_{i}$ and $\hat{p}_{i}$. In other words, when it's applied to a coherent state $|\alpha\rangle$, which will be defined in Section 2.5, it maps the state into another coherent state $\left|e^{i \phi} \alpha\right\rangle$.

A beam splitter acts on two different modes and is defined with

$$
S_{\mathrm{BS}}(\phi):=\left(\begin{array}{cccc}
\cos \phi & 0 & \sin \phi & 0 \\
0 & \cos \phi & 0 & \sin \phi \\
-\sin \phi & 0 & \cos \phi & 0 \\
0 & -\sin \phi & 0 & \cos \phi
\end{array}\right)
$$

for some angle $\phi$. The value $\cos ^{2} \phi$ is often called the transmittivity of the beam splitter. Again, if we consider the expectation values of the corresponding ladder operators, $\alpha_{i}$ and $\alpha_{j}$, a beam splitter rotates $\alpha_{i}$ and $\alpha_{j}$ of $i$-th and $j$-th modes to which it is applied.

## Squeezing transformations

Now, let us move on to the other type of transformations, $Z$ in Proposition 1. They are called squeezing transformations and in general require additional energy to be realised.

A single-mode squeezing transformations is characterised by the following symplectic transformation

$$
S_{1 \mathrm{SQ}}(r):=\left(\begin{array}{cc}
e^{-r} & 0  \tag{2.11}\\
0 & e^{r}
\end{array}\right)
$$

for some $r \in(-\infty, \infty)$. It contracts a canonical variable and expand the conjugate one. Single-
mode squeezing transformations are generated by the Hamiltonian $\hat{x} \hat{p}=i\left(\hat{a}^{\dagger 2}-\hat{a}^{2}\right) .{ }^{6}$ This Hamiltonian does not commute with the free Hamiltonian, and thus it requires energy from an external source to be realised. A two-mode squeezing transformation is defined with following symplectic matrix:

$$
S_{2 \mathrm{SQ}}(r):=\left(\begin{array}{cccc}
\cosh r & 0 & \sinh r & 0  \tag{2.12}\\
0 & \cosh r & 0 & -\sinh r \\
\sinh r & 0 & \cosh r & 0 \\
0 & -\sinh r & 0 & \cosh r
\end{array}\right)
$$

for some $r \in(-\infty, \infty)$. This transformation is generated by the Hamiltonian of the form $i\left(\hat{a}_{1}^{\dagger} \hat{a}_{2}^{\dagger}-\hat{a}_{1} \hat{a}_{2}\right)$ and requires an external energy source to be realised as well. When applied to the vacuum state, these single- and two-mode squeezing transformations generate the singleand two-mode squeezed vacuum states respectively, which will be defined in Section 2.5.

### 2.3.2 General Gaussian maps

## Tensor products and partial trace

Appending another Gaussian state and discarding a subsystem both preserve the Gaussianity of the initial state and thus are valid Gaussian operations. In the Hilbert space description, they correspond to tensor product and partial trace, respectively. Here, we give their effects on the finite dimensional description with the statistical moments $\overline{\mathbf{r}}$ and $\sigma$. Firstly, consider a two Gaussian states $\rho_{A}$ and $\rho_{B}$ which is characterised by $\left(\overline{\mathbf{r}}_{A}, \sigma_{A}\right)$ and $\left(\overline{\mathbf{r}}_{B}, \sigma_{B}\right)$, respectively. The

[^6]tensor-product of two states, $\rho_{A} \otimes \rho_{B}$, is characterised by
\[

$$
\begin{aligned}
& \overline{\mathbf{r}}=\binom{\overline{\mathbf{r}}_{A}}{\overline{\mathbf{r}}_{B}}=\overline{\mathbf{r}}_{A} \oplus \overline{\mathbf{r}}_{B} \\
& \sigma=\left(\begin{array}{cc}
\sigma_{A} & 0 \\
0 & \sigma_{B}
\end{array}\right)=\sigma_{A} \oplus \sigma_{B}
\end{aligned}
$$
\]

Thus, the tensor-product structure of the Hilbert space is converted into the direct-sum structure in the finite dimensional description. Now, let us consider a Gaussian state $\rho_{A B}$ which consists of two subsystems $A$ with $m$ modes and $B$ with $n$ modes and is described by

$$
\overline{\mathbf{r}}=\binom{\overline{\mathbf{r}}_{A}}{\overline{\mathbf{r}}_{B}} \quad \text { and } \quad \sigma=\left(\begin{array}{cc}
\sigma_{A} & \sigma_{A B}  \tag{2.13}\\
\sigma_{A B}^{\mathrm{T}} & \sigma_{B}
\end{array}\right)
$$

Then, the state $\operatorname{Tr}_{B}\left[\rho_{A B}\right]\left(\operatorname{Tr}_{A}\left[\rho_{A B}\right]\right)$ is a Gaussian state described by $\overline{\mathbf{r}}_{A}\left(\overline{\mathbf{r}}_{B}\right)$ and $\sigma_{A}\left(\sigma_{B}\right)$. It is worth remarking that the off-diagonal term $\sigma_{A B}$ in a CM with two subsystems of the form in Eq. (2.13) can be related to the correlation between two subsystems.

## Gaussian completely positive maps

Interaction with external environment, regardless of whether it was intended or not, often causes non-unitary dynamics on the quantum system. The dynamics of open CV quantum systems therefore allows us to define more general Gaussian operations, a class of Gaussian CP maps or channels, which physically describe decoherence or noisy channels. There are two types of Gaussian CP maps: trace-preserving maps and non-trace-preserving maps.

An $N$-mode Gaussian CPTP channel (or map) is a linear map $\mathcal{E}^{G}$ acting on quantum states, i.e., $\mathcal{E}^{G}: \rho \rightarrow \mathcal{E}^{G}(\rho) \in \mathcal{D}\left(\mathcal{H}^{N}\right)$. A simple way to understand this channel is to use the Stinespring dilation [21]. Consider a system in the input state $\rho$ and an ancillary mode in the vacuum state $|0\rangle_{E}$. The most general Gaussian CPTP channel can be represented by applying a Gaussian
unitary operation $\hat{U}_{G}$ to the system and the ancilla and then tracing out the ancillary mode:

$$
\mathcal{E}^{G}(\rho)=\operatorname{Tr}_{E}\left[\hat{U}_{G}\left(\rho \otimes|0\rangle\left\langle\left. 0\right|_{E}\right) \hat{U}_{G}^{\dagger}\right] .\right.
$$

It is shown that an ancilla with less than $2 N$ modes is sufficient to realise any $N$-mode Gaussian CPTP map [22, 23]. The action of a generic Gaussian CPTP channel on an $N$-mode Gaussian state $\rho_{G}$ with statistical moments $\overline{\mathbf{r}}$ and $\sigma$ can be fully described by two $2 N \times 2 N$ real matrices $X$ and $Y$ and a real vector $\mathbf{r} \in \mathbb{R}^{2 N}$ as follows [24]:

$$
\begin{align*}
\overline{\mathbf{r}} & \mapsto X \overline{\mathbf{r}}+\mathbf{r}  \tag{2.14}\\
\sigma & \mapsto X \sigma X^{\mathrm{T}}+Y .
\end{align*}
$$

The matrices $X$ and $Y$ must satisfy that

$$
Y+i \Omega \geq i X \Omega X^{\mathrm{T}}
$$

which ensures that the final CM satisfies the uncertainty relation Eq. (2.7).
For the case of non-trace preserving maps, we introduce one notable example in Section 2.3.3, and for more general information, we refer to references [13, Section 5.4] since we do not explicitly make use of them in this thesis.

## Classical mixing

The set of Gaussian states is not convex; in general, classically mixing Gaussian states does not preserve the Gaussianity of the initial state, and is a cheap way to create a non-Gaussian state. A convex combination of Gaussian states is Gaussian only if they are weighted by a Gaussian distribution. This Gaussian classical mixing, where the weight function is Gaussian, is one simple example of Gaussian CPTP channels.

Gaussian classical mixing $\mathcal{E}_{\text {mix }}^{G}$ is the Gaussian CPTP-map defined in Eq. (2.14) with $X=\mathbb{1}$, $Y \geq 0$ and $\mathbf{r}=\mathbf{0}$. When $Y>0$, the $N$-mode Gaussian classical mixing can be simply expressed
as

$$
\mathcal{E}_{\text {mix }}^{G}(\rho)=\iint_{\mathbb{R}^{2 N}} \mathrm{~d} \mathbf{r}^{\prime} \frac{e^{-\mathbf{r}^{\prime T} Y^{-1} \mathbf{r}^{\prime}}}{\pi^{N} \sqrt{\operatorname{det}(Y)}} \hat{D}_{\mathbf{r}^{\prime}} \rho \hat{D}_{\mathbf{r}^{\prime}}^{\dagger}
$$

### 2.3.3 Homodyne detection

Gaussian measurements refer to all POVMs which preserve the Gaussianity of the initial state after the detection, and form one example of non-trace-preserving Gaussian maps. In this section, we will briefly revise homodyne detection, which is one of the Gaussian measurements in the general-dyne category. This detection scheme will frequently appear in Part I.

A homodyne detection scheme performs the projective measurement of a canonical operator. In other words, for a given state $\rho$, it aims at measuring the general quadrature operator $\hat{x}_{\phi}=\cos \phi \hat{x}+\sin \phi \hat{p}$ with outcome probability

$$
p\left(x_{\phi}\right)=\left\langle x_{\phi}\right| \rho_{G}\left|x_{\phi}\right\rangle,
$$

where $\left\{\left|x_{\phi}\right\rangle\right\}$ are the eigenvectors of $\hat{x}_{\phi}$.
It is well known that homodyne detection can be implemented with a clever optical trick using a high intensity laser [20]. Let us assume that we want to measure the quadrature operator $\hat{x}_{\phi}=\cos \phi \hat{x}+\sin \phi \hat{p}$. In terms of the ladder operators, it can be written as

$$
\hat{x}_{\phi}=\frac{e^{-i \phi} \hat{a}+e^{i \phi} \hat{a}^{\dagger}}{\sqrt{2}} .
$$

We mix the initial state $\rho$ with a laser in a strong coherent state $|\alpha\rangle$ with $\alpha \gg 11^{7}$ at a 50:50 beam splitter. Then, the two output-modes of the beam splitter are described by

$$
\hat{a}_{1}=\frac{\hat{a}+\hat{b}}{\sqrt{2}}, \quad \hat{a}_{2}=\frac{\hat{a}-\hat{b}}{\sqrt{2}},
$$

where $\hat{b}$ denotes the ladder operator of the laser mode. What we want to do is to subtract the

[^7]detected intensities of these two output-modes. Let us see how this process results in quadrature measurements of $\hat{x}_{\phi}$. What we observe is the following operator:
$$
{ }_{b}\langle\alpha|\left(\frac{1}{2}(\hat{a}+\hat{b})^{\dagger}(\hat{a}+\hat{b})-\frac{1}{2}(\hat{a}-\hat{b})^{\dagger}(\hat{a}-\hat{b})\right)|\alpha\rangle_{b}={ }_{b}\langle\alpha|\left(\hat{a}^{\dagger} \hat{b}+\hat{a} \hat{b}^{\dagger}\right)|\alpha\rangle_{b} \approx \alpha \hat{a}^{\dagger}+\alpha^{*} \hat{a} .
$$

By adjusting the phase of the laser mode $|\alpha\rangle_{b}$ to $\phi$, we can measure the target quadrature operator up to a constant factor:

$$
\sqrt{2}|\alpha|\left(\frac{e^{-i \phi} \hat{a}+e^{i \phi} \hat{a}^{\dagger}}{\sqrt{2}}\right)=\sqrt{2}|\alpha| \hat{x}_{\phi} .
$$

### 2.4 Normal-mode decomposition of Gaussian states

In this section, we will introduce a useful mathematical tool called the normal-mode decomposition, which leads to some significant observation about Gaussian states.

According to Williamson's theorem [25], the following proposition holds:

Proposition 2. (Normal-mode decomposition) For any $2 N \times 2 N$ real positive definite matrix A, there exists a symplectic transformation $S \in S p(2 N)$ such that

$$
\begin{equation*}
S A S^{T}=D \quad \text { with } \quad D=\bigoplus_{i=1}^{N} d_{i} \mathbb{1}_{2} \tag{2.15}
\end{equation*}
$$

with $d_{i}>0 \forall i$, where $\mathbb{1}_{2}$ denotes the identity matrix of dimension 2.

Proof. Since $A$ and $D$ are real positive definite matrices, there exist the matrices $A^{-1 / 2}$ and $D^{1 / 2}$. Then, the real matrix $S$ satisfying Eq. (2.15) may be constructed by $S=D^{1 / 2} Q A^{-1 / 2}$ for some $Q \in O(2 N)$. The remaining thing to show is that there exists an orthogonal matrix $Q \in O(2 N)$ such that $S$ is symplectic, i.e.,

$$
S \Omega S^{\mathrm{T}}=\left(D^{1 / 2} Q A^{-1 / 2}\right) \Omega\left(D^{1 / 2} Q A^{-1 / 2}\right)^{\mathrm{T}}=D^{1 / 2} Q A^{-1 / 2} \Omega A^{-1 / 2} Q^{\mathrm{T}} D^{1 / 2}=\Omega
$$

where we used the facts that $D=D^{\mathrm{T}}$ and $A=A^{\mathrm{T}}$. Let us focus on the matrix $\Omega^{\prime} \equiv A^{-1 / 2} \Omega A^{-1 / 2}$ in the middle. $\Omega^{\prime}$ is a real, anti-symmetric matrix: (i) It is real since both $A$ and $\Omega$ are real
matrices, and (ii) it is anti-symmetric as

$$
\left(A^{-1 / 2} \Omega A^{-1 / 2}\right)^{\mathrm{T}}=A^{-1 / 2} \Omega^{\mathrm{T}} A^{-1 / 2}=-A^{-1 / 2} \Omega A^{-1 / 2}
$$

It is known that real anti-symmetric matrices are normal matrices and can be therefore diagonalised via the spectral theorem. More specifically, there exists a unitary matrix which transforms such matrix into a diagonal matrix with pure imaginary eigenvalues of the form $\left(i d_{1},-i d_{1}, \ldots, i d_{N},-i d_{N}\right)$. Due to this property, it is also known that there exists an orthogonal matrix $Q \in O(2 N)$ which brings a real anti-symmetric matrix $\Omega^{\prime}$ to the block-diagonal form $\bigoplus_{i=1}^{N} \frac{1}{d_{i}} \Omega_{1}$, where $\Omega_{1}=\left(\begin{array}{cc}0 & 1 \\ -1 & 0\end{array}\right)$ is the $2 \times 2$ block in the symplectic form defined in Eq. (2.2). If we choose such orthogonal matrix $Q$ for $\Omega^{\prime}$, then we have

$$
S \Omega S^{\mathrm{T}}=D^{1 / 2} Q\left(A^{-1 / 2} \Omega A^{-1 / 2}\right) Q^{\mathrm{T}} D^{1 / 2}=D^{1 / 2}\left(Q \Omega^{\prime} Q^{\mathrm{T}}\right) D^{1 / 2}=D^{1 / 2}\left(\bigoplus_{i=1}^{N} \frac{1}{d_{i}} \Omega_{1}\right) D^{1 / 2}
$$

If we choose $D=\bigoplus_{i=1}^{N} d_{i} \mathbb{1}_{2}$, then we finally obtain

$$
S \Omega S^{\mathrm{T}}=D^{1 / 2}\left(\bigoplus_{i=1}^{N} \frac{1}{d_{i}} \Omega_{1}\right) D^{1 / 2}=\bigoplus_{i=1}^{N} d_{i} d_{i}^{-1} \Omega_{1}=\Omega
$$

Note that $\left\{d_{i}\right\}_{i}$ are all strictly positive since $A$ is positive definite.

The diagonal elements $\left\{d_{i}\right\}_{i}$ are often called symplectic eigenvalues, and the normal mode decomposition is also known as symplectic diagonalisation. Let us now consider a Gaussian state $\rho_{G}$ characterised by a number $\beta>0$ and Hamiltonian $\hat{H}$ with Hamiltonian matrix $H>0$ and $\overline{\mathbf{r}}$ according to the parametrisation in Definition 1. We apply Proposition 2 to the Hamiltonian matrix $H$ as follows:

$$
\begin{equation*}
H=S^{-1}\left(\bigoplus_{i=1}^{N} \omega_{i} \mathbb{1}_{2}\right)\left(S^{-1}\right)^{\mathrm{T}} \tag{2.16}
\end{equation*}
$$

for some $S \in S p(2 N)$. Note that $S$ is the symplectic transformation that maps $H$ into the normal form as described in Proposition 2, and its inverse always exists since $H>0$. In the

Hilbert space description, Eq. (2.16) leads to

$$
\begin{align*}
\hat{\mathbf{r}}^{T} H \hat{\mathbf{r}} & =\hat{\mathbf{r}}^{T} S^{-1}\left(\bigoplus_{i=1}^{N} \omega_{i} \mathbb{1}_{2}\right)\left(S^{-1}\right)^{\mathrm{T}} \hat{\mathbf{r}}=\hat{U}_{S^{-1}} \hat{\mathbf{r}}^{\mathrm{T}}\left(\bigoplus_{i=1}^{N} \omega_{i} \mathbb{1}_{2}\right) \hat{\mathbf{r}} \hat{U}_{S^{-1}}^{\dagger}  \tag{2.17}\\
& =\hat{U}_{S^{-1}}\left(\sum_{i=1}^{N} \omega_{i}\left(\hat{x}_{i}^{2}+\hat{p}_{i}^{2}\right)\right) \hat{U}_{S^{-1}}^{\dagger},
\end{align*}
$$

where $\hat{U}_{S^{-1}}$ is the unitary operations defined as $\hat{U}_{S^{-1}} \hat{\mathbf{r}} \hat{U}_{S^{-1}}^{\dagger}=S^{-1} \hat{\mathbf{r}}$. It is worth remarking what Eqs. (2.16)-(2.17) imply. The symplectic transformation $S^{-1}$ defines a new canonical degrees of freedom since it preserves the CCRs. In this new set of degrees of freedom, all modes are dynamically decoupled and oscillate like independent free harmonic oscillators. Thus, every purely quadratic Hamiltonian with $H>0$ is unitarily equivalent to a set of decoupled free harmonic oscillators. The diagonal elements $\left\{\omega_{i}\right\}_{i}$ are the symplectic eigenvalues of $H$ and correspond to eigenfrequencies of the normal modes.

For our convenience, we will denote the Hamiltonian for a simple harmonic oscillator with eigenfrequency $\omega$ as

$$
\hat{H}_{\omega}=\frac{\omega}{2}\left(\hat{x}^{2}+\hat{p}^{2}\right) .
$$

Then, using the expression in Eq. (2.9), a generic second-order Hamiltonian can be written as

$$
\hat{H}=\frac{1}{2}(\hat{\mathbf{r}}-\overline{\mathbf{r}})^{\mathrm{T}} H(\hat{\mathbf{r}}-\overline{\mathbf{r}})=\hat{D}_{-\overline{\mathbf{r}}} \hat{U}_{S^{-1}}\left(\sum_{i=1}^{N} \hat{H}_{\omega_{i}}\right) \hat{U}_{S^{-1}}^{\dagger} \hat{D}_{\overline{\mathbf{r}}}
$$

Applying this result to the expression of the Gaussian state $\rho_{G}$ in Definition 1 obtains

$$
\begin{equation*}
\rho_{G}=\hat{D}_{-\overline{\mathbf{r}}} \hat{U}_{S^{-1}} \frac{\left(\bigotimes_{i=1}^{N} e^{-\beta \hat{H}_{\omega_{i}}}\right)}{\prod_{i=1}^{N} \operatorname{Tr}\left[e^{-\beta \hat{H}_{\omega_{i}}}\right]} \hat{U}_{S^{-1}}^{\dagger} \hat{D}_{\overline{\mathbf{r}}} \tag{2.18}
\end{equation*}
$$

which shows that every Gaussian state is also unitarily equivalent to a tensor-product of singlemode thermal states, where each mode has the eigenfrequency $\omega_{i}$. For the pure-state case, the
expression becomes

$$
\lim _{\beta \rightarrow \infty} \rho_{G}=\hat{D}_{-\overline{\mathbf{r}}} \hat{U}_{S^{-1}}\left(\bigotimes_{i=1}^{N}|0\rangle\langle 0|\right) \hat{U}_{S^{-1}}^{\dagger} \hat{D}_{\overline{\mathbf{r}}}
$$

Therefore, all pure Gaussian states are the consequences of applying Gaussian unitary operations to the vacuum state.

Now, let us look at the consequence of Eq. (2.18) for the CM of the Gaussian state $\rho_{G}$. Using the Fock basis and the fact that the Hamiltonian of a simple harmonic oscillator can be rephrased as $\hat{H}_{\omega}=\omega\left(\hat{n}+\frac{1}{2}\right)$, we can derive the representation of $\rho_{G}$ in the Fock basis as follows:

$$
\begin{equation*}
\rho_{G}=\left(\prod_{i=1}^{N}\left(1-e^{-\beta \omega_{i}}\right)\right) \hat{D}_{-\overline{\mathbf{r}}} \hat{U}_{S^{-1}}\left(\bigotimes_{i=1}^{N}\left(\sum_{n=0}^{\infty} e^{-\beta \omega_{i} n}|n\rangle\left\langle\left. n\right|_{j}\right)\right) \hat{U}_{S^{-1}}^{\dagger} \hat{D}_{\overline{\mathbf{r}}}\right. \tag{2.19}
\end{equation*}
$$

where $|n\rangle_{j}$ refers to the corresponding number state of $j$-th mode. Calculating the CM of $\rho_{G}$ using Eq. (2.19) obtains

$$
\begin{align*}
\sigma & =\operatorname{Tr}\left[\left\{(\hat{\mathbf{r}}-\overline{\mathbf{r}}),(\hat{\mathbf{r}}-\overline{\mathbf{r}})^{\mathrm{T}}\right\} \rho_{G}\right] \\
& =\left(\prod_{i=1}^{N}\left(1-e^{-\beta \omega_{i}}\right)\right) \operatorname{Tr}\left[\hat{U}_{S^{-1}}^{\dagger} \hat{D}_{\overline{\mathbf{r}}}\left\{(\hat{\mathbf{r}}-\overline{\mathbf{r}}),(\hat{\mathbf{r}}-\overline{\mathbf{r}})^{\mathrm{T}}\right\} \hat{D}_{-\overline{\mathbf{r}}} \hat{U}_{S^{-1}}\left(\bigotimes_{i=1}^{N}\left(\sum_{n=0}^{\infty} e^{-\beta \omega_{i} n}|n\rangle\left\langle\left. n\right|_{i}\right)\right)\right]\right. \\
& =\left(\prod_{i=1}^{N}\left(1-e^{-\beta \omega_{i}}\right)\right) \operatorname{Tr}\left[\hat{U}_{S^{-1}}^{\dagger}\left\{\hat{\mathbf{r}}, \hat{\mathbf{r}}^{\mathrm{T}}\right\} \hat{U}_{S^{-1}}\left(\bigotimes_{i=1}^{N}\left(\sum_{n=0}^{\infty} e^{-\beta \omega_{i} n}|n\rangle\left\langle\left. n\right|_{j}\right)\right)\right]\right. \\
& =\left(\prod_{i=1}^{N}\left(1-e^{-\beta \omega_{i}}\right)\right)\left(S^{-1}\right)^{\mathrm{T}} \operatorname{Tr}\left[\left\{\hat{\mathbf{r}}, \hat{\mathbf{r}}^{\mathrm{T}}\right\}\left(\bigotimes_{i=1}^{N}\left(\sum_{n=0}^{\infty} e^{-\beta \omega_{i} n}|n\rangle\left\langle\left. n\right|_{j}\right)\right)\right] S^{-1}\right. \\
& =\left(S^{-1}\right)^{\mathrm{T}}\left(\bigoplus_{i=1}^{N} \frac{1+e^{-\beta \omega_{i}}}{1-e^{-\beta \omega_{i}}} \mathbb{1}_{2}\right) S^{-1} \equiv\left(S^{-1}\right)^{\mathrm{T}}\left(\bigoplus_{i=1}^{N} \nu_{i} \mathbb{1}_{2}\right) S^{-1}, \tag{2.20}
\end{align*}
$$

where in the third line we exploited the property of displacement operators in Eq. (2.8), and in the last line we used the properties of the Fock basis to evaluate the expectation values of different combinations of the canonical operators ${ }^{8}$ and defined $\nu_{i} \equiv \frac{1+e^{-\beta \omega_{i}}}{1-e^{-\beta \omega_{i}}} \geq 1$ in the last line.

[^8]The last expression in Eq. (2.20) is referred to as the normal-mode decomposition of the CM, and $\left\{\nu_{i}\right\}_{i}$ are called the symplectic eigenvalues of the CM. The normal-mode decomposition of the CM indicates that the CM contains all the information about $S^{-1}$ and the symplectic eigenvalues of the Hamiltonian matrix $H$ generating the Gaussian state, which is consistent with the fact that any Gaussian state can be fully characterised by its $\overline{\mathbf{r}}$ and the CM $\sigma$.

The normal-mode decomposition of the CM in Eq. (2.20) can give more implication regarding the uncertainty relation $\sigma+i \Omega \geq 0$ introduced in Eq. (2.7). As any symplectic transformation preserves the symplectic form matrix $\Omega$, applying the symplectic transformation, which maps $\sigma$ into its normal form, to the both sides of the uncertainty relation obtains

$$
\left(\bigoplus_{j=1}^{N} \nu_{j} \mathbb{1}_{2}\right)+i \Omega=\bigoplus_{j=1}^{N}\left(\begin{array}{cc}
\nu_{j} & i \\
-i & \nu_{j}
\end{array}\right) \geq 0
$$

which is equivalent to

$$
\begin{equation*}
\nu_{j} \geq 1 \quad \forall j \tag{2.21}
\end{equation*}
$$

Also, if we look at a single-mode case, we can infer that

$$
\operatorname{det}\left(\begin{array}{cc}
\nu_{j} & i \\
-i & \nu_{j}
\end{array}\right) \geq 0 \quad \Longleftrightarrow \quad \nu_{j}^{2} \geq 1 \quad \Longleftrightarrow \quad \operatorname{det}\left(\nu_{j} \mathbb{1}_{2}\right) \geq 1
$$

Since (i) the determinant is a multiplicative map and (ii) all symplectic matrices have the determinant $1, \operatorname{det}\left(S\left(\nu_{j} \mathbb{1}_{2}\right) S^{T}\right) \geq 1$ still holds for a generic $\mathrm{CM} \sigma=S\left(\nu_{j} \mathbb{1}_{2}\right) S^{T}$. Then, we have $\sigma_{11} \sigma_{22}-\sigma_{12} \sigma_{21} \geq 1$ where $\sigma_{i j}$ denotes the $(i, j)$ th-element of $\sigma$, and it implies $\sigma_{11} \sigma_{22} \geq 1 .{ }^{9}$ If we remind ourselves that the diagonal elements of the CM is related to two times the variance of the corresponding canonical operator, $2 \Delta \hat{r}_{i}^{2}=2\left(\left\langle\hat{r}_{i}^{2}\right\rangle-\left\langle\hat{r}_{i}\right\rangle^{2}\right)$, the above relation recovers the usual Heisenberg uncertainty relation for a pair of the canonical operators:

$$
\begin{equation*}
\Delta \hat{x}^{2} \Delta \hat{p}^{2} \geq \frac{1}{4} \tag{2.22}
\end{equation*}
$$

[^9]
### 2.5 Examples of pure Gaussian states

In this section, we provide a few notable examples of pure Gaussian states frequently appeared in the literature.

## Coherent states

The coherent state $|\alpha\rangle$ of a single-mode CV quantum system is the eigenvector of the ladder operator $\hat{a}=(\hat{x}+i \hat{p}) / \sqrt{2}$ with eigenvalue $\alpha=(x+i p) / \sqrt{2}$ :

$$
\hat{a}|\alpha\rangle=\alpha|\alpha\rangle .
$$

It can be generated from the vacuum state $|0\rangle$ via the Displacement operator $\hat{D}_{-\mathbf{r}}=\hat{D}_{\alpha}=$ $e^{\alpha \hat{a}^{\dagger}-\alpha^{*} \hat{a}}$ where $\mathbf{r}=(x, p)^{\mathrm{T}}$ such that $\alpha=(x+i p) / \sqrt{2}$ :

$$
\hat{D}_{\alpha}|0\rangle=|\alpha\rangle .
$$

It is also well-known that coherent states can be expressed in the Fock basis as

$$
|\alpha\rangle=e^{-\frac{\mid \alpha \alpha^{2}}{2}} \sum_{n=0}^{\infty} \frac{\alpha^{n}}{\sqrt{n!}}|n\rangle .
$$

The coherent states form a non-orthogonal but complete set for the Hilbert space, in a sense that

$$
\frac{1}{\pi} \int_{\mathbb{C}}|\alpha\rangle\langle\alpha| \mathrm{d}^{2} \alpha=\hat{\mathbb{1}}
$$

In terms of the finite dimensional description, a coherent state $|\alpha\rangle$ is characterised by the CM $\sigma=\mathbb{1}$ and the first moment vector $\overline{\mathbf{r}}=\sqrt{2}(\operatorname{Re}(\alpha), \operatorname{Im}(\alpha))^{\mathrm{T}}$. Note that coherent states saturate the inequality in Eq. (2.21) as well as the uncertainty relation in Eq. (2.22).

## Squeezed states

A single-mode squeezed state is a state resulted from applying a single-mode squeezed transformation, defined via symplectic matrix $S_{1 S Q}(r)$ in Eq. (2.11), to a coherent state $|\alpha\rangle$;

$$
|\alpha, r\rangle_{1 \mathrm{SQ}}:=\hat{U}_{1 \mathrm{SQ}}(r) \hat{D}_{\alpha}|0\rangle,
$$

where $\hat{U}_{1 \mathrm{SQ}}(r)=e^{r\left(\hat{a}^{2}-\hat{a}^{\dagger 2}\right)}$ is the squeezing operator associated with $S_{1 \mathrm{SQ}}(r)$. When the original coherent state is the vacuum state, the state is particularly called the single-mode squeezed vacuum state. $|\alpha, r\rangle_{1 S Q}$ has the same first moment vector $\overline{\mathbf{r}}=\sqrt{2}(\operatorname{Re}(\alpha), \operatorname{Im}(\alpha))^{\mathrm{T}}$ as the original coherent state but a different CM

$$
\sigma_{1 \mathrm{SQ}}=\left(\begin{array}{cc}
e^{-2 r} & 0  \tag{2.23}\\
0 & e^{2 r}
\end{array}\right)
$$

Notice that one of the diagonal elements is squeezed, but the other one is anti-squeezed compared to the CM of coherent states. $\sigma_{1 \mathrm{SQ}}$ in Eq. (2.23) also saturates the uncertainty principle in Eq. (2.22), but one of the variances of the canonical operators is below the vacuum level. This property is desirable when one needs a small statistical deviation. For example, squeezed states are widely used to enhance precision measurements in a wide range of physical platforms including phase-squeezed light in interferometric measurements [26] and spin squeezed states for atomic clocks [27, 28].

A two-mode squeezed state can be generated by applying a two-mode squeezed transformation defined with Eq. (2.12) to a two-mode coherent state $|\boldsymbol{\alpha}\rangle$ :

$$
|\boldsymbol{\alpha}, r\rangle_{2 \mathrm{SQ}}:=\hat{U}_{2 \mathrm{SQ}}(r) \hat{D}_{\boldsymbol{\alpha}}|0\rangle
$$

with $\boldsymbol{\alpha}=\left(\alpha_{1}, \alpha_{2}\right)^{\mathrm{T}} . \hat{U}_{2 \mathrm{SQ}}(r)=e^{r\left(\hat{a}_{1} \hat{a}_{2}-\hat{a}_{1}^{\dagger} \hat{a}_{2}^{\dagger}\right)}$ is the two-mode squeezing operator in the Hilbert space. When the original coherent state is the vacuum state, the resulting state is often called the two-mode squeezed vacuum state. Two-mode squeezed states have the first moment vector
$\overline{\mathbf{r}}=\sqrt{2}\left(\operatorname{Re}\left(\alpha_{1}\right), \operatorname{Im}\left(\alpha_{1}\right), \operatorname{Re}\left(\alpha_{2}\right), \operatorname{Im}\left(\alpha_{2}\right)\right)^{\mathrm{T}}$ and the CM

$$
\sigma_{2 \mathrm{SQ}}=\left(\begin{array}{cccc}
\cosh (2 r) & 0 & \sinh (2 r) & 0  \tag{2.24}\\
0 & \cosh (2 r) & 0 & -\sinh (2 r) \\
\sinh (2 r) & 0 & \cosh (2 r) & 0 \\
0 & -\sinh (2 r) & 0 & \cosh (2 r)
\end{array}\right)
$$

The two-mode squeezed vacuum state can be decomposed into the Fock basis as

$$
|\mathbf{0}, r\rangle_{2 \mathrm{SQ}}=\frac{1}{\cosh r} \sum_{n=0}^{\infty}(\tanh r)^{n}|n, n\rangle
$$

This expression clearly shows that the two modes of the two-mode squeezed vacuum state are entangled. Indeed, we can establish this entanglement in another way. Let us define new operators $\hat{x}_{ \pm}:=\left(\hat{x}_{1} \pm \hat{x}_{2}\right) / \sqrt{2}$ and $\hat{p}_{ \pm}:=\left(\hat{p}_{1} \pm \hat{p}_{2}\right) / \sqrt{2}$. From the CM in Eq. (2.24), we can check that

$$
\Delta \hat{x}_{-}^{2}=\Delta \hat{p}_{+}^{2}=e^{-2 r}, \quad \Delta \hat{x}_{+}^{2}=\Delta \hat{p}_{-}^{2}=e^{2 r}
$$

Thus, each two canonical operators are correlated and have one direction with squeezed variance and the other with anti-squeezed variance. When $r \rightarrow \infty$, the state $|\mathbf{0}, r\rangle_{2 S Q}$ goes to the uniform superposition of the tensor products of the same number states, which is analogous to the maximally entangled state in discrete variable quantum systems. Therefore, this non-physical state with $r \rightarrow \infty$ is sometimes referred to as the EPR state in CV quantum systems.

### 2.6 Phase-space representation of Gaussian states

The finite-dimensional description of Gaussian states with their statistical moments of the canonical operators opens the door for us to develop another expedient description of quantum states: the phase space representation. In this section, we describe how we can characterise quantum states in the phase space associated with the canonical operators, which is remindful of the classical phase space.

The construction of the phase space starts from the Fourier-Weyl relation, which is a mathematical tool connecting operators on the Hilbert space of a CV quantum system to functions of $2 N$ variables. The relation is based on the fact that the displacement operators $\hat{D}_{\mathbf{r}}$ form an orthogonal complete set in $L^{2}\left(\mathbb{R}^{N}\right)$ with respect to the Hilbert-Schmidt scalar product. It can be stated as

$$
\begin{equation*}
\rho=\frac{1}{(2 \pi)^{N}} \int_{\mathbb{R}^{2 N}} \mathrm{~d} \mathbf{r} \operatorname{Tr}\left[\hat{D}_{-\mathbf{r}} \rho\right] \hat{D}_{\mathbf{r}} \equiv \frac{1}{(2 \pi)^{N}} \int_{\mathbb{R}^{2 N}} \mathrm{~d} \mathbf{r} \chi(\mathbf{r}) \hat{D}_{\mathbf{r}} \tag{2.25}
\end{equation*}
$$

where $\mathrm{d} \mathbf{r}=\mathrm{d} x_{1} \mathrm{~d} p_{1} \ldots \mathrm{~d} x_{N} \mathrm{~d} p_{N}$, and we define a function

$$
\chi(\mathbf{r}):=\operatorname{Tr}\left[\hat{D}_{-\mathbf{r}} \rho\right] .
$$

Note that Eq. (2.25) holds for not only density operators but for general bounded operators on the Hilbert space. The function $\chi(\mathbf{r})$ offers full information about the state $\rho$ and is often called the characteristic function of the quantum state. We can observe that

$$
\chi(\mathbf{0})=\operatorname{Tr}[\rho]=1,
$$

and thus the normalisation condition of the density operator indicates that the value of the characteristic function at the origin is 1 . Due to the hermiticity of $\rho$, the function $\chi(\mathbf{r})$ also satisfies

$$
\chi(\mathbf{r})=\chi^{*}(-\mathbf{r}) .
$$

Taking the Fourier transform of the characteristic function $\chi(\mathbf{r})$ gives us the quasi-probability distribution $W(\mathbf{r})$ in the phase space:

$$
W(\mathbf{r})=\frac{1}{(\sqrt{2} \pi)^{2 N}} \int_{\mathbb{R}^{2 N}} \mathrm{~d} \mathbf{r}^{\prime} e^{-i \mathbf{r}^{\mathrm{T}} \Omega^{\mathrm{T}} \mathbf{r}} \chi\left(\mathbf{r}^{\prime}\right) .
$$

The function $W(\mathbf{r})$ is often called the Wigner function named after Eugene Wigner who first introduced the concept [29]. To see the physical implication of $W(\mathbf{r})$, let us consider a simple
single-mode case. We can observe that

$$
\begin{aligned}
W(x, p) & =\frac{1}{2 \pi^{2}} \iint_{\mathbb{R}^{2}} \mathrm{~d} x^{\prime} \mathrm{d} p^{\prime} e^{i\left(p x^{\prime}-x p^{\prime}\right)} \chi\left(x^{\prime}, p^{\prime}\right)=\frac{1}{2 \pi^{2}} \iint_{\mathbb{R}^{2}} \mathrm{~d} x^{\prime} \mathrm{d} p^{\prime} e^{i\left(p x^{\prime}-x p^{\prime}\right)} \int_{\mathbb{R}} \mathrm{d} q\langle q| \hat{D}_{-\boldsymbol{r}} \rho|q\rangle \\
& =\frac{1}{2 \pi^{2}} \iint_{\mathbb{R}^{3}} \mathrm{~d} x^{\prime} \mathrm{d} p^{\prime} \mathrm{d} q e^{i\left(p x^{\prime}-x p^{\prime}\right)}\langle q| \hat{D}_{-\frac{r}{2}} \rho \hat{D}_{-\frac{r}{2}}|q\rangle,
\end{aligned}
$$

where we used the quadrature eigenvectors $\{|q\rangle\}$ to evaluate the trace. The displacement operator $\hat{D}_{-\frac{r}{2}}$ can be expressed as

$$
\hat{D}_{-\frac{r}{2}}=e^{-\frac{i}{2}\left(x^{\prime} \hat{p}-p^{\prime} \hat{x}\right)}=e^{-i \frac{x^{\prime}}{2} \hat{p}} e^{i \frac{p^{\prime}}{2}} \hat{x}^{i \frac{i^{\prime} p^{\prime}}{8}}=e^{i \frac{p^{\prime}}{2}} \hat{x}^{-i \frac{x^{\prime}}{2} \hat{p}} e^{-i \frac{x^{\prime} p^{\prime}}{8}}
$$

using the Baker-Campbell-Hausdorff formula

$$
e^{\hat{A}+\hat{B}}=e^{\hat{A}} e^{\hat{B}} e^{-[\hat{A}, \hat{B}] / 2} .
$$

Then, we have

$$
\begin{aligned}
W(x, p) & =\frac{1}{2 \pi^{2}} \iint_{\mathbb{R}^{3}} \mathrm{~d} x^{\prime} \mathrm{d} p^{\prime} \mathrm{d} q e^{i p x^{\prime}} e^{i p^{\prime}(q-x)}\left\langle q+\frac{x^{\prime}}{2}\right| \rho\left|q-\frac{x^{\prime}}{2}\right\rangle \\
& =\frac{1}{\pi} \int_{\mathbb{R}} \mathrm{d} x^{\prime} e^{i p x^{\prime}}\left\langle x+\frac{x^{\prime}}{2}\right| \rho\left|x-\frac{x^{\prime}}{2}\right\rangle=\frac{2}{\pi} \int_{\mathbb{R}} \mathrm{d} x^{\prime} e^{2 i p x^{\prime}}\left\langle x+x^{\prime}\right| \rho\left|x-x^{\prime}\right\rangle .
\end{aligned}
$$

This implies that

$$
\frac{1}{2} \int_{\mathbb{R}} \mathrm{d} p W(x, p)=\langle x| \rho|x\rangle
$$

and thus the integral of the Wigner function over one quadrature gives the physical probability distribution of measuring the conjugate quadrature, which is same as the outcome probability distribution of homodyne detection for the conjugate quadrature. This argument is phase invariant; it works for any general quadrature operator defined as $\hat{x}_{\theta}=\cos \theta \hat{x}-\sin \theta \hat{p}$.

The characteristic function and Wigner function of a generic $N$-mode Gaussian state $\rho_{G}$
have the forms

$$
\begin{aligned}
& \chi_{\rho_{G}}(\mathbf{r})=e^{-\frac{1}{4} \mathbf{r}^{\mathrm{T}} \Omega^{\mathrm{T}} \sigma \Omega \mathbf{r}} e^{i \mathbf{r}^{\mathrm{T}} \Omega^{\mathrm{T}} \overline{\mathbf{r}}} \\
& W_{\rho_{G}}(\mathbf{r})=\frac{2^{N}}{\pi^{N} \sqrt{\operatorname{det}(\sigma)}} e^{-(\mathbf{r}-\overline{\mathbf{r}})^{\mathrm{T}} \sigma^{-1}(\mathbf{r}-\overline{\mathbf{r}})}
\end{aligned}
$$

where $\overline{\mathbf{r}}$ and $\sigma$ are the first moment vector and the CM of $\rho_{G}$, respectively. This again demonstrates that every Gaussian state can be fully characterised by its first moments and CM and strengthens the connection with the classical Gaussian distributions.

### 2.7 Further reading

In this section, we introduce some selected topics on continuous variable quantum systems, which help us read through this thesis. We omitted some derivations and detailed calculations of mathematical results in this section when they are not particularly helpful for this thesis. CV quantum systems and the Gaussian framework are broad research topics which may require a whole text book to provide all relevant information. We kindly refer to a textbook [13] or review papers $[19,20]$ for more broad introductions on CV quantum systems and CV quantum information processing.

## Chapter 3

## Background theory: quantum resource theories

A physical task in general comes with restrictions, such as a limited set of tools or a distance between parties. These restrictions play an important role of determining which physical operations we can or cannot perform without requiring more resources. An optimal way to succeed the task would be to achieve the goal with these 'allowed' operations and as least additional resources as possible. Resource theories capture this basic structure existing in any physical task and provide a mathematical framework in which we can explore the physical implications that this structure draws.

In this chapter, we introduce a useful, versatile mathematical tool in quantum information theory called quantum resource theories (QRTs). Although the concept of resource theories can be applied far beyond quantum physics, in this thesis we only focus on the ones admitting the structure of quantum theory. The mathematical platform provided by QRTs helps us identify resourceful physical quantities or effects for given restrictions, and this allows us to derive state transformation conditions without using any identified resources. The state transformation conditions imply a hierarchical structure of quantum states which is then related to a question of quantifying the resource present in each state. QRTs have been successfully applied to various topics in quantum information theory including the most notable example of the resource theory of entanglement [30] as well as resource theories of magic [31, 32], asymmetry [33, 34]
and thermodynamics [35, 36]. Although the details depend on the specific choice of physical resources, QRTs admit a well-established universal structure, and there have been developed systematic ways of quantifying the resource of interest, such as entropic measures and geometric measures [30, 37]. There are also mathematical techniques proven to be useful in different QRTs like majorisation [38].

In Section 3.1, we first look at the general structure of quantum resource theories including their two key ingredients, allowed operations and free states, and the relation between the two. Then, we move on to the state-convertibility problem in Section 3.2, which is one of the basic physical tasks that can be studied via QRTs. We introduce a useful mathematical tool called majorisation whilst discussing about the single-shot convertibility problem. In Section 3.3, we study how to quantify the amount of resources present in a state and introduce an important concept called resource monotones. Lastly, we generalise all these concepts to the case when there exist more than one resource in Section 3.4. This chapter is written based on [37, 39].

### 3.1 The general structure of quantum resource theories

In this section, we describe the general mathematical structure of QRTs. We only focus on the construction of QRTs for static resources; in other words, QRTs involving interconversions between quantum states. For the case considering resource theories of quantum processes, we refer to [40, 41].

Our main objects in this section are quantum systems with some given physical restrictions and state transformations acting on these systems. A quantum system is associated by a Hilbert space $\mathcal{H}$ which can be either finite- or infinite-dimensional, and a possible state of the system is represented by a density operator acting on $\mathcal{H}$. Recall that we denote the set of all bounded operators on $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$, the set of all quantum states (density operators) in $\mathcal{H}$ by $\mathcal{D}(\mathcal{H})$, and the set of all CP maps from $\mathcal{B}\left(\mathcal{H}_{\text {in }}\right)$ to $\mathcal{B}\left(\mathcal{H}_{\text {out }}\right)$ by $\mathcal{C}\left(\mathcal{H}_{\text {in }}, \mathcal{H}_{\text {out }}\right)^{1}$. A quantum resource theory $\mathcal{R}$ of resource $f$ can be defined with two key ingredients: the set of free states $\mathcal{F}$ and the allowed operations $\mathbb{O}$. In simple words, the set of free states $\mathcal{F}$ is a subset of all quantum states $\mathcal{D}(\mathcal{H})$,

[^10]i.e., $\mathcal{F} \subseteq \mathcal{D}(\mathcal{H})$, consisting of all states which do not possess any resource $f$; i.e., it is the set of all states which can be produced for free under the given restrictions. The set of allowed operations $\mathbb{O}$ is a set of linear maps acting on $\mathcal{H}$ which preserve $\mathcal{F}$, i.e., which do not create any resource from a free state. We state the formal definition of the QRT as follows.

Definition 2. A quantum resource theory $\mathcal{R}$ of resource $f$ is defined by the tuple $\mathcal{R}=(\mathcal{F}, \mathbb{O})$, where each of $\mathcal{F}$ and $\mathbb{O}$ is a subset of $\mathcal{D}(\mathcal{H})$ and $\mathcal{C}(\mathcal{H})$ respectively, satisfying the following golden rule of QRTs: for any $\rho \in \mathcal{F}$ and $\Gamma \in \mathbb{O}$, it holds that $\Gamma(\rho) \in \mathcal{F}$.

Note that we allow the input and output Hilbert spaces to be different. In that case, a more rigorous description of $\mathcal{F}$ should be made. As any quantum state can be regarded as a quantum channel preparing the state, we can define the set of free states $\mathcal{F}(\mathcal{H})$ as a subset of $\mathcal{C}(\mathbb{C}, \mathcal{H})$ for the corresponding Hilbert space $\mathcal{H}$. Then, free states $\mathcal{F}\left(\mathcal{H}_{\text {in }}\right)$ and allowed operations $\mathbb{O}\left(\mathcal{H}_{\text {in }}, \mathcal{H}_{\text {out }}\right)$ are a subset of $\mathcal{C}\left(\mathbb{C}, \mathcal{H}_{\text {in }}\right)$ and $\mathcal{C}\left(\mathcal{H}_{\text {in }}, \mathcal{H}_{\text {out }}\right)$ respectively, and the golden rule can be stated that for any $\rho \in \mathcal{F}\left(\mathcal{H}_{\text {in }}\right)$ and $\Gamma \in \mathbb{O}\left(\mathcal{H}_{\text {in }}, \mathcal{H}_{\text {out }}\right)$, it holds that $\Gamma(\rho) \in \mathcal{F}\left(\mathcal{H}_{\text {out }}\right)$. Imposing the golden rule of QRTs is important as it prevents the QRT to become trivial; otherwise, there exists a pair of a free state $\rho_{\text {free }}$ and an allowed operation $\Gamma$ such that $\Gamma\left(\rho_{\text {free }}\right) \in \mathcal{D}(\mathcal{H}) \backslash \mathcal{F}(\mathcal{H})$, which means that the resource can be obtained for free.

By definition, we always include the identity channel in the allowed operations. We also assume that QRTs admit the tensor-product structure, which means that (i) all allowed operations are completely free, i.e., if $\Gamma_{A} \in \mathbb{O}\left(\mathcal{H}_{A}\right)$, then $\mathcal{I}_{E} \otimes \Gamma_{A} \in \mathbb{O}\left(\mathcal{H}_{E} \otimes \mathcal{H}_{A}\right)$ with identity map $\mathcal{I}_{E}$ on $\mathcal{B}\left(\mathcal{H}_{E}\right)$, (ii) appending an ancillary system in a free state $\rho_{E} \in \mathcal{F}\left(\mathcal{H}_{E}\right)$ is free, and (iii) discarding subsystems (partial trace) is free. These last two points imply the interconvertibility between free states via allowed operations; for any pair of free states, $\rho$ and $\sigma, \rho$ can be transformed into $\sigma$ by allowed operations, and vice versa.

When defining a QRT, one can start from either given free states or given allowed operations. In both cases, the other element should be chosen carefully so that the golden rule of QRTs is satisfied. For example, in the QRT of entanglement, the given physical restrictions, spatially separated parties without quantum communication channel, naturally determine the allowed operations as the set of local operations and classical communications (LOCC). Then, the free states are defined as all states which can be prepared with LOCC operations, which are separable
states. When the allowed operations are given first, the set of free states $\mathcal{F}(\mathcal{H})$ consistent with the golden rule of QRT s is unique ${ }^{2}$ and given by

$$
\begin{equation*}
\mathcal{F}_{\min }(\mathcal{H}):=\{\rho \mid \forall \sigma \in \mathcal{D}(\mathcal{H}), \exists \Gamma \in \mathbb{O} \text { s.t. } \Gamma(\sigma)=\rho\} \tag{3.1}
\end{equation*}
$$

On the other hand, in the QRT of coherence, it is more natural to define the set of free states as all quantum states without coherence in certain basis and then figure out the appropriate allowed operations. In this case, there can be many different choices of the allowed operations for given free states, which are consistent with the golden rule of QRTs. For example, we can consider as allowed operations the set of resource non-generating (RNG) operations defined as all CPTP maps $\Gamma \in \mathcal{C}\left(\mathcal{H}_{\text {in }}, \mathcal{H}_{\text {out }}\right)$ such that $\Gamma(\rho) \in \mathcal{F}\left(\mathcal{H}_{\text {out }}\right)$ for all $\rho \in \mathcal{F}\left(\mathcal{H}_{\text {in }}\right)$, or the set of 'completely' resource non-generating (CRNG) operations, which are all CPTP maps $\Gamma \in \mathcal{C}\left(\mathcal{H}_{\text {in }}, \mathcal{H}_{\text {out }}\right)$ such that $\mathcal{I}_{k} \otimes \Gamma \in \mathcal{C}\left(\mathbb{C}^{k} \otimes \mathcal{H}_{\text {in }}, \mathbb{C}^{k} \otimes \mathcal{H}_{\text {out }}\right)$ is a resource non-generating map for all positive numbers $k$. The two sets are not the same - in fact, it holds that CRNG $\subset$ RNG - and each is a valid set of allowed operations satisfying the golden rule.

We learned how to define QRTs and their general structure in this section. States belonging to $\mathcal{D}(\mathcal{H}) \backslash \mathcal{F}(\mathcal{H})$ are regarded as resourceful states. We now want to look at some simple physical task which can be studied based on the defined structure of QRTs: the convertibility between resourceful states.

### 3.2 State convertibility

One of the basic information-theoretic tasks that can be studied with QRTs is the convertibility between different quantum states via allowed operations; namely, under the given restrictions. Let us consider a QRT $\mathcal{R}=(\mathcal{F}, \mathbb{O})$. For simplicity, we henceforth assume that the input and output Hilbert spaces are the same and omit $\mathcal{H}$ in the notations of $\mathcal{F}$ and $\mathbb{O}$.

[^11]
### 3.2.1 Single-shot case

The most basic scenario of the convertibility problem is the conversion of one resourceful state into another using allowed operations. For some quantum states $\rho$ and $\sigma$, if there exists $\Gamma \in \mathbb{O}$ such that $\Gamma(\rho)=\sigma$, we write

$$
\begin{equation*}
\rho \xrightarrow{0} \sigma . \tag{3.2}
\end{equation*}
$$

We are interested in finding the necessary and sufficient conditions for the transformation of a given input state into a given output quantum state via allowed operations. In general, it is a difficult task to derive precise state-transformation conditions, and they highly depend on detailed properties of the free states and the allowed operations. Nevertheless, we present one example, the QRT of pure bipartite entanglement [38], which has well-established single-shot state-transformation conditions. It provides a good demonstration of the important role played by the mathematical tool called majorisation, which is also applicable to other QRTs.

Before looking at the state-transformation conditions in the QRT of pure bipartite entanglement, let us introduce the notion of majorisation. Majorisation is a preorder between real vectors defined as follows.

Definition 3. (Majorisation) Given two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{d}$ with non-negative elements, we say that $\boldsymbol{x}$ 'majorises' $\boldsymbol{y}$ written as $\boldsymbol{x} \succ \boldsymbol{y}$ if and only if

$$
\sum_{i=1}^{k} x_{i}^{\downarrow} \geq \sum_{i=1}^{k} y_{i}^{\downarrow} \quad \forall k \in\{1, \ldots, d\}
$$

where $\boldsymbol{z}^{\downarrow}$ for some real vector $\boldsymbol{z}$ denotes the vector with same components but sorted in descending order.

Majorisation has many applications in areas beyond physics such as gambling theory and economics, and it has also demonstrated its usefulness when applied to QRTs, mainly in the QRT of entanglement and thermodynamics.

The QRT of pure bipartite entanglement considers two spatially separated quantum systems without quantum communication channel. The allowed operations are naturally identified as

LOCC. As the allowed operations are given, we can easily find the consistent free states using $\mathcal{F}_{\text {min }}$ defined in Eq. (3.1) but with state space restricted to pure quantum states, i.e., unit vectors in $\mathcal{H}$. By construction, this set includes all pure quantum states which can be generated by LOCC starting from any other pure states and is given by all pure product states ${ }^{3}$. Then, using the Schmidt decomposition, any pure bipartite state $|\psi\rangle_{A B} \in \mathcal{H}_{A B}$ can be written in the form of

$$
|\psi\rangle_{A B}=\sum_{i=1}^{d} \sqrt{\lambda_{i}^{\psi}}\left|a_{i}\right\rangle_{A} \otimes\left|b_{i}\right\rangle_{B},
$$

where $\lambda_{i}^{\psi} \geq 0$ for all $i$ and $\sum_{i=1}^{d} \lambda_{i}^{\psi}=1$. We assumed that the two Hilbert spaces $\mathcal{H}_{A}$ and $\mathcal{H}_{B}$ have the same dimension $d$, i.e., $\operatorname{dim}\left(\mathcal{H}_{A}\right)=\operatorname{dim}\left(\mathcal{H}_{B}\right)=d$. The values $\left\{\lambda_{i}^{\psi}\right\}_{i}$ are often called Schmidt coefficients of the state $|\psi\rangle_{A B}$. For a given pure bipartite state $|\psi\rangle_{A B}$, let us define a vector of its Schmidt coefficients as $\boldsymbol{S}(|\psi\rangle)=\left(\lambda_{1}^{\psi}, \ldots, \lambda_{d}^{\psi}\right)^{\mathrm{T}}$. Then, Nielsen majorisation theorem [38] states that for given pure bipartite states $|\psi\rangle_{A B},|\phi\rangle_{A B} \in \mathcal{H}_{A B}$,

$$
\begin{equation*}
|\psi\rangle_{A B} \xrightarrow{L O C C}|\phi\rangle_{A B} \quad \text { iff } \quad \boldsymbol{S}(|\psi\rangle) \prec \boldsymbol{S}(|\phi\rangle) . \tag{3.3}
\end{equation*}
$$

Thus, the necessary and sufficient conditions of the single-shot convertibility in the QRT of pure bipartite entanglement are simply given by the majorisation relation between vectors of Schmidt coefficients.

The state transformation condition like Eq. (3.3) gives preorders to resourceful states, which implies a hierarchical structure of the state space in a QRT under the allowed operations. This allows us to compare resources present in different resourceful states even without defining a proper resource measure; for given two states $\rho, \sigma$, if $\rho \xrightarrow{0} \sigma$, it is safe to assume that $\rho$ is at least as resourceful as $\sigma$ since we can generate $\sigma$ from $\rho$ without any additional resource. However, comparing resources based on the convertibility is restricted as there are lots of states which are not interconvertible, and state transformation conditions usually fail to form a total order. This highlights a necessity of more sophisticated ways to quantify and compare resources present in arbitrary resourceful states, and this is the topic of Section 3.3. Before moving on to that, we

[^12]introduce a few important variants of the convertibility problem.

## Catalytic transformation

Majorisation also implies a remarkable concept in QRTs, the existence of catalysis. A catalyst, named after catalysts in chemical catalysis process, is a state of an ancillary system which makes a certain state transformation with allowed operations possible while remaining unchanged during the process. More specifically, a state $\tau$ is called a resource catalyst for the transformation from $\rho$ to $\sigma$ if

$$
\begin{equation*}
\rho \stackrel{\oplus}{\nrightarrow} \sigma \quad \text { and } \quad \rho \otimes \tau \xrightarrow{0} \sigma \otimes \tau . \tag{3.4}
\end{equation*}
$$

The existence of such catalysts in QRTs is not obvious but was first spotted by the authors of [42] for the QRT of pure bipartite entanglement. Let us consider two bipartite quantum states $|\psi\rangle_{A B}$ and $|\phi\rangle_{A B}$ with following vectors of Schmidt coefficients

$$
\boldsymbol{S}(|\psi\rangle)=\left(\frac{2}{5}, \frac{2}{5}, \frac{1}{10}, \frac{1}{10}\right)^{\mathrm{T}}, \quad \boldsymbol{S}(|\phi\rangle)=\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, 0\right)^{\mathrm{T}}
$$

Using the definition of majorisation in Definition 3, we can check that neither $\boldsymbol{S}(|\psi\rangle) \prec \boldsymbol{S}(|\phi\rangle)$ nor $\boldsymbol{S}(|\phi\rangle) \prec \boldsymbol{S}(|\psi\rangle)$; up to the first element, $2 / 5<1 / 2$ but up to the second element, $2 / 5+2 / 5>1 / 2+1 / 4$. Thus, they are not interconvertible, i.e., neither can be converted into each other via LOCC. However, if we introduce an ancillary system in $|\chi\rangle_{A^{\prime} B^{\prime}}$ with $\boldsymbol{S}(|\chi\rangle)=(3 / 5,2 / 5)^{\mathrm{T}}$, then it is not difficult to check that

$$
\boldsymbol{S}(|\psi\rangle) \otimes \boldsymbol{S}(|\chi\rangle) \prec \boldsymbol{S}(|\phi\rangle) \otimes \boldsymbol{S}(|\chi\rangle) .
$$

That is, the state $|\chi\rangle$ acts as a catalyst which enables the otherwise impossible transformation $|\psi\rangle \xrightarrow{\text { LOCC }}|\phi\rangle$.

As we have seen with the above example, allowing catalysts generally enlarges the set of allowed operations in QRTs. We call the new set of allowed operations with possibility of using catalysts as catalytic allowed operations. Catalysts play an important role in QRTs.

Especially, in the QRT of quantum thermodynamics, introducing catalysts is operationally motivated as it allows us to study the full picture of quantum thermodynamical processes involving experimental apparatus. It has been shown that the state transformation conditions for cyclic thermal processes including catalysts are characterised by a family of free energies, which generalises the traditional second law of thermodynamics [35]. Similar results, that the state transformation conditions under catalytic allowed operations are characterised by a family of functions, have been reported for the bipartite pure entanglement [43] and also for the more general case of the mathematical majorisation relation [44].

## Non-exact convertibility

We have so far studied exact state transformations as in Eq. (3.2) and (3.4). However, considering exact state transformations is usually too strict and not physically relevant as real experiments always involve errors or noise. There have been developed several variants of alternative resource transformations inspired by this issue, and we introduce a couple of notable examples in this section.

The first variant is to consider the approximate transformation in which the initial state is transformed to a state 'close' enough to the final state. The 'closeness' between states can be determined by fidelity or a mathematical distance measure of choice [35, 45]. A common technique for this generalisation is to introduce the smoothing. Let us consider a quantum state $\rho$. We define an $\epsilon$-smoothing ball around $\rho$ as

$$
B_{\epsilon}^{D}(\rho):=\{\tilde{\rho} \mid D(\rho, \tilde{\rho}) \leq \epsilon\},
$$

for some distance measure $D$. Then, we write

$$
\rho \stackrel{\mathbb{O}}{\rightarrow}_{\epsilon} \sigma
$$

if there exists $\Gamma \in \mathbb{O}$ such that $\Gamma(\rho)=\sigma^{\prime}$ where $\sigma^{\prime} \in B_{\epsilon}^{D}(\sigma)$. When also including catalysts to approximate transformation, one faces different ways to define the 'closedness', and this has been a fruitful source of interesting results across different QRTs. One notable example
is resource embezzling, first discovered in [45] in the QRT of entanglement, where any state transformation becomes possible under allowed operations with large enough catalysts if the bound on $\epsilon$ is not carefully chosen. Also, in the QRT of thermodynamics, it has been shown that one can recover a single inequality condition describing the second law of thermodynamics when one allows final correlation between the system and the catalyst and only requires the local catalytic state to return to the original state [46].

Another variant is to relax the condition that the state transformation should be deterministic. One can instead consider a probabilistic or stochastic transformation from one quantum state to another. For quantum states $\rho, \sigma$, we say a probabilistic transformation from $\rho$ to $\sigma$ is possible and write

$$
\rho \stackrel{\mathscr{G}}{p} \sigma
$$

if there exists $\Gamma \in \mathbb{O}$ such that $\Gamma(\rho)$ results in $\sigma$ with nonzero probability. A notable result in this direction is the work by Bennett et al. showing that every bipartite partially entangled state can be transformed into a maximally entangled state with nonzero probability [47].

### 3.2.2 Many-copy case

In this section, we consider the case when we have access to many copies of the initial state and investigate the convertibility of multiple copies of quantum states under allowed operations. Considering this many-copy scenario, which is also called the asymptotic scenario, can be seen as one approach to relax the exact state transformation and make more states interconvertible. The generalisation of $\mathcal{F}$ and $\mathbb{O}$ to the many-copy case is straightforward. We now consider an initial state of the form $\rho^{\otimes N} \in \mathcal{D}\left(\mathcal{H}^{\otimes N}\right)$. For this tensor-product Hilbert space $\mathcal{H}^{\otimes N}$, the sets of allowed operations $\mathbb{O}\left(\mathcal{H}^{\otimes N}\right)$ and free states $\mathcal{F}\left(\mathcal{H}^{\otimes N}\right)$ are subsets of $\mathcal{C}\left(\mathcal{H}^{\otimes N}\right)$ and $\mathcal{D}\left(\mathcal{H}^{\otimes N}\right)$ respectively and satisfy the golden rule of QRTs.

In the asymptotic scenario, the primary research topic is the rate of conversion between two quantum states. For given two quantum states $\rho, \sigma$ on $\mathcal{H}$, a rate of conversion from $\rho$ to $\sigma$ is a number $R$ such that for every $R^{\prime}<R$ and $\epsilon>0$ there exists a sufficiently large integer $N$
satisfying

$$
\rho^{\otimes N} \xrightarrow[\rightarrow]{\epsilon}_{\epsilon} \sigma^{\otimes\left\lfloor N R^{\prime}\right\rfloor} .
$$

The optimal rate $R(\rho \rightarrow \sigma)$ is then defined as the supremum of all $R$ s satisfying the above condition. As free states can be prepared for free by definition, $R(\rho \rightarrow \sigma)=+\infty$ for any $\sigma \in \mathcal{F}$. Also, one can expect that $R(\rho \rightarrow \sigma)=0$ for any $\rho \in \mathcal{F}$ and $\sigma \notin \mathcal{F}$.

An important protocol in the asymptotic scenario is resource distillation. In resource distillation protocols, one studies the possibility/optimal rate of conversions of multiple copies of a less resourceful state $\rho$ into less copies of a more resourceful state $\sigma$, in which $R(\rho \rightarrow \sigma)$ is less than 1. For example, in the QRT of bipartite entanglement, it is of interest to distill the maximally entangled state $|\Phi\rangle_{A B}$ from many copies of a partially entangled state with LOCC [47-49]. The quantity $R(\rho \rightarrow|\Phi\rangle\langle\Phi|)$ is particularly called the distillable entanglement of $\rho$, and the rate for the reverse protocol $R(|\Phi\rangle\langle\Phi| \rightarrow \rho)$ is called the entanglement cost of $\rho$. This asymptotic scenario also raises the interesting problem known as reversibility. For given quantum states $\rho, \sigma$ on $\mathcal{H}$, they are called reversible if $R(\rho \rightarrow \sigma) R(\sigma \rightarrow \rho)=1$. In this case, we can perform a cyclic transformation $\rho \rightarrow \sigma \rightarrow \rho$ without loss of any resources in the asymptotic rate. Except a few examples with specific sets of allowed operations [50, 51], most QRTs however exhibit resource irreversibility in their state transformations; $R(\rho \rightarrow \sigma) R(\sigma \rightarrow \rho)<1$ for some $\rho, \sigma$. The irreversibility in QRTs then revealed the existence of states with bound entanglement [52], which have non-zero entanglement cost but zero distillable entanglement, and analogs in other QRTs $[53,54]$.

### 3.3 Quantifying resources: monotones

In this section, our focus is on quantifying the amount of resources in an arbitrary quantum state for the given QRT $\mathcal{R}$. As we have already mentioned, state transformation conditions under allowed operations give us some guidance to how we develop resource measures. If a state $\rho$ can be transformed into another state $\sigma$ under allowed operations, we can rightfully assume that $\rho$ is at least as resourceful as $\sigma$. In addition, the structure of QRTs provides some insights
into desirable properties of resource measures.
A typical resource measure in QRTs is a resource monotone. A resource monotone is a non-negative function $M: \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}_{\geq 0}$ such that

$$
\begin{equation*}
M(\rho) \geq M(\Gamma(\rho)) \quad \forall \rho \in \mathcal{D}(\mathcal{H}), \forall \Gamma \in \mathbb{O} \tag{3.5}
\end{equation*}
$$

The property in Eq. (3.5) is called the monotonicity under allowed operations, from which the name 'monotone' came from. It assures that the amount of resource present in a quantum state never increases under allowed operations and reflects the golden rule of QRTs. Another fundamental property of any resource measure is that it should vanish for any free state:

$$
\begin{equation*}
M(\rho)=0 \quad \text { if } \rho \in \mathcal{F} \tag{3.6}
\end{equation*}
$$

Due to our assumption on the interconvertibility between free states, the monotonicity automatically implies $M(\rho)=M(\sigma)$ for any $\rho, \sigma \in \mathcal{F}$, and we can always make this property in Eq. (3.6) satisfied by adding a constant to $M$ so that $M(\rho)=0$ for any $\rho \in \mathcal{F}$.

There are a few non-essential but desirable properties for resource measures easily found in the literature. The first one is faithfulness saying that the converse of Eq. (3.6) is also true: $M(\rho)=0$ if and only if $\rho \in \mathcal{F}$. This property is convenient when one's task is to distinguish resourceful states from free states, but for example distillable entanglement famously does not satisfy the faithfulness [52]. The second property is the convexity defined as

$$
M\left(\sum_{i} p_{i} \rho_{i}\right) \leq \sum_{i} p_{i} M\left(\rho_{i}\right)
$$

for any ensemble of quantum states $\left\{\rho_{i}\right\}_{i}$ and probability distribution $\left\{p_{i}\right\}_{i}$. It is a desirable property as computing a resource monotone, which often involves optimisation, can take advantage of well-studied convex optimisation problems if the monotone satisfies the convexity [55]. Another desirable property is the subadditivity: a function $f$ is called subadditive if

$$
f(\rho \otimes \sigma) \leq f(\rho)+f(\sigma)
$$

for any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. If the equality holds, the function $f$ is called additive. The subadditivity is common among resource monotones in QRTs while the additivity is a very strong property and rare. Nevertheless, there is a procedure called regularisation for constructing a new resource monotone which is additive on multi-copies of the same state. For a given resource monotone $M: \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}_{\geq 0}$, the regularised version $M^{\infty}$ is defined as

$$
M^{\infty}(\rho)=\lim _{n \rightarrow \infty} \frac{1}{n} M\left(\rho^{\otimes n}\right)
$$

with assumption that the limit exists. By definition, $M^{\infty}$ is additive on the same state; $M^{\infty}\left(\rho^{\otimes n}\right)=n M^{\infty}(\rho)$.

Different QRTs have different resource monotones, and there can be also several different monotones for the same QRT. In the remaining of this section, we will introduce one famous type of resource monotones applicable to general QRTs.

## Distance-based monotones

Distance-based monotones are a family of monotones that can be constructed in any QRTs when given the set of free states. The idea is that we can quantify the amount of resource in a quantum state by measuring how far the state is from the set of free states. To measure this 'distance' between a given quantum state and the set of free states, we employ a contractive metric for the state space. A metric function $d:(\mathcal{D}(\mathcal{H}), \mathcal{D}(\mathcal{H})) \rightarrow \mathbb{R}_{\geq 0}$ is called contractive if for any pair of quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ and for any CPTP map $\Phi$, it holds that

$$
d(\Phi(\rho), \Phi(\sigma)) \leq d(\rho, \sigma)
$$

With such $d$, we can then define a resource monotone for any QRT as

$$
M_{d}(\rho)=\inf _{\sigma \in \mathcal{F}} d(\rho, \sigma)
$$

It is not difficult to check the monotonicity of $M_{d}$ : let us consider a quantum state $\rho$ and an allowed operation $\Gamma \in \mathbb{O}$. Then, we have

$$
\begin{aligned}
M_{d}(\Gamma(\rho)) & =\inf _{\sigma \in \mathcal{F}} d(\Gamma(\rho), \sigma) \\
& \leq \inf _{\sigma^{\prime} \in \mathcal{F}} d\left(\Gamma(\rho), \Gamma\left(\sigma^{\prime}\right)\right) \\
& \leq \inf _{\sigma^{\prime} \in \mathcal{F}} d\left(\rho, \sigma^{\prime}\right)=M_{d}(\rho),
\end{aligned}
$$

where we restricted the optimisation set to the image of $\Gamma$ in the second line as $\Gamma(\mathcal{F}) \subseteq \mathcal{F}$ for all $\Gamma \in \mathbb{O}$ and used the contractive property of $d$ in the last line. Typical contractive metrics used in QRTs are quantum relative Rényi entropies [56] of which the quantum relative entropy is an example: for $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ such that $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$, the quantum relative entropy between $\rho$ and $\sigma$ is defined as

$$
D(\rho \mid \sigma):=\operatorname{Tr}[\rho(\log \rho-\log \sigma)] .
$$

### 3.4 Multi-resource theories

We have so far considered the case when there exists only one resource (or one type of physical restrictions) with one pair of allowed operations $\mathbb{O}$ and free states $\mathcal{F}$. However, it is often true that a physical task involves more than one resource or more than one type of restrictions. For example, we later in Chapter 4 see the case when one's task is to distill some resource $f$, such as entanglement or purity, only using Gaussian states and Gaussian operations. This is the case when there are two different resources which are resource $f$ and the non-Gaussianity. Another example is quantum thermodynamics where one might regard energy and entropy as separate resources with different allowed operations and free states.

The general framework of QRTs that we have studied so far in this chapter can be generalised to the cases with more than one resource [57]. Let us consider a set of $n$ different QRTs $\left\{\mathcal{R}_{i}\right\}_{i=1}^{n}$ with corresponding sets of allowed operations $\left\{\mathbb{O}_{i}\right\}_{i=1}^{n}$ and free states $\left\{\mathcal{F}_{i}\right\}_{i=1}^{n}$. Then, the combined multi-resource theory $\mathcal{R}_{\text {multi }}$ is defined by a new set of allowed operations $\mathbb{O}_{\text {multi }}$
defined as the intersection between all different sets of allowed operations:

$$
\begin{equation*}
\mathbb{O}_{\text {multi }}:=\cap_{i=1}^{n} \mathbb{O}_{i} . \tag{3.7}
\end{equation*}
$$

Then, the new set of free states $\mathcal{F}_{\text {multi }}$ consistent with $\mathbb{O}_{\text {multi }}$ is followed using Eq. (3.1):

$$
\mathcal{F}_{\text {multi }}:=\left\{\rho \mid \forall \sigma \in \mathcal{D}(\mathcal{H}), \exists \Gamma \in \mathbb{O}_{\text {multi }} \text { s.t. } \Gamma(\sigma)=\rho\right\}
$$

It has been pointed out in [57] that there could be no free states in a quantum multi-resource theory. This is the case when all free states in one resource theory $\mathcal{R}_{i}$ are resourceful states in another resource theory $\mathcal{R}_{j}$, i.e., when the interception between $\mathcal{F}_{i}$ and $\mathcal{F}_{j}$ is an empty set. However, in this thesis, we only study the case when $\mathcal{F}_{\text {multi }}$ is not an empty set.

Starting from this general structure of multi-resource theories, we can repeat the same procedures as the ones for single-resource theories studied in the previous sections to develop, for example, appropriate resource measures. In multi-resource theories, new types of protocols, such as interconversion between different resources, become important. We refer to [57] for more detailed discussions of resource interconversion and the reversibility in quantum multi-resource theories.

### 3.5 Further reading

This chapter provides a brief introduction of quantum resource theories. We introduce some selected topics which are necessary for this thesis. For a more extensive review of QRTs, please see [37]. There are also separate reviews on QRTs of different topics, such as quantum entanglement [30], quantum thermodynamics [58], quantum reference frames [59] and Bell nonlocality [60].

## Chapter 4

## Convex Gaussian resource theories

The Gaussian framework is a preferable tool for exploring quantum theory due to its easy accessibility in experiments and concise mathematical description, which we have already seen in Section 2. Indeed, many early physical platforms for implementing quantum information processing tasks include systems described by the Gaussian framework such as optical and optomechanical systems [14-16, 61-63]. However, it was soon discovered that several limitations hold for the Gaussian framework when applied to various quantum information processing tasks. For example: quantum computation with Gaussian elements does not provide any advantage over classical computation $[6-8,64]$; it is impossible to correct Gaussian errors with Gaussian operations and measurements [9]; and entanglement distillation from Gaussian states using Gaussian operations and measurements is impossible [65-67].

Recently, more general resource distillation in the Gaussian framework has been discussed in [10], where the authors proposed Gaussian resource theories, quantum resource theories further restricted to Gaussian states and operations. They showed that resource distillation, not only entanglement distillation, is in general impossible in such Gaussian resource theories. One obvious way to overcome this limitation is to introduce non-Gaussian resources, and it is known that entanglement distillation [68], error correction [68, 69] and other tasks [6, 7, 9] become possible when non-Gaussian elements are introduced. Non-Gaussian elements are however expensive resources to be generated in experiments, and it is desirable to find a simpler solution to circumvent these limitations.

In this chapter, we examine whether classical randomness and post-selection can lift the limitations on resource distillation in Gaussian resource theories. We extend Gaussian resource theories in [10] to convex Gaussian resource theories so that the theories include these new tools, classical randomness and post-selection. These tools enable operations such as mixing Gaussian states or conditional Gaussian operations based on the outcomes of Gaussian measurements. The inclusion of these operations is well-motivated from an operational point of view since they are less difficult resources to be employed in a laboratory than non-Gaussian elements. Here, we formulate convex Gaussian resource theories based on the multi-resource perspective and then investigate resource distillation within convex Gaussian resource theories. Interestingly, we find that the restrictions on resource distillation can be relaxed in convex Gaussian resource theories; the impossibility of resource distillation is replaced by a limitation on the amount of distillable resources, which still leaves some possibility of resource distillation. We explore this possibility and construct explicit resource distillation protocols using the newly allowed tools in convex Gaussian resource theories. In particular, these examples show that our derived upper bound on the distillable resources is tight and can be saturated in special cases.

This chapter is based on the paper [H. H. Jee et al. Phys. Rev. A, 103 (2021)] and structured as follows. Firstly, in Section 4.1, we study the Gaussian resource theories introduced in [10] with the multi-resource theoretical perspective and revise the no-go theorem for Gaussian resource distillation (Proposition 3). In Section 4.2, we then define the notion of convex Gaussian resource theories with allowed operations defined in Definition 4 which particularly includes conditional Gaussian operations. In Section 4.3, we introduce a new resource measure for convex Gaussian resource theories and derive some of its useful properties (Lemma 4), which lead to our main result, Corollary 6, describing a general upper bound on the amount of distillable resources in convex Gaussian resource theories. Then, in Section 4.4, we construct as examples explicit distillation protocols showing that resource distillation becomes possible in a limited fashion in convex Gaussian resource theories. We summarise our results in Section 4.5.

### 4.1 Gaussian resource theories

Gaussian resource theories are QRTs further restricted to Gaussian states and operations. For examples, the Gaussian resource theory of entanglement is the QRT of entanglement restricted to those systems described by Gaussian states with allowed operations given by Gaussian LOCC. Gaussian resource theories aim at investigating capabilities of the Gaussian platform in various quantum information processing tasks. It is a practically relevant question as Gaussian elements are desirable tools for the physical implementations of such tasks due to their easy accessibility.

Gaussian resource theories introduced in [10] can be reformulated as an instance of multiresource theories (see Section 3.4). More specifically, a Gaussian resource theory of resource $f$ can be seen as a QRT with two resources: non-Gaussianity and resource $f$. Each of these two resources forms a separate resource theory. We denote the QRT of resource $f$ by $\mathcal{R}_{f}=\left(\mathcal{F}_{f}, \mathbb{O}_{f}\right)$ and the QRT of non-Gaussianity by $\mathcal{R}_{G}=\left(\mathcal{F}_{G}, \mathbb{O}_{G}\right) .{ }^{1}$ For example, if we consider the Gaussian resource theory of entanglement, then $\mathcal{R}_{f}$ is the QRT of entanglement with separable states as $\mathcal{F}_{f}$ and LOCC as $\mathbb{O}_{f}$. For the QRT of non-Gaussianity $\mathcal{R}_{G}$, the set of free states $\mathcal{F}_{G}$ is the set of Gaussian states, which we denote by $\mathcal{G}$, and the allowed operations $\mathbb{O}_{G}$ are all Gaussian maps introduced in Section 2.3. The Gaussian resource theory $\mathcal{R}_{f}^{G}$ is then the multi-resource theory built from the set $\left\{\mathcal{R}_{f}, \mathcal{R}_{G}\right\}$. The new allowed operations, denoted by $\mathbb{O}_{f}^{G}$, are defined as the interception between $\mathbb{O}_{f}$ and $\mathbb{O}_{G}$ following Eq. (3.7):

$$
\mathbb{O}_{f}^{G}:=\mathbb{O}_{f} \cap \mathbb{O}_{G} .
$$

We call these operations Gaussian allowed operations. We can also construct the set of new free states $\mathcal{F}_{f}^{G}$ simply by taking the interception between $\mathcal{F}_{f}$ and $\mathcal{F}_{G}$ (see Figure 4.1(a)). Therefore, we define $\mathcal{F}_{f}^{G}:=\mathcal{F}_{f} \cap \mathcal{F}_{G}$ as the set of free states consistent with $\mathbb{O}_{f}^{G}$ with respect to the golden rule of QRTs and call these states free Gaussian states.

The above construction of Gaussian resource theories are general so that it can be applied to any $\mathcal{R}_{f}$. However, to employ convenient mathematical results developed in [10], in this chapter,

[^13]we only consider Gaussian resource theories developed with $\mathcal{R}_{f}$ whose set of free states $\mathcal{F}_{f}$ satisfies the following assumptions:

1. The set $\mathcal{F}_{f}$ is closed under tensor products, partial traces, and permutations of subsystems.
2. The set $\mathcal{F}_{f}$ is convex.
3. The set $\mathcal{F}_{f}$ is norm-closed. ${ }^{2}$
4. The set $\mathcal{F}_{f}$ is invariant under displacement operations.

1-3 are standard assumptions in QRTs, but 4 is additionally introduced for the Gaussian setting. While there exist some exceptions such as the QRT of quantum thermodynamics [35] which fails to satisfy 4, most QRTs studied in the literature satisfy these assumptions.

As we have seen in Chapter 2, any Gaussian state can be fully represented by its first two statistical moments $(\overline{\mathbf{r}}, \sigma)$ of the canonical operators. Moreover, Assumption 4 implies that the first moment $\overline{\mathbf{r}}$ does not contribute to the amount of resources present in a state. Thus, we assume without loss of generality that $\overline{\mathbf{r}}=0$ from now on. We can now characterise the amount of resources in an arbitrary Gaussian state solely by its covariance matrix (CM) $\sigma$. It is therefore convenient to have an alternative representation of $\mathcal{F}_{f}^{G}$, which is the set of the CMs of all free Gaussian states and denoted by $\boldsymbol{F}_{f}^{G}$ - an arbitrary Gaussian state $\rho_{G}$ is a free state if its $\mathrm{CM} \sigma$ is in $\boldsymbol{F}_{f}^{G}$.

### 4.1.1 No-go theorem for Gaussian resource distillation

In [10], the authors derived the no-go theorem for Gaussian resource distillation within the formalism of Gaussian resource theories. Since this no-go theorem plays an important role in this chapter, we introduce it in more details in this section. In a Gaussian resource theory $\mathcal{R}_{f}^{G}$, one can show that $\boldsymbol{F}_{f}^{G}$, the set of free CMs, has a property called upward closedness. That is, for any $\mathrm{CMs} \sigma$ and $\tau$, if $\sigma \in \boldsymbol{F}_{f}^{G}$ and $\tau \geq \sigma$ with respect to the matrix inequality, then $\tau \in \boldsymbol{F}_{f}^{G}$. This property sometimes results in the existence of the threshold matrix $\sigma_{\mathrm{th}} \in \boldsymbol{F}_{f}^{G}$ such that for

[^14]

Figure 4.1: The geometrical illustration of the state spaces. (a) In Gaussian resource theories, the set of free states is the interception of $\mathcal{F}_{f}$, the set of free states for $\mathcal{R}_{f}$, and the set of Gaussian states $\mathcal{G}$ which is the set of free states for $\mathcal{R}_{G} ; \mathcal{F}_{f}^{G}=\mathcal{F}_{f} \cap \mathcal{F}_{G}$. (b) In convex Gaussian resource theories, the new set of free states is the convex hull of $\mathcal{F}_{f}^{G}$, i.e., $\mathcal{F}_{f}^{C}=\operatorname{conv}\left(\mathcal{F}_{f}^{G}\right)=\operatorname{conv}\left(\mathcal{F}_{f} \cap \mathcal{F}_{G}\right)$.
any $\mathrm{CM} \sigma$, if $\sigma \geq \sigma_{\text {th }}$ then $\sigma \in \boldsymbol{F}_{f}^{G}$. For example, in Gaussian entanglement theory, it is known that an $N$-mode Gaussian state with $\mathrm{CM} \sigma$ is a separable state (free state) if $\sigma \geq \mathbb{1}$ [13].

By exploiting the upward closedness of $\boldsymbol{F}_{f}^{G}$, one can define a resource measure for the Gaussian resource theory as

$$
\begin{equation*}
\kappa\left(\rho^{G}[\sigma]\right):=\min \left\{t \geq 1 \mid t \sigma \in \boldsymbol{F}_{f}^{G}\right\} \tag{4.1}
\end{equation*}
$$

where $\rho^{G}[\sigma]$ denotes the Gaussian state with $\mathrm{CM} \sigma$. This measure is a resource monotone for any Gaussian resource theory satisfying Assumptions 1-4 (see [10, Appendix B]) and also faithful in the sense that $\kappa\left(\rho^{G}\right)=1$ if and only if $\rho^{G} \in \mathcal{F}_{f}^{G}$. Furthermore, it has a property that

$$
\begin{equation*}
\kappa\left(\rho_{A}^{G} \otimes \rho_{B}^{G}\right)=\max \left\{\kappa\left(\rho_{A}^{G}\right), \kappa\left(\rho_{B}^{G}\right)\right\} \tag{4.2}
\end{equation*}
$$

for any Gaussian states $\rho_{A}^{G}, \rho_{B}^{G} \in \mathcal{G}$. This special property is often called the tensorisation property in classical probability theory [70] and plays the main role in the proof of the following no-go theorem for Gaussian resource distillation.

Proposition 3. (No-go theorem for Gaussian resource distillation [10, Theorem 1]) Consider a Gaussian resource theory satisfying Assumptions 1-4 and two Gaussian states $\rho^{G}, \gamma^{G} \in \mathcal{G}$. If $\kappa\left(\rho^{G}\right)<\kappa\left(\gamma^{G}\right)$, there does not exist $\Gamma \in \mathbb{O}_{f}^{G}$ such that $\Gamma\left(\left(\rho^{G}\right)^{\otimes N}\right)=\gamma^{G}$ for any $N \in \mathbb{N}$.

It is worth pointing out that Proposition 3 is shown for $\mathbb{O}_{f}^{G}$ including all completely positive
and trace non-increasing Gaussian allowed operations. Therefore, Proposition 3 rules out the possibility of not only deterministic resource distillations but also probabilistic ones. It is a crucial limitation as we cannot generate a more resourceful state even with small probability and as many copies of the initial state as we want.

### 4.2 Convex Gaussian resource theories

An important feature of the set of Gaussian states that we focus in this section is its non-convexity. In general, classically mixing two arbitrary Gaussian states results in a non-Gaussian state. Although working with non-Gaussian mixtures of Gaussian states can be more mathematically demanding than with Gaussian states, it is not difficult to create these states in practice using probabilistic Gaussian operations. For example, applying different Gaussian operations to a Gaussian state depending on the outcome of measurements performed on other modes can generate a non-Gaussian convex mixture of Gaussian states. Including non-Gaussian mixtures in free states is therefore well-motivated from an operational point of view. It is also worth remarking that this inclusion has been already discussed in the context of the QRT of non-Gaussianity in [71], and conditional Gaussian operations were shown to be useful for distillation of non-Gaussianity, which further motivated us to involve them for distillation of general resources.

In this section, we introduce convex Gaussian resource theories which take into account the easy accessibility of non-Gaussian convex mixtures of Gaussian states. We develop these theories based on Gaussian resource theories studied in the last section. We first introduce a new important set of quantum states in CV quantum systems, which is the convex hull ${ }^{3}$ of Gaussian states given by

$$
\operatorname{conv}(\mathcal{G})=\left\{\int \mathrm{d} \lambda p(\lambda) \rho_{\lambda}^{G} \mid \rho_{\lambda}^{G} \in \mathcal{G}, p(\lambda) \geq 0 \forall \lambda, \int \mathrm{~d} \lambda p(\lambda)=1\right\} .
$$

This set contains all Gaussian states as well as non-Gaussian convex mixtures of Gaussian states, i.e., $\mathcal{G} \subset \operatorname{conv}(\mathcal{G})$.

[^15]To define the convex Gaussian resource theory of resource $f$, we start from the corresponding Gaussian resource theory of resource $f, \mathcal{R}_{f}^{G}=\left(\mathcal{F}_{f}^{G}, \mathbb{O}_{f}^{G}\right)$. Then, the set of free states for the convex Gaussian resource theory of resource $f$, denoted by $\mathcal{F}_{f}^{C}$, is given by the convex hull of $\mathcal{F}_{f}^{G}$. i.e., $\mathcal{F}_{f}^{C}:=\operatorname{conv}\left(\mathcal{F}_{f}^{G}\right)$, as depicted in Figure 4.1(b). The allowed operations for the convex Gaussian resource theory, which we denote by $\mathbb{O}_{f}^{C}$, are defined as follows.

Definition 4. The allowed operations $\mathbb{O}_{f}^{C}$ of the convex Gaussian resource theory constructed from a Gaussian resource theory $\mathcal{R}_{f}^{G}=\left(\mathcal{F}_{f}^{G}, \mathbb{O}_{f}^{G}\right)$ are composed by the following two kinds of operations:

1. Appending an ancillary system described by a free state $\rho \in \mathcal{F}_{f}^{C}$.
2. Applying a mixture of free Gaussian operations conditioned on the outcome of a homodyne measurement; namely, a map from $\mathcal{H}_{A B}$ to $\mathcal{H}_{A}$ of the form

$$
\Gamma_{A B \mapsto A}\left(\rho_{A B}\right)=\int \mathrm{d} q \Phi_{A}^{q} \otimes M_{B}^{q}\left(\rho_{A B}\right),
$$

where $\Phi^{q} \in \mathbb{O}_{f}^{G}$ for all $q$, and $M^{q}=\langle q| \cdot|q\rangle$ is a selective homodyne measurement.

Note that we made an assumption that the partial selective homodyne measurement $M^{q}$ is an allowed operation of $\mathcal{R}_{f}^{G}$, i.e., $M^{q} \in \mathbb{O}_{f}^{G}$. For resource theories that we consider explicitly in this chapter, namely those of entanglement and squeezing, we show that this assumption is true in Appendix A.1. Physically, this means that applying the homodyne measurement to a part of the system cannot make the rest of the system more squeezed or entangled. Definition 4-2 also includes post-selection when the partial homodyne measurement on the input state has zero-probabilities for some values/ranges of $q$.

The convex Gaussian resource theory, denoted by $\mathcal{R}_{f}^{C}$, is then defined as $\mathcal{R}_{f}^{C}=\left(\mathcal{R}_{f}^{C}, \mathbb{O}_{f}^{C}\right)$. It is not difficult to show that $\mathcal{F}_{f}^{C}$ and $\mathbb{O}_{f}^{C}$ defined above fulfill the golden rule of QRTs. Indeed, a given generic free state

$$
\rho_{A}=\int \mathrm{d} \lambda p(\lambda) \rho_{A}^{G}(\lambda),
$$

where $\rho_{A}^{G}(\lambda) \in \mathcal{F}_{f}^{G}\left(\mathcal{H}_{A}\right)$ for all $\lambda$, appending a free state $\gamma_{B} \in \mathcal{F}_{f}^{G}\left(\mathcal{H}_{B}\right)$ (operation described in Definition 4-1) to $\rho_{A}$ does not map $\rho_{A}$ outside the set of free states $\mathcal{F}_{f}^{C}\left(\mathcal{H}_{A} \otimes \mathcal{H}_{B}\right)$. Also, applying a conditional operations described in Definition 4-2 to a free state $\rho_{A B}$ results in

$$
\begin{equation*}
\Gamma_{A B \mapsto A}\left(\rho_{A B}\right)=\int \mathrm{d} q \mathrm{~d} \lambda p(\lambda) \Phi_{A}^{q} \otimes M_{B}^{q}\left(\rho_{A B}^{G}(\lambda)\right) \tag{4.3}
\end{equation*}
$$

Since both $\Phi^{q}$ and $M^{q}$ are elements of $\mathbb{O}_{f}^{G}$, each state $\rho_{A B}^{G}(\lambda) \in \mathcal{F}_{f}^{G}\left(\mathcal{H}_{A B}\right)$ is mapped to a state in $\mathcal{F}_{f}^{G}\left(\mathcal{H}_{A}\right)$. Then, the state in Eq. (4.3) is a mixture of free states, which again belongs to $\mathcal{F}_{f}^{C}\left(\mathcal{H}_{A}\right)$.

We are now in a position where we can investigate resource distillation when new tools, classical randomness and post-selection, are allowed. The main question we aim to address hereafter is whether distillation is still impossible in convex Gaussian resource theories; i.e., whether Proposition 3 still holds for convex Gaussian resource theories. This is equivalent to asking whether we can distill resource $f$ from a resourceful state $\rho$ in $\operatorname{conv}(\mathcal{G}) \backslash \mathcal{F}_{f}^{C}$ using allowed operations in $\mathbb{O}_{f}^{C}$. It is worth remarking that the the other distillation scenario, distillation of non-Gaussianity from a state $\rho$ in $\mathcal{F}_{f} \backslash \mathcal{F}_{f}^{C}$ only using allowed operations, has already been studied in [71]. The authors showed that one can distill non-Gaussianity from any non-Gaussian state by filtering out the Gaussian part using post-selection. Their distillation protocol consists of a beam splitter and a homodyne measurement, which are allowed operations in most convex Gaussian resource theories.

### 4.3 Limitations on resource distillation in convex Gaussian resource theories

In this Section, we study resource distillation in convex Gaussian resource theories. We first need an appropriate resource measure for the theories. The resource measure $\kappa$ defined in Eq. (4.1) cannot be directly applied to convex Gaussian resource theories; a generic state considered in convex Gaussian resource theories is an element of $\operatorname{conv}(\mathcal{G})$, which is in general non-Gaussian, and $\kappa$ in Eq. (4.1) is not a valid resource monotone for non-Gaussian states. Therefore, we
define the following new resource measure for the convex hull of Gaussian states, $\operatorname{conv}(\mathcal{G})$, which serves as the state space in convex Gaussian resource theories: for any $\rho \in \operatorname{conv}(\mathcal{G})$, we define

$$
\begin{equation*}
\tilde{\kappa}(\rho):=\inf _{\substack{\left\{p(\lambda), \rho^{G}(\lambda)\right\}_{\lambda} \\ \text { s.t. } \rho=\int \mathrm{d} \lambda p(\lambda) \rho^{G}(\lambda)}} \int \mathrm{d} \lambda p(\lambda) \kappa\left(\rho^{G}(\lambda)\right) \tag{4.4}
\end{equation*}
$$

where the infimum is over all possible Gaussian decompositions of the state $\rho$ including discrete ones. This measure is the convex roof extension [72, 73] of the monotone $\kappa$ in Eq. (4.1). We can show that $\tilde{\kappa}$ is a valid resource monotone satisfying a few desirable properties for resource measures (see Section 3.3).

Lemma 4. The resource measure $\tilde{\kappa}$ defined in Eq. (4.4) has the following properties:

1. It coincides with the resource measure $\kappa$ for Gaussian states:

$$
\tilde{\kappa}\left(\rho^{G}\right)=\kappa\left(\rho^{G}\right) \quad \forall \rho^{G} \in \mathcal{G}
$$

2. It is convex:

$$
\tilde{\kappa}\left(\int d \lambda p(\lambda) \rho(\lambda)\right) \leq \int d \lambda p(\lambda) \tilde{\kappa}(\rho(\lambda))
$$

for any probability distribution $p(\lambda)$ and set of states $\{\rho(\lambda)\}_{\lambda}$ in $\operatorname{conv}(\mathcal{G})$.
3. It inherits the tensorisation property from $\kappa$ when applied to Gaussian states:

$$
\tilde{\kappa}\left(\rho^{G} \otimes \gamma^{G}\right)=\max \left\{\tilde{\kappa}\left(\rho^{G}\right), \tilde{\kappa}\left(\gamma^{G}\right)\right\} \quad \forall \rho^{G}, \gamma^{G} \in \mathcal{G} .
$$

4. It is monotonic under the allowed operations $\mathbb{O}_{f}^{C}$ :

$$
\tilde{\kappa}(\rho) \geq \tilde{\kappa}(\Gamma(\rho)) \quad \forall \Gamma \in \mathbb{O}_{f}^{C}
$$

5. It is faithful:

$$
\tilde{\kappa}(\rho)=1 \Longleftrightarrow \rho \in \mathcal{F}_{f}^{C}
$$

Proof. 1. Consider a Gaussian state $\rho^{G}$ with covariance matrix $\sigma$. Any Gaussian decomposition of a Gaussian state has the form

$$
\rho^{G}=\int \mathrm{d} \lambda p^{G}(\lambda) \rho^{G}(\lambda)
$$

where the probability distribution $p^{G}(\lambda)$ is Gaussian (if the state is mixed) or the Dirac delta distribution (if the state is pure). Furthermore, each state $\rho^{G}(\lambda)$ has a covariance matrix $\sigma(\lambda)$ such that $\sigma(\lambda) \leq \sigma$ for all $\lambda$ [74, Lemma 3]. This implies that $\kappa\left(\rho^{G}\right) \leq \kappa\left(\rho^{G}(\lambda)\right) \forall \lambda$ for any Gaussian decomposition of $\rho^{G}$. Given the optimal decomposition $\left\{p_{\text {opt }}(\lambda), \rho_{\text {opt }}^{G}(\lambda)\right\}_{\lambda}$ of $\rho^{G}$ with respect to the measure $\tilde{\kappa}$, we then have

$$
\begin{aligned}
\tilde{\kappa}\left(\rho^{G}\right) & =\int \mathrm{d} \lambda p_{\text {opt }}(\lambda) \kappa\left(\rho_{\text {opt }}^{G}(\lambda)\right) \\
& \geq \int \mathrm{d} \lambda p_{\text {opt }}(\lambda) \kappa\left(\rho^{G}\right)=\kappa\left(\rho^{G}\right) .
\end{aligned}
$$

As $\rho^{G}$ itself is a Gaussian decomposition of $\rho^{G}$, the converse also holds; $\tilde{\kappa}\left(\rho^{G}\right) \leq \kappa\left(\rho^{G}\right)$. Thus, $\tilde{\kappa}\left(\rho^{G}\right)=\kappa\left(\rho^{G}\right)$.
2. Consider a generic state $\rho=\int \mathrm{d} \lambda p(\lambda) \rho(\lambda)$; each $\rho(\lambda)$ does not need to be Gaussian. Let us denote $\left\{q_{\text {opt }}(\lambda, \mu), \gamma_{\text {opt }}^{G}(\lambda, \mu)\right\}_{\mu}$ as the optimal Gaussian decomposition of each $\rho(\lambda)$ such that

$$
\tilde{\kappa}(\rho(\lambda))=\int \mathrm{d} \mu q_{\mathrm{opt}}(\lambda, \mu) \kappa\left(\gamma_{\mathrm{opt}}^{G}(\lambda, \mu)\right) .
$$

Then, the set $\left\{p(\lambda) q_{\text {opt }}(\lambda, \mu), \gamma_{\text {opt }}^{G}(\lambda, \mu)\right\}_{\lambda, \mu}$ is a Gaussian decomposition of $\rho$ as well, i.e.,

$$
\iint \mathrm{d} \lambda \mathrm{~d} \mu p(\lambda) q_{\mathrm{opt}}(\lambda, \mu) \gamma_{\mathrm{opt}}^{G}(\lambda, \mu)=\int \mathrm{d} \lambda p(\lambda) \rho(\lambda)=\rho .
$$

It then follows that

$$
\begin{aligned}
\tilde{\kappa}(\rho) & \leq \iint \mathrm{d} \lambda \mathrm{~d} \mu p(\lambda) q_{\mathrm{opt}}(\lambda, \mu) \kappa\left(\gamma_{\mathrm{opt}}^{G}(\lambda, \mu)\right) \\
& =\int \mathrm{d} \lambda p(\lambda) \tilde{\kappa}(\rho(\lambda))
\end{aligned}
$$

3. The third property can be easily shown using Property 1 and the tensorisation property of the Gaussian resource measure $\kappa$ in Eq. (4.2). For any Gaussian states $\rho^{G}, \gamma^{G} \in \mathcal{G}$, we have

$$
\begin{array}{rlrl}
\tilde{\kappa}\left(\rho^{G} \otimes \gamma^{G}\right) & =\kappa\left(\rho^{G} \otimes \gamma^{G}\right) \quad(\because \text { Property } 1) \\
& =\max \left\{\kappa\left(\rho^{G}\right), \kappa\left(\gamma^{G}\right)\right\} & & (\because \text { Eq. }(4.2)) \\
& =\max \left\{\tilde{\kappa}\left(\rho^{G}\right), \tilde{\kappa}\left(\gamma^{G}\right)\right\} & & (\because \text { Property } 1) .
\end{array}
$$

4. Firstly, let us show that the monotonicity under appending an ancillary system in a free state (Definition 4-1). Consider a generic state $\rho \in \operatorname{conv}(\mathcal{G})$ with optimal Gaussian decomposition $\left\{p_{\text {opt }}(\lambda), \rho_{\text {opt }}^{G}(\lambda)\right\}_{\lambda}$ with respect to $\tilde{\kappa}$ and a free state $\gamma_{\text {free }}=\int \mathrm{d} \mu q(\mu) \gamma_{\text {free }}^{G}(\mu) \in \mathcal{F}_{f}^{C 4}$. We have

$$
\begin{aligned}
\tilde{\kappa}\left(\rho \otimes \gamma_{\text {free }}\right) & =\tilde{\kappa}\left(\int \mathrm{d} \lambda p_{\text {opt }}(\lambda) \rho_{\text {opt }}^{G}(\lambda) \otimes \int \mathrm{d} \mu q(\mu) \gamma_{\text {free }}^{G}(\mu)\right) \\
& =\tilde{\kappa}\left(\iint \mathrm{d} \lambda \mathrm{~d} \mu p_{\text {opt }}(\lambda) q(\mu)\left(\rho_{\text {opt }}^{G}(\lambda) \otimes \gamma_{\text {free }}^{G}(\mu)\right)\right) \\
& \leq \iint \mathrm{d} \lambda \mathrm{~d} \mu p_{\text {opt }}(\lambda) q(\mu) \tilde{\kappa}\left(\rho_{\text {opt }}^{G}(\lambda) \otimes \gamma_{\text {free }}^{G}(\mu)\right) \quad(\because \text { Property 2) } \\
& =\iint \mathrm{d} \lambda \mathrm{~d} \mu p_{\text {opt }}(\lambda) q(\mu) \max \left\{\tilde{\kappa}\left(\rho_{\text {opt }}^{G}(\lambda)\right), \tilde{\kappa}\left(\gamma_{\text {free }}^{G}(\mu)\right)\right\} \quad(\because \text { Property 3) } \\
& =\int \mathrm{d} \lambda p_{\text {opt }}(\lambda) \tilde{\kappa}\left(\rho_{\text {opt }}^{G}(\lambda)\right)=\tilde{\kappa}(\rho) .
\end{aligned}
$$

For the monotonicity under mixtures of conditional Gaussian free operations (Definition 4-2), let us consider again a generic state $\rho_{A B} \in \operatorname{conv}(\mathcal{G})$ with optimal Gaussian decomposition $\left\{p_{\text {opt }}(\lambda), \rho_{A B}^{G}(\lambda)\right\}_{\lambda}$ with respect to $\tilde{\kappa}$, and an operations $\Gamma_{A B \mapsto A}(\cdot)=\int \mathrm{d} q \Phi_{A}^{q} \otimes M_{B}^{q}(\cdot)$ with

[^16]$\Phi_{A}^{q} \in \mathbb{O}_{f}^{G} \forall \lambda$. We have
\[

$$
\begin{aligned}
\tilde{\kappa}\left(\Gamma_{A B \mapsto A}\left(\rho_{A B}\right)\right) & \leq \iint \mathrm{d} \lambda \mathrm{~d} q p_{\text {opt }}(\lambda) \tilde{\kappa}\left(\Phi_{A}^{q} \otimes M_{B}^{q}\left(\rho_{A B}^{G}(\lambda)\right)\right) \quad(\because \text { Property 2) } \\
& =\iint \mathrm{d} \lambda \mathrm{~d} q p_{\text {opt }}(\lambda) \tilde{\kappa}\left(\Phi_{A}^{q}\left(\mathbb{1}_{A} \otimes M_{B}^{q}\left(\rho_{A B}^{G}(\lambda)\right)\right)\right)
\end{aligned}
$$
\]

Note that the partial selective measurement $M^{q}$ is a Gaussian measurement and preserves the Gaussianity. Then, we obtain

$$
\begin{aligned}
\tilde{\kappa}\left(\Gamma_{A B \mapsto A}\left(\rho_{A B}\right)\right) & =\iint \mathrm{d} \lambda \mathrm{~d} q p_{\text {opt }}(\lambda) \kappa\left(\Phi_{A}^{q}\left(\mathbb{1}_{A} \otimes M_{B}^{q}\left(\rho_{A B}^{G}(\lambda)\right)\right)\right) \quad(\because \text { Property 1) } \\
& \leq \iint \mathrm{d} \lambda \mathrm{~d} q p_{\text {opt }}(\lambda) \kappa\left(\rho_{A B}^{G}(\lambda)\right) \quad\left(\because \text { monotonicity of } \kappa \text { under } \mathbb{O}_{f}^{G}\right) \\
& =\int \mathrm{d} \lambda p_{\text {opt }}(\lambda) \kappa\left(\rho_{A B}^{G}(\lambda)\right)=\tilde{\kappa}\left(\rho_{A B}\right) .
\end{aligned}
$$

5. Let us start from the direction $\tilde{\kappa}(\rho)=1 \Longrightarrow \rho \in \mathcal{F}_{f}^{C}$. Consider a state $\rho \in \operatorname{conv}(\mathcal{G})$ with optimal Gaussian decomposition $\left\{p_{\text {opt }}(\lambda), \rho_{\text {opt }}^{G}(\lambda)\right\}_{\lambda}$ with respect to $\tilde{\kappa}$, which satisfies

$$
\tilde{\kappa}(\rho)=\int \mathrm{d} \lambda p_{\mathrm{opt}}(\lambda) \kappa\left(\rho_{\mathrm{opt}}^{G}(\lambda)\right)=1 .
$$

Since $\kappa(\rho) \geq 1$ for all states $\rho \in \mathcal{G}$ by definition, $\tilde{\kappa}(\rho)=1$ implies that $\kappa\left(\rho_{\text {opt }}^{G}(\lambda)\right)=1$ for all $\lambda$. Due to the faithfulness of $\kappa$, it follows that $\rho_{\mathrm{opt}}^{G}(\lambda) \in \mathcal{F}_{f}^{G}$ for all $\lambda$. Then, $\rho$ is a convex combination of states in $\mathcal{F}_{f}^{G}$, and thus $\rho \in \mathcal{F}_{f}^{C}=\operatorname{conv}\left(\mathcal{F}_{f}^{G}\right)$.

For the other direction, let us assume that $\rho \in \mathcal{F}_{f}^{C}$. Then, by definition of the set of free states, a decomposition of $\rho$ in terms of free Gaussian states in $\mathcal{F}_{f}^{G}$ is possible:

$$
\begin{equation*}
\rho=\int \mathrm{d} \lambda p(\lambda) \rho_{\text {free }}^{G}(\lambda), \tag{4.5}
\end{equation*}
$$

where $\rho_{\text {free }}^{G} \in \mathcal{F}_{f}^{G}$ for all $\lambda$. Since (i) $\kappa\left(\rho_{\text {free }}^{G}(\lambda)\right)=1$ for all $\lambda$, (ii) Eq. (4.5) is one possible Gaussian decomposition of $\rho$, and (ii) $\tilde{\kappa}$ is the infimum of all decompositions, it follows that $\tilde{\kappa}(\rho) \leq 1$. Since the value of $\kappa$ is lower bounded by 1 , we have $\tilde{\kappa}(\rho)=1$.

Using the resource measure $\tilde{\kappa}$ and its properties in Lemma 4, we can prove the following
theorem for resource distillation in convex Gaussian resource theories. This theorem shows that no allowed operation can increase the value of the measure $\tilde{\kappa}$ above that of the most resourceful state composing the initial mixture.

Theorem 5. Let $\Gamma \in \mathbb{O}_{f}^{C}$ and $\rho=\int d \lambda p(\lambda) \rho^{G}(\lambda) \in \operatorname{conv}(\mathcal{G})$. Then, we have for any $N \in \mathbb{N}$ that

$$
\tilde{\kappa}\left(\Gamma\left(\rho^{\otimes N}\right)\right) \leq \tilde{\kappa}\left(\rho_{\max }^{G}\right),
$$

where $\rho_{\max }^{G}$ denotes the most resourceful element in the decomposition $\left\{\rho^{G}(\lambda)\right\}_{\lambda}$ with respect to $\tilde{\kappa}$.

Proof. Due to the monotonicity of $\tilde{\kappa}$ under the allowed operations $\mathbb{O}_{f}^{C}$, it holds that $\tilde{\kappa}\left(\Gamma\left(\rho^{\otimes N}\right)\right) \leq$ $\tilde{\kappa}\left(\rho^{\otimes N}\right)$. For the given decomposition of $\rho$, we have

$$
\rho^{\otimes N}=\int \mathrm{d} \boldsymbol{\lambda} \prod_{i=1}^{N} p\left(\lambda_{i}\right) \bigotimes_{i=1}^{N} \rho^{G}\left(\lambda_{i}\right),
$$

where $\boldsymbol{\lambda}=\left(\lambda_{1}, \ldots, \lambda_{N}\right)$. Then, it follows that

$$
\begin{aligned}
\tilde{\kappa}\left(\rho^{\otimes N}\right) & \leq \int \mathrm{d} \boldsymbol{\lambda} \prod_{i=1}^{N} p\left(\lambda_{i}\right) \tilde{\kappa}\left(\bigotimes_{i=1}^{N} \rho^{G}\left(\lambda_{i}\right)\right) \quad(\because \text { Property } 2) \\
& \leq \int \mathrm{d} \boldsymbol{\lambda} \prod_{i=1}^{N} p\left(\lambda_{i}\right) \tilde{\kappa}\left(\left(\rho_{\max }^{G}\right)^{\otimes N}\right) \\
& =\int \mathrm{d} \boldsymbol{\lambda} \prod_{i=1}^{N} p\left(\lambda_{i}\right) \tilde{\kappa}\left(\rho_{\max }^{G}\right)=\tilde{\kappa}\left(\rho_{\max }^{G}\right) \quad(\because \text { Property } 3) .
\end{aligned}
$$

We can finally state the following corollary which identify a limitation on resource distillation in convex Gaussian resource theories.

Corollary 6. Consider $\rho=\int d \lambda p(\lambda) \rho^{G}(\lambda) \in \operatorname{conv}(\mathcal{G})$ and $\gamma \in \operatorname{conv}(\mathcal{G})$ such that $\tilde{\kappa}\left(\rho_{\max }^{G}\right)<\tilde{\kappa}(\gamma)$ where $\rho_{\max }^{G}$ is the most resourceful element in the decomposition $\left\{\rho^{G}(\lambda)\right\}_{\lambda}$. For any $N \in \mathbb{N}$, there does not exist $\Gamma \in \mathbb{O}_{f}^{C}$ such that $\Gamma\left(\rho^{\otimes N}\right)=\gamma$.

Let us remark that this restriction is only applicable to states in $\operatorname{conv}(\mathcal{G})$, and it is not valid for general CV states. Corollary 6 states that one can never distill more resourceful states than those already present in the mixture in convex Gaussian resource theories, even with multiple copies of the initial state. Possible resource distillation protocols are thus equivalent to identifying and intensifying the resource already present in the mixture. As such, on the contrary to the case of the no-go theorem in Gaussian resource theories (Proposition 3), Corollary 6 leaves a possibility of resource distillation that might be useful in some specific scenarios. For example, it allows the possibilities for error-correction or purification.

In the next section, we show that one can indeed find some simple examples of resource distillation protocols in convex Gaussian resource theories. Particularly, we show that the upper bound on the amount of distillable resources identified in Corollary 6 can be achieved in special cases.

### 4.4 Convex distillation protocols

In this section, we provide explicit examples of possible resource distillation protocols in convex Gaussian resource theories exploiting classical randomness and conditional operations. These examples highlight the difference between the limitations in Gaussian resource theories (Proposition 3) and in convex Gaussian resource theories (Corollary 6). We start with the convex Gaussian resource theory of squeezing and then move on to discuss entanglement.

### 4.4.1 Squeezing distillation

In this section, we explore resource distillation when resource $f$ is squeezing; i.e., we consider resource distillation in the convex Gaussian resource theory of squeezing. Squeezing is often considered as a resource for metrological tasks as we discussed in Section 2.5

## One-shot deterministic case

We firstly look at the simplest scenario. Consider a single-mode squeezed vacuum (SMSV) state with squeezing parameter $r$ along the $\hat{x}$ quadrature, $|0, r\rangle_{\text {1SQ }}$, studied in Section 2.5.

Suppose the system experiences a random displacement noise with probability $p$, so that its state becomes a mixture of a SMSV state and a displaced SMSV state. For simplicity, let us assume that the displacement also occurs along the $\hat{x}$ quadrature. The displaced SMSV state is then $|d / \sqrt{2}, r\rangle_{1 S Q}$, which is the SMSV state displaced by $d$ along the $\hat{x}$ quadrature. The state after the displacement noise can be written as

$$
\begin{equation*}
\rho_{\mathrm{in}}=(1-p)|0, r\rangle_{1 \mathrm{SQ}}\left\langle 0,\left.r\right|_{1 \mathrm{SQ}}+p \mid d / \sqrt{2}, r\right\rangle_{1 \mathrm{SQ}}\left\langle d / \sqrt{2},\left.r\right|_{1 \mathrm{SQ}}\right. \tag{4.6}
\end{equation*}
$$

For a sufficiently large $d$, the state $\rho_{\text {in }}$ does not have any squeezing along the $\hat{x}$ quadrature which can be exploited for a quantum task. Our goal is to distill the hidden squeezing resource from this mixture.

A simple distillation protoco is shown in Figure 4.2(a). ${ }^{5}$ The main idea is to measure the $\hat{x}$ quadrature of the mixture via weak measurement and correct the displacement noise accordingly. To do so, we first need to correlate the main system with a pointer initialised in the vacuum state $|0\rangle$. The two systems interact with each other via a beam splitter with large transmissivity $t=\cos ^{2} \theta$. In this way, the pointer gets some information on whether the system was affected by the displacement noise or not and as a result disturb the system, which results in reducing the squeezing in the system. The larger the transmissivity of the beam splitter, the less the squeezing of the system is compromised, while at the same time the less is learned about the displacement affecting the system.

After the beam splitter, the $\hat{x}$ quadrature of the pointer is measured via homodyne detection. The measurement is described by the dichotomic POVM $\{\Pi, \mathbb{1}-\Pi\}$ where the measurement effect $\Pi$ is defined as

$$
\begin{equation*}
\Pi=\int_{-\infty}^{\Delta} \mathrm{d} x|x\rangle\langle x|, \tag{4.7}
\end{equation*}
$$

where $|x\rangle$ is the position eigenstate, and $\Delta=-\frac{d}{2} \sin \theta$. When the outcome of the POVM refers to $\Pi$, it means that with high probability the main system was displaced along $\hat{x}$, and then we

[^17](a)

(b)


Figure 4.2: (a) Schematic of the deterministic one-shot squeezing distillation protocol. The system, described by the state $\rho_{\text {in }}$, is fed into a beam splitter with transmissivity $t=\cos ^{2} \theta$ together with a pointer system in the vacuum state. After the interaction, a homodyne measurement is performed on the pointer system, and depending on the outcome the main system is displaced or not. (b) Schematic of the probabilistic multi-copy squeezing distillation protocol. The pointer system is now replaced by another copy of $\rho_{\text {in }}$, providing better performance due to the fact that it is a mixture of squeezed states (resourceful states). The system and the pointer interact via a 50:50 beam splitter, and a homodyne measurement is performed on the pointer. If the outcome falls within a desired range, the protocol is successful, otherwise it is aborted. The protocol can be iterated multiple times, with the output system of one iteration being the input state of the next one.
can correct the noise by performing a displacement operations of $-d \cos \theta$ on the system along the same quadrature. On the other hand, if the outcome of the POVM refers to $\mathbb{1}-\Pi$, with high probability the system was not affected by the noise, and thus we do not need to act on it.

Numerical simulation results, showing the performance of this protocol, are given in Figure 4.3. We characterise the performance in terms of the fidelity to the SMSV state $|0, r\rangle_{1 S Q}$ as a function of the distinguishability between the elements in the initial mixture; see Figure 4.3(a). The distinguishability can be quantified by the ratio $d / \sigma$, where $d$ is the displacement along the relevant quadrature while $\sigma=e^{-2 r}$ is the variance of $|0, r\rangle_{1 \mathrm{SQ}}$ along the same quadrature. When $d \approx \sigma$, the protocol is not able to distinguish the two elements in $\rho_{\text {in }}$ due to the large overlap, and therefore the outcome has a lower fidelity to the target state $|0, r\rangle_{1 S Q}$ than the initial one (dashed grey line). When instead $d \gg \sigma$, the protocol is able to distinguish the two elements, and the displacement noise can be effectively corrected - the final fidelity is closer to 1 than the


Figure 4.3: Numerical simulation of the deterministic one-shot squeezing distillation protocol, for two different beam splitter settings with transmissivity $t=90 \%$ and $65 \%$. (a) The fidelity between the target state $\rho_{\mathrm{sq}}=|0, r\rangle_{1 \mathrm{SQ}}\left\langle 0,\left.r\right|_{\text {1SQ }}\right.$ and the final state of the protocol as a function of $d / \sigma$, which is the ratio between the amount of the random displacement noise $d$ and the $\hat{x}$-variance $\sigma=e^{-2 r}$. While the low transmissivity $t=65 \%$ allows for correcting the displacement noise when $d / \sigma$ is small, it also degrades squeezing. Thus, it achieves a lower maximum fidelity for the high $d / \sigma$ region than the one achieved by the higher transmissivity $t=90 \%$. (b) The variance in the $\hat{x}$ quadrature for the final state of the protocol is shown as a function of $d / \sigma$. The protocol is regarded successful when the final state has a variance lower than the best free state, the vacuum state. For the simulation, we set the squeezing parameter $r=0.7$ and the probability of random displacement $p=0.5$.
initial state. We can also observe that the final fidelity has a higher fidelity to the target state than the best free state, the vacuum state (dotted grey line), in a sense that it is the closest free state to the target state $|0, r\rangle_{1 \mathrm{SQ}}$. With larger transmissivity it takes longer to reach higher fidelity as $d / \sigma$ increases, but higher fidelity is reached in the $d \gg \sigma$ region. This is because less information is extracted by the pointer from the initial state after the beam splitter with large transmissivity, which makes it harder to distinguish the two elements but preserve more squeezing.

Additionally, we compute the variance along the displaced quadrature (the $\hat{x}$-quadrature in
this example) in Figure 4.3(b) as the variance is a relevant measure for the QRT of squeezing [76]. We find that our protocol can indeed reduce the variance of the initial mixture when $d \gg \sigma$, reaching values well below the ones of the vacuum, which is the free state with minimum variance in all quadrature. Importantly, we can observe that the final state can never have a smaller variance than the maximally squeezed element in the initial mixture ('optimal squeezing' in the graph), in accordance with the upper bound of distillable resource derived in Corollary 6.

We remark that the variance is not a valid monotone for the convex Gaussian resource theory of squeezing when the allowed operations are of the form given in Definition 4. This is because conditional operations can reduce this measure, and thus it does not satisfy the monotonicity under the allowed operations. However, we show in Appendix A. 2 that variance is a meaningful monotone under a slightly different set of allowed operations, the set of all convex mixtures of Gaussian operations, which is still natural over the convex hull of Gaussian states.

It is worth pointing out that we do not use the measure $\tilde{\kappa}$ in the analysis of the numerical results, because the initial state and the target state $|0, r\rangle_{1 S Q}$ have the same resource with respect to $\tilde{\kappa}$. Indeed, this measure is useful to derive the general limitation on resource distillation in convex Gaussian resource theories, but it is not as good at reflecting the practical usefulness of the resource contained in a state. Note that the presented protocol is deterministic and requires only one copy of the initial state $\rho_{\text {in }}$.

## Multi-copy probabilistic case

As we have seen in Section 3.2.2, having access to multiple copies of the initial state often provides some benefit in QRTs. However, we have seen in [10] that it is in general not the case for Gaussian resource theories. In the following example, we explore whether the same is true for the convex Gaussian resource theory of squeezing as well. We show that the answer is negative; with multiple copies of the initial state, it is possible to get information on the displacement noise without compromising squeezing in the system, thus demonstrating some advantage over the single-shot case.

The multi-copy squeezing distillation protocol, depicted in Figure 4.2(b), is similar to the one in the one-shot case, but we now allow to use $N$ independent and identically distributed (i.i.d.)
copies of the initial state $\rho_{\text {in }}$, given in Eq. (4.6). In the first step, two copies of the system are considered and fed into a $50: 50$ beam splitter, i.e., a beam splitter with transmissivity $t=1 / 2$. Since both states are composed of a mixture of squeezed states with same parameter $r$, the beam splitter does not degrade the amount of squeezing contained in the two copies. In other words, every element in the initial mixture has the same CM, and thus mixing with each other via the beam splitter does not alter their CMs. After the beam splitter, the two systems are classically correlated, and the global state would look like

$$
\begin{align*}
\rho_{12}^{\prime}= & (1-p)^{2}|0, r\rangle\left\langle 0,\left.r\right|_{1} \otimes \mid 0, r\right\rangle\left\langle 0,\left.r\right|_{2}\right.  \tag{4.8}\\
& +p(1-p)|d / 2, r\rangle\left\langle d / 2,\left.r\right|_{1} \otimes \mid d / 2, r\right\rangle\left\langle d / 2,\left.r\right|_{2}\right. \\
& +p(1-p)|-d / 2, r\rangle\left\langle-d / 2,\left.r\right|_{1} \otimes \mid d / 2, r\right\rangle\left\langle d / 2,\left.r\right|_{2}\right. \\
& +(1-p)^{2}|0, r\rangle\left\langle 0,\left.r\right|_{1} \otimes \mid d, r\right\rangle\left\langle d,\left.r\right|_{2},\right.
\end{align*}
$$

where we omitted the subscription '1SQ' of the SMSV state $|\alpha, r\rangle_{1 S Q}$ for the concise notation. In Eq. (4.8), we can observe that by measuring the quadrature of the second system, we can infer the one of the first system.

The following step consists of a homodyne measurement on the second system, which is playing the role of the pointer in this multi-copy case, followed by post-selection. This measurement is composed by a dichotomic POVM $\{\Pi, \mathbb{1}-\Pi\}$ analogous to the one described in the previous section. In particular, the effect is equal to the one in Eq. (4.7) with an exception that the cut-off region is now given by $\left(-\Delta^{\prime}, \Delta^{\prime}\right]$ where $\Delta^{\prime}=e^{-r}$. If the outcome of the measurement is associated with the effect $\Pi$, the protocol is successful, and with high probability the remaining system is described by $|0, r\rangle_{1 S Q}$. Otherwise, the protocol fails and the remaining system is discarded. The protocol can be iterated over multiple copies, obtaining the final state that approaches the target state $|0, r\rangle_{1 S Q}$ while the success-probability decreases exponentially in the number of iterations performed. An interesting open question left to address is whether the displacement noise can be corrected for all values of $d$.

The numerical simulation is shown in Figure 4.4. Again, we plot the fidelity for $|0, r\rangle_{1 S Q}$ as well as the variance of the $\hat{x}$ quadrature of the final state as a function of $d / \sigma$ for different


Figure 4.4: Numerical simulation of the multi-copy squeezing distillation protocol for different numbers of copies $N$ of the initial state. (a) The fidelity between the target state $\rho_{\mathrm{sq}}=$ $|0, r\rangle_{1 \mathrm{SQ}}\left(0,\left.r\right|_{1 \mathrm{SQ}}\right.$ and the final state of the protocol as a function of $d / \sigma$, where $\sigma=e^{-2 r}$ is the $\hat{x}$-variance of $|0, r\rangle_{\text {SQ }}$. The success probability $p_{\text {succ }}$ of the protocol is shown in the inset. In most regions, more copies lead to more successful performance even though the success probability decreases significantly as $N$ increases. However, for smaller $d / \sigma$, using multiple copies of the initial state does not help to improve the performance. (b) The variance of the $\hat{x}$ quadrature for the final state as a function of $d / \sigma$. The protocol is regarded successful when the final state has a smaller variance than the best free state, the vacuum state (dotted grey line). The 'optimal squeezing' (dashed-dotted grey line) is the $\hat{x}$-variance of $|0, r\rangle_{1 S Q}$. For the simulation, we set the squeezing parameter $r=0.7$ and the probability of random displacement to $p=0.5$.
numbers of copies $N$. The fidelity graph also shows the success probability $p_{\text {succ }}$ in the inset. We observe that the performance is significantly improved compared to the one-shot case shown in Figure 4.3 in two ways: (i) The fidelity is higher than that of the vacuum state (dotted grey line) in most regions of $d / \sigma$, and (ii) the final state can reach fidelity 1 for large $d / \sigma$, which was not possible in the one-shot case. The point (ii) means that the final state is extremely close to $|0, r\rangle_{1 S Q}$, which is the most resourceful element in the initial state $\rho_{\mathrm{in}}$, and thus proves that the upper bound in Corollary 6 can be saturated - it is a tight bound. On the other hand, the success probability decreases significantly as $N$ and $d / \sigma$ increase; it decreases as $N$
increases because we perform post-selection after every iteration and as $d / \sigma$ increases because when the elements in the input state are more distinguishable, there is less overlap between the elements, which decreases the success probability in the post-selection (less contribution from other filtered-out elements). Again, no matter what the values of $N$ and $d / \sigma$ are, the final state can never be more squeezed than the 'optimal squeezing', the variance of $|0, r\rangle_{1 S Q}$ (dash-dotted grey line in Figure 4.4(b)), as stated in Corollary 6. In most regions larger $N$ results in a better performance, but for small $d / \sigma$ using more copies improves neither the $\hat{x}$-variance nor the fidelity. This is because when the two squeezed states in the initial state have a large overlap, the post-selection procedure does not filter out the unwanted parts as effectively, and more iterations just end up with more unwanted elements in the final state.

### 4.4.2 Entanglement distillation

We now move to the convex Gaussian resource theory of entanglement. In this section, we show that entanglement can be distilled when convex mixtures of Gaussian states and conditional Gaussian LOCC are allowed. This is achieved with a protocol very similar to the multi-copy squeezing distillation protocol studied in the last section.

Consider the following setting. Alice prepares pure two-mode squeezed vacuum (TMSV) states ${ }^{6}|\mathbf{0}, r\rangle_{\text {2SQ }}$, where $\mathbf{0}=(0,0)^{\mathrm{T}}$, and send one mode via a noisy channel to Bob. The noisy channel either applies a finite displacement along a certain quadrature with probability $p$ or leaves the mode unchanged. This noise model is a convenient idealisation of random displacement noise and produces a non-Gaussian mixture of Gaussian states from which entanglement can be distilled. However, it is worth pointing out that this noise model is not realistic, since for CV systems the most generic displacement noise would be continuous rather than discrete. We assume this idealised noise model here as our main goal is to provide an example of entanglement distillation using convex Gaussian resources, but it remains an interesting open question to see whether entanglement could be distilled under a more realistic noise model.

[^18]

Figure 4.5: The marginal probability distribution obtained from the Wigner function of the state $\rho_{\text {in }}$. In the $\hat{x}_{1}, \hat{x}_{2}$ space, the distribution is represented by two Gaussian blocks squeezed along the $\hat{x}_{1}+\hat{x}_{2}$ quadrature, one centred in the origin and the other displaced by $d$ along the $\hat{x}_{1}$ quadrature. If Bob was to measure his local quadrature $\hat{x}_{1}$ alone, the two Gaussian blocks would overlap and distinguishing them would be more difficult than measuring the non-local quadrature $\hat{x}_{1}+\hat{x}_{2}$, where we see that two distributions have a smaller overlap.

The initial state shared by Alice and Bob after the noisy channel is

$$
\rho_{\text {in }}=(1-p)|\mathbf{0}, r\rangle\left\langle\mathbf{0}, \left.\left.r\right|_{12}+p \hat{D}_{1}\left(\frac{d}{\sqrt{2}}\right) \right\rvert\, \mathbf{0}, r\right\rangle\left\langle\mathbf{0},\left.r\right|_{12} \hat{D}_{1}\left(\frac{d}{\sqrt{2}}\right)^{\dagger},\right.
$$

where we omitted the subscription ' 2 SQ' in the TMSV state $|\mathbf{0}, r\rangle_{2 S Q}$ for simplicity of the notation, and $\hat{D}_{1}(d)$ is the displacement operator $\hat{D}_{\alpha}$ on the system 1 with $\alpha=d$. A convenient way of visualising the above state is in the $\hat{x}_{1}-\hat{x}_{2}$ space, shown in Figure 4.5. It is clear that the displacement in the $\hat{x}_{1}$ quadrature can be better detected by measuring the non-local quadrature $\hat{x}_{+}=\left(\hat{x}_{1}+\hat{x}_{2}\right) / \sqrt{2}$ rather than $\hat{x}_{1}$ itself. We exploit this observation to construct a probabilistic entanglement distillation protocol.

The protocol is shown in Figure 4.6. Alice and Bob use two copies of the state $\rho_{\text {in }}$ at the time, and feed their local modes into a $50: 50$ beam splitter. This operation creates classical


Figure 4.6: Schematic of the probabilistic multi-copy entanglement distillation protocol. Alice and Bob share two copies of the initial state and feed their local modes into $50: 50$ beam splitters. Then, via classical communication they perform joint homodyne measurement on the quadrature $\hat{x}_{+}=\left(\hat{x}_{1}+\hat{x}_{2}\right) / \sqrt{2}$. If the outcome is in the desired region, the protocol is successful, otherwise it is aborted. The protocol can be iterated multiple times with the output state of one iteration being the input state of the next one.
correlations between the two local modes on each side and produce the following state

$$
\begin{aligned}
\rho^{\prime}= & (1-p)^{2}|\mathbf{0}, r\rangle\left\langle\mathbf{0},\left.r\right|_{12} \otimes \mid \mathbf{0}, r\right\rangle\left\langle\mathbf{0},\left.r\right|_{34}\right. \\
& +p(1-p) \hat{D}_{1}\left(\frac{d}{2}\right)|\mathbf{0}, r\rangle\left\langle\mathbf{0}, \left.\left.r\right|_{12} \hat{D}_{1}\left(\frac{d}{2}\right)^{\dagger} \otimes \hat{D}_{3}\left(\frac{d}{2}\right) \right\rvert\, \mathbf{0}, r\right\rangle\left\langle\mathbf{0},\left.r\right|_{34} \hat{D}_{3}\left(\frac{d}{2}\right)^{\dagger}\right. \\
& +p(1-p) \hat{D}_{1}\left(\frac{d}{2}\right)|\mathbf{0}, r\rangle\left\langle\mathbf{0}, \left.\left.r\right|_{12} \hat{D}_{1}\left(\frac{d}{2}\right)^{\dagger} \otimes \hat{D}_{3}\left(-\frac{d}{2}\right) \right\rvert\, \mathbf{0}, r\right\rangle\left\langle\mathbf{0},\left.r\right|_{34} \hat{D}_{3}\left(-\frac{d}{2}\right)^{\dagger}\right. \\
& +p^{2} \hat{D}_{1}(d)|\mathbf{0}, r\rangle\left\langle\mathbf{0},\left.r\right|_{12} \hat{D}_{1}(d)^{\dagger} \otimes \mid \mathbf{0}, r\right\rangle\left\langle\mathbf{0},\left.r\right|_{34},\right.
\end{aligned}
$$

where modes 1 and 3 are held by Bob while modes 2 and 4 by Alice. Alice and Bob now need to identify whether their systems are displaced or not by measuring the system on the modes 1 and 2 . Since the displacement was performed along the $\hat{x}_{1}$ quadrature, one way of detecting the displacement is for Bob to measure the mode 1 locally. However, as we discussed earlier, it is more advantageous to exploit the squeezing along the non-local quadrature $\hat{x}_{+}$when detecting this displacement; see Figure 4.5. Therefore, Alice and Bob can locally measure $\hat{x}_{2}$ and $\hat{x}_{1}$ respectively, communicate their outcomes classically, and process them. This makes it possible to measure the non-local quadrature $\hat{x}_{+}$. Note that the displacement by $\delta$ along $\hat{x}_{1}$ results in the displacement by $\delta / \sqrt{2}$ along $\hat{x}_{+}$, which could reduce the distinguishability when measuring $\hat{x}_{+}$. However, if the squeezing is significant, measuring $\hat{x}_{+}$is still able to provide an advantageous
over measuring $\hat{x}_{1}$. Alice and Bob then perform a dichotomic POVM $\{\Pi, \mathbb{1}-\Pi\}$ on modes 1 and 2 , where

$$
\begin{equation*}
\Pi=\int \mathrm{d} x_{1} \mathrm{~d} x_{2} \chi_{\left(-\Delta^{\prime}, \Delta^{\prime}\right]}\left(\frac{x_{1}+x_{2}}{\sqrt{2}}\right)\left|x_{1}\right\rangle\left\langle\left. x_{1}\right|_{1} \otimes \mid x_{2}\right\rangle\left\langle\left. x_{2}\right|_{2},\right. \tag{4.9}
\end{equation*}
$$

where the indicator function $\chi_{\left(-\Delta^{\prime}, \Delta\right]}\left(\frac{x_{1}+x_{2}}{\sqrt{2}}\right)$ is equal to 1 when $-\Delta^{\prime}<\frac{x_{1}+x_{2}}{\sqrt{2}} \leq \Delta^{\prime}$ and 0 otherwise, and $\Delta^{\prime}=e^{-r}$. The value $\frac{x_{1}+x_{2}}{\sqrt{2}}$ can be computed by Alice and Bob after they share their outcomes.

After measuring the above POVM, Alice and Bob can decide whether to continue the protocol or abort it. If $x_{+} \in\left(-\Delta^{\prime}, \Delta^{\prime}\right]$, then they keep the systems in Modes 3 and 4 , otherwise discard the systems and start again with a new copy. Even in the case of success, the final state is composed by a mixture, where the state $|\mathbf{0}, r\rangle_{2 S Q}$ occurs with high probability, and the displaced components have a much lower weight. Alice and Bob can iterate the protocol many times by suitably changing the interval of acceptance in Eq. (4.9) at each round, to increase the weight of the desired state $|\mathbf{0}, r\rangle_{2 S Q}$. Since this protocol is non-deterministic, the probability of success quickly decreases with the number of iterations. It should be stressed that Alice and Bob can construct a fully deterministic protocol based on the same idea with more measurement outcomes indicating different displacements; then one can correct the displacement depending on the measurement outcome. Although this increases the success probability to 1 , this new protocol is less straightforward to be implemented since the number of elements with different displacements increases at each iteration and requires more and more number of measurement outcomes after each iteration.

The numerical simulation of the protocol in Figure 4.7 shows the fidelity to $|\mathbf{0}, r\rangle_{2 \mathrm{SQ}}$ of the final state and the success probability $p_{\text {succ }}$ of the protocol as a function of $d / \sigma$ for different numbers of copies $N$. The results look similar to the multi-copy squeezing distillation case in Figure 4.4 as the idea is same. The performance of the entanglement distillation is generally worse than the squeezing case, but we can still achieve fidelity 1 for large $d / \sigma$. Again, for small $d / \sigma$, using more copies does not improve the performance as the filtering process via post-selection doe not work as effectively due to the large overlap between elements in the initial


Figure 4.7: Numerical simulation of the multi-copy entanglement distillation protocol for different $N$ copies of the initial state. The graph shows the fidelity between the target state $\rho_{\mathrm{sq}}=|\mathbf{0}, r\rangle_{2 S Q}\left\langle\mathbf{0},\left.r\right|_{2 S Q}\right.$ and the final state of the protocol as a function of $d / \sigma$. The success probability $p_{\text {succ }}$ of the protocol is shown in the inset. The protocol is regarded as successful when the final state has a larger fidelity than the initial state as well as the the vacuum state (the dotted grey line), which is the best free state, i.e., closest to $|\mathbf{0}, r\rangle_{2 \mathrm{SQ}}[1]$. The results are similar to the ones for the multi-copy squeezing distillation protocol. Using more copies leads to a better performance at the expense of the small success probability for larger $d / \sigma$, but not for small $d / \sigma$. For the simulation, we set the squeezing parameter $r=0.7$, and the probability of random displacement to $p=0.5$.
state. It is also worth noting that the upper bound in Corollary 6 can be again achieved for large $d / \sigma$; the fidelity reaches 1 . The described protocol is largely related to the protocol known as Gaussification, introduced in [77, 78], allowing for the distillation of Gaussian entanglement from initial non-Gaussian states. The main difference between our protocol and the Gaussification protocol is that we focus on a specific class of non-Gaussian states, convex mixtures of Gaussian states, as the initial state, and this specification allows us to obtain a more efficient protocol in terms of success probability.

### 4.5 Summary

In this chapter, we explored whether classical randomness and conditional operations can help to overcome known limitations on resource distillation in Gaussian resource theories. In Section 4.2, we introduced the notion of convex Gaussian resource theories, which is motivated
by the easy accessibility of non-Gaussian convex mixtures of Gaussian resources and their usefulness demonstrated in [71] in the context of non-Gaussianity distillation. We then defined an appropriate resource monotone in Section 4.3 using the convex roof extension and fully characterised the new limitation on resource distillation in convex Gaussian resource theories. In comparison with the no-go theorem for Gaussian resource distillation in [10], this new limitation does not rule out all possible resource distillation but restrict the amount of distillable resources, which leaves some possibility for a certain type of resource distillation. In Section 4.4, we explored this possibility and provided several explicit examples of resource distillation protocols for squeezing and entanglement. These examples in particular manifested that the derived upper bound on the amount of distillable resources is tight.

Although we showed that some limited resource distillation becomes possible in convex Gaussian resource theories, it does not seem that the non-convexity of Gaussian resource theories is the ultimate obstacle of resource distillation in the Gaussian platform. Even in convex Gaussian resource theories, the tensorisation property of the resource monotone in Lemma 4-4 still plays a crucial role in limiting more general resource distillation.

It is worth remarking that making the convex roof extension of Gaussian resource measures to the convex hull of Gaussian states seems quite general. Historically, the convex roof extension has been used to extend resource measures from pure states to mixed states [30]. Even though Gaussian states include mixed states, the relation between Gaussian states and convex mixtures of Gaussian states is not very different from the one between pure states and mixed states. Indeed, any monotone that is well-defined over the set of Gaussian states can be extended to the convex hull of Gaussian states by taking the convex roof extension.

## Chapter 5

## Catalytic Gaussian thermal operations

In this chapter, we will change our topic to a specific example of Gaussian resource theories, the Gaussian resource theory of thermodynamics. Since this example does not satisfy the assumptions described in Section $4.1^{1}$, the previous results in Chapter 4 cannot be applied to this specific example, and it is a new topic to explore.

The extension of thermodynamics outside of its original scope to include quantum systems has required the development of novel conceptual and theoretical approaches. In this regard, quantum resource theories have attracted attention for their application to thermodynamics. The resource theory approach enables precise definition of work and heat as well as novel insights into the role of quantum statistical quantities in small systems, such as entropy, entanglement, and coherence [34]. Moreover, it led to the appropriate mathematical formulation of the second law in quantum thermodynamics in terms of a family of free energy quantities [35]. However, these works have typically considered discrete-variable quantum systems and have limited applicability to continuous-variable (CV) quantum systems, some of the most important physical platforms for testing quantum thermodynamics. These include harmonic oscillators ranging from optics to mechanical systems, and in such systems the range of available operations and interactions is often limited to Gaussian elements. Thus, it is desirable to have a better adapted resource theory of quantum thermodynamics in the Gaussian regime. The first steps in this direction were made in $[79,80]$, which defined the resource theory of thermodynamics in Gaussian CV

[^19]systems. [80] identified a set of thermodynamical resources under passive linear interactions with a thermal environment and found some second law-like statements, while [79] analysed a slightly more general context with arbitrary quadratic Hamiltonian and provided a full characterisation of the possible dynamics of a single mode.

One element not yet considered in these works is the role played by catalysts. The notion of catalyst - a system that can be used but must be returned back to its initial state - is central even to classical thermodynamics; applies to any machine that operates in a cycle. The importance of catalysts to quantum resource theories beyond thermodynamics was demonstrated famously in the resource theory of entanglement (see Section 3.2.1). In discrete-variable thermodynamics, it is known that one recovers a single inequality describing the second law involving the Helmholtz free energy when looking at approximate transitions in the identical and independently distributed (i.i.d.) limit [35] or we allow for correlations between systems and catalysts in the final state [46].

In this paper, we characterise possible state transformations for Gaussian states under Gaussian thermal operations (GTOs) when catalysts are allowed; this is equivalent to characterise the action of allowed operations in the Gaussian resource theory of thermodynamics in the presence of catalysts. Our work can be considered an extension of the initial studies [79, 80] in two ways: (i) we now allow for catalysts, and (ii) we obtain state transition conditions for the multi-mode case. As the possible state transformations under non-catalytic Gaussian thermal operations characterised in $[79,80]$ were very limited, we ask whether catalysts can help us perform more interesting thermodynamic tasks in the Gaussian regime. We consider two different models for catalysts. The first type is called a strong catalyst, which must not only come back to its original state but also end up uncorrelated from the system. The second type, a weak catalyst, allows for final correlation with the system and requires the catalyst only locally to return to its initial state. Despite the general usefulness of catalysts, we find that strong catalysts do not enlarge the range of possible state transformations for a given initial state, while we can achieve more - yet still limited - state transformations with weak catalysts. Furthermore, we determine the full necessary and sufficient conditions for single-mode catalytic Gaussian thermal operations. For the multi-mode case, we introduce a new set of resource
monotones with clear physical interpretation and provide state transition conditions in terms of them.

To aid the reader, we provide a short summary of our results in this chapter here.

- Introduce a new set of resource monotones. We introduce a new set of resource monotones, principal mode temperatures and asymmetries ${ }^{2}$, and rephrase the state transition conditions under GTOs without catalysts in terms of these new monotones: Proposition 8 for the single-mode case and Theorem 9 for the multi-mode case.
- Define catalytic Gaussian thermal operations. We then formally define two types of catalysts in Definition 6: strong and weak catalysts. Strong catalysts are required to not only come back to its original state but also be uncorrelated to the system at the end of the process. Weak catalysts are allowed to have a remaining correlation with the system at the end of the process.
- Fully characterise single-mode catalytic Gaussian thermal operations. We fully characterise the possible state transformations under single-mode catalytic GTOs. Theorem 10 describes the state transition conditions for strong catalytic GTOs, which is exactly same as the conditions without catalyst in Proposition 8 - this implies that strong catalysts are not helpful. Theorem 11 describes the case for weak catalytic GTOs, which shows a limited improvement on the case without catalysts.
- Characterise multi-mode catalytic Gaussian thermal operations. For the multimode case, we derive several new state transition conditions in terms of principal mode temperatures and asymmetries. In particular, for strong catalysts, we expand the discussion to the approximate case and characterise approximate transformations with strong catalysts in Theorem 12. For weak catalysts, the state transition conditions are given in Theorem 13 and 14.

This chapter presents result from [Benjamin Yadin*, Hyejung H Jee* et al. Journal of Physics A. 55:325301 (2022)] and is structured as follows. In Section 5.1, we revise the definition

[^20]and known properties of Gaussian thermal operations and introduce a useful representation of Gaussian states and operations. We first look at state transitions under non-catalytic Gaussian thermal operations in Section 5.2 and then formally define two different types of catalytic Gaussian thermal operations in Section 5.3. The full characterisation of possible state transformations under single-mode catalytic Gaussian thermal operations is presented in Section 5.4, and the state transition conditions in the multi-mode case are discussed in Section 5.5. We discuss the physical implications of our results in Section 5.6 before summarising the chapter in Section 5.7.

As discussed in Chapter 2, we denote operators acting on the full Hilbert space with hat, e.g., $\hat{x}$, and operators acting on a finite dimensional space without hat, e.g., $x$. Also, the set of all passive transformation acting on an $N$-mode CV quantum system is denoted by $K(N) \equiv S p(2 N) \cap O(2 N)$.

### 5.1 Preliminaries

### 5.1.1 Gaussian thermal operations

Gaussian thermal operations (GTOs) are a sub-class of Gaussian operations which can be realised by energy-preserving interaction between a system and a thermal bath. More specifically, for a given second-order system Hamiltonian $\hat{H}_{S}$ and an inverse temperature $\beta=1 / k_{B} T$, GTOs are defined as operations obtained by

1. Preparing a thermal state with arbitrary second-order Hamiltonian $\hat{H}_{B}$, i.e.,

$$
\frac{e^{-\beta \hat{H}_{B}}}{\operatorname{Tr}\left[e^{-\beta \hat{H}_{B}}\right]}
$$

2. Applying an energy-preserving Gaussian unitary $\hat{U}_{S B}=e^{-i \hat{H}_{S B}}$ such that

$$
\left[\hat{H}_{S B}, \hat{H}_{S}+\hat{H}_{B}\right]=0
$$

The following theorem describes a representation of GTOs acting on covariance matrices.
Theorem 7. [79, Theorem 1] Let $\hat{H}_{S}=\frac{1}{2} \hat{\mathbf{r}}^{T} H_{S} \hat{\mathbf{r}}$ be a system Hamiltonian with normal form $S^{-1} H_{S}\left(S^{T}\right)^{-1}=\oplus_{l} \omega_{l} \mathbb{1}_{2 n_{l}}$ where $n_{l}$ is the mode degeneracy of an eigenfrequency $\omega_{l}$. For a given system $C M \sigma$, the transformation under generic GTOs at background inverse temperature $\beta$ can be characterised by

$$
\begin{equation*}
\sigma \mapsto S\left(\oplus_{l} W_{l} \circ \Phi_{l} \circ Z_{l}\left(S^{-1}(\sigma)\right)\right), \tag{5.1}
\end{equation*}
$$

where $W_{l}$ and $Z_{l}$ are passive transformations acting only on the modes with eigenfrequency $\omega_{l}$, and $\Phi_{l}$ are CP maps describing thermalisation as a result of interacting with thermal baths, i.e.,

$$
\Phi_{l}\left(\sigma_{l}\right)=X_{l} \sigma_{l} X_{l}^{T}+Y_{l},
$$

where $\sigma_{l}$ is the CM of the $l$-th degenerate sector, $X_{l}=\bigoplus_{k=1}^{n_{l}} \cos \theta_{l k} \mathbb{1}_{2}$ and $Y_{l}=\bigoplus_{k=1}^{n_{l}} \frac{\mathrm{e}^{\beta \omega_{l}}+1}{\mathrm{e}^{\beta \omega_{l}}-1} \sin ^{2} \theta_{l k} \mathbb{1}_{2}$, for some $\theta_{l k} \in \mathbb{R}$.

Note that in Eq. (5.1), apart from the symplectic transformation $S$ which brings the system Hamiltonian into the normal form, modes with different eigenfrequencies transform independently; GTOs cannot make modes with different eigenfrequencies interact with each other. is intuitive because GTOs are energy-preserving operations, and interactions between different energy levels would require some energy exchange. The characterisation in Theorem 7 says that, once the Hamiltonian is transformed to its normal form, a GTO can be implemented in three steps:

1. Adding thermal bath ancillae with same eigenfrequency and same number of modes to each different eigenfrequency-sector of the system.
2. Applying passive transformations to each eigenfrequency-sector separately.
3. Tracing out the bath modes.

Since non-degenerate modes do not interact, in this chapter we only consider the interesting case of degenerate modes, i.e., the case when all modes have the same eigenfrequency. Moreover, we assume a mode basis has already been chosen such that the Hamiltonian is in its normal
form $H_{S}=\omega \mathbb{1}_{2 n}$; equivalently $\hat{H}_{S}=\omega\left(\hat{n}_{S}+1 / 2\right)$ acting on the Hilbert space, where $\hat{n}_{S}$ is the total number operator. Under these conditions GTOs effectively reduce to only the three types of operations described above, which were introduced as bosonic linear thermal operations in [80].

### 5.1.2 Principal mode temperatures and principal mode asymmetries

As we have seen in Chapter 2, a general Gaussian state can be represented by its vector of first moments $\overline{\mathbf{r}}$ and its CM $\sigma$. Here, we consider the case of vanishing first moments, $\overline{\mathbf{r}}=0$, and concentrate only on the thermodynamical properties of the CM. We employ a special decomposition of the CM, first introduced in [81], that provides both a convenient mathematical description for analysing transformations under GTOs and a set of quantities with interesting physical interpretations. A $2 N \times 2 N$ covariance matrix can be decomposed into two $N \times N$ matrices:

$$
M_{i j}:=\left\langle\hat{a}_{j}^{\dagger} \hat{a}_{i}\right\rangle-\delta_{i j} \nu, \quad A_{i j}:=\left\langle\hat{a}_{j} \hat{a}_{i}\right\rangle,
$$

where $\nu=\frac{1+e^{-\beta \omega}}{1-e^{-\beta \omega}}$ is the variance of any quadrature in a thermal state at background inverse temperature $\beta$ (see Eq. (2.20)). They are both complex matrices; $M$ is Hermitian, and $A$ is symmetric. $M$ is sometimes known as the single-particle density matrix, up to a constant shift. In terms of these two matrices, we can recover the CM $\sigma$ via

$$
\begin{aligned}
\left\langle\hat{x}_{i} \hat{x}_{j}\right\rangle & =\operatorname{Re}\left[M_{i j}+A_{i j}\right]+\delta_{i j} \nu+\frac{1}{2}, \\
\left\langle\hat{p}_{i} \hat{p}_{j}\right\rangle & =\operatorname{Re}\left[M_{i j}-A_{i j}\right]+\delta_{i j} \nu+\frac{1}{2}, \\
\left\langle\left\{\hat{x}_{i}, \hat{p}_{j}\right\}\right\rangle & =2 \operatorname{Im}\left[A_{i j}-M_{i j}\right] .
\end{aligned}
$$

This representation is convenient since a thermal state at the background temperature is characterised by $M=A=\mathbf{0}$. Furthermore, these matrices evolve simply under the passive transformations on which GTOs are based. Under such an operation described by a unitary
matrix $U$, the matrices $M$ and $A$ transform as

$$
\begin{array}{ll}
M_{i j} \mapsto \sum_{k, l} U_{i k} M_{k l} U_{j l}^{*}, & M \mapsto U M U^{\dagger} \\
A_{i j} \mapsto \sum_{k, l} U_{i k} A_{k l} U_{j l}, \quad A \mapsto U A U^{T} .
\end{array}
$$

It is therefore possible to diagonalise $M$ with an appropriate choice of $U .^{3} \quad A$ can also be diagonalised with a generally different unitary $V$. An advantage of using $M$ and $A$ instead of CMs is that we can identify some quantities that are invariant under passive transformations and thus equal for states that are thermodynamically equivalent due to being unitarily interconvertible under GTOs.

In a basis where $M$ is diagonal, we have $M=\operatorname{diag}\left(\mu_{1}, \ldots, \mu_{N}\right) . A$ is not in general diagonal in this basis, but we are free to make arbitrary phase rotations of the form $U_{i j} \rightarrow U_{i j} e^{i \phi_{j}}$ while leaving $M$ unchanged, under which $A_{i j} \rightarrow e^{i\left(\phi_{i}+\phi_{j}\right)} A_{i j}$. We can thus always leverage this freedom to make the diagonals of $A$ real and non-negative, i.e., $A_{i i} \geq 0$. Ordering the canonical operators as ( $\left.\hat{x}_{1}, \hat{p}_{1}, \hat{x}_{2}, \hat{p}_{2}, \ldots\right)$, the resulting form of the CM is
$\sigma=2 \times\left(\begin{array}{cc|cc|cc}\mu_{1}+A_{11}+\nu+\frac{1}{2} & 0 & \operatorname{Re}\left[A_{12}\right] & \operatorname{Im}\left[A_{12}\right] & \ldots & \ldots \\ 0 & \mu_{1}-A_{11}+\nu+\frac{1}{2} & \operatorname{Im}\left[A_{12}\right] & -\operatorname{Re}\left[A_{12}\right] & \ldots & \cdots \\ \hline \operatorname{Re}\left[A_{12}\right] & \operatorname{Im}\left[A_{12}\right] & \mu_{2}+A_{22}+\nu+\frac{1}{2} & 0 & \ldots & \ldots \\ \operatorname{Im}\left[A_{12}\right] & -\operatorname{Re}\left[A_{12}\right] & 0 & \mu_{2}-A_{22}+\nu+\frac{1}{2} & \ldots & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots\end{array}\right)$.

Aside from the eigenvalues $\left\{\mu_{i}\right\}_{i}$ of $M$, another set of invariant quantities under passive transformations are the singular values $\alpha_{i} \geq 0$ of $A$, obtained by diagonalising $A$ with the unitary $V$ by congruence [82, Chapter 4]. In this chapter, we characterise states in terms of the parameter $\boldsymbol{\mu}=\left(\mu_{1}, \ldots, \mu_{N}\right)$ and $\boldsymbol{\alpha}=\left(\alpha_{1}, \ldots, \alpha_{N}\right)$. Without loss of generality, we assume

[^21]that they are arranged in descending order. Individually, $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ determine $M$ and $A$ up to passive transformations, but it should be noted that they do not generally provide a full characterisation of covariance matrices. The problem is that $M$ and $A$ are not in general simultaneously diagonalisable. However, when they are, these parameters are sufficient to determine $\sigma$ up to a passive transformations:

Definition 5. A covariance matrix $\sigma$ is called decouplable if there exists a passive transformation that makes all the modes uncorrelated. In this case, the normal form in Eq. (5.2) is fully diagonal and $A_{i i}=\alpha_{i}$ for all $i$.

The parameters $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ have interesting physical interpretations (see Figure 5.1). The quantities $\left\{\mu_{i}+\nu_{i}\right\}_{i}$ were named principal mode temperatures in [80] and described as the most extreme effective temperatures that can be found in any mode decomposition of a state. The mean energy of the $i$-th mode is a function of $\left\langle\hat{x}_{i}^{2}+\hat{p}_{i}^{2}\right\rangle / 2=M_{i i}+\nu+1 / 2$ in the case of vanishing first moments. Given the ordering $\mu_{1} \geq \mu_{2} \geq \ldots, \mu_{1}$ is the greatest effective temperature found in any mode decomposition. $\mu_{2}$ is the next greatest value found in any mode orthogonal to this, and so on. The $\left\{\alpha_{i}\right\}_{i}$ instead describe the rotational asymmetry of modes in the phase space, since $\left\langle\hat{x}_{i}^{2}-\hat{p}_{i}^{2}\right\rangle / 2=\operatorname{Re}\left[A_{i i}\right]=A_{i i}$ using the aforementioned phase freedom. We name them principal mode asymmetries. Again ordering $\alpha_{1} \geq \alpha_{2} \geq \ldots$, we see that $\alpha_{1}$ describes the mode with the greatest asymmetry, $\alpha_{2}$ describes the next greatest amongst modes orthogonal to that, and so on. These are somewhat similar to squeezing, except that $\alpha>0$ does not necessarily imply a sub-shot-noise quadrature variance.

For an example, let us consider a single-mode squeezed thermal state with the following CM:

$$
\sigma=\left(\begin{array}{cc}
2 t e^{-2 r} & 0 \\
0 & 2 t e^{2 r}
\end{array}\right)
$$

which can be generated by applying the single-mode squeezing transformation $S_{1 \mathrm{SQ}}(r)$ in Eq. (2.11) to a thermal state with $\mathrm{CM} \sigma_{\mathrm{th}}=\operatorname{diag}(2 t, 2 t)^{4}$. In this case, the matrices $M$ and $A$ are just numbers and naturally decouplable. Using the relation in Eq. (5.2), we can derive the

[^22]

Figure 5.1: Illustration of the parameters in Section 5.1.2 (shown for a single mode). We use a common visual representation of Gaussian states in the phase space, where an ellipse indicates the region with high Wigner probability distribution. Note that the quantities indicated are the squares of the lengths. Dashed circle: thermal state at the background temperature $T$, with variance $\nu$ in all quadrature directions. Solid ellipse: arbitrary Gaussian state (with vanishing first moments). $\mu$ is the principal mode temperature and $\alpha$ is the principal mode asymmetry.
principal mode temperature and asymmetry of this state:

$$
\mu=\frac{1}{2} t\left(e^{-2 r}+e^{2 r}\right)-\left(\nu+\frac{1}{2}\right), \quad \alpha=\frac{1}{2} t\left(e^{-2 r}-e^{2 r}\right),
$$

where $\nu$ is the variance of a thermal state at background temperature.

### 5.2 Gaussian thermal operations without catalysts

In general, our goal is to understand when one state $\rho$ of a system can be transformed into another state $\rho^{\prime}$ via a GTO in the presence of catalysts when all modes are degenerate. In this section, we first revise state transformations under normal GTOs without catalysts in terms of the principal mode temperatures and asymmetries.

When there exists a GTO which maps a state of a system to another state $\rho^{\prime}$, we write $\rho \xrightarrow{G T O} \rho^{\prime}$. For Gaussian states with vanishing first moments, this is equivalent to the conversion of one CM into another, denoted by $\sigma \xrightarrow{G T O} \sigma^{\prime}$. Alternatively, a simpler question can be asked using our decomposition of the CM: does a GTO exist that maps a state with matrix $M$ into one with matrix $M^{\prime}$, not concerned with what happens to the matrix $A$ ? When this is possible,
we write $M \xrightarrow{G T O} M^{\prime}$; likewise, the notation $A \xrightarrow{G T O} A^{\prime}$ implies the matrix $A$ can be transformed into $A^{\prime}$ via GTOs.

The following result from [79], which is restated in terms of the principal mode temperature $\mu$ and the principal mode asymmetry $\alpha$, answers when one single-mode CM can be transformed into another CM via GTO.

Proposition 8. [79] Let $\sigma, \sigma^{\prime}$ be covariance matrices of a single mode $C V$ system. Then $\sigma \xrightarrow{G T O} \sigma^{\prime}$ if and only if there exists $p \in[0,1]$ such that

$$
\begin{equation*}
\mu^{\prime}=p \mu, \quad \alpha^{\prime}=p \alpha \tag{5.3}
\end{equation*}
$$

Note that for the single-mode case, the matrices $A$ and $M$ reduce to single numbers and necessarily commute to each other. Using the relation between the CM and the matrices $A$ and $M$ in Eq. (5.2), one can state the condition in Eq. (5.3) alternatively as the existence of $p \in[0,1]$ and $S \in K(1)$ such that

$$
\begin{equation*}
\sigma^{\prime}=p S \sigma S^{T}+(1-p) \nu \mathbb{1}_{2} . \tag{5.4}
\end{equation*}
$$

In simple words, any single-mode GTO is equivalent to applying a phase rotation and mixing with thermal noise. ${ }^{5}$

### 5.2.1 Multi-mode case

For multiple modes, a set of necessary conditions in terms of principal mode temperatures were found in [80]. We derive them again here with a different approach which we can also use to derive analogous conditions for the principal mode asymmetries - the part two of Theorem 9 which describes the condition for principal mode asymmetries is our novel result. We first need to introduce some useful notation. The parameter $\boldsymbol{\mu}$ can be divided into super- and sub-thermal (or 'hot' and 'cold') parts, namely the positive and negative values:

[^23]$\boldsymbol{\mu}=\left(\mu_{1}^{+}, \ldots, \mu_{n_{+}}^{+}, 0, \ldots, 0,-\mu_{n_{-}}^{-}, \ldots,-\mu_{1}^{-}\right)$, where
\[

$$
\begin{aligned}
& \mu_{1}^{+} \geq \mu_{2}^{+} \geq \cdots \geq \mu_{n_{+}}^{+}>0 \\
& \mu_{1}^{-} \geq \mu_{2}^{-} \geq \cdots \geq \mu_{n_{-}}^{-}>0 .
\end{aligned}
$$
\]

It is sometimes convenient to extend these lists with zeroes, so that by convention $\mu_{i}^{ \pm}=0$ for $i>n_{ \pm}$. Also, for two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ with same length, we write $\boldsymbol{x} \leq \boldsymbol{y}$ when $x_{i} \leq y_{i}$ for all $i$.

Theorem 9. 1. (Necessity proved in [80]) For $M$ with eigenvalue vector $\boldsymbol{\mu}$ and $M^{\prime}$ with $\boldsymbol{\mu}^{\prime}$, $M \xrightarrow{G T O} M^{\prime}$ if and only if $\boldsymbol{\mu}^{++} \leq \boldsymbol{\mu}^{+}$and $\boldsymbol{\mu}^{\prime-} \leq \boldsymbol{\mu}^{-}$
2. (New result) For $A$ with singular-value vector $\boldsymbol{\alpha}$ and $A^{\prime}$ with $\boldsymbol{\alpha}^{\prime}, A \xrightarrow{G T O} A^{\prime}$ if and only if $\boldsymbol{\alpha}^{\prime} \leq \boldsymbol{\alpha}$.

The proof will be followed below. The main idea is that the matrices transform as $M^{\prime}=$ $P M P^{\dagger}, A^{\prime}=P A P^{\mathrm{T}}$, where $P$ is a sub-matrix of the unitary describing the coupling to the bath. Since $P$ is a contraction, it has a corresponding contractive effect on the eigenvalues and singular values. The sufficiency of the conditions results from finding a type of elementary GTO which applies what we call an L-transform on the parameters, which is able to perform the required transformations applied to each mode independently. An L-transform is achieved by simply mixing a single-mode system with a thermal mode at a beam splitter of arbitrary transmissivity, and scales the parameters $\mu, \alpha$ towards zero.

In simple words, this result says that the principal mode temperatures and asymmetries all converge individually towards the thermal values, i.e., zero. While the inequalities in Theorem 9 are necessary and sufficient for multi-mode GTO transformations between $M$ and $A$ matrices independently, we are not able to say anything general about whether both sets of conditions can be simultaneously satisfied. Thus, the inequalities are only necessary conditions for transformations of covariance matrices, but not sufficient.

Proof of Theorem 9. Part 1: The initial covariance matrix is $\sigma_{S} \oplus \sigma_{B}$ where $\sigma_{B}$ is thermal, so
$M_{S B}=M_{S} \oplus \mathbf{0}_{B}$. Let us partition the interacting passive unitary $U$ in the following way;

$$
U=\left(\begin{array}{ll}
P & Q \\
R & S
\end{array}\right)
$$

The submatrix $P$ is a contraction, meaning that $P^{\dagger} P \leq \mathbb{1}, P P^{\dagger} \leq \mathbb{1}$ with respect to the matrix inequality. Partitioning the final matrix $M^{\prime}=U M U^{\dagger}$ in the same way (with stars denoting unspecified elements),

$$
M^{\prime}=\left(\begin{array}{cc}
M_{S}^{\prime} & * \\
* & *
\end{array}\right)
$$

we find that $M_{S}^{\prime}=P M_{S} P^{\dagger}$. An extension of Ostrowski's theorem [82, Corollary 4.5.11] says that there exist $r_{i} \in\left[p_{n}^{2}, p_{1}^{2}\right]$, where $p_{1} \geq \cdots \geq p_{n}$ are the singular values of $P$, such that $\mu_{i}^{\prime}=r_{i} \mu_{i}$. The necessity of the conditions in Theorem 9-1 follows from $P$ being a contraction, so that $r_{i} \in[0,1]$.

Sufficiency is seen from the fact that $\boldsymbol{\mu}^{\prime}$ can be obtained from $\boldsymbol{\mu}$ by multiplying each element by some $r \in[0,1]$. We call such an operation on each element an $L$-transform. This can be performed on each mode by interacting with a single thermal mode via a beam splitter: concentrating on just this pair of modes, take

$$
U=\left(\begin{array}{cc}
\sqrt{r} & -\sqrt{1-r} \\
\sqrt{1-r} & \sqrt{r}
\end{array}\right)
$$

Without loss of generality, we can assume the initial matrix $M$ of the system to be diagonal, so for this pair of modes,

$$
M=\left(\begin{array}{ll}
\mu & 0 \\
0 & 0
\end{array}\right) \longrightarrow M^{\prime}=U M U^{\dagger}=\left(\begin{array}{cc}
r \mu & -\sqrt{r(1-r)} \mu \\
-\sqrt{r(1-r)} \mu & (1-r) \mu
\end{array}\right)
$$

and tracing out the bath just gives the diagonal element $r \mu$.
Part 2: We similarly have $A^{\prime}=U\left(A_{S} \oplus \mathbf{0}_{B}\right) U^{T}$ so that $A_{S}^{\prime}=P A_{S} P^{T}$. A related result for complex symmetric matrices [82, Theorem 4.5.13] says that $\alpha_{i}^{\prime}=r_{i} \alpha$, where $r_{i} \in\left[q_{n}, q_{1}\right]$ and
$q_{1} \geq \cdots \geq q_{n}$ are the eigenvalues of $P P^{\dagger}$ - which again lie in the interval $[0,1]$. Sufficiency follows from starting in a basis with $A=\operatorname{diag}(\boldsymbol{\alpha})$ and then applying the same $U$ as above for each mode. The calculation proceeds identically since all matrix elements are real.

### 5.3 Definition of catalytic Gaussian thermal operations

Our main interest in this chapter is state transformations under catalytic GTOs, namely GTOs with the presence of catalysts. In this section, we define two different cases of catalytic GTO transformations.

Definition 6. 1. (Strong catalysts) We say that a strong catalytic GTO transformation from $\rho_{S}$ to $\rho_{S}^{\prime}$ is possible if there exists a state $\rho_{C}$ such that

$$
\rho_{S} \otimes \rho_{C} \xrightarrow{G T O} \rho_{S}^{\prime} \otimes \rho_{C} .
$$

The same terminology applies to transformations at the level of the matrices $\sigma, M$ and $A$ - for example, $M_{S} \oplus M_{C} \xrightarrow{G T O} M_{S}^{\prime} \oplus M_{C}$.
2. (Weak catalysts) We say that a weak catalytic GTO transformation from $\rho_{S}$ to $\rho_{S}^{\prime}$ is possible if there exist a state $\rho_{C}$ and a final state $\rho_{S C}^{\prime}$ such that

$$
\rho_{S} \otimes \rho_{C} \xrightarrow{G T O} \rho_{S C}^{\prime} \quad \text { with } \rho_{C}^{\prime}=\operatorname{Tr}_{S}\left[\rho_{S C}^{\prime}\right]=\rho_{C} \text {. }
$$

The same terminology applies to transformations at the level of the matrices $\sigma, M$ and $A$ - for example, $A_{S} \oplus A_{C} \xrightarrow{G T O} A_{S C}^{\prime}$ with the catalyst sub-block $A_{C}^{\prime}=A_{C}$.

A strong catalyst must be not only returned to its original state but also uncorrelated from the system at the end. This agrees with the most common definition that has been used, for instance, in the resource theory of entanglement (see Section 3.2.1). For a weak catalyst, arbitrary correlations between the system and the catalyst in the final state are allowed. This weakening of the constraint has been studied recently and shown to result in sensible versions of the second law $[46,83]$. We would like to remark that although we use the same terminology for
catalytic conditions on either the density matrix or the covariance matrix and its submatrices, these conditions are not equivalent. Satisfying the catalytic condition on the density matrix picture implies catalysis at the level of second moments, but the reverse is in general not true. This is the case especially when one considers non-Gaussian inputs - one can have non-Gaussian states under GTO where the state transformation is not catalytic but the second moments are; for instance, a Fock state mixing at a beam splitter with a thermal mode with the same second moment. However, these conditions are indeed equivalent for Gaussian inputs, aside from first moments.

With the above definitions, a natural question is whether such catalytic GTO transformations allow for more state transformations to be performed on a system. For the rest of this chapter, we provide necessary or sufficient conditions for these two different catalytic GTO transformations in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ and compare them to the case where no catalyst is used.

### 5.4 Single-mode with catalysts: full characterisation

We first consider the simplest case of catalytic GTO transformations, where both the system and the catalyst are composed of a single mode. This single-mode case is a special case of our formulation - since the initial state $\rho_{S} \otimes \rho_{C}$ is naturally decouplable, $\boldsymbol{\mu}=\left(\mu_{S}, \mu_{C}\right)$ and $\boldsymbol{\alpha}=\left(\alpha_{S}, \alpha_{C}\right)$ fully characterise the initial state of the system and the catalyst. Indeed, the initial state $M_{S C}$ and $A_{S C}$ describing the system and the catalyst can be written as

$$
M_{S C}=\left(\begin{array}{cc}
\mu_{S} & 0 \\
0 & \mu_{C}
\end{array}\right), \quad A_{S C}=\left(\begin{array}{cc}
\alpha_{S} & 0 \\
0 & \alpha_{C}
\end{array}\right)
$$

Since we can perform any GTO transformation using as many bath modes as the total modes of the system and the catalyst [79], we add a two-mode thermal bath at inverse temperature $\beta$. Then, the joint initial state of the system, the catalyst and the thermal bath is described by the
following $M$ and $A$ matrices:

$$
M=\left(\begin{array}{cccc}
\mu_{S} & 0 & 0 & 0 \\
0 & \mu_{C} & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right)=M_{S C} \oplus \mathbf{0}_{B}, \quad A=\left(\begin{array}{cccc}
\alpha_{S} & 0 & 0 & 0 \\
0 & \alpha_{C} & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right)=A_{S C} \oplus \mathbf{0}_{B}
$$

where $\mathbf{0}_{B}$ denotes the $2 \times 2$ null matrix. As discussed in Section 5.1.1, any GTO transformation can be obtained by applying a passive transformation to the whole state and tracing out the bath modes, and any passive transformation corresponds to a unitary $U$ acting on $M$ and $A$ matrices. Let us divide the interacting passive unitary $U$ into four $2 \times 2$ blocks;

$$
U=\left(\begin{array}{ll}
P & Q \\
R & S
\end{array}\right)
$$

Again, the sub-matrix $P$ satisfies $P^{\dagger} P \leq \mathbb{1}$ and $P P^{\dagger} \leq \mathbb{1}$ with respect to the matrix inequality. With this notation, the evolution of the system and the catalyst is only affected by the sub-matrix $P$ :

$$
\begin{equation*}
M_{S C}^{\prime}=P M_{S C} P^{\dagger}, \quad A_{S C}^{\prime}=P A_{S C} P^{T} \tag{5.5}
\end{equation*}
$$

The main question is that for a given description of the initial system, $\mu_{S}$ and $\alpha_{S}$, which final parameters $\mu_{S}^{\prime}$ and $\alpha_{S}^{\prime}$ can be reached via catalytic GTO transformations.

We first find that a strong catalyst does not enlarge the range of possible transformations on a single mode:

Theorem 10. (Single-mode strong catalytic GTO) Let $\sigma$ and $\sigma^{\prime}$ be covariance matrices of a single-mode $C V$ quantum system. The state transition condition between $\sigma$ and $\sigma^{\prime}$ under single-mode strong catalytic GTOs is exactly the same as the one under non-catalytic GTOs described in Eq. (5.4): there exist $p \in[0,1]$ and $S \in K(1)$ such that

$$
\sigma^{\prime}=p S \sigma S^{T}+(1-p) \nu \mathbb{1}_{2} .
$$

In terms of principal mode temperatures and asymmetries, the condition becomes same as Eq. (5.3).

The proof is given at the end of this section.
On the other hand, with a weak catalyst, we show that a greater range of transformations become possible.

Theorem 11. (Single-mode weak catalytic GTO) Consider a single-mode CV quantum system described by the parameters $\mu_{S}$ and $\alpha_{S}$. It can be mapped to the final state described by $\mu_{S}^{\prime}$ and $\alpha_{S}^{\prime}$ via single-mode weak catalytic GTOs if and only if there exist $p, q \in[0,1]$ such that $p \geq q$, and

$$
\begin{equation*}
\mu_{S}^{\prime}=p \mu_{S}, \quad \alpha_{S}^{\prime}=q \alpha_{S} \tag{5.6}
\end{equation*}
$$

Again, the proof is given at the end of this section.
These two results are illustrated in Figure 5.2(a). With a strong catalyst (the red line), the only possible kind of transformation in $\mu-\alpha$ parameter space is to move along the line towards the origin, which represents the thermal state. Physically, this implies that, apart from the symplectic transformation $S$ to the normal form, all possible state transformations under single-mode strong catalytic GTOs can be described as thermalisation towards the bath mode, which is same as the case of non-catalytic single-mode GTOs. On the other hand, with a weak catalyst, a triangular region (the blue region) becomes accessible. This requires not only $\mu$ and $\alpha$ to be non-increasing, but also the ratio $\frac{\alpha}{\mu}$. This latter condition may be interpreted as the 'aspect ratio' of the Gaussian distribution in the phase space becoming less extreme. The same region is also redrawn in terms of the eigenvalues of the final CM, $\lambda_{1}, \lambda_{2}$ with $\lambda_{1} \leq \lambda_{2}$, in Figure 5.2(b).

In the case of weak catalysts, any point in the triangular region can be reached by concatenating two simple processes: a reduction in $\alpha$ with fixed $\mu$ and a non-catalytic thermalisation which scales both parameters similarly towards zero. The first process involves a catalyst but no bath. In order to carry out the transformation $\left(\mu_{S}, \alpha_{S}\right) \mapsto\left(\mu_{S}, \bar{\alpha}_{S}\right)$, we take $\mu_{C}=\mu_{S}$ and


Figure 5.2: Visualisation of possible state transformations under single-mode strong or weak catalytic GTOs, when the initial state of the system is described by the parameters $\mu_{S}, \alpha_{S}$ or the CM $\sigma_{S}$. The grey regions represent physical regions in both figures. (a) In the state space described by the principal mode temperature $\mu$ and asymmetry $\alpha$. Each point in the physical region (grey) fully describes a covariance matrix of the single-mode system up to passive transformations, and the origin represents the thermal bath. The red line describes all final states achievable by strong catalytic GTOs; it is same as the non-catalytic single-mode case (thermalisation). The blue region represents all possible final states with weak catalytic GTOs. In this case, we can additionally perform inverse squeezing to the system by coupling it to the catalyst. (b) In the state space described by the eigenvalues of the CM of the system, $\lambda_{1}, \lambda_{2}$ with $\lambda_{1} \geq \lambda_{2} . \sigma_{T}$ denotes the CM of the thermal bath. Again, the red line represents all possible state transformations with strong catalytic GTOs, and the blue region describes the same with weak catalytic GTOs. Note that the red line is the same as what we could achieve with non-catalytic single-mode GTOs discussed in Proposition 8.
$\alpha_{C}=\left(\alpha_{S}-\bar{\alpha}_{S}\right) / 2$, and apply the following unitary operation to the system and the catalyst:

$$
U=\left(\begin{array}{cc}
\sqrt{a} & i \sqrt{1-a} \\
\sqrt{1-a} & -i \sqrt{a}
\end{array}\right)
$$

where $a=\frac{\alpha_{S}+\bar{\alpha}_{S}}{3 \alpha_{S}-\bar{\alpha}_{S}} \in\left[\frac{1}{3}, 1\right]$, which can be performed with a suitable beam-splitter and phase shift.

Note that in the single-mode case the parameters $\mu_{S}$ and $\alpha_{S}$ fully characterise the system state up to passive transformations in the case of Gaussian input states. Thus, the conditions stated in Theorem 10 and 11 are necessary and sufficient conditions for state transformations under single-mode catalytic GTOs.

Now we present the proofs of the two theorems. As Theorem 10 can be derived as a special case of Theorem 11 with more constraints, we provide the proof of Theorem 11 first and then the one of Theorem 10.

Proof of Theorem 11. (Necessary conditions) We start from Eq. (5.5),

$$
\begin{aligned}
& M_{S C}^{\prime}=\left(\begin{array}{cc}
\left|p_{11}\right|^{2} \mu_{S}+\left|p_{12}\right|^{2} \mu_{C} & p_{11} p_{21}^{*} \mu_{S}+p_{12} p_{22}^{*} \mu_{C} \\
p_{11}^{*} p_{21} \mu_{S}+p_{12}^{*} p_{22} \mu_{C} & \left|p_{21}\right|^{2} \mu_{S}+\left|p_{22}\right|^{2} \mu_{C}
\end{array}\right), \\
& A_{S C}^{\prime}=\left(\begin{array}{cc}
p_{11}^{2} \alpha_{S}+p_{12}^{2} \alpha_{C} & p_{11} p_{21} \alpha_{S}+p_{12} p_{22} \alpha_{C} \\
p_{11} p_{21} \alpha_{S}+p_{12} p_{22} \alpha_{C} & p_{21}^{2} \alpha_{S}+p_{22}^{2} \alpha_{C}
\end{array}\right),
\end{aligned}
$$

where $p_{i j}=(P)_{i j}$ are the elements of the sub-matrix $P$. The final state is described by

$$
\begin{align*}
\mu_{S}^{\prime} & =\left|p_{11}\right|^{2} \mu_{S}+\left|p_{12}\right|^{2} \mu_{C}  \tag{5.7}\\
\alpha_{S}^{\prime} & =p_{11}^{2} \alpha_{S}+p_{12}^{2} \alpha_{C}
\end{align*}
$$

We want to find the relations between the initial parameters $\mu_{S}$ and $\alpha_{S}$ and the final parameters $\mu_{S}^{\prime}$ and $\alpha_{S}^{\prime}$. To do so, we need to exploit the catalytic conditions given by

$$
\begin{align*}
\left|p_{21}\right|^{2} \mu_{S}+\left|p_{22}\right|^{2} \mu_{C} & =\mu_{C},  \tag{5.8}\\
p_{21}^{2} \alpha_{S}+p_{22}^{2} \alpha_{C} & =\alpha_{C} .
\end{align*}
$$

Firstly, let us assume $\left|p_{22}\right|^{2}=1$. This is a trivial case as it implies $p_{12}=p_{21}=0$, so that $\mu_{S}^{\prime}=\left|p_{11}\right|^{2} \mu_{S}$ and $\alpha_{S}^{\prime}=p_{11}^{2} \alpha_{S}$, which satisfies the conditions in Eq. (5.6). In the following, we will assume $\left|p_{22}\right|^{2} \neq 1$.

Assuming $\mu_{S} \neq 0$ and $\alpha_{S} \neq 0$, we can rewrite the catalytic conditions in Eq. (5.8) as

$$
\frac{\mu_{C}}{\mu_{S}}=\frac{\left|p_{21}\right|^{2}}{1-\left|p_{22}\right|^{2}}, \quad \frac{\alpha_{C}}{\alpha_{S}}=\frac{p_{21}^{2}}{1-p_{22}^{2}}
$$

which implies that $\frac{p_{21}^{2}}{1-p_{22}^{2}}$ must be real and non-negative. ${ }^{6}$ Substituting these to the expressions

[^24]of the final state in Eq. (5.7) obtains
\[

$$
\begin{aligned}
& \mu_{S}^{\prime}=\left|p_{11}\right|^{2} \mu_{S}+\left|p_{12}\right|^{2} \mu_{C}=\left(\left|p_{11}\right|^{2}+\frac{\left|p_{12}\right|^{2}\left|p_{21}\right|^{2}}{1-\left|p_{22}\right|^{2}}\right) \mu_{S} \\
& \alpha_{S}^{\prime}=p_{11}^{2} \alpha_{S}+p_{12}^{2} \alpha_{C}=\left(p_{11}^{2}+\frac{p_{12}^{2} p_{21}^{2}}{1-p_{22}^{2}}\right) \alpha_{S}
\end{aligned}
$$
\]

which again implies that both expressions in the brackets are real. Then, we have

$$
\begin{aligned}
p_{11}^{2}+\frac{p_{12}^{2} p_{21}^{2}}{1-p_{22}^{2}} & \leq\left|p_{11}^{2}+\frac{p_{12}^{2} p_{21}^{2}}{1-p_{22}^{2}}\right| \leq\left|p_{11}^{2}\right|+\left|\frac{p_{12}^{2} p_{21}^{2}}{1-p_{22}^{2}}\right| \\
& =\left|p_{11}\right|^{2}+\left|p_{12}\right|^{2} \frac{\left|p_{21}^{2}\right|}{\left|1-p_{22}^{2}\right|} \leq\left|p_{11}\right|^{2}+\frac{\left|p_{21}\right|^{2}\left|p_{12}\right|^{2}}{1-\left|p_{22}\right|^{2}},
\end{aligned}
$$

where we used the reverse triangle inequality, $|x-y| \geq||x|-|y||$, and $\left|p_{22}\right| \leq 1$ in the last inequality. This shows $\alpha_{S}^{\prime} / \alpha_{S} \leq \mu_{S}^{\prime} / \mu_{S}$.

The remaining part is to prove $\mu_{S}^{\prime} / \mu_{S} \leq 1$. This can be easily shown using the condition of $P$ that $\mathbb{1}-P P^{\dagger} \geq 0$. We have that

$$
\mathbb{1}-P P^{\dagger}=\left(\begin{array}{cc}
1-\left|p_{11}\right|^{2}-\left|p_{12}\right|^{2} & -p_{11} p_{21}^{*}-p_{12} p_{22}^{*} \\
-p_{11}^{*} p_{21}-p_{12} p_{22}^{*} & 1-\left|p_{21}\right|^{2}-\left|p_{22}\right|^{2}
\end{array}\right) \geq 0
$$

which implies

$$
\begin{align*}
& 1-\left|p_{11}\right|^{2}-\left|p_{12}\right|^{2} \geq 0  \tag{5.10}\\
& 1-\left|p_{21}\right|^{2}-\left|p_{22}\right|^{2} \geq 0
\end{align*}
$$

Then, we obtain

$$
\left|p_{11}\right|^{2}+\frac{\left|p_{12}\right|^{2}\left|p_{21}\right|^{2}}{1-\left|p_{22}\right|^{2}} \leq\left|p_{11}\right|^{2}+\frac{\left(1-\left|p_{11}\right|^{2}\right)\left(1-\left|p_{22}\right|^{2}\right)}{1-\left|p_{22}\right|^{2}}=1
$$

where we used Eq. (5.10) in the first inequality. This completes the proof.
When $\mu_{S}=0$, the catalytic condition implies that either $\left|p_{22}\right|^{2}=1$ or $\mu_{C}=0$. If $\left|p_{22}\right|^{2}=1$, from the earlier argument on the case when $\left|p_{22}\right|^{2}=1$, we have $\mu_{S}^{\prime}=0$ and $\alpha_{S}^{\prime}=p_{11}^{2} \alpha_{S} \leq \alpha_{S}$. If $\mu_{C}=0$, then $\mu_{S}^{\prime}=0$, and all the results on $\alpha_{S}^{\prime}$ are same as above; $\alpha_{S}^{\prime} \leq \alpha_{S}$. A similar argument
works for the case when $\alpha_{S}=0$.
(Sufficient conditions) We would like to show that for any $\mu_{S}^{\prime}$ and $\alpha_{S}^{\prime}$ satisfying Eq. (5.6), we can find a weak catalytic GTO transformation which maps the initial state to the final state described by $\mu_{S}^{\prime}$ and $\alpha_{S}^{\prime}$. To do so, we provide an explicit weak catalytic GTO transformation which can map the initial state with $\mu_{S}$ and $\alpha_{S}$ to the final state with $\mu_{S}^{\prime}$ and $\alpha_{S}^{\prime}$. Consider a weak catalytic GTO transformation consisting of the following two steps:

1. By coupling the system with the catalyst, reduce the $\alpha_{S}$ to $\bar{\alpha}_{S}=\left(\mu_{S} / \mu_{S}^{\prime}\right) \alpha_{S}^{\prime}$ whilst keeping the same $\mu_{S}$. This can be done with the catalyst described by $\mu_{C}=\mu_{S}$ and $\alpha_{C}=\frac{\alpha_{S}-\bar{\alpha}_{S}}{2}$, and the following unitary operation $U_{1}$ acting on the matrices $M$ and $A$ :

$$
U_{1}=\left(\begin{array}{cc}
\sqrt{a} & i \sqrt{1-a}  \tag{5.11}\\
\sqrt{1-a} & -i \sqrt{a}
\end{array}\right)
$$

where $a=\frac{\alpha_{S}+\bar{\alpha}_{S}}{3 \alpha_{S}-\bar{\alpha}_{S}} \in\left[\frac{1}{3}, 1\right] . U_{1}$ can be realised by passive unitaries on the system and the catalyst; coupling the system and the catalyst via a beam splitter with transmissivity $a$ together with a phase shift.
2. Couple the system with the thermal bath and thermalise the system to the final state with $\mu_{S}^{\prime}$ and $\alpha_{S}^{\prime}$. This can be done by applying a beam splitter with transmissivity $\mu_{S}^{\prime} / \mu_{S}$ to the bath and the system, which is now described by $\mu_{S}$ and $\bar{\alpha}_{S}$, and tracing out the bath mode.

It is not difficult to check that above two steps transform $\mu_{S}, \alpha_{S}$ to $\mu_{S}^{\prime}, \alpha_{S}^{\prime}$, and the whole process is a GTO transformation. Geometrically, the first step describes moving vertically (reducing the principal mode asymmetry $\alpha$ ) in Fig. 5.2(a), and the second step is moving towards to the origin (pure thermalisation).

Now, we prove Theorem 10 borrowing a few results from the proof of Theorem 11.

Proof of Theorem 10. (Necessary conditions) We start again from Eq. (5.5), and the final state
is described by

$$
\begin{aligned}
\mu_{S}^{\prime} & =\left|p_{11}\right|^{2} \mu_{S}+\left|p_{12}\right|^{2} \mu_{C} \\
\alpha_{S}^{\prime} & =p_{11}^{2} \alpha_{S}+p_{12}^{2} \alpha_{C}
\end{aligned}
$$

Again, we want to find the relations between the initial parameters $\mu_{S}, \alpha_{S}$ and the final parameters $\mu_{S}^{\prime}, \alpha_{S}^{\prime}$. This time, we can exploit not only the catalytic conditions in Eq. (5.8) but also the no-correlation conditions given by vanishing off-diagonal elements,

$$
\begin{align*}
& p_{11} p_{21}^{*} \mu_{S}+p_{12} p_{22}^{*} \mu_{C}=0  \tag{5.12}\\
& p_{11} p_{21} \alpha_{S}+p_{12} p_{22} \alpha_{C}=0 .
\end{align*}
$$

As we are looking at a more restricted case (with more conditions) than the one with weak catalysts, the conditions for weak catalysts must hold regardless of the no-correlation conditions; the condition in Eq. (5.6) automatically holds due to the proof of Theorem 11. Now, we just need to show that the ratios $p=\frac{\mu_{S}^{\prime}}{\mu_{S}}$ and $q=\frac{\alpha_{S}^{\prime}}{\alpha_{S}}$ are the same.

Firstly, let us assume $\mu_{S} \neq 0$ and $\alpha_{S} \neq 0$. We notice from the no-correlation conditions in Eq. (5.12) that

$$
\frac{\mu_{C}}{\mu_{S}}=-\frac{p_{11} p_{21}^{*}}{p_{12} p_{22}^{*}} \in \mathbb{R}, \quad \frac{\alpha_{C}}{\alpha_{S}}=-\frac{p_{11} p_{21}}{p_{12} p_{22}} \in \mathbb{R},
$$

which implies that $\mu_{C} / \mu_{S}= \pm \alpha_{C} / \alpha_{S}$ depending on the phase. Combining this result with the catalytic conditions in Eq. (5.8) gives us

$$
\begin{gathered}
\frac{\mu_{C}}{\mu_{S}}=\frac{\left|p_{21}\right|^{2}}{1-\left|p_{22}\right|^{2}}=-\frac{p_{11} p_{21}^{*}}{p_{12} p_{22}^{*}} \Longrightarrow p_{11}=\frac{p_{21} p_{12} p_{22}^{*}}{\left|p_{22}\right|^{2}-1} \\
\frac{\alpha_{C}}{\alpha_{S}}=\frac{p_{21}^{2}}{1-p_{22}^{2}}=-\frac{p_{11} p_{21}}{p_{12} p_{22}} \Longrightarrow p_{11}=\frac{p_{21} p_{12} p_{22}}{p_{22}^{2}-1}
\end{gathered}
$$

where we assume that $\left|p_{22}\right| \neq 1, p_{12} \neq 0$, and $p_{22} \neq 0$. Then, using this together with the
catalytic conditions, we can obtain from Eq. (5.7) that

$$
\begin{aligned}
\mu_{S}^{\prime} & =\left(\frac{\left|p_{21}\right|\left|p_{12}\right|\left|p_{22}\right|}{\left|p_{22}\right|^{2}-1}\right)^{2} \mu_{S}+\left|p_{12}\right|^{2}\left(\frac{\left|p_{21}\right|^{2}}{1-\left|p_{22}\right|^{2}} \mu_{S}\right) \\
& =\frac{\left|p_{21}\right|^{2}\left|p_{12}\right|^{2}\left|p_{22}\right|^{2}+\left(1-\left|p_{22}\right|^{2}\right)\left|p_{12}\right|^{2}\left|p_{21}\right|^{2}}{\left(1-\left|p_{22}\right|^{2}\right)^{2}} \mu_{S}=\frac{\left|p_{12}\right|^{2}\left|p_{21}\right|^{2}}{\left(1-\left|p_{22}\right|^{2}\right)^{2}} \mu_{S}
\end{aligned}
$$

Also, we can obtain that

$$
\alpha_{S}^{\prime}=\left(\frac{p_{21} p_{12} p_{22}}{p_{22}^{2}-1}\right)^{2} \alpha_{S}+p_{12}^{2}\left(\frac{p_{21}^{2}}{1-p_{22}^{2}} \alpha_{S}\right)=\frac{p_{12}^{2} p_{21}^{2}}{\left(1-p_{22}^{2}\right)^{2}} \alpha_{S}=\frac{p_{12}^{2}}{p_{21}^{2}} \times \frac{p_{21}^{4}}{\left(1-p_{22}^{2}\right)^{2}} \alpha_{S}
$$

where we assume $p_{21} \neq 0$ here $^{7}$. Since $p_{21}^{2} /\left(1-p_{22}^{2}\right)=\alpha_{C} / \alpha_{S}= \pm \mu_{C} / \mu_{S}= \pm\left|p_{21}\right|^{2} /\left(1-\left|p_{22}\right|^{2}\right) \in$ $\mathbb{R}$, this leads to

$$
\alpha_{S}^{\prime}= \pm \frac{p_{12}^{2}}{p_{21}^{2}} \frac{\left|p_{21}\right|^{4}}{\left(1-\left|p_{22}\right|^{2}\right)^{2}} \alpha_{S}= \pm\left|\frac{p_{12}^{2}}{p_{21}^{2}}\right| \frac{\left|p_{21}\right|^{4}}{\left(1-\left|p_{22}\right|^{2}\right)^{2}} \alpha_{S}=+\frac{\left|p_{12}\right|{ }^{2}\left|p_{21}\right|^{2}}{\left(1-\left|p_{22}\right|^{2}\right)^{2}} \alpha_{S}
$$

where we used the fact that the factor must be real in the second equality, and we chose a + sign instead of a $-\operatorname{sign}$ at the end as $\alpha$ is always non-negative. This proves that $\alpha_{S}^{\prime} / \alpha_{S}=\mu_{S}^{\prime} / \mu_{S}$.

We have to look at the remaining cases: (i) When $p_{12}=0$, then either $p_{11}$ or $p_{21}$ must be zero because of the no-correlation conditions in Eq. (5.12). When $p_{12}=p_{11}=0, \mu_{S}^{\prime}=\alpha_{S}^{\prime}=0$, which means that the final state is thermal. When $p_{12}=p_{21}=0$, then $\mu_{S}^{\prime}=\left|p_{11}\right|^{2} \mu_{S}$, and $\alpha_{S}^{\prime}=\left|p_{11}\right|^{2} \alpha_{S}$ with $\left|p_{11}\right|^{2} \leq 1$. (ii) When $p_{22}=0$, again either $p_{11}$ or $p_{21}$ must be zero. If $p_{22}=p_{11}=0, \mu_{S}^{\prime}=\left|p_{12}\right|^{2}\left|p_{21}\right|^{2} \mu_{S}$, and $\alpha_{S}^{\prime}=\left|p_{12}\right|^{2}\left|p_{21}\right|^{2} \alpha_{S}$ with $\left|p_{12}\right|^{2}\left|p_{21}\right|^{2} \leq 1$. If $p_{22}=p_{21}=0$, the catalyst is thermal, resulting in a normal non-catalytic single-mode GTO transformation. (iii) When $\left|p_{22}\right|^{2}=1$, this implies $p_{12}=p_{21}=0$, so that $\mu_{S}^{\prime}=\left|p_{11}\right|^{2} \mu_{S}$, and $\alpha_{S}^{\prime}=\left|p_{11}\right|^{2} \alpha_{S}$ with $\left|p_{11}\right|^{2} \leq 1$. All three cases satisfy the conditions in Theorem 10 .

When $\mu_{S}=0$, from the catalytic conditions, we have either $\left|p_{22}\right|^{2}=1$ or $\mu_{C}=0$, and from the no-correlation conditions we have either $p_{12}=0$ or $p_{22}=0$ or $\mu_{C}=0$. It is not difficult to check that in all possible cases, it holds that $\mu_{S}^{\prime}=0$ and $\alpha_{S}^{\prime} \leq \alpha_{S}$, which satisfies the conditions. A similar argument also holds for the case when $\alpha_{S}=0$.
(Sufficient conditions) If we define $\gamma \leq 1$ such that $\left|\mu_{S}^{\prime}\right|=\gamma\left|\mu_{S}\right|$ and $\left|\alpha_{S}^{\prime}\right|=\gamma\left|\alpha_{S}\right|$, the state

[^25]transition condition can be expressed in terms of CMs as
$$
\sigma_{S}^{\prime}=\gamma \sigma_{S}+(1-\gamma) \sigma_{B}
$$
where $\sigma_{B}$ is the CM of the bath mode. Thus, the state transformation described by strong catalytic GTO transformations is mixing the initial CM of the system with the thermal bath, which can be achieved by a non-catalytic single-mode GTO transformation. As non-catalytic single-mode GTOs are included in single-mode strong catalytic GTOs, any final state satisfying the conditions in Eq. (5.3) or Eq. (5.4) can be also achieved by single-mode strong catalytic GTO transformations.

### 5.5 Multi-mode with catalysts

In the last section, we looked at possible state transformations under catalytic GTOs when each of the system and the catalyst is single-mode. We now want to explore state transformations under GTOs involving systems and catalysts consisting of an arbitrary number of modes. Given the difficulty of completely characterising the equivalence of covariance matrices under the set of passive transformations, the same analysis used for solving the single-mode case does not generalise to the multi-mode case. Thus, in this section, we employ different approaches to determine necessary conditions for catalytic GTO transformations in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$.

### 5.5.1 Approximate transformations with strong catalysts

In the last section, it was found that a strong catalyst permits no more state transformations than are possible without a catalyst. Would it be because of the strict requirement that the catalyst should be returned exactly to its initial state? To investigate this, we relax the condition to allow for an error with respect to the trace distance between covariance matrices. This choice of error measure is partly for mathematical convenience, but it is also more experimentally relevant as the statistical moments, rather than the states represented by a density operator, are the accessible quantities in experiments.

Definition 7. (Approximate transformations) We write $M \xrightarrow[\delta]{G T O} M^{\prime}$ if there exists $\widetilde{M}$ such that $M \xrightarrow{G T O} \widetilde{M}$ and $\left\|M^{\prime}-\widetilde{M}\right\|_{1} \leq \delta$, where $\|\cdot\|_{1}$ is the trace norm. An approximate strong catalytic transformation then takes the form

$$
M_{S} \oplus M_{C} \xrightarrow[\delta]{G T O} M_{S}^{\prime} \oplus M_{C}
$$

Note that $\delta$-closedness in terms of the matrix $M$ implies $\delta$-closedness in terms of the eigenvalues since the Wielandt-Hoffman inequality [82, Corollary 7.4.9] for the trace norm says

$$
\sum_{i=1}^{n}\left|\mu_{i}^{\prime}-\mu_{i}\right| \leq\left\|M^{\prime}-M\right\|_{1},
$$

where as usual we assume non-increasing ordering for $\boldsymbol{\mu}$ and $\boldsymbol{\mu}^{\prime}$. In the following theorem, we identify the necessary conditions for state transformations under approximate strong catalytic GTOs when the system and the catalyst consist of multiple modes.

Theorem 12. Let $[x]^{+}:=\max \{x, 0\}$ denote the positive part of a real number. Then, for given matrix $M$ and $M^{\prime}$, there exists $M_{C}$ such that $M \oplus M_{C} \xrightarrow[\delta]{G T O} M^{\prime} \oplus M_{C}$ if and only if

$$
\begin{equation*}
\sum_{i}\left[\mu_{i}^{\prime+}-\mu_{i}^{+}\right]^{+}+\sum_{i}\left[\mu_{i}^{\prime-}-\mu_{i}^{-}\right]^{+} \leq \delta . \tag{5.13}
\end{equation*}
$$

Similarly, there exists $A_{C}$ such that $A \oplus A_{C} \xrightarrow[\delta]{G T O} A^{\prime} \oplus A_{C}$ if and only if

$$
\begin{equation*}
\sum_{i}\left[\alpha_{i}^{\prime}-\alpha_{i}\right]^{+} \leq \delta \tag{5.14}
\end{equation*}
$$

The proof can be found in Appendix B.1.
Note that setting $\delta=0$ recovers the same transformation laws as the non-catalytic case described in Theorem 9. Otherwise, the LHSs of Eqs. (5.13) and (5.14) quantify the total amounts by which the laws are violated - and these totals are bounded by the error $\delta$. Also, note that there is no possibility for embezzlement in this case - such a phenomenon occurs, for instance, in entanglement theory [84], whereby allowing for a small error in returning the catalyst state can permit arbitrary transformations on the system. Rather, the result that a
strong catalyst cannot enable any additional transformation is stable with respect to errors in the catalyst.

### 5.5.2 Necessary majorisation conditions for weak catalysts

For a weak catalyst, instead of the ordering relation denoted by the inequality symbol ( $\leq$ or $\geq$ ) we find that majorisation ( $\succ$ or $\prec$ ) and weak majorisation $\left(\succ_{w}\right.$ or $\prec_{w}$ ) become relevant. For given two non-increasingly ordered vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of length $n$, we say that $\boldsymbol{y}$ weakly majorises $\boldsymbol{x}$, written as $\boldsymbol{x} \prec_{w} \boldsymbol{y}$, if [85]

$$
\sum_{i=1}^{k} x_{i} \leq \sum_{i=1}^{k} y_{i} \quad \forall k=1,2, \ldots, n
$$

When in addition the total sum is the same, i.e.,

$$
\sum_{i=1}^{n} x_{i}=\sum_{i=1}^{n} y_{i}
$$

then recall that we say $\boldsymbol{y}$ majorises $\boldsymbol{x}$, denoted by $\boldsymbol{x} \prec \boldsymbol{y}$.
The transformation laws for $M$ and $A$ now differ from each other and from the non-catalytic case. A weak catalyst on its own, without involving a thermal bath, is also able to make some influence on $\boldsymbol{\mu}$ as follows.

Theorem 13. 1. A weak catalytic transformation from $M$ to $M^{\prime}$ is possible without use of a thermal bath if and only if

$$
\boldsymbol{\mu}^{\prime} \prec \boldsymbol{\mu} .
$$

If the system has $n$ modes, then the catalyst only needs at most $n-1$ modes.
2. Allowing for the use of a thermal bath, the transformation is possible if and only if

$$
\boldsymbol{\mu}^{\prime+} \prec_{w} \boldsymbol{\mu}^{+} \text {and } \boldsymbol{\mu}^{\prime-} \prec_{w} \boldsymbol{\mu}^{-} .
$$

The catalyst only needs at most $n-1$ modes, and the bath needs at most $n$ modes.

The proof is given in Appendix B.2. Note that no statement can be made about the number of positive or negative principal mode temperatures. It is also worth noting that the construction of the operation in Theorem 13-2 shows that we may interact the system with just the catalyst first, and then just with the bath as in the single-mode case in the last section.

For the principal mode asymmetries $\boldsymbol{\alpha}$, we find instead that the same set of transformation is possible, whether or not a thermal bath is involved, as described in the following theorem:

Theorem 14. A weak catalytic transformation from $A$ to $A^{\prime}$ is possible either with or without use of a thermal bath if and only if

$$
\boldsymbol{\alpha}^{\prime} \prec_{w} \boldsymbol{\alpha} .
$$

If the system has $n$ modes, we require at most $2 n-1$ catalyst modes or otherwise $n$ catalyst modes plus $n-1$ bath modes.

The proof is given in Appendix B.3.
The key additional transformation provided by a weak catalyst is an operation on a given pair of modes. For instance, the result on the $M$ matrix is a so-called T-transform [85, Lemma 2.B.1] on its eigenvalues, resulting in $\mu_{1}^{\prime}=(1-t) \mu_{1}+t \mu_{2}, \mu_{2}^{\prime}=(1-t) \mu_{2}+t \mu_{1}$ for some $t \in[0,1]$.

Again it is important to remark that the above conditions are only necessary conditions for state transformation due to the fact that $M$ and $A$ are in general not simultaneously diagonalisable. Also, the derived multi-mode conditions do not fully recover the single-mode results presented in Section 5.4 - we are missing the conditions on the ratio between $\mu$ and $\alpha$. It is an open question to see whether we can derive a generalised version of the condition on the ratio in the multi-mode case.

### 5.6 Physical implications

In this section, we would like to discuss the physical implications of the monotonicity results found for the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$, the principal mode temperatures and asymmetries. We summarise our results for different cases in Table 5.1 and also illustrate them in Figure 5.3.

Table 5.1: Summary of ordering relations on the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ for different types of GTO transformations.

| Transformation type | Ordering |
| :---: | :---: |
| No catalyst | $\boldsymbol{\mu}^{\prime \pm} \leq \boldsymbol{\mu}^{ \pm}, \quad \boldsymbol{\alpha}^{\prime} \leq \boldsymbol{\alpha}$ |
| Strong catalyst | $\boldsymbol{\mu}^{\prime \pm} \leq \boldsymbol{\mu}^{ \pm}, \quad \boldsymbol{\alpha}^{\prime} \leq \boldsymbol{\alpha}$ |
| Weak catalyst, no bath | $\boldsymbol{\mu}^{\prime} \prec \boldsymbol{\mu}, \quad \boldsymbol{\alpha}^{\prime} \prec_{w} \boldsymbol{\alpha}$ |
| Weak catalyst, with bath | $\boldsymbol{\mu}^{\prime \pm} \prec_{w} \boldsymbol{\mu}^{ \pm}, \quad \boldsymbol{\alpha}^{\prime} \prec_{w} \boldsymbol{\alpha}$ |



Figure 5.3: Lorenz curves (i.e., partial sums) demonstrating the different kinds of orderings encountered here. If the top curve (solid, black) is $\boldsymbol{x}$, then the lower curves from top to bottom satisfy respectively $\boldsymbol{x}^{\prime} \prec \boldsymbol{x}$ (solid, red), $\boldsymbol{x}^{\prime} \prec_{w} \boldsymbol{x}$ (dashed, blue), and $\boldsymbol{x}^{\prime} \leq \boldsymbol{x}$ (dotted, green).

### 5.6.1 Concentration of athermal resources

Let us first consider the matrix $M$. Given a choice of orthogonal modes, the diagonal $M_{i i}$ in the corresponding basis are related to the mean energy in each mode, under the assumption of vanishing first moments. The principal mode temperature $\mu_{i}$ are the diagonals in a basis where $M$ is diagonal. Due to the majorisation relation between the eigenvalues of a matrix and its diagonal, the distribution of $\left\{M_{i i}\right\}_{i}$ is generally more uniform than that of $\left\{\mu_{i}\right\}_{i}$. For a set of initially uncorrelated modes isolated from the thermal environment and which interact via a passive transformation, this relation describes an approach to equilibrium via what could be described as heat exchange, since these components of energy are related to the quadrature fluctuations. The $\left\{\mu_{i}\right\}_{i}$ are intrinsic to the closed system, so are unchanged over this unitary evolution.

An important common feature apparent in the relations in Table 5.1 is a restriction on the concentration of resources. For instance, all the orderings $\leq, \prec, \prec_{w}$ require $\mu_{1}^{\prime+} \leq \mu_{1}^{+}$.

This means that it is impossible to concentrate the energy of multiple modes in such a way that the 'hottest' mode becomes hotter. The same rule applies inversely to modes below the background temperature; the coldest sub-thermal mode cannot be cooled. Therefore, an absorption refrigerator is an impossible machine within the Gaussian framework. This observation was also made in other literature [86-88] - however, we also see further that more subtle conditions apply to concentration into larger subsets of modes. For instance, either with no catalyst or with a strong catalyst, $\mu_{2}^{++} \leq \mu_{2}^{+}$, and thus the second hottest mode also cannot become hotter. A weak catalyst instead opens up the possibility of heating this mode, subject to a different constraint that the two hottest modes in total do not heat up, i.e., $\mu_{1}^{\prime+}+\mu_{2}^{\prime+} \leq \mu_{1}^{+}+\mu_{2}^{+}$. That is, energy in the first mode can be traded for heating of the second mode.

Note that the case of weak catalysts without interaction with a thermal bath makes no distinction between super- and sub-thermal modes; this is to be expected since there is no background reference temperature in this case. The effect of including a thermal bath is then to separate the conditions on super- and sub-thermal modes, and also to allow for losses, in the sense that the total $\sum_{i=1}^{n_{+}} \mu_{i}^{+}$can decrease, for instance. So the super-thermal modes can in total lose energy to the bath, and the sub-thermal modes can in total gain energy from the bath.

Similar considerations apply to the principal mode asymmetries $\alpha_{i}$. A notable difference compared with $\mu_{i}$ is that the 'lossy' property exists for a weak catalyst with no bath interaction. Intuitively, this is due to the existence of operations that result in components of the asymmetry being transferred to correlations between the system and the catalyst.

### 5.6.2 General limitations of Gaussian systems

In addition to the constraints described above, there is a deeper limitation preventing Gaussian systems from making useful thermodynamical machines. In order to develop this argument, we first need to understand how to describe work in the GTO framework. In the quantum resource theory of discrete thermal operations [35, 89], a work battery is typically defined as a system that transitions from one pure energy eigenstate to another so that it does not change entropy, and a definite, non-fluctuating amount of work is exchanged with other systems. This definition cannot be used in GTOs since no energy eigenstate other than the vacuum is Gaussian. Given
the significance of quadratures in the Gaussian setting, it seems reasonable to instead define a Gaussian work battery as a system that transitions from one Gaussian state to another under the action of a displacement - thus, only its first moments can change, but not its CM. The work is then defined as the change in mean energy of the battery. Apart from the change in first moments, the energy distribution changes too, so this is necessarily a weakening of the usual requirements. However, its entropy is a function of the CM and is constant.

Depending on whether the work battery is allowed to become correlated with the system, it therefore functions at the level of the CM as either a strong or a weak catalyst. Either way, the final CM of the system is only a function of the initial CM and completely independent of the first moments. Thus, the possible transformations of the system's CM are the same no matter how much energy is stored in the battery. In this Gaussian setting, therefore, no useful machine can be constructed whereby work is traded for resources at the CM level. What can a Gaussian work battery do then? Only non-trivial transformations involving first moments alone are possible. In general, given an $N_{S}$-mode system $S$ and an $N_{W}$-mode work battery $W$ with first moments $\overline{\mathbf{r}}_{S}$ and $\overline{\mathbf{r}}_{W}$, a joint passive transformation represented by $S \in K\left(N_{S}+N_{W}\right)$ has the action

$$
\binom{\overline{\mathbf{r}}_{S}^{\prime}}{\overline{\mathbf{r}}_{W}^{\prime}}=S\binom{\overline{\mathbf{r}}_{S}}{\overline{\mathbf{r}}_{W}}
$$

i.e., the point in the combined phase pace is simply rotated about the origin by an orthogonal symplectic matrix. For example, with one mode each, a beam-splitter with reflectivity $r$ has the effect on a pair of coherent states: $\left|\alpha_{1}\right\rangle_{S}\left|\alpha_{2}\right\rangle_{W} \rightarrow\left|r \alpha_{1}+\sqrt{1-r^{2}} \alpha_{2}\right\rangle_{S}\left|r \alpha_{1}-\sqrt{1-r^{2}} \alpha_{2}\right\rangle_{W}$. In general, a displacement can be applied to the system $S$ by choosing a battery $W$ with the same CM as the system's and suitable chosen first moments along with a set of beam-splitters. Thus, while displacements are not free within GTO, they can be accomplished with an additional system that exchanges energy in the form of first moments but is strongly catalytic in its second moments. Alternatively, if we do not want to adapt the system $W$ to fit the CM of the system $S$, we can use an arbitrary CM and a large displacement $\alpha$; in the limit of $\alpha \rightarrow \infty$ with $r \times|\alpha|$ held constant, a displacement is performed on the system $S$ [90]. For finite $\alpha$, this is an instance of an approximate strong catalytic transformation, so Theorem 12 guarantees that the thermal
monotone laws found in the last section hold approximately.
At a broader level, it is also worthwhile noting that the direct-sum structure of multi-mode Gaussian systems seems crucial for these general limitations in the Gaussian setting. The main reason why introducing catalysts does not have as a dramatic impact on the ordering relations of $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ as in other resource theories is that catalysts do not change the pre-existing elements of $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ but just add more elements to such vectors. In contrast, in the tensor-product structure of the full Hilbert space, the catalyst's vector elements are multiplied with the system's elements, which can significantly affect vector ordering. The direct-sum structure of multi-mode Gaussian systems also plays an important role in other limitations to Gaussian resource theories as we have seen in Chapter 4; the tensorisation property of a resource monotone (Part 3 in Lemma 4 in Chapter 4), due to the direct-sum structure of CMs, results in the limitations for Gaussian resource distillation - having multiple input states does not help us distill more resources.

### 5.7 Summary

In this chapter, we looked at possible state transformations under Gaussian thermal operations when catalysts are allowed. We defined two different types of catalysts: a strong catalyst which must not only come back to the original state but also end up uncorrelated to the system, and a weak catalyst whose local state only needs to be reset. We asked whether exploiting these catalysts permits more state transformations than the non-catalytic case. Despite the general benefit of catalysts in other QRTs, we found that strong catalysts do not enable more state transformations while weak catalysts can achieve more but limited state transformations compared to the non-catalytic case. Alongside with the principal mode temperature introduced in [80], we defined another resource monotone, the principal mode asymmetry, to describe the state transition conditions. We fully characterised the necessary and sufficient conditions for state transformations under single-mode catalytic GTOs for both types of catalysts, and also provided new necessary conditions for the multi-mode case in terms of principal mode temperatures and symmetries. We discussed physical implications of our results on Gaussian
thermal machines such as Gaussian refrigerators and Gaussian work batteries.
In general, due to the stringent requirement imposed by the direct-sum composition of the phase space, resource theories in the Gaussian regime allow for very limited operations. It is therefore natural to inject non-Gaussian elements into the theory. In this regard, it would be interesting to see what transformation would be unlocked if Gaussian thermal operations are combined with resourceful non-Gaussian catalysts. The conditions determined in this chapter would still hold at the level of second moments (since Gaussian states allow one to reproduce all physical CMs, and their transformations would still be same as described here), but more interesting dynamics could be allowed at the full Hilbert space level. More fundamentally, however, non-Gaussian operations need to be included as resources to break the constraints we have found on the manipulation of second moments.

## Part II

## Semidefinite Programming Hierarchies

## for Quantum Information

## Chapter 6

## Background theory

In Part I, we considered continuous-variable quantum systems associated with infinite-dimensional Hilbert spaces. In Part II, we go back to the more standard setting of quantum information science and consider discrete-variable quantum systems with finite-dimensional Hilbert spaces.

In this chapter, we dive into new topics: semidefinite programming and quantum de Finetti theorems. These two concepts are used as main technical tools in Chapter 7 to characterise quantum correlations of fixed dimensions in the context of non-local games. The first topic, semidefinite programming, is a special type of mathematical optimisation problem involving positive semidefinite matrices. As quantum states are mathematically represented by positive semidefinite matrices with trace one, semidefinite programming can be easily found in many different problems of quantum information theory when optimisation is involved. Indeed, it is known that some of the famous problems in quantum information theory, such as quantum separability problem and characterising quantum correlations, admit semidefinite programming relaxations which can be effectively implemented. In this chapter, we formally introduce the concept of semidefinite programming and the two important, widely-exploited semidefinite programming hierarchies in quantum information science; the Doherty-Parrilo-Spedalieri (DPS) and Navascués-Pironio-Acín (NPA) hierarchy. Then, we move to the second topic, quantum de Finetti theorems. Quantum de Finetti theorems address the close relation between quantum separable states and effective outer-approximations of separable states known as extendible states. Extendible states and quantum de Finetti theorems are key ingredients in the next


Figure 6.1: The schematic of how each ingredient will be used in the next chapter, where we construct a new hierarchy of semidefinite programming relaxations for quantum correlations with fixed dimension.
chapter, and we provide a short introduction on them in this chapter.
To motivate the reader through this chapter, we provide a brief summary (see Figure 6.1) explaining how each ingredient introduced in this chapter will play their role in the next chapter. The goal of Chapter 7 is to develop a new way to characterise quantum correlations which can be realised by measuring quantum states on a Hilbert space of fixed dimension. We achieve this by constructing a new hierarchy of semidefinite programming relaxations (the topic of Section 6.3) for quantum correlations of fixed dimension. Firstly, we rephrase the problem as an instance of quantum separability problem and construct a new semidefinite programming hierarchy using the DPS hierarchy (Section 6.3.1), which is well-known semidefinite programming relaxations for quantum separability problem. Then, we derive the convergence rate of the newly constructed hierarchy using quantum de Finetti theorem (Section 6.4) - this is possible since quantum de Finetti theorems can be related to the quantitative accuracy of the DPS hierarchy. Lastly, we also exploit the NPA hierarchy (Section 6.3.2), which is a different hierarchy of semidefinite programming relaxations for general quantum correlations without dimension constraint, and combine it with our hierarchy to achieve a better result.

This chapter is structured as follows. In Section 6.1, we first provide some mathematical background to study convex optimisation. In Section 6.2, we formally define general optimisation
problems and convex optimisation, a sub-class of mathematical optimisation problem which admits an effective algorithm for solving them, and introduce semidefinite programming as a special case of it. We also discuss their useful mathematical properties. Then, we move on to the topic of semidefinite programming relaxations in Section 6.3 and introduce two famous examples, the DPS and NPA hierarchy. Finally, in Section 6.4, we provide a brief introduction of quantum de Finetti theorems.

### 6.1 Convex sets and convex functions

## Convex sets

In geometry, a set $C$ is called convex if all points 'between' any two points in $C$ lies in $C$, i.e., if for any $\boldsymbol{x}_{1}, \boldsymbol{x}_{2} \in C$ and any $\mu \in[0,1]$, we have

$$
\begin{equation*}
\mu \boldsymbol{x}_{1}+(1-\mu) \boldsymbol{x}_{2} \in C . \tag{6.1}
\end{equation*}
$$

A closed and bounded convex set is particularly called a convex body.
We call a point of the form in Eq. (6.1) a convex combination of $\boldsymbol{x}_{1}$ and $\boldsymbol{x}_{2}$, and such points represent the points between or mixtures of $\boldsymbol{x}_{1}$ and $\boldsymbol{x}_{2}$. The concept of convex combination can be straightforwardly extended to an arbitrary number of points: for given points $\boldsymbol{x}_{1}, \ldots, \boldsymbol{x}_{n}$, we define convex combinations of these points as $\boldsymbol{x}$ such that

$$
\boldsymbol{x}=\mu_{1} \boldsymbol{x}_{1}+\ldots+\mu_{n} \boldsymbol{x}_{n}, \quad \mu_{1}+\ldots+\mu_{n}=1, \quad \mu_{i} \geq 0 \forall i .
$$

When an infinite number of points are given, we can generalise the concept to include infinite sums or integrals: for given $\left\{\boldsymbol{x}_{i}\right\}_{i}$,

$$
\sum_{i=1}^{\infty} \mu_{i} \boldsymbol{x}_{i}, \quad \sum_{i=1}^{\infty} \mu_{i}=1, \quad \mu_{i} \geq 0 \forall i
$$

or for some probability distribution $p(\boldsymbol{x})^{1}$ over points $\boldsymbol{x}$ in a given set $S$

$$
\int_{S} p(\boldsymbol{x}) \boldsymbol{x} \mathrm{d} \boldsymbol{x}
$$

The convex hull of a set $S$, denoted by $\operatorname{conv}(S)$, is defined as the set of all convex combinations of points in $S$, i.e.,

$$
\operatorname{conv}(S):=\left\{\mu_{1} \boldsymbol{x}_{1}+\ldots+\mu_{n} \boldsymbol{x}_{n} \mid \boldsymbol{x}_{i} \in S, \mu_{1}+\ldots+\mu_{n}=1, \mu_{i} \geq 0 \forall i\right\}
$$

A convex hull $\operatorname{conv}(S)$ is always convex and the smallest convex set containing $S$.
A set $C$ is called a cone, if for every $\boldsymbol{x} \in C$ and $\lambda \geq 0$, we have $\lambda \boldsymbol{x} \in C$. A cone is called a convex cone if it is convex, which implies that for any $\boldsymbol{x}_{1}, \boldsymbol{x}_{2} \in C$ and $\lambda_{1}, \lambda_{2} \geq 0$, we have $\lambda_{1} \boldsymbol{x}_{1}+\lambda_{2} \boldsymbol{x}_{2} \in C$. This concept can be extended to a generalised inequality such as the matrix inequality. An important example is the cone of positive semidefinite matrices which can be defined as the set of all matrices $A$ such that $A \succeq 0$, where we denote the matrix inequality by $\succeq$ (or $\preceq$ ) to avoid the confusion with the scalar inequality. It is easy to check that this set satisfies the property of convex cones with respect to the matrix inequality.

Convex bodies always have some special points which cannot be generated by convex combinations of other points. We call these special points extreme points, or often pure points in the physics community. The extreme points are very useful for characterising a convex set thanks to the following two theorems we introduce here.

Theorem 15. (Minkowski's theorem) Any convex body is the convex hull of its extreme points.

Theorem 16. (Carathéodory's theorem) Any point in an n-dimensional convex set $C$ can be expressed as a convex combination of at most $n+1$ extreme points of $C$.

The dimension of a convex set is defined as the smallest dimension of an affine subspace containing the convex set.

Convex sets are an important concept in quantum information theory. Firstly, the set of quantum states, the primary object of quantum mechanics, is a convex set. This reflects the

[^26]physical observation that any probabilistic mixture of valid states of a quantum system should be also a physical state. Also, sets of free states in many QRTs are convex sets, which is a result of treating classical mixing as an allowed operation for quantum tasks. Various sets of correlations considered in the non-locality theory [60], such as the sets of local correlations, quantum correlations and no-signalling correlations, are all examples of convex sets.

## Convex functions

A real valued function $f: \mathbb{C}^{n} \rightarrow \mathbb{R}$ is called convex if the domain of $f$, denoted by $\operatorname{dom}(f)$, is a convex set, and if for all $\boldsymbol{x}_{1}, \boldsymbol{x}_{2} \in \operatorname{dom}(f)$ and $\mu \in[0,1]$ we have

$$
\begin{equation*}
f\left(\mu \boldsymbol{x}_{1}+(1-\mu) \boldsymbol{x}_{2}\right) \leq \mu f\left(\boldsymbol{x}_{1}\right)+(1-\mu) f\left(\boldsymbol{x}_{2}\right) . \tag{6.2}
\end{equation*}
$$

If the strict inequality in Eq. (6.2) holds whenever $\boldsymbol{x}_{1} \neq \boldsymbol{x}_{2}$, the function is called strictly convex. Similarly, a function is called concave if the function $-f$ is convex, and strictly concave if $-f$ is strictly convex. Geometrically, the convexity in Eq. (6.2) means that for any two points $\boldsymbol{x}_{1}, \boldsymbol{x}_{2} \in \operatorname{dom}(f)$ the line connecting between $\left(\boldsymbol{x}_{1}, f\left(\boldsymbol{x}_{1}\right)\right)$ and $\left(\boldsymbol{x}_{2}, f\left(\boldsymbol{x}_{2}\right)\right)$ lies above the graph of $f$.

One relevant class of functions over the real numbers is affine functions which are defined as functions consisting of a sum of linear functions and a constant; for example, functions $f: \mathbb{R}^{n} \rightarrow \mathbb{R}^{m}$ of the form $f(\boldsymbol{x})=A \boldsymbol{x}+\boldsymbol{b}$ where $A \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^{m}$. Affine functions satisfy the equality in Eq. (6.2) and thus are both convex and concave. The converse holds as well - any function that is convex and concave at the same time is affine. As affine functions include all linear functions, linear functions are also convex and concave. Affine functions have a nice property that for some convex set $S \subset \operatorname{dom}(f)$, the image of $S$ under $f$ is convex.

### 6.2 Semidefinite programming

In mathematics, an optimisation problem is a problem of finding an optimal solution from all feasible solutions. Such problems can be naturally seen in all kinds of disciplines including not only physics and computer science but also engineering, economics, product development
etc. Finding good methods for optimisation problems has been therefore of interest for many centuries. As a result of the long-running research, we now have some types of optimisation problems which are well-studied and admit effective methods for solving them. We look at one example in this section, called convex optimisation, and its important sub-class, semidefinite programming.

## Convex Optimisation

We consider a problem of the form

$$
\begin{align*}
\min _{\boldsymbol{x} \in \mathbb{C}^{N}} & f_{0}(\boldsymbol{x}) \\
\text { subject to } & f_{i}(\boldsymbol{x}) \leq 0 \quad i \in(1, \ldots, m)  \tag{6.3a}\\
& h_{i}(\boldsymbol{x})=0 \quad i \in(1, \ldots, p), \tag{6.3b}
\end{align*}
$$

which describes the problem of finding an optimal $\boldsymbol{x}$ that minimises $f_{0}(\boldsymbol{x})$ and satisfy the constraints in Eq. (6.3a)-(6.3b) at the same time. We call $\boldsymbol{x} \in \mathbb{C}^{N}$ the optimisation variable and the function $f_{0}: \mathbb{C}^{N} \rightarrow \mathbb{R}$ the objective function. Eq. (6.3a)-(6.3b) are called the optimisation constraints or simply constraints, and the functions $\left\{f_{i}\right\}_{i}$ and $\left\{h_{i}\right\}_{i}$ are called the constraint functions. The domain $D$ of the optimisation problem is defined as $D=\left(\bigcap_{i=0}^{m} \operatorname{dom}\left(f_{i}\right)\right) \cap$ $\left(\bigcap_{i=1}^{p} \operatorname{dom}\left(h_{i}\right)\right)$. A point $\boldsymbol{x} \in D$ is called feasible if it satisfies constraints Eq. (6.3a)-(6.3b). The feasible set is then defined as

$$
\left\{\boldsymbol{x} \mid f_{i}(\boldsymbol{x}) \leq 0, i \in(1, \ldots, m), h_{i}(\boldsymbol{x})=0, i \in(1, \ldots, p)\right\}
$$

which describes the set of all feasible points. The optimal value of the optimisation problem in Eq. (6.3) is the value $\omega$ such that

$$
\omega=\inf \left\{f_{0}(\boldsymbol{x}) \mid f_{i}(\boldsymbol{x}) \leq 0, i \in(1, \ldots, m), h_{i}(\boldsymbol{x})=0, i \in(1, \ldots, p)\right\}
$$

$\omega$ can take the values $\pm \infty$, and if the problem is infeasible, we have $\omega=\infty .^{2}$ Some $\tilde{\boldsymbol{x}}$ is an optimal solution or optimal point when $\tilde{\boldsymbol{x}}$ is feasible and $f_{0}(\tilde{\boldsymbol{x}})=\omega$. There can be more than one optimal solutions, and we call the set of all optimal solutions the optimal set.

While mathematical optimisation is in general difficult to solve, there are some classes of optimisation which are known to admit effective algorithms. One example is convex optimisation. A convex optimisation problem is a problem of the form in Eq. (6.3) where (i) the objective function $f_{0}$ is convex, (ii) the constraint functions $\left\{f_{i}\right\}_{i=1}^{m}$ in Eq. (6.3a) are convex, and (iii) the constraint functions $\left\{h_{i}\right\}_{i=1}^{p}$ are affine functions, i.e., $h_{i}(\boldsymbol{x})=\boldsymbol{a}_{i}^{\mathrm{T}} \boldsymbol{x}-b_{i}$. The immediate implication of these conditions is that the feasible set of a convex optimisation problem must be convex. Thus, a convex optimisation problem is a problem of optimising a convex objective function over a convex set. Another convenient property of convex optimisation is that any local optimal point is also a global optimum, which makes finding an optimal solution easier. We can also consider a problem of the form

$$
\begin{align*}
\max _{\boldsymbol{x} \in \mathbb{C}^{N}} & f_{0}(\boldsymbol{x}) \\
\text { subject to } & f_{i}(\boldsymbol{x}) \leq 0 \quad i \in(1, \ldots, m)  \tag{6.4}\\
& \boldsymbol{a}_{i}^{\mathrm{T}} \boldsymbol{x}=b_{i} \quad i \in(1, \ldots, p) .
\end{align*}
$$

If the objective function $f_{0}$ is concave, and the constraint functions $\left\{f_{i}\right\}_{i}$ in Eq. (6.4) are convex, we also treat the problem in Eq. (6.4) as a convex optimisation problem. This is because this concave maximisation problem can be solved by minimising the convex function $-f_{0}$ over the same convex feasible set.

Although there is in general no analytical formula to solve such convex optimisation problems, there exist methods which are known to be very effective for them, such as interior-point method. These methods work very well in practice, and for some special cases, it can be proven that these methods provide polynomial-time algorithms for solving the cases. It is worth remarking that an optimal value of a convex optimisation problem with affine objective function can be obtained by an extreme point of the feasible set. To show this, let us consider a convex

[^27]optimisation problem of the form in Eq. (6.3) with affine function $f_{0}$ and the set of the extreme points of its feasible set, $\left\{\boldsymbol{e}_{i}\right\}_{i}$. We denote by $\tilde{\boldsymbol{e}}$ the optimal extreme point among $\left\{\boldsymbol{e}_{i}\right\}_{i}$ such that $f_{0}(\tilde{\boldsymbol{e}}) \leq f_{0}\left(\boldsymbol{e}_{i}\right)$ for all $\boldsymbol{e}_{i} \in\left\{\boldsymbol{e}_{i}\right\}_{i}$. Then, since the objective function $f_{0}$ is affine, for any convex combination of the extreme points, $\boldsymbol{x}=\sum_{i} p_{i} \boldsymbol{e}_{i}$, we have
$$
f_{0}(\boldsymbol{x})=\sum_{i} p_{i} f_{0}\left(\boldsymbol{e}_{i}\right) \geq f_{0}(\tilde{\boldsymbol{e}}) .
$$

Due to Carathéodory's theorem (Theorem 16), any feasible point can be expressed as a convex combination of $n+1$ extreme points where $n$ is the dimension of the feasible set, and thus $f_{0}(\tilde{\boldsymbol{e}})$ is the global lower bound for the whole feasible set, which proves the claim. This observation implies that we can focus on the extreme points rather than the whole feasible set when we solve a convex optimisation problem with affine objective function.

Famous examples of convex optimisation problems are linear programming and least-squares problems [91] as well as semidefinite programming, which is the main topic of the rest of this section.

## Semidefinite programming

Semidefinite programming (SDP) is a sub-class of convex optimisation where the optimisation problem includes positive semidefinite constraints for matrices; namely, the optimisation set is not only a convex set but also a subset of the cone of positive semidefinite matrices. The standard form of SDP can be written as follows: for some Hilbert space $\mathcal{H}$

$$
\begin{array}{cl}
\min _{X \in \operatorname{Herm}(\mathcal{H})} & \operatorname{Tr}[C X] \\
\text { subject to } & \Phi(X)=B  \tag{6.5}\\
& X \succeq 0,
\end{array}
$$

where (i) the matrix $C$ in the objective function is Hermitian, (ii) the constraint function $\Phi$ is a Hermiticity-preserving linear map, and (iii) $B$ is Hermitian. We can easily see that the optimisation problem of the form in Eq. (6.5) is an instance of convex optimisation. The
objective function is a linear function, which is convex, and the feasible set is the intersection between the cone of positive semidefinite matrices and the affine space satisfying the equality constraints, which is a convex set.

SDPs have an important property called the duality. Let us call the standard form of SDP in Eq. (6.5) the primal problem. Then, every primal SDP problem has an associated dual problem in the form of

$$
\begin{array}{ll}
\max _{Y \in \operatorname{Herm}(\mathcal{H})} & \operatorname{Tr}[B Y] \\
\text { subject to } & \Phi^{\dagger}(Y) \preceq C,
\end{array}
$$

where $\Phi^{\dagger}$ is the adjoint of $\Phi$ defined as $\Phi^{\dagger}$ such that $\operatorname{Tr}\left[\Phi^{\dagger}(X) A\right]=\operatorname{Tr}[X \Phi(A)]$ for all linear operators $X$ and $A$. In particular, if an operator $X$ is feasible for the primal problem, we call it primal feasible, and if an operator $Y$ is feasible for the dual problem, we call it dual feasible. Then, we have the following proposition.

Proposition 17. (Weak Duality) Consider an SDP of the form in Eq. (6.5). Let $\alpha$ be the optimal value of the primal problem and $\beta$ be the optimal value of the dual problem. Then, we have

$$
\alpha \geq \beta .
$$

Proof. If the SDP is infeasible, i.e., $\alpha=\infty$ and $\beta=-\infty$, then the proposition is trivial. If the problem is feasible, for every primal feasible $X$ and dual feasible $Y$, we have

$$
\operatorname{Tr}[C X] \geq \operatorname{Tr}\left[\Phi^{\dagger}(Y) X\right]=\operatorname{Tr}[Y \Phi(X)]=\operatorname{Tr}[Y B] .
$$

This relation holds for all primal feasible $X$ and dual feasible $Y$, and thus taking the supremum over all $X$ on the LHS and the infimum over all $Y$ on the RHS proves the claim.

When the equality condition holds, i.e., $\alpha=\beta$, we say that the problem has the strong duality.

### 6.3 Semidefinite programming relaxations

In the last section, we have studied convex optimisation, a special type of mathematical optimisation, and SDP as its sub-class. Such optimisation problems are useful tools since there exist known effective algorithms for solving many of these problems. While a general mathematical optimisation is NP-hard to solve, it is sometimes possible to derive relaxations for such difficult optimisation problem, which can provide an approximation (lower or upper bound) to the true solution but admit an effective algorithm. For example, we may replace each non-convex constraint in a given general optimisation problem with a looser but convex constraint. The resulting problem is then a convex optimisation problem which can be efficiently solved, for example, using interior point method, and may provide a lower or upper bound on the optimal value of the original general optimisation problem. In particular, when the relaxed problem is an SDP, we call it semidefinite programming relaxation. In this section, we introduce two famous semidefinite programming relaxations on important problems in quantum information theory.

### 6.3.1 The Doherty-Parrilo-Spedalieri (DPS) hierarchy

## Quantum separability problem

Entanglement is one of the key features in quantum mechanics for both practical and fundamental reasons. Thus, the problem of identifying entanglement has been one of the long-standing problems in quantum information theory. A bipartite quantum state is called separable if it can be decomposed as a convex combination of pure product states ${ }^{3}$ :

$$
\rho_{A B}=\sum_{i} p_{i}\left|\psi_{i}\right\rangle\left\langle\left.\psi_{i}\right|_{A} \otimes \mid \phi_{i}\right\rangle\left\langle\left.\phi_{i}\right|_{B},\right.
$$

for some pure quantum states $\left|\psi_{i}\right\rangle_{A} \in \mathcal{H}_{A}$ and $\left|\phi_{i}\right\rangle_{B} \in \mathcal{H}_{B}$ and probability distribution $\left\{p_{i}\right\}_{i}$. Any state that is not separable is called entangled. Identifying whether a given state is entangled

[^28]or not is then equivalent to find whether the given state admits such separable decomposition or not. This problem is often referred to as quantum separability problem although the concept of quantum separability problem is more broad and covers any problem involving identifying separable/entangled states; for example, an optimisation over the set of separable states is also called quantum separability problem since it involves the problem of identifying separable states. However, solving quantum separability problem is not very straightforward. In fact, it has been shown by Gurvits [92] that it is NP-hard. It is therefore unlikely to be able to find a computationally efficient method for quantum separability problem which scales well with the Hilbert space dimension.

There still exist not complete but effective criteria for the separability problem, and these are usually given by necessary but not sufficient conditions for quantum separability. The most well-known example is the positive partial transpose (PPT) criteria introduced by [93]. It is not difficult to observe that if a state is separable, it must have a PPT:

$$
\begin{equation*}
\rho_{A B}^{\mathrm{T}_{A}}=\sum_{i} p_{i}\left|\psi_{i}^{*}\right\rangle\left\langle\left.\psi_{i}^{*}\right|_{A} \otimes \mid \phi_{i}\right\rangle\left\langle\left.\phi_{i}\right|_{B},\right. \tag{6.6}
\end{equation*}
$$

since $\left|\psi_{i}^{*}\right\rangle$ is also a valid pure state. This implies that any state that does not have a PPT is necessarily entangled. The PPT criteria is very easy to check computationally and has been widely used as a simple entanglement test in quantum information science. Furthermore, it was shown that they form both necessary and sufficient conditions for quantum separability in $\mathcal{H}_{A} \otimes \mathcal{H}_{B}$ when $\operatorname{dim}\left(\mathcal{H}_{A}\right)=2, \operatorname{dim}\left(\mathcal{H}_{B}\right)=2$ and $\operatorname{dim}\left(\mathcal{H}_{A}\right)=3, \operatorname{dim}\left(\mathcal{H}_{B}\right)=2$ [94]. For general dimension, however, there are entangled states with PPTs, which are called bound entangled states, ${ }^{4}$ and the PPT criteria fails to give a complete characterisation of separable states.

## The DPS hierarchy

In this section, we introduce another useful criteria for quantum separability, introduced in [95]. These new criteria are easily implementable and form a 'complete' family of separability criteria in a sense that any entangled state can be identified by an element in this family.

[^29]Let us consider a bipartite separable state $\rho_{A B}$ on $\mathcal{H}_{A} \otimes \mathcal{H}_{B}$ as in Eq. (6.6). We introduce a new quantum state $\rho_{A B^{2}}$ on $\mathcal{H}_{A} \otimes \mathcal{H}_{B} \otimes \mathcal{H}_{B}$ given by

$$
\rho_{A B^{2}}=\sum_{i} p_{i}\left|\psi_{i}\right\rangle\left\langle\left.\psi_{i}\right|_{A} \otimes \mid \phi_{i}\right\rangle\left\langle\left.\phi_{i}\right|_{B} \otimes \mid \phi_{i}\right\rangle\left\langle\left.\phi_{i}\right|_{B} .\right.
$$

This $\rho_{A B^{2}}$ has some special properties: (i) $\rho_{A B^{2}}$ can be seen as an extension of the original state $\rho_{A B}$ to three parties such that

$$
\operatorname{Tr}_{B}\left[\rho_{A B^{2}}\right]=\rho_{A B}
$$

(ii) $\rho_{A B^{2}}$ is invariant (symmetric) under permutations between the second and third parties, i.e.,

$$
\rho_{A B^{2}}=P_{B \mid B} \rho_{A B^{2}} P_{B \mid B},
$$

where $P_{B \mid B}$ is the swap operator between the two copies of system $B$. (iii) $\rho_{A B^{2}}$ has a PPT for any partial transposition.

Therefore, we call $\rho_{A B^{2}}$ a PPT symmetric extension of $\rho_{A B}$ to two copies of system $B$. The same idea can be easily extended to the general $n$-copy case: a PPT symmetric extension of $\rho_{A B}$ to $n$ copies of system $B$ can be written as

$$
\rho_{A B^{n}}=\sum_{i} p_{i}\left|\psi_{i}\right\rangle\left\langle\left.\psi_{i}\right|_{A} \otimes \mid \phi_{i}\right\rangle\left\langle\left.\phi_{i}\right|_{B} \otimes \mid \phi_{i}\right\rangle\left\langle\left.\phi_{i}\right|_{B} ^{\otimes n-1} .\right.
$$

As we have just demonstrated, we can construct PPT symmetric extensions for any separable state, and thus the existence of such extensions is a necessary condition for separable states, namely a separability criterion. We call this property of separable states extendibility. As the extendibility condition for each $n$ is an individual separability criterion, we now have a countable-infinite family of separability criteria. Furthermore, this family of separability criteria has a hierarchical structure. It is easy to see if some bipartite state $\rho_{A B}$ has a PPT symmetric extension to $n$ copies of system B, denoted by $\rho_{A B^{n}}$, then automatically has a PPT symmetric extension to $n-1$ copies, which is just $\operatorname{Tr}_{B}\left[\rho_{A B^{n}}\right]$. This hierarchy of separability criteria is
known as the Doherty-Parrilo-Spedalieri (DPS) hierarchy.
Let us now formally define the extendibility as follows.

Definition 8. (Extendibility) A bipartite quantum state $\rho_{A B}$ is $n$-extendible if there exists a PPT symmetric extension to $n$ copies of system $B$, i.e., $\rho_{A B^{n}}$ such that

$$
\operatorname{Tr}_{B^{n-1}}\left[\rho_{A B^{n}}\right]=\rho_{A B}, \quad\left(\mathcal{I}_{A} \otimes \mathcal{U}_{B^{n}}^{\pi}\right)\left(\rho_{A B^{n}}\right)=\rho_{A B^{n}} \quad \forall \pi \in \mathcal{S}\left(B^{n}\right),
$$

and $\rho_{A B^{n}}$ has a PPT for any partial transposition, where $\mathcal{S}\left(B^{n}\right)$ is the symmetric group over $B^{n}, \mathcal{U}_{B^{n}}^{\pi}(\cdot)=U_{B^{n}}^{\pi}(\cdot)\left(U_{B^{n}}^{\pi}\right)^{\dagger}$ is the adjoint representation of the group, and $U_{B^{n}}^{\pi}$ is a unitary permutation operator acting on $B^{n}$.

As we have discussed above, these extendible states form good approximations for separable quantum states. This is related to the concept known as the monogamy of entanglement [96]; if two quantum system $A$ and $B$ are strongly entangled, they cannot be entangled with another quantum system. As the correlation between $A$ and $B$ in an $n$-extendible state $\rho_{A B}$ can be shared between $A$ and $n$ different $B$ systems (due to the first condition in Definition 8), the state $\rho_{A B}$ cannot be very entangled and should be close to the set of separable states. Furthermore, a quantum state is $n$-extendible for every $n \geq 2$ if and only if it is separable - this was shown in [97, 98] where the authors were interested in proving quantum version of de Finetti theorem, which is the main topic of Section 6.4. As a result, the DPS hierarchy provides a complete family of separability criteria. Since the set of $(n+1)$-extendible states is a subset of the set of $n$-extendible states due to the natural hierarchical structure (see Figure 6.2), these extendible states form a hierarchy of outer approximations for separable states, which converges to the set of separable states when $n$ goes to infinity.

The DPS hierarchy has another useful property that each test can be formulated as an SDP. For example, given a bipartite state $\rho_{A B}$, the test of two-extendibility of $\rho_{A B}$ can be


Figure 6.2: Extendible states. Extendible states form a converging hierarchy of outer approximations for the set of separable states. As $n$ becomes larger, the set of $n$-extendible states becomes tighter and closer to the set of separable states. It actually coincides with the separable set when $n$ goes to infinity.
implemented with the following optimisation problem: (with a slight abuse of the notation)

$$
\begin{array}{cl}
\min _{\sigma} & 0 \\
\text { subject to } & \sigma \equiv \sigma_{A B^{2}} \oplus \sigma_{A B^{2}}^{\mathrm{T}_{A}} \oplus \sigma_{A B^{2}}^{\mathrm{T}_{B}} \succeq 0 \\
& \operatorname{Tr}_{B}\left[\sigma_{A B^{2}}\right]=\rho_{A B} \\
& P_{B \mid B} \sigma_{A B^{2}} P_{B \mid B}=\sigma_{A B^{2}} .
\end{array}
$$

If such extension $\sigma_{A B^{2}}$ exists, the above problem outputs 0 , and if $\rho_{A B}$ is not extendible so that such extension does not exist, the problem is infeasible and outputs $\infty$. The above optimisation problem is an SDP as it optimises the linear objective function over the intersection of the cone of positive semidefinite matrices with the affine space satisfying the equality constraints. Thus, the DPS hierarchy is an instance of SDP relaxations.

It is interesting to remark that the original quantum separability problem is in principle a convex optimisation problem since the set of separable states on a finite-dimensional Hilbert space is convex; optimising a linear function over separable states is then a convex optimisation problem. The reason why it is hard to solve despite being a convex optimisation problem is that there is no simple characterisation of the set of separable states in terms of linear equations and inequalities. As in the DPS hierarchy, one needs an infinite number of constraints to characterise
the feasible set, the set of separable states. Thus, quantum separability problem is a good example showing that not all convex optimisation is efficiently solvable - one still needs an effective characterisation of the convex feasible set.

We can extend the same idea to the multipartite case.

Definition 9. (Multipartite extendibility) A k-partite state $\rho_{A_{1} A_{2} \cdots A_{k}} \in \mathcal{H}_{A_{1} A_{2} \cdots A_{k}}$ is $\left(n_{1}, \cdots, n_{k-1}\right)$ -


$$
\begin{aligned}
\operatorname{Tr}_{A_{2}^{n_{1}-1} A_{3}^{n_{2}-1} \ldots A_{k}^{n_{k-1}-1}}\left[\rho_{\left.A_{1} A_{2}^{n_{1} \ldots A_{k}^{n_{k-1}}}\right]}=\rho_{A_{1} A_{2} \cdots A_{k}}\right. \\
\left(\mathcal{I}_{A_{1}} \otimes \mathcal{U}_{A_{2}^{n_{1}}}^{\pi_{1}} \otimes \cdots \otimes \mathcal{U}_{A_{k}^{\pi_{k-1}}}^{\pi_{k-1}}\right)\left(\rho_{\left.A_{1} A_{2}^{n_{1} \ldots A_{k}^{n_{k-1}}}\right)}=\rho_{A_{1} A_{2}^{n_{1} \ldots A_{k}^{n_{k-1}}}}\right.
\end{aligned}
$$

for any permutations $\pi_{1} \in \mathcal{S}\left(A_{2}^{n_{1}}\right), \cdots, \pi_{k-1} \in \mathcal{S}\left(A_{k}^{n_{k-1}}\right)$, and $\rho_{A_{1} A_{2}^{n_{1}} \ldots A_{k}^{n_{k-1}}}$ has a PPT for any partial transpose.

It is shown in [99] that it is sufficient to consider the case when $n \equiv n_{1}=\ldots=n_{k}$ for the convergence of multipartite extendible states to the set of separable states. As in the bipartite case, the set of $(n, \ldots, n)$-extendible states converges to the set of $k$-partite separable states when $n \rightarrow \infty$.

In this section, we have constructed a converging hierarchy of SDP relaxations for quantum separability, which can be effectively implemented. We can see that the relaxations are formed in a way that each level is as powerful as the previous one, and the first level corresponding to $n=1$ is equivalent to the PPT criteria. The DPS hierarchy has a more fruitful structure that the dual problem allows us to build an explicit entanglement witness to detect the entanglement present in the given state. We refer to the original paper [95] of the DPS hierarchy for more details.

### 6.3.2 The Navascués-Pironio-Acín (NPA) hierarchy

In this section, we introduce another well-known hierarchy of SDP relaxations widely used in the quantum information field. It was introduced in [100-102] and is called the Navascués-Pironio-Acín (NPA) hierarchy. The NPA hierarchy provides effective relaxations for optimisation
problems of the form

$$
\begin{align*}
\min _{\boldsymbol{X}, \psi, \mathcal{H}} & \langle\psi| P(\boldsymbol{X})|\psi\rangle  \tag{6.7}\\
\text { subject to } & Q_{i}(\boldsymbol{X}) \succeq 0 \quad i \in(1, \ldots, m),
\end{align*}
$$

where $\boldsymbol{X}=\left(X_{1}, \ldots, X_{N}\right)$ is the set of non-commuting bounded operators on the Hilbert space $\mathcal{H}$, $|\psi\rangle$ is a normalised vector in $\mathcal{H}$, and $P(\boldsymbol{X})$ and $\left\{Q_{i}(\boldsymbol{X})\right\}_{i}$ are polynomials in the variables $\boldsymbol{X} .{ }^{5}$ Note that the Hilbert space $\mathcal{H}$ is also subject to optimisation, so the dimension of the Hilbert space is not fixed. The optimisation of the form in Eq. (6.7) is very general and has numerous applications beyond the quantum theory.

The NPA hierarchy has an important application in quantum information science, which is related to the problem of characterising quantum correlations in the context of non-locality. Let us consider the well-known Bell scenario. There are two distant parties, Alice and Bob, who perform measurements on a shared physical system. Alice performs a measurement $q_{1} \in Q_{1}$ on her part of the system and obtains an output $a_{1} \in A_{1}$. Similarly, Bob performs a measurement $q_{2} \in Q_{2}$ and obtains an output $a_{2} \in A_{2}$. Then, the joint conditional probability $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ observed by Alice and Bob represents the correlation shared by them via the physical system. Quantum correlations are the correlations which can be realised by quantum states and measurements and have the form

$$
\begin{equation*}
p_{Q}\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right):=\left\langle\left.\psi\right|_{A B}\left(M_{A}\left(a_{1} \mid q_{1}\right) \otimes N_{B}\left(a_{2} \mid q_{2}\right)\right) \mid \psi\right\rangle_{A B}, \tag{6.8}
\end{equation*}
$$

where $|\psi\rangle$ is the shared quantum state, and $\left\{M\left(a_{1} \mid q_{1}\right)\right\}_{a_{1}}$ and $\left\{N\left(a_{2} \mid q_{2}\right)\right\}_{a_{2}}$ for $q_{1}, q_{2}$ are the POVMs performed by Alice and Bob respectively. Then, a typical problem in the non-locality theory is to derive a quantum Bell inequality: for a given Bell function $B$, which is a linear function of $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ taking the form

$$
B(p)=\sum_{a_{1}, a_{2}, q_{1}, q_{2}} B_{a_{1}, a_{2}}^{q_{1}, q_{2}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)
$$

[^30]with some constants $B_{a_{1}, a_{2}}^{q_{1}, q_{2}} \in \mathbb{R}$, we want to find a quantum bound $B_{Q}$ such that $B\left(p_{Q}\right) \leq B_{Q}$ for all quantum correlations $p_{Q}\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ of the form in Eq. (6.8). A conventional Bell inequality with Bell function $B$ can be then defined as $B\left(p_{C}\right) \leq B_{C}$ with classical bound $B_{C}$ satisfied by all classical correlations $p_{C}\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$. We study a more general form of Bell inequalities later in Section 7.1.

Deriving such quantum Bell inequality involves finding the quantum bound $B_{Q}$ for a given Bell function $B$, which is equivalent to maximising the Bell function $B$ over all quantum correlations of the form in Eq. (6.8):

$$
\begin{equation*}
B_{Q}:=\max _{\psi, M, N, \mathcal{H}} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} B_{a_{1}, a_{2}}^{q_{1}, q_{2}}\left\langle\left.\psi\right|_{A B}\left(M_{A}\left(a_{1} \mid q_{1}\right) \otimes N_{B}\left(a_{2} \mid q_{2}\right)\right) \mid \psi\right\rangle_{A B} . \tag{6.9}
\end{equation*}
$$

Note that the Hilbert space $\mathcal{H}$ is again subject to optimisation. We can draw a connection between the optimisation problems in Eq. (6.9) and Eq. (6.7). If we define new notations $\left\{E_{A B}\left(a_{1}, q_{1}\right)\right\}_{a_{1}, q_{1}} \equiv\left\{M_{A}\left(a_{1} \mid q_{1}\right) \otimes \mathbb{1}_{B}\right\}_{a_{1}, q_{1}}$ and $\left\{E_{A B}\left(a_{2}, q_{2}\right)\right\}_{a_{2}, q_{2}} \equiv\left\{\mathbb{1}_{A} \otimes N_{B}\left(a_{2} \mid q_{2}\right)\right\}_{a_{2}, q_{2}}$, then we can rewrite the optimisation problem in Eq. (6.9) as

$$
B_{Q}=\max _{\psi, E, \mathcal{H}} \quad\langle\psi|\left(\sum_{a_{1}, a_{2}, q_{1}, q_{2}} B_{a_{1}, a_{2}}^{q_{1}, q_{2}} E\left(a_{1}, q_{1}\right) E\left(a_{2}, q_{2}\right)\right)|\psi\rangle
$$

where we omitted the system subscriptions for simplicity. Then, the expression in the brackets is a polynomial of the variables $\boldsymbol{E} \equiv\left\{E\left(a_{1}, q_{1}\right)\right\} \cup\left\{E\left(a_{2}, q_{2}\right)\right\}$, which are bounded operators on $\mathcal{H}$, and thus the problem is exactly in the form of Eq. (6.7) considered by the NPA hierarchy.

Therefore, we have shown that the problem of calculating the quantum bound for a quantum Bell inequality can take advantage of the NPA hierarchy. Although the set of quantum correlations $p_{Q}\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ in Eq. (6.8) is convex, there is no method to characterise the set with finite number of constraints as in the case of quantum separability studied in the last section. The NPA hierarchy provides practical relaxations for this problem and is also complete in the sense that the set of correlations characterised by the $n$-th level relaxation of the NPA hierarchy converges to the set of quantum correlations when $n \rightarrow \infty$.

## The NPA hierarchy

To formally introduce what the NPA hierarchy is, we consider the same Bell scenario described above: quantum correlations of the form

$$
p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=\operatorname{Tr}\left[E\left(a_{1}, q_{1}\right) E\left(a_{2}, q_{2}\right) \rho_{A B}\right]
$$

for a quantum state $\rho_{A B}$ and local measurements $\left\{E\left(a_{1}, q_{1}\right)\right\}_{a_{1}, q_{1}}$ and $\left\{E\left(a_{2}, q_{2}\right)\right\}_{a_{2}, q_{2}}$ performed by Alice and Bob respectively. In the NPA hierarchy, we can always assume without loss of generality that the measurements in $\left\{E\left(a_{i}, q_{i}\right)\right\}_{a_{i}, q_{i}}$ are composed by orthogonal projectors, and $\rho_{A B}$ is a pure quantum state, since the dimension of the system is unbounded. ${ }^{6}$ We assume the following properties are satisfied by the measurement operators in $\left\{E\left(a_{i}, q_{i}\right)\right\}_{a_{i}, q_{i}}$ :

1. hermiticity: $E\left(a_{i}, q_{i}\right)^{\dagger}=E\left(a_{i}, q_{i}\right)$,
2. orthogonality: $E\left(a_{i}, q_{i}\right) E\left(\bar{a}_{i}, \bar{q}_{i}\right)=\delta_{a_{i}, \bar{a}_{i}} E\left(a_{i}, q_{i}\right)$ if $q_{i}=\bar{q}_{i}$,
3. completeness: $\sum_{a_{i}} E\left(a_{i}, q_{i}\right)=\mathbb{1}_{A B}$ for all $q_{i}$,
4. commutativity: $\left[E\left(a_{1}, q_{1}\right), E\left(a_{2}, q_{2}\right)\right]=0$ for all $a_{1}, a_{2}, q_{1}, q_{2}$.

The orthogonality describes that the measurements are composed by orthogonal projectors. The commutativity came from the condition that $E\left(a_{1}, q_{1}\right)$ acts on Alice's side, and $E\left(a_{2}, q_{2}\right)$ acts on Bob's side although the two conditions are not equivalent ${ }^{7}$. We define the sets of indices $\boldsymbol{\alpha}:=\left\{\left(a_{1}, q_{1}\right)\right\}_{a_{1}, q_{1}}$ and $\boldsymbol{\beta}:=\left\{\left(a_{2}, q_{2}\right)\right\}_{a_{2}, q_{2}}$. Then, we can simply denote the correlation as $P_{\alpha \beta}:=p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ where $\alpha=\left(a_{1}, q_{1}\right) \in \boldsymbol{\alpha}$ and $\beta=\left(a_{2}, q_{2}\right) \in \boldsymbol{\beta}$. The set $\boldsymbol{\alpha}$ has $m_{1}=\left|A_{1}\right|\left|Q_{1}\right|$ elements, and the set $\boldsymbol{\beta}$ has $m_{2}=\left|A_{2}\right|\left|Q_{2}\right|$ elements.

Let us assume that a correlation $P_{\alpha \beta}$ is a quantum correlation which has a quantum state $|\psi\rangle$ and measurements $\left\{E\left(a_{i}, q_{i}\right)\right\}_{a_{i}, q_{i}}$. By taking product of $E\left(a_{i}, q_{i}\right)$ or linear combinations of such products, we can construct a set of $n$ operators, $\boldsymbol{S}=\left\{S_{1}, \ldots, S_{n}\right\}$; in other words, we build

[^31]a set $\boldsymbol{S}$ of polynomials in the variables $\left\{E\left(a_{i}, q_{i}\right)\right\}_{a_{i}, q_{i}}$. For a set $\boldsymbol{S}$, we can construct an $n \times n$ matrix $\Gamma$ of the form
\[

$$
\begin{equation*}
\Gamma_{i j}=\operatorname{Tr}\left[S_{i}^{\dagger} S_{j} \rho_{A B}\right] . \tag{6.10}
\end{equation*}
$$

\]

From the construction, $\Gamma$ is a Hermitian matrix which satisfies that

$$
\begin{equation*}
\text { if } \sum_{i, j} c_{i j} S_{i}^{\dagger} S_{j}=0 \quad \text { then } \quad \sum_{i, j} c_{i j} \Gamma_{i j}=0 \tag{6.11}
\end{equation*}
$$

as well as that

$$
\begin{equation*}
\text { if } \sum_{i, j} c_{i j} S_{i}^{\dagger} S_{j}=\sum_{\alpha, \beta} d_{\alpha \beta} E(\alpha) E(\beta) \quad \text { then } \quad \sum_{i, j} c_{i j} \Gamma_{i j}=\sum_{\alpha, \beta} d_{\alpha \beta} P_{\alpha \beta}, \tag{6.12}
\end{equation*}
$$

where $\alpha \in \boldsymbol{\alpha}$ and $\beta \in \boldsymbol{\beta}$. The coefficients $c_{i j}$ and $d_{\alpha \beta}$ are determined by the set $\boldsymbol{S}$ - they are related to any linear constraints on the set $\boldsymbol{S}$. Moreover, we have

$$
\begin{equation*}
\Gamma \succeq 0 \tag{6.13}
\end{equation*}
$$

As a result, we get the following necessary conditions for quantum correlations.

Lemma 18. [100, 101] For a given set $\boldsymbol{S}=\left\{S_{1}, \ldots, S_{n}\right\}$ constructed by polynomials in $\left\{E\left(a_{i}, q_{i}\right)\right\}_{a_{i}, q_{i}}$ satisfying the conditions 1-4, a necessary condition for a correlation $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ to be quantum is that there exists a Hermitian $n \times n$ matrix $\Gamma$ which satisfies the conditions Eqs. (6.11)-(6.13).

The simplest example is when we choose $\boldsymbol{S}$ as $\left\{E\left(a_{1}, q_{1}\right)\right\}_{a_{1}, q_{1}} \cup\left\{E\left(a_{2}, q_{2}\right)\right\}_{a_{2}, q_{2}} \equiv\left\{E\left(a_{i}, q_{i}\right)\right\}_{a_{i}, q_{i}}$. For any given correlation $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$, the corresponding $\Gamma_{\boldsymbol{S}}$ for $\boldsymbol{S}=\left\{E\left(a_{i}, q_{i}\right)\right\}_{a_{i}, q_{i}}$ constructed via Eq. (6.10) takes the form

$$
\Gamma_{\boldsymbol{S}}=\left(\begin{array}{cc}
Q & C  \tag{6.14}\\
C^{\mathrm{T}} & R
\end{array}\right)
$$

where the $m_{1} \times m_{2}$ sub-matrix $C$ has the form

$$
C_{\alpha \beta}=P_{\alpha \beta}=\operatorname{Tr}\left[E(\alpha) E(\beta) \rho_{A B}\right] \quad \forall \alpha \in \boldsymbol{\alpha}, \beta \in \boldsymbol{\beta}
$$

and the $m_{1} \times m_{1}$ sub-matrix Q and $m_{2} \times m_{2}$ sub-matrix $R$ satisfy

$$
\begin{align*}
& Q_{\alpha \alpha}=p\left(a_{1} \mid q_{1}\right)=\operatorname{tr}\left[E\left(a_{1}, q_{1}\right) \rho_{A B}\right], \quad \alpha=\left(a_{1}, q_{1}\right) \in \boldsymbol{\alpha} \\
& Q_{\alpha \alpha^{\prime}}=0, \quad \alpha=\left(a_{1}, q_{1}\right), \alpha^{\prime}=\left(a_{1}^{\prime}, q_{1}^{\prime}\right) \in \boldsymbol{\alpha} \quad \text { when } q_{1}=q_{1}^{\prime}  \tag{6.15}\\
& R_{\beta \beta}=p\left(a_{2} \mid q_{2}\right)=\operatorname{tr}\left[E\left(a_{2}, q_{2}\right) \rho_{A B}\right], \quad \beta=\left(a_{2}, q_{2}\right) \in \boldsymbol{\beta} \\
& R_{\beta \beta^{\prime}}=0, \quad \beta=\left(a_{2}, q_{2}\right), \beta^{\prime}=\left(a_{2}^{\prime}, q_{2}^{\prime}\right) \in \boldsymbol{\beta} \quad \text { when } q_{2}=q_{2}^{\prime},
\end{align*}
$$

where $p\left(a_{1} \mid q_{1}\right)=\sum_{a_{2}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ and $p\left(a_{2} \mid q_{2}\right)=\sum_{a_{1}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ are the marginal probabilities. ${ }^{8}$ Note that we are suppressing the conditioning arguments $q_{2}$ and $q_{1}$ in the marginals since we assume the no-signalling condition:

$$
\begin{array}{ll}
p\left(a_{1} \mid q_{1}, q_{2}\right)=p\left(a_{1} \mid q_{1}, q_{2}^{\prime}\right) & \forall a_{1}, q_{1}, q_{2}, q_{2}^{\prime}  \tag{6.16}\\
p\left(a_{2} \mid q_{1}, q_{2}\right)=p\left(a_{2} \mid q_{1}^{\prime}, q_{2}\right) & \forall a_{2}, q_{1}, q_{1}^{\prime}, q_{2} .
\end{array}
$$

With this construction, we can specify all entries of $\Gamma_{S}$ in terms of the given correlation $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ except the entries of $Q_{\alpha \alpha^{\prime}}$ when $q_{1} \neq q_{1}^{\prime}$ and $R_{\beta \beta^{\prime}}$ when $q_{2} \neq q_{2}^{\prime}$. The matrix $\Gamma_{S}$ automatically satisfies the conditions in Eq. (6.11) and (6.12) due to Eq. (6.15). The only remaining check is whether we can find a matrix of the form $\Gamma_{S}$ in Eq. (6.14) which is positive semidefinite by giving appropriate values to the undecided entries of $\Gamma_{S}$. In other words, for a given correlation $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$, we need to decide whether we can find $\Gamma_{S}$ of the form in Eq. (6.14) such that

$$
\Gamma_{S} \succeq 0
$$

As different set $\boldsymbol{S}^{\prime}$ provide a different necessary condition for quantum correlations, we can

[^32]build a hierarchy of necessary conditions by using a sequence of independent sets $\left\{\boldsymbol{S}_{k}\right\}_{k}$ such that $\boldsymbol{S}_{k} \subset \boldsymbol{S}_{k+1}$ for all $k$. Conventionally, in the NPA hierarchy, the first level is chosen as $\boldsymbol{S}_{1}=\boldsymbol{S}_{0} \cup\left\{E\left(a_{i}, q_{i}\right)\right\}$ where $\boldsymbol{S}_{0}=\{\mathbb{1}\}$, the second level is $\boldsymbol{S}_{2}=\boldsymbol{S}_{0} \cup \boldsymbol{S}_{1} \cup\left\{E\left(a_{i}, q_{i}\right) E\left(a_{j}, q_{j}\right)\right\},{ }^{9}$ the third level is $\boldsymbol{S}_{3}=\boldsymbol{S}_{0} \cup \boldsymbol{S}_{1} \cup \boldsymbol{S}_{2} \cup\left\{E\left(a_{i}, q_{i}\right) E\left(a_{j}, q_{j}\right) E\left(a_{k}, q_{k}\right)\right\}$, and so on. We denote the $k$-th level NPA matrix by $\Gamma_{k}$;
\[

$$
\begin{equation*}
\left(\Gamma_{k}\right)_{i j}=\operatorname{Tr}\left[S_{i}^{\dagger} S_{j} \rho_{A B}\right] \tag{6.17}
\end{equation*}
$$

\]

for the set $\boldsymbol{S}_{k}$. As we mentioned earlier, the completeness of the NPA hierarchy was shown in [101]. That is, a correlation which satisfies the $k$-th level NPA condition for all $k \geq 1$ is a quantum correlation.

### 6.4 Quantum de Finetti theorems

In this last section, we give a short introduction of quantum de Finetti theorems which play an important role in the next chapter. The original de Finetti theorem in classical probability theory concerns the close relation between an infinite permutation-invariant sequence of random variables and a mixture of independent and identically distributed (i.i.d.) sequences. There have been developed multiple variants of this theorem since then, and the version in which we are interested is the de Finetti theorem for finite symmetric quantum states. This type of quantum de Finetti theorems says that if a global quantum state of a group of quantum systems is invariant under permutations of the quantum systems, then the reduced state for some sub-group must be close to a convex combination of identical product states. A simple version of finite quantum de Finetti theorems can be stated as follows.

Theorem 19 (quantum de Finetti theorem [104]). Let $A_{1}, \ldots, A_{n}$ be identical quantum registers, each having a d-dimensional associated Hilbert space, and let $\rho_{A_{1} \ldots A_{n}}$ be a quantum state on the Hilbert space $\mathcal{H}_{A_{1}} \otimes \ldots \otimes \mathcal{H}_{A_{n}}$ which is invariant under any permutations between $A_{1}, \ldots, A_{n}$. For any choice of $k \in\{1, \ldots, n\}$, there exists a finite set $\Gamma$, a probability distribution $\left\{p_{i}\right\}_{i}$, and a

[^33]collection of density operators $\left\{\sigma_{i}\right\}_{i}$ such that
$$
\left\|\rho_{A_{1} \ldots A_{k}}-\sum_{i \in \Gamma} p_{i} \sigma_{i}^{\otimes k}\right\|_{1}<\frac{4 d^{2} k}{n}
$$

The theorem states that for any reduced state on $k \leq n$ subsystems of the permutationinvariant quantum state $\rho_{A_{1} \ldots A_{n}}$, there exists a separable state, more precisely a convex mixtures of identical product states, which is arbitrarily close to the given reduced state. Moreover, it provides a quantitative bound on the trace-norm distance between the quantum state and the separable state in terms of the dimension $d$ and the numbers $n$ and $k$.

One can make a connection from Theorem 19 to the extendible states introduced in Section 6.3.1. Recall that an $n$-extendible state $\rho_{A B}$ is a state which has a PPT symmetric extension to $n$ copies of system $B$, i.e., $\rho_{A B_{1} \cdots B_{n}}$ such that $\operatorname{Tr}_{B^{n-1}}\left[\rho_{A B_{1} \cdots B_{n}}\right]=\rho_{A B}$, and $\rho_{A B_{1} \cdots B_{n}}$ is invariant under permutations of $B_{1} \cdots B_{n}$. If we discard the side quantum system $A$ and consider the reduced state $\rho_{B_{1} \cdots B_{n}}$ of the PPT symmetric extension, it satisfies the condition of Theorem 19, and we can apply the theorem to $\rho_{B_{1} \cdots B_{n}}$. Then, Theorem 19 tells us that any reduced state of $\rho_{B_{1} \cdots B_{n}}$ is very close to the set of separable states. Although this does not automatically imply that the original state $\rho_{A B_{1} \cdots B_{n}}$ with side quantum system $A$ has the same property, we can deduce that it would also have a similar property as $\rho_{A B_{1} \cdots B_{n}}$ has a permutation-invariant property on $B_{1} \cdots B_{n}$. Indeed, this version of quantum de Finetti theorem with side quantum system has been studied in the literature [105], and it was shown that this deduction is correct; the extendible state $\rho_{A B}$, which is a reduced state of a permutation-invariant extension $\rho_{A B_{1} \cdots B_{n}}$, is close to the set of separable states with respect to the bipartition $A \mid B$, and we can provide a quantitative bound on the closedness via quantum de Finetti theorems. In the next chapter, we make use of the multipartite version of this case; multipartite quantum de Finetti theorems with quantum side information.

## Chapter 7

## Characterising quantum correlations of fixed dimension

Thanks to the celebrated discovery of John Bell [106], it is now well-known that quantum correlations can be used as a resource to overcome locality constraints, which are often expressed in the form of Bell inequalities. This was one of the earliest examples of quantum advantages over classical ones and led to the development of a vast number of quantum information processing tasks that make use of quantum correlations. A prominent example is given by the field of deviceindependent quantum information processing, where the violation of a certain locality constraint is used to certificate the 'quantumness' of the underlying state describing the system even when we have no a priori knowledge about the device used [12]. In a semi-device-independent task, there is often an additional constraint on the dimensionality of the device. Since the maximum amount of violation of a locality constraint could depend on the dimensionality, it has been pointed out that this feature could be exploited as a dimension witness of the underlying system (e.g., see [107]) if strong tools to characterise quantum correlations of fixed dimension are available.

A common way to formulate locality constraints is a non-local game, which could be seen as a generalisation of Bell scenario (see Section 7.1). In non-local games, the strength of a certain set of correlations, such as classical or quantum correlations, is measured by the maximum winning probability that can be achieved by the given set of correlations. For example, the classical value
$\omega_{C}(G)$ of a given non-local game $G$ is the maximum winning probability achievable by classical correlations shared between the distant players, and the quantum value $\omega_{Q}(G)$ is the maximum winning probability achievable by arbitrary quantum state shared between the players. The dimension-bounded quantum value $\omega_{Q(T)}(G)$, which is our main topic in this chapter, is then the maximum winning probability achievable by using quantum states of dimension $T \times T$ shared between the players. It holds that $\omega_{Q(1)}=\omega_{C}(G)^{1}$ and $\omega_{Q}(G)=\sup _{T \geq 1} \omega_{Q(T)}(G)$. For a given game $G$, if $\omega_{C}(G)<\omega_{Q(T)}(G) \leq \omega_{Q}(G)$, it implies the quantum violation of the locality condition associated with $G$ - using a quantum correlation allows us to exceed the classical upper bound on the winning probability of $G$.

Given the description of a non-local game $G$, computing these values is often computationally hard, but nevertheless several methods have been developed. For the case of unbounded dimension, the NPA hierarchy [100-102], studied in Section 6.3.2, provides a sequence of asymptotically converging SDP upper bounds on $\omega_{Q}(G)$, which has been widely used within the quantum information field. However, the NPA hierarchy may give only loose bounds for problems with dimensional constraint on quantum correlations. There have been a few results trying to target this specific problem, the characterisation of quantum correlations with fixed dimension: In [108], the authors constructed hierarchies of SDP upper bounds on $\omega_{Q(T)}(G)$ by exploiting a connection to the standard quantum separability problem; also, in [109, 110], the authors employed the moment-matrix technique similar to the NPA hierarchy to derive SDP relaxations with improved numerical performance compared to [108]. However, for these works the asymptotic convergence speed is either not analytically quantified or at best exponential in the number of possible questions $Q$ and the number of possible answers $A$ of the game.

In this chapter, we show how to construct algorithms for computing $\omega_{Q(T)}(G)$ whose running time has a significantly improved dependence on $A$ and $Q$. More specifically, we construct a new hierarchy of SDP upper bounds on $\omega_{Q(T)}(G)$ and derive analytical bounds on the convergence speed of the hierarchy. This result leads to bounds on the computational complexity of the problem of calculating $\omega_{Q(T)}(G)$ as a function of the size of the game $G: A, Q$ and $T$. For the case of two-player free games, i.e., games where the questions of the players are chosen

[^34]independently of each other, we provide a semidefinite program of size
$$
\exp \left(\mathcal{O}\left(\frac{T^{12}}{\epsilon^{2}} \log (A T)(\log (Q)+\log (A T))\right)\right)
$$
for computing approximation of $\omega_{Q(T)}(G)$ for an additive error $\epsilon$. Note that the dependence is quasi-polynomial in $A$ and polynomial in $Q$, thereby improving on previously mentioned approximation algorithms for which worst-case run-time guarantees are at best exponential in $Q$ and $A$ [108-110]. We also give a program for general games, which is still quasi-polynomial in $A$ but exponential in $Q$. The main idea for deriving these algorithms is to make a connection of the problem $\omega_{Q(T)}(G)$ to a variant of the quantum separability problem where the states involved are subject to additional linear constraints. Then, we obtain improved multipartite quantum de Finetti theorems with linear constraints which allow us to quantify the convergence speed of our hierarchy of SDP relaxations.

To help the readers, here we provide a short overview of this chapter emphasising each important step.

- Connection with quantum separability. Our first step is to relate the problem of calculating $\omega_{Q(T)}(G)$ for a given non-local game $G$ to an instance of tripartite quantum separability problems with linear constraints. This is done in Lemma 20.
- Hierarchy of semidefinite programming relaxations We then construct our new hierarchy of semidefinite programming relaxations for $\omega_{Q(T)}(G)$ by simply employing the DPS hierarchy for tripartite quantum separability problem but with additional linear constraints. The resulting hierarchy is written in Eq. (7.7)-(7.12).
- Multipartite quantum de Finetti theorem with linear constraints. The standard quantum de Finetti theorem can provide a quantitative bound on the accuracy of the DPS hierarchy, but this is not possible in our setting due to the existence of additional linear constraints. Thus, we derive an adapted one, namely multipartite quantum de Finetti theorems with linear constraints, in Theorem 23 to prove the convergence rate of our new hierarchy.


Figure 7.1: Non-local game. The referee gives Alice and Bob questions $q_{1} \in Q_{1}$ and $q_{2} \in Q_{2}$ according to the question probability distribution $\pi\left(q_{1}, q_{2}\right)$, and then Alice and Bob give answers $a_{1} \in A_{1}$ and $a_{2} \in A_{2}$ back to the referee depending on the questions they received. The referee decides whether Alice and Bob win or lose according to the rule function $V: A_{1} \times A_{2} \times Q_{1} \times Q_{2} \rightarrow$ $\{0,1\}$, where 0 denotes losing the game, and 1 denotes winning the game. Alice and Bob cannot communicate with each other during the game, but they can agree on a strategy beforehand. We assume that $\left|Q_{1}\right|=\left|Q_{2}\right|=Q$ and $\left|A_{1}\right|=\left|A_{2}\right|=A$.

- Convergence rate of the hierarchy. Using the multipartite quantum de Finetti theorem with linear constraints, we can derive an upper bound on the asymptotic convergence rate of our hierarchy. This is described in Theorem 28.

This chapter is written based on the paper [Hyejung H Jee et al., ICALP 2021, 198, 82 (2021)] and structured as follows. In Section 7.1, we introduce the notion of non-local games and define the maximum winning probabilities of such games for different sets of correlations. In Section 7.2, we derive the constraints for characterising fixed-dimensional quantum correlations and state a new hierarchy of SDP relaxations for $\omega_{Q(T)}(G)$ of two-player free games. In Section 7.3, we analytically prove the convergence rate of the derived hierarchy exploiting tripartite quantum de Finetti theorems with additional linear constraints which we also prove in the same section. Then, in Section 7.4, we discuss how to combine our derived constraints with the existing dimension-agnostic NPA hierarchy. We compare our results with previous work in the literature in Section 7.5 before concluding the chapter with some remarks in Section 7.6.

### 7.1 Non-local games

There exist various mathematical formulations of the correlations between distant parties, but in this chapter we use the notion of non-local games. Specifically, we consider two-player games where two distant parties are considered. In this formulation, we consider the correlation between two parties as a resource to win some specific games.

Let us consider two spatially separated players, called Alice and Bob, and a referee. When a non-local game starts, the referee chooses questions $q_{1}$ and $q_{2}$ from the question set $Q_{1} \times Q_{2}$ according to a given question probability distribution $\pi\left(q_{1}, q_{2}\right)$ and sends them to Alice and Bob, respectively. Then, Alice and Bob must provide answers $a_{1}$ and $a_{2}$ for their questions to the referee; see Figure 7.1. The correct answers are determined by a given rule function

$$
V: A_{1} \times A_{2} \times Q_{1} \times Q_{2} \rightarrow\{0,1\}
$$

where 0 means the answers are wrong, and 1 means the answers are correct. A specific twoplayer game $G$ can be represented by a question probability distribution $\pi\left(q_{1}, q_{2}\right)$ and a rule function $V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$; thus, we hereafter represent a two-player game $G$ by $(V, \pi)$. Alice and Bob cannot communicate with each other during the game, but they can agree on a strategy beforehand, which can be represented by a conditional probability distribution $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$. Depending on the resources the players can access to, this conditional probability distribution describes different kinds of correlations. When no resources are shared, each player's answer only depends on the question they receive, so that the distribution takes the form

$$
\begin{equation*}
p_{C}\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=e\left(a_{1} \mid q_{1}\right) d\left(a_{2} \mid q_{2}\right) \tag{7.1}
\end{equation*}
$$

where $e\left(a_{1} \mid q_{1}\right)$ and $d\left(a_{2} \mid q_{2}\right)$ are conditional probability distributions representing the strategy of Alice and Bob, respectively. When quantum resources are allowed, the distribution takes the more general form (recall from Eq. (6.8))

$$
p_{Q}\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=\operatorname{Tr}\left[\rho_{T \hat{T}}\left(E_{T}\left(a_{1} \mid q_{1}\right) \otimes D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right)\right],
$$

where $\rho_{T \hat{T}}$ is the (possibly entangled) quantum state shared between Alice and Bob, and for fixed questions $q_{1}, q_{2},\left\{E_{T}\left(a_{1} \mid q_{1}\right)\right\}_{a_{1}}$ and $\left\{D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right\}_{a_{2}}$ are POVMs performed by Alice and Bob, respectively. Then, the difference in performance between classical and quantum correlations can be quantified in terms of a non-zero gap between the maximum winning probabilities achieved using the two different resources. For a given game $G=(V, \pi)$, the classical value, namely the classical maximum winning probability, is given by

$$
\begin{equation*}
\omega_{C}(V, \pi):=\max _{e, d} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) e\left(a_{1} \mid q_{1}\right) d\left(a_{2} \mid q_{2}\right), \tag{7.2}
\end{equation*}
$$

where $\sum_{a_{1}} e\left(a_{1} \mid q_{1}\right)=1 \forall q_{1}$, and $\sum_{a_{2}} d\left(a_{2} \mid q_{2}\right)=1 \forall q_{2}$, while the quantum value is given by

$$
\begin{equation*}
\omega_{Q}(V, \pi):=\sup _{E, D, \rho, \mathcal{H}_{T \vec{T}}} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\rho_{T \hat{T}}\left(E_{T}\left(a_{1} \mid q_{1}\right) \otimes D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right)\right] \tag{7.3}
\end{equation*}
$$

where $\operatorname{Tr}\left[\rho_{T \hat{T}}\right]=1, \sum_{a_{1}} E_{T}\left(a_{1} \mid q_{1}\right)=\mathbb{1}_{T} \forall q_{1}$, and $\sum_{a_{2}} D_{\hat{T}}\left(a_{2} \mid q_{2}\right)=\mathbb{1}_{\hat{T}} \forall q_{2}$. Note that the supremum is taken over all possible POVMs and quantum states as well as the Hilbert space $\mathcal{H}_{T \hat{T}}$ on which they act on. Then, denoting the winning probability for a given game ( $V, \pi$ ) with some classical strategy $p_{C}$ as $p_{\text {win }}\left(V, \pi, p_{C}\right)$, the Bell inequality of the game $(V, \pi)$ can be expressed as

$$
p_{\operatorname{win}}\left(V, \pi, p_{C}\right) \leq \omega_{C}(V, \pi)
$$

for all classical correlations $p_{C}$ of the form in Eq. (7.1). The violation of such classical upper bound with quantum correlations are called the violation of the Bell inequality. For example, the CHSH inequality [111] can be derived from the CHSH game, which can be represented by the uniform distribution $\pi_{\text {uni }}\left(q_{1}, q_{2}\right)$ as $\pi\left(q_{1}, q_{2}\right)$ and $V_{\mathrm{CHSH}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$ as the rule function, where

$$
V_{\mathrm{CHSH}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)=\left\{\begin{array}{ll}
1 & \text { if } q_{1} \cdot q_{2}=a_{1} \oplus a_{2} \\
0 & \text { otherwise }
\end{array} .\right.
$$

It is known that $\omega_{C}\left(V_{\mathrm{CHSH}}, \pi_{\mathrm{uni}}\right)=0.75$, and thus the CHSH inequality can be expressed as

$$
p_{\mathrm{win}}\left(V_{\mathrm{CHSH}}, \pi_{\mathrm{uni}}, p_{C}\right) \leq \omega_{C}\left(V_{\mathrm{CHSH}}, \pi_{\mathrm{uni}}\right)=0.75
$$

for all classical correlations $p_{C}$. It is also found that $\omega_{Q}\left(V_{\mathrm{CHSH}}, \pi_{\mathrm{uni}}\right)=\cos ^{2}\left(\frac{\pi}{8}\right) \approx 0.85355$, and thus we can violate the CHSH inequality with a suitable quantum correlation.

We can also think of a situation when Alice and Bob are only allowed to share a fixeddimensional quantum state such as a qubit- or qutrit-state; in other words, when the Hilbert space $\mathcal{H}_{T \hat{T}}$ is fixed. In this case, the relevant quantity is the maximum winning probability for a given two-player game $(V, \pi)$ with quantum assistance of fixed dimension $T \times T$ defined as

$$
\begin{equation*}
\omega_{Q(T)}(V, \pi):=\max _{E, D, \rho} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\rho_{T \hat{T}}\left(E_{T}\left(a_{1} \mid q_{1}\right) \otimes D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right)\right] . \tag{7.4}
\end{equation*}
$$

In this chapter, we are interested in approximating $\omega_{Q(T)}$ for a given two-player game.

### 7.2 Semidefinite programming relaxations for $\omega_{Q(T)}(V, \pi)$

In this section, we show how to derive a hierarchy of semidefinite programming relaxations for approximating $\omega_{Q(T)}(G)$ for a given two-player game $G$. We restrict ourselves to two-player free games, where $\pi\left(q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$, in the main context of this thesis and address general games in Appendix C.1. This is because most non-local games considered in physics are indeed free games (such as the CHSH game) and restricting to free games allows us to obtain a better result in terms of the computational complexity compared to the general case.

### 7.2.1 Connection to quantum separability

In this section, as the first step to derive the SDP hierarchy, we connect the original formulation of $\omega_{Q(T)}$ for a two-player free game to the quantum separability problem studied in Section 6.3.1. In the following lemma, we show that computing $\omega_{Q(T)}$ in Eq. (7.4) for a given two-player free game $(V, \pi)$ can be rephrased as an instance of tripartite quantum separability problems subject to additional linear constraints.

Lemma 20. For a two-player free game with $V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$ and $\pi\left(q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$ assisted by $T \times T$-dimensional quantum correlation, we have

$$
\begin{align*}
\omega_{Q(T)}(V, \pi) & =T^{2} \cdot \max _{(E, D, \rho)} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}} \otimes \rho_{S \hat{S}}\right)\right] \\
\text { s.t. } & \rho_{S \hat{S}} \geq 0, \quad \operatorname{Tr}\left[\rho_{S \hat{S}}\right]=1 \\
& E_{A_{1} Q_{1} T}=\sum_{a_{1}, q_{1}} \pi_{1}\left(q_{1}\right)\left|a_{1} q_{1}\right\rangle\left\langle\left. a_{1} q_{1}\right|_{A_{1} Q_{1}} \otimes \frac{E_{T}\left(a_{1} \mid q_{1}\right)}{T} \geq 0\right. \\
& D_{A_{2} Q_{2} \hat{T}}=\sum_{a_{2}, q_{2}} \pi_{2}\left(q_{2}\right)\left|a_{2} q_{2}\right\rangle\left\langle\left. a_{2} q_{2}\right|_{A_{2} Q_{2}} \otimes \frac{D_{\hat{T}}\left(a_{2} \mid q_{2}\right)}{T} \geq 0\right.  \tag{7.5}\\
& \operatorname{Tr}_{A_{1}}\left[E_{A_{1} Q_{1} T}\right]=\sum_{q_{1}} \pi_{1}\left(q_{1}\right)\left|q_{1}\right\rangle\left\langle\left. q_{1}\right|_{Q_{1}} \otimes \frac{\mathbb{1}_{T}}{T}\right. \\
& \operatorname{Tr}_{A_{2}}\left[D_{A_{2} Q_{2} \hat{T}}\right]=\sum_{q_{2}} \pi_{2}\left(q_{2}\right)\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}} \otimes \frac{\mathbb{1}_{\hat{T}}}{T},\right.
\end{align*}
$$

where $\Phi_{T \hat{T} \mid S \hat{S}}=|\Phi\rangle\left\langle\left.\Phi\right|_{T \hat{T} \mid S \hat{S}} \text { is the (non-normalised) maximally-entangled state, i.e., } \mid \Phi\right\rangle_{T \hat{T} \mid S \hat{S}}=$ $\sum_{i}|i\rangle_{T \hat{T}}|i\rangle_{S \hat{S}}$, and $V_{A_{1} A_{2} Q_{1} Q_{2}}$ is a diagonal matrix whose entries are given by the rule function $V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$.

Note that the optimisation in Eq. (7.5) is now taken over all product states with respect to the tripartition $A_{1} Q_{1} T\left|A_{2} Q_{2} \hat{T}\right| S \hat{S}$, which satisfy the stated linear constraints. Since product states are extreme points in the set of separable states, we can equivalently think of the above as an optimisation over the convex hull of the feasible states, which are all product states satisfying the linear constraints - the optimisation over the convex hull gives the same solution as the one over product states. Thus, this new form of the optimisation problem can be regarded as an instance of tripartite quantum separability problems with additional linear constraints.

To prove Lemma 20, we first need a simple mathematical trick called the swap trick. We prove a slightly modified version here.

Lemma 21. Let $M_{A B}$ be an operator acting on $\mathcal{H}_{A} \otimes \mathcal{H}_{B}$, and $N_{A}$ be an operator on $\mathcal{H}_{A}$. Then, it holds that

$$
\operatorname{Tr}\left[\left(N_{A} \otimes \mathbb{1}_{B}\right) M_{A B}\right]=\operatorname{Tr}\left[\left(F_{\hat{A} \mid A} \otimes \mathbb{1}_{B}\right)\left(N_{\hat{A}} \otimes M_{A B}\right)\right]
$$

where $F_{\hat{A} \mid A}$ denotes the swap operator between systems $\hat{A}$ and $A$, i.e., $F_{\hat{A} \mid A}=\sum_{i, j}|i j\rangle\left\langle\left. j i\right|_{\hat{A} A}\right.$.

Proof. By inspection, we have that

$$
\begin{aligned}
& \operatorname{Tr}\left[\left(F_{\hat{A} \mid A} \otimes \mathbb{1}_{B}\right)\left(N_{\hat{A}} \otimes M_{A B}\right)\right] \\
= & \operatorname{Tr}\left[\left(F_{\hat{A} \mid A} \otimes \mathbb{1}_{B}\right)\left(\sum_{i, j} n_{i j}|i\rangle\left\langle\left. j\right|_{\hat{A}} \otimes \sum_{k, \ell, s, t} m_{(k \ell)(s t)} \mid k\right\rangle\left\langle\left.\ell\right|_{A} \otimes \mid s\right\rangle\left\langle\left. t\right|_{B}\right)\right]\right. \\
= & \operatorname{Tr}\left[\sum_{i, j, k, \ell, s, t} n_{i j} m_{(k \ell)(s t)}|k\rangle\left\langle\left. j\right|_{\hat{A}} \otimes \mid i\right\rangle\left\langle\left.\ell\right|_{A} \otimes \mid s\right\rangle\left\langle\left. t\right|_{B}\right]\right. \\
= & \sum_{i, j, s} n_{i j} m_{(j i)(s s)}=\operatorname{Tr}\left[\left(N_{A} \otimes \mathbb{1}_{B}\right) M_{A B}\right],
\end{aligned}
$$

where we used $N_{A}=\sum_{i, j} n_{i j}|i\rangle\left\langle\left. j\right|_{A}\right.$ and $\left.M_{A B}=\sum_{k, \ell, s, t} m_{(k \ell)(s t)} \mid k\right\rangle\left\langle\left.\ell\right|_{A} \otimes \mid s\right\rangle\left\langle\left. t\right|_{B}\right.$.

We are now ready to prove Lemma 20.

Proof of Lemma 20. Let us start from the expression for $\omega_{Q(T)}$ in Eq. (7.5). For free games, we have $\pi\left(q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$. By substituting the free-game condition and replacing all classical systems with classical (diagonal) quantum systems, we can have

$$
\begin{equation*}
\omega_{Q(T)}(V, \pi)=T^{2} \max _{E, D, \rho} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \rho_{T \hat{T}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}}\right)\right] \tag{7.6}
\end{equation*}
$$

s.t. $\rho_{T \hat{T}} \geq 0, \quad \operatorname{Tr}\left[\rho_{T \hat{T}}\right]=1$

$$
\begin{aligned}
& E_{A_{1} Q_{1} T}=\sum_{a_{1}, q_{1}} \pi_{1}\left(q_{1}\right)\left|a_{1} q_{1}\right\rangle\left\langle\left. a_{1} q_{1}\right|_{A_{1} Q_{1}} \otimes \frac{E_{T}\left(a_{1} \mid q_{1}\right)}{T} \geq 0\right. \\
& D_{A_{2} Q_{2} \hat{T}}=\sum_{a_{2}, q_{2}} \pi_{2}\left(q_{2}\right)\left|a_{2} q_{2}\right\rangle\left\langle\left. a_{2} q_{2}\right|_{A_{2} Q_{2}} \otimes \frac{D_{\hat{T}}\left(a_{2} \mid q_{2}\right)}{T} \geq 0\right. \\
& \operatorname{Tr}_{A_{1}}\left[E_{A_{1} Q_{1} T}\right]=\sum_{q_{1}} \pi_{1}\left(q_{1}\right)\left|q_{1}\right\rangle\left\langle\left. q_{1}\right|_{Q_{1}} \otimes \frac{\mathbb{1}_{T}}{T}\right. \\
& \operatorname{Tr}_{A_{2}}\left[D_{A_{2} Q_{2} \hat{T}}\right]=\sum_{q_{2}} \pi_{2}\left(q_{2}\right)\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}} \otimes \frac{\mathbb{1}_{\hat{T}}}{T},\right.
\end{aligned}
$$

where we define $V_{A_{1} A_{2} Q_{1} Q_{2}}:=\sum_{a_{1}, a_{2}, q_{1}, q_{2}}\left|a_{1}, a_{2}, q_{1}, q_{2}\right\rangle\left\langle a_{1}, a_{2}, q_{1}, q_{2}\right|$. Then, using Lemma 21 we
can rewrite the objective function in Eq. (7.6) as

$$
\begin{align*}
& \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \rho_{T \hat{T}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}}\right)\right] \\
& =\operatorname{Tr}\left[\left(\mathbb{1}_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \rho_{T \hat{T}}\right)\left(\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \mathbb{1}_{T \hat{T}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}}\right)\right)\right] \\
& =\operatorname{Tr}\left[\left(\mathbb{1}_{A_{1} A_{2} Q_{1} Q_{2}} \otimes F_{T \hat{T} \mid S \hat{S}}\right)\left(\left(\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \mathbb{1}_{T \hat{T}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}}\right)\right) \otimes \rho_{S \hat{S}}\right)\right] \\
& \quad \quad \text { (by Lemma 21) }  \tag{byLemma21}\\
& =\operatorname{Tr}\left[\left(\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes F_{T \hat{T} \mid S \hat{S}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}} \otimes \rho_{S \hat{S}}\right)\right)\right],
\end{align*}
$$

which has a similar form to the objective function in Lemma 20 with the exception that $F_{T \hat{T} \mid S \hat{S}}$ replaces $\Phi_{T \hat{T} \mid S \hat{S}}$. To complete the proof, we need to express the swap operator $F_{A \mid \hat{A}}$ in terms of the (non-normalised) maximally entangled state $\Phi_{A \mid \hat{A}}=|\Phi\rangle\left\langle\left.\Phi\right|_{A \mid \hat{A}} \text {, where } \mid \Phi\right\rangle_{A \mid \hat{A}}=\sum_{i=1}^{d_{A}}|i\rangle_{A} \otimes|i\rangle_{\hat{A}}$. It is not difficult to see that

$$
\Phi_{A \mid \hat{A}}^{T_{A}}=\left(\sum_{i, j}|i i\rangle\left\langle\left. j j\right|_{A \hat{A}}\right)^{T_{A}}=\sum_{i, j}|j i\rangle\left\langle\left. i j\right|_{A \hat{A}}=F_{A \mid \hat{A}} .\right.\right.
$$

Redefining the variable $\rho_{S \hat{S}}$ as $\rho_{S \hat{S}}^{T}$, we then immediately obtain Eq. (7.5) in Lemma 20 as this last step leaves the constraints invariant.

### 7.2.2 Semidefinite programming relaxations

In the last section, we showed that $\omega_{Q(T)}(V, \pi)$ for a given free game $G=(V, \pi)$ can be rephrased as a variant of quantum separability problems which is subject to additional constraints. As we have studied in Section 6.3.1, we have effective relaxations for quantum separable states; the DPS hierarchy. In this section, we construct a hierarchy of SDP relaxations for the optimisation problem derived in Lemma 20 using the DPS hierarchy.

The way to employ the DPS hierarchy is straightforward. We can simply replace the product optimisation variable with $\left(n_{1}, n_{2}\right)$-extendible states with respect to the appropriate tripartition.

Here is our new approximation algorithm for $\omega_{Q(T)}(V, \pi)$ of a two player free game $(V, \pi)$ :

$$
\begin{equation*}
\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T):=T^{2} \max _{\rho} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right) \rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)(S \hat{S})}\right] \tag{7.7}
\end{equation*}
$$

s.t. $\quad \rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}} \geq 0, \quad \operatorname{Tr}\left[\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}\right]=1$

$$
\begin{equation*}
\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}} \text { perm. inv. on }\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}} \text { w.r.t. }\left(A_{1} Q_{1} T\right)(S \hat{S})^{n_{2}} \tag{7.8}
\end{equation*}
$$

$$
\begin{equation*}
\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}} \text { perm. inv. on }(S \hat{S})^{n_{2}} \text { w.r.t. }\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}} \tag{7.9}
\end{equation*}
$$

$$
\begin{equation*}
\operatorname{Tr}_{A_{1}}\left[\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}\right]=\left(\sum_{q_{1}} \pi_{1}\left(q_{1}\right)\left|q_{1}\right\rangle\left\langle\left. q_{1}\right|_{Q_{1}} \otimes \frac{\mathbb{1}_{T}}{T}\right) \otimes \rho_{\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}\right. \tag{7.10}
\end{equation*}
$$

$$
\begin{equation*}
\operatorname{Tr}_{A_{2}}\left[\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}\right]=\left(\sum_{q_{2}} \pi_{2}\left(q_{2}\right)\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}} \otimes \frac{\mathbb{1}_{\hat{T}}}{T}\right) \otimes \rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{\left(n_{1}-1\right)}(S \hat{S})^{n_{2}}}\right. \tag{7.11}
\end{equation*}
$$

$\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}^{T_{A_{1}} Q_{1} T} \geq 0, \quad \rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}^{T_{\left(A_{2} Q_{2} \hat{n_{2}}\right.} \geq 0, \ldots,}$
where the last line Eq. (7.12) contains all PPT conditions with respect to all the cuts. The first line states that the optimisation variable $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$ must be a quantum state (density operator). Eqs. (7.8)-(7.9) came from the ( $n_{1}, n_{2}$ )-extendibility conditions, permutation invariance on $\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}$ and $(S \hat{S})^{n_{2}}$. We also arrive at the additional linear constraints in Eqs. (7.10)-(7.11) originated from the constraints in Eq. (7.5). The last line contains all PPT conditions for the $\left(n_{1}, n_{2}\right)$-extendible state $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$.

The derived algorithms are a hierarchy of semidefinite programs whose each level gives a valid upper bound on $\omega_{Q(T)}(V, \pi)$. As the set of $\left(n_{1}, n_{2}\right)$-extendible states gets closer to the set of separable states when $n_{1}$ and $n_{2}$ increase, the higher-level relaxation gives a tighter upper bound than the ones from lower-level relaxations; $\operatorname{sdp}_{n_{1}, n_{2}} \leq \operatorname{sdp}_{n_{1}^{\prime}, n_{2}^{\prime}}$ if $n_{1} \geq n_{1}^{\prime}$ and $n_{2} \geq n_{2}^{\prime}$. However, implementing a high-level relaxation requires more computational resources, such as memory size and run-time. In the next section, we discuss how to find the minimum level of the relaxations for achieving a desired accuracy.

Before moving on to the next section, it is worth remarking that the value of $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ is naturally upper bounded by 1 .

Proposition 22. Let $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ be the $\left(n_{1}, n_{2}\right)$-th level SDP relaxation for the $T \times T$ -
dimensional two-player free game with rule matrix $V$ and question probability distribution $\pi\left(q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$. Then, we have that

$$
0 \leq \operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T) \leq 1
$$

Proof. Let $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$ be the state optimising the $\left(n_{1}, n_{2}\right)$-th level SDP relaxation. As subsystems $A_{1}, A_{2}, Q_{1}, Q_{2}$ are classical, we can write its marginal appeared in the objective function in Eq. (7.7) in the block-diagonal form:

$$
\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)(S \hat{S})}=\sum_{a_{1}, a_{2}, q_{1}, q_{2}}\left|a_{1}, a_{2}, q_{1}, q_{2}\right\rangle\left\langle a_{1}, a_{2}, q_{1}, q_{2}\right| \otimes \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right) .
$$

Then, we can simplify the $A_{1}, A_{2}, Q_{1}, Q_{2}$ part of the trace in the objective function and obtain

$$
\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)=T^{2} \max _{\rho} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right] .
$$

Using $V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \leq 1$ for all $a_{1}, a_{2}, q_{1}, q_{2}$, we find

$$
\begin{align*}
\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T) & \leq T^{2} \max _{\rho} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right] \\
& =T^{2} \operatorname{Tr}_{T \hat{T} S \hat{S}}\left[\Phi_{T \hat{T} \mid S \hat{S}} \operatorname{Tr}_{A_{1} Q_{1} A_{2} Q_{2}}\left[\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)(S \hat{S})}\right]\right] . \tag{7.13}
\end{align*}
$$

We can now use the linear constraints to write $\rho_{T \hat{T} S \hat{S}}$ explicitly. Consider the first linear constraint in Eq. (7.10), and taking the trace on the subsystem $A_{1}$ gives us

$$
\operatorname{Tr}_{A_{1} A_{2}}\left[\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}\right]=\left(\sum_{q_{1}} \pi_{1}\left(q_{1}\right)\left|q_{1}\right\rangle\left\langle\left. q_{1}\right|_{Q_{1}} \otimes \frac{\mathbb{1}_{T}}{T}\right) \otimes \operatorname{Tr}_{A_{2}}\left[\rho_{\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}\right] . .14\right)
$$

If we now consider the second linear constraint in Eq. (7.11) and take the trace on the subsystems $A_{1} Q_{1} T$ and $A_{2}$, then we get

$$
\begin{equation*}
\operatorname{Tr}_{A_{2}}\left[\rho_{\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}\right]=\left(\sum_{q_{2}} \pi_{2}\left(q_{2}\right)\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}} \otimes \frac{\mathbb{1}_{\hat{T}}}{T}\right) \otimes \rho_{\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}-1}(S \hat{S})^{n 2}}\right. \tag{7.15}
\end{equation*}
$$

We combine Eq. (7.14) and (7.15) and take the trace on the remaining $n_{1}-1$ subsystems $A_{2} Q_{2} \hat{T}$ and $n_{2}-1$ subsystems $S \hat{S}$. Then, we can conclude that

$$
\begin{aligned}
& \operatorname{Tr}_{A_{1} A_{2}}\left[\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)(S \hat{S})}\right] \\
& =\left(\sum _ { q _ { 1 } } \pi _ { 1 } ( q _ { 1 } ) | q _ { 1 } \rangle \langle q _ { 1 } | _ { Q _ { 1 } } \otimes \frac { \mathbb { 1 } _ { T } } { T } ) \otimes \left(\sum_{q_{2}} \pi_{2}\left(q_{2}\right)\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}} \otimes \frac{\mathbb{1}_{\hat{T}}}{T}\right) \otimes \rho_{S \hat{S}}\right.\right.
\end{aligned}
$$

If we take the trace over subsystems $Q_{1} Q_{2}$, we obtain the following explicit form of $\rho_{T \hat{T} S \hat{S}}$ :

$$
\rho_{T \hat{T} S \hat{S}}=\operatorname{Tr}_{A_{1} Q_{1} A_{2} Q_{2}}\left[\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)(S \hat{S})}\right]=\frac{\mathbb{1}_{T \hat{T}}}{T^{2}} \otimes \rho_{S \hat{S}}
$$

By substituting the above expression in Eq. (7.13) and the swap trick, we find that

$$
\begin{aligned}
\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T) & \leq \operatorname{Tr}_{T \hat{T} S \hat{S}}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\mathbb{1}_{T \hat{T}} \otimes \rho_{S \hat{S}}\right)\right] \\
& =\operatorname{Tr}_{T \hat{T} S \hat{S}}\left[F_{T \hat{T} \mid S \hat{S}}\left(\mathbb{1}_{T \hat{T}} \otimes \rho_{S \hat{S}}\right)\right] \\
& =\operatorname{Tr}_{T \hat{T}}\left[\rho_{T \hat{T}}\right]=1,
\end{aligned}
$$

where we have taken the partial transpose over $T \hat{T}$ in the second line and used $\operatorname{Tr}\left[M_{A B} N_{A B}^{\mathrm{T}_{A}}\right]=$ $\operatorname{Tr}\left[M_{A B}^{\mathrm{T}_{A}} N_{A B}\right]$ in the last line.

### 7.3 Convergence of the hierarchy

We have constructed approximation algorithms for calculating $\omega_{Q(T)}(V, \pi)$ for a given two-player free game ( $V, \pi$ ) in the last section, which has the form of a hierarchy of SDP relaxations. A natural following-up question is then how fast these SDP relaxations converge to the actual solution, and to answer this question is the main topic of this section. Our approach is to find the lowest level of the SDP hierarchy which can achieve a desired accuracy, which can be done by the help of quantum de Finetti theorems. Then, this information can be further related to the convergence speed of the hierarchy.

### 7.3.1 Preliminary: conditional quantum mutual information

Before diving into proving the convergence speed of the SDP hierarchy, we would like to provide some mathematical concepts necessary for the next section. In particular, we introduce the mathematical quantity called conditional quantum mutual information whose multipartite version is one of the main technical tools in the next section.

Let us consider a quantum state $\rho_{A}$ on a Hilbert space $\mathcal{H}_{A}$. The von Neumann entropy of the quantum state $\rho_{A}$ is defined as

$$
S(A)_{\rho}:=-\operatorname{Tr}\left[\rho_{A} \log \rho_{A}\right]
$$

where we define $0 \log 0 \equiv 0$. When considering a quantum state with subsystems, such as $\rho_{A_{1} \cdots A_{k}}$, the quantum joint entropy is nothing but the von Neumann entropy of the whole system:

$$
S\left(A_{1} \cdots A_{k}\right)_{\rho}=-\operatorname{Tr}\left[\rho_{A_{1} \cdots A_{k}} \log \rho_{A_{1} \cdots A_{k}}\right] .
$$

Then, for the bipartite setting, we can define the conditional entropy and quantum mutual information of a quantum state $\rho_{A B}$ on $\mathcal{H}_{A B}$ by

$$
\begin{align*}
& S(A \mid B)_{\rho}:=S(A B)_{\rho}-S(B)_{\rho} \\
& I(A: B)_{\rho}:=S(A)_{\rho}+S(B)_{\rho}-S(A B)_{\rho} . \tag{7.16}
\end{align*}
$$

There exist a few different ways to generalise the definition of quantum mutual information in Eq. (7.16) to the multipartite setting [112]. In this chapter, we use the following definition: given $k$ quantum systems $A_{1} \cdots A_{k}$ described by $\rho_{A_{1} A_{2} \cdots A_{k}}$, the multipartite quantum mutual information is defined as

$$
I\left(A_{1}: A_{2}: \cdots: A_{k}\right)_{\rho}:=\sum_{i=1}^{k} S\left(A_{i}\right)_{\rho}-S\left(A_{1} A_{2} \cdots A_{k}\right),
$$

where $S\left(A_{i}\right)=-\operatorname{Tr}\left[\rho_{A_{i}} \log \rho_{A_{i}}\right]$ is the von Neumann entropy of the reduced state $\rho_{A_{i}}$. This definition is useful as it is equivalent to another measure of correlation between subsystems, the
relative entropy distance between the given state and the tensor product of its marginals:

$$
\begin{equation*}
I\left(A_{1}: A_{2}: \cdots: A_{k}\right)_{\rho}=D\left(\rho_{A_{1} \cdots A_{k}} \| \rho_{A_{1}} \otimes \cdots \otimes \rho_{A_{i}}\right) \tag{7.17}
\end{equation*}
$$

where $D(\rho \| \sigma):=\operatorname{Tr}[\rho(\log \rho-\log \sigma)]$ is the relative entropy between two state $\rho$ and $\sigma$ such that $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$. Given $k$ quantum systems $A_{1} \cdots A_{k}$ and a classical system $R$, described by the global state $\rho_{A_{1} A_{2} \cdots A_{k} R}$, the conditional multipartite quantum mutual information is defined as

$$
I\left(A_{1}: A_{2}: \cdots: A_{k} \mid R\right)_{\rho}:=\sum_{i=1}^{k} S\left(A_{i} R\right)-S\left(A_{1} A_{2} \cdots A_{k} R\right)-S(R)
$$

The conditional multipartite quantum mutual information with classical system $R$ can be expressed in terms of the bipartite ones [113, Lemma 13]:

$$
\begin{equation*}
I\left(A_{1}: \cdots: A_{k} \mid R\right)_{\rho}=I\left(A_{1}: A_{2} \mid R\right)_{\rho}+I\left(A_{1} A_{2}: A_{3} \mid R\right)_{\rho}+\ldots+I\left(A_{1} \cdots A_{k-1}: A_{k} \mid R\right)_{\rho} \tag{7.18}
\end{equation*}
$$

Another convenient property of quantum mutual information is the chain rule:

$$
\begin{equation*}
I(A B: C)=I(B: C)+I(A: C \mid B) \tag{7.19}
\end{equation*}
$$

which holds for the conditional quantum mutual information as well:

$$
\begin{equation*}
I(A B: C \mid D)=I(B: C \mid D)+I(A: C \mid B D) \tag{7.20}
\end{equation*}
$$

### 7.3.2 Tripartite quantum de Finetti theorem with linear constraints

As we discussed in Section 6.4, quantum de Finetti theorems can be related to a quantitative bound on how close extendible states are to the set of separable states. This information can then be converted to an upper bound on the accuracy of the SDP relaxations exploiting the DPS hierarchy. However, the existences of both the side quantum system in extendible states and the additional linear constraints in Eq. (7.5) prevent us to be able to exploit the standard
quantum de Finetti theorem like Theorem 19, and thus we need an adapted version. What we need is an upper bound on how close $n$-extendible states satisfying additional linear constraints are to the convex hull of product states satisfying the same linear constraints.

In this section, we derive multipartite quantum de Finetti theorems with additional linear constraints employing the information-theoretic proof technique based on quantum entropy inequalities $[113,114]$. Using the adapted quantum de Finetti theorem with improved distance bound is crucial to obtain the improved complexity bounds on the problem of approximating $\omega_{Q(T)}(V, \pi)$ in the next section. Here, we state the tripartite version of the theorem.

Theorem 23. Let $\rho_{A B^{n_{1}} C^{n_{2}}}$ be a quantum state which is invariant under permutations on $B^{n_{1}}$ with respect to $A C^{n_{2}}$ and on $C^{n_{2}}$ with respect to $A B^{n_{1}}$, satisfying for linear maps $\mathcal{E}_{A \rightarrow \tilde{A}}, \Lambda_{B \rightarrow \tilde{B}}$, and $\Gamma_{C \rightarrow \tilde{C}}$ and operators $\mathbf{X}_{\tilde{A}}, \mathbf{Y}_{\tilde{B}}$, and $\mathbf{Z}_{\tilde{C}}$ that

$$
\begin{aligned}
\left(\mathcal{E}_{A \rightarrow \tilde{A}} \otimes \mathcal{I}_{B^{n_{1}} C^{n_{2}}}\right)\left(\rho_{\left.A B^{n_{1} C^{n_{2}}}\right)}\right)=\mathbf{X}_{\tilde{A}} \otimes \rho_{B^{n_{1}} C^{n_{2}}} & \text { linear constraint on } A \\
\left(\Lambda_{B \rightarrow \tilde{B}} \otimes \mathcal{I}_{B^{n_{1}-1} C^{n_{2}}}\right)\left(\rho_{\left.B^{n_{1} C^{n_{2}}}\right)}\right)=\mathbf{Y}_{\tilde{B}} \otimes \rho_{B^{n_{1}-1} C^{n_{2}}} & \text { linear constraint on } B \\
\left(\mathcal{I}_{B^{n_{1} C^{n_{2}-1}}} \otimes \Gamma_{C \rightarrow \tilde{C}}\right)\left(\rho_{B^{n_{1} C^{n_{2}}}}\right)=\mathbf{Z}_{\tilde{C}} \otimes \rho_{B^{n_{1} C^{n_{2}-1}}} & \text { linear constraint on } .
\end{aligned}
$$

Then, there exist a probability distribution $\left\{p_{i}\right\}_{i \in I}$ and sets of quantum states $\left\{\sigma_{A}^{i}\right\}_{i \in I},\left\{\omega_{B}^{i}\right\}_{i \in I}$ and $\left\{\tau_{C}^{i}\right\}_{i \in I}$ such that we have that

$$
\begin{align*}
& \left\|\rho_{A B C}-\sum_{i \in I} p_{i} \sigma_{A}^{i} \otimes \omega_{B}^{i} \otimes \tau_{C}^{i}\right\|_{1} \\
& \quad \leq \min \left\{18^{3 / 2} \sqrt{|A B C|, 4|B C|}\right\} \times \sqrt{2 \ln 2}\left(\sqrt{\frac{\log |A|+8 \log |B|}{n_{2}}+\frac{\log |A|}{n_{1}}}\right)  \tag{7.21}\\
& \quad \mathcal{E}_{A \rightarrow \tilde{A}}\left(\sigma_{A}^{i}\right)=\mathbf{X}_{\tilde{A}}, \quad \Lambda_{B \rightarrow \tilde{B}}\left(\omega_{B}^{i}\right)=\mathbf{Y}_{\tilde{B}}, \quad \Gamma_{C \rightarrow \tilde{C}}\left(\tau_{C}^{i}\right)=\mathbf{Z}_{\tilde{C}} \quad \forall i \in I .
\end{align*}
$$

To prove the above theorem, instead of directly working with the trace distance, we exploit quantum entropy inequalities and their chain rules. This approach allows us to carefully quantify how correlations are divided between different partitions of the extendible states.

For the first ingredient, we derive a general upper bound on the conditional multipartite
quantum mutual information of a state with classical subsystems. Before then, we prove the following lemma which is needed to prove the first ingredient.

Lemma 24. For a quantum state $\rho_{A Z X}$ with classical systems $Z$ and $X$, it holds that

$$
I(A: Z \mid X) \leq \log |A|
$$

Proof. For classical system $X$, the conditional quantum mutual information can be written as $I(A: Z \mid X)=\sum_{x} p_{x} I(A: Z)_{\rho_{A Z \mid x}}$. We know that $I(A: Z) \leq \log |A|$ when $Z$ is classical, and hence

$$
I(A: Z \mid X)=\sum_{x} p_{x} I(A: Z)_{\rho_{A Z \mid x}} \leq \log |A|
$$

Note that we denote the conditional state $\rho_{A \mid z}$ of a state $\rho_{A Z}$ with classical system $Z$ as

$$
\rho_{A \mid z}:=\frac{\operatorname{Tr}_{Z}\left[\rho_{A Z}\left(\mathbb{1}_{A} \otimes|z\rangle\left\langle\left. z\right|_{Z}\right)\right]\right.}{\operatorname{Tr}\left[\rho_{A Z}\left(\mathbb{1}_{A} \otimes|z\rangle\left\langle\left. z\right|_{Z}\right)\right]\right.} .
$$

This describes the state after measurement on system $Z$, when the measurement outcome is $z$. Now, we are ready to prove the first ingredient.

Lemma 25. Consider a quantum state $\rho_{A Z^{n_{1}} W^{n_{2}}}$ classical on systems $Z$ and $W$. Then, there exist $\bar{m} \in\left\{0, \ldots, n_{1}-1\right\}$ and $\bar{l} \in\left\{0, \ldots, n_{2}-1\right\}$ such that

$$
\begin{equation*}
I\left(A: Z_{\bar{m}+1}: W_{\bar{l}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right) \leq \frac{\log |A|}{n_{1}}+\frac{\log |A|+\log |Z|}{n_{2}} \tag{7.22}
\end{equation*}
$$

Moreover, by Pinsker's inequality, this implies that

$$
\begin{align*}
& \mathbb{E}_{z^{\bar{m}} w^{\bar{l}}}\left\{\left\|\rho_{A Z_{\bar{m}+1} W_{\bar{l}+1} \mid z^{\bar{m}} w^{\bar{l}}}-\rho_{A \mid z^{\bar{m}} w^{\bar{l}}} \otimes \rho_{Z_{\bar{m}+1} \mid z^{\bar{m}} w^{\bar{l}}} \otimes \rho_{W_{\bar{l}+1} \mid z^{\bar{m}} w^{\bar{\iota}}}\right\|_{1}^{2}\right\}  \tag{7.23}\\
& \leq 2 \ln 2\left(\frac{\log |A|}{n_{1}}+\frac{\log |A|+\log |Z|}{n_{2}}\right) .
\end{align*}
$$

Proof. The multipartite quantum mutual information $I\left(A: Z_{\bar{m}+1}: W_{\bar{l}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right)$ can be expressed in terms of bipartite ones using Eq. (7.18):

$$
\begin{equation*}
I\left(A: Z_{\bar{m}+1}: W_{\bar{l}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right)=I\left(A: Z_{\bar{m}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right)+I\left(A Z_{\bar{m}+1}: W_{\bar{l}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right) \tag{7.24}
\end{equation*}
$$

The two terms in the RHS are the bipartite mutual information between quantum and classical systems, and this allows us to find an upper bound for each term using the chain rule in Eq. (7.20) and Lemma 24.

First term: For any $l$, it holds that

$$
I\left(A: Z^{n_{1}} \mid W^{l}\right)=\sum_{m=0}^{n_{1}-1} I\left(A: Z_{m+1} \mid Z^{m} W^{l}\right) \leq \log |A|,
$$

where the first equality is the chain rule in Eq. (7.20) and the second inequality is a result of applying Lemma 24 to $I\left(A: Z^{n_{1}} \mid W^{l}\right)$. Then, summing over all $l$ gives us

$$
\begin{equation*}
\sum_{m=0}^{n_{1}-1} \sum_{l=0}^{n_{2}-1} I\left(A: Z_{m+1} \mid Z^{m} W^{l}\right) \leq n_{2} \log |A| \tag{7.25}
\end{equation*}
$$

Second term: Using the same argument, for any $m$, it holds that

$$
I\left(A Z_{m+1}: W^{n_{2}} \mid Z^{m}\right)=\sum_{l=0}^{n_{2}-1} I\left(A Z_{m+1}: W_{l+1} \mid Z^{m} W^{l}\right) \leq \log |A||Z|
$$

and summing over $m$ gives us

$$
\begin{equation*}
\sum_{m=0}^{n_{1}-1} \sum_{l=0}^{n_{2}-1} I\left(A Z_{m+1}: W_{l+1} \mid Z^{m} W^{l}\right) \leq n_{1}(\log |A|+\log |Z|) \tag{7.26}
\end{equation*}
$$

Combining Eq. (7.25) and Eq. (7.26) gives

$$
\begin{aligned}
& n_{2} \log |A|+n_{1}(\log |A|+\log |Z|) \\
& \geq \sum_{m=0}^{n_{1}-1} \sum_{l=0}^{n_{2}-1}\left(I\left(A: Z_{m+1} \mid Z^{m} W^{l}\right)+I\left(A Z_{m+1}: W_{l+1} \mid Z^{m} W^{l}\right)\right) \\
& \geq n_{1} n_{2}\left(I\left(A: Z_{\bar{m}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right)+I\left(A Z_{\bar{m}+1}: W_{\bar{l}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right)\right),
\end{aligned}
$$

where $\bar{m}$ and $\bar{l}$ are the indices of the smallest element in the sum. Dividing both sides by $n_{1} n_{2}$ gives us the desired relation,

$$
\begin{aligned}
I\left(A: Z_{\bar{m}+1}: W_{\bar{l}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right) & =I\left(A: Z_{\bar{m}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right)+I\left(A Z_{\bar{m}+1}: W_{\bar{l}+1} \mid Z^{\bar{m}} W^{\bar{l}}\right) \\
& \leq \frac{\log |A|}{n_{1}}+\frac{\log |A|+\log |Z|}{n_{2}}
\end{aligned}
$$

This ends the proof of Eq. (7.22). Then, using Eq. (7.17) and Pinsker's inequality ${ }^{2}$ we can obtain Eq. (7.23).

We remark that the resulting bound is not symmetric under exchanging the systems $Z$ and $W$ even though the conditions are symmetric. This is because we used the particular bipartite expression for the conditional multipartite quantum mutual information in Eq. (7.24). The same proof works with the roles of the systems $Z$ and $W$ exchanged, and one can freely choose whichever bound is stronger. We also remark that it is straightforward to extend the above result in Lemma 25 to the general $k$-partite case.

The next ingredient is a bound on the loss in distinguishability between quantum states when they are processed with measurements - called minimal distortion.

Lemma 26. 1. There exist fixed measurements $\mathcal{M}_{A}, \mathcal{M}_{B}$, and $\mathcal{M}_{C}$ with at most $|A|^{8},|B|^{8}$, and $|C|^{8}$ outcomes, respectively, such that for every traceless Hermitian operator $\gamma_{A B C}$ on $\mathcal{H}_{A B C}$

$$
\left\|\gamma_{A B C}\right\|_{1} \leq 18^{3 / 2} \sqrt{|A B C|} \cdot\left\|\left(\mathcal{M}_{A} \otimes \mathcal{M}_{B} \otimes \mathcal{M}_{C}\right)\left(\gamma_{A B C}\right)\right\|_{1}
$$

2. There exists a fixed measurement $\mathcal{M}_{B}$ with at most $|B|^{6}$ outcomes such that for every traceless Hermitian operator $\gamma_{A B}$ on $\mathcal{H}_{A B}$

$$
\begin{equation*}
\left\|\gamma_{A B}\right\|_{1} \leq 2|B| \cdot\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B}\right)\left(\gamma_{A B}\right)\right\|_{1} \tag{7.28}
\end{equation*}
$$

[^35]Note that as the trace distance is a contractive metric, it in general holds that $\|\mathcal{M}(\gamma)\|_{1} \leq\|\gamma\|$, which represents the loss of information during the process $\mathcal{M}$. The above lemma concerns lower bounds, instead of this trivial upper bound, on the loss of information.

The first part is straightforward from [114, Lemma 14]. We remark that when a traceless Hermitian operator already has a classical subsystem, for example $\gamma_{A B C Z}$ with classical system $Z$, the dimension factor only includes the dimension of the quantum part; i.e.,

$$
\begin{equation*}
\left\|\gamma_{A B C Z}\right\|_{1} \leq 18^{3 / 2} \sqrt{|A B C|} \cdot\left\|\left(\mathcal{M}_{A} \otimes \mathcal{M}_{B} \otimes \mathcal{M}_{C} \otimes \mathcal{I}_{Z}\right)\left(\gamma_{A B C Z}\right)\right\|_{1} \tag{7.29}
\end{equation*}
$$

This follows from the property of classical-quantum states that

$$
\| \sum_{z} \rho_{A}^{z} \otimes|z\rangle\langle z|\left\|_{1}=\sum_{z}\right\| \rho_{A}^{z} \|_{1} .
$$

The proof of the second part is given in Appendix C.2. We note that the second part improves on the factor $\sqrt{18}|B|^{3 / 2}$ given in [113, Eq. (68)]. As there exist quantum states $\rho_{A B}$ and $\sigma_{A B}$ such that [115]

$$
\left\|\rho_{A B}-\sigma_{A B}\right\|_{1}=2 \quad \text { and } \quad \sup _{\mathcal{M}_{B}}\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B}\right)\left(\gamma_{A B}\right)\right\|_{1}=\frac{2}{|B|+1},
$$

our result establishes the dimension dependence $\Omega(|B|)$ for minimal distortion relative to quantum side information. This answers a question left open in [116].

## Proof of Theorem 23

Now, we prove Theorem 23. The idea is to apply one of the optimal measurements specified in Lemma 26 to the given extendible state $\rho_{A B^{n_{1} C^{n_{2}}}}$ in Theorem 23 so that the resulting state is partially classical, and then apply Lemma 25 to the resulting classical-quantum state. Here is the proof.

Proof of Theorem 23. Let $\mathcal{M}_{B \rightarrow Y}$ be a quantum-to-classical measurement from $B$ to the classical system $Y$, and $\mathcal{M}_{C \rightarrow Z}$ be a quantum-to-classical measurement from $C$ to the classical system $Z$. We apply these measurements to the quantum state $\rho_{A B^{n_{1} C^{n_{2}}}}$ and will denote the outcome
classical-quantum state as $\rho_{A Y^{n_{1}} Z^{n_{2}}}$. Then, according to Lemma 25 , there exist $m \in\left\{0, \cdots, n_{1}-\right.$ $1\}$ and $\ell \in\left\{0, \cdots, n_{2}-1\right\}$ such that

$$
\begin{align*}
\mathbb{E}_{y^{m} z^{\ell}}\left\{\| \rho_{A Y_{m+1} Z_{\ell+1} \mid y^{m} z^{\ell}}-\rho_{A \mid y^{m} z^{\ell}}\right. & \left.\otimes \rho_{Y_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{Z_{\ell+1} \mid y^{m} z^{\ell}} \|_{1}^{2}\right\} \\
& \leq 2 \ln 2\left(\frac{\log |A|}{n_{1}}+\frac{\log |A|+\log |Y|}{n_{2}}\right) \tag{7.30}
\end{align*}
$$

As $\rho_{A B^{n_{1}} C^{n_{2}}}$ is invariant under permutations of the systems $B^{n_{1}}$ and $C^{n_{2}}$, we can always find $m$ and $l$ satisfying Eq. (7.30).

Now, let us define a traceless Hermitian operator

$$
\gamma_{A B C} \equiv \rho_{A B_{m+1} C_{\ell+1} \mid y^{m} z^{\ell}}-\rho_{A \mid y^{m} z^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{C_{\ell+1} \mid y^{m} z^{\ell}}
$$

Note that

$$
\mathcal{I}_{A} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow Z}\left(\gamma_{A B C}\right)=\rho_{A Y_{m+1} Z_{\ell+1} \mid y^{m} z^{\ell}}-\rho_{A \mid y^{m} z^{\ell}} \otimes \rho_{Y_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{Z_{\ell+1} \mid y^{m} z^{\ell}}
$$

Using the second part of Lemma 26 iteratively, we can obtain

$$
\begin{aligned}
\left\|\gamma_{A B C}\right\|_{1} & \leq 2|C|\left\|\left(\mathcal{I}_{A B} \otimes \mathcal{M}_{C \rightarrow Z}\right)\left(\gamma_{A B C}\right)\right\|_{1} \\
& \leq 2|B| \times 2|C|\left\|\left(\mathcal{I}_{A C} \otimes \mathcal{M}_{B \rightarrow Y}\right)\left(\mathcal{I}_{A B} \otimes \mathcal{M}_{C \rightarrow Z}\right)\left(\gamma_{A B C}\right)\right\|_{1} \\
& =4|B C|\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow Z}\right)\left(\gamma_{A B C}\right)\right\|_{1}
\end{aligned}
$$

with $|Y| \leq|B|^{6}$. We can also exploit the first part of Lemma 26 to obtain

$$
\begin{aligned}
\left\|\gamma_{A B C}\right\|_{1} & \leq \sqrt{18^{3}|A B C|}\left\|\left(\mathcal{M}_{A} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow Z}\right)\left(\gamma_{A B C}\right)\right\|_{1} \\
& \leq \sqrt{18^{3}|A B C|}\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow Z}\right)\left(\gamma_{A B C}\right)\right\|_{1}
\end{aligned}
$$

with $|Y| \leq|B|^{8}$, where the second inequality follows from the monotonicity of the trace norm under CPTP maps. Depending on the dimensions, we can freely choose the tighter bound
between the two cases. Combining Eq. (7.30) with the above two results we obtain

$$
\begin{aligned}
\mathbb{E}_{y^{m} z^{\ell}} & \left\{\left\|\rho_{A B_{m+1} C_{\ell+1} \mid y^{m} z^{\ell}}-\rho_{A \mid y^{m} z^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{C_{\ell+1} \mid y^{m} z^{\ell}}\right\|_{1}^{2}\right\} \\
& \leq \min \left\{\sqrt{18^{3}|A B C|}, 4|B C|\right\}^{2} \times 2 \ln 2\left(\frac{\log |A|}{n_{1}}+\frac{\log |A|+8 \log |B|}{n_{2}}\right)
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
& \left\|\rho_{A B_{m+1} C_{\ell+1}}-\mathbb{E}_{y^{m} z^{\ell}}\left\{\rho_{A \mid y^{m} z^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{C_{\ell+1} \mid y^{m} z^{\ell}}\right\}\right\|_{1} \\
& \leq \mathbb{E}_{y^{m} z^{\ell}}\left\{\left\|\rho_{A B_{m+1} C_{\ell+1} \mid y^{m} z^{\ell}}-\rho_{A \mid y^{m} z^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{C_{\ell+1} \mid y^{m} z^{\ell}}\right\|_{1}\right\} \\
& \leq \sqrt{\mathbb{E}_{y^{m} z^{\ell}}\left\{\left\|\rho_{A B_{m+1} C_{\ell+1} \mid y^{m} z^{\ell}}-\rho_{A \mid y^{m} z^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{C_{\ell+1} \mid y^{m} z^{\ell}}\right\|_{1}^{2}\right\}} \\
& \leq \min \left\{\sqrt{18^{3}|A B C|}, 4|B C|\right\} \times \sqrt{2 \ln 2}\left(\sqrt{\frac{\log |A|}{n_{1}}+\frac{\log |A|+8 \log |B|}{n_{2}}}\right)
\end{aligned}
$$

where we used the triangular inequality for Schatten $p$-norms in the second line and the concavity of the square function in the third line. As $\mathbb{E}_{y^{m} z^{\ell}}\left\{\rho_{A \mid y^{m} z^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} z^{\ell}} \otimes \rho_{C+1} \mid y^{m} z^{\ell}\right\}$ is a separable state with respect to the tripartition $A|B| C$, this proves the first half of the theorem.

The remaining part is to check whether $\rho_{A \mid y^{m} z^{\ell}}, \rho_{B_{m+1} \mid y^{m} z^{\ell}}$ and $\rho_{C_{\ell+1} \mid y^{m} z^{\ell}}$ satisfy the desired linear constraints. Let us denote $M_{B_{i}}^{y_{i}}$ and $M_{C_{i}}^{z_{i}}$ as the POVM elements of the measurements $\mathcal{M}_{B_{i} \rightarrow Y_{i}}$ and $\mathcal{M}_{C_{i} \rightarrow Z_{i}}$ corresponding to the measurement outcomes $y_{i}$ and $z_{i}$, respectively. Then, we find

$$
\begin{aligned}
& \mathcal{E}_{A \rightarrow \tilde{A}}\left(\sigma_{A}^{i}\right)=\mathcal{E}_{A \rightarrow \tilde{A}}\left(\rho_{A \mid y^{m} z^{\ell}}\right) \\
& =\frac{\operatorname{Tr}_{B^{m} C^{\ell}}\left[\left(\mathbb{1}_{A} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right) \mathcal{E}_{A \rightarrow \tilde{A}}\left(\rho_{A B^{m} C^{\ell}}\right)\right]}{\operatorname{Tr}\left[\left(\mathbb{1}_{A} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right) \rho_{A B^{m} C^{\ell}}\right]} \\
& =\frac{\operatorname{Tr}_{B^{m} C^{\ell}}\left[\left(\mathbb{1}_{A} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{\ell_{\ell}}\right)\left(\mathcal{X}_{\tilde{A}} \otimes \rho_{B^{m} C^{\ell}}\right)\right]}{\operatorname{Tr}\left[\left(\mathbb{1}_{A} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right) \rho_{A B^{m} C^{\ell}}\right]} \\
& =\mathcal{X}_{\tilde{A}} .
\end{aligned}
$$

$$
\begin{aligned}
& \Lambda_{B \rightarrow \tilde{B}}\left(\omega_{B}^{i}\right)=\Lambda_{B \rightarrow \tilde{B}}\left(\rho_{B_{m+1} \mid y^{m} z^{\ell}}\right) \\
& =\frac{\operatorname{Tr}_{B^{m} C^{\ell}}\left[\left(\mathbb{1}_{\tilde{B}} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right) \Lambda_{B \rightarrow \tilde{B}}\left(\rho_{B^{m+1} C^{\ell}}\right)\right]}{\operatorname{Tr}\left[\left(M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes \mathbb{1}_{B_{m+1}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right) \rho_{B^{m+1} C^{\ell}}\right]} \\
& =\frac{\operatorname{Tr}_{B^{m} C^{\ell}}\left[\left(\mathbb{1}_{\tilde{B}} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right)\left(\mathcal{Y}_{\tilde{B}} \otimes \rho_{B^{m} C^{\ell}}\right)\right]}{\operatorname{Tr}\left[\left(M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes \mathbb{1}_{B_{m+1}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right) \rho_{B^{m+1} C^{\ell}}\right]} \\
& =\mathcal{Y}_{\tilde{B}} . \\
& =\frac{\Gamma_{C \rightarrow \tilde{C}}\left(\tau_{C}^{i}\right)=\Gamma_{C \rightarrow \tilde{C}}\left(\rho_{C_{\ell+1} \mid y^{m} z^{\ell}}\right)}{} \\
& =\frac{\operatorname{Tr}_{B^{m} C^{\ell}}\left[\left(\mathbb{1}_{\tilde{C}} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right) \Gamma_{C \rightarrow \tilde{C}}\left(\rho_{B^{m} C^{\ell+1}}\right)\right]}{\operatorname{Tr}\left[\left(M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}} \otimes \mathbb{1}_{C_{\ell+1}}\right)\left(\rho_{B^{m} C^{\ell+1}}\right)\right]} \\
& =\frac{\operatorname{Tr}_{B^{m} C^{\ell}}\left[\left(\mathbb{Z}_{\tilde{C}} \otimes M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}}\right)\left(\mathcal{Z}_{\tilde{C}} \otimes \rho_{B^{m} C^{\ell}}\right)\right]}{\operatorname{Tr}\left[\left(M_{B_{1}}^{y_{1}} \otimes \cdots \otimes M_{B_{m}}^{y_{m}} \otimes M_{C_{1}}^{z_{1}} \otimes \cdots \otimes M_{C_{\ell}}^{z_{\ell}} \otimes \mathbb{1}_{C_{\ell+1}}\right)\left(\rho_{B^{m} C^{\ell+1}}\right)\right]} \\
&
\end{aligned}
$$

Theorem 23 describes a general setting; both the extendible state and the linear constraints do not have any refined structures. However, if we have more information about the state and the constraint, we can exploit the information to improve the dimension dependence of the bound in Eq. (7.21). It is indeed our case. The extendible state $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$ in $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$, to which we apply the de Finetti theorem, already has some classical systems, and the linear constraints are partial trace constraints. This information helps us obtain a better bound in the quantum de Finetti theorem, and we state this special case as a lemma with proof in Appendix C.3.

Lemma 27. Let $\rho_{(A X \tilde{X}) B^{n_{1}}(C Z \tilde{Z})^{n_{2}}}$ be a quantum state with classical $X \tilde{X}$ - and $Z \tilde{Z}$-systems invariant under permutation on $B^{n_{1}}$ and $(C Z \tilde{Z})^{n_{2}}$ with respect to the other systems, satisfying

$$
\begin{align*}
& \operatorname{Tr}_{X}\left[\rho_{(A X \tilde{X}) B^{n_{1}}(C Z \tilde{Z})^{n_{2}}}\right]=\mathcal{X}_{A \tilde{X}} \otimes \rho_{B^{n_{1}}(C Z \tilde{Z})^{n_{2}}}  \tag{7.31}\\
& \operatorname{Tr}_{Z}\left[\rho_{(A X \tilde{X}) B^{n_{1}}(C Z \tilde{Z})^{n_{2}}}\right]=\mathcal{Z}_{C \tilde{Z}} \otimes \rho_{(A X \tilde{X}) B^{n_{1}}(C Z \tilde{Z})^{n_{2}-1}}
\end{align*}
$$

for some operators $\mathcal{X}_{A \tilde{X}}$, and $\mathcal{Z}_{C \tilde{Z}}$. Then, there exist a probability distribution $\left\{p_{i}\right\}_{i \in I}$ and sets
of quantum states $\left\{\sigma_{A X \tilde{X}}^{i}\right\}_{i \in I},\left\{\omega_{B}^{i}\right\}_{i \in I}$ and $\left\{\tau_{C Z \tilde{Z}}^{i}\right\}_{i \in I}$ such that

$$
\begin{align*}
& \left\|\rho_{(A X \tilde{X}) B(C Z \tilde{Z})}-\sum_{i \in I} p_{i} \sigma_{A X \tilde{X}}^{i} \otimes \omega_{B}^{i} \otimes \tau_{C Z \tilde{Z}}^{i}\right\|_{1} \\
& \leq \min \left\{18^{3 / 2} \sqrt{|A B C|}, 4|B C|\right\} \times \sqrt{4 \ln 2}\left(\sqrt{\frac{\log |X|+8 \log |B|}{n_{2}}+\frac{\log |X|}{n_{1}}}\right) \tag{7.32}
\end{align*}
$$

with $\operatorname{Tr}_{X}\left[\sigma_{A X \tilde{X}}^{i}\right]=\mathcal{X}_{A \tilde{X}}$ and $\operatorname{Tr}_{Z}\left[\tau_{C Z \tilde{Z}}^{i}\right]=\mathcal{Z}_{C \tilde{Z}}$ for all $i \in I$.

### 7.3.3 Convergence rate of the hierarchy

The variants of quantum de Finetti theorems derived in the last section allows us to find an upper bound on the accuracy of the SDP relaxations in Eq. (7.7). We derive analytical bounds on the convergence speed of our SDP hierarchy in terms of the quantum dimension $T$ and the size of the given two-player free game, $|A| \equiv A$ and $|Q| \equiv Q$.

Theorem 28. Let $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ be the $\left(n_{1}, n_{2}\right)$-th level $S D P$ relaxation for the $T \times T$ dimensional two-player free game with rule matrix $V$ and question probability distribution $\pi\left(q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$. Then, we have

$$
0 \leq \operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)-\omega_{Q(T)}(V, \pi) \leq O\left(T^{6} \sqrt{\frac{\log (T A)}{n_{2}}+\frac{\log A}{n_{1}}}\right)
$$

Proof. Let $\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}$ be the optimal state of the $\left(n_{1}, n_{2}\right)$-th level relaxation $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$. The state should be $\left(n_{1}, n_{2}\right)$-extendible since all feasible states must be $\left(n_{1}, n_{2}\right)$-extendible states

# CHAPTER 7. CHARACTERISING QUANTUM CORRELATIONS OF FIXED DIMENSION 

satisfying the linear constraints. Then, we have

$$
\begin{aligned}
& \operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)=T^{2} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right) \rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}\right] \\
& =T^{2} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right)\left(\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right)\right] \\
& \quad+T^{2} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right)\left(\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}-\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right)\right] \\
& \leq \omega_{Q(T)}(V, \pi)+T^{2} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right)\left(\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}-\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right)\right],
\end{aligned}
$$

where $\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}$ is one of the close separable states to $\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}$ specified by Theorem 23. As $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ is an upper bound for $\omega_{Q(T)}(V, \pi)$ we obtain

$$
\begin{aligned}
& \left|\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)-\omega_{Q(T)}(V, \pi)\right| \\
& \quad \leq T^{2}\left|\operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right)\left(\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}-\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right)\right]\right| \\
& \leq T^{2}\left\|V_{A_{1} A_{2} Q_{1} Q_{2}} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right\|_{\infty}\left\|\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}-\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right\|_{1}
\end{aligned}
$$

( $\because$ Hölder's inequality)
$=T^{2}\left\|V_{A_{1} A_{2} Q_{1} Q_{2}}\right\|_{\infty}\left\|\Phi_{T \hat{T} \mid S \hat{S}}\right\|_{\infty}\left\|\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}-\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right\|_{1}$
$=T^{4}\left\|\rho_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}-\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right\|_{1}$

$$
\left(\because\left\|V_{A_{1} A_{2} Q_{1} Q_{2}}\right\|_{\infty}=1,\left\|\Phi_{T \hat{T} \mid S \hat{S}}\right\|_{\infty}=T^{2}\right)
$$

$\leq T^{4}\left[18^{3 / 2} T^{2}(\sqrt{2 \ln 2})\left(\sqrt{\frac{(\log |A|+8 \log |S \hat{S}|)}{n_{2}}+\frac{\log |A|}{n_{1}}}\right)\right]$
( $\because$ Lemma 27)
$=18^{3 / 2} T^{6}(\sqrt{2 \ln 2})\left(\sqrt{\frac{(\log A+16 \log T)}{n_{2}}+\frac{\log A}{n_{1}}}\right)$.
Here, we set $A=T, X=A_{1}, \tilde{X}=Q_{1} B=S \hat{S}, C=\hat{T}, Z=A_{2}$, and $\tilde{Z}=Q_{2}$ when we applied Lemma 27.

Corollary 29. Let $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ be the $\left(n_{1}, n_{2}\right)$-th level relaxation for the $T \times T$-dimensional two-player free game with rule matrix $V$ and question probability distribution $\pi\left(q_{1}, q_{2}\right)=$ $\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$. Then, we have

$$
\omega_{Q(T)}(V, \pi)=\lim _{n_{1}, n_{2} \rightarrow \infty} \operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)
$$

Theorem 28 provides us the convergence speed of our SDP relaxations. For simplicity, let us assume $n_{1}=n_{2}=n$. To achieve a constant error $\epsilon$, we need to go up to the following level of the hierarchy:

$$
\mathcal{O}\left(T^{6} \sqrt{\frac{\log (T A)}{n}}\right) \leq \epsilon \quad \Longleftrightarrow \quad n \geq \mathcal{O}\left(T^{12} \frac{\log (T A)}{\epsilon^{2}}\right)
$$

When $n_{1}=n_{2}=\mathcal{O}\left(T^{12} \frac{\log (T A)}{\epsilon^{2}}\right)$, the size of the variables in the SDP is

$$
\begin{equation*}
(A Q T)^{\mathcal{O}\left(T^{12} \frac{\log (T A)}{\epsilon^{2}}\right)}, \text { which is } \exp \left(\mathcal{O}\left(\frac{T^{12}}{\epsilon^{2}}\left(\log ^{2}(A T)+\log Q \log (T A)\right)\right)\right) \tag{7.33}
\end{equation*}
$$

This implies that approximating $\omega_{Q(T)}(V, \pi)$ of a two-player free game with fixed-dimensional quantum correlation can be solved within the additive error $\epsilon>0$ in quasi-polynomial time in terms of the sizes of answers and questions of the game. Note that this convergence rate is derived only from the linear constraints in $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$, and we did not use the PPT constraints.

### 7.4 Combining with the NPA hierarchy

In this section, we discuss how to combine our $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ in Eq. (7.7) with the NPA hierarchy studied in Section 6.3.2. Combining these two hierarchies is advantageous since the resulting SDP relaxation always returns a bound which is equal or tighter than the ones obtained with the individual relaxations.

As we have seen in Section 6.3.2, the NPA hierarchy provides a different set of necessary conditions satisfied by quantum correlations in the form of the positive semidefinite conditions
[100, 101]. Our aim in this section is to develop an appropriate procedure for constructing the NPA matrix $\Gamma_{k}$, given in Eq. (6.17), in terms of the optimisation variable $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$ in our SDP relaxation $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ in Eq. (7.7). As an example, we show how to build the first-level NPA matrix in this section and refer to Appendix C. 4 for the general case.

We begin by recalling that the optimal winning probability for a two-player game $(V, \pi)$ with quantum correlations of dimension $T \times T$ can be written as,

$$
\begin{equation*}
\omega_{Q(T)}(V, \pi)=\max _{p \in \mathcal{Q}_{T}} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right), \tag{7.34}
\end{equation*}
$$

where $\mathcal{Q}_{T}$ is the set of all quantum correlations of dimension $T \times T$. When we take into account that the systems $A_{1}, A_{2}, Q_{1}, Q_{2}$ are classical, the objective function of $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$, given in Eq. (7.7), can be expressed as

$$
\begin{equation*}
\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)=T^{2} \max _{\rho} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right] \tag{7.35}
\end{equation*}
$$

by using the following classical-quantum form of $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}(S \hat{S})^{n_{2}}}}$ :

$$
\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}=\sum_{a_{1}, \boldsymbol{a}_{2}, q_{1}, \boldsymbol{q}_{2}}\left|a_{1}, \boldsymbol{a}_{2}, q_{1}, \boldsymbol{q}_{2}\right\rangle\left\langle a_{1}, \boldsymbol{a}_{2}, q_{1}, \boldsymbol{q}_{2}\right| \otimes \rho_{T \hat{T}^{n_{1}}(S \hat{S})^{n_{2}}}\left(a_{1}, \boldsymbol{a}_{2}, q_{1}, \boldsymbol{q}_{2}\right),
$$

where $\boldsymbol{a}_{2}$ and $\boldsymbol{q}_{2}$ are the $n_{1}$-dimensional vectors of indices for $A_{2}^{n_{1}}$ and $Q_{2}^{n_{1}}$ respectively, and $\rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$ is the marginal state of $\rho_{T \hat{T}^{n_{1}}(S \hat{S})^{n_{2}}}\left(a_{1}, \boldsymbol{a}_{2}, q_{1}, \boldsymbol{q}_{2}\right) .^{3}$ Comparing EqS. (7.34) and (7.35) allows us to derive the relation between $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ and the variables $\rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$, which is

$$
\begin{equation*}
p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=\frac{T^{2}}{\pi\left(q_{1}, q_{2}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right] . \tag{7.36}
\end{equation*}
$$

[^36]The marginal probabilities for Alice are given by

$$
\begin{aligned}
p\left(a_{1} \mid q_{1}\right) & =\sum_{a_{2}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=\frac{T^{2}}{\pi\left(q_{1}, q_{2}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\sum_{a_{2}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right)\right] \\
& =\frac{T}{\pi_{1}\left(q_{1}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\mathbb{1}_{\hat{T}} \otimes \rho_{T S \hat{S}}\left(a_{1}, q_{1}\right)\right)\right],
\end{aligned}
$$

where we used the linear constraints of the program and the fact that we consider free games in the second line. Similarly, the marginal probability distribution for Bob is given by

$$
p\left(a_{2} \mid q_{2}\right)=\frac{T}{\pi_{2}\left(q_{2}\right)} \operatorname{tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\mathbb{1}_{T} \otimes \rho_{\hat{T} S \hat{S}}\left(a_{2}, q_{2}\right)\right)\right]
$$

Then, using these joint and marginal probabilities, we can construct the first level NPA matrix $\Gamma_{1}$, which has the form of Eq. (6.14). The sole entries of $\Gamma_{1}$ which we cannot relate to the optimisation variable are those in the sub-matrices $Q$ and $R$ with different inputs $q$; these entries become new variables to the SDP relaxation. To impose the NPA constraint on our $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$, we just simply add the positive semidefinite condition $\Gamma_{1} \succeq 0$, where $\Gamma_{1}$ is expressed in terms of the optimisation variables.

For any $k \geq 2$, the $k$-th level NPA matrix involves some entries of the form

$$
p\left(\boldsymbol{a}_{1}, \boldsymbol{a}_{2} \mid \boldsymbol{q}_{1}, \boldsymbol{q}_{2}\right)=\operatorname{Tr}\left[\left(\prod_{i=1}^{m} E\left(\boldsymbol{a}_{1}(i), \boldsymbol{q}_{1}(i)\right)\right)\left(\prod_{j=1}^{\ell} E\left(\boldsymbol{a}_{2}(j), \boldsymbol{q}_{2}(j)\right)\right) \rho_{A B}\right]
$$

with $m+\ell \leq 2 k$, where $\rho_{A B}$ is the shared state between the players, and $\left\{E\left(a_{1}, q_{1}\right)\right\}_{a_{1}, q_{1}}$ and $\left\{E\left(a_{2}, q_{2}\right)\right\}_{a_{2}, q_{2}}$ are the measurement operators for each player respectively (we followed the notation in Section 6.3.2 in this equation). $\left|\boldsymbol{a}_{1}\right|=\left|\boldsymbol{q}_{1}\right|=m$ and $\left|\boldsymbol{a}_{2}\right|=\left|\boldsymbol{q}_{2}\right|=\ell$, and $\boldsymbol{a}_{i}(j)$ $\left(\boldsymbol{q}_{i}(j)\right)$ is the $j$-th entry of the vector $\boldsymbol{a}_{i}\left(\boldsymbol{q}_{i}\right)$. Among these entries, the ones that we can express in terms of the optimisation variable $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$ are those with $m=1$ (since we are not extending the sub-system $\left.A_{1} Q_{1} T\right)$ and $\ell \leq n_{1}$. The relation between these entries is that
for $\left|\boldsymbol{a}_{2}\right|=\left|\boldsymbol{q}_{2}\right|=\ell$
$p\left(a_{1}, \boldsymbol{a}_{2} \mid q_{1}, \boldsymbol{q}_{2}\right)$

$$
\begin{equation*}
=\frac{T^{2 \ell}}{\pi_{1}\left(q_{1}\right) \prod_{j=1}^{\ell} \pi_{2}\left(\boldsymbol{q}_{2}(j)\right)} \operatorname{Tr}\left[\left(P_{\text {cyclic }}^{\ell+1}\right)^{\mathrm{T}} \mathrm{~T}_{S \hat{S}}\left(\left(\bigotimes_{i=2}^{\ell} \mathbb{1}_{T_{i}}\right) \otimes \rho_{T \hat{T}^{\ell} S \hat{S}}\left(a_{1}, \boldsymbol{a}_{2}, q_{1}, \boldsymbol{q}_{2}\right)\right)\right] \quad \forall \ell: \ell \leq k \tag{7.37}
\end{equation*}
$$

where $P_{\text {cyclic }}^{n}$ is a unitary operator acting on $n$ copies of the subsystem $T \hat{T}$, implementing a cyclic permutation over these copies described by the following action over tuples, $(1,2,3, \ldots, n) \rightarrow$ $(2,3, \ldots, n, 1)$. The above equation is explicitly derived in Appendix C.4. Then, to impose the $k$-th level NPA constraint, we can simply add the positive semidefinite condition $\Gamma_{k} \succeq 0$ to the SDP relaxation.

### 7.5 Comparison with previous work

In the last section, we introduced a procedure to construct the NPA matrix with the optimisation variable of $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ so that the corresponding NPA constraint can be added to $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$. We denote these new SDP relaxations with additional $k$-th level NPA constraint by $\operatorname{sdp}_{n_{1}, n_{2}}^{\mathrm{NPA}_{k}}(V, \pi, T)$; namely, $\operatorname{sdp}_{n_{1}, n_{2}}^{\mathrm{NPA}_{k}}(V, \pi, T)$ is $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ in Eq. (7.7) with an additional NPA constraint of the form $\Gamma_{k} \succeq 0$. In this section, we investigate the performance of these combined SDP relaxations $\operatorname{sdp}_{n_{1}, n_{2}}^{\mathrm{NPA}_{k}}(V, \pi, T)$ and compare it with previous work in the literature.

What we want is that the combined SDP relaxations $\operatorname{sdp}_{n_{1}, n_{2}}^{\mathrm{NPA}_{k}}(V, \pi, T)$ can provide tighter bounds for dimension-constrained settings compared to the ones we can get only from the dimension-agnostic NPA hierarchy. Firstly, we remark that the following relation holds

$$
\operatorname{sdp}_{1,1}(V, \pi, T)=\omega_{\mathrm{NS}}(V, \pi)
$$

where $\omega_{\mathrm{NS}}(V, \pi)$ is the optimal winning probability of a two-player game $(V, \pi)$ with the assistance of general no-signalling correlations satisfying Eq. (6.16) - the so-called no-signalling value. The

Table 7.1: The rule matrix $V_{\text {ex }}$ with $\operatorname{sdp}_{2,1}^{\mathrm{NPA}_{1}}\left(V_{\text {ex }}, \pi_{\text {uni }}, 2\right)<\mathrm{NPA}_{1}\left(V_{\text {ex }}, \pi_{\text {uni }}\right)$, where $\mathrm{NPA}_{1}$ denotes the value given by the first level of the NPA hierarchy. Here, $\pi_{\text {uni }}$ is the uniform probability distribution, $|A|=3$, and $|Q|=2$. The table only shows the winning answers for each question set, and all the other answers lose.

|  | $\mathbf{q}_{\mathbf{1}}=\mathbf{0}, \mathbf{q}_{\mathbf{2}}=\mathbf{0}$ | $\mathbf{q}_{\mathbf{1}}=\mathbf{0}, \mathbf{q}_{\mathbf{2}}=\mathbf{1}$ | $\mathbf{q}_{\mathbf{1}}=\mathbf{1}, \mathbf{q}_{\mathbf{2}}=\mathbf{0}$ | $\mathbf{q}_{\mathbf{1}}=\mathbf{1}, \mathbf{q}_{\mathbf{2}}=\mathbf{1}$ |
| :--- | :---: | :---: | :---: | :---: |
|  | $a_{1}=0, a_{2}=1$ | $a_{1}=0, a_{2}=2$ | $a_{1}=0, a_{2}=0$ | $a_{1}=1, a_{2}=2$ |
| Winning | $a_{1}=1, a_{2}=2$ | $a_{1}=1, a_{2}=0$ | $a_{1}=0, a_{2}=2$ |  |
| answers | $a_{1}=2, a_{2}=1$ |  | $a_{1}=1, a_{2}=2$ |  |
|  | $a_{1}=2, a_{2}=2$ |  | $a_{1}=2, a_{2}=1$ |  |

proof is provided in Appendix C.5. As the NPA hierarchy already has no-signalling conditions built-in, we need to combine a higher level of $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ with NPA constraints to see any improvements over the plain NPA hierarchy. For the second-level relaxation $\operatorname{sdp}_{2,1}^{\mathrm{NPA}_{1}}(V, \pi, T)$, which is $\operatorname{sdp}_{2,1}(V, \pi, T)$ combined with the first-level NPA constraint, we give in Table 7.1 a rule matrix $V_{\text {ex }}$ which improves for $2 \times 2$-dimensional quantum assistance on the plain first NPA level. That is, we find that

$$
\operatorname{sdp}_{2,1}^{\mathrm{NPA}_{1}}\left(V_{\mathrm{ex}}, \pi_{\mathrm{uni}}, 2\right)=0.79888<0.80157=\mathrm{NPA}_{1}\left(V_{\mathrm{ex}}, \pi_{\mathrm{uni}}\right)
$$

where $\mathrm{NPA}_{1}\left(V_{\text {ex }}, \pi_{\text {uni }}\right)$ denotes the value given by the first-level of the NPA hierarchy. The computations were carried out in Python using the SDP solver MOSEK [117], with at least 5 digits accuracy as given from the value obtained from the primal versus the corresponding dual program.

We give in Appendix C.6.1 an adapted SDP hierarchy $\overline{\operatorname{sdp}}_{n_{1, n}}^{\mathrm{proj}}(V, \pi, T)$ which is valid under the assumption of projective rank-one measurements for the special case $|A|=|T|=2$. These SDPs have smaller program sizes compared to the original $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ and are thus advantageous for the implementation of higher levels. In particular, for $|A|=|T|=2$, the optimal measurements are necessarily projective rank-one, and we find that for $|Q|$ questions on each side distributed uniformly according to $\pi_{\text {uni }}$

$$
\overline{\operatorname{sdp}}_{n_{1}=n_{2}=|Q|}^{\operatorname{proj}}\left(V, \pi_{\text {uni }}, 2\right) \leq \frac{1}{|Q|^{2}} \cdot \operatorname{sdp}_{\mathrm{PPT}}(V, T)
$$

where $\operatorname{sdp}_{\mathrm{PPT}}(V, T)$ denotes the PPT type SDP relaxations previously given in [108, Eq. (6)].

Hence, for such settings, our SDP relaxations can be understood as adding more constraints to the relaxations given in [108]. In the same spirit, we discuss in Appendix C.6.2 another adapted SDP hierarchy $\operatorname{sdp}_{n_{1}, n_{2}}^{\text {proj }}(V, \pi, 2)$ which improves the asymptotic scaling of the hierarchy in [108] from exponential to polynomial in $|Q|$ for additive $\epsilon$-approximations.

As an example, for the $I_{3322}$ Bell inequality with the (non-binary) rule matrix $V_{3322}$ in the form of $[118,119],{ }^{4}$ we can reproduce the dimension witness

$$
\omega_{Q(2)}\left(V_{3322}, \pi_{\mathrm{uni}}\right) \leq \overline{\operatorname{sdp}}_{|Q|,|Q|}^{\mathrm{proj}}\left(V_{3322}, \pi_{\mathrm{uni}}, 2\right) \leq 0.25
$$

Overall, we note that adding the projective rank-one measurement assumption is typically very useful for practical performance. Going beyond that, numerical tests for low-level relaxations reveal that it seems challenging to compete with the further methods from [109, 110] that were designed with practical purposes in mind but lack analytical bounds on the convergence speed.

### 7.6 Summary

In this chapter, we explored the characterisation of quantum correlations with dimension constraints. More specifically, we gave a converging hierarchy of SDP relaxations for the set of quantum correlations with fixed dimension and provided analytical bounds on the convergence speed by means of multipartite quantum de Finetti theorems with linear constraints. In particular, we showed that one can compute additive $\epsilon$-approximations on the value of a twoplayer free game with $T \times T$-dimensional quantum correlations with a program of size that scales polynomially in the number of questions $|Q|$ and quasi-polynomially in the number of answers $|A|$. This improves on previously known approximation algorithms for which worst-case run-time guarantees are at best exponential in $|Q|$ and $|A|$.

In the main text, we always considered free games where the given probability distribution of the questions for Alice and Bob is not correlated, i.e., $\pi\left(q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi\left(q_{2}\right)$. However, we can also derive lower bounds on the computational complexity for general games using the

[^37]same procedure presented in this chapter. The key difference is to absorb $\pi\left(q_{1}, q_{2}\right)$ into the rule matrix $V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$ instead of $E_{A_{1} Q_{1} T}$ and $D_{A_{2} Q_{2} \hat{T}}$ when we convert the problem into a quantum separability problem. We refer to Appendix C. 1 for the derivation, where we find that, in contrast to the case of free games, the scaling becomes exponential in terms of $|Q|$.

Our approach can be applied to other problems of interest. The most obvious example would be the multipartite extension of two-player games. In addition, as pointed out in [116], finding the maximum success probability for transmitting a message under a given noisy channel can be formalised as a quantum separability problem. The techniques presented in this chapter can be applied to the multipartite generalisation of this quantum error-correction problem.

Compared to a given level of the NPA hierarchy, the corresponding level of $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ has the optimisation variable with larger size and more optimisation constraints. One way to improve this aspect is to exploit the symmetry embedded in the semidefinite program. Our SDP relaxations have a few symmetries: (i) $\Phi_{T \hat{T} \mid S \hat{S}}$ in the objective function is invariant under any local unitary transformation, and (ii) depending on the game, the rule matrix $V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$ may be invariant under some group actions. Employing the symmetry-reduction technique like [120] can potentially reduce the program size and improve the performance of $\operatorname{sdp}_{n_{1}, n_{2}}^{\mathrm{NPA}}(V, \pi, T)$ in numerical implementations.

## Chapter 8

## Conclusions

In this thesis, we explored what we can achieve with currently available tools for quantum information science. In Part I, we investigated the capability of Gaussian states and operations, which are often regarded as easily accessible resources, for two particular quantum tasks: resource distillations and quantum thermodynamical processes with catalysts. In Part II, we changed the topic and developed a classical algorithm in the form of semidefinite programming hierarchy for characterising quantum correlations of fixed dimension. In this section, we give overall discussions and outlooks for each of these two parts.

In Part I, we presented two novel results within the topic of Gaussian resource theories. In Chapter 4, we examined whether classical randomness and conditional Gaussian operations can enable resource distillations in Gaussian platforms, which is known to be impossible only using Gaussian states and operations. We found that some resource distillations indeed become possible when using conditional Gaussian operations but in a limited fashion - the amount of distillable resource is restricted. In Chapter 5, we changed the topic to the Gaussian resource theory of thermodynamics and characterised the possible state transformations under the allowed operations (Gaussian thermal operations) when catalysts are allowed. From the characterisation, we found that catalysts are not much helpful for this particular setting, and possible thermodynamical processes realisable by Gaussian thermal operations are limited even with assistance of resourceful catalysts. In both results, we consistently encountered that the direct-sum structure of multi-mode Gaussian systems leads to the limitations of Gaussian
platforms: in Chapter 4, the tensorisation property of the Gaussian resource monotone, which was originated from the direct-sum composition of covariance matrices, restricted our ability to distill more resources even when starting with multiple input states; in Chapter 5 , the direct-sum structure of Gaussian systems prevented catalysts from having a dramatic impact on the ordering relations of thermodynamical resources. It would be an interesting research direction to see which other limitations of Gaussian platforms also stem from the direct-sum structure and to formalise these general limitations of Gaussian systems in a comprehensive theory. Also, there is another well-known class of quantum states and operations whose dynamics can be described by the direct-sum structure: stabilizer states and Clifford operations. It would be an interesting question to answer whether these limitations in Gaussian platforms due to the direct-sum structure also hold for stabilizer states and Clifford operations.

In Part II, we developed a new hierarchy of semidefinite programming relaxations for twoplayer non-local games and derived the analytical asymptotic convergence rate of the developed hierarchy, which led to an upper bound on the computational complexity of calculating dimensionbounded quantum values for two-player non-local games. We also showed how one can combine our hierarchy with the NPA hierarchy to get a better result; this is particularly interesting as we found a use-case where one can combine the DPS hierarchy and the NPA hierarchy, two very different SDP hierarchies widely used in quantum information science. Despite having quasi-polynomial scaling in the size of the game for free non-local games, our new hierarchy could not prove itself being practical in numerical implementations. Thus, an obvious direction of future research is to improve the numerical performance of our hierarchy so that it can give tighter upper bounds for non-local games of interest than the currently best-known bounds calculated by other methods [108-110]. As we already mentioned in the last section, one possible way is to exploit symmetries embedded in the problem to reduce the size of the SDPs by restricting the search to the set invariant under those symmetries. This method has been already proven to be effective in other settings [121-123]. Also, as it is in general more effective to use the bipartite DPS hierarchy, it would be interesting to see whether we can formulate the problem in a way that the optimisation variables are described by bipartite separable states rather than tripartite ones.

## Appendix A

## Additional information for convex

## Gaussian resource theories

In this appendix we provide some additional information for discussions in Chapter 2.

## A. 1 Partial selective homodyne measurement

In this appendix, we prove that the partial selective homodyne measurement introduced within Definition 4 is an allowed operation for the Gaussian resource theories of squeezing and entanglement. The operation, defined as

$$
\begin{aligned}
M_{B}^{q}: & \mathcal{H}_{A} \otimes \mathcal{H}_{B} \\
& \rightarrow \mathcal{H}_{A} \\
\rho_{A B} \longmapsto & \operatorname{Tr}_{B}\left[\left(\mathbb{1}_{A} \otimes|q\rangle\left\langle\left. q\right|_{B}\right) \rho_{A B}\right],\right.
\end{aligned}
$$

consists of applying the partial projector on the eigenvector $|q\rangle$ of a quadrature operator $\hat{q}$ and tracing out the projected mode. Notice that the subsystems $A$ and $B$ can both be composed by multiple modes. Even though the quadrature eigenvector $|q\rangle$ is an infinitely squeezed state, one can show that this does not increase the amount of squeezing (or entanglement) in the state. Let us first recall how the partial selective homodyne measurement acts on the covariance matrix of a state.

Lemma 30. [66, 67] Let $\sigma$ be the covariance matrix of an state describing a $(n+m)$-mode system, which can be written as

$$
\sigma=\left(\begin{array}{cc}
A & B \\
B^{T} & C
\end{array}\right)
$$

where $A, B$, and $C$ are real matrices with size $2 n \times 2 n, 2 n \times 2 m$, and $2 m \times 2 m$ respectively. The covariance matrix $\sigma^{\prime}$ of the state after a selective homodyne measurement on the last modes is then given by

$$
\begin{equation*}
\sigma^{\prime}=A-B(P C P)^{M P} B^{T} \tag{A.1}
\end{equation*}
$$

where $P=\mathbb{1}_{m} \oplus \mathbf{0}_{m}$ is a projection operator, and MP denotes the Moore Penrose pseudo-inverse.

Using this lemma, we can show that partial selective homodyne measurement is an allowed operation for both the Gaussian resource theories of squeezing and entanglement.

Proposition 31. Partial selective homodyne measurements on Gaussian states are an allowed operation for the Gaussian theory of squeezing. That is, we have for all $\rho_{A B} \in \mathcal{F}_{\mathrm{sq}}^{G}$ that $M_{B}^{q}\left(\rho_{A B}\right) \in \mathcal{F}_{\mathrm{sq}}^{G}$.

Proof. Instead of considering the set of free states $\mathcal{F}_{\mathrm{sq}}^{G}$, we focus on the set of free covariance matrices, characterised by $\boldsymbol{F}_{\mathrm{sq}}^{G}=\{\sigma \mid \sigma \geq \mathbb{1}\}$ for the theory of squeezing. Then, for any $(n+m)$-mode free covariance matrix $\sigma \in \boldsymbol{F}_{\mathrm{sq}}^{G}$, we can write

$$
\sigma=\left(\begin{array}{cc}
A & B \\
B^{T} & C
\end{array}\right) \geq\left(\begin{array}{cc}
\mathbb{1}_{2 n} & \mathbf{0} \\
\mathbf{0} & \mathbb{1}_{2 m}
\end{array}\right) \geq\left(\begin{array}{cc}
\mathbb{1}_{2 n} & \mathbf{0} \\
\mathbf{0} & \mathbf{0}_{2 m}
\end{array}\right)
$$

The last inequality imply that the following matrix

$$
W=\left(\begin{array}{cc}
A-\mathbb{1}_{2 n} & B \\
B^{T} & C
\end{array}\right)
$$

is positive semidefinite. We can now construct a different matrix by applying the operator
$\mathbb{1}_{2 n} \oplus P$ to $W$ by congruence, where $P=\mathbb{1}_{m} \oplus \mathbf{0}_{m}$.

$$
V=\left(\mathbb{1}_{2 n} \oplus P\right) W\left(\mathbb{1}_{2 n} \oplus P\right)^{\mathrm{T}}=\left(\begin{array}{cc}
A-\mathbb{1}_{2 n} & B P \\
P B^{T} & P C P
\end{array}\right)
$$

Clearly, this matrix is positive semidefinite as well. Then, its Schur complement must be positive semidefinite,

$$
\left(A-\mathbb{1}_{2 n}\right)-B P(P C P)^{M P} P B^{T} \geq 0
$$

Furthermore, one can easily show that $P(P C P)^{M P} P$ satisfies all four conditions to be the Moore Penrose pseudo-inverse of $P C P$ [104]. Due to the uniqueness of the pseudo-inverse, this implies that $P(P C P)^{M P} P=(P C P)^{M P}$. Then, we conclude that

$$
A-B(P C P)^{M P} B^{T} \geq \mathbb{1}_{2 n}
$$

In Lemma 30, we have seen that this is in fact the resulting covariance matrix after a partial selective homodyne measurement is performed on a Gaussian state with covariance matrix $\sigma$. This closes the proof.

The case of entanglement is also not difficult to deal with since the partial selective homodyne measurement needs to be a local operation. Any separable Gaussian state has a covariance matrix $\sigma_{A B}=\sigma_{A} \oplus \sigma_{B}$, where $\sigma_{A}\left(\sigma_{B}\right)$ is the covariance matrix of the state owned by Alice (Bob). Then, if we act with $M^{q}$ on part of the subsystem of Bob, we obtain a final covariance matrix $\sigma_{A B}^{\prime}=\sigma_{A}^{\prime} \oplus \sigma_{B}^{\prime}$, where $\sigma_{B}^{\prime}$ has the form given in Eq. (A.1), and similarly for Alice. Clearly, the output covariance matrix is that of a separable state.

## A. 2 Monotones for squeezing

In Chapter 4, we use, together with the fidelity to the target state, the variance with respect to a specific quadrature as a measure for squeezing. In this appendix, we aim to justify its use
over convex mixtures of Gaussian states. Let us first introduce the measure

$$
\begin{equation*}
M_{\mathrm{var}}(\rho)=\min _{\hat{q}} \operatorname{Var}[\hat{q}]_{\rho}, \tag{A.2}
\end{equation*}
$$

where the minimization is performed over all quadrature operators $\hat{q}$, and the variance is defined as

$$
\operatorname{Var}[\hat{q}]_{\rho}=\left\langle\hat{q}^{2}\right\rangle_{\rho}-\langle\hat{q}\rangle_{\rho}^{2},
$$

where $\langle\hat{A}\rangle_{\rho}=\operatorname{Tr}[\hat{A} \rho]$. It is worth noting that the above measure coincide with the minimum eigenvalue of the covariance matrix of the state $\rho$, a squeezing measure for pure Gaussian states [124]. The above measure is monotone under the allowed operations of the Gaussian resource theory of pure squeezing, denoted by $\mathbb{O}_{\mathrm{sq}}^{G}$, which are composed by the following fundamental operations:

1. Appending modes in the vacuum state
2. Performing passive operations (beam splitter, phase shift, and displacements)
3. Trace out a subset of the system's modes.

When the state space is extended to the set of all Gaussian states including mixed states, the measure in Eq. (A.2) can be modified as

$$
\begin{equation*}
\bar{M}_{\mathrm{var}}(\rho)=\min \left\{1, M_{\mathrm{var}}(\rho)\right\} . \tag{A.3}
\end{equation*}
$$

Notice that we have to modify the measure since $M_{\text {var }}$ is not monotonic over states whose covariance matrix has all eigenvalues higher than 1, for instance thermal states. Indeed, one could always decrease the measure by replacing these states with the vacuum. To avoid this problem, the minimization in Eq. (A.3) has been added. Finally, it is worth noting that in the main text we do not explicitly consider $\bar{M}_{\text {var }}$, but rather we focus on the relevant quadrature for the problem.

From the examples in Sec. 4.4 of the main text, it is clear that $\bar{M}_{\text {var }}$ is not a monotone under the allowed operations in Def. 4 (since it can be decreased under allowed operations). However, we would like to argue that this measure is still a meaningful one over the set of convex mixtures of Gaussian states. To do so we take a different notion of allowed operations, that are still relevant for the state space we are considering, but do not allow for conditional operations. ${ }^{1}$ The operations we consider here are mixtures of allowed operations for squeezing theory. The most general map in this set can be written as

$$
\begin{equation*}
\Gamma(\rho)=\int \mathrm{d} \lambda p(\lambda) \Phi_{\lambda}(\rho), \tag{A.4}
\end{equation*}
$$

where $p(\lambda)$ is a probability distribution, and $\Phi_{\lambda} \in \mathbb{O}_{\mathrm{sq}}^{G}$.
Let us first show that the measure $\bar{M}_{\text {var }}$ is monotone under $\mathbb{O}_{\mathrm{sq}}^{G}$ alone, when the state space is given by the convex hull of Gaussian states. If a mixture of Gaussian states $\rho_{\text {mix }}$ is such that $M_{\text {var }}\left(\rho_{\text {mix }}\right)<1$, then it is easy to see that neither (i) appending ancillary systems in the vacuum state nor (ii) performing passive operations can modify this value. Indeed, passive operations are represented by orthogonal matrices acting over the covariance matrix by congruence (irrespectively of whether the state is Gaussian or not) thus preserving the eigenvalues of the matrix. Furthermore, (iii) partial tracing the system corresponds to selecting a principal sub-matrix of the covariance matrix and discarding the rest. It is known [82] that the smallest eigenvalue of the sub-matrix cannot be lower than the original matrix. Thus, when $M_{\text {var }}\left(\rho_{\text {mix }}\right)<1$, we have that

$$
\begin{equation*}
M_{\mathrm{var}}\left(\Phi\left(\rho_{\text {mix }}\right)\right) \geq M_{\text {var }}\left(\rho_{\text {mix }}\right), \tag{A.5}
\end{equation*}
$$

where $\Phi \in \mathbb{O}_{\mathrm{sq}}^{G}$. When $M_{\mathrm{var}}\left(\rho_{\text {mix }}\right) \geq 1$, instead, it is easy to see that only operation (i) can decrease the variance, since the vacuum's covariance matrix is $\mathbb{1}$. However, the variance cannot be reduced further than the unit, so that $\bar{M}_{\text {var }}$ is monotonic under the allowed operations $\mathbb{O}_{\text {sq }}^{G}$ over mixtures of Gaussian states.

Monotonicity of this measure under maps of the form given in Eq. (A.4) follows from the

[^38]concavity of the variance. When the output state of the map has a minimum variance that is smaller than one, we have that
\[

$$
\begin{aligned}
\bar{M}_{\text {var }}\left(\Gamma\left(\rho_{\text {mix }}\right)\right) & =\operatorname{Var}[\bar{Q}]_{\Gamma\left(\rho_{\text {mix }}\right)} \\
& \geq \int \mathrm{d} \lambda p(\lambda) \operatorname{Var}[\bar{Q}]_{\Phi_{\lambda}\left(\rho_{\text {mix }}\right)} \\
& \geq \int \mathrm{d} \lambda p(\lambda) \bar{M}_{\text {var }}\left(\Phi_{\lambda}\left(\rho_{\text {mix }}\right)\right) \\
& \geq \int \mathrm{d} \lambda p(\lambda) \bar{M}_{\text {var }}\left(\rho_{\text {mix }}\right)=\bar{M}_{\text {var }}\left(\rho_{\text {mix }}\right),
\end{aligned}
$$
\]

where $\bar{Q}$ is the quadrature operator minimising the variance of $\Gamma\left(\rho_{\text {mix }}\right)$, and the first inequality follows from the concavity of the variance and by the definition of $\Gamma$. The second inequality follows from the fact that $\bar{M}_{\text {var }}$ involves a minimisation over all quadrature operators (plus a cut-off for values higher than the unit), and the last inequality follows from Eq. (A.5). When the minimum variance of the output state is already higher than unit, the monotonicity relation is trivially true. Thus, variance is a good measure of squeezing over convex mixtures of Gaussian states, although it is not monotonic when the allowed operations include conditional maps.

## Appendix B

## Additional information for catalytic Gaussian thermal operations

## B. 1 Proof of Theorem 12

In this section, we present the proof of Theorem 12, which describes the necessary conditions for state transformations under approximate strong catalytic GTOs when the catalyst and the system involve multiple modes.

Firstly, we provide two intermediate results which will be used in the main proof.

Lemma 32. Let $\boldsymbol{z}, \boldsymbol{z}^{\prime}$ and $\tilde{\boldsymbol{z}}$ be non-increasingly ordered lists of length $l$ such that

$$
\tilde{\boldsymbol{z}} \leq \boldsymbol{z} \quad \text { and } \quad \sum_{i=1}^{l}\left|\tilde{z}_{i}-z_{i}^{\prime}\right| \leq \delta
$$

Then

$$
\sum_{i=1}^{l}\left[z_{i}^{\prime}-z_{i}\right]^{+} \leq \delta .
$$

Again, $[x]^{+}:=\max \{x, 0\}$ denotes the positive part of a real number.

Proof.

$$
\begin{aligned}
\sum_{i=1}^{l}\left[z_{i}^{\prime}-z_{i}\right]^{+} & =\sum_{i: z_{i}^{\prime}>z_{i}} z_{i}^{\prime}-z_{i} \\
& \leq \sum_{i: z_{i}^{\prime}>z_{i}} z_{i}^{\prime}-\tilde{z}_{i} \\
& \leq \sum_{i: z_{i}^{\prime}>z_{i}}\left|z_{i}^{\prime}-\tilde{z}_{i}\right| \\
& \leq \sum_{i=1}^{l}\left|z_{i}^{\prime}-\tilde{z}_{i}\right|
\end{aligned}
$$

$$
\leq \delta
$$

Theorem 33. Let $\boldsymbol{x}=\left(x_{1}, \ldots, x_{n}\right), \boldsymbol{x}^{\prime}=\left(x_{1}^{\prime}, \ldots, x_{n}^{\prime}\right)$ and $\boldsymbol{y}=\left(y_{1}, \ldots, y_{m}\right)$ be non-increasing lists of real numbers, and define the composite ordered lists $\boldsymbol{z}=(\boldsymbol{x}, \boldsymbol{y})^{\downarrow}, \boldsymbol{z}^{\prime}=\left(\boldsymbol{x}^{\prime}, \boldsymbol{y}\right)^{\downarrow}$. Suppose that there exists $\tilde{\boldsymbol{z}}$ such that

$$
\tilde{\boldsymbol{z}} \leq \boldsymbol{z} \quad \text { and } \quad \sum_{i=1}^{n+m}\left|\tilde{z}_{i}-z_{i}^{\prime}\right| \leq \delta
$$

Then

$$
\sum_{i=1}^{n}\left[x_{i}^{\prime}-x_{i}\right]^{+} \leq \delta
$$

Proof. The main useful idea is to construct a partition of the indices of $\boldsymbol{z}, \boldsymbol{z}^{\prime}$ into contiguous blocks $B_{1}, B_{2}, \ldots$ such that in each block, the $x_{i}$ values either all increase or all decrease when going from $\boldsymbol{z}$ to $\boldsymbol{z}^{\prime}-$ see Fig. B. 1 for an illustration. This will let us separate out those values that break the monotonicity condition and bound how much they do so. Let $p\left(x_{i}\right)$ denote the index position of $x_{i}$ in $\boldsymbol{z}$, and similarly $p^{\prime}\left(x_{i}^{\prime}\right)$ for $x_{i}^{\prime}$ in $\boldsymbol{z}^{\prime}$. Note that a unique designation of indices is possible if we adopt the convention that, if $y_{j}=x_{i}$, then $y_{j}$ appears before $x_{i}$ in the list $\boldsymbol{z}$.

The end result of all this is that the set of indices $i$ of all $x_{i}$ in $B_{b}$ and $x_{i}^{\prime}$ in $B_{b}$ are the


Figure B.1: Illustration of the block partitioning of non-increasingly ordered lists $\boldsymbol{z}=(\boldsymbol{x}, \boldsymbol{y})^{\downarrow}$, $\boldsymbol{z}^{\prime}=\left(\boldsymbol{x}^{\prime}, \boldsymbol{y}\right)^{\downarrow}$. In this case, $B_{1}$ is an increasing block, $B_{2}$ is decreasing, and so on alternately. Therefore $x_{1}^{\prime}>x_{1}, x_{2}^{\prime}>x_{2}, x_{3}^{\prime} \leq x_{3}$ and $x_{4}^{\prime} \leq x_{4}$.
same for each block $B_{b}$. Within an "increasing" block, $x_{i}^{\prime}>x_{i}$ and within a 'decreasing' block, $x_{i}^{\prime} \leq x_{i} \forall i$. This claim will be proved separately below by showing how to construct the blocks. Using this, it is easy to see that the set of $y_{j}$ is also the same in each corresponding block. In other words, to send $\boldsymbol{z} \rightarrow \boldsymbol{z}^{\prime}$, in each block we apply a permutation of the elements and replace $x_{i} \rightarrow x_{i}^{\prime}$.

Since the $y_{j}$ values match up in each block $B_{b}$, we have

$$
\sum_{k \in B_{b}} z_{k}^{\prime}-z_{k}=\sum_{i: p\left(x_{i}\right) \in B_{b}} x_{i}^{\prime}-x_{i} .
$$

Summing over all increasing blocks and using Lemma 32,

$$
\begin{aligned}
\sum_{i=1}^{n}\left[x_{i}^{\prime}-x_{i}\right]^{+} & =\sum_{B_{b} \text { increasing }} \sum_{i: p\left(x_{i}\right) \in B_{b}} x_{i}^{\prime}-x_{i} \\
& =\sum_{B_{b} \text { increasing }} \sum_{k \in B_{b}} z_{k}^{\prime}-z_{k} \\
& \leq \sum_{k=1}^{n+m}\left[z_{k}^{\prime}-z_{k}\right]^{+} \\
& \leq \delta .
\end{aligned}
$$

Construction of the block partition:
Starting from the beginning of the list, suppose that $x_{1}^{\prime}>x_{1}$, and find the lowest $j$ such that
$x_{j}^{\prime} \leq x_{j}$. The list $\left\{1,2, \ldots, p\left(x_{j-1}\right)\right\}$ is then called the first block $B_{1}-$ this is an "increasing" block, such that $x_{i}^{\prime}>x_{i} \forall p\left(x_{i}\right) \in B_{1}$. If instead $x_{1}^{\prime} \leq x_{1}$, then we find the lowest $j$ such that $x_{j}^{\prime}>x_{j}$ and $B_{1}:=\left\{1,2, \ldots, p^{\prime}\left(x_{j-1}^{\prime}\right)\right\}$ is a 'decreasing' block. Moving along the list and repeating the process, we can then partition $\boldsymbol{z}$ into blocks $B_{1}, B_{2}, \ldots$ which are alternately increasing and decreasing.

We can say something about the index positions of the $x_{i}$ according to whether they are increasing or decreasing:

$$
\begin{align*}
& x_{i}^{\prime}>x_{i} \Rightarrow p^{\prime}\left(x_{i}^{\prime}\right) \leq p\left(x_{i}\right),  \tag{B.1}\\
& x_{i}^{\prime} \leq x_{i} \Rightarrow p^{\prime}\left(x_{i}^{\prime}\right) \geq p\left(x_{i}\right) . \tag{B.2}
\end{align*}
$$

To see this: suppose first that $p^{\prime}\left(x_{i}^{\prime}\right)<p\left(x_{i}\right)$, then there exists $y_{j}$ such that $y_{j} \geq x_{i}$ and $x_{i}^{\prime}>y_{j}$ because $x_{i}^{\prime}$ has moved earlier in the list and thus displaced $y_{j}$-therefore $x_{i}^{\prime}>x_{i}$. This proves (B.2) via its contrapositive. The same argument works for the opposite inequality (B.1). If $p^{\prime}\left(x_{i}^{\prime}\right)=p\left(x_{i}\right)$, then either case is possible.

Now our main claim is that the blocks $B_{b}$ have the following property:

$$
\forall B_{b}, \quad p\left(x_{i}\right) \in B_{b} \Leftrightarrow p^{\prime}\left(x_{i}^{\prime}\right) \in B_{b} .
$$

In other words, the blocks simultaneously partition $\boldsymbol{z}$ and $\boldsymbol{z}^{\prime}$ such that all pairs of $x_{i}$ and $x_{i}^{\prime}$ are in the same blocks in their respective lists. We prove this claim by construction.

- If $B_{1}$ is increasing, then for all $x_{i}$ with $p\left(x_{i}\right) \in B_{1}$, we must have $p^{\prime}\left(x_{i}^{\prime}\right) \in B_{1}$ due to (B.1). Also, whenever $p^{\prime}\left(x_{i}^{\prime}\right) \in B_{1}$, we must have $p\left(x_{i}\right) \in B_{1}$, otherwise $p^{\prime}\left(x_{i}^{\prime}\right)$ would be strictly less than $p\left(x_{i}\right)$. Then, by (B.2) $x_{i}^{\prime}>x_{i}$, and there would have been another increasing $x_{i}$ not included in the block. (It is also not possible to have some increasing $x_{i}$ in some later increasing block, for example $B_{3}$, since then there would be some $x_{j}$ in $B_{2}$ such that $x_{j}>x_{i}$ but $x_{i}^{\prime}>x_{j}^{\prime}-$ and this is incompatible with the ordering $j<i$.)
- For the next block $B_{2}$, which is decreasing, by definition all $x_{i}^{\prime}$ are included. We already know from the first step that there cannot be $x_{i}$ in $B_{1}$ such that $x_{i}^{\prime}$ is in $B_{2}$. And by (B.2),
for any $x_{1}^{\prime}$ in $B_{2}, p\left(x_{i}\right) \leq p^{\prime}\left(x_{i}^{\prime}\right)$, so $x_{i}$ cannot be outside of $B_{2}$.
- Iterate the procedure over all remaining blocks.
- If instead $B_{1}$ is decreasing, then the argument proceeds similarly.

Using the above intermediate results, we now prove Theorem 12:

Proof of Theorem 12. First note that the sufficiency of the conditions in Eq. (5.13) is straightforward: for each $\mu_{i}^{\prime} \leq \mu_{i}$, we use the same construction as in Theorem 9, while for any violation of that inequality, we simply do nothing to the corresponding mode. We can likewise prove the sufficiency of the conditions in Eq. (5.14) for $A$ matrices.

The proof of necessity for $A$ is an immediate application of Theorem 33, setting $\boldsymbol{x}=\boldsymbol{\alpha}_{S}$, $\boldsymbol{x}^{\prime}=\boldsymbol{\alpha}_{S}^{\prime}, \boldsymbol{y}=\boldsymbol{\alpha}_{C}$ and $\tilde{\boldsymbol{z}}=\tilde{\boldsymbol{\alpha}}_{S C}$.

For $M$ we have more work to do because the eigenvalues must be divided into positive and negative parts. Unlike the non-catalytic case, it is possible that $\boldsymbol{\mu}^{\prime}$ contains more positive or negative values than $\boldsymbol{\mu}_{S}$. If the number $n_{+}^{\prime}$ of positive $\boldsymbol{\mu}_{S}^{\prime}$ is greater than the number $n_{+}$ of positive $\boldsymbol{\mu}_{S}$, then we set $\boldsymbol{x}=\left(\boldsymbol{\mu}_{S}^{+}, 0, \ldots\right)$, padding with zeroes to a total length of $n_{+}^{\prime}$. We straightforwardly set $\boldsymbol{x}^{\prime}=\boldsymbol{\mu}_{S}^{\prime+}$ and $\boldsymbol{y}=\boldsymbol{\mu}_{C}^{+}$. We also choose $\tilde{\boldsymbol{z}}=\left(\tilde{\boldsymbol{\mu}}_{S C}^{+}, 0, \ldots\right)$, similarly padded to the same length as $\boldsymbol{z}$ - and need to check that this satisfies the assumptions of Theorem 33. Since $M_{S} \oplus M_{C} \xrightarrow{G T O} \tilde{M}_{S C}$, Theorem 9 ensures that the number of positive elements $\tilde{l}_{+}$of $\tilde{\boldsymbol{\mu}}_{S C}$ is no more than the length $l_{+}$of $\boldsymbol{z}$. Thus, we have $\tilde{\boldsymbol{z}} \leq \boldsymbol{z}$. Moreover,

$$
\begin{aligned}
\sum_{i=1}^{l_{+}}\left|\tilde{z}_{i}-z_{i}^{\prime}\right| & =\sum_{i \leq \tilde{l}_{+}}\left|\tilde{z}_{i}-z_{i}^{\prime}\right|+\sum_{i>\tilde{l}_{+}}\left|\tilde{z}_{i}-z_{i}^{\prime}\right| \\
& =\sum_{i \leq \tilde{l}_{+}}\left|\tilde{\mu}_{S C, i}^{+}-\mu_{S C, i}^{\prime+}\right|+\sum_{i>\tilde{l}_{+}} \mu_{S C, i}^{\prime+}-0 \\
& \leq \sum_{i \leq \tilde{l}_{+}}\left|\tilde{\mu}_{S C, i}^{+}-\mu_{S C, i}^{+}\right|+\sum_{i>\tilde{l}_{+}} \mu_{S C, i}^{\prime+}-\tilde{\mu}_{S C, i} \\
& =\sum_{i=1}^{l_{+}}\left|\mu_{S C, i}^{\prime}-\tilde{\mu}_{S C, i}\right|=: \delta_{+},
\end{aligned}
$$

where the third line uses $\tilde{\mu}_{S C, i} \leq 0$ for $i>\tilde{l}_{+}$. So Theorem 33 can now be applied with the distance bound $\delta_{+}$. The same calculation can be done in exactly the same way for the negative values; summing the two results gives the claimed result since $\delta_{+}+\delta_{-} \leq \delta$

## B. 2 Proof of Theorem 13

In this appendix, we provide the proof of Theorem 13, the state transformation conditions for matrix $M$ under weak catalytic GTOs in the general multi-mode case. We first prove the following useful mathematical property of majorisation with an additive catalyst.

Lemma 34. For any $\boldsymbol{y},\left(\boldsymbol{x}^{\prime}, \boldsymbol{y}\right) \prec_{w}(\boldsymbol{x}, \boldsymbol{y}) \Leftrightarrow \boldsymbol{x}^{\prime} \prec_{w} \boldsymbol{x}$ and $\left(\boldsymbol{x}^{\prime}, \boldsymbol{y}\right) \prec(\boldsymbol{x}, \boldsymbol{y}) \Leftrightarrow \boldsymbol{x}^{\prime} \prec \boldsymbol{x}$.

Proof. We use the following equivalent characterisation of majorisation [85, 4.B.3]: $\boldsymbol{x} \prec_{w} \boldsymbol{y}$ if and only if

$$
g_{a}(\boldsymbol{x}) \leq g_{a}(\boldsymbol{y}) \forall a \in \mathbb{R}, \text { where } g_{a}(\boldsymbol{x}):=\sum_{i}\left[x_{i}-a\right]^{+}
$$

For $\boldsymbol{x} \prec \boldsymbol{y}$ we just have to add in the condition $\sum_{i} x_{i}=\sum_{i} y_{i}$. Using additivity of $g_{a}$,

$$
\begin{aligned}
\left(\boldsymbol{x}^{\prime}, \boldsymbol{y}\right) \prec_{w}(\boldsymbol{x}, \boldsymbol{y}) & \Leftrightarrow g_{a}\left(\boldsymbol{x}^{\prime}\right)+g_{a}(\boldsymbol{y}) \leq g_{a}(\boldsymbol{x})+g_{a}(\boldsymbol{y}) \forall a \\
& \Leftrightarrow g_{a}\left(\boldsymbol{x}^{\prime}\right) \leq g_{a}(\boldsymbol{x}) \forall a \\
& \Leftrightarrow \boldsymbol{x}^{\prime} \prec_{w} \boldsymbol{x} .
\end{aligned}
$$

For $\boldsymbol{x} \prec \boldsymbol{x}^{\prime}$, the sum condition is also clearly equivalent.

We now provide the main proof of Theorem 13

Proof of Theorem 13. For part (1): The initial $M$ matrix is of the form $M=M_{S} \oplus M_{C}$, with eigenvalues $\left(\boldsymbol{\mu}_{S}, \boldsymbol{\mu}_{C}\right)$. The final matrix is of the block form

$$
M^{\prime}=U M U^{\dagger}=\left(\begin{array}{cc}
M_{S}^{\prime} & * \\
* & M_{C}
\end{array}\right)
$$

Using local unitary rotations, we can diagonalise the principal blocks to obtain $\left(\boldsymbol{\mu}_{S}^{\prime}, \boldsymbol{\mu}_{C}^{\prime}\right)$ on the diagonals of $M^{\prime}$ (although the off-diagonal blocks need not vanish). The Schur-Horn theorem [82, Theorems 4.3.45, 4.3.48] says that the eigenvalues of a matrix majorise its diagonals in any basis (and conversely that a basis can always be found giving any set of diagonals allowed by this condition). Hence $\left(\boldsymbol{\mu}_{S}^{\prime}, \boldsymbol{\mu}_{C}\right) \prec\left(\boldsymbol{\mu}_{S}, \boldsymbol{\mu}_{C}\right)$ and Lemma 34 gives the claimed condition as necessary.

For the converse, we use the fact [85, 2.B.1] that if $\boldsymbol{x} \prec \boldsymbol{y}$, then $\boldsymbol{x}$ can be obtained from $\boldsymbol{y}$ by a finite number (in fact, at most $n-1$ ) $T$-transforms, namely partial swaps of pairs of modes - represented by matrices of the form

$$
T=t \mathbb{1}+(1-t) Q
$$

where $t \in[0,1]$ and $Q$ swaps two modes. In order to show that any T-transform can be performed, we focus on an arbitrary pair of modes, labelling the eigenvalues without loss of generality as $\mu_{1} \geq \mu_{2}$. This is equivalent to proving that any $\mu_{1}^{\prime} \geq \mu_{2}^{\prime}$ satisfying $\mu_{1}^{\prime} \leq \mu_{1}$ and $\mu_{1}^{\prime}+\mu_{2}^{\prime}=\mu_{1}+\mu_{2}$ can be achieved with a single catalyst mode. This requires a unitary $U$ such that

$$
U\left(\begin{array}{ccc}
\mu_{1} & 0 & 0 \\
0 & \mu_{2} & 0 \\
0 & 0 & \mu_{C}
\end{array}\right) U^{\dagger}=\left(\begin{array}{ccc}
\mu_{1}^{\prime} & 0 & * \\
0 & \mu_{2}^{\prime} & * \\
* & * & \mu_{C}
\end{array}\right)
$$

Choose any $\mu_{C} \in\left[\mu_{2}^{\prime}, \mu_{1}^{\prime}\right]$, which by assumption is also contained in $\left[\mu_{2}, \mu_{1}\right]$. The necessary and sufficient conditions for the existence of the eigenvalues $\mu_{1}^{\prime}, \mu_{2}^{\prime}$ of the two-dimensional upper-left block are [82, Theorem 4.3.21], bearing in mind the ordering $\mu_{1} \geq \mu_{C} \geq \mu_{2}$,

$$
\mu_{1} \geq \mu_{1}^{\prime} \geq \mu_{C} \geq \mu_{2}^{\prime} \geq \mu_{2}
$$

Evidently these are always satisfied under the assumed conditions.

Part (2) uses part (1) after extending the system to include the bath modes. This immediately gives the necessary condition

$$
\begin{equation*}
\boldsymbol{\mu}_{S B}^{\prime} \prec \boldsymbol{\mu}_{S B}=\left(\boldsymbol{\mu}_{S}, \mathbf{0}_{B}\right) . \tag{B.3}
\end{equation*}
$$

Since $\boldsymbol{\mu}_{S}^{\prime}$ are eigenvalues of a principal submatrix of $M_{S B}^{\prime}$, for any $k \leq n_{+}^{\prime}$ we have [82, Corollary 4.3.34] $\sum_{i=1}^{k} \mu_{S, i}^{\prime} \leq \sum_{i=1}^{k} \mu_{S B, i}^{\prime}$. Therefore

$$
\begin{aligned}
\sum_{i=1}^{k} \mu_{S, i}^{\prime+} & =\sum_{i=1}^{k} \mu_{S, i}^{\prime} \\
& \leq \sum_{i=1}^{k} \mu_{S B, i}^{\prime} \\
& \leq \sum_{i=1}^{k} \mu_{S B, i} \\
& \leq \sum_{i=1}^{k}\left[\mu_{S B, i}\right]^{+} \\
& =\sum_{i=1}^{k} \mu_{S, i}^{+}
\end{aligned}
$$

where the second inequality follows from (B.3). Hence $\boldsymbol{\mu}^{+}{ }_{S} \prec_{w} \boldsymbol{\mu}^{+}{ }_{S}$. The corresponding necessary condition for the negative values follows by symmetry.

For sufficiency, we use the following result [85, 2.C.6.a] analogous to that used in part (1) but for weak majorisation: for $\boldsymbol{x}, \boldsymbol{y}$ composed of non-negative elements, $\boldsymbol{x} \prec_{w} \boldsymbol{y}$ if and only if $\boldsymbol{x}$ can be derived from $\boldsymbol{y}$ by a finite number $(\leq n-1)$ of T-transforms, followed by a finite number $(\leq n)$ of L-transforms. The T-transforms were dealt with above; an L-transform can be performed simply by mixing a single mode at a beam splitter with a thermal mode as in Theorem 9. The same construction works for the negative values independently.

## B. 3 Proof of Theorem 14

In this appendix, we prove the necessary condition of matrix $A$ for state transformations under weak catalytic GTOs for the general multi-mode case.

Proof of Theorem 14. We prove the necessary condition including a bath as this is the more general case. As in Theorem 13, we have $A^{\prime}=U A U^{T}$ for some unitary $U$ and the singular values of $A$ are $\left(\boldsymbol{\alpha}_{S}, \mathbf{0}_{B}, \boldsymbol{\alpha}_{C}\right)$. The parameters $\left(\boldsymbol{\alpha}_{S B}^{\prime}, \boldsymbol{\alpha}_{C}\right)$ are just diagonals of $A^{\prime}$. A set of necessary conditions relating the diagonals $d_{i}$ of a complex symmetric matrix to its singular
values $s_{i}$ are [125]

$$
\sum_{i=1}^{k}\left|d_{i}\right| \leq \sum_{i=1}^{k} s_{i} \quad \forall k
$$

This implies $\left(\boldsymbol{\alpha}_{S B}^{\prime}, \boldsymbol{\alpha}_{C}\right) \prec_{w}\left(\boldsymbol{\alpha}_{S}, \mathbf{0}_{B}, \boldsymbol{\alpha}_{C}\right)$ and by Lemma 34 ,

$$
\boldsymbol{\alpha}_{S B}^{\prime} \prec_{w}\left(\boldsymbol{\alpha}_{S}, \mathbf{0}_{B}\right) .
$$

Therefore

$$
\begin{aligned}
\sum_{i=1}^{k} \alpha_{S, i}^{\prime} & \leq \sum_{i=1}^{k} \alpha_{S B, i}^{\prime} \\
& \leq \sum_{i=1}^{k} \alpha_{S B, i} \\
& =\sum_{i=1}^{k} \alpha_{S, i} .
\end{aligned}
$$

Sufficiency is proved without using a bath, using the same statement about weak majorisation used in Theorem 13. It is easy to see that T-transforms can be performed for each mode by just considering real matrix elements, so that $U$ is orthogonal and $A^{\prime}=U A U^{T}$ is real and symmetric. The same statement [82, Theorem 4.3.21] also guarantees that the required sets of eigenvalues can be achieved with purely real matrix elements. Then, L-transforms can be performed as before using a thermal bath mode. Alternatively, a catalyst mode can be used instead with the same unitary as in Eq. (5.11) for the single-mode case. So with no bath, we need $2 n-1$ catalyst modes, otherwise we need $n$ catalyst modes plus $n-1$ bath modes.

## Appendix C

## Additional information for

## characterising quantum correlations

## C. 1 General games

In this appendix, we consider general games when the questions for Alice and Bob are correlated, i.e., $\pi\left(q_{1}, q_{2}\right) \neq \pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$. Recall that the value for a two-player game with quantum assistance of dimension $T \times T$ can be written as

$$
\omega_{Q(T)}=\max _{E, D, \rho} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\rho_{T \hat{T}}\left(E_{T}\left(a_{1} \mid q_{1}\right) \otimes D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right)\right] .
$$

Let us divide this with $T^{2} Q_{1} Q_{2}$, where $Q_{1} \equiv\left|Q_{1}\right|$ and $Q_{2} \equiv\left|Q_{2}\right|$, and then multiply by it again:

$$
\omega_{Q(T)}=Q_{1} Q_{2} T^{2} \max _{E, D, \rho} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\rho_{T \hat{T}}\left(\frac{E_{T}\left(a_{1} \mid q_{1}\right)}{T Q_{1}} \otimes \frac{D_{\hat{T}}\left(a_{2} \mid q_{2}\right)}{T Q_{2}}\right)\right] .
$$

Writing classical systems as diagonal quantum systems, we obtain

$$
\begin{align*}
\omega_{Q(T)}(V, \pi)= & Q_{1} Q_{2} T^{2} \max _{E, D, \rho} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}}^{\pi} \otimes \rho_{T \hat{T}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}}\right)\right] \\
\text { s.t. } & \rho_{T \hat{T}} \geq 0, \quad \operatorname{Tr}\left[\rho_{T \hat{T}}\right]=1 \\
& E_{A_{1} Q_{1} T}=\sum_{a_{1}, q_{1}}\left|a_{1} q_{1}\right\rangle\left\langle\left. a_{1} q_{1}\right|_{A_{1} Q_{1}} \otimes \frac{E_{T}\left(a_{1} \mid q_{1}\right)}{T Q_{1}} \geq 0\right. \\
& E_{Q_{1} T}=\sum_{q_{1}}\left|q_{1}\right\rangle\left\langle\left. q_{1}\right|_{Q_{1}} \otimes \frac{\mathbb{1}_{T}}{T Q_{1}}=\frac{\mathbb{1}_{Q_{1} T}}{T Q_{1}}\right.  \tag{C.1}\\
& D_{A_{2} Q_{2} \hat{T}}=\sum_{a_{2} q_{2}}\left|a_{2} q_{2}\right\rangle\left\langle\left. a_{2} q_{2}\right|_{A_{2} Q_{2}} \otimes \frac{D_{\hat{T}}\left(a_{2} \mid q_{2}\right)}{T Q_{2}} \geq 0\right. \\
& D_{Q_{2} \hat{T}}=\sum_{q_{2}}\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}} \otimes \frac{\mathbb{1}_{\hat{T}}}{T Q_{2}}=\frac{\mathbb{1}_{Q_{2} \hat{T}}}{T Q_{2}}\right.
\end{align*}
$$

where $V_{A_{1} A_{2} Q_{1} Q_{2}}^{\pi}=\sum_{a_{1}, q_{1}, a_{2}, q_{2}} \pi\left(q_{1}, q_{2}\right) V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\left|a_{1}, a_{2}, q_{1}, q_{2}\right\rangle\left\langle a_{1}, a_{2}, q_{1}, q_{2}\right|$. If we use the same modified swap trick from Lemma 21, we can again rewrite this as an instance of quantum separability problems with linear constraints: the optimisation with objective function

$$
\omega_{Q(T)}(V, \pi)=Q_{1} Q_{2} T^{2} \max _{E, D, \rho} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}}^{\pi} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right)\left(E_{A_{1} Q_{1} T} \otimes D_{A_{2} Q_{2} \hat{T}} \otimes \rho_{S \hat{S}}\right)\right]
$$

with the same constraints in Eq. (C.1). Then, using extendible states, we can write a hierarchy of SDP relaxations for general two-player games as

$$
\begin{aligned}
& \operatorname{sdp}_{n_{1}, n_{2}}^{G}(V, \pi, T):=Q_{1} Q_{2} T^{2} \max _{\rho} \operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}}^{\pi} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right) \rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)(S \hat{S})}\right] \\
& \text { s.t. } \quad \rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})} \geq 0, \quad \operatorname{tr}\left[\rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})}\right]=1 \\
& \mathcal{U}_{\left(A_{1} Q_{1} T\right)^{n_{1}}}^{\pi}\left(\rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})}\right)=\rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})} \quad \forall \pi \in \mathcal{S}\left(\left(A_{1} Q_{1} T\right)^{n_{1}}\right) \\
& \mathcal{U}_{\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}}^{\pi}\left(\rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})}\right)=\rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}(S \hat{S})}} \quad \forall \pi \in \mathcal{S}\left(\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}\right) \\
& \operatorname{tr}_{A_{1}}\left[\rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})}\right]=\left(\frac{\mathbb{1}_{Q_{1} T}}{T Q_{1}}\right) \otimes \rho_{\left(A_{1} Q_{1} T\right)^{n_{1}-1}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})} \\
& \operatorname{tr}_{A_{2}}\left[\rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})}\right]=\left(\frac{\mathbb{1}_{Q_{2} \hat{T}}}{T Q_{2}}\right) \otimes \rho_{\left(A_{1} Q_{1} T\right)^{n_{1}}\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}-1}(S \hat{S})} \\
& \rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}^{T_{A_{1} Q_{1} T}} \geq 0, \quad \rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}^{T_{\left(A_{2} Q_{2} \hat{T}\right.}^{n_{1}}} \geq 0, \quad \ldots
\end{aligned}
$$

It is again easy to bound the computational complexity of these SDP relaxations using the multipartite quantum de Finetti theorem derived in Section 7.3.2. We find that

$$
\begin{aligned}
& \left|\operatorname{sdp}_{n_{1}, n_{2}}^{G}(V, \pi, T)-\omega_{Q(T)}(V, \pi)\right| \\
& \leq Q_{1} Q_{2} T^{2}\left|\operatorname{Tr}\left[\left(V_{A_{1} A_{2} Q_{1} Q_{2}}^{\pi} \otimes \Phi_{T \hat{T} \mid S \hat{S}}\right)\left(\tilde{\rho}_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}-\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}\right)\right]\right| \\
& \leq Q_{1} Q_{2} T^{2}\left(18^{3 / 2} T^{4}(\sqrt{4 \ln 2})\left(\sqrt{\frac{\left(\log A_{1}+16 \log T\right)}{n_{2}}+\frac{\log A_{1}}{n_{1}}}\right)\right)
\end{aligned}
$$

where we assume that $\tilde{\rho}_{A_{1} Q_{1} T\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$ is the optimal solution of $\operatorname{sdp}_{n_{1}, n_{2}}^{G}(V, \pi, T)$, and $\sum_{i} p_{i} \sigma_{A_{1} Q_{1} T}^{i} \otimes \omega_{A_{2} Q_{2} \hat{T}}^{i} \otimes \tau_{S \hat{S}}^{i}$ is one of the close separable states to the reduced state $\tilde{\rho}_{A_{1} Q_{1} T A_{2} Q_{2} \hat{T} S \hat{S}}$ specified by Lemma 27. The intermediate step is the same as in the proof of Theorem 28. The additional $Q_{1} Q_{2}$ factor leads to a worse convergence rate; setting $n:=n_{1}=n_{2}$, the size of the SDP for achieving a desired accuracy $\epsilon$ becomes

$$
\exp \left(\mathcal{O}\left(\frac{T^{12} Q^{4}\left(\log ^{2} A T+\log A T \log Q\right)}{\epsilon^{2}}\right)\right)
$$

where $Q_{1}=Q_{2} \equiv Q$ and $\left|A_{1}\right|=\left|A_{2}\right| \equiv A$. In contrast to the case for free games, Eq. (7.33), this is exponential in terms of $Q$.

## C. 2 Distortion relative to quantum side information

In this appendix, we prove the second part of Lemma 26 which states that for a traceless Hermitian operator $\gamma_{A B}$ on $\mathcal{H}_{A B}$, there exists a measurement $\mathcal{M}_{B}$ on $\mathcal{H}_{B}$ with at most $|B|^{6}$ outcomes such that $\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B}\right)\left(\gamma_{A B}\right)\right\|_{1} \geq \frac{1}{2|B|}\left\|\gamma_{A B}\right\|_{1}$. The proof is inspired by [126, Theorem $16]$.

Proof of Lemma 26. Let us start with the maximally entangled state

$$
\Phi_{A^{\prime} \mid B^{\prime}}=|\Phi\rangle\left\langle\left.\Phi\right|_{A^{\prime} \mid B^{\prime}} \text { where } \mid \Phi\right\rangle_{A^{\prime} \mid B^{\prime}}=\frac{1}{\left|A^{\prime}\right|\left|B^{\prime}\right|} \sum_{i}|i\rangle_{A^{\prime}}|i\rangle_{B^{\prime}}, \text { and }\left|A^{\prime}\right|=\left|B^{\prime}\right|
$$

We can create a separable state $\omega_{A^{\prime} B^{\prime}}$ by mixing $\Phi_{A^{\prime} \mid B^{\prime}}$ with another separable state $\sigma_{A^{\prime} B^{\prime}}=$
$\frac{\mathbb{1}_{A^{\prime} B^{\prime}}-\Phi_{A^{\prime} \mid B^{\prime}}}{\left|B^{\prime}\right|^{2}-1}$ as

$$
\omega_{A^{\prime} B^{\prime}}=\frac{1}{\left|B^{\prime}\right|} \Phi_{A^{\prime} B^{\prime}}+\frac{\left|B^{\prime}\right|-1}{\left|B^{\prime}\right|} \sigma_{A^{\prime} B^{\prime}} \in \operatorname{SEP}\left(\mathrm{A}^{\prime}: \mathrm{B}^{\prime}\right),
$$

where $\operatorname{SEP}\left(\mathrm{A}^{\prime}: \mathrm{B}^{\prime}\right)$ denotes the set of separable states with respect to the bipartition $A^{\prime} \mid B^{\prime}$. Hence, we can write $\omega_{A^{\prime} B^{\prime}}=\sum_{i} p_{i} \omega_{A^{\prime}}^{i} \otimes \omega_{B^{\prime}}^{i}$ for some probability distribution $\left\{p_{i}\right\}_{i}$ and states $\left\{\omega_{A^{\prime}}^{i}\right\}_{i}$ and $\left\{\omega_{B^{\prime}}^{i}\right\}_{i}$ with at most $\left|A^{\prime} B^{\prime}\right|^{2}$ elements [127]. Next, we define a measurement $\mathcal{M}_{B}$ with operators $\left\{\tilde{M}_{B}(i, k)\right\}_{i, k}$, as well as a set of measurements $\left\{\mathcal{M}_{A}^{i, k}\right\}_{i, k}$ with operators $\left\{\tilde{M}_{A}^{i, k}(j)\right\}_{j}$ as

$$
\begin{align*}
\tilde{M}_{B}(i, k) & =\operatorname{Tr}_{B^{\prime}}\left[p_{i} U_{B}^{\dagger}(k) \sqrt{\omega_{B^{\prime}}^{i}} \Phi_{B B^{\prime}} \sqrt{\omega_{B^{\prime}}^{i}} U_{B}(k)\right] \text { and }  \tag{C.2}\\
\tilde{M}_{A}^{i, k}(j) & =\operatorname{Tr}_{A^{\prime}}\left[\sqrt{\omega_{A^{\prime}}^{i}} U_{A^{\prime}}^{\dagger}(k) N_{A A^{\prime}}(j) U_{A^{\prime}}(k) \sqrt{\omega_{A^{\prime}}^{i}}\right] \tag{C.3}
\end{align*}
$$

where $U(k)$ denote generalised Pauli operators, $\omega_{A^{\prime}}^{i}$ and $\omega_{B^{\prime}}^{i}$ are the elements of the decomposition of $\omega_{A^{\prime} B^{\prime}}$, and $\left\{N_{A A^{\prime}}(j)\right\}_{j}$ are measurement operators defined later. We can check that both definitions indeed correspond to valid measurements:

$$
\sum_{i, k} \tilde{M}_{B}(i, k)=\mathbb{1}_{B}, \sum_{j} \tilde{M}_{A}^{i, k}(j)=\mathbb{1}_{A}, \text { and } \tilde{M}_{B}(i, k), \tilde{M}_{A}^{i, k}(j) \geq 0 \forall i, k, j
$$

The goal is to show that $\mathcal{M}_{B}$ defined in (C.2) gives rise to (7.28). Before showing that, however, it is helpful to understand where these measurements came from. They are related to the quantum teleportation protocol from [128]. Without loss of generality, let us assume that $|A| \geq|B|=\left|A^{\prime}\right|=\left|B^{\prime}\right|$. Then, the quantum teleportation protocol from $B$ to $A$ is a quantum channel defined as [128]

$$
\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}(\cdot)=\sum_{k=1}^{|B|^{2}} U_{A^{\prime}}(k) \operatorname{Tr}_{B B^{\prime}}\left[(\cdot)\left(\mathbb{1}_{A A^{\prime}} \otimes U_{B}(k) \Phi_{B B^{\prime}} U_{B}^{\dagger}(k)\right)\right] U_{A^{\prime}}^{\dagger}(k)
$$

For a traceless Hermitian operator $\gamma_{A B}$, we then consider

$$
\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}}\right)\right\|_{1}=\sum_{j}\left|\operatorname{Tr}\left[N_{A A^{\prime}}(j)\left(\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}}\right)\right)\right]\right|,
$$

where we used the expression $\left\|X_{A}\right\|_{1}=\max _{\left\{M_{A}(i)\right\}_{i}} \sum_{i}\left|\operatorname{tr}\left[M_{A}(i) X_{A}\right]\right|$ for the trace norm with corresponding arg max $\left\{N_{A A^{\prime}}(j)\right\}_{j}$ to be used in (C.3). We have

$$
\begin{align*}
& \left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}}\right)\right\|_{1} \\
& =\sum_{j}\left|\sum_{k} \operatorname{Tr}\left[N_{A A^{\prime}}(j)\left(U_{A^{\prime}}(k) \operatorname{tr}_{B B^{\prime}}\left[\left(\gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}}\right)\left(\mathbb{1}_{A A^{\prime}} \otimes U_{B}(k) \Phi_{B B^{\prime}} U_{B}^{\dagger}(k)\right)\right] U_{A^{\prime}}^{\dagger}(k)\right)\right]\right| \\
& =\sum_{j}\left|\sum_{k} \operatorname{Tr}\left[\left(U_{A^{\prime}}^{\dagger}(k) N_{A A^{\prime}}(j) U_{A^{\prime}}(k) \otimes \mathbb{1}_{B B^{\prime}}\right)\left(\left(\gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}}\right)\left(\mathbb{1}_{A A^{\prime}} \otimes U_{B}(k) \Phi_{B B^{\prime}} U_{B}^{\dagger}(k)\right)\right)\right]\right| \\
& =\sum_{j}\left|\sum_{k} \operatorname{Tr}\left[\left(U_{A^{\prime}}^{\dagger}(k) N_{A A^{\prime}}(j) U_{A^{\prime}}(k) \otimes U_{B}^{\dagger}(k) \Phi_{B B^{\prime}} U_{B}(k)\right)\left(\gamma_{A B} \otimes\left(\sum_{i} p_{i} \omega_{A^{\prime}}^{i} \otimes \omega_{B^{\prime}}^{i}\right)\right)\right]\right| \\
& =\sum_{j} \mid \sum_{i, k} \operatorname{Tr}\left[\left(\left(\sqrt{\omega_{A^{\prime}}^{i}} U_{A^{\prime}}^{\dagger}(k) N_{A A^{\prime}}(j) U_{A^{\prime}}(k) \sqrt{\omega_{A^{\prime}}^{i}}\right)\right.\right. \\
& \left.\left.\quad \otimes\left(p_{i} U_{B}^{\dagger}(k) \sqrt{\omega_{B^{\prime}}^{i}} \Phi_{B B^{\prime}} \sqrt{\omega_{B^{\prime}}^{i}} U_{B}(k)\right)\right)\left(\gamma_{A B} \otimes \mathbb{1}_{A^{\prime} B^{\prime}}\right)\right] \mid \\
& =\sum_{j}\left|\sum_{i, k} \operatorname{Tr}\left[\gamma_{A B}\left(\tilde{M}_{A}^{i, k}(j) \otimes \tilde{M}_{B}(i, k)\right)\right]\right| . \tag{C.4}
\end{align*}
$$

The measurement $\mathcal{M}_{B}$ defined in (C.2) now gives rise to

$$
\begin{align*}
& \left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B}\right)\left(\gamma_{A B}\right)\right\|_{1} \\
& =\sum_{i, k}\left\|\operatorname{Tr}_{B}\left[\left(\mathbb{1}_{A} \otimes \tilde{M}_{B}(i, k)\right) \gamma_{A B}\right]\right\|_{1} \\
& =\sum_{i, k} \max _{\left\{M_{A}^{i, k}(j)\right\}_{j}} \sum_{j}\left|\operatorname{Tr}\left[\left(M_{A}^{i, k}(j) \otimes \tilde{M}_{B}(i, k)\right) \gamma_{A B}\right]\right| \\
& \geq \sum_{i, k} \sum_{j}\left|\operatorname{Tr}\left[\left(\tilde{M}_{A}^{i, k}(j) \otimes \tilde{M}_{B}(i, k)\right) \gamma_{A B}\right]\right| \\
& \geq \sum_{j}\left|\sum_{i, k} \operatorname{Tr}\left[\left(\tilde{M}_{A}^{i, k}(j) \otimes \tilde{M}_{B}(i, k)\right) \gamma_{A B}\right]\right|  \tag{C.5}\\
& =\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}}\right)\right\|_{1} \quad(\text { by }(\mathrm{C} .4)) \\
& =\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes\left(\frac{1}{|B|} \Phi_{A^{\prime} B^{\prime}}+\frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}}\right)\right)\right\|_{1} \\
& =\left\|\frac{1}{|B|} \tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \Phi_{A^{\prime} B^{\prime}}\right)+\frac{|B|-1}{|B|} \tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \sigma_{A^{\prime} B^{\prime}}\right)\right\|_{1} \\
& \geq \frac{1}{|B|}\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \Phi_{A^{\prime} B^{\prime}}\right)\right\|_{1}-\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes\left(\frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}}\right)\right)\right\|_{1}, \tag{C.6}
\end{align*}
$$

where in the third line we substituted the measurement operators $\left\{\tilde{M}_{A}^{i, k}(j)\right\}_{j}$ instead of the maximisation, and in the last line we used the reverse triangular inequality. Note that the first term in the last line is equivalent to $\left\|\gamma_{A B}\right\|_{1}$ since $\Phi_{A^{\prime} B^{\prime}}$ is the maximally entangled state. Let us investigate the second term more closely. We have the chain of elementary implications

$$
\begin{aligned}
& \frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}} \leq \frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}}+\frac{1}{|B|} \Phi_{A^{\prime} B^{\prime}}=\omega_{A^{\prime} B^{\prime}} \\
& \Rightarrow \gamma_{A B} \otimes \frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}} \leq \gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}} \\
& \Rightarrow\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes\left(\frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}}\right)\right)\right\|_{1} \leq\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes \omega_{A^{\prime} B^{\prime}}\right)\right\|_{1} \\
& \Rightarrow\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes\left(\frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}}\right)\right)\right\|_{1} \leq \sum_{j}\left|\sum_{i, k} \operatorname{Tr}\left[\gamma_{A B}\left(\tilde{M}_{A}^{i, k}(j) \otimes \tilde{M}_{B}(i, k)\right)\right]\right|
\end{aligned}
$$

$$
\begin{equation*}
\Rightarrow\left\|\tau_{A B A^{\prime} B^{\prime} \rightarrow A A^{\prime}}\left(\gamma_{A B} \otimes\left(\frac{|B|-1}{|B|} \sigma_{A^{\prime} B^{\prime}}\right)\right)\right\|_{1} \leq\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B}\right)\left(\gamma_{A B}\right)\right\|_{1} \tag{C.4}
\end{equation*}
$$

and substituting this into (C.6) yields the claim

$$
\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B}\right)\left(\gamma_{A B}\right)\right\|_{1} \geq \frac{1}{|B|}\left\|\gamma_{A B}\right\|_{1}-\left\|\left(\mathcal{I}_{A} \otimes \mathcal{M}_{B}\right)\left(\gamma_{A B}\right)\right\|_{1} .
$$

It remains to quantify the number of measurement outcomes of $\mathcal{M}_{B}$ with operators $\left\{\tilde{M}_{B}(i, k)\right\}_{i, k}$ defined in (C.2). The index $i$ came from the number of elements in the separable state $\omega_{A^{\prime} B^{\prime}}$, which is at most $\left|A^{\prime} B^{\prime}\right|^{2}=|B|^{4}$, and the index $k$ came from the number of generalised Pauli operators, which is $|B|^{2}$. Therefore, the number of outcomes is at most $|B|^{6}$.

## C. 3 Tripartite quantum de Finetti theorem with partial trace constraints

In this appendix, we prove Lemma 27, the special case of tripartite quantum de Finetti theorem with linear constraints when (i) the linear constraints are partial trace constraints, and (ii) the given extendible state includes some classical subsystems. This requires to prove an improved version of Lemma 25 when there exists some specific partial trace constraint. Let us first consider the bound on the conditional bipartite quantum mutual information for this special case.

Lemma 35. Let $\rho_{A B C D}$ be a quantum state such that $\operatorname{Tr}_{A}\left[\rho_{A B C D}\right]=\rho_{B} \otimes \rho_{C D}$. Then, we have that

$$
I(A B: C \mid D)_{\rho} \leq 2 \log |A|
$$

Proof. Using the chain rule for the conditional quantum mutual information in Eq. (7.20), we have

$$
I(A B: C \mid D)_{\rho}=I(B: C \mid D)_{\rho}+I(A: C \mid B D)_{\rho}=I(A: C \mid B D)_{\rho}
$$

where we used the fact that the systems $B C$ are uncorrelated after tracing out the system $A$.

Using the chain rule of the mutual information in Eq. (7.19) we get that

$$
I(A: C \mid B D)=I(A: B C D)-I(A: B D) \leq I(A: B C D) \leq 2 \log |A|
$$

Exploiting Lemma 35, we prove the following lemma, which is an improved version of Lemma 25 for this special case.

Lemma 36. Consider a quantum-classical state $\rho_{(A B) Z_{1}^{n_{1}} W_{1}^{n_{2}}}$ with classical systems $Z$ and $W$, which satisfies the linear constraint

$$
\operatorname{Tr}_{A}\left[\rho_{(A B) Z_{1}^{n_{1}} W_{1}^{n_{2}}}\right]=\rho_{B} \otimes \rho_{Z_{1}^{n_{1}} W_{1}^{n_{2}}} .
$$

Then, there exist integers $\bar{m} \in\left[0, n_{1}\right)$ and $\bar{\ell} \in\left[0, n_{2}\right)$ such that

$$
I\left(A B: Z_{\bar{m}+1}: W_{\bar{\ell}+1} \mid Z_{1}^{\bar{m}} W_{1}^{\bar{\ell}}\right) \leq 2\left(\frac{\log |A|}{n_{1}}+\frac{(\log |A|+\log |Z|)}{n_{2}}\right)
$$

and

$$
\begin{aligned}
& \mathbb{E}_{z_{1}^{\bar{m}} w_{1}^{\bar{l}}}\left[\left\|\rho_{(A B) Z_{\bar{m}+1} W_{\bar{\ell}+1} \mid z_{1}^{\bar{m}} w_{1}^{\bar{l}}}-\rho_{A B \mid z_{1}^{\bar{m}} w_{1}^{\bar{l}}} \otimes \rho_{Z_{\bar{m}+1} \mid z_{1}^{\bar{m}} w_{1}^{\bar{q}}} \otimes \rho_{W_{\bar{\ell}+1} \mid z_{1}^{\bar{m}} w_{1}^{\bar{l}}}\right\|_{1}^{2}\right] \\
& \leq 4 \ln 2\left(\frac{\log |A|}{n_{1}}+\frac{\log |A|+\log |Z|}{n_{2}}\right) .
\end{aligned}
$$

Proof. As the first step, notice that for all $m$ and $\ell$, using the relation between multipartite mutual information and bipartite ones in Eq. (7.18), we can express the multipartite conditional mutual information as,

$$
\begin{equation*}
I\left(A B: Z_{m+1}: W_{\ell+1} \mid Z_{1}^{m} W_{1}^{\ell}\right)=I\left(A B: Z_{m+1} \mid Z_{1}^{m} W_{1}^{\ell}\right)+I\left(A B Z_{m+1}: W_{\ell+1} \mid Z_{1}^{m} W_{1}^{\ell}\right) \tag{C.7}
\end{equation*}
$$

We now derive a bound for the RHS of the above equation, valid for a specific choice of $m$ and
$\ell$. For any $\ell \in\left[0, n_{2}\right)$, we have that

$$
\sum_{m=0}^{n_{1}-1} I\left(A B: Z_{m+1} \mid Z_{1}^{m} W_{1}^{\ell}\right)=I\left(A B: Z_{1}^{n_{1}} \mid W_{1}^{\ell}\right) \leq 2 \log |A|,
$$

where we used the chain rule in Eq. (7.20) for the first equality and Lemma 35 for the second inequality. By summing over $\ell$, we obtain

$$
\begin{equation*}
\sum_{m=0}^{n_{1}-1} \sum_{\ell=0}^{n_{2}-1} I\left(A B: Z_{m+1} \mid Z_{1}^{m} W_{1}^{\ell}\right) \leq 2 n_{2} \log |A| \tag{C.8}
\end{equation*}
$$

Similarly, for any $m \in\left[0, n_{1}\right)$, we find that

$$
\sum_{\ell=0}^{n_{2}-1} I\left(A B Z_{m+1}: W_{\ell+1} \mid Z_{1}^{m} W_{1}^{\ell}\right)=I\left(A B Z_{m+1}: W_{1}^{n_{2}} \mid Z_{1}^{m}\right) \leq 2 \log \left|A Z_{m+1}\right|
$$

where the inequality follows from Lemma 35 when the first partition of the system is $A Z_{m+1}$. Summing over $m$ gives us

$$
\begin{equation*}
\sum_{m=0}^{n_{1}-1} \sum_{\ell=0}^{n_{2}-1} I\left(A B Z_{m+1}: W_{\ell+1} \mid Z_{1}^{m} W_{1}^{\ell}\right) \leq 2 n_{1}(\log |A|+\log |Z|) \tag{C.9}
\end{equation*}
$$

Combining Eq. (C.8) and Eq. (C.9), we obtain

$$
\begin{aligned}
& 2 n_{2} \log |A|+2 n_{1}(\log |A|+\log |Z|) \\
& \geq \sum_{m=0}^{n_{1}-1} \sum_{\ell=0}^{n_{2}-1}\left[I\left(A B: Z_{m+1} \mid Z_{1}^{m} W_{1}^{\ell}\right)+I\left(A B Z_{m+1}: W_{\ell+1} \mid Z_{1}^{m} W_{1}^{\ell}\right)\right] \\
& \geq n_{1} n_{2}\left[I\left(A B: Z_{\bar{m}+1} \mid Z_{1}^{\bar{m}} W_{1}^{\bar{\ell}}\right)+I\left(A B Z_{\bar{m}+1}: W_{\bar{\ell}+1} \mid Z_{1}^{\bar{m}} W_{1}^{\bar{\ell}}\right)\right] \\
& =n_{1} n_{2} I\left(A B: Z_{\bar{m}+1}: W_{\bar{\ell}+1} \mid Z_{1}^{\bar{m}} W_{1}^{\bar{\ell}}\right)
\end{aligned}
$$

where $\bar{m}$ and $\bar{\ell}$ are the indices of the smallest element in the sum, and the last equality follows from Eq. (C.7). This proves the first part of the theorem. The second part is obtained as follows. First notice that, when the conditioning system is classical, we can write the conditional mutual
information as

$$
I\left(A B: Z_{\bar{m}+1}: W_{\bar{\ell}+1} \mid Z_{1}^{\bar{m}} W_{1}^{\bar{\ell}}\right)=\mathbb{E}_{z_{1}^{\bar{m}} w_{1}^{\bar{\ell}}}\left[I\left(A B: Z_{\bar{m}+1}: W_{\bar{\ell}+1}\right)_{A B Z_{\bar{m}+1} W_{\bar{\ell}+1} \mid z_{1}^{\overline{\bar{m}}} w_{1}^{\overline{\overline{ }}}}\right]
$$

Then, by using the fact that the mutual information can be expressed in terms of the relative entropy, as in Eq. (7.17), and by using Pinsker's inequality in Eq. (7.27), we can derive the second part of the theorem.

Note that now we are missing the dimension factor of system $B$, which is uncorrelated from the classical systems $Z_{1}^{n_{1}} W_{1}^{n_{2}}$ after tracing out system $A$, in the two upper bounds in Lemma 35 . This leads to an improved bound in the special quantum de Finetti theorem in Lemma 27:

Proof of Lemma 27. Let $\mathcal{M}_{B \rightarrow Y}\left(\mathcal{M}_{C \rightarrow W}\right)$ be a quantum-to-classical channel, i.e. a measurement, from the quantum system $B(C)$ to the classical system $Y(W)$. By measuring both $B$ and $C$, we obtain the following quantum-classical state,

$$
\rho_{(A X \tilde{X}) Y^{n_{1}}(W Z \tilde{Z})^{n_{2}}}=\left(\mathcal{I}_{A X \tilde{X}(Z \tilde{Z})^{n_{2}}} \otimes \mathcal{M}_{B \rightarrow Y}^{\otimes n_{1}} \otimes \mathcal{M}_{C \rightarrow W}^{\otimes n_{2}}\right)\left(\rho_{(A X \tilde{X}) B^{n_{1}}\left(C Z \tilde{)^{n_{2}}}\right.}\right)
$$

where $\mathcal{M}_{B \rightarrow Y}^{\otimes n_{1}}$ is composed of $n_{1}$ independent and identical measurements $\mathcal{M}_{B \rightarrow Y}$, each one acting on a different $B$ system, and $\mathcal{M}_{C \rightarrow W}^{\otimes n_{2}}$ is defined similarly as well. It is easy to see that the post-measurement state still satisfies the linear constraint Eq. (7.31)

$$
\operatorname{Tr}_{X}\left[\rho_{(A X \tilde{X}) Y^{n_{1}}(W Z \tilde{Z})^{n_{2}}}\right]=\mathcal{X}_{A \tilde{X}} \otimes \rho_{Y^{n_{1}}(W Z \tilde{Z})^{n_{2}}}
$$

and therefore is compatible with the condition of Lemma 36. Then, we can find $m \in\left[0, n_{1}\right)$ and $\ell \in\left[0, n_{2}\right)$ such that

$$
\begin{array}{r}
\mathbb{E}_{y^{m} w^{\ell}}\left[\left\|\rho_{(A X \tilde{X}) Y_{m+1}(W Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}-\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{Y_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(W Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}\right\|_{1}^{2}\right]  \tag{C.10}\\
\leq 4 \ln 2\left(\frac{\log |X|}{n_{1}}+\frac{\log |X|+\log |Y|}{n_{2}}\right)
\end{array}
$$

by Lemma 36. For given $y^{m}$ and $w^{\ell}$, let us define the traceless Hermitian operator

$$
\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})} \equiv \rho_{(A X \tilde{X}) B_{m+1}(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}-\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}
$$

This operator is related to the one in the LHS of Eq. (C.10) by a measurement on the $B$ and $C \tilde{C}$ systems via

$$
\begin{aligned}
& \left(\mathcal{I}_{A X \tilde{X} Z \tilde{Z}} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow W}\right)\left(\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right) \\
& =\rho_{(A X \tilde{X}) Y_{m+1}(W Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}-\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{Y_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(W Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}
\end{aligned}
$$

Again, we consider two ways to relate the trace norm of $\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}$ to $\left(\mathcal{I}_{A X \tilde{X} Z \tilde{Z}} \otimes \mathcal{M}_{B \rightarrow Y} \otimes\right.$ $\left.\mathcal{M}_{C \rightarrow W}\right)\left(\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right)$. Firstly, if we let the measurements $\mathcal{M}_{B \rightarrow Y}$ and $\mathcal{M}_{C \rightarrow W}$ be the measurement described in the second part of Lemma 26, we get

$$
\begin{aligned}
& \left\|\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right\|_{1} \\
& \leq 2|C|\left\|\left(\mathcal{I}_{A X \tilde{X} B Z \tilde{Z}} \otimes \mathcal{M}_{C \rightarrow W}\right)\left(\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right)\right\|_{1} \\
& \leq 2|B| \times 2|C|\left\|\left(\mathcal{I}_{A X \tilde{X} W Z \tilde{Z}} \otimes \mathcal{M}_{B \rightarrow Y}\right)\left(\left(\mathcal{I}_{A X \tilde{X} B Z \tilde{Z}} \otimes \mathcal{M}_{C \rightarrow W}\right)\left(\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right)\right)\right\|_{1} \\
& =4|B C|\left\|\left(\mathcal{I}_{A X \tilde{X} Z \tilde{Z}} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow W}\right)\left(\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right)\right\|_{1},
\end{aligned}
$$

and $|Y| \leq|B|^{6}$. Secondly, if we use the first part of Lemma 26 instead, we obtain

$$
\begin{aligned}
\left\|\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right\|_{1} & \leq \sqrt{18^{3}|A B C|}\left\|\left(\mathcal{M}_{A X \tilde{X} Z \tilde{Z}} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow W}\right)\left(\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right)\right\|_{1} \\
& \leq \sqrt{18^{3}|A B C|}\left\|\left(\mathcal{I}_{A X \tilde{X} Z \tilde{Z}} \otimes \mathcal{M}_{B \rightarrow Y} \otimes \mathcal{M}_{C \rightarrow W}\right)\left(\gamma_{(A X \tilde{X}) B(C Z \tilde{Z})}\right)\right\|_{1}
\end{aligned}
$$

with $|Y| \leq|B|^{8}$, where the second inequality follows from the monotonicity of the trace norm under CPTP maps. Note that the classical systems do not contribute to the dimension factor in the first inequality as we have discussed in Eq. (7.29). Combining Eq. (C.10) with the above
two results gives

$$
\begin{align*}
\mathbb{E}_{y^{m} w^{\ell}} & {\left[\left\|\rho_{(A X \tilde{X}) B_{m+1}(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}-\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}\right\|_{1}^{2}\right] } \\
& \leq \min \left\{18^{3 / 2} \sqrt{|A B C|}, 4|B C|\right\}^{2}(4 \ln 2)\left(\frac{\log |X|+8 \log |B|}{n_{2}}+\frac{\log |X|}{n_{1}}\right) \tag{C.11}
\end{align*}
$$

Depending on the dimensions, we can choose the tighter one between the two bounds. The following chain of inequalities combining with Eq. (C.11) concludes the proof of the first part of the theorem,

$$
\begin{aligned}
& \left\|\rho_{(A X \tilde{X}) B_{m+1}(C Z \tilde{Z})_{\ell+1}}-\mathbb{E}_{y^{m} w^{\ell}}\left[\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}\right]\right\|_{1} \\
& \leq \mathbb{E}_{y^{m} w^{\ell}}\left[\left\|\rho_{(A X \tilde{X}) B_{m+1}(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}-\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}\right\|_{1}\right]
\end{aligned}
$$

( $\because$ Triangular Inequality)

$$
\leq \sqrt{\mathbb{E}_{y^{m} w^{\ell}}\left[\left\|\rho_{(A X \tilde{X}) B_{m+1}(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}-\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}\right\|_{1}^{2}\right]}
$$

$(\because$ Concavity),
where the quantum state $\mathbb{E}_{y^{m} w^{\ell}}\left[\rho_{A X \tilde{X} \mid y^{m} w^{\ell}} \otimes \rho_{B_{m+1} \mid y^{m} w^{\ell}} \otimes \rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}\right]$ is separable over the tripartite cut $A X \tilde{X}|B| C Z \tilde{Z}$. It is worth noting that, since the state under consideration is permutation invariant over $B^{n_{1}}$ and $(C Z \tilde{Z})^{n_{2}}$, the result we obtain is independent of the specific $m$ and $\ell$ considered. This closes the proof of Eq. (7.32).

To conclude the proof, we need to show that each state in the mixture still satisfies the corresponding linear constraint. For a state $\rho_{A X \tilde{X} \mid y^{m} w^{\ell}}$ describing the $A X \tilde{X}$ system, we have

$$
\begin{aligned}
\operatorname{Tr}_{X}\left[\rho_{A X \tilde{X} \mid y^{m} w^{\ell}}\right] & =\frac{\operatorname{Tr}_{Y^{m}(W Z \tilde{Z})^{\ell}}\left[\left(\mathbb{1}_{A \tilde{X}\left(Z \tilde{)^{\ell}}\right.} \otimes M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}}\right) \operatorname{Tr}_{X}\left[\rho_{(A X \tilde{X}) B^{m}(C Z \tilde{Z})}\right]\right]}{\operatorname{Tr}\left[\left(\mathbb{1}_{A X \tilde{X}(Z \tilde{Z})^{\ell}} \otimes M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}}\right) \rho_{(A X \tilde{X}) B^{m}(C Z \tilde{Z})^{\ell}}\right]} \\
& =\frac{\operatorname{Tr}_{Y^{m}(W Z \tilde{Z})^{\ell}}\left[\left(\mathbb{1}_{A \tilde{X}(Z \tilde{Z})^{\ell}} \otimes M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}}\right)\left(\mathcal{X}_{A \tilde{X}} \otimes \rho_{B^{m}(C Z \tilde{Z})^{\ell}}\right)\right]}{\operatorname{Tr}\left[\left(M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}}\right) \rho_{B^{m} C^{\ell}}\right]} \\
& =\mathcal{X}_{A \tilde{X}},
\end{aligned}
$$

where $M_{B^{m} \rightarrow Y^{m}}^{y^{m}}$ is the measurement effect over $B^{m}$ corresponding to the outcome $y^{m}$, while $M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}}$ is the effect over $C^{\ell}$ corresponding to the outcome $w^{\ell}$. For a state $\rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}$, we have

$$
\begin{aligned}
& \operatorname{Tr}_{Z}\left[\rho_{(C Z \tilde{Z})_{\ell+1} \mid y^{m} w^{\ell}}\right] \\
& =\frac{\operatorname{Tr}_{Y^{m}(W Z \tilde{Z})^{\ell}}\left[\left(M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}} \otimes \mathbb{1}_{(Z \tilde{Z})^{\ell}(C \tilde{Z})_{\ell+1}}\right) \operatorname{Tr}_{Z}\left[\rho_{B^{m}(C Z \tilde{Z})^{\ell+1}}\right]\right]}{\operatorname{Tr}\left[\left(M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}} \otimes \mathbb{1}_{(Z \tilde{Z})^{\ell}(C Z \tilde{Z})_{\ell+1}}\right) \rho_{B^{m}\left(C Z \tilde{)^{\ell+1}}\right.}\right]} \\
& =\frac{\operatorname{Tr}_{Y^{m}(W Z \tilde{Z})^{\ell}}\left[\left(M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}} \otimes \mathbb{1}_{(Z \tilde{Z})^{\ell}(C \tilde{Z})_{\ell+1}}\right)\left(\mathcal{Z}_{(C \tilde{Z})_{\ell+1}} \otimes \rho_{B^{m}(C Z \tilde{Z})^{\ell}}\right)\right]}{\operatorname{Tr}\left[\left(M_{B^{m} \rightarrow Y^{m}}^{y^{m}} \otimes M_{C^{\ell} \rightarrow W^{\ell}}^{w^{\ell}}\right) \rho_{B^{m}(C Z \tilde{Z})^{\ell}}\right]} \\
& =\mathcal{Z}_{(C \tilde{Z})_{\ell+1}} .
\end{aligned}
$$

## C. 4 Constructing NPA matrices with the optimisation variable

In the main text, we explained how to add the NPA constraint to our SDP relaxations, and we noticed that some of the entries of NPA matrices can be expressed in terms of linear combinations of the optimisation variable $\rho_{\left(A_{1} Q_{1} T\right)\left(A_{2} Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}}$ in $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$. In this appendix, we explicitly show how these linear combinations are derived, for an alternative SDP relaxation obtained by extending the subsystems $A_{1} Q_{1} T$ and $A_{2} Q_{2} \hat{T}$ instead of $A_{2} Q_{2} \hat{T}$ and $S \hat{S}$. Then, using the obtained result, we derive the relation with the original SDP relaxation $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ in Eq. (7.7), described in Eq. (7.37).

We define the following alternative SDP relaxations for $\omega_{Q(T)}(V, \pi, T)$ of a two-player free
game $(V, \pi)$ :

$$
\begin{array}{ll} 
& \overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T):=T^{2} \max _{\rho} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right] \\
\text { s.t. } & \rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right) \geq 0, \sum_{a_{1}^{n_{1}}, a_{2}^{n_{2}, q_{1}^{n_{1}}, q_{2}^{n_{2}}}} \operatorname{Tr}\left[\rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)\right]=1 \\
& \rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right) \text { perm. inv. on }\left(A_{1} Q_{1} T\right)^{n_{1}} \text { w.r.t. to other systems } \\
& \rho_{T^{n_{1}} \hat{T}^{n_{2} S \hat{S}}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right) \text { perm. inv. on }\left(A_{2} Q_{2} \hat{T}\right)^{n_{2}} \quad \text { w.r.t. to other systems } \quad \text { (C. }  \tag{C.12}\\
& \sum_{a_{1}} \rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)=\pi_{1}\left(q_{1}\right) \frac{\mathbb{1}_{T}}{T} \otimes \rho_{T^{n_{1}-1} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}-1}, a_{2}^{n_{2}}, q_{1}^{n_{1}-1}, q_{2}^{n_{2}}\right) \\
& \sum_{a_{2}} \rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)=\pi_{2}\left(q_{2}\right) \frac{\mathbb{1}_{\hat{T}}}{T} \otimes \rho_{T^{n_{1}} \hat{T}^{n_{2}-1} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}-1}, q_{1}^{n_{1}}, q_{2}^{n_{2}-1}\right) \\
& \text { PPT conditions w.r.t. } T_{1}|\ldots| T_{n_{1}}\left|\hat{T}_{1}\right| \ldots\left|\hat{T}_{n_{2}}\right| S \hat{S},
\end{array}
$$

where we introduced the new notations $a_{i}^{n}$ for the $n$-dimensional vector $\boldsymbol{a}_{i}$ and $q_{i}^{n}$ for the $n$-dimensional vector $\boldsymbol{q}_{i}$. Note that in the above relaxation we extend subsystem $A_{1} Q_{1} T$ rather than $S \hat{S}$. This relaxation has slightly worse analytical convergence bounds than $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ described in the main text, since we can no more exploit the improved quantum de Finetti theorem with partial trace constraint in Lemma 27 if we extend the subsystem $A_{1} Q_{1} T$. Nonetheless, we consider it here for the following reasons:

1. From the point of view of numerical implementation, the relaxation $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$ requires a smaller number of optimisation variables than the original $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$, particularly when we add the NPA constraints. This is because, as we show in the remaining of the appendix, we can rewrite more entries of $\Gamma_{k}$ as linear combinations of the optimisation variable $\rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)$ when considering $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$. For the relaxation $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$, only a subset of these entries can be expressed as a linear combination of the optimisation variable, and the remaining entries need to be accounted as new variables.
2. The relation between the entries of $\Gamma_{k}$ and the optimisation variable of the relaxation $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$ is more straightforward than the one for the original relaxation $\operatorname{sdp}_{n_{1, n}}(V, \pi, T)$ used in the main text. For this reason, in this appendix, we explicitly
derive the formal relation and briefly comment on how to re-purpose it for the original $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$.
3. The relaxation $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$ scale better than $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ in numerical implementations. Extending the subsystem $A_{1} Q_{1} T$ only increases the total dimension of the quantum system in the variable by $T$, but extending the subsystem $S \hat{S}$ increases it by $T^{2}$. Thus, even though $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ has a better analytical convergence speed, it is practically more advantageous to use $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$.

Our goal is to provide a set of instructions that allows us to express the elements of the $k$-th level NPA matrix $\Gamma_{k}$ as functions of the optimisation variable in the new relaxations $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$, where $k \leq \min \left\{n_{1}, n_{2}\right\}$. Let us first recall that our relaxation aims to approximate the $\omega_{Q(T)}(V, \pi, T)$ of a two-player free game $(V, \pi)$, which we rewrite here for convenience

$$
\omega_{Q(T)}(V, \pi)=\max _{E, D, \rho} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\rho_{T \hat{T}}\left(E_{T}\left(a_{1} \mid q_{1}\right) \otimes D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right)\right] .
$$

By direct comparison with the objective function in Eq. (C.12), we see that, for the $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$ to obtain the optimal value $\omega_{Q(T)}(V, \pi)$, the optimisation variable reduced to the system $T \hat{T} S \hat{S}$ needs to satisfy

$$
\begin{equation*}
\rho_{T \hat{T} S \hat{S}}\left(a_{1}, q_{1}, a_{2}, q_{2}\right)=\frac{\pi\left(q_{1}, q_{2}\right)}{T^{2}} E_{T}\left(a_{1} \mid q_{1}\right) \otimes D_{\hat{T}}\left(a_{2} \mid q_{2}\right) \otimes \rho_{S \hat{S}}^{T} \quad \forall a_{1}, q_{1}, a_{2}, q_{2}, \tag{C.13}
\end{equation*}
$$

where $\left\{E_{T}\left(a_{1} \mid q_{1}\right)\right\}_{a_{1}, q_{1}}$ and $\left\{D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right\}_{a_{2}, q_{2}}$ are the optimal measurements, and $\rho_{S \hat{S}}$ is the optimal quantum state to be shared. In order to derive the full optimisation variable $\rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)$, we need to first make use of the assumption that $(V, \pi)$ is a free game. Thus, for a fixed value of $a_{1}, q_{1}, a_{2}$ and $q_{2}$ we can rewrite Eq. (C.13) as

$$
\rho_{T \hat{T} S \hat{S}}\left(a_{1}, q_{1}, a_{2}, q_{2}\right)=E_{T}\left(a_{1}, q_{1}\right) \otimes D_{\hat{T}}\left(a_{2}, q_{2}\right) \otimes \rho_{S \hat{S}}^{\mathrm{T}},
$$

where $E_{T}\left(a_{1}, q_{1}\right) \equiv \frac{\pi_{1}\left(q_{1}\right)}{T} E_{T}\left(a_{1} \mid q_{1}\right)$ and $D_{\hat{T}}\left(a_{2}, q_{2}\right) \equiv \frac{\pi_{2}\left(q_{2}\right)}{T} D_{\hat{T}}\left(a_{2} \mid q_{2}\right)$.

We can extend this state by taking $n_{1}$ i.i.d copies of the system in $T$, and $n_{2}$ i.i.d copies of the system in $\hat{T}$, obtaining the following assignment for the objective variable which is optimal and satisfies all the constraints in $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$,

$$
\begin{aligned}
& \rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right) \\
& =E_{T_{1}}\left(a_{1}^{(1)}, q_{1}^{(1)}\right) \otimes \ldots \otimes E_{T_{n_{1}}}\left(a_{1}^{\left(n_{1}\right)}, q_{1}^{\left(n_{1}\right)}\right) \otimes D_{\hat{T}_{1}}\left(a_{2}^{(1)}, q_{2}^{(1)}\right) \otimes \ldots \otimes D_{\hat{T}_{n_{2}}}\left(a_{2}^{\left(n_{2}\right)}, q_{2}^{\left(n_{2}\right)}\right) \otimes \rho_{S \hat{S}}^{\mathrm{T}},
\end{aligned}
$$

where $a_{i}^{(j)}\left(q_{i}^{(j)}\right)$ is the $j$-th element in the vector $a_{i}^{n_{i}}\left(q_{i}^{n_{i}}\right)$. We can now make use of the explicit form of the optimal variable to derive the NPA constraints for the SDP relaxations. The highest level of the NPA hierarchy we can fully implement is given by the minimum between $n_{1}$ and $n_{2}$, and in the following we assume without loss of generality that $n_{1} \geq n_{2}$. For a given $k \leq n_{2}$, the element of the NPA matrix $\Gamma_{k}$ which we can express as a function of the optimisation variable are of the form,

$$
\begin{align*}
& p\left(a_{1}^{m}, a_{2}^{\ell} \mid q_{1}^{m}, q_{2}^{\ell}\right)  \tag{C.14}\\
& =\operatorname{Tr}\left[\left(E_{T}\left(a_{1}^{(1)} \mid q_{1}^{(1)}\right) \ldots E_{T}\left(a_{1}^{(m)} \mid q_{1}^{(m)}\right) \otimes D_{\hat{T}}\left(a_{2}^{(1)} \mid q_{2}^{(1)}\right) \ldots D_{\hat{T}}\left(a_{2}^{(\ell)} \mid q_{2}^{(\ell)}\right)\right) \rho_{T \hat{T}}\right]
\end{align*}
$$

where $m, \ell \leq k$. The other elements of the matrix are either zeros, if we make the additional assumption that $\left\{E_{T}\left(a_{1} \mid q_{1}\right)\right\}_{a_{1}}$ and $\left\{D_{\hat{T}}\left(a_{2} \mid q_{2}\right)\right\}_{a_{2}}$ are projective measurements (PVMs) for each $q_{1}$ and $q_{2}$ respectively, or they need to be considered as new variables of the problem.

To rewrite the elements of the NPA matrix in Eq. (C.14) in terms of $\rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)$ we first need to introduce the following lemma which generalises Lemma 21 in the main text.

Lemma 37. Consider a set of operators $\left\{M_{i}\right\}_{i=1}^{n}$, each of them acting over the Hilbert space $\mathcal{H}$. Then, it holds that

$$
\begin{equation*}
\operatorname{Tr}\left[P_{\text {cyclic }}^{n}\left(M_{1} \otimes M_{2} \otimes \ldots \otimes M_{n}\right)\right]=\operatorname{Tr}\left[M_{1} M_{2} \ldots M_{n}\right] \tag{C.15}
\end{equation*}
$$

where $P_{\text {cyclic }}^{n} \in \mathcal{B}(\mathcal{H})$ is the unitary operator associated with the cyclic permutation $\pi$, which acts over the $n$-tuple as $\pi(1,2,3, \ldots, n)=(2,3, \ldots, n, 1)$.

Proof. The lemma is proven by explicitly computing the RHS and LHS of Eq. (C.15). For the RHS we have,

$$
\begin{aligned}
\operatorname{Tr}\left[M_{1} M_{2} \ldots M_{n}\right] & =\operatorname{Tr}\left[\sum_{i_{1}, j_{1}, i_{2}, j_{2}, \ldots, i_{n}, j_{n}} m_{i_{1}, j_{1}}^{(1)} m_{i_{2}, j_{2}}^{(2)} \ldots m_{i_{n}, j_{n}}^{(n)}\left|i_{1}\right\rangle\left\langle j_{1} \mid i_{2}\right\rangle\left\langle j_{2}\right| \ldots\left|i_{n}\right\rangle\left\langle j_{n}\right|\right] \\
& =\sum_{i_{1}, j_{1}, i_{2}, j_{2}, \ldots, i_{n}, j_{n}} m_{i_{1}, j_{1}}^{(1)} m_{i_{2}, j_{2}}^{(2)} \ldots m_{i_{n}, j_{n}}^{(n)} \delta_{j_{1}, i_{2}} \delta_{j_{2}, i_{3}} \ldots \delta_{j_{n}, i_{1}} \\
& =\sum_{i_{1}, i_{2}, \ldots, i_{n}} m_{i_{1}, i_{2}}^{(1)} m_{i_{2}, i_{3}}^{(2)} \ldots m_{i_{n}, i_{1}}^{(n)} .
\end{aligned}
$$

To compute the LHS, we first need the explicit form of the operator $P_{\text {cyclic }}^{n}$,

$$
P_{\text {cyclic }}^{n}=\sum_{k_{1}, k_{2}, k_{3}, k_{4}, \ldots, k_{n}}\left|k_{2}\right\rangle\left\langle\left. k_{1}\right|_{1} \otimes \mid k_{3}\right\rangle\left\langle\left. k_{2}\right|_{2} \otimes \mid k_{4}\right\rangle\left\langle\left. k_{3}\right|_{3} \otimes \ldots \otimes \mid k_{1}\right\rangle\left\langle\left. k_{n}\right|_{n},\right.
$$

which takes a vector on the subsystem $k$ and maps it to the subsystem $k-1$, with the exception of $k=1$ which is mapped into the $n$-th subsystem. The LHS of Eq. (C.15) is then

$$
\begin{aligned}
& \operatorname{Tr}\left[P_{\text {cyclic }}^{n}\left(M_{1} \otimes M_{2} \otimes \ldots \otimes M_{n}\right)\right] \\
& =\sum_{i_{1}, j_{1}, i_{2}, j_{2}, \ldots, i_{n}, j_{n}} m_{i_{1}, j_{1}}^{(1)} m_{i_{2}, j_{2}}^{(2)} \ldots m_{i_{n}, j_{n}}^{(n)} \operatorname{Tr}\left[P_{\text {cyclic }}^{n}\left|i_{1}\right\rangle\left\langle\left. j_{1}\right|_{1} \otimes \mid i_{2}\right\rangle\left\langle\left. j_{2}\right|_{2} \otimes \ldots \otimes \mid i_{n}\right\rangle\left\langle\left. j_{n}\right|_{n}\right]\right. \\
& =\sum_{i_{1}, j_{1}, i_{2}, j_{2}, \ldots, i_{n}, j_{n}} m_{i_{1}, j_{1}}^{(1)} m_{i_{2}, j_{2}}^{(2)} \ldots m_{i_{n}, j_{n}}^{(n)} \operatorname{Tr}\left[\left|i_{2}\right\rangle\left\langle\left. j_{1}\right|_{1} \otimes \mid i_{3}\right\rangle\left\langle\left. j_{2}\right|_{2} \otimes \ldots \otimes \mid i_{1}\right\rangle\left\langle\left. j_{n}\right|_{n}\right]\right. \\
& =\sum_{i_{1}, j_{1}, i_{2}, j_{2}, \ldots, i_{n}, j_{n}} m_{i_{1}, j_{1}}^{(1)} m_{i_{2}, j_{2}}^{(2)} \ldots m_{i_{n}, j_{n}}^{(n)} \delta_{i_{2}, j_{1}} \delta_{i_{3}, j_{2}} \ldots \delta_{i_{1}, j_{n}}=\sum_{i_{1}, i_{2}, \ldots, i_{n}} m_{i_{1}, i_{2}}^{(1)} m_{i_{2}, i_{3}}^{(2)} \ldots m_{i_{n}, i_{1}}^{(n)},
\end{aligned}
$$

which concludes the proof.

We can now derive the map between the elements of $\Gamma_{k}$ and the optimisation variable in $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$. Without loss of generality, let us assume that $m \geq \ell$ in Eq. (C.14); by
reordering the operators we get

$$
\begin{aligned}
& p\left(a_{1}^{m}, a_{2}^{\ell} \mid q_{1}^{m}, q_{2}^{\ell}\right) \\
& =\operatorname{Tr}\left[\left(\left(\prod_{i=1}^{\ell} E_{T}\left(a_{1}^{(i)} \mid q_{1}^{(i)}\right) \otimes D_{\hat{T}}\left(a_{2}^{(i)} \mid q_{2}^{(i)}\right)\right)\left(\prod_{j=\ell+1}^{m} E_{T}\left(a_{1}^{(j)} \mid q_{1}^{(j)}\right) \otimes \mathbb{1}_{\hat{T}}\right)\right) \rho_{T \hat{T}}\right] \\
& =\operatorname{Tr}\left[P_{\text {cyclic }}^{m+1}\left(\bigotimes_{i=1}^{\ell} E_{T_{i}}\left(a_{1}^{(i)} \mid q_{1}^{(i)}\right) \otimes D_{\hat{T}_{i}}\left(a_{2}^{(i)} \mid q_{2}^{(i)}\right)\right) \otimes\left(\bigotimes_{j=\ell+1}^{m} E_{T_{j}}\left(a_{1}^{(j)} \mid q_{1}^{(j)}\right) \otimes \mathbb{1}_{\hat{T}_{j}}\right) \otimes \rho_{S \hat{S}}\right],
\end{aligned}
$$

where the second equality follows from Lemma 37 with $\mathcal{H}_{T \hat{T}}$. The operator in the above equation is close to the optimisation variable $\rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)$, whose dimension can be reduced by summing over the classical variables while tracing out the quantum degrees of freedom

$$
\begin{aligned}
& \rho_{T^{m} \hat{T}^{m} S \hat{S}}\left(a_{1}^{m}, a_{2}^{\ell}, q_{1}^{m}, q_{2}^{\ell}\right) \\
& =\sum_{\substack{a_{1}^{(m+1)}, q_{1}^{(m+1)}, \ldots, a_{1}^{\left(n_{1}\right)}, q_{1}^{\left(n_{1}\right)}, a_{2}^{(++1)}, q_{2}^{(L+1)}, \ldots, a_{2}^{\left(n_{2}\right)}, q_{2}^{\left(n_{2}\right)}}} \operatorname{Tr}_{T_{m+1} \ldots T_{n_{1}} \hat{T}_{m+1} \ldots \hat{T}_{n_{2}}}\left[\rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(a_{1}^{n_{1}}, a_{2}^{n_{2}}, q_{1}^{n_{1}}, q_{2}^{n_{2}}\right)\right] \\
& =\left(\bigotimes_{i=1}^{\ell} E_{T_{i}}\left(a_{1}^{(i)}, q_{1}^{(i)}\right) \otimes D_{\hat{T}_{i}}\left(a_{2}^{(i)}, q_{2}^{(i)}\right)\right) \otimes\left(\bigotimes_{j=\ell+1}^{m} E_{T_{j}}\left(a_{1}^{(j)}, q_{1}^{(j)}\right) \otimes \frac{\mathbb{1}_{\hat{T}_{j}}}{T}\right) \otimes \rho_{S \hat{S}}^{\mathrm{T}},
\end{aligned}
$$

where we have used the fact that $\sum_{a_{1}, q_{1}} E_{T}\left(a_{1}, q_{1}\right)=\frac{1_{T}}{T}$ and $\sum_{a_{2}, q_{2}} E_{\hat{T}}\left(a_{2}, q_{2}\right)=\frac{1_{\hat{T}}}{T}$. By combining together the two equations above, we find that the elements of $\Gamma_{k}$ can be expressed in terms of the optimisation variable as

$$
\begin{aligned}
& p\left(a_{1}^{m}, a_{2}^{\ell} \mid q_{1}^{m}, q_{2}^{\ell}\right) \\
& = \begin{cases}\frac{T^{2 m}}{\prod_{i=1}^{m} \pi_{1}\left(q_{1}^{(i)}\right) \prod_{j=1}^{\ell} \pi_{2}\left(q_{2}^{(j)}\right)} \operatorname{Tr}\left[\left(P_{\text {cyclic }}^{m+1}\right)^{\mathrm{T}_{S} \hat{S}} \rho_{T^{m} \hat{T}^{m} S \hat{S}}\left(a_{1}^{m}, a_{2}^{\ell}, q_{1}^{m}, q_{2}^{\ell}\right)\right], & \forall \ell, m: \ell \leq m \leq k \\
\frac{T^{2 \ell}}{\prod_{i=1}^{m} \pi_{1}\left(q_{1}^{(i)}\right) \prod_{j=1}^{\ell} \pi_{2}\left(q_{2}^{(j)}\right)} \operatorname{Tr}\left[\left(P_{\text {cyclic }}^{\ell+1}\right)^{\mathrm{T}}{ }^{\hat{S} \hat{S}} \rho_{T^{\ell} \hat{T}^{\ell} S \hat{S}}\left(a_{1}^{m}, a_{2}^{\ell}, q_{1}^{m}, q_{2}^{\ell}\right)\right], & \forall \ell, m: m \leq \ell \leq k,\end{cases}
\end{aligned}
$$

where we have included the case in which $\ell \geq m$, that can be derived analogously.
When the NPA constraints are applied to the original $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ in the main text, the optimization variable is given by $\rho_{T(\hat{T})^{n_{1}}(S \hat{S})^{n_{2}}}\left(a_{1}, a_{2}^{n_{1}}, q_{1}, q_{2}^{n_{1}}\right)$. The relation between this variable and the entries of the NPA matrix $\Gamma_{k}$ can be obtained by following the same procedure presented in this appendix. The main difference is that the optimisation variable is defined
over a single subsystem $T$. As a result, when $\ell \geq 2$ we need to pad the variable with the $\ell-1$ copies of the maximally-mixed state, so as to be able to apply the operator $P_{\text {cyclic }}^{\ell+1}$. The relation between entries of $\Gamma_{k}$ and the optimisation variable of the original $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ is thus given by,

$$
\begin{aligned}
& p\left(a_{1}, a_{2}^{\ell} \mid q_{1}, q_{2}^{\ell}\right) \\
& =\frac{T^{2 \ell}}{\pi_{1}\left(q_{1}\right) \prod_{j=1}^{\ell} \pi_{2}\left(q_{2}^{(j)}\right)} \operatorname{Tr}\left[\left(P_{\text {cyclic }}^{\ell+1}\right)^{\mathrm{T}} \hat{S}\left(\bigotimes_{i=2}^{\ell} \mathbb{1}_{T_{i}}\right) \otimes \rho_{T \hat{T}^{\ell} S \hat{S}^{\prime}}\left(a_{1}, a_{2}^{\ell}, q_{1}, q_{2}^{\ell}\right)\right] \quad \forall \ell: \ell \leq k .
\end{aligned}
$$

## C. 5 No-signalling value

Apart from the classical and quantum values defined in Eqs. (7.2) and (7.3) respectively, we can define another quantity called the no-signalling value. This is the optimal winning probability achieved by no-signalling correlations such as

$$
\omega_{\mathrm{NS}}(V, \pi):=\max _{p \in \mathrm{NS}} \sum_{q_{1}, q_{2}} \pi\left(q_{1}, q_{2}\right) \sum_{a_{1}, a_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right),
$$

where NS denotes the set of all no-signalling correlations such that

$$
\sum_{a_{1}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=p\left(a_{2} \mid q_{2}\right) \forall a_{2}, q_{2} \text { and } \sum_{a_{2}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=p\left(a_{1} \mid q_{1}\right) \forall a_{1}, q_{1}
$$

As any classical or quantum correlation satisfies the no-signalling condition, the no-signalling value gives an upper bound to the classical and quantum values

$$
\omega_{C}(V, \pi) \leq \omega_{Q(T)}(V, \pi) \leq \omega_{Q}(V, \pi) \leq \omega_{\mathrm{NS}}(V, \pi)
$$

In this section, we show that $\operatorname{sdp}_{1,1}(V, \pi, T)$ is equal to $\omega_{\mathrm{NS}}(V, \pi)$ for any $T$.
Lemma 38. Let $\operatorname{sdp}_{1,1}(V, \pi, T)$ be the first-level SDP relaxation for the game with $V$ and $\pi\left(q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)$. Then, we have for all $T$ that

$$
\operatorname{sdp}_{1,1}(V, \pi, T)=\omega_{\mathrm{NS}}(V, \pi)
$$

Proof. We first show that $\operatorname{sdp}_{1,1}(V, \pi, T) \leq \omega_{\mathrm{NS}}(V, \pi)$. From $\operatorname{sdp}_{1,1}(V, \pi, T)$ we have the linear constraint

$$
\sum_{a_{2}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)=\pi_{2}\left(q_{2}\right) \frac{1_{\hat{T}}}{T} \otimes \rho_{T S \hat{S}}\left(a_{1}, q_{1}\right)
$$

Then, using the expression for $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ in Eq. (7.36), we can write the sum of $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$ over $a_{2}$ as

$$
\begin{aligned}
\sum_{a_{2}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right) & =\frac{T^{2}}{\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\sum_{a_{2}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right)\right] \\
& =\frac{T}{\pi_{1}\left(q_{1}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\mathbb{1}_{\hat{T}} \otimes \rho_{T S \hat{S}}\left(a_{1}, q_{1}\right)\right)\right]
\end{aligned}
$$

Using a more general formula for $p\left(a_{1} \mid q_{1}\right)$ without assuming the no-signalling, we obtain

$$
\begin{aligned}
p\left(a_{1} \mid q_{1}\right) & =\sum_{a_{2}, q_{2}} p\left(q_{2}\right) p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)=\frac{T^{2}}{\pi_{1}\left(q_{1}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\sum_{a_{2}, q_{2}} \rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\right)\right] \\
& =\frac{T^{2}}{\pi_{1}\left(q_{1}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\sum_{q_{2}} \pi_{2}\left(q_{2}\right) \frac{\mathbb{1}_{\hat{T}}}{T} \otimes \rho_{T S \hat{S}}\left(a_{1}, q_{1}\right)\right)\right] \\
& =\frac{T}{\pi_{1}\left(q_{1}\right)} \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(\mathbb{1}_{\hat{T}} \otimes \rho_{T S \hat{S}}\left(a_{1}, q_{1}\right)\right)\right],
\end{aligned}
$$

which is same as the expression for $\sum_{a_{2}} p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$. Similarly, we can also show the nosignalling condition for $a_{1}$ using the linear constraint on $A_{1}$ in $\operatorname{sdp}_{1,1}(V, \pi, T)$. Thus, all the states satisfying the linear constraints of $\operatorname{sdp}_{1,1}(V, \pi, T)$ are no-signalling and hence form a smaller set for optimisation than the set of no-signalling correlations.

Next, we show that $\operatorname{sdp}_{1,1}(V, \pi, T) \geq \omega_{\mathrm{NS}}(V, \pi)$. For any no-signalling correlation $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$, we can construct the Ansatz state

$$
\rho_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}, q_{1}, q_{2}\right)=\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right) p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right) \frac{\mathbb{1}_{T \hat{T}}}{T^{2}} \otimes \rho_{S \hat{S}},
$$

where $\rho_{S \hat{S}}$ is an arbitrary state. We can easily check that this state is a valid feasible state of
$\operatorname{sdp}_{1,1}(V, \pi, T)$, i.e. it satisfies all conditions in $\operatorname{sdp}_{1,1}(V, \pi, T)$. The objective function becomes

$$
\begin{aligned}
& \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right) p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right) \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}} \mathbb{1}_{T \hat{T}} \otimes \rho_{S \hat{S}}\right] \\
= & \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right) p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right),
\end{aligned}
$$

which is a success probability of the game with the strategy $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right)$. This implies that the feasible states of $\operatorname{sdp}_{1,1}(V, \pi, T)$ can cover all no-signalling cases, and hence $\operatorname{sdp}_{1,1}(V, \pi, T) \geq$ $\omega_{\mathrm{NS}}(V, \pi)$.

## C. 6 Rank-one projective measurements

## C.6.1 Adapted SDP hierarchies

When implementing the SDP relaxations, it is always advantageous if we take into account any additional structure available in order to reduce the computational resources for the implementation. In particular, when considering rank-one projective measurements, we can connect our results to the previous work [108] covering such cases. To exemplify this, we consider here the restricted case $A_{1}=A_{2}=\{0,1\}$ with $|T|=2$.

Let us recall the expression of the optimal winning probability $\omega_{Q(T)}(V, \pi)$ after the swap trick

$$
\omega_{Q(2)}(V, \pi)=\max _{(E, D, \rho)} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\Phi_{T \hat{T} \mid S \hat{S}}\left(E_{T}\left(a_{1}, q_{1}\right) \otimes D_{\hat{T}}\left(a_{2}, q_{2}\right) \otimes \rho_{S \hat{S}}\right)\right],
$$

where we redefine $E_{T}\left(a_{1}, q_{1}\right) \equiv \pi_{1}\left(q_{1}\right) E_{T}\left(a_{1} \mid q_{1}\right)$ and similarly for $D_{\hat{T}}\left(a_{2}, q_{2}\right)$. Using that $E_{T}\left(1 \mid q_{1}\right)=\mathbb{1}_{T}-E_{T}\left(0 \mid q_{1}\right)$ and $D_{\hat{T}}\left(1 \mid q_{2}\right)=\mathbb{1}_{\hat{T}}-D_{\hat{T}}\left(0 \mid q_{2}\right)$, we have

$$
\omega_{Q(2)}(V, \pi)=\max _{(E, D, \rho)} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\Lambda_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}\right)\left(E_{T}\left(q_{1}\right) \otimes D_{\hat{T}}\left(q_{2}\right) \otimes \rho_{S \hat{S}}\right)\right]
$$

where $E_{T}\left(q_{1}\right):=E_{T}\left(0, q_{1}\right)$ and $D_{\hat{T}}\left(q_{2}\right):=D_{\hat{T}}\left(0, q_{2}\right)$, and we defined the matrices

$$
\Lambda_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}\right):=\left(a_{1} \mathbb{1}_{T S}+(-1)^{a_{1}} \Phi_{T \mid S}\right) \otimes\left(a_{2} \mathbb{1}_{\hat{T} \hat{S}}+(-1)^{a_{2}} \Phi_{\hat{T} \mid \hat{S}}\right)
$$

When converting the $Q_{1}, Q_{2}$ systems to diagonal quantum systems, we obtain

$$
\omega_{Q(2)}(V, \pi)=\max _{(E, D, \rho)} \sum_{a_{1}, a_{2}} \operatorname{Tr}\left[\left(V_{Q_{1} Q_{2}}\left(a_{1}, a_{2}\right) \otimes \Lambda_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}\right)\right)\left(E_{Q_{1} T} \otimes D_{Q_{2} \hat{T}} \otimes \rho_{S \hat{S}}\right)\right]
$$

s.t. $\quad \rho_{S \hat{S}} \geq 0, \quad \operatorname{tr}\left[\rho_{S \hat{S}}\right]=1$

$$
\begin{array}{ll}
E_{Q_{1} T}=\sum_{q_{1}} \pi_{1}\left(q_{1}\right)\left|q_{1}\right\rangle\left\langle\left. q_{1}\right|_{Q_{1}} \otimes E_{T}\left(0 \mid q_{1}\right),\right. & \operatorname{Tr}_{T}\left[E_{Q_{1} T}\right]=\sum_{q_{1}} \pi_{1}\left(q_{1}\right)\left|q_{1}\right\rangle\left\langle\left. q_{1}\right|_{Q_{1}}\right. \\
D_{Q_{2} \hat{T}}=\left.\sum_{q_{2}} \pi_{2}\left(q_{2}\right)\left|q_{2}\right\rangle q_{2}\right|_{Q_{2}} \otimes D_{\hat{T}}\left(0 \mid q_{2}\right), & \operatorname{Tr}_{\hat{T}}\left[D_{Q_{2} \hat{T}}\right]=\sum_{q_{2}} \pi_{2}\left(q_{2}\right)\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}}\right.
\end{array}
$$

where $V_{Q_{1} Q_{2}}\left(a_{1}, a_{2}\right):=\sum_{q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)\left|q_{1}, q_{2}\right\rangle\left\langle q_{1},\left.q_{2}\right|_{Q_{1}, Q_{2}}\right.$. We used the fact that the trace of a rank-one projector is one. This then motivates the SDP hierarchy

$$
\begin{aligned}
& \overline{\operatorname{sdp}}_{n_{1}, n_{2}}^{\operatorname{proj}}(V, \pi, 2):=\max _{\rho} \sum_{a_{1}, a_{2}} \operatorname{Tr}\left[\left(V_{Q_{1} Q_{2}}\left(a_{1}, a_{2}\right) \otimes \Lambda_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}\right)\right) \rho_{Q_{1} T Q_{2} \hat{T} S \hat{S}}\right] \\
& \text { s.t. } \rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2}} S \hat{S}} \geq 0, \quad \operatorname{Tr}\left[\rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2}} S \hat{S}}\right]=1 \\
& \rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2}} S \hat{S}} \text { perm. inv. on }\left(Q_{1} T\right)^{n_{1}} \text { w.r.t. }\left(Q_{2} \hat{T}\right)^{n_{2}} S \hat{S} \\
& \rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2}} S \hat{S}} \text { perm. inv. on }\left(Q_{2} \hat{T}\right)^{n_{2}} \text { w.r.t. }\left(Q_{1} T\right)^{n_{1}} S \hat{S} \\
& \operatorname{Tr}_{T}\left[\rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2} S \hat{S}}}\right]=\left.\sum_{q_{1}} \pi\left(q_{1}\right)\left|q_{1}\right\rangle q_{1}\right|_{Q_{1}} \otimes \operatorname{Tr}_{Q_{1} T}\left[\rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2} S \hat{S}}}\right] \\
& \operatorname{Tr}_{\hat{T}}\left[\rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2} S \hat{S}}}\right]=\sum_{q_{2}} \pi\left(q_{2}\right)\left|q_{2}\right\rangle\left\langle\left. q_{2}\right|_{Q_{2}} \otimes \operatorname{Tr}_{Q_{2} \hat{T}}\left[\rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2}} S \hat{S}}\right]\right. \\
& \rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2} S \hat{S}}} \geq 0, \quad \rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2} S \hat{S}}}^{T_{2}} \geq 0, \quad \rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2} S \hat{S}}}^{T_{S}} \geq 0, \ldots
\end{aligned}
$$

We note that we do no longer have the $T^{2}$ pre-factor and that the size of the optimisation variable is smaller compared to $\overline{\operatorname{sdp}}_{n_{1}, n_{2}}(V, \pi, T)$ in Eq. (C.12). This allows for more efficient numerical implementations.

In fact, again writing out the block-diagonal structure explicitly

$$
\rho_{\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})}=: \sum_{q_{1}^{n_{1}, q_{2}^{n_{2}}}}\left|q_{1}^{n_{1}}, q_{2}^{n_{2}}\right\rangle\left\langle q_{1}^{n_{1}}, q_{2}^{n_{2}}\right| \otimes \rho_{T^{n_{1}} \hat{T}^{n_{2}} S \hat{S}}\left(q_{1}^{n_{1}}, q_{2}^{n_{2}}\right),
$$

and taking $n_{1}=n_{2}=|Q|$, one can check that the objective function can be expressed in terms of the single renormalised block

$$
W_{T|Q| \hat{T}^{|Q|} \mid \hat{S}}:=\frac{\rho_{T^{|Q| \hat{T}}|Q| S \hat{S}}(1,2, \cdots,|Q|, 1,2, \cdots,|Q|)}{\pi_{1}(1) \pi_{1}(2) \ldots \pi_{1}(|Q|) \pi_{2}(1) \pi_{2}(2) \ldots \pi_{2}(|Q|)}
$$

where $\pi_{1}\left(q_{1}\right) \pi_{2}\left(q_{2}\right)=\pi\left(q_{1}, q_{2}\right)$. Then, for $|Q|$ questions on each side and, e.g., the uniform distribution $\pi_{\text {unif }}=\frac{1}{|Q|^{2}}$ for the questions, ignoring all other blocks and only enforcing the positivity, normalisation, and PPT constraints give the outer relaxation such that

$$
\begin{equation*}
\overline{\operatorname{sdp}}_{|Q|}^{\mathrm{proj}}\left(V, \pi_{\text {unif }}, 2\right) \leq \frac{1}{|Q|^{2}} \cdot \operatorname{sdp}_{\mathrm{PPT}}(V) \tag{C.16}
\end{equation*}
$$

with the SDP from [108, Equation (6)]

$$
\begin{aligned}
\operatorname{sdp}_{\mathrm{PPT}}(V) & :=\max _{W} \sum_{a_{1}, a_{2}, q_{1}, q_{2}} V\left(a_{1}, a_{2}, q_{1}, q_{2}\right) \operatorname{Tr}\left[\Lambda_{T \hat{T} S \hat{S}}\left(a_{1}, a_{2}\right) W_{T_{q_{1}} \hat{T}_{q_{2}} \hat{S}}\right] \\
\text { s.t. } & W_{T|Q| \hat{T}|Q| S \hat{S}} \geq 0, \quad \operatorname{Tr}\left[W_{T^{|Q|}|\hat{T}| Q \mid S \hat{S}}\right]=1 \\
& W_{T|Q| Q \mid}^{T_{T}|Q| S \hat{S}} \geq 0, W_{T|Q| \hat{T}|Q| S \hat{S}} \geq 0, \ldots \text { (PPT constraints), }
\end{aligned}
$$

where the $|Q|^{-2}$ coefficient arises as we are considering a joint probability distribution $p\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$ rather than a conditional one $p\left(a_{1}, a_{2} \mid q_{1}, q_{2}\right) .{ }^{1}$ Hence, in this scenario for $|Q|$ questions on each side, the $|Q|$-th level of our hierarchy is never a worse upper bound than the previously studied $\operatorname{sdp}_{\mathrm{PPT}}(V)$.

[^39]
## C.6.2 Asymptotic convergence analysis

Even in this specific setting with $|A|=|T|=2$, an advantage of our techniques compared to the previous $\operatorname{sdp}_{\mathrm{PPT}}(V)$ and its extendibility extensions discussed in [108] is that for a slight variation we can give an approximation scheme scaling polynomially in $|Q|$. Namely, when extending as $\left(Q_{1} T\right)\left(Q_{2} \hat{T}\right)^{n_{1}}(S \hat{S})^{n_{2}}$ instead of $\left(Q_{1} T\right)^{n_{1}}\left(Q_{2} \hat{T}\right)^{n_{2}}(S \hat{S})$, we get for the resulting SDP hierarchy denoted by $\operatorname{sdp}_{n_{1}, n_{2}}^{\text {proj }}(V, \pi, 2)$ that

$$
\left|\operatorname{sdp}_{n_{1}, n_{2}}^{\text {proj }}(V, \pi, 2)-\omega_{Q(2)}(V, \pi)\right| \leq 128 \cdot \sqrt{\frac{17}{n_{1}}+\frac{1}{n_{2}}} .
$$

The asymptotic convergence analysis is done similarly as for the general $\operatorname{sdp}_{n_{1}, n_{2}}(V, \pi, T)$ in the main text.

## Bibliography

[1] Paulina Marian, Tudor A Marian, and Horia Scutaru. Bures distance as a measure of entanglement for two-mode squeezed thermal states. Physical Review A, 68(6):062309, 2003.
[2] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM review, 41(2):303-332, 1999.
[3] Lov K Grover. Quantum mechanics helps in searching for a needle in a haystack. Physical review letters, 79(2):325, 1997.
[4] Charles H Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. arXiv preprint arXiv:2003.06557, 2020.
[5] John Preskill. Quantum computing in the NISQ era and beyond. Quantum, 2:79, 2018.
[6] Seth Lloyd and Samuel L Braunstein. Quantum computation over continuous variables. In S.L. Braunstein and A.K. Pati, editors, Quantum information with continuous variables, pages 9-17. Springer, Dordrecht, 1999.
[7] Stephen D Bartlett and Barry C Sanders. Universal continuous-variable quantum computation: requirement of optical nonlinearity for photon counting. Physical Review A, 65(4):042304, 2002.
[8] M Ohliger, K Kieling, and J Eisert. Limitations of quantum computing with Gaussian cluster states. Physical Review A, 82(4):042336, 2010.
[9] Julien Niset, Jaromír Fiurášek, and Nicolas J Cerf. No-go theorem for Gaussian quantum error correction. Physical Review Letters, 102(12):120501, 2009.
[10] Ludovico Lami, Bartosz Regula, Xin Wang, Rosanna Nichols, Andreas Winter, and Gerardo Adesso. Gaussian quantum resource theories. Physical Review A, 98(2):022335, 2018.
[11] Stefano Pirandola, Ulrik L Andersen, Leonardo Banchi, Mario Berta, Darius Bunandar, Roger Colbeck, Dirk Englund, Tobias Gehring, Cosmo Lupo, Carlo Ottaviani, et al. Advances in quantum cryptography. Advances in optics and photonics, 12(4):1012-1236, 2020.
[12] Antonio Acín and Lluis Masanes. Certified randomness in quantum physics. Nature, 540(7632):213-219, 2016.
[13] Alessio Serafini. Quantum continuous variables: a primer of theoretical methods. CRC press, 2017.
[14] Nicolas J Cerf, Marc Levy, and Gilles Van Assche. Quantum distribution of Gaussian keys using squeezed states. Physical Review A, 63(5):052311, 2001.
[15] Frédéric Grosshans and Philippe Grangier. Continuous variable quantum cryptography using coherent states. Physical Review Letters, 88(5):057902, 2002.
[16] Frédéric Grosshans, Gilles Van Assche, Jérôme Wenger, Rosa Brouri, Nicolas J Cerf, and Philippe Grangier. Quantum key distribution using Gaussian-modulated coherent states. Nature, 421(6920):238-241, 2003.
[17] Christian Weedbrook, Andrew M Lance, Warwick P Bowen, Thomas Symul, Timothy C Ralph, and Ping Koy Lam. Quantum cryptography without switching. Physical review letters, 93(17):170504, 2004.
[18] Andrew M Lance, Thomas Symul, Vikram Sharma, Christian Weedbrook, Timothy C Ralph, and Ping Koy Lam. No-switching quantum key distribution using broadband modulated coherent light. Physical review letters, 95(18):180503, 2005.
[19] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J Cerf, Timothy C Ralph, Jeffrey H Shapiro, and Seth Lloyd. Gaussian quantum information. Reviews of Modern Physics, 84(2):621, 2012.
[20] Samuel L Braunstein and Peter Van Loock. Quantum information with continuous variables. Reviews of modern physics, 77(2):513, 2005.
[21] W Forrest Stinespring. Positive functions on c*-algebras. Proceedings of the American Mathematical Society, 6(2):211-216, 1955.
[22] Filippo Caruso, Jens Eisert, Vittorio Giovannetti, and Alexander S Holevo. Multi-mode bosonic Gaussian channels. New Journal of Physics, 10(8):083030, 2008.
[23] Filippo Caruso, Jens Eisert, Vittorio Giovannetti, and Alexander S Holevo. Optimal unitary dilation for bosonic gaussian channels. Physical Review A, 84(2):022306, 2011.
[24] Alexander S Holevo and Reinhard F Werner. Evaluating capacities of bosonic gaussian channels. Physical Review A, 63(3):032312, 2001.
[25] John Williamson. On the algebraic problem concerning the normal forms of linear dynamical systems. American journal of mathematics, 58(1):141-163, 1936.
[26] Junaid Aasi, Joan Abadie, BP Abbott, Richard Abbott, TD Abbott, MR Abernathy, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikari, et al. Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light. Nature Photonics, 7(8):613-619, 2013.
[27] Ian D Leroux, Monika H Schleier-Smith, and Vladan Vuletić. Orientation-dependent entanglement lifetime in a squeezed atomic clock. Physical review letters, 104(25):250801, 2010.
[28] Anne Louchet-Chauvet, Jürgen Appel, Jelmer J Renema, Daniel Oblak, Niels Kjaergaard, and Eugene S Polzik. Entanglement-assisted atomic clock beyond the projection noise limit. New Journal of Physics, 12(6):065032, 2010.
[29] Eugene P Wigner. On the quantum correction for thermodynamic equilibrium. In Part I: Physical Chemistry. Part II: Solid State Physics, pages 110-120. Springer, 1997.
[30] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. Reviews of Modern Physics, 81(2):865-942, 2009.
[31] Victor Veitch, SA Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. New Journal of Physics, 16(1):013009, 2014.
[32] Mehdi Ahmadi, Hoan Bui Dang, Gilad Gour, and Barry C Sanders. Quantification and manipulation of magic states. Physical Review A, 97(6):062332, 2018.
[33] G Gour and RW Spekkens. Fundamental limitations for quantum and nanoscale thermodynamics. New Journal of Physics, 10:033023, 2008.
[34] Matteo Lostaglio, Kamil Korzekwa, David Jennings, and Terry Rudolph. Quantum coherence, time-translation symmetry, and thermodynamics. Physical Review X, 5(2):021001, 2015.
[35] Fernando Brandao, Michał Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. The second laws of quantum thermodynamics. Proceedings of the National Academy of Sciences, 112(11):3275-3279, 2015.
[36] Nelly Huei Ying Ng and Mischa Prebin Woods. Resource theory of quantum thermodynamics: Thermal operations and second laws. In Thermodynamics in the Quantum Regime, pages 625-650. Springer, 2018.
[37] Eric Chitambar and Gilad Gour. Quantum resource theories. Reviews of Modern Physics, 91(2):025001, 2019.
[38] Michael A Nielsen. Conditions for a class of entanglement transformations. Physical Review Letters, 83(2):436, 1999.
[39] Carlo Sparaciari. Multi-resource theories and applications to quantum thermodynamics. PhD thesis, UCL (University College London), 2018.
[40] Zi-Wen Liu and Andreas Winter. Resource theories of quantum channels and the universal role of resource erasure. arXiv preprint arXiv:1904.04201, 2019.
[41] Yunchao Liu and Xiao Yuan. Operational resource theory of quantum channels. Physical Review Research, 2(1):012035, 2020.
[42] Daniel Jonathan and Martin B Plenio. Entanglement-assisted local manipulation of pure quantum states. Physical Review Letters, 83(17):3566, 1999.
[43] Sadi Turgut. Catalytic transformations for bipartite pure states. Journal of Physics A: Mathematical and Theoretical, 40(40):12185, 2007.
[44] Matthew Klimesh. Inequalities that collectively completely characterize the catalytic majorization relation. arXiv preprint arXiv:0709.3680, 2007.
[45] Wim van Dam and Patrick Hayden. Universal entanglement transformations without communication. Physical Review A, 67(6):060302, 2003.
[46] Markus P Müller. Correlating thermal machines and the second law at the nanoscale. Physical Review X, 8(4):041051, 2018.
[47] Charles H Bennett, Herbert J Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. Physical Review A, 53(4):2046, 1996.
[48] Charles H Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A Smolin, and William K Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. Physical review letters, 76(5):722, 1996.
[49] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Inseparable two spin-1/2 density matrices can be distilled to a singlet form. Physical Review Letters, 78(4):574, 1997.
[50] Fernando GSL Brandao and Gilad Gour. Reversible framework for quantum resource theories. Physical review letters, 115(7):070503, 2015.
[51] Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Reversible transformations from pure to mixed states and the unique measure of information. Physical Review A, 67(6):062104, 2003.
[52] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? Physical Review Letters, 80(24):5239, 1998.
[53] Matteo Lostaglio, David Jennings, and Terry Rudolph. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. Nature communications, $6(1): 1-9,2015$.
[54] Qi Zhao, Yunchao Liu, Xiao Yuan, Eric Chitambar, and Andreas Winter. One-shot coherence distillation: towards completing the picture. IEEE Transactions on Information Theory, 65(10):6441-6453, 2019.
[55] Mark W Girard, Gilad Gour, and Shmuel Friedland. On convex optimization problems in quantum information theory. Journal of Physics A: Mathematical and Theoretical, 47(50):505302, 2014.
[56] Dénes Petz. Quasi-entropies for finite quantum systems. Reports on mathematical physics, 23(1):57-65, 1986.
[57] Carlo Sparaciari, Lídia Del Rio, Carlo Maria Scandolo, Philippe Faist, and Jonathan Oppenheim. The first law of general quantum resource theories. Quantum, 4:259, 2020.
[58] John Goold, Marcus Huber, Arnau Riera, Lídia Del Rio, and Paul Skrzypczyk. The role of quantum information in thermodynamics - a topical review. Journal of Physics A: Mathematical and Theoretical, 49(14):143001, 2016.
[59] Stephen D Bartlett, Terry Rudolph, and Robert W Spekkens. Reference frames, superselection rules, and quantum information. Reviews of Modern Physics, 79(2):555, 2007.
[60] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. Reviews of Modern Physics, 86(2):419, 2014.
[61] Lev Vaidman. Teleportation of quantum states. Physical Review A, 49(2):1473, 1994.
[62] Samuel L Braunstein and H Jeff Kimble. Teleportation of continuous quantum variables. Physical Review Letters, 80(4):869, 1998.
[63] Akira Furusawa, Jens Lykke Sørensen, Samuel L Braunstein, Christopher A Fuchs, H Jeff Kimble, and Eugene S Polzik. Unconditional quantum teleportation. Science, 282(5389):706-709, 1998.
[64] Nicolas C Menicucci, Peter Van Loock, Mile Gu, Christian Weedbrook, Timothy C Ralph, and Michael A Nielsen. Universal quantum computation with continuous-variable cluster states. Physical Review Letters, 97(11):110501, 2006.
[65] Jaromír Fiurášek. Gaussian transformations and distillation of entangled Gaussian states. Physical Review Letters, 89(13):137904, 2002.
[66] Géza Giedke and J Ignacio Cirac. Characterisation of Gaussian operations and distillation of Gaussian states. Physical Review A, 66(3):032316, 2002.
[67] Jens Eisert, Stefan Scheel, and Martin B Plenio. Distilling Gaussian states with Gaussian operations is impossible. Physical Review Letters, 89(13):137903, 2002.
[68] Daniel Gottesman, Alexei Kitaev, and John Preskill. Encoding a qubit in an oscillator. Physical Review A, 64(1):012310, 2001.
[69] Kyungjoo Noh, S. M. Girvin, and Liang Jiang. Encoding an oscillator into many oscillators. Physical Review Letters, 125(8):080503, 2020.
[70] Salman Beigi and Amin Gohari. $\Phi$-entropic measures of correlation. IEEE Transactions on Information Theory, 64(4):2193-2211, 2018.
[71] Ryuji Takagi and Quntao Zhuang. Convex resource theory of non-Gaussianity. Physical Review A, 97(6):062337, 2018.
[72] William K Wootters. Entanglement of formation of an arbitrary state of two qubits. Physical Review Letters, 80(10):2245, 1998.
[73] Armin Uhlmann. Roofs and convexity. Entropy, 12(7):1799-1832, 2010.
[74] Michael M Wolf, Geza Giedke, O Krüger, R F Werner, and J Ignacio Cirac. Gaussian entanglement of formation. Physical Review A, 69(5):052320, 2004.
[75] J Heersink, Ch Marquardt, R Dong, R Filip, S Lorenz, G Leuchs, and U L Andersen. Distillation of squeezing from non-Gaussian quantum states. Physical Review Letters, 96(25):253601, 2006.
[76] Benjamin Yadin, Felix C Binder, Jayne Thompson, Varun Narasimhachar, Mile Gu, and MS Kim. Operational resource theory of continuous-variable nonclassicality. Physical Review X, 8(4):041038, 2018.
[77] Daniel E Browne, Jens Eisert, Stefan Scheel, and Martin B Plenio. Driving non-Gaussian to Gaussian states with linear optics. Physical Review A, 67(6):062320, 2003.
[78] Jens Eisert, DE Browne, S Scheel, and MB Plenio. Distillation of continuous-variable entanglement with optical means. Annals of Physics, 311(2):431-458, 2004.
[79] A Serafini, M Lostaglio, S Longden, U Shackerley-Bennett, C-Y Hsieh, and G Adesso. Gaussian thermal operations and the limits of algorithmic cooling. Physical Review Letters, 124(1):010602, 2020.
[80] Varun Narasimhachar, Syed Assad, Felix C. Binder, Jayne Thompson, Benjamin Yadin, and Mile Gu. Thermodynamic resources in continuous-variable quantum systems. npj Quantum Information, 7(1):9, dec 2021.
[81] R Simon, N. Mukunda, and Biswadeb Dutta. Quantum-noise matrix for multimode systems: $\mathrm{U}(\mathrm{n})$ invariance, squeezing, and normal forms. Physical Review A, 49(3):15671583, mar 1994.
[82] Roger A. Horn and Charles R. Johnson. Matrix Analysis. Cambridge University Press, New York, USA, 1985.
[83] Paul Boes, Jens Eisert, Rodrigo Gallego, Markus P. Müller, and Henrik Wilming. Von Neumann entropy from unitarity. Physical Review Letters, 122(21):210402, may 2019.
[84] Wim van Dam and Patrick Hayden. Universal entanglement transformations without communication. Physical Review A, 67(6):060302, jun 2003.
[85] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. Inequalities: Theory of Majorization and Its Applications. Springer Series in Statistics. Springer New York, New York, NY, 2011.
[86] Esteban A. Martinez and Juan Pablo Paz. Dynamics and thermodynamics of linear quantum open systems. Physical Review Letters, 110(13):130406, mar 2013.
[87] Nahuel Freitas and Juan Pablo Paz. Fundamental limits for cooling of linear quantum refrigerators. Physical Review E, 95(1):012146, jan 2017.
[88] Tomáś Opatrný, Šimon Bräuer, Avijit Misra, Nilakantha Meher, Ofer Firstenberg, Eilon Poem, and Gershon Kurizki. Nonlinear interferometry enables coherent heat machine operation. arXiv:2108.10157, aug 2021.
[89] Michał Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. Nature Communications, 4(1):2059, oct 2013.
[90] Matteo G.A. Paris. Displacement operator by beam splitter. Physics Letters A, 217:78-80, 71996.
[91] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004.
[92] Leonid Gurvits. Classical deterministic complexity of Edmonds' problem and quantum entanglement. In Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, pages 10-19, 2003.
[93] Asher Peres. Separability criterion for density matrices. Physical Review Letters, 77(8):1413, 1996.
[94] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. Physics Letters A, 223(1):1-8, 1996.
[95] Andrew C Doherty, Pablo A Parrilo, and Federico M Spedalieri. Complete family of separability criteria. Physical Review A, 69(2):022308, 2004.
[96] Barbara M Terhal. Is entanglement monogamous? IBM Journal of Research and Development, 48(1):71-78, 2004.
[97] Mark Fannes, John T Lewis, and André Verbeure. Symmetric states of composite systems. Letters in Mathematical Physics, 15(3):255-260, 1988.
[98] GA Raggio and RF Werner. Quantum statistical mechanics of general mean field systems. Helvetica Acta Physica, 62(DIAS-STP-88-49):980, 1989.
[99] Andrew C Doherty, Pablo A Parrilo, and Federico M Spedalieri. Detecting multipartite entanglement. Physical Review A, 71(3):032333, 2005.
[100] Miguel Navascués, Stefano Pironio, and Antonio Acín. Bounding the set of quantum correlations. Physical Review Letters, 98(1):010401, 2007.
[101] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. New Journal of Physics, 10(7):073013, 2008.
[102] Stefano Pironio, Miguel Navascués, and Antonio Acin. Convergent relaxations of polynomial optimization problems with noncommuting variables. SIAM Journal on Optimization, 20(5):2157-2180, 2010.
[103] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP*= RE. arXiv:2001.04383, 2020.
[104] John Watrous. The theory of quantum information. Cambridge University Press, 2018.
[105] Matthias Christandl, Robert König, Graeme Mitchison, and Renato Renner. One-and-ahalf quantum de Finetti theorems. Communications in mathematical physics, 273(2):473498, 2007.
[106] John S Bell. On the Einstein Podolsky Rosen paradox. Physics Physique Fizika, 1(3):195, 1964.
[107] Rodrigo Gallego, Nicolas Brunner, Christopher Hadley, and Antonio Acín. Deviceindependent tests of classical and quantum dimensions. Physical Review Letters, 105(23):230501, 2010.
[108] Miguel Navascués, Gonzalo de la Torre, and Tamás Vértesi. Characterization of quantum correlations with local dimension constraints and its device-independent applications. Physical Review X, 4(1):011011, 2014.
[109] Miguel Navascués and Tamás Vértesi. Bounding the set of finite dimensional quantum correlations. Physical Review Letters, 115(2):020501, 2015.
[110] Miguel Navascués, Adrien Feix, Mateus Araújo, and Tamás Vértesi. Characterizing finite-dimensional quantum behavior. Physical Review A, 92(4):042117, 2015.
[111] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. Physical review letters, 23(15):880, 1969.
[112] Dong Yang, Karol Horodecki, Michal Horodecki, Pawel Horodecki, Jonathan Oppenheim, and Wei Song. Squashed entanglement for multipartite states and entanglement measures based on the mixed convex roof. IEEE Transactions on Information Theory, 55(7):33753387, 2009.
[113] Fernando GSL Brandão and Aram W Harrow. Quantum de Finetti theorems under local measurements with applications. Communications in Mathematical Physics, 353(2):469506, 2017.
[114] Fernando GSL Brandão and Aram W Harrow. Product-state approximations to quantum states. Communications in Mathematical Physics, 342(1):47-80, 2016.
[115] William Matthews, Stephanie Wehner, and Andreas Winter. Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding. Communications in Mathematical Physics, 291(3):813-843, 2009.
[116] Mario Berta, Francesco Borderi, Omar Fawzi, and Volkher Scholz. Semidefinite programming hierarchies for quantum error correction. arXiv:1810.12197, 2018.
[117] MOSEK. The MOSEK optimization toolbox for MATLAB manual. Version 8.1., 2017.
[118] Daniel Collins and Nicolas Gisin. A relevant two qubit Bell inequality inequivalent to the CHSH inequality. Journal of Physics A: Mathematical and General, 37(5):1775, 2004.
[119] Marcel Froissart. Constructive generalization of Bell's inequalities. Il Nuovo Cimento B, 64(241-251), 1981.
[120] Christine Bachoc, Dion C Gijswijt, Alexander Schrijver, and Frank Vallentin. Invariant semidefinite programs. In Handbook on semidefinite, conic and polynomial optimization, pages 219-269. Springer, 2012.
[121] Debbie Leung and William Matthews. On the power of PPT-preserving and non-signalling codes. IEEE Transactions on Information Theory, 61(8):4486-4499, 2015.
[122] Xin Wang, Kun Fang, and Runyao Duan. Semidefinite programming converse bounds for quantum communication. IEEE Transactions on Information Theory, 65(4):2583-2592, 2018.
[123] Omar Fawzi, Ala Shayeghi, and Hoang Ta. A hierarchy of efficient bounds on quantum capacities exploiting symmetry. IEEE Transactions on Information Theory, 2022.
[124] Barbara Kraus, Klemens Hammerer, Géza Giedke, and J Ignacio Cirac. Entanglement generation and Hamiltonian simulation in continuous-variable systems. Physical Review A, 67(4):042314, 2003.
[125] Robert C. Thompson. Singular values and diagonal elements of complex symmetric matrices. Linear Algebra and its Applications, 26(C):65-106, aug 1979.
[126] Ludovico Lami, Carlos Palazuelos, and Andreas Winter. Ultimate data hiding in quantum mechanics and beyond. Communications in Mathematical Physics, 361(2):661-708, 2018.
[127] Pawel Horodecki. Separability criterion and inseparable mixed states with positive partial transposition. Physics Letters A, 232:333, 1997.
[128] Charles H Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. Physical review letters, 70(13):1895, 1993.


[^0]:    $1 *$ : co-first authors

[^1]:    ${ }^{1}$ Even though we use the eigenstates of $\hat{x}$ and $\hat{p}$ as basis of $L^{2}(\mathbb{R})$ here, they are not parts of $L^{2}(\mathbb{R})$. They are Dirac delta distributions centred at their eigenvalues $x$ and $p$ in the position and momentum space respectively, which do not belong to $L^{2}(\mathbb{R})$.

[^2]:    ${ }^{2}$ This means that $\mathcal{H}$ has infinitely many but countable orthonormal basis.

[^3]:    ${ }^{3} \mathrm{~A} c$-number refers to the identity operator times a complex number.

[^4]:    ${ }^{4}$ Note that $\hat{D}_{\mathbf{r}}$ is an operator acting on the full Hilbert space and thus acts on each canonical operator $\hat{r}_{i}$.

[^5]:    ${ }^{5} O(2 N)$ refers to the orthogonal group in dimension $2 N$.

[^6]:    ${ }^{6}$ Such Hamiltonian does not have a positive definite Hamiltonian matrix, i.e., $H \ngtr 0$, and thus its thermal state is not well-defined.

[^7]:    ${ }^{7}$ Coherent states are eigenvectors of the ladder operator $\hat{a}$, i.e., $\hat{a}|\alpha\rangle=\alpha|\alpha\rangle$, and can be generated by applying the displacement operator to the vacuum state. They will be discussed in more details in Section 2.5.

[^8]:    ${ }^{8}$ For more details, please see [13, Section 3.3].

[^9]:    ${ }^{9}$ Any CM is a symmetric matrix, and thus $\sigma_{12}=\sigma_{21}$.

[^10]:    ${ }^{1}$ If the input and output Hilbert spaces are the same, i.e., $\mathcal{H}_{\text {in }}=\mathcal{H}_{\text {out }}=\mathcal{H}$, we simply denote it by $\mathcal{C}(\mathcal{H})$.

[^11]:    ${ }^{2}$ This is true when we assume that QRTs admit the tensor-product structure.

[^12]:    ${ }^{3}$ For the case of general bipartite entanglement, the free states $\mathcal{F}_{\text {min }}$ identified by LOCC are separable states.

[^13]:    ${ }^{1}$ Here, we assume that the input and output Hilbert spaces are the same, $L^{2}\left(\mathbb{R}^{N}\right)$ where $N \in \mathbb{N}$ is the number of modes (degrees of freedom).

[^14]:    ${ }^{2}$ It means that the set $\mathcal{F}$ is closed with respect to the topology induced by the trace norm.

[^15]:    ${ }^{3}$ For the definition of the convex hull of a set, see Section 6.1.

[^16]:    ${ }^{4}$ Any free state in $\mathcal{F}_{f}^{C}$ is a convex combination of free Gaussian states in $\mathcal{F}_{f}^{G}$.

[^17]:    ${ }^{5}$ It is worth noting that a similar squeezing distillation protocol has been reported in [75]. The difference is that our protocol is deterministic while the protocol in [75] is probabilistic.

[^18]:    ${ }^{6}$ As we studied in Section 2.5, two-mode squeezed states play the role of entangled states in the Gaussian framework.

[^19]:    ${ }^{1}$ In quantum thermodynamics, the free states are no longer invariant under displacement operations.

[^20]:    ${ }^{2}$ Principal mode temperatures have been already introduced in [80], but principal mode asymmetries are newly introduced in our work.

[^21]:    ${ }^{3}$ Note that diagonalisation of the CM with passive transformation is not in general possible, since $S p(2 N) \cap$ $O(2 N)$ is strictly smaller than the orthogonal group $O(2 N)$.

[^22]:    ${ }^{4}$ This means that the variance of any quadrature of this thermal state is $t$.

[^23]:    ${ }^{5}$ This is up to the symplectic transformation which brings the system Hamiltonian into the normal form, as we have discussed in Section 5.1.1.

[^24]:    ${ }^{6}$ Recall the parameters $\alpha$ are real and non-negative - see Section 5.1.2.

[^25]:    ${ }^{7}$ If $p_{21}=0$, then $\left|p_{22}\right|^{2}=1$ due to the catalytic conditions in Eq. (5.8), but we already assumed $\left|p_{22}\right|^{2} \neq 1$.

[^26]:    ${ }^{1} p(\boldsymbol{x})$ such that $p(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in S$ and $\int_{S} p(\boldsymbol{x}) \mathrm{d} \boldsymbol{x}=1$.

[^27]:    ${ }^{2}$ If the optimisation problem involves maximising the objective function and is infeasible, then $\omega=-\infty$.

[^28]:    ${ }^{3}$ Alternatively, one can define separable states as quantum states that can be written as convex combinations of general product states (not pure states). Since any mixed quantum state can be written as a convex combination of pure states, the two definitions are equivalent.

[^29]:    ${ }^{4}$ We have looked at bound entanglement in Section 3.2.2 in the context of entanglement distillation.

[^30]:    ${ }^{5}$ For example, they could be linear polynomials, such as $\sum_{i} a_{i} X_{i}$ for $a_{i} \in \mathbb{R} \forall i$, or quadratic polynomials, such as $\sum_{i}\left(a_{i} X_{i}+b_{i}\right)^{2}$ for $a_{i}, b_{i} \in \mathbb{R} \forall i$.

[^31]:    ${ }^{6}$ Thanks to Stinespring dilation theorem, any POVMs can be realised by a projective measurement on a larger Hilbert space, and any mixed quantum state can be viewed as a marginal state of a pure state in a larger system.
    ${ }^{7}$ The two conditions are known to be equivalent in finite dimensions, but it has been recently shown that this is no longer true in infinite dimensions [103].

[^32]:    ${ }^{8}$ It is worth noting that we have used the assumption that the measurements are composed by orthogonal projectors to set some of the entries of the $Q$ and $R$ matrices to zero. When this assumption cannot be made, one should replace these entries with new variables of the problem. We will see such example in the next chapter.

[^33]:    ${ }^{9} \boldsymbol{S}_{2}$ represents the set of all products in $\left\{E\left(a_{i}, q_{i}\right)\right\}$ up to the second order plus the identity.

[^34]:    ${ }^{1}$ A classical system can be seen as a quantum system with dimension 1.

[^35]:    ${ }^{2}$ For given two quantum states $\rho$ and $\sigma$ such that $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$, Pinsker's inequality states that

    $$
    \begin{equation*}
    \|\rho-\sigma\|_{1}^{2} \leq 2 \ln 2 \cdot D(\rho \| \sigma) . \tag{7.27}
    \end{equation*}
    $$

[^36]:    ${ }^{3}$ It is the marginal state of $\rho_{T \hat{T}^{n_{1}}(S \hat{S})^{n_{2}}}\left(a_{1}, \boldsymbol{a}_{2}, q_{1}, \boldsymbol{q}_{2}\right)$ after tracing out $\hat{T}^{n_{1}-1}$ systems and $(S \hat{S})^{n_{2}-1}$ systems, and summing over $n_{1}-1$ indices of each $\boldsymbol{a}_{2}$ and $\boldsymbol{q}_{2}$.

[^37]:    ${ }^{4}$ For general Bell inequalities the rule matrix $V\left(a_{1}, a_{2}, q_{1}, q_{2}\right)$ is not necessarily binary but can take general real values.

[^38]:    ${ }^{1}$ These are the ones allowing for a reduction of the variance in the examples in Sec. 4.4.

[^39]:    ${ }^{1}$ The inequality in Eq. (C.16) holds as $\overline{\operatorname{sdp}}_{|Q|}^{\mathrm{proj}}\left(V, \pi_{\text {unif }}, 2\right)$ imposes more constraints than $\operatorname{sdp}_{\mathrm{PPT}}(V)$.

