Semidefinite Programming Techniques for the Quantum Causal Compatibility Problem

LAURENS LIGTHART

aus Alkmaar

Köln, 2023

This dissertation has been accepted by the Faculty of Mathematics and Natural Sciences of the University of Cologne.

Gutachter:	Prof. Dr. David Gross
	Prof. Dr. Frank Vallentin
Tag der Disputation:	30.11.2023

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen:

- Laurens T Ligthart, Mariami Gachechiladze, and David Gross. "A convergent inflation hierarchy for quantum causal structures". Communications in Mathematical Physics **401** (2023), doi:10.1007/s00220-023-04697-7 [1]

- Laurens T Ligthart and David Gross. "The inflation hierarchy and the polarization hierarchy are complete for the quantum bilocal scenario". Journal of Mathematical Physics **64** (2023), doi:10.1063/5.0143792 [2]

- Owidiusz Makuta, Laurens T Ligthart, and Remigiusz Augusiak. "No graph state is preparable in quantum networks with bipartite sources and no classical communication". npj Quantum Information **9** (2023), doi:10.1038/s41534-023-00789-3 [3]

Thursday 25th January, 2024, Laurens Ligthart

Abstract

Characterizing the set of correlations that can arise from performing measurements in a quantum description of Nature is a relevant, but challenging task. Such a characterization, known as the quantum causal compatibility problem, provides us with insights in the advantages of quantum theory over a classical theory, but can also show its limitations. This problem becomes particularly challenging when the quantum states and measurements are required to be compatible with a given causal structure. A causal structure dictates the causal dependencies of the parties and systems involved in the experiment. One can think of the Bell scenario as one of the simplest causal structures, in which two spatially separated parties, Alice and Bob, are assumed to perform a measurement on a shared source. In more general causal structures we might have more parties, and more sources that are independent of each other.

Recently, a systematic way of analyzing the correlations in classical and quantum causal structures was proposed in the form of the inflation technique. In the inflation technique the causal dependencies, which are difficult to encode algorithmically, are relaxed to easy-to-encode symmetry constraints on a larger number of parties. For the classical case, this provides a converging hierarchy of linear programs for the causal compatibility problem. For the quantum case, it instead yields a hierarchy of increasingly restrictive semidefinite programming relaxations. It is, however, unknown whether this hierarchy is also complete.

One of the main results of this thesis is to show that a modified version of the quantum inflation technique is convergent for the quantum causal compatibility problem. This modified hierarchy introduces an additional parameter, r, that restricts the Schmidt rank of the observables. For each value of r we provide a hierarchy of compatibility tests that is complete in the sense that it will detect, at some finite level, any probability distribution that is incompatible with the causal model under the Schmidt rank constraint. Such compatibility tests are formulated as non-commutative polynomial optimization problems, of which we provide a C^*-algebraic description.

Additionally, we develop a separate hierarchy of semidefinite programs, which

we call the polarization hierarchy. It is shown that the polarization hierarchy, as well as the original quantum inflation hierarchy are complete for the causal structure known as the bilocal scenario. In the bilocal scenario there are three distant parties, Alice, Bob and Charlie, performing measurements on two independent sources: one shared by Alice and Bob and the other by Bob and Charlie. We show that a model for Bob's algebra, which consists of two commuting subalgebras, can be constructed from the commutants of representations of Alice's and Charlie's algebras. This construction also gives insight into a bilocal version of Tsirelson's problem. We show that if Alice's and Charlie's systems can be modeled with a finite dimensional representation, the commuting observables model and tensor product model of locality in quantum theory coincide.

Our convergence results rely centrally on the fact that certain symmetries in the limit imply independence of random variables or quantum states. Such statements are collectively known as de Finetti theorems. For the specific setup considered in this thesis, namely the C^*-algebraic description of quantum mechanics, a de Finetti theorem had not yet been proven beyond the special case of the minimal tensor product. Another result of the thesis is therefore the proof that a quantum de Finetti theorem also holds for general tensor products of C^*-algebras.

The quantum causal compatibility problem can be seen as a version of the quantum network compatibility problem – in which we ask which quantum states can be produced in a certain causal structure – where the output is assumed to be a classical state. We show that the techniques developed for the causal compatibility problem can be adapted to the more general setting of quantum networks. Furthermore, an analytic proof is given of the fact that graph states cannot be produced in bipartite quantum networks. This proof again relies on the inflation technique by linking correlations of different inflations of the network to each other. By assuming that the correlations arise from a graph state, it can be shown that the bound of a particular inequality can be violated, which leads to a contradiction.

Lastly, we show that the polarization hierarchy can be used to optimize over a large class of optimization problems known as state polynomial optimization. In such problems, the goal is to optimize an objective function that is a polynomial in the expectation values of observables, under similar polynomial constraints. This allows us, for example, to optimize over covariances or non-linear Bell inequalities. We also give an alternative version of a recently developed semidefinite programming hierarchy that solves this problem, in which we incorporate the polarization trick.

Contents

Symbols

Symbol	Meaning
$\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$	Unital C^*-algebra
$\mathcal{A}_{\mathrm{sa}}$	Self-adjoint elements of $\mathcal{A}$
$\mathcal{A}_+$	Positive elements of $\mathcal{A}$
$\mathcal{C}_{\mathrm{SE}}$	Commutative scalar extension algebra
$A_{a|x}, E_a^x$	POVM element with outcome a and setting x
$B(\mathcal{H})$	Bounded operators on a Hilbert space $\mathcal{H}$
$C^*(\mathcal{G}|\mathcal{R})$	Universal C^*-algebra generated by $\mathcal{G}$ subject to $\mathcal{R}$
$\mathcal{D}^n$	Algebra associated with level n inflation
$\mathcal{D}^\infty$	Quasi-local algebra
$\mathcal{F}(\mathcal{G})$	Free $*$-algebra in the generators $\mathcal{G}$
$\mathcal{F}^{(k)}(\mathcal{G})$	Subset of $\mathcal{F}(\mathcal{G})$ with words in the generators $\mathcal{G}$ up to length k
$f^\star, f^*, f^n$	Optimal value of an optimization problem
$\mathcal{G}$	Set of generators
$\mathcal{H}$	Hilbert space
$\mathcal{R}$	Set of relations
$K(\mathcal{A})$	State space of a C^*-algebra $\mathcal{A}$
$K_s(\mathcal{D}^\infty)$	Symmetric states on the quasi-local algebra
Λ	Quantum channel; Local hidden variable
$\rho, \sigma, \phi, \omega$	Quantum state (as density matrix or functional)
$|\Omega\rangle$	Cyclic GNS state on a Hilbert space
$\langle\,.\,\rangle$	Expectation value
$p(abc|xyz)$	Conditional probability
P, p	Probability distribution
p_i, q	Polynomials in generators
π	Representation of a C^*-algebra; Permutation of n elements
S_n	Symmetric group of n elements
α_π	Action of $\pi \in S_n$ on elements of an algebra
Π_σ	Infinite symmetric product state on the quasi-local algebra
Sym^n	Symmetrization operator on n elements
$\mathrm{Sym}^\infty(\mathcal{D})$	Symmetric quasi-local algebra
tr, Tr	Trace
$\|\cdot\|_\gamma$	C^*-(semi-)norm
$\star$	Abelian product on the symmetric (quasi-)local algebra
$\otimes$	(Algebraic) tensor product
$\otimes_{\mathrm{min}}$	Minimal tensor product
$\otimes_{\mathrm{max}}$	Maximal tensor product
$\bar{\otimes}$	Von Neumann tensor product
$\overline{\cdots}$	Completion w.r.t. a (semi-)norm or a topology

Introduction

One of the most noteworthy discoveries in physics of recent times is that, with very high likelihood, Nature cannot both be local (causal influences do not travel faster than the speed of light) and real (the value of physical quantities is predictable with certainty, without disturbing the system). This counter intuitive result was shown in a series of experiments performed by Alain Aspect and collaborators in 1982 [4, 5] and has since been repeated many times over, in order to improve its statistical significance and to get rid of any loopholes [6, 7]. The idea behind the experiment is relatively simple, but its results have profound implications on any theory of Nature. The setup is based on an observation made by Einstein, Podolski and Rosen [8] that quantum theory is not simultaneously local and real, as well as on the subsequent derivation of an inequality by Bell [9] that any local, real theory should obey. Later, a similar inequality was derived by Clauser, Horne, Shimony and Holt [10], which is now known as the CHSH inequality. It can be formulated as follows. Let Alice and Bob be two experimenters whose laboratories are spatially separated (see Fig. 1). Alice has a choice between two measurements with outcomes ± 1 each, denoted by A_1 and A_2, and the same is true for Bob, with measurements B_1 and B_2. Then for any local, real theory

$$\langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle \leq 2. \tag{1}$$

The experiments of Aspect showed that this inequality could be violated by a pair of photons, confirming that with very high likelihood Nature cannot be fully explained by a local hidden variable model, and putting a nail in the coffin of local realism.

Quantum theory is one way of explaining the violation of the CHSH inequality. Indeed, by choosing the right quantum states and measurements, the CHSH inequality (1) can be violated up to $2\sqrt{2}$. These results show that quantum theory is in some sense more powerful than classical local hidden variable models. In what sense quantum theory rejects local realism is a discussion that is well beyond the goal of this thesis. We are, however, interested in the consequences, advantages and
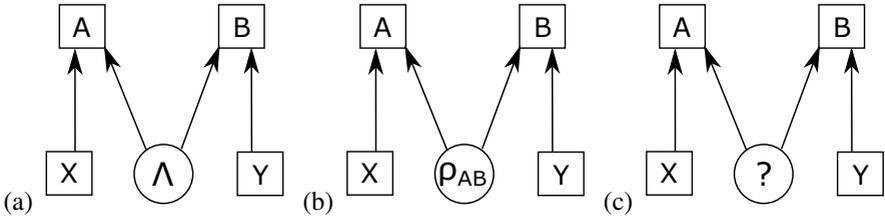
Figure 1: The famous Bell scenario, in which two distant observers, Alice and Bob, perform measurements on a system received from a shared source. The random variable X serves as a measurement setting for Alice, while Y serves as a setting for Bob. In Fig. (a), the source is assumed to be a classical local hidden variable Λ. The correlations that arise from this are bounded by Bell inequalities such as the CHSH inequality. In Fig. (b), the source distributes one part of a bipartite quantum state ρ_{AB} to each of the observers. This allows for a larger set of correlations, but is still dictated by the limitations of quantum theory. It is possible that there exist more exotic sources, such as the unknown source of Fig. (c), but so far the experiments agree with the predictions of quantum theory.

limitations of choosing quantum theory as the description of Nature.

A logical follow-up question to the Bell experiment would be how these results can be extended to more general *causal structures*. That is, what are the correlations that can be achieved in more general experiments, involving more parties, more states, and more measurements? This question, known as the *causal compatibility problem*, is one of the central topics of this thesis.

Both in classical and in quantum theory, causal relationships can be conveniently depicted in the form of a directed acyclic graph (DAG) G. We have already seen an example of a DAG in the form of the Bell scenario in Fig. 1. In a DAG, we distinguish between unobserved systems, depicted in circles, and observed systems, depicted in boxes. The causal dependencies are denoted by arrows. In the quantum case, these arrows can be thought of as quantum channels [11].

In order to analyze the power of different causal structures, we are interested in finding the answer to the following problem, known as the causal compatibility problem, for particular choices of probability distributions P.

Problem 1 (Causal compatibility)**.** *Given a (conditional) probability distribution P and a causal structure G, determine whether P can be produced in a model compatible with the causal structure.*

If the model in Problem 1 is restricted to a classical description of nature, we call the problem the *classical causal compatibility problem*. Similarly, if the model

is allowed to be a quantum model, the problem is known as the *quantum causal compatibility problem*. What it means to be a "quantum model" turns out to be a subtle question. In this thesis, quantum theory will be modeled from the algebraic perspective. That means that the central object for a quantum system is a C^*-algebra of bounded observables, with states acting as normalized functionals on these observables. If there are multiple systems, they are commuting subalgebras of a larger algebra. This way of modeling locality in quantum theory gained traction after a series of influential papers by Murray and Von Neumann [12, 13, 14, 15, 16]. The second way of modeling quantum mechanics, which is more common in the last decades, involves vector states on Hilbert spaces, and observables that are taken to be the bounded linear operators on the Hilbert space. Subsystems are then combined by simply taking the tensor product of the Hilbert spaces. For a long time, it was an open problem whether these two models were equivalent. This problem, known as Tsirelson's problem, has recently been answered in the negative for infinite dimensional systems [17]. In Section 1.2 we will therefor motivate our decision to adopt the algebraic model.

Figuring out the causal relationship among random variables has applications in many fields of study that deal with statistics, such as economics and medicine [18, 19, 20]. Arguably, the quantum version of this problem is not of great interest in these research areas. In the field of quantum information theory the motivation is therefore often a bit different. There, the goal is to distinguish quantum theory from other theories in a device-independent way, either by showing that the quantum model is capable of beating classical models (cf. the Bell scenario), or by showing that it is less powerful than other generalized probabilistic theories (GPTs). A convenient way to do this is to find witnesses of incompatibility, for example in the form of inequalities such as the CHSH inequality of Eq. (1). Similar inequalities can be derived to bound the set of quantum correlations. Each such bound gives us an experiment that can be performed in a lab and violations of the bounds force us to reconsider the assumptions of the theory.

In general it is considered very challenging to analyze causal structures with unobserved systems, even classically [18]. Recently, a new method, called the inflation technique, has been proposed for both classical [21] and quantum causal structures [22]. It is formulated as a hierarchy of compatibility tests in the form of linear programs in the classical case, and semidefinite programs in the quantum case. At each level of the hierarchy, the causal relationships implied by the causal structure are relaxed to a set of symmetry constraints on a larger number of variables. The classical inflation hierarchy has been shown to be complete in the sense that any probability distribution that is incompatible with the causal hypothesis will be detected at some level of the hierarchy [23]. The result relies on a de Finetti theorem, which (roughly) states that the symmetry constraints in the limit are sufficient to enforce the causal

relationships. For the quantum case no such result is known.

In Chapter 3 we partially solve this problem: we provide a hierarchy of semidefinite programming relaxations that is inspired by the inflation technique, but which is provably convergent. Our proof also relies on a de Finetti theorem. In the quantum case, such theorems state that symmetry constraints are, in the limit, sufficient to force a quantum state to be separable. However, no such theorem was proven for the very general algebraic model that we work with. We therefore adapt the proof technique of Ref. [24] to this more general case, resulting in a quantum de Finetti theorem for the maximal tensor product (Theorem 2.1.1).

This newly developed inflation method has some drawbacks over the original version of the inflation technique: it requires a much larger number of variables, and the compatibility tests are dependent on the Schmidt rank of the measurements that are performed, which dictates how entangled the measurements can be. Therefore, it remains relevant to investigate the convergence of the original inflation method.

The simplest generalization of the Bell scenario that involves independence constraints is the bilocal scenario. It consists of three parties, Alice, Bob and Charlie, and two independent quantum sources, one shared by Alice and Bob, and one shared by Bob and Charlie (see Fig. 3.7). This scenario forms the basis for entanglement swapping experiments [25]. As we show in Sec. 3.3, it is also one of the few causal structures for which we know that the original quantum inflation technique is convergent. One of the biggest hurdles in proving convergence consists of constructing all the local subalgebras from the information given to and provided by the computer in the SDP relaxations of the problem. For the bilocal scenario, we show that this can be done by constructing Bob's system from the commutants of Alice's and Charlie's algebras. Apart from the inflation hierarchy, we develop a different convergent hierarchy, which we call the polarization hierarchy.

Additionally we might want to go beyond correlations and ask what happens if the output is not a probability distribution, but a quantum state? Which states can we produce in quantum networks that are under certain causal restrictions? Since such networks form the backbone of many quantum information processing tasks – think of quantum key distribution [26, 27, 28], clock synchronisation [29], parallel computing [30] or even a quantum internet [31, 32, 33, 34] – such questions form a very relevant line of inquiry. The problem can be stated as a generalization of Problem 1:

Problem 2 (Network compatibility). *Given a d^N-dimensional quantum state ρ on N parties and a causal structure G with N end nodes, determine whether ρ can be produced in a quantum model compatible with the causal structure.*

In Section 4.1, we show that the SDP hierarchies developed for the causal compatibility problem can relatively easily be adapted to this more general setting as

well.

It is also possible to use the inflation technique as an analytic tool for showing incompatibility [21]. This is the approach in Section 4.2, where it is shown that graph states cannot be produced in bipartite networks in the so-called local operations and shared randomness (LOSR) setup, i.e. without classical communication. Since graph states are a very promising set of states due to their easy description in terms of the stabilizer formalism, and due to their error-correcting potential [35, 36], this can be seen as an argument against the LOSR setup.

In the final chapter of the thesis, we momentarily forget about the causal compatibility and network compatibility problems and treat state polynomial optimization problems. This type of problems is tailored to optimization over objective functions and constraints that are polynomials in the state. It provides a way to optimize over non-linear Bell inequalities, which also includes e.g. optimization over covariances [37, 38, 39], generalizing the non-commutative polynomial optimization (NPO) method of Ref. [40]. We show that the polarization hierarchy is convergent for this type of problems, and we compare it to the recently developed SDP hierarchy of Ref. [39], of which we give an alternative version and proof. This second hierarchy is closely related to what is known as scalar extension in the quantum information literature [41].

Outline

In Chapter 1 we treat the theoretical preliminaries that are necessary to understand the subsequent chapters. In particular, some important notions from the theory of operator algebras is discussed in Sec. 1.1. The finer details of locality and causality are treated in Sec. 1.2. Here the notion of a causal structure is also introduced. In order to analyze such causal structures, we use convex optimization methods in the form of semidefinite programming, which is the topic of Sec. 1.3.

The causal compatibility problem that we have introduced above in Problem 1 involves statements about product states. We prove a de Finetti theorem that fits the causal compatibility problem in Chapter 2. Additionally, some closely related statements are derived for the case of scalar extension and quantum inflation.

Chapter 3 discusses the quantum causal compatibility problem, and our efforts in solving it. Indeed, in Secs. 3.1 and 3.2 the difficulties in proving convergence of any kind of hierarchy to tackle this problem are explained, and a way to overcome them is proposed. This involves an SDP hierarchy that is heavily inspired by the quantum inflation technique, but which is provably complete. In Sec. 3.3, it is shown that for the special case of the bilocal scenario, the extra efforts of Secs. 3.1 and 3.2 are unnecessary.

In Chapter 4, we show that the inflation and polarization hierarchies can also

be used to make statements about states that are produced in quantum networks. In Sec. 4.1 we show that many of the convergence results from Chapter 3 carry over to this case. Additionally, in Sec. 4.2, we give an analytic proof of the fact that non-trivial graph states cannot be produced in bipartite quantum networks if we only allow for local operations and shared randomness (LOSR). This result also holds for states that are close to a graph state.

Lastly, in Chapter 5 a large class of problems, known as state polynomial optimization (SPO), is considered. This involves the polarization hierarchy, which is described in Sec. 5.1, as well as the scalar extension hierarchy, which was developed in Ref. [39] and of which we give an alternative version in Sec. 5.2.

We end the thesis with some concluding remarks and an outlook on some open problems.

Contributions

- **Paper [1]:** The theory of a maximal tensor product de Finetti theorem, stated in Sec. 2.1, was developed together with my collaborators Mariami Gachechiladze and David Gross. I developed the completeness argument presented in Secs. 3.1 and 3.2 of this thesis. I wrote the manuscript, with later revisions by my collaborators.

- **Paper [2]:** The proof strategy for Theorem 3.3.10 and the corollaries that follow from it were developed and worked out in discussions and collaboration with David Gross. The completeness results of Sec. 3.3.6, in particular Theorems 3.3.15 and 3.3.16 were derived by me, based on the results of Paper [1]. I wrote the manuscript, with later revisions by my collaborator.

- **Paper [3]:** The majority of this paper was written by the first author, Owidiusz Makuta. My contributions are the following

    - Independently formulating the problem description.

    - Stating and proving Lemma 4.2.3.

    - Stating and providing an independent proof for Lemma 4.2.4.

    - Developing the robustness results of Sec. 4.2.3.

- All other results in this thesis, most notably the unpublished results of Sec. 2.3, Sec. 4.1 and Chapter 5, were developed by me unless stated otherwise.

Chapter 1

Preliminaries

> *The procedure that creates the universe out of the vacuum is called the Genesis Construction (often abbreviated as GNS). This is no coincidence, if for no other reason that nothing ever is! It might seem peculiar that the Genesis Construction starts with $|\Omega\rangle$ which should, by rights, be the end of it. But on reflection, things are wholly consistent. Genesis is meant to be used with the Heisenberg picture, where time flows backwards and calculations do indeed start with an element A (of an observable algebra) and are contracted against $|\Omega\rangle$ in the very end.*
>
> From a discussion between David Gross and me

In this chapter the mathematical, physical and computational topics that are relevant for the subsequent chapters are introduced. We start the discussion in section 1.1 with the treatment of C^*-algebras as generalizations of algebras of observables on Hilbert spaces. This will also involve some details on von Neumann algebras, a special type of C^*-algebras.

Secondly, one of the main goals of this thesis is to study the consequences of assuming a certain causal structure, either for correlations (Chapter 3) or for network states (Chapter 4). It turns out that there are multiple different notions of locality that will play a highly important role in this analysis. The relevant theory of the surprisingly subtle study of locality is treated in section 1.2. Subsequently, we use these definitions of locality to formally set up the definitions for causal structures in the form of directed acyclic graphs (DAGs).

A common way to analyze the correlations in causal structures is by convex optimization in the form of semidefinite programming. Therefore, in section 1.3 an

exposition of some of the basic notions in semidefinite programming is given, with some extra emphasis on the problem of non-commutative polynomial optimization.

1.1 C^*-algebras & topologies

Von Neumann algebras and C^*-algebras were first introduced approximately a century ago by von Neumann and Murray in a series of extremely influential papers on what they called "rings of operators" and "operator algebras" [12, 13, 14, 15, 16]. Their goal was to study infinite dimensional objects and abstract operator theory on Hilbert spaces. This line of research was very much motivated by the recent developments in quantum theory, where such objects arose naturally, and where the lack of understanding of such objects proved to be a major difficulty.

The study of C^*-algebras essentially consists of two parts [42]: 1). The intrinsic structure and abstract description of the algebra, and 2). the concrete realizations of such an algebra as representations on Hilbert spaces. One can compare this to the abstract characterization of groups, and the classification of their representations. As we will see below, the theory of operator algebras therefore consists of a mixture of algebraic, as well as analytic results.

Definition 1.1.1. A C^*-algebra $\mathcal{A}$ is a Banach algebra with an involution * and a norm $\| \, . \, \|$ that satisfies the C^* property

$$\|xx^*\| = \|x\|^2 = \|x^*\|^2. \tag{1.1}$$

A C^*-algebra is said to be unital if it contains an identity element $\mathbb{1}$.

Throughout this thesis, the C^*-algebras we encounter are assumed to be unital.

The subset of $\mathcal{A}$ that is self-adjoint is denoted by $\mathcal{A}_{\text{sa}}$. The *dual* of $\mathcal{A}$ is denoted by $\mathcal{A}^*$ and is given by all linear maps $\mathcal{A} \to \mathbb{C}$ (or $\mathcal{A} \to \mathbb{R}$ in the case of a real C^*-algebra). Such maps are also known as *functionals*.

Definition 1.1.2. An element $x \in \mathcal{A}_{\text{sa}}$ is *positive*, denoted by $x \geq 0$ (or sometimes $x \succeq 0$), if it can be written as $x = y^*y$ for some $y \in \mathcal{A}$. The set of positive elements of $\mathcal{A}$ is denoted by $\mathcal{A}_+$.

Positivity defines a partial order on $\mathcal{A}_{\text{sa}}$ in the sense that $x \geq y$ if and only if $x - y \geq 0$. It also allows us to define positivity on the dual $\mathcal{A}^*$, from which we can define the notion of a state.

Definition 1.1.3. A *state* on a unital C^*-algebra $\mathcal{A}$ is a functional $\rho : \mathcal{A} \to \mathbb{C}$ that is positive in the sense that

$$\rho(x) \geq 0 \quad \forall x \in \mathcal{A}_+,$$

and that is normalized as

$$\rho(\mathbb{1}) = 1.$$

We denote the *state space*, i.e. the set of all states, of $\mathcal{A}$ by $K(\mathcal{A})$.

We started the discussion of C^*-algebras by motivating them as generalizations of algebras of bounded operators on Hilbert spaces. Conversely, the pairing of a C^*-algebra with a state gives us a way of constructing representations on a Hilbert space via the GNS construction detailed below.

Theorem 1.1.4. GNS construction (see e.g. [42, Thm. 9.14] or [43, II.6.4.1]) *Let $\mathcal{A}$ be a unital C^*-algebra. Then, to any state $\omega \in K(\mathcal{A})$ there corresponds uniquely (up to an isometry) a representation $(\pi_\omega, \mathcal{H}_\omega)$ of $\mathcal{A}$ with a vector $|\Omega\rangle$ such that*

1. *$\pi_\omega(\mathcal{A})|\Omega\rangle$ is dense in $\mathcal{H}_\omega$ (i.e. $|\Omega\rangle$ is a cyclic vector for $\pi_\omega(\mathcal{A})$),*

2. *$\omega(x) = \langle\Omega|\pi_\omega(x)|\Omega\rangle$.*

Proof. Each state ω defines a pre-inner product on $\mathcal{A}$ through

$$\langle x, y \rangle := \omega(x^*y). \tag{1.2}$$

Let $N_\omega := \{x \in \mathcal{A} : \langle x, x \rangle = 0\}$. This is a closed ideal of $\mathcal{A}$. Define $\mathcal{H}_\omega$ as the completion of the pre-Hilbert space $\mathcal{A}/N_\omega$ with respect to the inner product (1.2). For any $a \in \mathcal{A}$, let $\pi_\omega^0(a)$ be the left multiplication operator on $\mathcal{A}/N_\omega$, i.e. the operator that maps

$$\pi_\omega^0(a)(x + N_\omega) = ax + N_\omega.$$

$\pi_\omega^0(a)$ is bounded due to

$$x^*a^*ax \leq \|a\|^2 x^*x,$$

which follows from $a^*a \leq \|a\|^2 \mathbb{1}$. Hence, $\left\|\pi_\omega^0(a)\right\| \leq \|a\|$ and thus $\pi_\omega^0(a)$ can be extended to a bounded operator $\pi_\omega(a)$ on $\mathcal{H}_\omega$. By the Riesz representation theorem, there exists a unique vector $|\Omega\rangle$ such that $\omega(x) = \langle\Omega|\pi_\omega(x)|\Omega\rangle$. Identifying $|\Omega\rangle$ with the identity element shows that $\pi_\omega(a)|\Omega\rangle$ is indeed dense in $\mathcal{H}_\omega$.

To show uniqueness, let $(\pi_\omega, \mathcal{H}_\omega, |\Omega\rangle)$ and $(\pi_\omega', \mathcal{H}_\omega', |\Omega'\rangle)$ be two GNS representations obeying the assumptions of the theorem. Then, define a map V_0 as

$$V_0 \pi_\omega'(a)|\Omega'\rangle = \pi_\omega(a)|\Omega\rangle.$$

This yields

$$\langle \Omega'|(V_0\pi'_\omega(a))^* V_0\pi'_\omega(b)|\Omega'\rangle = \langle \Omega|(\pi_\omega(a))^* \pi_\omega(b)|\Omega\rangle$$
$$= \omega(a^*b) = \langle \Omega'|(\pi'_\omega(a))^* \pi'_\omega(b)|\Omega'\rangle \,,$$

so that V_0 is a well-defined isometry for $\pi'_\omega |\Omega'\rangle$ onto $\pi_\omega |\Omega\rangle$ and can be extended to an isometry for the Hilbert spaces. V_0 also defines an isometry for the representations, since

$$\pi_\omega(a)V_0\pi'_\omega(b)\,|\Omega'\rangle = V_0\pi'_\omega(a)\pi'_\omega(b)\,|\Omega'\rangle \,,$$

so that $V_0\pi'_\omega(a) = \pi_\omega(a)V_0$ for all $a \in \mathcal{A}$. $\square$

1.1.1   Universal C^*-algebras

The following section corresponds to section 2.2.2 from Paper [1].

Let $\mathcal{G} = \{g_i\}_i$ be a countable set of symbols. Denote by $\mathcal{F}(\mathcal{G})$ the free complex $*$-algebra generated by the elements of $\mathcal{G}$. Put differently, $\mathcal{F}(\mathcal{G})$ is the set of finite complex linear combinations of words in the symbols g_i and g_i^*, with multiplication defined by concatenation of words. Choose a countable set $\mathcal{R} \subset \mathcal{F}(\mathcal{G})$. We aim to define the "largest C^*-algebra with generators $\mathcal{G}$, subject to the constraint that each $q \in \mathcal{R}$ is positive". We will refer to the elements of $\mathcal{R}$ as *relations*.

To make this notion precise, define a *representation of* $(\mathcal{G}|\mathcal{R})$ to be a homomorphism $\pi : \mathcal{F}(\mathcal{G}) \to \mathcal{B}(\mathcal{H})$ from the free algebra into the set of bounded operators on some Hilbert space $\mathcal{H}$, such that $\pi(q)$ is a positive operator for every $q \in \mathcal{R}$. On $\mathcal{F}(\mathcal{G})$, define

$$\|x\| := \sup\{\|\pi(x)\| \mid \pi \text{ is a representation of } (\mathcal{G}|\mathcal{R})\}. \qquad (1.3)$$

Now assume that the relations imply $\|x\| < \infty$ for all $x \in \mathcal{F}(\mathcal{G})$. In this case, $\|\cdot\|$ is a seminorm on $\mathcal{F}(\mathcal{G})$. The *universal C^*-algebra on* $(\mathcal{G}|\mathcal{R})$, abbreviated as $C^*(\mathcal{G}|\mathcal{R})$, is then the completion of $\mathcal{F}(\mathcal{G})$ with respect to this C^*-seminorm.[1] We will not differentiate notationally between an element $x \in \mathcal{F}(\mathcal{G})$ and its image in the completion $C^*(\mathcal{G}|\mathcal{R})$.

In our applications, we will mostly consider two types of relations:

---

[1] Recall that, despite the everyday connotations of the word, the process of "completing" a metric space does not necessarily only add elements to it. The completion can be defined as the set of Cauchy sequences, with two considered equivalent if the norm of their differences converges to zero. Every $x \in \mathcal{F}(\mathcal{G})$ gives rise to an element in $C^*(\mathcal{G}|\mathcal{R})$, represented by the sequence that is constant and equal to x. Two elements $x, y \in \mathcal{F}(\mathcal{G})$ induce the same element in $C^*(\mathcal{G}|\mathcal{R})$ if and only if $\|x - y\| = 0$. Put differently, the completion adds elements to the quotient space $\mathcal{F}(\mathcal{G})/\{x \mid \|x\| = 0\}$.

Equality constraints. Let $x, y \in \mathcal{F}(\mathcal{G})$ and assume that $\mathcal{R}$ contains both $x - y$ and $y - x$. It then follows easily that $x = y$ in the universal C^*-algebra $C^*(\mathcal{G}|\mathcal{R})$. We will always assume that $g_1 =: \mathbb{1}$ is constrained to commute with all the others and obeys $\mathbb{1}x = x\mathbb{1} = x$ for all $x \in \mathcal{F}(\mathcal{G})$, so that $C^*(\mathcal{G}|\mathcal{R})$ is unital.

Norm constraints. Let $x \in \mathcal{F}(\mathcal{G}), C \in \mathbb{R}_+$ and assume that $\mathcal{R}$ contains $C^2\mathbb{1} - x^*x$. Then $\|x\| \leq C$ in $C^*(\mathcal{G}|\mathcal{R})$.

We will prove the following lemma, which justifies our characterization of $C^*(\mathcal{G}|\mathcal{R})$ as the "largest C^*-algebra such that each $q \in \mathcal{R}$ is positive".

Lemma 1.1.5. *In $C^*(\mathcal{G}|\mathcal{R})$, any element $q \in \mathcal{R}$ is positive.*

Proof. We first show that for every representation ϕ of $C^*(\mathcal{G}|\mathcal{R})$, it holds that $\phi(q)$ is a positive operator for each $q \in \mathcal{R}$. (Using the terminology introduced above, this says that a representation of $C^*(\mathcal{G}|\mathcal{R})$ is also a representation of $(\mathcal{G}|\mathcal{R})$).

Fix a $q \in \mathcal{R}$. There is no loss of generality in assuming $\|q\| \leq 2$.

It holds that $q = q^*$, because

$$\|q - q^*\| = \sup_{\pi \text{ representation of } (\mathcal{G}|\mathcal{R})} \|\pi(q - q^*)\| = 0,$$

as $\pi(q)$ is positive (and hence self-adjoint) by definition of representations of $(\mathcal{G}|\mathcal{R})$.

For every representation π of $(\mathcal{G}|\mathcal{R})$, it holds that $\|\pi(\mathbb{1} - q)\| \leq 1$ [43, Proposition II.3.1.2(iv)]. From the definition of the seminorm, this implies that $\|\mathbb{1} - q\| \leq 1$.

Now assume for the sake of reaching a contradiction that for some representation ϕ of $C^*(\mathcal{G}|\mathcal{R})$, the operator $\phi(q)$ is not positive. Using again [43, Proposition II.3.1.2(iv)], $\|\phi(\mathbb{1} - q)\| > 1 \geq \|\mathbb{1} - q\|$, which is a contradiction, as representations are norm-contractions.

Next, let $\rho \in K(C^*(\mathcal{G}|\mathcal{R}))$. By the above, using the GNS representation,

$$\rho(q) = \langle \Omega | \pi_\omega(q) | \Omega \rangle \geq 0.$$

Hence q is positive by [43, Corollary II.6.3.5]. $\qquad\qquad\square$

1.1.2 Tensor products of C^*-algebras

The following section is largely taken from section 2.2.4 of Paper [1].

If $\mathcal{A}$ and $\mathcal{B}$ are the C^*-algebras of observables on one subsystem each, then the joint system should come with an observable algebra $\mathcal{C}$ that contains copies of $\mathcal{A}$ and $\mathcal{B}$ as commuting subalgebras and is generated by them. Unfortunately, these two requirements are not quite enough to uniquely determine $\mathcal{C}$. To understand the

freedom we have in defining the set of global observables, start with the *algebraic tensor product* $\mathcal{C}_0 = \mathcal{A} \otimes_{\text{alg}} \mathcal{B}$. This is the $*$-algebra of elements x of the form

$$x = \sum_i a_i \otimes b_i \qquad a_i \in \mathcal{A}, \ b_i \in \mathcal{B} \tag{1.4}$$

with multiplication and involution in $\mathcal{C}_0$ defined factor-wise as

$$(a_1 \otimes b_1)(a_2 \otimes b_2) = a_1 a_2 \otimes b_1 b_2, \qquad (a_1 \otimes b_1)^* = a_1^* \otimes b_1^*. \tag{1.5}$$

To promote the $*$-algebra $\mathcal{C}_0$ to a C^*-algebra $\mathcal{C}$, we have to endow it with a norm satisfying the C^*-norm property in Eq. (1.1) and complete it with respect to this norm. The choice of this norm is not unique [42, 43]. There are two distinguished norms: minimal and maximal, named-so because they constrain the value of any C^*-norm on the algebraic tensor product by

$$\|x\|_{\text{min}} \leq \|x\| \leq \|x\|_{\text{max}}. \tag{1.6}$$

In the more general case of n tensor factors, they are defined via their respective values on elements of $\mathcal{A}_1 \otimes_{\text{alg}} \ldots \otimes_{\text{alg}} \mathcal{A}_n$ as

$$\|x_1 \otimes \cdots \otimes x_n\|_{\text{min}} = \sup\{\|\pi_1(x_1)\| \ldots \|\pi_n(x_n)\| \mid \pi_i \text{ a rep. of } \mathcal{A}_i\}, \tag{1.7}$$

$$\|x_1 \otimes \cdots \otimes x_n\|_{\text{max}} = \sup\{\|\pi(x_1 \otimes \cdots \otimes x_n)\| \mid \pi \text{ a rep. of } \mathcal{A}_1 \otimes_{\text{alg}} \ldots \otimes_{\text{alg}} \mathcal{A}_n\}, \tag{1.8}$$

where the suprema are taken over representations of the respective C^*-algebras as operator algebras on a Hilbert space. We denote the C^*-algebra generated by the tensor product of $\mathcal{A}$ and $\mathcal{B}$ and completed with respect to the norm $\|\cdot\|_{\gamma}$ by $\mathcal{A} \otimes_{\gamma} \mathcal{B}$.

If the $\mathcal{A}_i$'s arise as bounded operators on Hilbert spaces, $\mathcal{A}_i = B(\mathcal{H}_i)$, the approach from elementary quantum mechanics corresponds to their natural embedding into $B(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$. The operator norm in this picture corresponds to the minimal tensor product. In this way, the elementary approach reappears as a special case of the algebraic construction.

Note that due to the definition in Eq. (1.8), convergence with respect to the maximal norm implies convergence for *any* operator representation of the global observable algebra.

The tensor product structure now allows us to define two special subsets of states on tensor product algebras [1, Def. 2].

Definition 1.1.6. Let $\mathcal{A}$ and $\mathcal{B}$ be C^*-algebras and let $\mathcal{A} \otimes_{\gamma} \mathcal{B}$ be the completion of the algebraic tensor product with respect to the C^*-norm $\|\cdot\|_{\gamma}$. A state $\rho \in$

$K(\mathcal{A} \otimes_\gamma \mathcal{B})$ is a *product state* across $\mathcal{A} \,|\, \mathcal{B}$ if

$$\rho(ab) = \rho(a)\rho(b) \qquad \forall a \in \mathcal{A}, b \in \mathcal{B}. \tag{1.9}$$

A state $\omega \in K(\mathcal{A} \otimes_\gamma \mathcal{B})$ is said to be *separable* across the partition $\mathcal{A} \,|\, \mathcal{B}$ if it is a convex combination of product states, i.e. if

$$\omega(ab) = \int \mathrm{d}\mu(\sigma)\sigma(ab) = \int \mathrm{d}\mu(\sigma)\sigma(a)\sigma(b), \tag{1.10}$$

for some probability measure μ over product states.

1.1.3 Topologies and von Neumann algebras

A large part of this thesis deals with the concept of convergence. While quite often it is sufficient to consider convergence with respect to a certain norm, or equivalently convergence with respect to the topology induced by that norm, we will sometimes require convergence with respect to different topologies. Therefore, this subsection is meant to recall some notions about (functional) analysis that will be useful in subsequent chapters.

Definition 1.1.7. Let X be a topological vector space and let $X^* : X \to \mathbb{C}$ be a separating dual. Then we call the X^*-topology on X the *weak topology*. That is, the weak topology is the collection of all unions of finite intersections of sets $f^{-1}(V)$, with $f \in X^*$ and $V \subset \mathbb{C}$ an open set.

Similarly, if X is separating for X^*, define the *weak*-topology* as the X-topology on X^* by noting that every $x \in X$ defines a linear functional f_x on X^* via $f_x(\Lambda) = \Lambda(x)$ for $\Lambda \in X^*$.

The bounded linear functionals on some Hilbert space $\mathcal{H}$ are an insightful example to which these definitions can be applied. If ϕ is a bounded linear functional on $\mathcal{H}$, then it can be uniquely identified with a vector $|\tau\rangle \in \mathcal{H}$ such that $\phi(|\omega\rangle) = \langle\tau|\omega\rangle$. Hence, $\mathcal{H}$ can be identified with its dual and the weak- and weak*-topologies are both induced by the inner product. That is, a sequence $\{|\tau_i\rangle\}_i$ converges weakly to $|\tau\rangle$, denoted $|\tau_i\rangle \to |\tau\rangle$ weakly, if $\langle\tau_i|\omega\rangle \to \langle\tau|\omega\rangle$ for all $|\omega\rangle \in \mathcal{H}$.

There are several useful ways to topologize the bounded linear operators on a Hilbert space $B(\mathcal{H})$. Since $B(\mathcal{H})$ comes equipped with a norm, the most obvious topology is given by the norm topology: $X_i \to X$ in norm if $\|X - X_i\| \to 0$.

The following topologies make use of the norm topology and weak topology on $\mathcal{H}$ respectively.

Definition 1.1.8. [43, I.3.1.1] A sequence $X_i \to X$ in the *strong operator topology*

if $X_i \ket{\omega} \to X \ket{\omega}$ for all $\ket{\omega} \in \mathcal{H}$. Similarly, $X_i \to X$ in the *weak operator topology* if $\braket{\tau|X_i|\omega} \to \braket{\tau|X|\omega}$ for all $\ket{\tau}, \ket{\omega} \in \mathcal{H}$.

If there is no confusion, the strong and weak operator topologies are sometimes also simply called the strong and weak topologies in the literature. In this thesis, however, we try to avoid confusion and will always use the full names.

The following theorem shows that in many relevant cases, the closures with respect to the strong and weak operator topologies coincide.

Theorem 1.1.9. [44, 5.1.2] *If $K \subset B(\mathcal{H})$ is a convex subset, then the weak and strong operator closures of K coincide.*

The strong and weak operator topology often appear in the context of *von Neumann algebras*.

Definition 1.1.10. A von Neumann algebra is a unital C^*-algebra $\mathcal{A}$, acting on a Hilbert space $\mathcal{H}$, that is closed in the weak operator topology.

We denote by $\mathcal{A} \bar{\otimes} \mathcal{B}$ the von Neumann tensor product of two algebras, i.e. the closure of the algebraic tensor product with respect to the weak operator topology.

There is another useful property that defines von Neumann algebras in a way that is equivalent to definition 1.1.10. Let the *commutant* $\mathcal{A}'$ of $\mathcal{A} \subset B(\mathcal{H})$ be the set

$$\mathcal{A}' = \{x \in B(\mathcal{H}) : [x, a] = 0 \quad \forall a \in \mathcal{A}\},$$

with $[x, a] = xa - ax$. Then a von Neumann algebra $\mathcal{A}$ is a C^*-algebra acting on a Hilbert space with the property $\mathcal{A}'' = \mathcal{A}$. Lemma 1.1.11 below summarizes some useful relations between commutants, von Neumann algebras, and weak (strong) operator closures.

Lemma 1.1.11.   1. For any set S it holds that $(S'')' = S'$ [43, I.2.5.3].

2. (Bicommutant Theorem [44, 5.3.1]) If $\mathcal{A}$ is a unital $*$-algebra acting on a Hilbert space $\mathcal{H}$, then the weak and strong operator closures of $\mathcal{A}$ coincide with $\mathcal{A}''$.

3. (Commutation Theorem for tensor products [43, II.4.5.8]) Let $\mathcal{A}$ and $\mathcal{B}$ be von Neumann algebras on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. Then

$$(\mathcal{A} \bar{\otimes} \mathcal{B})' = \mathcal{A}' \bar{\otimes} \mathcal{B}'. \tag{1.11}$$

4. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. Then $B(\mathcal{H}_1) \bar{\otimes} B(\mathcal{H}_2) = B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ [43, III.1.5.4].

The results stated above mainly pertain to the structure of the algebras. The lemma below instead summarizes some results that will be relevant, e.g. in Chapter 3, for the classification of states, which are the positive normalized elements of the dual of the algebras.

Lemma 1.1.12.   *1. (Hahn-Banach Theorem [45, Thm. 3.3]) Let M be a subspace of a vector space X, p a seminorm on X, and f a linear functional on M such that*

$$|f(x)| \leq p(x) \quad \forall x \in M.$$

Then, f extends to a linear functional Λ on X that satisfies

$$|\Lambda(x)| \leq p(x) \quad \forall x \in X.$$

2. *(Krein-Milman Theorem [45, Thm 3.23]) Let X be a topological vector space with a dual X^* that separates points in X. If K is a nonempty compact convex subset of K, then K is the closed convex hull of the set of its extreme points.*

3. *(Banach-Alaoglu [45, Thm. 3.15]) If V is a neighborhood of 0 in a topological vector space X and if*

$$K = \{\Lambda \in X^* : |\Lambda(x)| \leq 1 \quad \forall x \in V\}$$

then K is weak*-compact.

The Hahn-Banach theorem can be applied to functionals (e.g. states) that are only defined on a subspace of a C^*-algebra in order to show that they can be extended to the entire C^*-algebra. Krein-Milman can be used to decompose states into extremal, pure states. The Banach-Alaoglu theorem can also be applied to the state space of a C^*-algebra, since the norm of states is bounded by 1.

1.2   Causal structures and models of locality

This section sets up the formalism for causal structures and causal networks. To define non-trivial causal models, one needs a notion of (independent) local subsystems. Though intuitively clear, it turns out that locality is a surprisingly subtle concept to formalize mathematically, depending on the theory of nature one prescribes to. From the perspective of special relativity, two events in space-time cannot be causally related if they are spatially separated. Since we believe that the predictions of special relativity hold true, our theories should also reflect this property. Which other restrictions we impose depends on the theory we choose. Below are two such models:

classical and quantum theory. Since this thesis mainly pertains to the latter, quantum theory will receive the most attention in the subsequent sections. It is also possible to define causal models in the broader context of generalized probabilistic theories (GPTs) [46], but this is outside the scope of this thesis.

We will assume some familiarity with probability theory and graph theory. We will write random variables in upper case letters, while specific outcomes are denoted by lower case letters. For example, the random variable X can take values in the set $\{x\}_x$. For probabilities, we will often use the shorthand notation $P(x) = P(X = x)$.

1.2.1 Classical causal models

To analyze causality in statistical models we will make use of a graphical notation. Our focus will be on directed acyclic graphs (DAGs)[2]. In such graphs, each of the vertices represents a random variable, while each directed edge represents a functional relationship. The outcome of each variable X_i is completely determined by its graph theoretical parents, $\mathrm{PA}(X_i)$, and some independent local randomness. That is, if for a variable X_i with outcomes $\{x_i\}$ we denote the outcomes of the parents of X_i by $pa(X_i)$, and the local randomness of X_i by Λ_{X_i}, then there is a function f_i such that

$$x_i = f_i(pa(X_i), \lambda_{X_i}). \tag{1.12}$$

The set of such *structural equations* completely determines the causal model [18]. It allows us to write the joint probability distribution as

$$P(x_1, \ldots, x_n) = \prod_i P(x_i|pa(X_i)). \tag{1.13}$$

Here, the local randomness has been suppressed, but could be added back in via the identification

$$P(x_i|pa(X_i)) = \sum_{\lambda_i} P(x_i|pa(X_i), \lambda_i)P(\lambda_i).$$

We will use the connection between the graphical model and the functional dependencies as our definition of a classical causal structure.

Definition 1.2.1. cf. [18, Defs. 1.2.2 and 2.2.1] A *causal structure* of a set of (classical) variables V is a DAG G in which each node corresponds to a unique element of

---

[2] More general graphs with undirected or bidirected edges also have their use in causal inference, but will not be very relevant for our discussions here. See e.g. [18] for more details.

V and each edge represents a direct functional relationship among the corresponding variables.

If a probability distribution admits the factorization of (1.13) relative to G, we say that P is *compatible* or *Markov compatible* with G.

In some abuse of terminology, we will sometimes also refer to a causal structure as a causal model. Technically speaking, a causal model requires both a causal structure and a set of functional dependencies as in Eq. (1.12).

In general, when analyzing data, one does not have access to all of the variables in a causal structure. The ones that can be accessed are called *observed* (or observable) variables, while those that cannot be measured are called *unobserved* or *latent* variables. In the graphical notation, a boxed variable corresponds to an observed variable, while a circled one corresponds to a latent variable (see e.g. Fig. 1.1).

As explained in the introduction, it is a very relevant, but in general quite difficult problem to determine whether a probability distribution over observed variables is compatible with a causal structure. Determining which distributions are compatible is known as the *causal compatibility problem* or *causal hypothesis testing*.

The functional dependencies also imply certain (conditional) independence relations between variables. For example, if there is no directed path connecting two variables to each other, they must be independent. A more general conclusion can be drawn from a property called *d-separation* (where d stems from the word 'directional').

Definition 1.2.2. cf. [18, Def. 1.2.3] A path p is said to be d-separated by a set of nodes Z if and only if

1. p contains a chain $i \rightarrow m \rightarrow j$ or a fork $i \leftarrow m \rightarrow j$ such that $m \in Z$, or

2. p contains an inverted fork, also known as a collider, $i \rightarrow m \leftarrow j$ such that neither m nor any of its descendants are in Z.

A set Z is said to d-separate X from Y if and only if Z d-separates all paths from X to Y.

Theorem 1.2.3. [18, Thm. 1.2.4] *If sets X and Y are d-separated by Z in a DAG G, then X is independent of Y conditional on Z in every distribution compatible with G.*

Conversely, if X and Y are not d-separated by Z, then there exists a distribution compatible with G such that X and Y are dependent conditional on Z.

In fact, for the converse part of theorem 1.2.3, almost all compatible distributions will have a conditional dependency of X and Y. It would require quite precise fine-tuning of the variables for conditional independence to still hold. This motivates

us to apply Occam's razor to the causal compatibility problem: In principle, there is an unbounded set of compatible causal explanations for any observed probability distribution, since extending a causal structure by adding more variables and dependencies only allows for more compatible distributions. However, since observed (conditional) independence would require precise fine-tuning we generally consider it reasonable to choose the "simplest" model that allows the observed statistics. To be more precise, a causal structure G is preferred over a different structure G', sometimes written $G \leq G'$, if G' can mimic all of the statistics compatible with G. If $G \leq G'$ and $G' \leq G$, then G and G' are said to be (observationally) equivalent [18, Def. 2.3.4].

Let us illustrate the formalism with an example: Consider the graph in Fig. 1.1(a), known as the triangle scenario. The value of the random variable A is determined by the values that X and Z take (up to some local randomness). Similarly, B is determined by X and Y, and C is determined by Y and Z. Since X, Y and Z have no parents, they only depend on local randomness, and are independent of each other. This also allows us to write down the form that the joint probability distribution of all the variables must take. We get

$$P(a,b,c,x,y,z) = P(a|b,c,x,y,z)P(b|c,x,y,z)P(c|x,y,z)p(x,y,z) \quad (1.14)$$
$$= P(a|x,z)P(b|x,y)P(c|y,z)p(x)p(y)p(z), \quad (1.15)$$

where the second equality follows from the d-separation properties of the triangle graph and indeed corresponds to what we defined as a compatible distribution according to definition 1.2.1.

To refer back to the discussion of locality at the start of this section: In the case of classical causal structures locality is enforced by requiring that the joint probability distribution of all the variables in the causal structure factorizes according to Eq. 1.13. This then has implications on the dependencies of the observed (conditional) distribution. Even though Eq. (1.13) is a fairly simple equation to interpret, determining whether an observable probability distribution is compatible with a global distribution of this form remains a highly nontrivial problem.

1.2.2 Quantum causal models

The following section contains excerpts of sections 1.2, 2.1 and 2.3 of Paper [1] and combines the notions introduced there to define quantum causal structures.

It is natural to generalize the causal hypothesis testing problem to *quantum causal structures* [11, 22, 47, 48, 49, 50]. A conceptual difference to classical causal models is that, due to the no-broadcasting theorem, quantum states cannot be both measured and also serve as an input for further processing. Here, we mainly focus on the subset
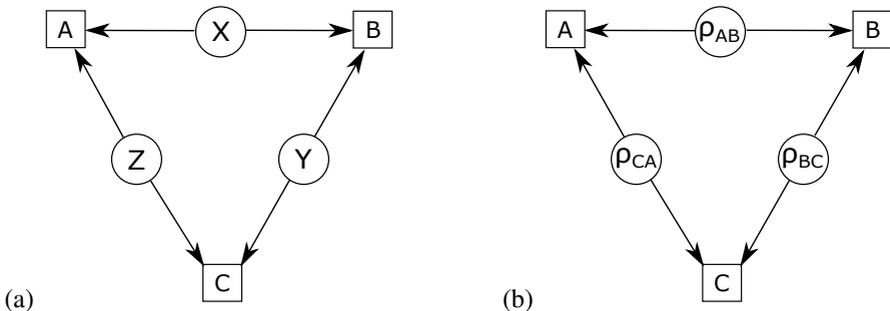
Figure 1.1: The *triangle scenario* will be an important example of a causal structure that we will refer to multiple times in this thesis. Round vertices denote latent, or unobserved systems, while square vertices denote observed ones. Arrows represent causal relationships. Even though the triangle features only three observed and three unobserved variables, it is nevertheless highly non-trivial to analyze. This is in part due to the fact that its description lacks independence constraints on the level of the observed variables. Fig. (a) depicts the classical case, in which the unobserved systems are assumed to be classical random variables. In the quantum case depicted in Fig. (b), the unobserved systems are bipartite quantum states. Each of the quantum systems is then distributed over a channel that is represented by an outgoing arrow. At each square vertex, a measurement is performed simultaneously on all incoming quantum systems at that node, resulting in a classical random variable.

of quantum causal structures known as *correlation scenarios* [47]. These comprise one layer of hidden nodes and one layer of observed nodes, with arrows pointing from hidden to observed ones (more general causal structures are discussed in Section 3.2.3).

The input to the causal hypothesis test for correlation scenarios is a bipartite directed graph and a joint probability distribution with one classical variable corresponding to every observed node (see Fig. 1.1(b) and Fig. 3.7 for examples). The problem is then to decide whether the classical distribution could have arisen from the following process:

1. For each hidden node, prepare a quantum state on as many systems as there are outgoing arrows from that node. The quantum state can be entangled among the subsystems, but the states associated with different latent nodes must be independent. Then, distribute the subsystems along the arrows to the observed nodes.

2. At each observed node, perform a global measurement on all incoming quan-

tum systems. Assign the result to the observed random variable.

To give an example, we again take the triangle scenario, which is depicted for the quantum case in Fig. 1.1(b). The latent variables are now quantum systems with quantum states that are labeled according to the observer to which they are sent. For example, ρ_{AB} is a bipartite quantum state, of which the first part is sent to Alice and the second part to Bob. The arrows indicate independent quantum channels and whenever an arrow ends in a classical observable node, a measurement is performed at that node. Nodes that do not have any incoming edges are called root nodes and are assumed to be prepared in independent initial states. Note the abuse of notation that is commonly used in the graphical representation of quantum causal structures: The hidden nodes are labeled by quantum states, as opposed to the quantum systems on which they live. This is in contrast to the classical case, where the hidden nodes are labeled by the random variables and not by their probability distributions. Here we have opted to adopt this commonly used abuse of notation, as the quantum state is generally considered as the more central object.

The formal definition of a quantum causal structure depends crucially on the notion of a *subsystem*. Here, we describe two subtle modeling decisions that have to be made when making this term precise.

In elementary quantum mechanics, the central object that characterizes a quantum system is its Hilbert space. In this framework, one thus associates to each subsystem a Hilbert space $\mathcal{H}_i$ and takes the joint Hilbert space to be their tensor product $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$. The set of observables is then derived from the Hilbert space structure. For the individual subsystems, the observables are the linear operators $\mathcal{A}_i = L(\mathcal{H}_i)$. They can be embedded into $\mathcal{A}_{12} = L(\mathcal{H}_{12})$ by taking the tensor product with identities on the other subsystem:

$$\mathcal{A}_1 \simeq \mathcal{A}_1 \otimes \mathbb{1}, \qquad \mathcal{A}_2 \simeq \mathbb{1} \otimes \mathcal{A}_2. \tag{1.16}$$

In contrast, in *algebraic quantum mechanics* (cf. [51, Chapter 8], [52]), the set of observables is seen as being more central. Consequently, one associates an observable algebra $\mathcal{A}_i$ with each subsystem. A joint system is then any algebra $\mathcal{A}_{12}$ that contains $\mathcal{A}_1$ and $\mathcal{A}_2$ as commuting subalgebras and is generated by them. Clearly, the construction in (1.16) provides an example of algebras standing in such a relation, but it turns out that there are more general scenarios that cannot be realized using Hilbert space tensor products.

Commutativity has physical consequences, e.g. in terms of joint measurability. So if we accept the quantum-mechanical description of observable phenomena in terms of operators, we are forced to conclude that measurements in space-like separated regions are described in terms of commuting operators. However, the stronger requirement that the underlying Hilbert space forms a tensor product is not obviously

physically motivated.

For a long time, it was an open problem to decide whether there are correlations that can be realized by performing measurements on commuting operators, but not on operators acting on distinct factors of a tensor product Hilbert space. This relates to the question when the minimal and maximal tensor product of two algebras coincide. In quantum information theory, this question has been known as *Tsirelson's problem* and was shown to be equivalent to other long-standing open problems in operator theory, most notably *Connes' embedding problem* [53, 54, 55]. In a recent breakthrough result, these questions have been decided: The commuting-operator model *does* capture more general correlations than the tensor-product model [17].

The above raises the question which of the two mathematical models to adopt. Here, we take a pragmatic approach. It has long been realized (and in fact, has historically triggered Tsirelson to speculate) that commutativity is easily encoded as a constraint in SDPs that give outer approximations to the set of quantum correlations [56]. The same is *not* true for the tensor product property. Since either model is legitimate, but one is a better fit for the SDP hierarchies we want to make a statement about, we opt for the approach in which locality is modeled by commutativity. Therefore, in this work, we will assume throughout that one can associate an algebra of observables with each party and that these algebras commute.

The second subtlety deals with the boundedness of operators. In this thesis, we are mostly interested in observable probabilities that describe measurements on quantum systems. Probabilities are associated with elements of a positive operator-valued measure (POVM), which are bounded: Their operator norm does not exceed 1. It follows that the entire observable algebra generated by POVM elements consists of bounded operators. Many problems – e.g. the problem of characterizing the set of correlations compatible with a Bell scenario – can be described solely in terms of this algebra. From a technical point of view, this property can provide significant simplifications. For example, the convergence proofs of the NPO hierarchy [40] or the Quantum de Finetti Theorem for infinite-dimensional quantum systems [24] make central use of the fact that operators are bounded.

It thus comes as bad news that this simplifying property is not obviously available for the causal compatibility problem.

Indeed, consider a node of a quantum causal structure, say the one that gives rise to the random variable A in the triangle scenario. Each possible outcome $A = a$ is associated with a POVM element E_a. As there are two incoming arrows to this vertex, E_a acts on two quantum systems. We therefore assume that the observable algebra $\mathcal{A}$ of the joint system is generated by two commuting subalgebras $\mathcal{A}_-, \mathcal{A}_+$. These local algebras play an important role in the definition of the causal structure: It is with respect to them that the state is required to factorize (see Eq. (1.18) below).

But, while E_a is bounded, we are not aware of any result that would imply that one can assume the same is true for elements of $\mathcal{A}_-, \mathcal{A}_+$.

More concretely, we cannot exclude the possibility that there is a mathematical model of "local quantum systems" in which one can assign a precise meaning to the series

$$E_a = \sum_{\alpha=1}^{\infty} e_-(a, \alpha) e_+(a, \alpha) \qquad (1.17)$$

for suitable *unbounded* operators $e_-(a, \alpha) \in \mathcal{A}_-, e_+(a, \alpha) \in \mathcal{A}_+$, but where no such expression for E_a exists if the e_-, e_+'s are required to be bounded.

In our precise definition of a quantum causal model, we will *assume* that it is not necessary to allow for such singular situations. The convergence proofs in later sections make use of this assumption (implicitly, by virtue of being phrased in terms of C^*-algebras, which model bounded operators).

While it is an interesting question about operator algebras whether the assumption is actually necessary, it seems that under mild *physical* conditions, observed correlations can be approximated using models for which it is valid. For example, if each subsystem is endowed with a non-degenerate Hamiltonian and the state has finite energy, one can always compress the local observables to finite-dimensional low-energy subspaces on which they are obviously bounded. So as long as not *both* the observable and the state display rather singular behavior, an approximate bounded model should always be possible in physical situations.

We can now state the definition of quantum causal structures in mathematically precise terms. Here we restrict attention again to the triangle scenario (Fig. 1.1(b)) as a guiding example.

Let A, B and C be random variables. We say that a probability distribution $P(A, B, C)$ is *compatible* with the quantum triangle scenario, if it can be realized in the following mathematical model.

Assume that there is a C^*-algebra $\mathcal{D}$ that is generated by commuting subalgebras $\mathcal{A}, \mathcal{B}, \mathcal{C}$ that are each associated with a vertex of the triangle. Each of the algebras $\mathcal{A}, \mathcal{B}, \mathcal{C}$ is in turn generated by two commuting subalgebras: $\mathcal{A}$ by $\mathcal{A}_-, \mathcal{A}_+$; $\mathcal{B}$ by $\mathcal{B}_-, \mathcal{B}_+$; and $\mathcal{C}$ by $\mathcal{C}_-, \mathcal{C}_+$. They model the observables measurable on the subsystems that enter the respective node from either side in the diagram (that is, $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$ are C^*-tensor products in the sense of Sec. 1.1.2 – but we take no stance on which particular one). Next, we assume that there is a state ρ on $\mathcal{D}$ that factorizes according to the independence implied by the causal structure, i.e. a product state of the form

$$\rho(A_- A_+ B_- B_+ C_- C_+) = \rho(C_+ A_-)\rho(A_+ B_-)\rho(B_+ C_-) \qquad (1.18)$$

where $A_- \in \mathcal{A}_-, A_+ \in \mathcal{A}_+, B_- \in \mathcal{B}_-$ and so on. Finally, we assume that there are POVMs

$$\{E_a\}_a \subset \mathcal{A}, \quad \{F_b\}_b \subset \mathcal{B}, \quad \{G_c\}_c \subset \mathcal{C} \tag{1.19}$$

such that the joint distribution can be realized as

$$P(a, b, c) = \rho(E_a F_b G_c). \tag{1.20}$$

The definition for a quantum causal structure is now readily extended to other correlation scenarios.

Remark. The GNS construction applied to two commuting algebras acting on a product state gives rise to a tensor product Hilbert space. Thus, by (1.18), there is no loss of generality in assuming that $\mathcal{A} = \mathcal{A}_- \otimes_{\min} \mathcal{A}_+$ and likewise for $\mathcal{B}$ and $\mathcal{C}$. However, the same is *not* true for the tensor products between $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$. (Certainly the same argument doesn't apply – as ρ does not factorize as a state between the nodes. One can combine the results from Refs. [17, 47] to see that there are correlations P that cannot be modeled using minimal tensor products between nodes at all). This observation does not obviate the need for a generalized Quantum de Finetti Theorem, as we will apply it to the state that is extracted from the SDP hierarchy, and there is no semidefinite constraint that can express that $\mathcal{A}$ is a minimal, rather than a general, tensor product of its constituents.

Remark. More generally, one can think of nodes with incoming edges as performing a quantum channel on the systems they receive. In the case of correlation scenarios, where such systems are always leaf nodes that perform a measurement in the end, this generalization does not add much, since the effect of the quantum channel can be absorbed into the POVM. Later, in Chapter 4, where the output of the quantum channel can also be a quantum state, this distinction will become more relevant.

1.3 Semidefinite Programming

The third main topic of this thesis is semidefinite programming. A semidefinite program (SDP) is a conic convex optimization program, where one optimizes over positive semidefinite (PSD) matrices under linear matrix inequalities. Such optimization programs have many practically relevant use cases, including, but not limited to, system and control theory [57]; experiment design [58, 59]; combinatorial optimization [60, 61], among which the well-known NP-hard MAX-CUT [62] and traveling

salesperson [63] problems; and of course problems in quantum information theory [56, 64, 65, 66].

Though convex conic optimization is in general NP-hard[3], SDPs can be solved in polynomial time in the input size within a fixed approximation error using interior point methods under some mild assumptions [68]. Additionally, if the problem turns out to be infeasible, we can extract a seperating hyperplane as a certificate of infeasibility, which often has a useful interpretation as a so-called 'witness', separating the feasible region from (a part of) the infeasible region.

Here we explain some of the basic concepts of semidefinite programming. Some special attention will be given to SDP relaxations for non-commutative polynomial optimization, which is a central ingredient in our approach to analyze the quantum causal compatibility problem.

1.3.1   Standard form of SDPs

In semidefinite programming, we optimize some linear objective function over the cone of positive semidefinite $n \times n$ matrices, which is a subset of the real[4] symmetric matrices $\mathcal{S}^n$. The PSD cone is defined as

$$\mathcal{S}^n_{\succeq 0} = \{X \in \mathcal{S}^n : X \text{ is PSD}\}.$$

Furthermore, we use $\langle A, B \rangle$ to denote the trace inner product $\operatorname{tr}(AB)$ on $\mathcal{S}^n$. There are several ways to characterize PSD matrices, as outlined in the theorem below.

Theorem 1.3.1. [69, Thm. 1.7.2] *Let $X \in \mathcal{S}^n$ be a symmetric $n \times n$ matrix. Then the following properties are equivalent.*

1. *X is positive semidefinite.*

2. *$x^T X x \geq 0$ for all $x \in \mathbb{R}^n$.*

3. *The smallest eigenvalue of X is non-negative.*

4. *$X = LL^T$ for some $L \in \mathbb{R}^{n \times k}$. This is called a Cholesky decomposition of X.*

5. *$X_{ij} = v_i^T v_j$ for some set of vectors $\{v_i\}_{i=1}^n$ in $\mathbb{R}^n$. This is called a Gram decomposition.*

---

[3] This can be concluded from the fact that one can rewrite any quadratic optimization problem as a convex conic problem [67]. Since there exist NP-hard quadratic problems (e.g. max clique), this implies that convex conic optimization is also NP-hard.

[4] One can in principle also consider complex matrices. Most of the statements in this section carry over to this case, where 'symmetric' should then be replaced by 'self-adjoint'. Additionally, in practice we are only optimizing over rational numbers.

6. *All principal minors of X are non-negative.*

Additionally, we have

7. *X is positive semidefinite if $\langle X, Y \rangle \geq 0$ for all $Y \in \mathcal{S}^n_{\succeq 0}$.*

We use 2. as the defining property of PSD matrices. Note that 7. implies that the PSD cone is self-dual.

Additionally, in our optimization we may wish to restrict to an affine subspace of the cone of PSD matrices. This leads to an optimization problem of the form

$$
\begin{aligned}
p^* = {}& \sup_{X} \ \langle C, X \rangle \\
\text{s. t.} \quad & \langle A_j, X \rangle = b_j, \\
& X \succeq 0,
\end{aligned}
\tag{1.21}
$$

where $X, A_j \in \mathcal{S}^n$ and $b_j \in \mathbb{R}$ for all j. The matrix C represents the objective function that we are trying to optimize. The intersection of the PSD cone with the affine subspace defined by the matrices $\{A_j\}_j$ and data $b = (b_j)_j$ is called the feasible region. If there is a solution such that X is positive definite, then the program is said to be *strictly feasible*.

Equation (1.21) is referred to as the *primal* SDP in standard form. There is also a *dual* SDP in standard form, which looks like

$$
\begin{aligned}
d^* = {}& \inf_{y} \ b^T y \\
\text{s. t.} \quad & \sum_{j} y_j A_j - C \succeq 0,
\end{aligned}
\tag{1.22}
$$

where y is a real vector indexed in the number of linear constraints. The matrices $\{A_j\}_j$ and C, together with the vector b are thus sufficient to define both the primal and the dual standard form of a semidefinite program.

Under certain conditions, the values p^* and d^* coincide. This is summarized in the theorem below.

Theorem 1.3.2. *[69, Lemma 2.1.1 and Thm. 2.1.2]*

1. **(weak duality)** *Let (X, y) be a primal/dual pair of feasible solutions for the SDP defined by $\{A_j\}_j, b, C$. Then $\langle C, X \rangle \leq b^T y$ and thus $p^* \leq d^*$.*

2. **(strong duality)** *If (1.21) is bounded from above and strictly feasible, then the dual attains its infimum and there is no duality gap, i.e. $p^* = d^*$. Similarly if (1.22) is bounded from below and strictly feasible, the primal attains its supremum and there is no duality gap.*

It is not difficult to prove weak duality. If X and y are feasible points to a pair of primal and dual SDPs, then

$$\langle C, X \rangle \leq \langle \sum_j y_j A_j, X \rangle = \sum_j y_j b_j = b^T y.$$

Strong duality is often much more difficult to prove. Though it turns out that in many practically relevant use cases the duality gap does indeed vanish.

Let us briefly comment on the mild assumptions that are needed for polynomial-time convergence of semidefinite programming [68]. Let F denote the feasible region of a semidefinite programming problem with rational input. Furthermore, we denote by $B(X, r)$ the unit ball (in the Frobenius norm, i.e. the norm derived from the trace inner product) around $X \in \mathcal{S}^n$ of radius r. If we know a rational point $X_0 \in F$ and radii r and R such that

$$B(X_0, r) \subseteq F \subseteq B(X_0, R), \tag{1.23}$$

then we can find a feasible point X^* such that $\langle C, X^* \rangle - p^* \leq \varepsilon$ in time polynomial in the dimension n, the number of constraints m, $\log_2(R/r)$, $\log_2(1/\varepsilon)$ and the bit size of the input data [68, 70].

In other words, the feasible region cannot be too small or too big. In the SDP hierarchies that we will use in this thesis, the operators we optimize over are bounded in such a way that the feasible region is indeed not too big [71]. However, we make no statement about the smaller radius r. In the literature on semidefinite programming, some results are known on when the radius r is provably large enough (see e.g. Ref. [71]), but it is not clear whether the requirements for such statements always hold in the SDP hierarchies that are used in this thesis. Instead, we will simply *assume* that the feasible region is not more than exponentially small. Alternatively, many of these SDPs can be phrased as a type of eigenvalue optimization problem for which the feasible region is the entire state space of the C^*-algebra, such that X_0 and r can be chosen large enough for $\log_2(R/r)$ to be at most polynomial.

1.3.2 Non-commutative polynomial optimization

In a series of papers [40, 56, 72] Navascués, Pironio and Acín developed a method for solving non-commutative polynomial optimization (NPO) problems, using a convergent hierarchy of SDP relaxations. This was a generalization of the hierarchy developed by Lasserre [73] for the commutative case.

Here we outline an equivalent formulation in terms of optimization problems over states on C^*-algebras, as developed in Paper [1, Sec. 2.2.3].

Given generators $\mathcal{G}$ and relations $\mathcal{R}$, we are interested in certain linear optimization problems over states on the algebra they generate.

In Ref. [40], the NPO problem is phrased as an optimization over representations π of the free $*$-algebra $\mathcal{F}(\mathcal{G})$ and normalized vectors $|\phi\rangle$ in the representation space. Concretely, choose an element $y_0 \in \mathcal{F}(\mathcal{G})$ and a countable set $\{y_1, y_2, \ldots\} = \mathcal{Y} \subset \mathcal{F}(\mathcal{G})$ and consider:

$$
\begin{aligned}
f_{\text{NPO}}^\star = \min_{\pi, |\phi\rangle} \quad & \langle\phi|\pi(y_0)|\phi\rangle \\
\text{s. t.} \quad & \pi(q) \succeq 0 \qquad q \in \mathcal{R}, \\
& \langle\phi|\pi(y_i)|\phi\rangle \geq 0 \qquad y_i \in \mathcal{Y}.
\end{aligned}
\tag{1.24}
$$

We prefer to think of this problem more abstractly, as an optimization over the state space of the universal algebra $C^*(\mathcal{G}|\mathcal{R})$:

$$
\begin{aligned}
f_{\text{uni}}^\star = \min_{\rho \in K(C^*(\mathcal{G}|\mathcal{R}))} \quad & \rho(y_0) \\
\text{s. t.} \quad & \rho(y_i) \geq 0 \qquad y \in \mathcal{Y}.
\end{aligned}
\tag{1.25}
$$

We may write $\min$ instead of $\inf$, because the Banach-Alaoglu Theorem implies that the state space is weak*-compact and thus that the infimum over states evaluated on any fixed element of the algebra is attained. Following [40, Section 3.6], one can in addition impose constraints of the form $\rho(\cdot z) = 0$ for a countable set of $z \in \mathcal{F}(\mathcal{G})$. We have omitted this type of constraint from the discussion, as it is not needed for our use cases.

Lemma 1.3.3. *The solutions of (1.24) and (1.25) coincide.*

Proof. Let ω be an optimizer of (1.25). Let π_ω be the GNS representation and $|\Omega\rangle \in \mathcal{H}_\omega$ the vector that implements ω. By Lemma 1.1.5, (π_ω, Ω) is feasible for (1.24) and achieves the optimal value of (1.25).

Conversely, let (π, ϕ) be an optimizer of (1.24). For $x \in \mathcal{F}(\mathcal{G})$, define $\rho(x) := \langle\phi|\pi(x)|\phi\rangle$. If the seminorm vanishes on x, $\|x\| = 0$, then, in particular, $\|\pi(x)\| = 0$ and hence $\pi(x) = 0$. Thus, ρ is constant on cosets of the ideal of elements on which the seminorm vanishes and therefore well-defined as a functional on $C^*(\mathcal{G}|\mathcal{R})$. As such, it is feasible for (1.25) and achieves the optimal value of (1.24). $\qquad\square$

We now briefly describe the semidefinite programming hierarchy introduced in Ref. [40] and sketch the completeness proof.

Of course, the difficulty in solving (1.25) lies in the fact that $C^*(\mathcal{G}|\mathcal{R})$ is, in general, infinite-dimensional. The broad idea behind the NPO hierarchy is to partition $\mathcal{F}(\mathcal{G})$ into an increasing family of finite-dimensional subspaces $\mathcal{F}^{(k)} \subset \mathcal{F}(\mathcal{G})$.

At the k-th level of the hierarchy, one imposes the conditions that ρ be a state and that the relations be fulfilled only to the extent to which they can be expressed using elements from $\mathcal{F}^{(2k)}$.

To carry out this program, let $\mathcal{F}^{(k)}$ be the space of all elements $x \in \mathcal{F}(\mathcal{G})$ that can be expressed as a polynomial in the generators and their adjoints of degree at most k. We fix some basis $\{b_1^{(k)}, \ldots, b_{d_k}^{(k)}\}$ of each $\mathcal{F}^{(k)}$.

Recall that a linear functional ρ on $C^*(\mathcal{G}|\mathcal{R})$ is a state if and only if $\rho(\mathbb{1}) = 1$ and $\rho(x^*x) \geq 0$ for all $x \in C^*(\mathcal{G}|\mathcal{R})$. We impose a related condition by demanding that the matrix $\Gamma^{(k)}$ with elements

$$\Gamma_{ij}^{(k)} = \rho\big(b_i^{(k)*} b_j^{(k)}\big), \qquad i,j = 1, \ldots, d_k \tag{1.26}$$

be positive semidefinite and that $\Gamma_{1,1}^{(k)} = \rho(\mathbb{1}) = 1$. A matrix that obeys Eq. (1.26) is referred to as a *moment matrix*.

Next, consider a relation $q \in \mathcal{R}$. Let l be the smallest integer such that $q \in \mathcal{F}^{(2l)}$. We relax the requirement that q be positive to demanding that the matrix

$$(\Lambda_q^{(k)})_{ij} = \rho(b_i^{(k-l)*} q b_j^{(k-l)}), \qquad i,j = 1, \ldots, d_{k-l}$$

be positive semidefinite. The matrices Λ_q are called *localizing matrices*.

Let k_0 be such that $x \in \mathcal{F}^{(2k_0)}$. For each $k \geq k_0$, one thus arrives at a relaxation of Eq. (1.25) in terms of the semidefinite program

$$
\begin{aligned}
f^k = \min_{\rho \in (\mathcal{F}^{2k})^*} \quad & \rho(y_0), \\
\text{s. t.} \quad & \rho(\mathbb{1}) = 1, \\
& \Gamma^{(k)} \succcurlyeq 0, \\
& \Lambda_q^{(k)} \succcurlyeq 0 \qquad q \in \mathcal{R} \cap \mathcal{F}^{(2k)}, \\
& \rho(y_i) \geq 0 \qquad y_i \in \mathcal{Y} \cap \mathcal{F}^{(2k)}.
\end{aligned}
\tag{1.27}
$$

The completeness result of Ref. [40] states that, in the case where $|\mathcal{Y}| \leq \infty$ is finite, the optimal values f^k of the relaxations (1.27) converge to $f_{\mathrm{NPO}}^\star = f_{\mathrm{uni}}^\star =: f^\star$ from below.

Lemma 1.3.4. *The completeness result* $\lim_{k \to \infty} f^k = f^\star$ *extends to the case of a countably infinite number of inequality constraints* $\rho(y) \geq 0$.

Proof. Choose some enumeration $y_1, y_2, \ldots$ for the countable set $\mathcal{Y}$. Let $\mathcal{Y}_s = \{y_1, \ldots, y_s\}$. Assume that (1.27) with $\mathcal{Y}$ replaced by $\mathcal{Y}_s$ is feasible for every k, s, with optimal value f_s^k. Using the convergence proof of Ref. [40] and Lemma 1.3.3,

there exists a sequence of states $\rho_s^\star \in K(C^*(\mathcal{G}|\mathcal{R}))$ that are feasible for (1.25) with inequality constraints $\mathcal{Y}_s$ and attain $f_s^\star := \lim_{k \to \infty} f_s^k$. By the Banach-Alaoglu Theorem, there is a convergent subsequence. Let $\rho^\star$ be its limit point. Then, for each $y_i \in \mathcal{Y}$, $\rho^\star(y_i) \geq 0$, as this constraint is fulfilled by all but a finite number of the $\rho_s^\star$. Thus, $\rho^\star$ is feasible for (1.25) with all inequality constraints $\mathcal{Y}$ taken into account and attains $f^\star = \lim_{s \to \infty} f_s^\star$. $\qquad\qquad\qquad\qquad\qquad\square$

Remark. If a space $N \subset \mathcal{F}(\mathcal{G})$ of elements with vanishing seminorm is known, then one can replace $\mathcal{F}(\mathcal{G})$ by the quotient space $\mathcal{F}(\mathcal{G})/N$ in the constructions above, while retaining convergence. This can result in significantly smaller matrices that need to be treated. In particular, every equality constraint $x = y$ gives rise to an element $x - y \in N$.

Chapter 2

Quantum de Finetti Theorems

Symmetry is one of the most relevant notions in all of physics. It often allows us to extract information about a system, e.g. in the form of conservation laws, or to reduce the size or complexity of a problem. Symmetries in the form of groups also enable us to use the large body of knowledge that has been accumulated over the centuries on group theory and representation theory. Perhaps it then doesn't come as a surprise that the symmetries that are present in the causal compatibility problems that we encounter in this thesis will also provide us with a strategy to solve those problems. Additionally, symmetries can help to reduce the size of the semidefinite programs that are used to analyze these compatibility problems.

This chapter shows how symmetry of a state relates to the entanglement of that state. More precisely, if a state is symmetric over an infinite number of copies of an algebra, it must be a separable state over any finite number of those copies. Such statements are broadly known as quantum de Finetti theorems [24, 74, 75, 76, 77]. They form a generalization to the classical de Finetti theorems that state that infinitely symmetric distributions must be convex mixtures of product distributions [78, 79]. Similar statements can be made for theories that are even more general than quantum theory. In particular, there has been recent progress on a de Finetti-type theorem for convex cones that can be applied to generalized probabilistic theories (GPTs) [80].

Here we present several quantum de Finetti theorems. In Sec. 2.1 we prove a de Finetti theorem for the maximal tensor product of C^*-algebras. This is a generalization of the works of Størmer [74], and Raggio and Werner [24], in which only the minimal tensor product was treated.

In Sec. 2.2 we prove that the scalar extension method described in Sec. 5.2 also gives rise to a separable state. In fact, the proof is much simpler, due to the fact

that the scalar extensions already form an abelian algebra; something that we have to work hard for to achieve in the other cases.

Lastly, in Sec. 2.3 it is shown that one can still use parts of the construction and proof of the existing quantum de Finetti theorems to analyze the state that arises from the original formulation of quantum inflation. Such a state will be separable over many of the subsystems. While this gives us some insight in the type of states one optimizes over in the inflation method, it is not immediately clear what to do with this information in regard to completeness proofs.

2.1   The Maximal Tensor Product

The following section is largely taken from Paper [1, Sec. 3]

To the best of our knowledge, the existing literature on de Finetti Theorems for infinite systems is phrased only in terms of the minimal tensor product [24, 74]. These results are not directly applicable to the quantum models that result from the NPO hierarchy. Indeed, the latter naturally guarantees the existence of a representation π_ρ of the algebraic tensor product as operators on a Hilbert space that arise from a state ρ on $\mathcal{A}_1 \otimes_{\text{alg}} \ldots \otimes_{\text{alg}} \mathcal{A}_n$ via the GNS construction. While the resulting operator norm $\|\pi_\rho(x)\|$ constitutes a C^*-norm on the tensor product, we have no a priori control over its value beyond the constraints in Eq. (1.6).

The purpose of this section is therefore to retrace the arguments given by Raggio and Werner in Ref. [24] to verify that the infinite de Finetti Theorem established there generalizes to arbitrary C^*-norms on algebraic tensor products. We also present a somewhat simpler formulation that is sufficient for our purposes.

In fact, we state the results only in terms of the maximal C^* tensor product norm. A priori, it is possible that one can derive stronger results for the GNS norm $\|\pi_\rho(\cdot)\|$, in particular if the state ρ is known to have symmetries. We leave this possible improvement open for later investigations.

Let $\mathcal{D}$ be a unital C^*-algebra and let

$$\mathcal{D}^n = \mathcal{D}^{\otimes_{\max} n}$$

be the completion of the algebraic tensor product of n copies of $\mathcal{D}$ with respect to the maximal C^*-norm. The infinite maximal tensor product is defined as the *inductive limit*

$$\mathcal{D}^\infty = \lim_{n \to \infty} \mathcal{D}^n. \tag{2.1}$$

We recall the definition of an inductive limit using the tensor product structure [43,

Section II.8.2]: For any $n, k \in \mathbb{N}$, there is a natural embedding

$$\mathcal{D}^n \to \mathcal{D}^{n+k}, \qquad x \mapsto x \otimes \mathbb{1}^{\otimes k}.$$

It allows us to define addition and multiplication between elements of the union

$$\bigcup_{n=1}^{\infty} \mathcal{D}^n \tag{2.2}$$

by embedding the element living in the smaller tensor power into the larger one and performing the operations there. The resulting $*$-algebra is the *local algebra*, called so as each of its elements lives in a finite tensor power. The inductive limit $\mathcal{D}^\infty$, the *quasi-local algebra*, is the completion of the local algebra with respect to the C^*-norm $\| \cdot \|_{\mathcal{D}^\infty}$ on (2.2) which assigns to every $x \in \mathcal{D}^n$ the value

$$\|x\|_{\mathcal{D}^\infty} = \|x\|_{\mathcal{D}^n} = \lim_{k \to \infty} \|x \otimes \mathbb{1}^k\|_{\mathcal{D}^{n+k}}.$$

We will not notationally distinguish between an element $x \in \mathcal{D}^n$ and its embedding in $\mathcal{D}^\infty$. Note that any $x \in \mathcal{D}^n$ and its extensions $x \otimes \mathbb{1}^k$ are identified in $\mathcal{D}^\infty$.

For any n and permutation $\pi \in S_n$, there is an automorphism α_π on $\mathcal{D}^n$ which acts by permuting tensor factors

$$\alpha_\pi(x_1 \otimes \cdots \otimes x_n) = x_{\pi(1)} \otimes \cdots \otimes x_{\pi(n)}.$$

It extends to any $\mathcal{D}^{n+k}$ by letting π act on the first n tensor factors, and by continuity to $\mathcal{D}^\infty$.

A state ρ on $\mathcal{D}^\infty$ is *symmetric* if $\rho(x) = \rho(\alpha_\pi(x))$ for every $x \in \mathcal{D}^\infty$. Denote the set of symmetric states by $K_s(\mathcal{D}^\infty)$. We aim to show:

Theorem 2.1.1 (Max tensor product Quantum de Finetti Theorem). *Let $\rho \in K_s(\mathcal{D}^\infty)$ be a symmetric state on an infinite maximal tensor product*

$$\mathcal{D}^\infty = \lim_{n \to \infty} \mathcal{D}^{\otimes_{\max} n}.$$

Then there exists a unique probability measure μ over states on $\mathcal{D}$ such that for all $x \in \mathcal{D}^\infty$,

$$\rho(x) = \int_{K(\mathcal{D})} \Pi_\sigma(x) \, d\mu(\sigma),$$

where Π_σ is the infinite symmetric product state on $\mathcal{D}^\infty$ associated with the state σ

on $\mathcal{D}$.

Key to the proof is to show that symmetric states ρ define a state on the *abelian C*-algebra* of *symmetric observables* whose multiplication law is derived from the tensor product. It is established in the general theory of C^*-algebras [52, Chapter 4] that pure states ϕ of abelian algebras are homomorphisms, i.e. that $\phi(xy) = \phi(x)\phi(y)$, and that general states of abelian algebras are unique convex combinations of pure states (cf. the Krein-Milman Theorem of Lemma 1.1.12). Using the fact that in our case, the product xy is related to the tensor product $x \otimes y$, we will obtain the claimed decomposition of ρ as a convex combination of symmetric product states.

To construct the symmetric algebra, define the *symmetrization map*

$$\mathrm{Sym}^n(x) = \frac{1}{n!} \sum_{\pi \in S_n} \alpha_\pi(x), \tag{2.3}$$

and let $\mathrm{Sym}^n(\mathcal{D})$ be the image of $\mathcal{D}^n$ under Sym^n. Define the *symmetric local algebra* to be the set

$$\bigcup_{n=1}^{\infty} \mathrm{Sym}^n(\mathcal{D}) \tag{2.4}$$

with an associative and abelian multiplication law given by the *symmetrized tensor product*

$$\star : \mathrm{Sym}^n(\mathcal{D}) \times \mathrm{Sym}^m(\mathcal{D}) \to \mathrm{Sym}^{n+m}(\mathcal{D}), \qquad x \star y = \mathrm{Sym}^{n+m}(x \otimes y).$$

Our aim is to mimic the construction of $\mathcal{D}^\infty$ to arrive at a *symmetric quasi-local algebra* $\mathrm{Sym}^\infty(\mathcal{D})$. To this end, define embeddings

$$\mathrm{Sym}^n(\mathcal{D}) \to \mathrm{Sym}^{n+k}(\mathcal{D}), \qquad x \mapsto x \star \mathbb{1}^{\otimes k}. \tag{2.5}$$

As was the case for $\mathcal{D}^\infty$, addition between two symmetric local elements can now be defined by embedding the lower power into the higher power and performing the addition there. This convention turns the symmetric local algebra into an abelian *-algebra. One can endow it with a C^*-seminorm [81, Section 6.1] so that the completion $\mathrm{Sym}^\infty(\mathcal{D})$ is an abelian C^*-algebra:

Lemma 2.1.2. *The limit*

$$\|x\|_{\mathrm{Sym}} := \lim_{k \to \infty} \|x \star \mathbb{1}^{\otimes k}\|_{\mathcal{D}^\infty} \tag{2.6}$$

defines a C^*-seminorm on the symmetric local algebra $\cup_n \mathrm{Sym}^n(\mathcal{D})$ fulfilling

$$\|x\|_{\mathrm{Sym}} \leq \|x\|_{\mathcal{D}^\infty}. \tag{2.7}$$

The completion $\mathrm{Sym}^\infty(\mathcal{D})$ is an abelian C^*-algebra.

The central ingredient to the proof is the following combinatorial lemma, which shows that the multiplication on $\cup_n \mathrm{Sym}^n(\mathcal{D})$ inherited from $\cup_n \mathcal{D}^n$ and the newly defined $\star$-multiplication are asymptotically equivalent.

Lemma 2.1.3. *Let $x \in \mathrm{Sym}^m(\mathcal{D})$ and $y \in \mathrm{Sym}^n(\mathcal{D})$. Then*

$$\lim_{k \to \infty} \left\| (x \star \mathbb{1}^{\otimes(k-m)})(y \star \mathbb{1}^{\otimes(k-n)}) - (x \star y \star \mathbb{1}^{\otimes(k-m-n)}) \right\|_{\mathcal{D}^\infty} = 0.$$

Proof. Choose two sets of respective size m, n uniformly at random from $[k] := \{1, \ldots, k\}$. The probability that any given element is contained in both sets is $\frac{m}{k}\frac{n}{k}$. By the union bound, the probability that these two sets intersect at all is not larger than $\frac{mn}{k}$. Thus

$$\|(x \star \mathbb{1}^{\otimes(k-m)})(y \star \mathbb{1}^{\otimes(k-n)}) - (x \star y \star \mathbb{1}^{\otimes(k-m-n)})\|_{\mathcal{D}^\infty}$$

$$= \left\| \frac{1}{(k!)^2} \sum_{\substack{\pi,\pi' \in S_k \\ \pi([m]) \cap \pi'([n]) \neq \emptyset}} \alpha_\pi(x \otimes \mathbb{1}^{\otimes(k-m)}) \alpha_{\pi'}(y \otimes \mathbb{1}^{\otimes(k-n)}) \right\|_{\mathcal{D}^\infty}$$

$$\leq \frac{mn}{k} \|x\|_{\mathcal{D}^\infty} \|y\|_{\mathcal{D}^\infty}.$$

The claim then follows from taking the limit $k \to \infty$ and from the boundedness of x and y. $\qquad\square$

Proof of Lemma 2.1.2. For $x \in \mathrm{Sym}^n(\mathcal{D})$, we have the estimate

$$\|x \star \mathbb{1}\|_{\mathcal{D}^\infty} = \left\| \frac{1}{(n+1)!} \sum_{\pi \in S_{n+1}} \alpha_\pi(x \otimes \mathbb{1}) \right\|_{\mathcal{D}^\infty}$$

$$\leq \frac{1}{(n+1)!} \sum_{\pi \in S_{n+1}} \|\alpha_\pi(x \otimes \mathbb{1})\|_{\mathcal{D}^\infty} = \|x \otimes \mathbb{1}\|_{\mathcal{D}^\infty} \leq \|x\|_{\mathcal{D}^\infty}.$$

Using this estimate repeatedly shows that the sequence $\|x \star \mathbb{1}^{\otimes k}\|_{\mathcal{D}^\infty} = \|(x \star \mathbb{1}^{\otimes(k-1)}) \star \mathbb{1}\|_{\mathcal{D}^\infty}$ is non-increasing and hence convergent. Subadditivity, absolute homogeneity, and invariance under involution of $\|\cdot\|_{\mathrm{Sym}}$ on $\cup_n \mathrm{Sym}^n(\mathcal{D})$ follow directly from the same properties of $\|\cdot\|_{\mathcal{D}^\infty}$. For $x \in \mathrm{Sym}^m(\mathcal{D})$ and $y \in \mathrm{Sym}^n(\mathcal{D})$,

Lemma 2.1.3 implies the C^*-norm property

$$
\begin{aligned}
\|x \star y\|_{\mathrm{Sym}} &= \lim_{k \to \infty} \|x \star y \star \mathbb{1}^{\otimes(k-n-m)}\|_{\mathcal{D}^\infty} \\
&= \lim_{k \to \infty} \|(x \star \mathbb{1}^{\otimes(k-n)})(y \star \mathbb{1}^{\otimes(k-m)})\|_{\mathcal{D}^\infty} \\
&\leq \lim_{k \to \infty} \|(x \star \mathbb{1}^{\otimes(k-n)})\|_{\mathcal{D}^\infty} \|(y \star \mathbb{1}^{\otimes(k-m)})\|_{\mathcal{D}^\infty} \\
&= \|x\|_{\mathrm{Sym}} \|y\|_{\mathrm{Sym}}
\end{aligned}
$$

with equality if $y = x^*$.

We have thus verified the C^*-seminorm properties, and the second advertised claim follows from the general theory [81, Section 6.1]. $\qquad\square$

Next, we aim to set up a bijection between the space of symmetric states $K_s(\mathcal{D}^\infty)$ and the state space $K(\mathrm{Sym}^\infty(\mathcal{D}))$ of the abelian algebra. The connection revolves around $\cup_n \mathrm{Sym}^n(\mathcal{D})$, as it can be interpreted as a subspace of either algebra. We will thus look for natural ways of extending a state ρ from $\cup_n \mathrm{Sym}^n(\mathcal{D})$ to $\mathrm{Sym}^\infty(\mathcal{D})$ and to $\mathcal{D}^\infty$ respectively.

For the former case, we can use the fact that $\cup_n \mathrm{Sym}^n(\mathcal{D})$ is dense in $\mathrm{Sym}^\infty(\mathcal{D})$. Thus, if $\rho \in K_s(\mathcal{D}^\infty)$, it is natural to try to extend it by continuity from $\cup_n \mathrm{Sym}^n(\mathcal{D})$ to a state on all of $\mathrm{Sym}^\infty(\mathcal{D})$. Lemma 2.1.4 shows that this ansatz indeed leads to a well-defined map

$$
E : K_s(\mathcal{D}^\infty) \to K(\mathrm{Sym}^\infty(\mathcal{D})). \tag{2.8}
$$

Conversely, in order to evaluate a state $\rho \in K(\mathrm{Sym}^\infty(\mathcal{D}))$ on an element of $x \in \mathcal{D}^\infty$, our approach is to map x to a symmetrized version $\mathrm{Sym}(x) \in \mathrm{Sym}^\infty(\mathcal{D})$ and then to apply ρ to $\mathrm{Sym}(x)$. To define the symmetrization operation, note that any element of $\mathcal{D}^\infty$ can be represented by a Cauchy sequence $(x_n)_n$ with $x_n \in \mathcal{D}^n$ and set

$$
\mathrm{Sym} : (x_n)_n \mapsto (\mathrm{Sym}^n(x_n))_n. \tag{2.9}
$$

Lemma 2.1.4 establishes that the result lies in $\mathrm{Sym}^\infty(\mathcal{D})$ and that the adjoint

$$
(\mathrm{Sym}^*(\rho))(x) = \rho(\mathrm{Sym}(x)) \tag{2.10}
$$

defines a map

$$
\mathrm{Sym}^* : K(\mathrm{Sym}^\infty(\mathcal{D})) \to K_s(\mathcal{D}). \tag{2.11}
$$

Lemma 2.1.4. *The maps* Sym^* *and* E *are well-defined and inverses of each other. What is more,* Sym^* *is weakly continuous.*

Proof. We will repeatedly make use of the fact [43, Prop. II.6.2.5] that the states on a C^*-algebra are exactly those functionals ρ that satisfy

$$\rho(\mathbb{1}) = 1, \qquad |\rho(x)| \leq \|x\|. \tag{2.12}$$

Eq. (2.9) indeed defines a map from $\mathcal{D}^\infty \to \mathrm{Sym}^\infty(\mathcal{D})$: If $(x_n)_n, x_n \in \mathcal{D}^n$ is a Cauchy sequence with respect to $\| \cdot \|_{\mathcal{D}^\infty}$, then by Eq. (2.7), the sequence $(\mathrm{Sym}^n(x_n))_n$ is Cauchy with respect to $\| \cdot \|_{\mathrm{Sym}}$ and therefore an element of $\mathrm{Sym}^\infty(\mathcal{D})$. Next, let $\rho \in K(\mathrm{Sym}^\infty(\mathcal{D}))$. Then

$$\rho(\mathrm{Sym}(\mathbb{1})) = \rho(\mathbb{1}) = 1, \qquad |\rho(\mathrm{Sym}(x))| \leq \|\mathrm{Sym}(x)\|_{\mathrm{Sym}} \leq \|x\|_{\mathcal{D}^\infty},$$

thus $\mathrm{Sym}^*(\rho)$ is a state. Because $\mathrm{Sym} \circ \alpha_\pi = \mathrm{Sym}$ for any permutation π, $\mathrm{Sym}^*(\rho)$ is symmetric. The map Sym^* is weakly continuous: If a net ρ_λ in $K(\mathrm{Sym}^\infty(\mathcal{D}))$ converges weakly to ρ, then in particular $\rho_\lambda(\mathrm{Sym}(x)) \to \rho(\mathrm{Sym}(x))$ for all $x \in \mathcal{D}$. Thus $\mathrm{Sym}^*(\rho_\lambda)$ converges weakly to $\mathrm{Sym}^*(\rho)$.

To prove that E is well-defined, start with a state $\rho \in K_s(\mathcal{D}^\infty)$. For $x \in \cup_n \mathrm{Sym}^n(\mathcal{D})$, using symmetry and Eq. (2.12),

$$|\rho(x)| = \lim_{k \to \infty} |\rho(x \star \mathbb{1}^{\otimes k})| \leq \lim_{k \to \infty} \|x \star \mathbb{1}^{\otimes k}\|_{\mathcal{D}^\infty} = \|x\|_{\mathrm{Sym}}.$$

In other words, on $\cup_n \mathrm{Sym}^n(\mathcal{D})$, ρ is bounded with respect to the $\| \cdot \|_{\mathrm{Sym}}$-norm and can thus be uniquely extended by continuity to a functional $E(\rho)$ on $\mathrm{Sym}^\infty(\mathcal{D})$. Using Eq. (2.12) once more, the preceding estimate also shows that $E(\rho)$ is a state.

Finally, for each $\rho \in K_s(\mathcal{D}^\infty)$, $x \in \mathcal{D}^\infty$ and $\sigma \in K(\mathrm{Sym}^\infty(\mathcal{D}))$, $y \in \cup_n \mathrm{Sym}^n(\mathcal{D})$,

$$\mathrm{Sym}^*(E(\rho))(x) = E(\rho)(\mathrm{Sym}(x)) = \rho(\mathrm{Sym}(x)) = \rho(x),$$
$$E(\mathrm{Sym}^*(\sigma))(y) = (\mathrm{Sym}^* \sigma)(y) = \sigma(y),$$

which shows that the two maps are inverses of each other. $\qquad\square$

Proof of Theorem 2.1.1. Consider $E(\rho) \in K(\mathrm{Sym}^\infty(\mathcal{D}))$. By [52, Example 4.1.30 and Proposition 2.3.27], because $\mathrm{Sym}^\infty(\mathcal{D})$ is abelian, there exists a unique non-negative measure $\tilde{\mu}$ over pure states $K_{\mathrm{pure}}(\mathrm{Sym}^\infty(\mathcal{D}))$, such that

$$E(\rho)(x) = \int_{K_{\mathrm{pure}}(\mathrm{Sym}^\infty(\mathcal{D}))} \tilde{\sigma}(x) \, \mathrm{d}\tilde{\mu}(\tilde{\sigma}).$$

But then, for $x_i \in \mathcal{D}$,

$$
\begin{aligned}
\rho(x_1 \otimes \ldots \otimes x_n) &= \rho(x_1 \star \ldots \star x_n) \\
&= (E(\rho))(x_1 \star \ldots \star x_n) \\
&= \int_{K_{\text{pure}}(\text{Sym}^\infty(\mathcal{D}))} \tilde{\sigma}(x_1 \star \ldots \star x_n) \, \mathrm{d}\tilde{\mu}(\tilde{\sigma}).
\end{aligned}
$$

Consider one $\tilde{\sigma} \in K_{\text{pure}}(\text{Sym}^\infty(\mathcal{D}))$, let $R : K_s(\mathcal{D}^\infty) \to K(\mathcal{D})$ be the map that restricts states to $\mathcal{D} \subset \mathcal{D}^\infty$, and let $\sigma = R \circ \text{Sym}^*(\tilde{\sigma})$. Because pure states of abelian algebras are homomorphisms,

$$
\tilde{\sigma}(x_1 \star \ldots \star x_n) = \tilde{\sigma}(x_1) \ldots \tilde{\sigma}(x_n) = \sigma(x_1) \ldots \sigma(x_n) = \Pi_\sigma(x_1 \otimes \cdots \otimes x_n).
$$

The restriction R is the adjoint of the symmetric embedding $\mathcal{D} \to \bigcup_n \text{Sym}^n(\mathcal{D}^n)$ $\subset \mathcal{D}^\infty$. As the adjoint of a bounded map, it is weak*-continuous by the same argument as the one used in the proof of Lemma 2.1.4. Thus the concatenation $R \circ \text{Sym}^* : \tilde{\sigma} \mapsto \sigma$ is continuous and hence measurable. We can therefore define a non-negative measure μ on $K(\mathcal{D})$ by

$$
\mu(S) = \tilde{\mu}\big((R \circ \text{Sym}^*)^{-1}(S)\big).
$$

Then

$$
\rho(x_1 \otimes \ldots \otimes x_n) = \int_{K(\mathcal{D})} \Pi_\sigma(x_1 \otimes \ldots \otimes x_n) \, \mathrm{d}\mu(\sigma),
$$

which proves the claim, as $\cup_n \mathcal{D}^n$ is dense in $\mathcal{D}^\infty$. $\square$

2.2 Scalar Extension

In sections 3.3.6 and 5.2 we will discuss a way of analyzing causal scenarios, and more generally of solving certain polynomial optimization problems, that does not involve copying the entire algebra, but instead uses a hierarchy of so-called *scalar extensions* [41, 82, 83]. In Ref. [39] the authors develop a way of proving convergence of such a scalar extension hierarchy for state polynomial optimization problems. In section 5.2 we will prove a similar convergence statement by re-expressing their proof idea in the C^*-algebra language, and combining this with some techniques for proving convergence for the polarization method that we developed in Papers [1, 2].

Here we prove a corollary that can be used for proving convergence of scalar extension hierarchies and is motivated by the proof strategy of the de Finetti theo-

rem 2.1.1, as well as by lemmas IV.4 and IV.5 of Ref. [24]. It will turn out to be simpler than Theorem 2.1.1, due to the fact that we will be dealing with abelian algebras.

Let $\mathcal{A}$ be the universal C^*-algebra generated by a set of generators $\mathcal{G}$ and relations $\mathcal{R}$. Furthermore, let $\mathcal{B}_i$ be the abelian C^*-algebra generated by a single generator b_i (and the identity), with the relation $C_i^2 \mathbb{1} - b_i^* b_i \geq 0$ for some constant $C_i \in \mathbb{R}$ that we leave unspecified for now. In the most general case we will have a countable number of such algebras $\mathcal{B}_i$.

Since each $\mathcal{B}_i$ is an abelian algebra, they are nuclear [43, II.9.4.4], i.e. for every C^*-algebra $\mathcal{C}$ the completion of the algebraic tensor product $\mathcal{B}_i \otimes_{\text{alg}} \mathcal{C}$ with respect to a C^*-norm is unique, so that we can simply write $\mathcal{B}_i \otimes \mathcal{C}$. In particular, this means that $\mathcal{A} \otimes \mathcal{B}_i = \mathcal{A} \otimes_{\min} \mathcal{B}_i$ for every $\mathcal{B}_i$ so that it seems likely that we can use the proof strategy of the existing de Finetti theorem of [24]. Indeed the proof follows directly from this paper and its references.

We start by defining a second type of inductive limit of C^*-algebras in order to work more easily with an algebra that is defined by a countable set of generators, and subalgebras thereof. Indeed the construction will follow the standard definition of inductive limits as given in e.g. Ref. [43, II.9.4.5]. Define $\mathcal{D}^K = \mathcal{A} \otimes B_1 \otimes \ldots \otimes B_K$, where each of the tensor products is unique due to the fact that nuclearity is preserved under tensor products (and also under inductive limits, as we will use in a moment) [43, II.8.2.1]. $\mathcal{D}^K$ can be embedded into $\mathcal{D}^{K+1}$ in the obvious way, so that we can define $\mathcal{D}^{K \to \infty}$ as the completion of

$$\bigcup_K \mathcal{D}^K \tag{2.13}$$

with respect to the norm $\|.\|_{\mathcal{D}^{K \to \infty}}$, which is defined for each $x \in \mathcal{D}^K$ as

$$\|x\|_{\mathcal{D}^{K \to \infty}} = \|x\|_{\mathcal{D}^K}. \tag{2.14}$$

In $\mathcal{D}^{K \to \infty}$ the elements $x \in \mathcal{D}^K$ and $x \otimes \mathbb{1} \in \mathcal{D}^{K+1}$ are identified, similar to how $y \in \mathcal{D}^n$ and $y \otimes \mathbb{1} \in \mathcal{D}^{n+1}$ are identified in $\mathcal{D}^\infty$ in the previous section. We denote by $\mathcal{B}^{K \to \infty}$ the special case where $\mathcal{A}$ is trivial.

Corollary 2.2.1. *A state $\omega \in K(\mathcal{B}^{K \to \infty})$ has a weak*-integral decomposition*

$$\omega = \int \mathrm{d}\mu(\sigma)\sigma, \tag{2.15}$$

where σ is a pure state on the abelian algebra $\mathcal{B}^{K \to \infty}$, i.e. a unital *-homomorphism and μ is a probability measure on pure states of $\mathcal{B}^{K \to \infty}$.

Proof. This follows directly from [52, Example 4.1.30]: Since $\mathcal{B}^{K\to\infty}$ is an abelian C^*-algebra, there exists for any state $\rho \in K(\mathcal{B}^{K\to\infty})$ a unique measure μ over pure states $K_{\text{pure}}(\mathcal{B}^{K\to\infty})$ such that

$$\rho(x) = \int_{K_{\text{pure}}(\mathcal{B}^{K\to\infty})} \sigma(x)\mathrm{d}\mu(\sigma),$$

where each pure state is a *-homomorphism.   □

It is possible to prove a more general corollary, namely the one below, which follows directly from the proofs of lemma IV.4 and IV.5 of Ref. [24]. However, since we will not need this result, we will only state it here and refer the interested reader to Ref. [24].

Corollary 2.2.2. *A state $\omega \in K(\mathcal{D}^{K\to\infty})$ has a weak*-integral decomposition*

$$\omega = \int \mathrm{d}\mu(\sigma)\phi_\sigma \otimes \sigma, \tag{2.16}$$

where $\phi_\sigma \in K(\mathcal{A})$, σ is a pure state on the abelian algebra $\mathcal{B}^{K\to\infty}$, i.e. a unital *-homomorphism and μ is a probability measure on pure states of $\mathcal{B}^{K\to\infty}$. Moreover, for each $a \in \mathcal{A}$, $\phi_\sigma(a)$ is uniquely determined by ϕ almost everywhere with respect to μ. If $\mathcal{A}$ is separable, the ϕ_σ is uniquely determined by σ almost everywhere.

Remark. In a way, the corollaries 2.2.1 and 2.2.2 are different from a "standard" de Finetti theorem, since they do not require a combinatorial argument. The conclusion follows almost directly from the decomposition of states on abelian algebras.

2.3   Quantum inflation

In the case of the original formulation of the inflation technique in Ref. [22], the global algebra is not a (maximal) tensor product of copies of the same algebra, and as such it is not possible to directly apply Theorem 2.1.1 to it. Nevertheless, the algebra contains many isomorphic subalgebras for which the Theorem does apply. In this section, we will derive all these separability statements from one "de Finetti-like" theorem for the global algebra.

We start by formalizing the definition for a C^*-algebra that corresponds to the limiting procedure of taking infinite inflation levels. We will phrase this in more general terms, without referring to the inflation technique, with the goal of making the statement more widely applicable.

The result of the Theorem can loosely be seen as an interpolation between "standard" quantum de Finetti theorems and the result of e.g. Ref. [84], which is a de Finetti theorem for exchangeable sequences in free products of $*$-algebras.

Let $\mathcal{E}$ be a universal C^*-algebra generated by a set of generators $\mathcal{G}$ and relations $\mathcal{R}$. Let $\mathcal{A}, \mathcal{B}, \dots$ be commuting subalgebras of $\mathcal{E}$, such that $\mathcal{E} = \mathcal{A} \otimes_{\max} \mathcal{B} \otimes_{\max} \dots$. Denote the subset of generators that generate the algebra of system X by $\mathcal{G}_X$. Let $\{i, j, \dots\}$ be a set of indices of size L and assign a subset I_X of the indices to each subalgebra X, e.g. if $I_A = \{i, j\}$, $I_B = \{j, k\}$, we write $\mathcal{A}^{ij}, \mathcal{B}^{jk}$, and so on[1]. We denote an assignment of values for the indices of system X with ℓ_X, e.g. $\ell_A = (i = 1, j = 2) = (1, 2)$ leads to $\mathcal{A}^{\ell_A} = \mathcal{A}^{i=1, j=2} = \mathcal{A}^{12}$. The generators of such a copy of the subalgebra X are denoted by $\mathcal{G}_{X^{\ell_X}}$. In a similar way, we denote by $\ell_{\mathcal{E}}$ an assignment of values between 1 and n to the indeces of all parties, with corresponding generators $\mathcal{G}^{\ell_{\mathcal{E}}}$ and relations $\mathcal{R}^{\ell_{\mathcal{E}}}$. With this definition $C^*(\mathcal{G}^{\ell_{\mathcal{E}}} | \mathcal{R}^{\ell_{\mathcal{E}}})$ is isomorphic to $\mathcal{E}$.

We now define the C^*-algebra $\mathcal{E}^n$ to be the universal C^*-algebra $C^*(\mathcal{G}^n | \mathcal{R}^n)$, where

$$\mathcal{G}^n := \bigcup_{\ell_{\mathcal{E}} \in \mathbb{Z}_n^L} \mathcal{G}^{\ell_{\mathcal{E}}}, \qquad \mathcal{R}^n := \left(\bigcup_{\ell^{\mathcal{E}} \in \mathbb{Z}_n^L} \mathcal{R}^{\ell_{\mathcal{E}}} \right) \cup \mathcal{R}_{\text{com}}, \tag{2.17}$$

$$\mathcal{R}_{\text{com}} := \left\{ [x^{\ell_X}, y^{\ell'_Y}] : x^{\ell_X} \in \mathcal{G}_{X^{\ell_X}}, y^{\ell'_Y} \in \mathcal{G}_{Y^{\ell'_Y}} \text{ s.t. } \begin{cases} X \neq Y, \text{ or} \\ (\ell_X)_i \neq (\ell'_Y)_i \, \forall i \end{cases} \right\}, \tag{2.18}$$

As noted before, this algebra contains many isomorphic copies of the subalgebras, most of which commute. However, if copies of the same subalgebra have overlapping indices, their elements do not commute. Hence, contrary to the algebra $\mathcal{D}^n$ in section 2.1, $\mathcal{E}^n$ as defined above is not a tensor product of n algebras. The commutation relations that do hold are represented by $\mathcal{R}_{\text{com}}$. These are the commutation relations between copies of different commuting subalgebras, and between copies of the same subalgebra that have no overlapping indices.

In order to arrive at a de Finetti theorem, we will first show that there is a natural way of taking an inductive limit in the parameter n. An inductive limit can be constructed from a set of norm-decreasing $*$-homomorphisms $\phi_{mn} : \mathcal{E}^m \to \mathcal{E}^n$ such that $\phi_{mn} = \phi_{kn} \circ \phi_{mk}$. We will go one step further and show that the elements from $\mathcal{F}(\mathcal{G}^{n-1}) \subset \mathcal{E}^{n-1}$ can be identified with the elements $\mathcal{F}(\mathcal{G}^{n-1}) \subset \mathcal{E}^n$. The construction will follow from Lemma 2.3.1 below.

Lemma 2.3.1. $\mathcal{E}^{n-1}$ is a C^*-subalgebra of $\mathcal{E}^n$. That is, a sequence $(x_k)_k$ with

---

[1] In the inflation technique these indices correspond to copies of the sources.

$x_k \in \mathcal{F}(\mathcal{G}^{n-1})$ *converges to* 0 *in the norm* $\|\cdot\|_{\mathcal{E}^{n-1}}$ *if and only if it converges to* 0 *in the norm* $\|\cdot\|_{\mathcal{E}^n}$.

Proof. Recall the definition for the norm of a universal C^*-algebra

$$\|x\| := \sup\{\|\pi(x)\| \,|\, \pi \text{ is a representation of } (\mathcal{G}|\mathcal{R})\}, \tag{2.19}$$

so that this is the largest algebra for which all $q \in \mathcal{R}$ are positive (cf. lemmma 1.1.5).

$\lim_k \|x_k\|_{\mathcal{E}^{n-1}} = 0 \implies \lim_k \|x_k\|_{\mathcal{E}^n} = 0$: This is the easy direction. We have that $\mathcal{R}^{n-1} \subset \mathcal{R}^n$. Hence, from Eq. (2.19) it follows that $\|x\|_{\mathcal{E}^n} \leq \|x\|_{\mathcal{E}^{n-1}}$.

$\lim_k \|x_k\|_{\mathcal{E}^n} = 0 \implies \lim_k \|x_k\|_{\mathcal{E}^{n-1}} = 0$: For the converse direction let $\mathcal{K} = \mathcal{G}^n \setminus \mathcal{G}^{n-1}$ and $\mathcal{Q} = \mathcal{R}^n \setminus \mathcal{R}^{n-1}$. The set $\mathcal{Q}$ consists of

1. the relations $\mathcal{R}^{\ell_{\mathcal{E}}}$ that involve generators that have at least one index with value n,

2. commutation relations of generators with at least one index with value n.

More precisely, the only commutation relations that are "missing" to make the algebra into a tensor product of algebras are commutation relations between copies of the same system with overlapping indices, where at least one index has value n. Indeed, if we define

$$\tilde{\mathcal{R}}^n = \mathcal{R}^n \cup \{[x^{\ell_X}, y^{\ell'_X}] : x \in \mathcal{G}_{X^{\ell_X}}, y \in \mathcal{G}_{X^{\ell'_X}} \text{ s.t. } n \in \ell_X \text{ and } n \notin \ell'_X\}$$
$$\tilde{\mathcal{E}}^n = C^*(\mathcal{G}^n | \tilde{R}^n),$$

then it holds that $\tilde{\mathcal{E}}^n = \mathcal{E}^{n-1} \otimes_{\max} C^*(K | \cup_{\ell_{\mathcal{E}}:n \in \ell_{\mathcal{E}}} \mathcal{R}^{\ell_{\mathcal{E}}})$. Hence, we can embed x into $\tilde{\mathcal{E}}^n$ in the obvious way by tensoring with identity. From this it follows that

$$\|x\|_{\mathcal{E}^{n-1}} = \|x\|_{\tilde{\mathcal{E}}^n} \leq \|x\|_{\mathcal{E}^n},$$

where the last inequality follows again from the fact that $\mathcal{R}^n \subset \tilde{\mathcal{R}}^n$. □

Lemma 2.3.1 shows that it makes sense to define an inductive limit via embeddings, similar to the construction of $\mathcal{D}^n$ in section 2.1. That is, we define addition and multiplication in

$$\bigcup_n \mathcal{E}^n$$

by embedding into the larger algebra. We then define $\mathcal{E}^\infty$ as the completion of the
$*$-algebra $\bigcup_n \mathcal{E}^n$ with respect to the norm $\| \cdot \|_{\mathcal{E}^\infty}$ defined for every $x \in \mathcal{E}^n$ as

$$\|x\|_{\mathcal{E}^\infty} := \|x\|_{\mathcal{E}^n}.$$

To get to a de Finetti theorem, let us start by defining the symmetrization map on $\mathcal{E}^n$. Let $\pi_i \in S_n$ be a permutation for the index i, and similar for $j, k, \ldots$. Then we define for every $x \in \mathcal{E}^n$

$$\text{Sym}^n(x) = \left(\frac{1}{n!} \right)^L \sum_{\pi_i, \pi_j, \ldots \in S_n} \alpha_{\pi_i, \pi_j, \ldots}(x), \tag{2.20}$$

where L is the total number of indices. Sym^n thus symmetrizes over each of the indices individually. Again, the maps Sym^n can be extended to $\mathcal{E}^\infty$ via their action on cauchy sequences $(x_n)_n$ with $x_n \in \mathcal{E}^n$ as

$$\text{Sym} : (x_n)_n \mapsto (\text{Sym}^n(x_n))_n.$$

We call a state $\rho \in K(\mathcal{E}^\infty)$ symmetric if $\rho(x) = \rho(\text{Sym}(x))$. The set of symmetric states is denoted by $K_s(\mathcal{E}^\infty)$.

Theorem 2.3.2. *Let ρ be a symmetric state on $\mathcal{E}^\infty$. Then there exists a decomposition*

$$\rho = \int \mathrm{d}\mu(\phi)\phi, \tag{2.21}$$

where μ is a unique probability measure over symmetric states on $\mathcal{E}^\infty$. Additionally, if X and Y are subalgebras of $\mathcal{E}^\infty$ that commute and whose assignments of indeces ℓ_X and ℓ'_Y do not have an index with the same value (i.e. $(\ell_{I_X \cap I_Y})_i \neq (\ell'_{I_X \cap I_Y})_i$ for all i), then for all $x \in X, y \in Y$

$$\phi(xy) = \phi(x)\phi(y) \qquad \mu - almost\ everywhere. \tag{2.22}$$

The proof of this theorem follows the same structure as that of Theorem 2.1.1, with some details changed.

We start by defining a product $\star_{n,m}$ that acts on elements $x \in \mathcal{E}^n$ and $y \in \mathcal{E}^m$ as

$$x \star_{n,m} y = \text{Sym}^{n+m}(x \otimes y),$$

where $x \otimes y$ is understood as its embedding in $\mathcal{E}^{n+m}$. We write the similar product acting on elements in $\mathcal{E}^\infty$ simply as $\star$.

With this product we get once more a symmetric local algebra as the set

$$\bigcup_n \text{Sym}^n(\mathcal{E}^n).$$

Defining embeddings

$$\text{Sym}^n(\mathcal{E}^n) \to \text{Sym}^{n+k}(\mathcal{E}^{n+k}), \qquad x \mapsto x \star_{n,k} \mathbb{1}_k,$$

with $\mathbb{1}_k$ the identity on $\mathcal{E}^k$, and identifying x with $x \star_{n+k} \mathbb{1}_k$ for all k, makes $(\bigcup_n \text{Sym}^n(\mathcal{E}^n), \star)$ into an abelian $*$-algebra. The following lemmas, that are similar to lemmas 2.1.3 and 2.1.2, show that completion with respect to the norm defined as

$$\|x\|_{\text{Sym}} = \lim_{k\to\infty} \|x \star \mathbb{1}_k\|_{\mathcal{E}^\infty}$$

turns $(\bigcup_n \text{Sym}^n(\mathcal{E}^n), \star)$ into an abelian C^*-algebra.

Lemma 2.3.3. *For $x \in \text{Sym}^m(\mathcal{E}^m)$ and $y \in \text{Sym}^n(\mathcal{E}^n)$ it holds that*

$$\lim_{k\to\infty} \|(x \star \mathbb{1}_{k-m})(y \star \mathbb{1}_{k-n}) - (x \star y \star \mathbb{1}_{k-n-m})\|_{\mathcal{E}^\infty} = 0. \tag{2.23}$$

Proof. The probability that a permutation $\pi_i([m]) \in S_k$ and $\pi_i'([n]) \in S_k$ intersect is upper bounded by $\frac{mn}{k}$ (see lemma 2.1.3). Then, by the union bound the probability that any of the indices of the L permutations overlap is upper bounded by $\frac{Lmn}{k}$. We use this to write

$$\lim_{k\to\infty} \|(x \star \mathbb{1}_{k-m})(y \star \mathbb{1}_{k-n}) - (x \star y \star \mathbb{1}_{k-n-m})\|_{\mathcal{E}^\infty}$$

$$= \lim_{k\to\infty} \left\| \left(\frac{1}{k!}\right)^{2L} \sum_{\substack{\pi_i, \pi_i' \in S_k \\ \pi_i([m]) \cap \pi_i'([n]) \neq \emptyset}} \sum_{\substack{\pi_j, \pi_j' \in S_k \\ \pi_j([m]) \cap \pi_j'([n]) \neq \emptyset}} \cdots \right.$$

$$\left. \alpha_{\pi_i \pi_j \ldots}(x \otimes \mathbb{1}_{k-m}) \alpha_{\pi_i' \pi_j' \ldots}(y \otimes \mathbb{1}_{k-n}) \right\|_{\mathcal{E}^\infty}$$

$$\leq \lim_{k\to\infty} \frac{mnL}{k} \|x\|_{\mathcal{E}^\infty} \|y\|_{\mathcal{E}^\infty} = 0,$$

since x and y are bounded.   □

Using this combinatorial lemma, we can prove that $\|\cdot\|_{\text{Sym}}$ is a C^*-norm (cf. lemma 2.1.2).

Lemma 2.3.4. *The limit*

$$\|x\|_{\mathrm{Sym}} := \lim_{k \to \infty} \|x \star \mathbb{1}_k\|_{\mathcal{E}^\infty}$$

defines a C^*-seminorm on the symmetric local algebra $\bigcup_n \mathrm{Sym}^n(\mathcal{E}^n)$ fulfilling

$$\|x\|_{\mathrm{Sym}} \leq \|x\|_{\mathcal{E}^\infty}.$$

The completion $\mathrm{Sym}(\mathcal{E}^\infty)$ is an abelian C^*-algebra.

Proof. See lemma 2.1.2.   □

Once again, the aim is to set up a bijection between the state space of the abelian algebra $\mathrm{Sym}(\mathcal{E}^\infty)$ and the symmetric states of the quasi-local algebra $\mathcal{E}^\infty$. The construction is the same as in section 2.1. Define a map

$$E : K_s(\mathcal{E}^\infty) \to K(\mathrm{Sym}(\mathcal{E}^\infty)), \tag{2.24}$$

which extends a state on $\bigcup_n \mathcal{E}^n$ to $\mathrm{Sym}(\mathcal{E}^\infty)$ by continuity. Secondly, define the map

$$\mathrm{Sym}^* : K(\mathrm{Sym}(\mathcal{E}^\infty)) \to K_s(\mathcal{E})$$

as

$$(\mathrm{Sym}^*(\rho))(x) = \rho(\mathrm{Sym}(x)). \tag{2.25}$$

Lemma 2.3.5. *The maps Sym^* and E are well-defined and inverses of each other. Additionally, Sym^* is weakly continuous.*

Proof. See lemma 2.1.4.   □

We are now ready to prove Theorem 2.3.2. Since most of the ingredients of the proof are identical to those of theorem 2.1.1, the proof strategy will be exactly the same as well. The most notable difference is that we will not get a measure over states on the algebra $\mathcal{E}$, but only over symmetric states on the inflated algebra $\mathcal{E}^\infty$. We will see that all separability conditions that one would expect to hold from the algebra derived from the inflated causal structure, indeed follow from this theorem.

Proof of Theorem 2.3.2. Since $\mathrm{Sym}(\mathcal{E}^\infty)$ is an abelian algebra, any state $E(\rho) \in \mathcal{K}(\mathrm{Sym}(\mathcal{E}^\infty))$ can be written as

$$E(\rho) = \int_{K_{\mathrm{pure}}(\mathrm{Sym}(\mathcal{E}^\infty))} \mathrm{d}\tilde{\mu}(\tilde{\sigma}) \, \tilde{\sigma},$$

where $\tilde{\mu}$ is a probability measure over pure states on $\mathrm{Sym}(\mathcal{E}^\infty)$.

Then, using the fact that the map Sym^* is weakly continuous and the inverse of E, we can write

$$\rho = \int_{K_s(\mathcal{E}^\infty)} \mathrm{d}\mu(\sigma)\, \sigma,$$

where

$$\mu(S) = \tilde{\mu}((\mathrm{Sym}^*)^{-1}(S)).$$

Now, for any choice of $x, y \in \bigcup_n \mathcal{E}^n$ obeying the conditions stated in the theorem, we get

$$\rho(xy) = \rho(x \star y) \tag{2.26}$$
$$= (E(\rho))(x \star y) \tag{2.27}$$
$$= \int \mathrm{d}\tilde{\mu}(\tilde{\sigma})\, \tilde{\sigma}(x \star y) \tag{2.28}$$
$$= \int \mathrm{d}\tilde{\mu}(\tilde{\sigma})\, \tilde{\sigma}(x)\tilde{\sigma}(y) \tag{2.29}$$
$$= \int \mathrm{d}\mu(\sigma)\, \sigma(x)\sigma(y), \tag{2.30}$$

where it was used that $\tilde{\sigma}$ is a pure state on an abelian algebra, i.e. a $*$-homomorphism. This proves the claim, as $\bigcup_n \mathcal{E}^n$ is dense in $\mathcal{E}^\infty$. $\square$

If x and y would not obey the conditions of the theorem, the equality (2.26) would not hold, because x and y would share at least one index with the same value. In the context of quantum inflation, this would correspond to two operators acting on the same copy of a quantum state, for which it is indeed not expected that the state factorizes over these operators.

The statement of the theorem can roughly be interpreted as follows. If the subalgebra generated by two subalgebras $\mathcal{A}$ and $\mathcal{B}$ of an inflated causal structure is a tensor product $\mathcal{A} \otimes \mathcal{B}$, the restriction of the state to these subalgebras is separable. In fact, since a state $\phi \in K_s(\mathcal{E}^\infty)$ is equivalent to some pure state on an abelian algebra for all x, y with unequal or non-overlapping indices, it holds that ϕ μ-almost surely *simultaneously* factorizes.

By restricting the states $\sigma \in K_s(\mathcal{E}^\infty)$ to n levels of inflation, it might be possible to draw non-trivial conclusions on lower levels of the inflation from the factorization conditions on higher levels. It is, however, not immediately clear how to do this.

Chapter 3

Quantum causal compatibility

One of the main motivations of this thesis is the quantum causal compatibility problem, which was informally stated in the introduction as Problem 1. Here we will state it more precisely, in such a way that we can make mathematically precise statements that take, for example, machine precision into account. Afterwards, we rephrase this problem as a hierarchy of optimization problems in the form of NPO (cf. Sec. 1.3.2) inspired by the quantum inflation technique of Ref. [22].

The main contribution of this chapter is to rephrase the quantum causal compatibility problem as a hierarchy of NPO problems that is also *complete* in the sense that any incompatible probability distribution will ultimately be detected as such. In Sections 3.1 and 3.2 we will focus on a modified version of the inflation hierarchy and prove it to be complete. In Section 3.3 we treat the special case of the bilocal scenario. For this causal structure it turns out that such a modification to the inflation technique is not necessary. In fact, we will show that there exist *several* convergent hierarchies for this special scenario. The following is largely taken from Paper [1, Sec. 1.3].

We follow the notation of Ref. [22] and introduce a precision parameter ϵ to define the approximate quantum causal compatibility problem below.

Problem 3 (Approximate quantum causal compatibility). *Given $\epsilon \geq 0$, a causal structure and a probability distribution over observable variables P, determine whether there exists a distribution $\tilde{P}$ that can be produced by a quantum model compatible with the causal structure, such that $\|\tilde{P} - P\|_2^2 \leq \epsilon$.*

Similar to Refs. [22, 23], we will phrase Problem 3 as a special case of a more general *causal polynomial optimization problem*. For this problem, the goal is to minimize a polynomial function $f_0(\rho)$ over states ρ that are compatible with the causal structure and in addition satisfy a set of polynomial constraints $f_i(\rho) = 0$.

The notion of a "polynomial function in states" needs to be explained. A quantum model associates with each party an algebra of observables. We assume that these algebras possess a finite set of *generators* $\mathcal{G} = \{g_i\}_i$ (we will describe how to derive a suitable set of generators for each causal structure and any given number of outcomes per party later). For now, we treat the g_i as (non-commutative) variables – later we will optimize over all possible assignments of concrete operators to the g_i. Let $\mathcal{F}^{(k)}(\mathcal{G})$ be the vector space formed by complex linear combinations of words of length k in the symbols g_i and g_i^*. Each element $x \in \mathcal{F}^{(k)}$ defines a "linear function f on states" in the following sense: Choose an assignment of operators to the g_i. Let $\mathcal{D}$ be the resulting algebra of observables. Then x can be understood as an element of $\mathcal{D}$. Let ρ be a state on $\mathcal{D}$. Then the linear function is just $f : \rho \mapsto \rho(x)$. To define a degree-g polynomial, start with an element x in the g-fold tensor product $\mathcal{F}^{(k)} \otimes \cdots \otimes \mathcal{F}^{(k)}$ and set

$$f : \rho \mapsto \rho^{\otimes g}(x). \tag{3.1}$$

We say that x is a *polarization* of the polynomial f. Note that polarizations x are defined independently of any assignments of concrete operators to the generators – so they can be used to specify polynomial objective functions for problems that optimize over such assignments.[1]

Problem 4 (Quantum causal polynomial optimization). *Given a causal structure, the number of possible measurement outcomes for each party, a polynomial function f_0 on quantum states (as defined above), and a countable set of polynomial functions $f_1, f_2, \ldots$ that are non-negative on quantum states compatible with the causal structure. Find*

$$f^\star = \min_\rho f_0(\rho)$$

$$s.\ t. \quad f_i(\rho) = 0 \quad i \geq 1$$

$$\rho \text{ is compatible with the causal structure.}$$

Problem 3 reduces to Problem 4 by choosing f_0 to be the 2-norm distance between the observed data P and the one produced by the state. More general objective

---

[1] In the context of quantum inflations, there are two distinct sources of "polynomials" that must not be confused. First, general operators arise as (non-commutative) polynomials in the generators. The word "polynomial" in the *non-commutative polynomial optimization* (NPO) framework refers to this sense. But NPO objective functions are still linear in the state. In contrast, the term "polynomial" in the *quantum causal polynomial optimization problem* – both as treated here and in Refs. [22, 23] – indicates that we are allowing for objective functions that are polynomials in the state (in the sense explained above). The degree of polynomial expressions in the generators of the algebra is connected to the level k of the NPO hierarchy, while the degree of polynomial functions of the states corresponds to the level n of the inflation or polarization hierarchy.

functions can be important in applications – see [22, Section VII] for examples.

As we have seen in Sec. 1.2.2, the rigorous definition of a quantum causal structure depends on a mathematical model of the notion of a "local subsystem". Here we model local subsystems via commuting observable algebras of bounded operators. One could also consider alternative models: Instead of via commuting algebras, locality could be formalized in terms of tensor products of Hilbert spaces [17, 53, 54, 55]. Additionally, even though the causal compatibility problem only involves bounded operators (the algebra generated by the POVM elements that give rise to the observed probabilities), one could allow for unbounded operators in the local algebras on which they act. A detailed discussion of these modeling decisions was given in Sec. 1.2.2.

The SDP hierarchy that we describe in Sections 3.1 and 3.2 differs slightly from the original quantum inflation hierarchy of [22]. Most importantly, we add two new parameters: r, C, which are related to the *Schmidt decomposition* of the measurement operators. To define them, consider a node of a quantum causal structure, say the one that gives rise to the random variable A in the triangle scenario (Fig. 1.1(b)). Each possible outcome is associated with a POVM element E. As there are two incoming arrows to this vertex, E acts on two quantum systems. Call the observable algebras acting on the respective subsystems $\mathcal{A}_-, \mathcal{A}_+$. For fixed values of r, C, we assume that E is of the form

$$E = \sum_{\alpha=1}^{r} e_-(\alpha) e_+(\alpha)$$

for suitable operators $e_-(\alpha) \in \mathcal{A}_-, e_+(\alpha) \in \mathcal{A}_+$ such that

$$\|e_-(\alpha)\|, \|e_+(\alpha)\| \leq C.$$

Call models like this *rank-constrained with parameters r, C*.

Problem 5 (Rank-constrained quantum causal polynomial optimization). *With the notation of Problem 4, find*

$$f_{r,C}^{\star} = \min_{\rho} f_0(\rho)$$

s. t. $\quad f_i(\rho) = 0 \quad i \geq 1$

$\qquad \rho$ is a state on a model that is rank-constrained with parameters r, C

$\qquad \rho$ is compatible with the causal structure.

Here, we construct semidefinite programming relaxations for these problems:

Theorem 3.0.1. *Use the notation of Problem 4 and Problem 5. For every r, C, there*

exists a hierarchy of semidefinite programs indexed by an inflation parameter n *and an* NPO parameter k. *Denote the optimal values by* $f^\star_{r,C,n,k}$.

The hierarchy is complete in the sense that, as $n, k \to \infty$, the $f^\star_{r,C,n,k}$ converge to $f^\star_{r,C}$ from below. What is more, as $r, C \to \infty$, the $f^\star_{r,C}$ converge to $f^\star$ from above.

Increasing the parameter C does not come with significant computational cost (cf. the discussion of the *Archimedean property* in [40], and the discussion on convergence speed of SDPs in Sec. 1.3). Larger values of r, in contrast, do correspond to a larger number of variables and constraints in the SDP formulation. The decision to add these additional degrees of freedom must therefore be well-justified. While we cannot prove that they are strictly necessary (which would in particular imply that the original quantum inflation hierarchy is not convergent), we identify some challenges that any constructive convergence proof that does not include these extra variables would face in Secs. 3.1.1 and 3.1.2.

3.1   Inflation for the quantum causal compatibility problem

Here we give a brief overview of the quantum inflation technique, adapted from Sec. 2.3.1 of Paper [1].

The quantum inflation technique was first introduced in Ref. [22] and generalizes classical inflation [21, 23] to the case of quantum causal structures. This section serves to motivate the construction of the hierarchy in Sec. 3.2 – but in the rest of the thesis, we will not rely on results and notation introduced here. We again focus on the triangle scenario (Fig. 1.1(b)).

Given a joint distribution $P(A, B, C)$, assume that there is a quantum model compatible with the triangle scenario. Using the terminology of Sec. 1.2.2, we thus know there exist a global observable algebra $\mathcal{D}$ generated by local algebras $\mathcal{A}_-$, $\mathcal{A}_+$, $\mathcal{B}_-$, $\mathcal{B}_+$, $\mathcal{C}_-$, $\mathcal{C}_+$, a state ρ factorizing as in (1.18), and POVM elements $\{E_a\}_a$, $\{F_b\}_b$, $\{G_c\}_c$ that reproduce the correlations P as in (1.20).

Denoting the restrictions of ρ to the subalgebra $\langle \mathcal{C}_+ \cdot \mathcal{A}_- \rangle$ generated by $\mathcal{C}_+, \mathcal{A}_-$ as ρ_{CA}, to $\langle \mathcal{A}_+ \cdot \mathcal{B}_- \rangle$ as ρ_{AB}, and to $\langle \mathcal{B}_+ \cdot \mathcal{C}_- \rangle$ as ρ_{BC}, Eq. (1.18) is equivalent to demanding that ρ factorizes as

$$\rho = \rho_{CA} \otimes \rho_{AB} \otimes \rho_{BC}.$$

For any *level n*, we construct an *inflated* model as follows. Distribute n independent copies of the original states ρ_{AB}, ρ_{BC} and ρ_{CA} among the three nodes A, B, and C. At each node, we consider n^2 POVMs $\{E^{ij}_a\}_a$, $\{F^{kl}_b\}_b$, $\{G^{pq}_c\}_c$. The POVM element E^{ij}_a replicates the original E_a, but acts on the i-th copy of the state ρ_{CA} and
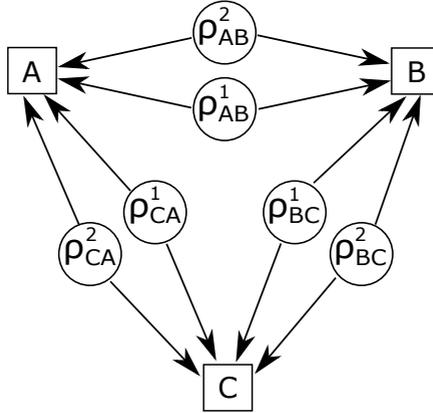
Figure 3.1: The level 2 inflation of the quantum triangle scenario. The latent quantum systems have been copied individually. Alice, Bob and Charlie now each have 4 choices of which combination of the incoming states to measure. This is reflected in the labeling of the corresponding POVMs and algebras. For example, Alice's POVM elements are denoted by $\{E_a^{ij}\}$, where i denotes the copy of the state ρ_{CA} and j denotes the copy of the state ρ_{AB} that the operator is acting on.

j-th copy of the state ρ_{AB}. The other two cases are defined analogously. As a result, POVM elements E_a^{ij}, F_b^{jl}, and G_c^{li} reproduce the original probabilities $P(a,b,c)$. Fig. 3.1 depicts the level 2 inflation of the triangle scenario.

We now list a number of properties of the inflated model. These properties can be directly imposed as constraints in an NPO program. It follows that if P is compatible with the causal model, then the resulting NPO problem is feasible for any inflation level n [22]. In Sec. 3.2 we construct a variant of this NPO hierarchy for which we supply a proof of the converse implication.

First, in Ref. [22] it is assumed that the $\{E_a\}_a$, $\{F_b\}_b$ and $\{G_c\}_c$ are orthogonal projective measurements, rather than more general POVMs. This simplifies the SDP and can be done without loss of generality, because we do not restrict dimension and possible dilations would still be compatible with the causal structure.

$$(E_a^{ij})^* = E_a^{ij} \qquad\qquad \forall i,j,a, \qquad (3.2)$$

$$E_a^{ij} E_{a'}^{ij} = \delta_{a,a'} E_a^{ij} \qquad\qquad \forall i,j,a,a', \qquad (3.3)$$

$$\sum_a E_a^{ij} = \mathbb{1} \qquad\qquad \forall i,j, \qquad (3.4)$$

and similar for $\{F_b^{kl}\}$ and $\{G_c^{pq}\}$. In later sections we will drop this restriction, and will only assume that we have POVM elements, i.e. non-negative operators that sum to the identity.

Second, operators that act on different subsystems commute:

$$\left[E_a^{ij}, F_b^{kl}\right] = \left[E_a^{ij}, G_c^{kl}\right] = \left[F_b^{ij}, G_c^{kl}\right] = 0 \qquad\qquad \forall i, j, k, l, a, b, c, \tag{3.5}$$

$$\left[E_a^{ij}, E_{a'}^{i'j'}\right] = \left[F_b^{ij}, F_{b'}^{i'j'}\right] = \left[G_c^{ij}, G_{c'}^{i'j'}\right] = 0 \quad \forall i \neq i', j \neq j', a, a', b, b', c, c'. \tag{3.6}$$

Third, there is a permutation symmetry, resulting from the fact that the global state is built out of independent copies of the original one. For any polynomial Q in the measurement operators $\{E_a^{ij}\}, \{F_b^{kl}\}, \{G_c^{pq}\}$ up to inflation level n and for all permutations π, π', π'' of n elements, the following must hold:

$$\rho\Big(Q(\{E_a^{ij}, F_b^{kl}, G_c^{pq}\})\Big) = \rho\Big(Q(\{E_a^{\pi(i)\pi'(j)}, F_b^{\pi'(k)\pi''(l)}, G_c^{\pi''(p)\pi(q)}\})\Big). \tag{3.7}$$

For example,

$$\rho(E_a^{11} F_b^{12} F_{b'}^{21} G_c^{21}) = \rho(E_a^{12} F_b^{22} F_{b'}^{11} G_c^{21})$$
$$= \rho(E_a^{12} F_b^{21} F_{b'}^{12} G_c^{11})$$
$$= \rho(E_a^{22} F_b^{21} F_{b'}^{12} G_c^{12}),$$

where we have swapped $\rho_{AB}^1 \leftrightarrow \rho_{AB}^2$ in the first step, $\rho_{BC}^1 \leftrightarrow \rho_{BC}^2$ in the second and $\rho_{CA}^1 \leftrightarrow \rho_{CA}^2$ in the third.

Fourth and finally, for the specific problem of causal compatibility the authors of Ref. [22] include constraints of the marginal distribution over $g \leq n$ copies of the triangle scenario. In particular, for the triangle scenario it must hold that

$$\rho\Big(\prod_{i=1}^{g} E_{a^i}^{ii} F_{b^i}^{ii} G_{c^i}^{ii}\Big) = \prod_{i=1}^{g} P(a^i, b^i, c^i), \tag{3.8}$$

since these variables describe g independent copies of the triangle causal structure. We will not be needing these types of constraints for our quantum inflation hierarchy, since we can already show convergence without them. Instead, for the approximate causal compatibility problem we will choose an objective function that, if the optimal value is ϵ-close to 0, ensures that Eq. (3.8) approximately holds. If $\epsilon = 0$ Eq. (3.8) will hold exactly.

As noted before, the classical version of inflation has been shown to be complete in Ref. [23]. Let us briefly comment on why the approach of this paper does not directly carry over to the quantum case. The theorem the authors of Ref. [23] derive to prove this statement, relies heavily on the observations that any distribution can be decomposed as a convex combination of deterministic distributions, and that any deterministic distribution can always be produced in any network. It is then shown that symmetric deterministic distributions are close to product distributions. Those product distributions allow one to construct polynomials over which it is possible to optimize. In particular, one can minimize the 2-norm distance between these deterministic product distributions and the observed statistics.

In the quantum case, this argument fails not because the symmetric extremal states are not product states (in fact, they are if we choose appropriate subalgebras, as we will see in later sections), but because the extremal states are not clearly producible in any network. Whereas the set of extremal points in the classical case consisted of deterministic distributions, in the quantum case this set is replaced by the set of pure states. By now we have many examples of pure states that cannot be produced in the triangle network, such as the GHZ state and the W state [22]. One would thus additionally have to show that the constraints of the inflation technique are sufficient to restrict the optimization to a smaller set of states that *is* compatible with the causal structure. This requirement relates strongly to challenge 3 of the following section.

3.1.1 Challenges

In our view, to prove that the quantum inflation hierarchy, or any similar hierarchy like the polarization or scalar extension hierarchies of Chapter 5, is complete for the quantum causal optimization problem one needs to find solutions to the three challenges presented below. These challenges were first outlined in Sections 2.1, 2.4 and 2.5 of Paper [1].

Challenge 1: Mathematical models of subsystems. The first decision that has to be made relates to which framework for quantum theory one will use. As explained in Section 1.2, this choice between elementary quantum mechanics, where joint systems are constructed through Hilbert space tensor products, and algebraic quantum mechanics, where subsystems are described by commuting subalgebras, turns out to be relevant for infinite dimensional systems [17]. While most physicists prefer to work with the Hilbert space tensor product model, there seems to be no clear physical reason to prefer it over the commuting operator model.

In the quantum causal compatibility problem, no assumption is made about the dimension of the sources in the causal structure. In particular, the states could be

infinite dimensional systems such that the distinction between the commuting observables model and the Hilbert space tensor product model could actually become relevant.

In what follows, we choose to adopt the algebraic model. This is a pragmatic approach, for two reasons. First, it is already known that the quantum causal compatibility problem in the Hilbert space tensor product model is undecidable [85]. And secondly, commutativity is a constraint that can be imposed algebraically, in contrast to the requirement that the underlying Hilbert space is of tensor product form.

To reason numerically about observable algebras, one needs to express them in a format that can be processed by a computer. Both the original quantum inflation hierarchy [22] and our work [1, 2] are built on the *non-commutative polynomial optimization* (NPO) hierarchy [40], which was explained in Sec. 1.3.2. The NPO hierarchy allows us to optimize an expectation value over the set of states on a C^*-algebra, under constraints that are linear in the state, via a convergent hierarchy of SDP relaxations. This, however, only solves a part of our problem, since we need to be able to express independence constraints of the form of Eq. (1.18) that are polynomial in the state. This brings us to the second challenge.

Challenge 2: Polynomials in the state and a quantum de Finetti Theorem for general C^* tensor products.

In general, the independence constraints that follow from the causal structure are not convex and therefore cannot directly be phrased as SDP constraints. Optimizing over the set of states that obey these constraints is therefore highly non-trivial.

The basic idea underlying the quantum inflation hierarchy is to relax *independence conditions* to *symmetry conditions*, which are linear in elements of the algebra and easily incorporated into an SDP. It is easily seen that independence implies symmetries in the inflated causal structure. Central to convergence arguments is that sometimes, in an asymptotic sense, the converse is also true. This is the case for classical causal structures [23]. Such converse results that obtain independence from symmetries are known as *de Finetti Theorems* and have been formulated both for classical [78, 79] and for quantum [24, 74, 75, 76] probability theories, as well as for more general settings [80, 84].

In addition to the conceptual problems mentioned in Challenge 1 and in Sec. 1.2, switching to a more general notion of locality raises additional technical challenges. Because the generalized Quantum de Finetti Theorem that is central to our convergence proof requires C^*-algebraic methods, we have rephrased the NPO framework of Ref. [40] in this language. We employ the terminology of *universal C^*-algebras* as described in Sec. 1.1.1 (see also [43, Section II.8.3]).

A problem we faced was that the literature on de Finetti Theorems for C^*-algebras to our best knowledge only pertained to minimal tensor products – too

narrow for our use case. One of the main technical contributions of this work is
the realization that the construction and proof of the Quantum de Finetti Theorem in
[24] carries over from the minimal tensor product case for which it was formulated,
to general C^*-tensor products. The proof of this statement was given in Section 2.1.

Challenge 3: Identifying the local observable algebras. We now present what
we consider to be the most difficult challenge in deciding completeness of quantum
inflation hierarchies.

Assume that a joined probability distribution $P(A, B, C)$ passes all levels of the
original quantum inflation hierarchy, as outlined at the start of this section. We thus
know that there is a C^*-algebra $\mathcal{D}$ generated by the observables $\{E_a^{ij}, F_b^{ij}, G_c^{ij}\}$
and a state ρ that reproduces the observed correlations (Eq. (3.8)) and is symmetric
(Eq. (3.7)).

We now need to verify that this quantum model fulfills the causal constraints, e.g.
Eq. (1.18) for the triangle (see also Fig. 3.2). This involves, in particular, showing
that

1. one can embed the algebra $\langle\{E_a^{ii}\}_a\rangle$ containing the measured POVM elements
   into a potentially larger algebra $\mathcal{A}^{ii}$ of all observables associated with the ver-
   tex A, such that $\mathcal{A}^{ii}$ is generated by two commuting subalgebras $\mathcal{A}_-^i, \mathcal{A}_+^i$,
   and

2. that the state ρ can be extended to all of $\mathcal{A}^{ii}$ and that it factorizes in the sense
   that for each $A_- \in \mathcal{A}_-^i, A_+ \in \mathcal{A}_+^i$ we have $\rho(A_-A_+) = \rho(A_-)\rho(A_+)$.

The second condition can be addressed using the generalized Quantum de Finetti
Theorem 2.1.1 proven in Sec. 2.1. The first condition, however, seems much more
challenging: *There is no obvious ansatz for constructing $\mathcal{A}^{ii}$ and its commuting gen-*
erators $\mathcal{A}_-^i, \mathcal{A}_+^i$ from the algebra $\mathcal{D}$ that results from the original quantum inflation
hierarchy of Ref. [22]. In fact, in the subsection just below, we will give an argument
that suggests that $\mathcal{D}$ does not in general contain local observable algebras $\mathcal{A}_-^i, \mathcal{A}_+^i$
that satisfy the two conditions above. It would then follow that if the original quan-
tum inflation hierarchy is complete, any constructive proof of that fact would neces-
sarily have to introduce additional operators that are not generated by the measured
POVMs and their copies.[2]

The modified quantum inflation hierarchy we construct in this chapter follows
such an approach (for details, see Sec. 3.2). It contains generators $e_-^i(a, \alpha), e_+^j(a, \alpha)$,
which are constrained to commute unless both the upper and lower indices coincide.

---

[2] The construction in chapter 3.3 for the specific case of the bilocal scenario indeed introduces such
operators by taking commutants of representations of subalgebras on certain Hilbert spaces.
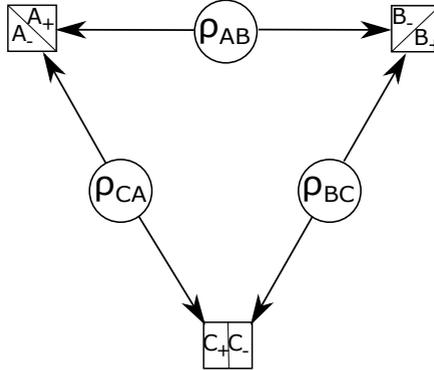
Figure 3.2: We associate to each quantum system an observable algebra. This means that the POVM elements of Alice generate a subalgebra of a larger algebra $\mathcal{A}_- \otimes \mathcal{A}_+$, and similar for Bob and Charlie. It is with respect to the splitting $\mathcal{C}_+\mathcal{A}_- \mid \mathcal{A}_+\mathcal{B}_- \mid \mathcal{B}_+\mathcal{C}_-$ that the global state is supposed to factorize.

One can then *define*

$$\mathcal{A}^i_- = \langle \{e^i_-(a,\alpha)\}_{a,\alpha}\rangle, \tag{3.9}$$

$$\mathcal{A}^i_+ = \langle \{e^i_+(a,\alpha)\}_{a,\alpha}\rangle, \tag{3.10}$$

$$\mathcal{A}^{ij} = \langle \mathcal{A}^i_- \cdot \mathcal{A}^j_+\rangle, \tag{3.11}$$

$$E_a = \sum_{\alpha=1}^{r} e_-(a,\alpha)e_+(a,\alpha). \tag{3.12}$$

The observables at the other two vertices are treated analogously. This modified hierarchy thus fulfills condition 1. listed above by construction. Theorem 3.2.1 then shows that the generalized Quantum de Finetti Theorem implies that there exists a state ρ such that condition 2. holds as well.

3.1.2 Example of measurement operators that do not generate elements from the local algebras

In this subsection, corresponding to Sec. 2.5.1 in Paper [1], we provide evidence for the claim that the algebra $\mathcal{D}$ that results from the original quantum inflation hierarchy does not in general contain the local observable algebras satisfying the two conditions laid out above. The purpose of the material presented here is to motivate our

ansatz and guide future research.

The strategy is to give a natural example of mutually commuting observable algebras $\mathcal{A}^i_-, \mathcal{A}^j_+$ and POVMs $E^{ij}_a \in \mathcal{A}^i_- \cdot \mathcal{A}^j_+$ such that the algebra generated by the $\{E^{ij}_a\}_{ija}$ does not contain any non-trivial local observable, i.e. no element in any of the $\mathcal{A}^i_-$ or $\mathcal{A}^j_+$ other than $\mathbb{1}$. This does not constitute a *proof* of the claim made above: We do not know whether there are correlations $P(A, B, C)$ that will cause the original inflation hierarchy to output such a model. But it does show that there are natural choices for the operators E^{ij}_a that fulfill all the constraints of the hierarchy, while failing to generate the local observables with respect to which the factorization properties of the causal structure are defined.

The model is very simple: For $i, j = 1, \ldots, n$, let $\mathcal{A}^i_-, \mathcal{A}^j_+$ be the observable algebra of one qubit each. Consider the maximally entangled *magic basis*

$$|\psi_1\rangle^{ij} = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)^{ij},$$

$$|\psi_2\rangle^{ij} = \frac{1}{\sqrt{2}}\big(|01\rangle + |10\rangle\big)^{ij},$$

$$|\psi_3\rangle^{ij} = \frac{1}{\sqrt{2}}\big(|00\rangle - |11\rangle\big)^{ij},$$

$$|\psi_4\rangle^{ij} = \frac{1}{\sqrt{2}}\big(|01\rangle - |10\rangle\big)^{ij},$$

and define POVMs

$$E^{ij}_a = |\psi_a\rangle\langle\psi_a|^{ij}.$$

Lemma 3.1.1. *The algebra $\mathcal{D}$ generated by $\{E^{ij}_a\}$ for $i, j = 1, \ldots n; a = 1, \ldots, 4$ does not contain any non-trivial local operator.*

Proof. The magic basis is a stabilizer basis, and we can thus express the projection operator onto each vector by summing over the respective stabilizer group. In terms of the usual Pauli operators, this gives

$$E^{ij}_1 = \frac{1}{4}(\mathbb{1} + X^i_- X^j_+ + Z^i_- Z^j_+ - Y^i_- Y^j_+), \tag{3.13}$$

$$E^{ij}_2 = \frac{1}{4}(\mathbb{1} + X^i_- X^j_+ - Z^i_- Z^j_+ + Y^i_- Y^j_+), \tag{3.14}$$

$$E^{ij}_3 = \frac{1}{4}(\mathbb{1} - X^i_- X^j_+ + Z^i_- Z^j_+ + Y^i_- Y^j_+), \tag{3.15}$$

$$E^{ij}_4 = \frac{1}{4}(\mathbb{1} - X^i_- X^j_+ - Z^i_- Z^j_+ - Y^i_- Y^j_+). \tag{3.16}$$

Let

$$\bar{X} = \prod_i^n X_-^i X_+^i, \qquad \bar{Z} = \prod_i^n Z_-^i Z_+^i.$$

Because distinct Pauli operators on the same system anti-commute,

$$[\bar{X}, E_a^{ij}] = [\bar{Z}, E_a^{ij}] = 0$$

and thus $\mathcal{D}$ is contained in the commutant of $\bar{X}, \bar{Z}$. But there is no non-trivial local operator that commutes with both $\bar{X}$ and $\bar{Z}$. $\square$

We note that Eqs. (3.13)-(3.16) imply that the effects E_a^{ij} have Schmidt-rank ≤ 4 and a product decomposition with factors of operator norm $C \leq \frac{1}{4}$.

3.2 A convergent hierarchy

The entirety of Section 3.2 corresponds to Section 4 of Paper [1].

Motivated by the difficulties that were outlined in Section 3.1.1, we propose a modified hierarchy of semidefinite programs for a Schmidt rank-constrained version of the quantum causal optimization problem that is provably complete. We show that by increasing the Schmidt rank, one can approximate any POVM arbitrarily well.

We use the triangle causal structure without settings as a guiding example to demonstrate the technique and to keep the notation relatively legible. More general scenarios can be accommodated – e.g. it is straight-forward to add additional generators to describe several possible POVMs per party. Extensions of these methods to arbitrary quantum causal structures are discussed in Section 3.2.3.

3.2.1 Construction of the hierarchy

The universal algebra of the quantum causal structure

First, we define generators and relations for the universal C^*-algebra $\mathcal{D}^n$ modeling the most general set of observables for the n-th inflation level of the causal structure. The algebra depends on a number of parameters:

1. The causal structure (taken to be the triangle scenario for now);

2. The number of outcomes M per vertex;

3. The inflation level n;

4. A bound r on the Schmidt rank of the measurement operators;

5. A bound C on the norm of the generators of the local algebra.

The dependency of $\mathcal{D}^n$ on the parameters will not be made explicit, with the exception of the inflation level.

From this data, define the set $\mathcal{G}^n$ of $6(M-1)rn+1$ generators to be

$$\{\mathbb{1}\} \cup \bigcup_{a,\alpha,i} \{e^i_-(a,\alpha), e^i_+(a,\alpha), f^i_-(a,\alpha), f^i_+(a,\alpha), g^i_-(a,\alpha), g^i_+(a,\alpha)\},$$

where

$$a \in \{1,\ldots,M-1\}, \quad \alpha \in \{1,\ldots,r\}, \quad i \in \{1,\ldots,n\}.$$

We will use the abbreviations

$$E^{ij}_a := \sum_{\alpha=1}^{r} e^i_-(a,\alpha)e^j_+(a,\alpha),$$

$$F^{ij}_a := \sum_{\alpha=1}^{r} f^i_-(a,\alpha)f^j_+(a,\alpha), \tag{3.17}$$

$$G^{ij}_a := \sum_{\alpha=1}^{r} g^i_-(a,\alpha)g^j_+(a,\alpha)$$

and

$$X^{ij}_M := \mathbb{1} - \sum_{a=1}^{M-1} X^{ij}_a, \qquad X \in \{E,F,G\}. \tag{3.18}$$

Four types of constraints are imposed.

1. Locality constraints: for all $x \in \{e,f,g\}$, $y \in \{-,+\}$, $a \in \{1,\ldots,M\}$, $\alpha \in \{1,\ldots,r\}$, $i \in \{1,\ldots,n\}$:

$$[x^i_y(a,\alpha), x'^{i'}_{y'}(a',\alpha')] = 0 \qquad \text{unless } x = x', y = y', \text{ and } i = i'. \tag{3.19}$$

2. Measurement constraints:

$$X^{ij}_a \quad \text{is positive,} \qquad X \in \{E,F,G\}. \tag{3.20}$$

3. Norm constraints:

$$\|g\| \leq C, \qquad \forall \, \mathbb{1} \neq g \in \mathcal{G}^n. \tag{3.21}$$

4. And finally that

$$\mathbb{1}x = x\mathbb{1} = x \tag{3.22}$$

for all generators x.

Together, these constraints define the set of relations $\mathcal{R}^n$. The NPO problem will run over states on the universal C^*-algebra $\mathcal{D}^n = C^*(\mathcal{G}^n | \mathcal{R}^n)$.

Polynomial constraints and objective function

The quantum causal polynomial optimization problem minimizes a polynomial function f_0 over compatible states $\rho \in K(\mathcal{D}^1)$ that also fulfill a number of polynomial constraints $f_i(\rho) = 0$. Here, we construct these objects precisely.

Choose some $g, k \in \mathbb{N}$. Recall the definition of the finite vector space $\mathcal{F}^{(k)}(\mathcal{G})$ of polynomials of order k in the generators $\mathcal{G}$ from Sec. 1.3.2. We assume that the functions are such that for every f_i, there exists a y_i in the g-fold algebraic tensor product $\mathcal{F}^{(k)} \otimes_{\mathrm{alg}} \cdots \otimes_{\mathrm{alg}} \mathcal{F}^{(k)}$ such that $f_i(\rho)$ equals the evaluation of the product state $\rho^{\otimes g}$ on y_i:

$$f_i(\rho) = \rho^{\otimes g}(y_i) \tag{3.23}$$

For our purposes, it will be enough to take Eq. (3.23) as the definition of the type of functions we allow for. We remark, though, that passing from a degree-g polynomial function f_i on $\mathcal{F}^{(k)}$ to a $y_i \in (\mathcal{F}^{(k)})^{\otimes_{\mathrm{alg}} g}$ such that Eq. (3.23) holds is known as a *polarization* in multi-linear algebra. In this context, it is proven that a unique suitable y_i always exists. Polarizations will be discussed in a more general context in Section 3.3 and Chapter 5.

As an example, consider the 2-norm distance that allows one to reduce Problem 3 to Problem 4 as we will see in Corollary 3.2.3. The objective function is then given by

$$\sum_{a,b,c} \left(\rho(E_a^{11} F_b^{11} G_c^{11}) - P(a,b,c) \right)^2. \tag{3.24}$$

To find the polarization, note that for a compatible state ρ it holds that

$$
\begin{aligned}
&\sum_{a,b,c} \left(\rho(E_a^{11} F_b^{11} G_c^{11}) - P(a,b,c) \right)^2 \\
&= \sum_{a,b,c} \rho(E_a^{11} F_b^{11} G_c^{11}) \rho(E_a^{22} F_b^{22} G_c^{22}) - 2P(a,b,c)\rho(E_a^{11} F_b^{11} G_c^{11}) + P(a,b,c)^2 \\
&= \rho^{\otimes 2} \left(\sum_{a,b,c} E_a^{11} F_b^{11} G_c^{11} E_a^{22} F_b^{22} G_c^{22} - 2P(a,b,c) E_a^{11} F_b^{11} G_c^{11} + P(a,b,c)^2 \right),
\end{aligned}
$$
(3.25)

which is indeed of the form $\rho^{\otimes g}(y_i)$.

We have now assigned a precise meaning to every object that appeared in the quantum causal polynomial optimization problem (Problem 4), which we restate here with constraints on the Schmidt rank of the POVM elements and the norm of the generators (i.e. as in Problem 5): Given a causal structure, a choice for the parameters M, r, C, and a family of polynomial functions f_i on $K(\mathcal{D}^1)$ as defined above and such that the $f_i, i \geq 1$ are non-negative on states that are compatible with the causal structure. Find

$$
\begin{aligned}
f_{r,C}^\star = \min_{\rho \in K(\mathcal{D}^1)} &\quad f_0(\rho) \\
\text{s.t.} \quad &\quad f_i(\rho) = 0 \quad i \geq 1 \\
&\quad \rho \text{ is compatible with the causal structure.}
\end{aligned}
$$
(3.26)

We adopt the common convention that $f_{r,C}^\star$ is ∞ in case the problem is infeasible.

NPO formulation

We now pass to an NPO problem, which we will show is asymptotically equivalent to the causal optimization problem in Eq. (3.26). To do so, we will replace the polynomial functions f_i by their polarizations y_i, and replace the causal constraint on ρ by symmetry constraints on a degree-n inflation.

Choose some n larger than or equal to the degree of y_0. For permutations $\pi, \pi', \pi'' \in S_n$ define an action on generators:

$$
\begin{aligned}
e_+^i &\mapsto e_+^{\pi(i)}, & f_-^i &\mapsto f_-^{\pi(i)}, \\
f_+^i &\mapsto f_+^{\pi'(i)}, & g_-^i &\mapsto g_-^{\pi'(i)}, \\
g_+^i &\mapsto g_+^{\pi''(i)}, & e_-^i &\mapsto e_-^{\pi''(i)}.
\end{aligned}
$$
(3.27)

Let $\alpha_{\pi,\pi',\pi''}$ be the extension of this action to $\mathcal{F}(\mathcal{G}^n)$.

The NPO problem relaxation of (3.26) at inflation level n is

$$f_{r,C}^n = \min_{\rho \in K(\mathcal{D}^n)} \rho(y_0)$$
$$\text{s. t.} \quad \rho(y_i) = 0, \qquad i \geq 1, y_i \in \mathcal{F}^{(n)}(\mathcal{G}), \tag{3.28}$$
$$\rho(b_j^{(n)} - \alpha_{\pi,\pi',\pi''}(b_j^{(n)})) = 0,$$

where the final symmetry constraint ranges over a basis $\{b_j^{(n)}\}$ for $\mathcal{F}^{(n)}(\mathcal{G})$ and a generating set of permutations in $S_n^{\times 3}$.

We note that (3.28) is not yet directly a semidefinite program. Instead, every instance gives rise to the infinite (but complete) hierarchy of SDP relaxations discussed in Section 1.3.2.

3.2.2   Proof of completeness

Now follows the proof that the inflation hierarchy of Eq. (3.28) is complete, i.e. that in the limit of $n \to \infty$, Eq. (3.28) and Eq. (3.26) are equivalent. Afterwards, we show that for every ϵ in the approximate quantum causal compatibility problem (Problem 3), there exist a Schmidt rank r and a norm bound C such that any compatible distribution can be ϵ-approximated by one that can be realized in a model that respects the bounds on r, C.

Theorem 3.2.1. *The hierarchy* (3.28) *is complete for the problem* (3.26) *in the sense that*

$$f_{r,C}^\infty := \lim_n f_{r,C}^n = f_{r,C}^\star.$$

Proof. Since each level of the hierarchy is a relaxation of the original problem, it holds that

$$f_{r,C}^n \leq f_{r,C}^\star \qquad \forall n. \tag{3.29}$$

The converse inequality is more involved. We start by constructing a state ω_n on $\mathcal{D}^\infty$ for each n by taking the infinite tensor product of some optimizing state of the problem in Eq. (3.28). By the Banach-Alaoglu Theorem applied to the state space $K(\mathcal{D}^\infty)$, there exists a weak*-convergent subsequence of the ω_n. Let ω be its limit point.

For each $i \geq 1$, y_i has a finite degree n_i. The constraint $\rho(y_i) = 0$ in (3.28) implies that $\omega_n(y_i) = 0$ for every $n \geq n_i$, and therefore the same is true for the

limit:

$$\omega(y_i) = 0.$$

Because ω_n is chosen to be an optimizer, $\omega_n(y_0) = f_{r,C}^n$ and thus

$$\omega(y_0) = \lim_{n \to \infty} f_{r,C}^n = f_{r,C}^\infty.$$

Likewise, the symmetry constraints in (3.28) imply that

$$\omega \circ \alpha_{\pi,\pi',\pi''} = \omega. \tag{3.30}$$

Restricting to the diagonal case $\pi = \pi' = \pi''$, we conclude that the limit ω is a symmetric state on $\mathcal{D}^\infty$, so that Theorem 2.1.1 applies.

Next, for each $1 \le n \le \infty$, introduce the algebras

$$(\mathcal{C}_+\mathcal{A}_-)^n, \qquad (\mathcal{A}_+\mathcal{B}_-)^n, \qquad (\mathcal{B}_+\mathcal{C}_-)^n,$$

where $(\mathcal{C}_+\mathcal{A}_-)^n \subset \mathcal{D}^n$ is the subalgebra generated by $\bigcup_{i \le n} \bigcup_{a,\alpha} \{g_+^i(a,\alpha), e_-^i(a,\alpha)\}$, and similar for $(\mathcal{A}_+\mathcal{B}_-)^n$ and $(\mathcal{B}_+\mathcal{C}_-)^n$. As n ranges over all natural numbers, the linear span of elements of the form

$$x = uvw,$$

with $u \in (\mathcal{C}_+\mathcal{A}_-)^n$, $v \in (\mathcal{A}_+\mathcal{B}_-)^n$ and $w \in (\mathcal{B}_+\mathcal{C}_-)^n$ is dense in $\mathcal{D}^\infty$. We therefore lose no generality by restricting the analysis of the action of ω to elements of this form.

Fix one $n \in \mathbb{N}$ and $x = uvw \in \mathcal{D}^n$. Using the cycle notation, define the permutations

$$\pi = (1, n+1)\,(2, n+2)\,\dots\,(n, 2n),$$
$$\pi' = (1, 2n+1)\,(2, 2n+2)\,\dots\,(n, 3n),$$

i.e. π exchanges the 1st block of n symbols with the 2nd block of n symbols and π' exchanges the 1st block of n symbols with the 3rd block of n symbols. Then

$$\omega(x) = \omega(\alpha_{\mathbb{1},\pi,\pi'}(x)) \tag{3.31}$$
$$= \omega\big(u\,\alpha_\pi(v)\,\alpha_{\pi'}(w)\big) \tag{3.32}$$

$$= \int \mathrm{d}\mu(\sigma)\, \Pi_\sigma\big(u\, \alpha_\pi(v)\, \alpha_{\pi'}(w)\big) \tag{3.33}$$

$$= \int \mathrm{d}\mu(\sigma)\, \Pi_\sigma(u)\, \Pi_\sigma\big(\alpha_\pi(v)\big)\, \Pi_\sigma\big(\alpha_{\pi'}(w)\big) \tag{3.34}$$

$$= \int \mathrm{d}\mu(\sigma)\, \Pi_\sigma(u)\, \Pi_\sigma(v)\, \Pi_\sigma(w), \tag{3.35}$$

where Eq. (3.32) follows from Eq. (3.30), Eq. (3.33) from Theorem 2.1.1, and in Eqs. (3.34) and (3.35) we have used that Π_σ is a symmetric product state for disjoint sets of layers of the inflation.

For each σ, the integrand in Eq. (3.35) factorizes (cf. Fig. 3.2). The respective marginals of Π_σ will be denoted as

$$\Lambda_\sigma^{\mathcal{C}_+\mathcal{A}_-} := \Pi_\sigma|_{(\mathcal{C}_+\mathcal{A}_-)^\infty}, \quad \Lambda_\sigma^{\mathcal{A}_+\mathcal{B}_-} := \Pi_\sigma|_{(\mathcal{A}_+\mathcal{B}_-)^\infty}, \quad \Lambda_\sigma^{\mathcal{B}_+\mathcal{C}_-} := \Pi_\sigma|_{(\mathcal{B}_+\mathcal{C}_-)^\infty}, \tag{3.36}$$

so that the product state appearing in the integrand is

$$\Lambda_\sigma := \Lambda_\sigma^{\mathcal{C}_+\mathcal{A}_-} \otimes \Lambda_\sigma^{\mathcal{A}_+\mathcal{B}_-} \otimes \Lambda_\sigma^{\mathcal{B}_+\mathcal{C}_-}. \tag{3.37}$$

Therefore, ω is a convex combination

$$\omega(x) = \int \mathrm{d}\mu(\sigma)\, \Lambda_\sigma(x). \tag{3.38}$$

of states Λ_σ that are compatible with the causal structure.

It remains to be shown that we can choose one σ, such that $\Lambda_\sigma(y_0) = f_{r,C}^\infty$ and $\Lambda_\sigma(y_i) = 0$ for all $i \geq 1$.

By the definition of Problem 4, the y_i are non-negative on states compatible with the causal structure, i.e. $\Lambda_\sigma(y_i) \geq 0$ for all $i \geq 1$. Because $\omega(y_i) = 0$ as well, the constraints must be fulfilled on a set $E \subset K(\mathcal{D})$ of measure $\mu(E)$ equal to one. For every Λ_σ with $\sigma \in E$ it must hold that $\Lambda_\sigma(y_0) \geq f_{r,C}^\infty$, for else one could have chosen μ to be the point measure on a state $\sigma' \in E$ with $\Lambda_{\sigma'}(y_0) < f_{r,C}^\infty$, which contradicts the fact that $f_{r,C}^\infty$ is a minimum. As before, there must be a subset $F \subset E$ of measure $\mu(F)$ equal to one such that $\Lambda_\sigma(y_0) = f_{r,C}^\infty$ for all $\sigma \in F$. Therefore, any state Λ_σ such that $\sigma \in F$ is compatible with the constraints of Eq. (3.26), so that we can conclude

$$f_{r,C}^\infty = \Lambda_\sigma(y_0) \geq f_{r,C}^\star \quad \forall \sigma \in F. \tag{3.39}$$

Combining Eqs. (3.29) and (3.39) yields $f_{r,C}^\infty = f_{r,C}^\star$.   □

Remarks.

1. It is not obviously possible to extract a compatible state from the SDP.

2. The proof of Theorem 3.2.1 shows that it is in general not possible to add additional constraints of the form $\rho(x) \geq 0$ to the program for elements x that are not necessarily positive on compatible states: the states Λ_σ that are compatible with the causal structure might not obey these constraints, since they only apply to ω. That is, the set E defined in the proof will in general not have full measure. However, if the optimization problem is a feasibility problem, i.e. if it has a trivial objective function, it is possible to put one such constraint as the objective function and reject the solution if the optimal value does not obey the inequality. We will apply this to the constraints of the causal compatibility problem in Corollary 3.2.3 below.

Lemma 3.2.2. *Consider a probability distribution P that is compatible with a given causal structure. Choose $\epsilon > 0$. There exist constants C, r such that there is a distribution $\tilde{P}$ that approximates P in the sense that*

$$\|P - \tilde{P}\|_2^2 \leq \epsilon$$

which can be realized using only POVM elements of the form

$$\tilde{E} = \sum_{\alpha=1}^{r} e_-(\alpha) \cdot e_+(\alpha), \qquad \text{such that } \|e_-(\alpha)\|, \|e_+(\alpha)\| \leq C. \tag{3.40}$$

Proof. Consider the original model that gives rise to P. By the definition of a C^*-tensor product, for each $a = 1, \ldots, M - 1$, there is a convergent series

$$E_a = \sum_{\alpha=1}^{\infty} e_-(a, \alpha) \cdot e_+(a, \alpha), \qquad e_-(a, \alpha) \in \mathcal{A}_-, e_+(a, \alpha) \in \mathcal{A}_+.$$

Let $E_a^{(r)}$ be the truncation of the series to the first r terms. Convergence implies that for every $\delta > 0$, there exists an r such that

$$\|E_a^{(r)} - E\| \leq \delta \qquad a = 1, \ldots, M - 1.$$

What remains to be proven is that one can turn these partial sums into an exact POVM. To this end, set

$$\tilde{E}_a = \frac{1}{1 + 2M\delta}(\delta \mathbb{1} + E_a^{(r)}) \qquad a = 1, \ldots, M - 1. \tag{3.41}$$

Then the $\tilde{E}_a$ are positive. What is more,

$$
\begin{aligned}
\left\| \sum_{a=1}^{M-1} \tilde{E}_a \right\| &\leq \frac{1}{1+2M\delta} \left(\sum_{a=1}^{M-1} \|\delta \mathbb{1}\| + \left\| \sum_{a=1}^{M-1} (E_a^{(r)} + E_a - E_a) \right\| \right) \\
&\leq \frac{1}{1+2M\delta} \left((M-1)\delta + \left\| \sum_{a=1}^{M-1} E_a \right\| + \left\| \sum_{a=1}^{M-1} (E_a^{(r)} - E_a) \right\| \right) \\
&\leq \frac{1}{1+2M\delta} \big((M-1)\delta + 1 + (M-1)\delta \big) < 1
\end{aligned}
$$

so that

$$
\tilde{E}_M := \mathbb{1} - \sum_{a=1}^{M-1} \tilde{E}_a
$$

is also positive. Therefore $\{\tilde{E}_1, \ldots, \tilde{E}_M\}$ forms a POVM. Repeating the construction, one arrives at approximations $\tilde{F}_b$ to F_b and $\tilde{G}_c$ to G_c.

From Eq. (3.41), the approximating POVM elements converge to the original ones in operator norm as $\delta \to 0$. The same is thus true for all polynomial expressions in the POVM elements. Therefore,

$$
\tilde{P}(a,b,c) := \rho(\tilde{E}_a \tilde{F}_b \tilde{G}_c) \to P(a,b,c) \qquad (\delta \to 0)
$$

and, because there are only finitely many outcomes,

$$
\|\tilde{P} - P\|_2^2 \to 0 \qquad (\delta \to 0).
$$

Thus, choosing r sufficiently high, an arbitrarily good approximation can be achieved. The advertised claim follows by choosing C to be the largest operator norm of any factor of the partial sums involved.

$\square$

Corollary 3.2.3. *Given a probability distribution P over observed variables, the SDP hierarchy that corresponds to the optimization problem of Eq. (3.28) can solve the approximate quantum causal compatibility problem described in Problem 3.*

Proof. In order to solve Problem 3, we need to show that there exists a state ρ that is compatible with the description of the causal structure and that produces statistics that are close in 2-norm to the observed statistics. In particular, for the triangle

scenario it must hold that

$$\sum_{a,b,c} \left(\rho(E_a^{11} F_b^{11} G_c^{11}) - P(a,b,c)\right)^2 \leq \epsilon, \tag{3.42}$$

where ρ is a compatible state. Once again we can polarize the expression, which yields the objective function

$$\min_{\rho \in K(\mathcal{D}^2)} \sum_{a,b,c} \rho(E_a^{11} F_b^{11} G_c^{11} E_a^{22} F_b^{22} G_c^{22}) - 2P(a,b,c)\rho(E_a^{11} F_b^{11} G_c^{11}) + P(a,b,c)^2, \tag{3.43}$$

as was shown in Eq. (3.25).

    If the NPO hierarchy attains the optimal value $f_{r,C}^\infty$ for this objective function, there also exists a product state $\Pi_\sigma \in K(\mathcal{D}^\infty)$ that is compatible with the infinitely inflated causal structure that attains the same optimal value by Theorem 3.2.1. If $f_{r,C}^n > \epsilon$ for any n, we reject the hypothesis that the given description of the causal structure with measurement operators of rank r and generators with a norm-bound of C can produce the observed statistics. If there does exist a quantum description of $P(A, B, C)$ Lemma 3.2.2 ensures that there exist r and C such that the optimal value is not rejected for any n. In that case the restriction of Π_σ to $\mathcal{D}$ is a product state that is compatible with the triangle causal structure and that approximately produces the probability distribution $P(A, B, C)$. $\square$

Remark. Though in the limit of $n \to \infty$ the objective function (3.43) is equivalent to Eq. (3.42), Eq. (3.43) is likely to be impractical to detect incompatibility. For low values of n the state will not be separable and this objective function can become negative. Fortunately, once convergence is proven, it is possible to add additional constraints that are as strict as, or relaxations of, Problem 3: Denoting the optimal value of such an adjusted hierarchy by $\tilde{f}_{r,C}^n$, it holds that $f_{r,C}^n \leq \tilde{f}_{r,C}^n \leq f_{r,C}^*$. Then $\lim_{n\to\infty} f_{r,C}^n = f_{r,C}^*$ implies $\lim_{n\to\infty} \tilde{f}_{r,C}^n = f_{r,C}^*$. This means in particular that we can simply add the linear equality constraints (3.8), while retaining convergence. This should allow for detection of incompatibility at much lower levels than without such constraints.

3.2.3   Arbitrary quantum causal structures

In this section we will generalize the results presented above to arbitrary causal structures for which all leaf nodes are observed classical variables[3]. Ref. [22, Sec. V] dis-

---

[3] Recall that one can also define quantum causal structures that give rise to quantum states rather than classical variables [11]. Such scenarios are beyond the scope of this chapter, but will be discussed in the

cusses a number of transformation rules that bring such general causal structures into a form amenable to the hierarchy introduced above. It is argued there that these rules are sufficient to treat every structure that gives rise to classical random variables. We will not repeat this argument here. However, for each of their transformation rules, we will explain how they can be applied in our modified framework.

We will first argue that the approach described in the previous sections is applicable to all *network scenarios*, which are two-layered causal structures with observed leaf nodes, i.e. nodes without children. In the second step, we will map any *latent exogenous* causal structure to such a network scenario. Latent exogenous causal structures are those in which the only unobservable nodes are root nodes, i.e. nodes that have no parents. The final step is to map non-exogenous causal structures to exogenous ones by introducing a new type of node that can also be treated in our model.

Network scenarios

It is not difficult to see that Theorem 3.2.1 and Corollary 3.2.3 can be extended to arbitrary *correlation scenarios*, which are the causal structures that only have a bottom layer of independent latent systems with arrows pointing to a top layer of observed variables. The triangle scenario is an example of a correlation scenario. If the causal structure has L latent (quantum) variables, one can employ the proof strategy of Theorem 3.2.1 by writing an operator on n inflation levels as an operator on $L \cdot n$ levels of inflation. Indeed, for the triangle scenario one has 3 latent quantum systems, and thus $3n$ levels of inflation were sufficient in the reasoning of Eqs.(3.31)-(3.35). The proof that causal polynomial optimization can be solved with an SDP hierarchy as described above remains nearly the same, with the only difference that the algebra $\mathcal{D}^n$ modeling the level-n inflated causal structure has to be defined in accordance with the proposed causal structure. To show that causal compatibility is also solved, one just needs to write down a similar objective function as in Eq. (3.43) for the given probability distribution.

By allowing classical root nodes that only have one child in two-layered causal structures, one obtains so called *network scenarios* (see e.g the Bell scenario of Fig. 1). Whenever a classical, observed variable is an input to another observed variable, we give the POVM elements of the latter variable an extra index. For example in the Bell scenario, where Alice has the input variable X, the POVM elements of Alice become $\{E_a^x\}$, with $\sum_a E_a^x = \mathbb{1}$ for every x, so that x can be interpreted as a measurement setting. POVM elements with different measurement settings, e.g. E_a^x and $E_a^{x'}$, need no longer commute. Though this description introduces more variables to the model, nothing major changes in the proofs.
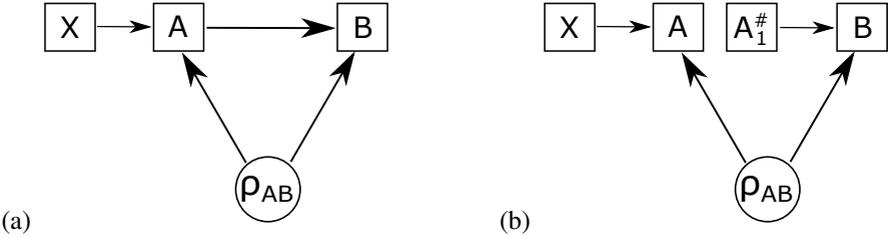
---

network scenarios of Chapter 4.

Figure 3.3: (a) The instrumental scenario as an example of a latent exogenous causal structure. The variable A is both a parent and a child and thus the causal structure is not a network scenario. By splitting A into A and $A_1^{\#}$, as in (b), and post-selecting $A_1^{\#}$ on the outcome of A, the instrumental scenario can be modeled by a network scenario, which happens to be the Bell scenario. This process is an example of maximal interruption.

Latent exogenous causal structures

It is also possible to extend our result to quantum causal structures with more than two layers. The reduction of the general case to the proof methods considered here is not immediate. We follow the approach of Wolfe *et al.* [22], who generalize to arbitrary causal structures in two steps. In both cases, they offer a solution for how to alter the description of these causal structures such that they fit in the framework of inflation and NPO. We will adopt the first and alter the second method to adhere to our formalism. One then has to show that these descriptions still obey all the results of the previous sections. Here we briefly outline why this is indeed the case and give the general transformation rules to map those causal structures to equivalent network scenarios.

   In the first generalization step, the causal structure is also allowed to contain observed nodes that have both one or more parents and one or more children. However, all unobserved nodes remain root nodes. Such causal structures are called *latent exogenous*. An example is the instrumental scenario in Fig. 3.3. The probability distribution of the instrumental scenario is denoted by $P_{IS}(A, B, X)$. However, it is more common to express the statistics in terms of conditional probabilities. For a set of N random variables $\{A_1, \ldots, A_N\}$ conditioned on K independent variables $\{X_1, \ldots, X_K\}$, the statistics are fully captured by the combination of the conditional probabilities

$$P(A_1, \ldots, A_N \mid X_1 \ldots X_K) = \frac{P(A_1, \ldots, A_N, X_1 \ldots X_K)}{P(X_1 \ldots X_K)}, \qquad (3.44)$$

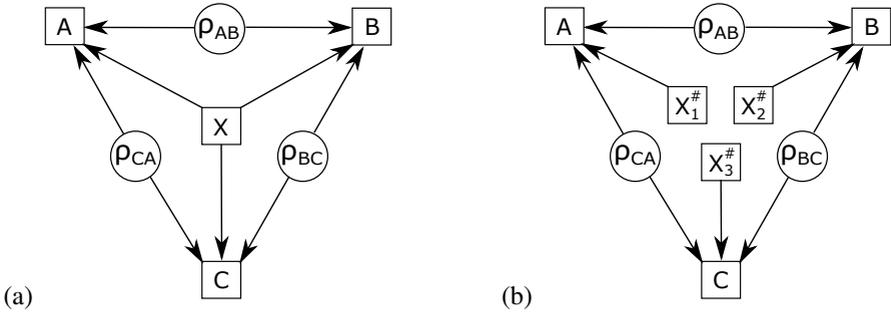(a)   (b)

Figure 3.4: (a) The triangle scenario with a shared setting X. Via maximal interruption the observed variable X is split into three independent and identically distributed random variables, producing the network scenario (b). The allowed probability distributions of the original causal structure can be obtained from the network scenario by post-selection.

and the requirement that the setting-associated variables factorize:

$$P(X_1, \ldots, X_K) = P(X_1) \ldots P(X_K). \tag{3.45}$$

We refer to Ref. [86] for a fuller discussion on why this is necessary and how this works in more general (classical) setups. Hence, checking for compatibility will consist of two separate steps: First one has to confirm that variables that are being conditioned on form a product distribution; Secondly, one checks compatibility with the causal structure via the procedure outlined below.

In particular, for the instrumental scenario this simply reduces to checking compatibility of

$$P_{IS}(a, b \mid x) = \frac{P_{IS}(a, b, x)}{P_{IS}(x)}. \tag{3.46}$$

The classical random variable A is both a child of ρ_{AB} and X, as well as a parent of B. This problem is resolved by a process known as *maximal interruption* [87, 88, 89, 90]: the variable A is split into two random variables A and $A_1^{\#}$, where A is only a child and $A_1^{\#}$ is only a parent. The causal structure is then effectively mapped to the Bell scenario. By post-selecting $A_1^{\#}$ on the outcome $A = a$, i.e. by setting $P_{IS}(A = a, B = b \mid X = x) = P_{Bell}(A = a, B = b \mid X = x, A_1^{\#} = a)$, one can still obtain the allowed distributions of the original graph.

In the case that an observable node has multiple children, one applies a similar

splitting and post-selection procedure. Consider for example the triangle scenario with a shared setting X depicted in Fig 3.4(a). Via maximal interruption the observed variable X is split up into three independent and identically distributed random variables $X_1^{\#}$, $X_2^{\#}$ and $X_3^{\#}$ as depicted in Fig 3.4(b). The resulting causal structure is a network scenario. By post-selecting on $X_1^{\#} = X_2^{\#} = X_3^{\#} = x$ one can link the (conditional) probability distribution of the network scenario to the (conditional) probability distribution $P(A, B, C \mid X = x)$ of the triangle with a shared setting and find the allowed distributions in the original causal structure.

In this way, one can map all latent exogenous causal structures to network scenarios. The general rule is then as follows: Whenever an observed node is not a leaf node and is directly connected to multiple other nodes, split the node into as many copies as there are outgoing arrows. Remove every outgoing arrow from the original node and attach it to a copy. If there are no incoming arrows to the original node, remove it. Check whether the setting-associated variables factorize into a product distribution. Finally, analyze this causal structure, which is now a network scenario, and apply post-selection on the copies.

Non-exogenous causal structures

The second step in the generalization is more involved. The causal structures are now also allowed to have latent (quantum) variables with parents. In classical causal structures it is possible to transform these *non-exogenous* causal structures into exogenous ones, for which it is then possible to apply maximal interruption as described above, if necessary [91]. However, for quantum causal structures this is in general not possible.

Ref. [22] gives a clear example (credited there to Stefano Pironio), which is depicted in Fig. 3.5. In this example, the quantum system ρ_{BC} in Fig. 3.5(a) is non-exogenous. The structure in Fig. 3.5(b) has been exogenized as if it were a classical causal structure, by removing ρ_{BC} and drawing arrows from the parents of ρ_{BC} to its children. It can be seen that in the original causal structure it is possible to maximally violate a Bell inequality for either the systems A and B or A and C, based on the setting determined by S. However, after the exogenization depicted on the right, the setting S cannot determine anymore which pair maximally violates a Bell inequality. Since it is impossible that both A and B, as well as A and C maximally violate a Bell inequality due to monogamy of entanglement [92], the causal structure on the right cannot produce the same statistics as the one on the left.

We will split the treatment of non-exogenous causal structures in two parts: First, we will consider unobservable systems with one or more observed parents and at most one unobservable parent. Secondly, we will regard unobservable systems with multiple unobservable parents, but no observed parents. These two solutions can

be combined to form a general set of rules for treating non-exogenous causal structures. It will turn out that, even though we cannot directly apply the exogenization procedure for classical causal structures, each observable leaf node will instead get an index associated to each of its parents in the classically exogenized causal structure. These indices then contain the information about commutation relations and independence constraints for the causal structure.

In Ref. [22] unobservable systems are eliminated by instead regarding such a system as a quantum channel applied to its unobservable parents and regarding any observed parents as a classical control for this quantum channel. If there is no observed parent to a non-exogenous system, the quantum channel that replaces it does not have such a control variable.

We opt for a slightly different treatment of non-exogenous systems with a similar interpretation. Instead of acting with a quantum channel on a state, we alter the POVM elements. Consider again the causal structure of Fig. 3.5(a) as an example. The intermediate state ρ_{BC} has the interpretation of redistributing the S subsystem of ρ_{AS} among Bob and Charlie, based on the observed variable S. Hence, for different outcomes s of S, Bob and Charlie will perform measurements over different parts of the S subsystem of ρ_{AS}. However, for every specific outcome s, the measurements operators will be given by commuting POVMs for Bob and Charlie. We therefore define for every outcome of S the commuting algebras $\mathcal{B}^{(s)}$ and $\mathcal{C}^{(s)}$ with elements $\{F_b^{(s)}\}_b$ and $\{G_c^{(s)}\}_c$ respectively. Elements from these algebras obey the commutation relations

$$[F_b^{(s)}, G_c^{(s)}] = 0, \tag{3.47}$$

but whenever two operators have different indices s and s', they are no longer required to commute. The difference between this description and the measurement settings in network scenarios is that in network scenarios all POVM elements of Bob commute with all POVM elements of Charlie, while that is no longer the case here. We therefore propose a new graphical notation for this exogenization procedure, in which the nodes of Bob and Charlie are initially combined and only become commuting POVMs after the measurement setting s has been processed (see Fig. 3.5(c)). We will call such nodes *endogenous* nodes.

One can still apply inflation to this causal structure as well. The algebras $\mathcal{A}$, $\mathcal{B}^{(s)}$ and $\mathcal{C}^{(s)}$ will be copied and get the index i corresponding to the i'th copy of ρ. Different inflation levels will be modeled by commuting subalgebras with an exchange symmetry and one can show, using the de Finetti theorem, that in the limit a symmetric global state is separable across copies of ρ. The proofs of Theorem 3.2.1 and Corollary 3.2.3 then also follow.

In general, one can apply the following rule to remove a non-exogenous system
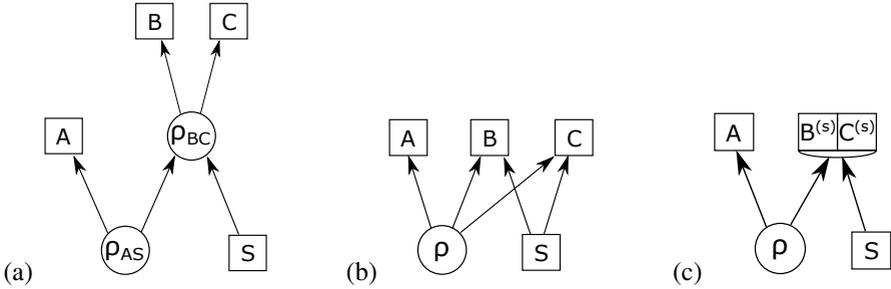
Figure 3.5: (a) An example of a non-exogenous causal structure. It is not possible to exogenize such a quantum causal structure as depicted in (b), like one would do for classical causal structures. This can be seen by noting that structure (a) can maximally violate a Bell inequality for A and B *or* for A and C based on the measurement setting S. Due to monogamy of entanglement this is not possible for structure (b). In structure (c) this is solved by first regarding Bob and Charlie as one observer and only choosing the commuting algebras $\mathcal{B}$ and $\mathcal{C}$ after the setting S has been received. This procedure is represented by a new type of node.

with observed parents and at most one latent parent, starting with non-exogenous systems that are closest to a leaf node: Split up each leaf node according to the structure of its local algebras, similar to the triangle scenario. Combine all leaf nodes that have a directed path from the non-exogenous system to that leaf node into one endogenous node. For every observed variable that is a parent of the non-exogenous node, introduce an index to the elements of the endogenous node. Elements of the algebra of the endogenous node commute if all such setting indices are the same and the elements originated from spatially separated systems (e.g. Bob and Charlie).

The final class of causal structures that has not been discussed yet, is the one where there are multiple unobserved parents to a latent variable. We will again first treat an example and then give the general rule.

Consider the causal structure in Fig. 3.6(a). The intermediate node ρ_{BC} is non-exogenous and has two latent parents. We start by splitting up the algebras $\mathcal{C}$ and $\mathcal{D}$ into their minus and plus sub-algebras, similar to the triangle scenario. Then we remove the non-exogenous node by taking those algebras together into an endogenous node that are its descendants, as was done in Fig. 3.5(c). In this case, that will remove ρ_{BC} and combine $\mathcal{B}$ and $\mathcal{C}_-$ into an endogenous node.

When the causal structure is inflated, the root nodes are copied and given inflation indices i, j, k respectively. The intermediate state ρ_{BC} would then have gotten
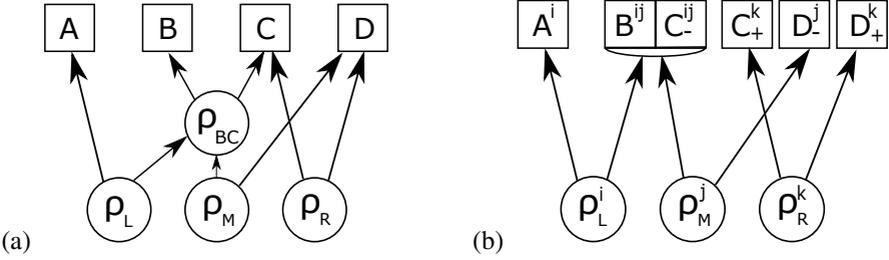
Figure 3.6: (a) A more complicated causal structure, where ρ_{BC} has two latent parents and two observable children. The inflated version of this structure can alternatively be depicted as in Fig. (b). The algebras $\{\mathcal{B}^{ij}\}_{ij}$ and $\{\mathcal{C}^{ij}_-\}_{ij}$ are taken together and do not commute if there is only one overlapping index.

the two inflation indices i, j, which in turn are *both* passed down to $\mathcal{B}$ and $\mathcal{C}_-$ (see Fig. 3.6(b)). We thus have the following algebras after inflation: $\mathcal{A}^i, \mathcal{B}^{ij}, \mathcal{C}^{ij}_-, \mathcal{C}^k_+, \mathcal{D}^j_-$ and $\mathcal{D}^k_+$. Operators from these algebras will be denoted in a similar way. Let $\mathcal{E}^n$ be the algebra describing the level-n inflated causal structure. Though each of these algebras is a subalgebra of $\mathcal{E}^n$, it is no longer true that $\mathcal{E}^n$ is the tensor product of all these algebras, because some of them do not commute. This is due to the "mixing" of the root nodes ρ_L^i and ρ_M^j by the intermediate nodes ρ_{BC}^{ij}. In particular, for $B^{ij} \in \mathcal{B}^{ij}$ and $C_-^{i'j'} \in \mathcal{C}_-^{i'j'}$

$$[B^{ij}, C_-^{i',j'}] \begin{cases} = 0 & \text{if } i = i', j = j' \text{ or } i \neq i', j \neq j', \\ \neq 0 & \text{if } i = i', j \neq j' \text{ or } i \neq i', j = j'. \end{cases} \tag{3.48}$$

By requiring that the POVM elements of Bob and Charlie commute when they perform a measurement on the same state ρ_{BC}^{ij}, or on independent copies of the state, ρ_{BC}^{ij} and $\rho_{BC}^{i'j'}$ with $i \neq i', \ j \neq j'$, we ensure that we still model spatially separated measurements in a physical scenario.

The question is now which independence relations hold and how to properly apply the quantum de Finetti theorem to show that these relations hold asymptotically in the inflation formalism.

To answer this question, note that the state ρ_{BC} masks the independence of the parts of ρ_L and ρ_M that are sent to Bob and Charlie: The correlations in this causal structure could have also been produced by a four-partite state, of which the first and fourth subsystems are required to be independent, combined with the independent bi-partite state ρ_R. The independence requirements that still have to hold after the

quantum channel that produces ρ_{BC} are thus

$$\rho(A^i B^{ij} C^{ij}_- C^k_+ D^j_- D^k_+) = \rho(A^i B^{ij} C^{ij}_- D^j_-)\rho(C^k_+ D^k_+) \qquad \forall i,j,k, \qquad (3.49)$$

$$\rho(A^i D^j_-) = \rho(A^i)\rho(D^j_-), \qquad\qquad\qquad \forall i,j \qquad (3.50)$$

$$\rho(\prod_{i=1}^{n} A^i B^{ii} C^{ii}_- C^i_+ D^i_- D^i_+) = \rho(\prod_{i=1}^{n} A^{\pi(i)} B^{\pi(i)\pi'(i)} C^{\pi(i)\pi'(i)}_- C^{\pi''(i)}_+ D^{\pi'(i)}_- D^{\pi''(i)}_+)$$

$$= \prod_{i=1}^{n} \rho(A^i B^{ii} C^{ii}_- C^i_+ D^i_- D^i_+),$$

$$(3.51)$$

where the first two equalities signify independence within layers of inflation (namely the independence of ρ_R with respect to ρ_L and ρ_M, and the independence of A and D), while the last equality corresponds to independence between layers of inflation. Note, however, that the algebras $\mathcal{B}^{ij}$ and $\mathcal{C}^{ij}_-$ do not have this independence between inflation layers if the state is evaluated over products of operators of either algebra for which only one of the two indices coincides, e.g. $\rho(B^{ij} C^{i'j}_-) \neq \rho(B^{ij})\rho(C^{i'j}_-)$ and $\rho(B^{ij} B^{i'j}) \neq \rho(B^{ij})\rho(B^{i'j})$.

Relaxing the independence conditions of Eqs. (3.49)-(3.51) to their corresponding symmetry constraints and combining this with the commutation relations and the de Finetti Theorem, we can still asymptotically solve the (rank-constrained) causal compatibility problem even for this causal structure.

The general rule to map each inflation level of a causal structure with non-exogenous systems that have multiple latent parents, but no observed parents to an equivalent latent exogenous causal structure is then as follows: Start with the non-exogenous system that is closest to a leaf node. Split up each leaf node according to the structure of its local algebras. Combine all leaf nodes that have a directed path from the non-exogenous system to that leaf node into one endogenous node. For every root node that is an ancestor of the non-exogenous node, attach the inflation index of that root node to the elements of the endogenous node. Elements of the algebra of the endogenous node commute if either (1) all inflation indices are pair-wise the same, i.e. $i = i', j = j', \ldots$, and the elements originated from spatially separated systems, e.g. Bob and Charlie, or if (2) all inflation indices are pair-wise different, i.e. $i \neq i', j \neq j', \ldots$.

More succinctly: to reduce a non-exogenous causal structure to an exogenous one, we attach to every child of a non-exogenous node an index for all of the root nodes of the non-exogenous system, and apply the appropriate commutation relations.

Our approach is thus applicable to all relevant quantum causal structures, by se-

quentially applying the techniques outlined above, i.e. by first turning a non-exogenous causal structure into an exogenous one with a new type of node that can also be treated in our model, and then applying maximal interruption to turn it into a network scenario. Checking the feasibility of the SDP hierarchy of this network scenario, as well as the required factorization of the setting-associated variables is then a necessary and sufficient procedure for checking causal compatibility.

3.3   The bilocal scenario

This section will focus on the *bilocal scenario*. The text largely coincides with that of Paper [2].

   In the bilocal scenario (Fig. 3.7), we are concerned with the set of correlations that can be obtained by three parties (Alice, Bob, and Charlie) performing measurements on pairs of quantum particles originating from two independent sources: one distributing a pair between Alice and Bob, and one between Bob and Charlie. We assume that each party can choose among a finite number of measurement settings, their choices being labeled by numbers x, y, z. Each then obtains one of a finite number of possible outcomes. We represent their respective outcomes by α, β, γ. The statistics of such an experiment are then described by a collection $p(\alpha\beta\gamma|xyz)$ of conditional probabilities.

   The bilocal scenario is one of the most fundamental causal structures: It is the simplest non-trivial structure in which source states are assumed to be independent. It is also a straightforward generalization of the Bell scenario. Nevertheless it allows for new behaviour such as entanglement swapping [25] and is surprisingly hard to analyze. Here, we are primarily concerned with the *bilocal causal compatibility problem*: Given a collection of conditional probabilities $p(\alpha\beta\gamma|xyz)$, decide whether it is compatible with an experiment of the form described above.

   Several techniques to answer this question have already been developed. These include, but are not limited to, (non-linear) Bell inequalities [93], machine learning techniques [94], information-theoretic methods [11], scalar extension [41, 82, 83] and the inflation technique that is also considered in this chapter [1, 21, 22, 23]. For a more complete list, both on the bilocal scenario and more general network scenarios, we refer to the excellent review of Ref. [95].

   Recently, the authors of Ref. [83] asked whether the *quantum inflation technique* is complete for the bilocal compatibility problem. One of the main results of this chapter, partly building on their constructions, is to answer this question in the affirmative. We develop two complete semidefinite programming hierarchies that are closely related. The first, which we call the *polarization hierarchy*, uses symmetric product states to linearize the non-convex independence constraint. The second is a

version of the quantum inflation hierarchy, which relaxes the independence condition to a family of linear symmetry constraints. Along the way, we obtain a number of equivalent characterizations of bilocal quantum correlations, which might be of independent interest.

3.3.1   Quantum models of locality

In order to give a precise definition of the set of *bilocal quantum correlations*, one needs to fix a quantum model of locality. As noted before in Sections 1.2 and 3.1, this turns out to be a surprisingly subtle issue. Here we briefly repeat part of the discussion to motivate the several definitions for bilocal quantum correlations.

There are two commonly used "pictures" on which a formalization of quantum mechanical descriptions of Nature can be based.

In elementary quantum mechanics (related to the Schrödinger picture), the fundamental mathematical object associated with a quantum system is a Hilbert space $\mathcal{H}$. The set of observables is then derived as the algebra of bounded operators $B(\mathcal{H})$ acting on $\mathcal{H}$.

Alternatively, in *algebraic quantum mechanics* [52, 96, 97] (related to the Heisenberg picture), quantum systems are primarily described via an algebra $\mathcal{A}$ of observables. A Hilbert space is then a secondary object, which can be derived e.g. via the GNS construction [43].

The two points of view are mostly equivalent as a basis for describing natural phenomena. Differences are commonly associated with finer technical points, e.g. in the rigorous description of the thermodynamic limit [52]. One would thus assume that the choice of which point of view to adopt becomes a matter of taste and convenience. While most working physicists prefer the Schrödinger picture, the algebraic model is easier to reason about algorithmically, which explains its use in completeness proofs such as those of Refs. [1, 40, 56].

However, the two approaches suggest different formalizations of the notion of "locality", which is obviously relevant for the problem treated in this chapter.

Indeed, consider two spatially separated subsystems A, B of some composite system. Separation implies that physical properties of A and B can be simultaneously measured, which means that the associated observable algebras $\mathcal{A}$, $\mathcal{B}$ must mutually commute, $[a, b] = 0, a \in \mathcal{A}, b \in \mathcal{B}$. In algebraic quantum mechanics, this assumption (sometimes referred to as *Einstein locality* [96, Sec. 8.5]) is the only one made.

In contrast, the Schrödinger picture-approach is to associate one Hilbert space

$\mathcal{H}_A, \mathcal{H}_B$ with each subsystem and to take the observable algebras to be

$$\begin{aligned}
\mathcal{A} &= B(\mathcal{H}_A) \otimes \mathbb{1} \subset B(\mathcal{H}_A \otimes \mathcal{H}_B), \\
\mathcal{B} &= \mathbb{1} \otimes B(\mathcal{H}_B) \subset B(\mathcal{H}_A \otimes \mathcal{H}_B)
\end{aligned} \tag{3.52}$$

respectively.

The surprisingly technically complex theory of tensor products of operator algebras [42] shows that not every pair of commuting algebras can be realized on a tensor product of Hilbert spaces as in Eq. (3.52). For a considerable time, it was an open question (known as *Tsirelson's Problem* [53, 54, 55]), whether these operator-theoretic subtleties would manifest themselves at the level of finite sets of observable correlation functions (as made precise in Sec. 3.3.2). Unfortunately, it has now become clear that this is indeed the case [17]. Thus, whenever one speaks about "quantum correlations", one has to be specific as to whether one is working in the more restrictive tensor product Hilbert space model or the more general commuting observable model.

At present, there does not seem to be strong evidence indicating which of the two approaches is more relevant for the description of natural phenomena. Both are legitimate targets of inquiry, as long as authors indicate clearly (as we have tried to do) which model they are working with at any time.

3.3.2   Quantum correlations

Two-party quantum correlations

As a warm-up, we can now state precisely the two well-known distinct models of *two-party quantum correlations*, i.e. the set of conditional probabilities $p(\alpha\beta|xy)$ obtainable by two parties performing local measurements on a shared quantum state.

Definition 3.3.1 (Two-party correlations, tensor product model). A set $p(\alpha\beta|xy)$ of conditional probabilities is a *bipartite quantum distribution in the Hilbert space tensor product model* if the following holds. There are

- Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$,

- for each of Alice's settings x a POVM $\{A_{\alpha|x}\}_\alpha \subset B(\mathcal{H}_A)$, and for each of Bob's settings y a POVM $\{B_{\beta|y}\}_\beta \subset B(\mathcal{H}_B)$,

- a density operator ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$

such that

$$p(\alpha\beta|xy) = \mathrm{tr}\big(\rho\, A_{\alpha|x} \otimes B_{\beta|y}\big).$$

In order to highlight the essential difference, we first give a version of the commuting observables model that is phrased as closely as possible to the tensor product model.

Definition 3.3.2 (Two-party quantum correlations, commuting observables model)**.**
A set $p(\alpha\beta|xy)$ of conditional probabilities is a *bipartite quantum distribution in the commuting observables model* if the following holds. There is

- a Hilbert space $\mathcal{H}$

- for each of Alice's settings x a POVM $\{A_{\alpha|x}\}_\alpha \subset B(\mathcal{H})$, and for each of Bob's settings y a POVM $\{B_{\beta|y}\}_\beta \subset B(\mathcal{H})$, such that all of Alice's operators commute with all of Bob's,

- a density operator ρ on $\mathcal{H}$

such that

$$p(\alpha\beta|xy) = \mathrm{tr}\big(\rho\, A_{\alpha|x} B_{\beta|y}\big).$$

As alluded to before, *Tsirelson's problem* asked whether the two definitions characterize the same set of correlations [53, 54, 55, 98]. This problem has since been answered in the negative [17].

There is an equivalent way of characterizing the commuting observable model. This version refers only to the observable algebras and not directly to any Hilbert space:

Definition 3.3.3 (Two-party quantum correlations, commuting operator model: algebraic formulation)**.** A set $p(\alpha\beta|xy)$ of conditional probabilities is a bipartite quantum distribution *in the commuting observable model* if the following holds. There are

- a C^*-algebra $\mathcal{D}$ (of global observables),

- two mutually commuting C^*-subalgebras $\mathcal{A}, \mathcal{B} \subset \mathcal{D}$ (the observables measurable by the respective parties),

- for each of Alice's settings x a POVM $\{A_{\alpha|x}\}_\alpha \subset \mathcal{A}$, and for each of Bob's settings y a POVM $\{B_{\beta|y}\}_\beta \subset \mathcal{B}$,

- a state ρ on $\mathcal{D}$

such that

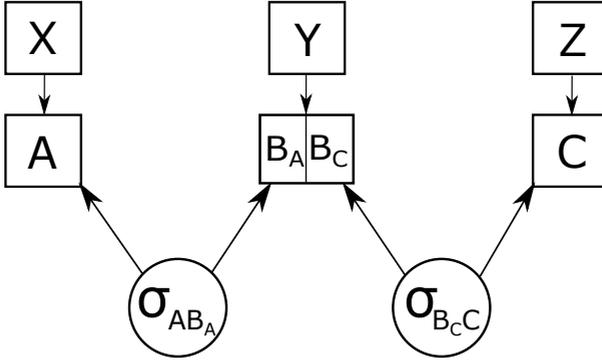$$p(\alpha\beta|xy) = \rho(A_{\alpha|x} B_{\beta|y}).$$

Figure 3.7: The bilocal scenario. Alice and Bob share a bipartite quantum state σ_{AB_A} and Bob and Charlie share a bipartite quantum state $\sigma_{B_C C}$. Alice performs a measurement with the POVM $\{A_{\alpha|x}\}_\alpha$ based on the setting measurement setting x. Bob and Charlie perform a similar measurement. The conditional probabilities $p(\alpha\beta\gamma|xyz)$ that can arise in this way are called bilocal correlations.

Proving the equivalence between Def. 3.3.2 and Def. 3.3.3 amounts to an application of the GNS construction [43].

Bilocal correlations

When modeling locality using tensor products of Hilbert spaces, the set of bilocal correlations is defined as follows.

Definition 3.3.4 (Tensor product model)**.** A set $p(\alpha\beta\gamma|xyz)$ of conditional probabilities is a bilocal quantum distribution *in the tensor product model* if the following holds. There are

- Hilbert spaces $\mathcal{H}_A, \mathcal{H}_{B_A}, \mathcal{H}_{B_C}, \mathcal{H}_C$,

- for each of the settings x, y, z POVMs

$$\{A_{\alpha|x}\}_\alpha \subset B(\mathcal{H}_A),$$
$$\{B_{\beta|y}\}_\beta \subset B(\mathcal{H}_{B_A})\bar{\otimes}B(\mathcal{H}_{B_C}),$$
$$\{C_{\gamma|z}\}_\gamma \subset B(\mathcal{H}_C),$$

- density operators σ_{AB_A} on $\mathcal{H}_A \otimes \mathcal{H}_{B_A}$ and $\sigma_{B_C C}$ on $\mathcal{H}_{B_C} \otimes \mathcal{H}_C$,

such that

$$p(\alpha\beta\gamma|xyz) = \text{tr}\left((\sigma_{AB_A} \otimes \sigma_{B_C C})(A_{\alpha|x} \otimes B_{\beta|y} \otimes C_{\gamma|z})\right).$$

In the commuting observables-model, bilocality takes on the following form:

Definition 3.3.5 (Commuting observables model)**.** A set $p(\alpha\beta\gamma|xyz)$ of conditional probabilities is said to be a *bilocal quantum distribution in the commuting observables model* if the following holds. There are

- a C^*-algebra $\mathcal{D}$,

- mutually commuting C^*-subalgebras $\mathcal{A}, \mathcal{B}_A, \mathcal{B}_C, \mathcal{C} \subset \mathcal{D}$,

- for each of the settings x, y, z POVMs

$$\{A_{\alpha|x}\}_\alpha \subset \mathcal{A},$$
$$\{B_{\beta|y}\}_\beta \subset \mathcal{B}_A \cdot \mathcal{B}_C,$$
$$\{C_{\gamma|z}\}_\gamma \subset \mathcal{C},$$

  with $\mathcal{B} := \mathcal{B}_A \cdot \mathcal{B}_C$ the subalgebra of $\mathcal{D}$ generated by $\mathcal{B}_A$ and $\mathcal{B}_C$,

- a state ρ on $\mathcal{D}$ that acts as a product state in the sense

$$\rho(ab_A b_C c) = \rho(ab_A)\rho(b_C c) \tag{3.53}$$

  for all $a \in \mathcal{A}, b_A \in \mathcal{B}_A, b_C \in \mathcal{B}_C, c \in \mathcal{C}$,

such that

$$p(\alpha\beta\gamma|xyz) = \rho(A_{\alpha|x} B_{\beta|y} C_{\gamma|z}).$$

3.3.3 Equivalent characterizations of bilocal quantum correlations

We will give three further characterizations of the set of bilocal quantum correlations in the commuting operator model. Some of these equivalences are integral to our completeness proof – but they might also be of independent interest.

All statements made here are corollaries of the technical Theorem 3.3.10 proven in Sec. 3.3.5.

A reduced factorization condition

To motivate the first reformulation in Corollary 3.3.6 below, let us try to see which part of Def. 3.3.5 might be the most difficult to work with algorithmically. In our assessment, this is the "hidden factorization condition" of Eq. (3.53). It is "hidden" in the sense that it involves product operators $b_A b_C, b_A \in \mathcal{B}_A, b_C \in \mathcal{B}_C$ that *need not lie in the algebra generated by the POVMs* (cf. Example 3.1.2). But it is properties of precisely this algebra that methods building on the non-commutative polynomial optimization (NPO) hierarchy [40], used in e.g. the quantum inflation method of [22], typically optimize over. As argued in Challenge 3 of Sec. 3.1.1, this poses a barrier against proving completeness for such methods, including the original quantum inflation scheme.

In Sec. 3.2 we circumvented this problem by explicitly adding generators for the algebras $\mathcal{B}_A, \mathcal{B}_C$ to the input of the NPO hierarchy, and expressing the POVM elements as finite-rank superpositions of those. The price to pay for this workaround consists of additional computational costs, as well as the necessity to upper-bound this "Schmidt rank" of the POVM elements.

The following corollary shows that in the special case of the bilocal scenario, these difficulties can fortunately be avoided. Indeed, the weaker factorization condition (3.54), involving only operators generated by the measured POVMs, suffices to imply the *a priori* more general (3.53). We will refer to the weaker constraints as the *reduced model*.

Corollary 3.3.6 (Reduced model)**.** *A set $p(\alpha\beta\gamma|xyz)$ of conditional probabilities is bilocal in the commuting observables model of Def. 3.3.5 if and only if there are*

- *a C^*-algebra $\mathcal{D}$,*

- *mutually commuting C^*-subalgebras $\mathcal{A}, \mathcal{B}, \mathcal{C} \subset \mathcal{D}$,*

- *for each of the settings x, y, z POVMs*

$$\{A_{\alpha|x}\}_\alpha \subset \mathcal{A},$$
$$\{B_{\beta|y}\}_\beta \subset \mathcal{B},$$
$$\{C_{\gamma|z}\}_\gamma \subset \mathcal{C},$$

- *a state ρ on $\mathcal{D}$ that acts as a product state in the sense*

$$\rho(ac) = \rho(a)\rho(c) \tag{3.54}$$

for all $a \in \mathcal{A}, c \in \mathcal{C}$,

such that

$$p(\alpha\beta\gamma|xyz) = \rho(A_{\alpha|x}\, B_{\beta|y}\, C_{\gamma|z}).$$

We note that the reduced model arises implicitly from the *factorisation bilocal NPA hierarchy* of [83].

Bilocal Tsirelson's Problem

The second reformulation, specified in Corollary 3.3.7 below, clarifies the differences between the Hilbert space tensor product model and the commuting operator model of bilocal correlations.

The two approaches are obviously different: To see this, one can simply embed the two-party scenario into the bilocal one, e.g. by taking the A system or C system to be trivial.

It could be surmised that there are "genuine bilocal differences" between the two approaches, and that the bilocal scenario could teach us about Tsirelson's Problem in a way that goes beyond the two-party case. We will, however, show that this is not the case. More precisely, consider the *mixed model* formalized below, where the bipartition $AB_A|B_C C$ is described by a Hilbert space tensor product, while all we can say about the bipartitions $A|B_A$ and $B_C|C$ is that they are associated with commuting observable algebras.

Corollary 3.3.7 (Mixed model)**.** *A set $p(\alpha\beta\gamma|xyz)$ of conditional probabilities is bilocal in the commuting observables model of Def. 3.3.5 if and only if there are*

- *Hilbert spaces $\mathcal{H}_{AB_A}, \mathcal{H}_{B_C C}$,*

- *mutually commuting C^*-algebras*

$$\mathcal{A}, \mathcal{B}_A \subset B(\mathcal{H}_{AB_A}), \qquad \mathcal{B}_C, \mathcal{C} \subset B(\mathcal{H}_{B_C C}),$$

- *for each of the settings x, y, z POVMs*

$$\{A_{\alpha|x}\}_\alpha \subset \mathcal{A},$$
$$\{B_{\beta|y}\}_\beta \subset \mathcal{B}_A \bar{\otimes} \mathcal{B}_C,$$
$$\{C_{\gamma|z}\}_\gamma \subset \mathcal{C},$$

- *density operators σ_{AB_A} on $\mathcal{H}_{AB_A}$ and $\sigma_{B_C C}$ on $\mathcal{H}_{B_C C}$,*

such that

$$p(\alpha\beta\gamma|xyz) = \mathrm{tr}\left((\sigma_{AB_A} \otimes \sigma_{B_C C})\,(A_{\alpha|x}\, B_{\beta|y}\, C_{\gamma|z})\right),$$

where all operators act on $B(\mathcal{H}_{AB_A} \otimes \mathcal{H}_{B_C C})$ in the natural way.

The Renou-Xu formulation

Finally, we consider the formulation used in Ref. [83]. It could be described as the bilocal analogue of Def. 3.3.2, in the sense that it formalizes a commuting observables-model while avoiding to explicitly introduce the local observable algebras. While in the two-party case, the equivalence of Def. 3.3.2 and Def. 3.3.3 was a direct consequence of the GNS construction, the relation between the Renou-Xu model and commuting operator models defined above may not be as obvious. However, we will show:

Corollary 3.3.8 (Renou-Xu model [83]). *A set $p(\alpha\beta\gamma|xyz)$ of conditional probabilities is bilocal in the commuting observables model of Def. 3.3.5 if and only if there are*

- *a Hilbert space $\mathcal{H}$,*

- *commuting projection operators $P, Q \in B(\mathcal{H})$, such that PQ is a normalized rank-one projection,*

- *for each of Alice's settings x, a POVM $\{A_{\alpha|x}\}_\alpha \subset B(\mathcal{H})$, and likewise for Bob and Charlie, such that: (1) operators belonging to different parties commute, and (2)*

$$[A_{\alpha|x}, Q] = [P, C_{\gamma|z}] = 0,$$

such that

$$p(\alpha\beta\gamma|xyz) = \mathrm{tr}\left(PQ\, A_{\alpha|x} B_{\beta|y} C_{\gamma|z}\right).$$

3.3.4 The finite-dimensional case

In the two-party case, the distinction between the tensor product model and the commuting observable model ceases to exist if either can be realized in finite dimensions. Reference [83] asked whether the same is true for the bilocal scenario. Here, we answer this question in the affirmative. In fact, the equivalence already holds when both Alice and Charlie can be associated with a finite-dimensional system.

Corollary 3.3.9. *Assume $p(\alpha\beta\gamma|xyz)$ is compatible with* any *of the models given in Def. 3.3.4, Def. 3.3.5, Cor. 3.3.6, Cor. 3.3.7, Cor. 3.3.8, and is such that the C^*-algebra generated by Alice's and Charlie's POVMs are finite-dimensional.*

Then $p(\alpha\beta\gamma|xyz)$ is compatible with all *these models, and all operator algebras and Hilbert spaces can be chosen to be finite-dimensional.*

3.3.5   Proof of the equivalences

The claimed equivalences derive from the following theorem. We state it in general terms (i.e. not yet specific to the various models of bilocality).

Recall [43, Sec. II.6.4]. that the GNS construction associates with every C^*-algebra $\mathcal{F}$ and state $\sigma \in K(\mathcal{F})$ a triple $(\mathcal{H}, \pi, |\Omega\rangle)$, where $\mathcal{H}$ is a Hilbert space, $\pi : \mathcal{F} \to B(\mathcal{H})$ a $*$-representation, and $|\Omega\rangle \in \mathcal{H}$ a cyclic vector that implements the state in the sense that $\sigma(f) = \langle \Omega | \pi(f) | \Omega \rangle$ for every $f \in \mathcal{F}$.

Theorem 3.3.10. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be mutually commuting C^*-subalgebras of some C^*-algebra $\mathcal{D}$. Let ρ be a state on $\mathcal{D}$ such that*

$$\rho(ac) = \rho(a)\rho(c) \qquad \forall a \in \mathcal{A}, c \in \mathcal{C}. \tag{3.55}$$

Let $(\mathcal{H}_{AB_A}, \pi_A, |\Omega_A\rangle)$ be the GNS representation of $\mathcal{A}$ associated with the state ρ. Let $\mathcal{B}_A$ be the commutant of $\pi_A(\mathcal{A})$ in $B(\mathcal{H}_{AB_A})$. Define $(\mathcal{H}_{B_CC}, \pi_C, |\Omega_C\rangle)$ and $\mathcal{B}_C \subset B(\mathcal{H}_{B_CC})$ analogously.

Then there exists a completely positive unital map

$$\Lambda : \mathcal{B} \to \mathcal{B}_A \bar{\otimes} \mathcal{B}_C \subset B(\mathcal{H}_{AB_A} \otimes \mathcal{H}_{B_CC})$$

such that for all $a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}$

$$\rho(abc) = \text{tr}\left(|\Omega_A\rangle\langle\Omega_A| \otimes |\Omega_C\rangle\langle\Omega_C| \; \pi_A(a)\Lambda(b)\pi_C(c)\right),$$

where all operators act on $B(\mathcal{H}_{AB_A} \otimes \mathcal{H}_{B_CC})$ in the natural way.

The spaces $\mathcal{H}_{AB_A}, \mathcal{H}_{B_CC}$ have previously appeared in the proof of Thm. 3.2 in Ref. [83] (as V_{AB_L}, V_{B_RC}). In fact, this inspired our formulation of Thm. 3.3.10. We go beyond this prior result by showing that they give rise to a tensor product structure on the global Hilbert space.

To prove the theorem, consider in addition the GNS representation $(\pi_\mathcal{D}, \mathcal{H}_\mathcal{D}, |\Omega_\mathcal{D}\rangle)$ of $\mathcal{D}$ associated with ρ.

Lemma 3.3.11. *There is an isometric embedding $V : \mathcal{H}_{AB_A} \otimes \mathcal{H}_{B_CC} \to \mathcal{H}_\mathcal{D}$ which fulfills*

$$V |\Omega_A\rangle \otimes |\Omega_C\rangle = |\Omega_\mathcal{D}\rangle, \tag{3.56}$$
$$V \pi_A(a) \otimes \pi_C(c) = \pi_\mathcal{D}(ac)V. \tag{3.57}$$

Equation (3.57) says that V intertwines $\pi_A \otimes \pi_C$ and $\pi_\mathcal{D}$ as representations of the C^*-algebra generated by $\mathcal{A}$ and $\mathcal{C}$.

Proof. The factorization property (3.55) implies that $|\Omega_A\rangle \otimes |\Omega_C\rangle$ and $|\Omega_\mathcal{D}\rangle$ induce the same state on the C*-algebra generated by $\mathcal{A}$ and $\mathcal{C}$:

$$\langle\Omega_A| \otimes \langle\Omega_C| \, (\pi_A(a) \otimes \pi_C(c)) \, |\Omega_A\rangle \otimes |\Omega_C\rangle$$
$$=\rho(a)\rho(c) = \rho(ac) = \langle\Omega_\mathcal{D}|\pi_\mathcal{D}(ac)|\Omega_\mathcal{D}\rangle.$$

By the uniqueness property of the GNS construction [44, Proposition 4.5.3], there exists a unitary

$$V : \mathcal{H}_{AB_A} \otimes \mathcal{H}_{B_CC} \to \overline{\pi_\mathcal{D}(\mathcal{AC})|\Omega_\mathcal{D}\rangle} =: \mathcal{K}$$

such that

$$V|\Omega_A\rangle \otimes |\Omega_C\rangle = |\Omega_\mathcal{D}\rangle,$$
$$V\pi_A(a) \otimes \pi_C(c)V^* = \pi_\mathcal{D}(ac) \restriction \mathcal{K},$$

where the final symbol denotes the restriction of $\pi_\mathcal{D}$ to $\mathcal{K} = \text{range } V$. The advertised intertwining relation follows by multiplying the last line with V from the right and finally re-interpreting V as a map to all of $\mathcal{H}_\mathcal{D}$. $\square$

Lemma 3.3.12. *It holds that*

$$V^*\pi_\mathcal{D}(\mathcal{B})V \subset \mathcal{B}_A \bar{\otimes} \mathcal{B}_C.$$

Proof. We first claim that

$$V^*\pi_\mathcal{D}(\mathcal{B})V \subset (\pi_A(\mathcal{A}) \otimes_{\text{alg}} \pi_C(\mathcal{C}))'.$$

Indeed, for $a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}$, Eq. (3.57) and its adjoint give

$$[V^*\pi_\mathcal{D}(b)V, \pi_A(a) \otimes \pi_C(c)]$$
$$=V^*\pi_\mathcal{D}(b)V\pi_A(a) \otimes \pi_C(c) - \pi_A(a) \otimes \pi_C(c)V^*\pi_\mathcal{D}(b)V$$
$$=V^*\pi_\mathcal{D}(b)\pi_\mathcal{D}(ac)V - V^*\pi_\mathcal{D}(ac)\pi_\mathcal{D}(b)V$$
$$=V^*[\pi_\mathcal{D}(b), \pi_\mathcal{D}(ac)]V = 0.$$

Now use the fact that the commutator of a set equals the commutator of its weak operator closure [43, I.2.5.3], the Bicommutant Theorem [44, Theorem 5.3.1], and

tensor product
commuting observables $\Big\}$ $\Rightarrow$ reduced
Renou-Xu

$$\underset{\text{Thm. 3.3.10}}{\Longrightarrow} \text{mixed} \Rightarrow \begin{cases} \text{[if } \dim\mathcal{A}, \dim\mathcal{C} < \infty] \text{ tensor product} \\ \text{commuting observables} \\ \text{Renou-Xu} \end{cases}$$
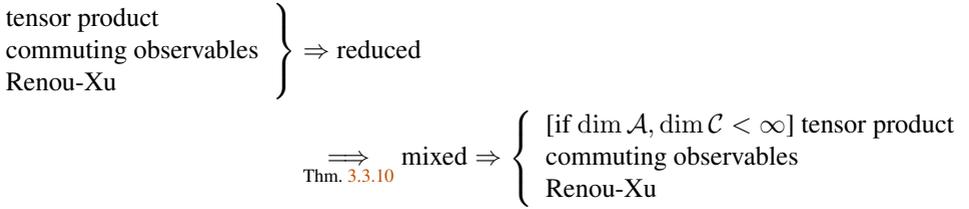
Figure 3.8: Logical structure of the proof given in Sec. 3.3.5. The equivalences claimed in Sec. 3.3.3 follow from this chain of implications among the various models of bilocal quantum correlations.

the Commutation Theorem for von Neumann algebras [43, III.4.5.8] to conclude

$$\begin{aligned}
(\pi_A(\mathcal{A}) \otimes_{\text{alg}} \pi_C(\mathcal{C}))' &= (\pi_A(\mathcal{A})\bar{\otimes}\pi_C(\mathcal{C}))' \\
&= (\pi_A(\mathcal{A})''\bar{\otimes}\pi_C(\mathcal{C})'')' \\
&= (\pi_A(\mathcal{A})'''\bar{\otimes}\pi_C(\mathcal{C})''') \\
&= (\pi_A(\mathcal{A})'\bar{\otimes}\pi_C(\mathcal{C})').
\end{aligned}$$

$\square$

Proof (of Theorem 3.3.10). Set

$$\Lambda : b \mapsto V^*\pi_\mathcal{D}(b)V$$

and compute, using Lemma 3.3.11 repeatedly,

$$\begin{aligned}
&\langle\Omega_A|\langle\Omega_C|\,\pi_A(a)(V^*\pi_\mathcal{D}(b)V)\pi_C(c)|\Omega_A\rangle|\Omega_C\rangle \\
=&\langle\Omega_A|\langle\Omega_C|\,V^*\pi_\mathcal{D}(a)\pi_\mathcal{D}(b)\pi_\mathcal{D}(c)V|\Omega_A\rangle|\Omega_C\rangle \\
=&\langle\Omega_\mathcal{D}|\,\pi_\mathcal{D}(abc)|\Omega_\mathcal{D}\rangle = \rho(abc).
\end{aligned}$$

$\square$

After these preparations, we can now proceed to prove the equivalences claimed to hold in Sec. 3.3.3. The proof's chain of implications among the various models is visualized in Fig. 3.8. The implication "reduced model $\Rightarrow$ Renou-Xu model" also follows from Thm. 3.2 of [83].

Proof (of the equivalences stated in Sec. 3.3.3).
Step 1: We claim that if $p(\alpha\beta\gamma|xyz)$ is compatible with the tensor product model of Def. 3.3.4, the commuting observables model of Def. 3.3.5, or the Renou-Xu model of Cor. 3.3.8, then it is also compatible with the reduced model (Cor. 3.3.6).

This is straightforward to verify, except perhaps for the Renou-Xu model, which we treat explicitly. Indeed, consider a Renou-Xu model realization for $p(\alpha\beta\gamma|xyz)$. Let $\mathcal{A}$ be the C^*-algebra generated by Alice's POVM elements, and likewise for Bob and Charlie. Let $\mathcal{D}$ be the C^*-algebra generated by $\mathcal{A}, \mathcal{B}, \mathcal{C}$. By assumption, there is a normalized vector $|\psi\rangle \in \mathcal{H}$, such that $PQ = |\psi\rangle\langle\psi|$. Let ρ be the associated vector state $\rho(x) = \langle\psi|x|\psi\rangle$ on $\mathcal{D}$. We have now constructed all objects that enter the reduced model. Using the commutation relations of the Renou-Xu model, one verifies the factorization constraint (3.54) for all $a \in \mathcal{A}, c \in \mathcal{C}$:

$$
\begin{aligned}
\rho(ac) &= \operatorname{tr}(PQac) \\
&= \operatorname{tr}(PPQQac) \\
&= \operatorname{tr}(PQaQPc) \\
&= \langle\psi|a|\psi\rangle\langle\psi|c|\psi\rangle \\
&= \rho(a)\rho(c)
\end{aligned}
$$

(a similar calculation appears in the proof of Thm. 3.2 of Ref. [83]).

Step 2 is to show that, by Thm. 3.3.10, the reduced model implies the mixed model. Assume a reduced model description with elements $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \rho, A_{\alpha|x}, \ldots$ is given. They satisfy the assumptions of Thm. 3.3.10, so that we can use the objects whose existence it guarantees in the construction of the mixed model. Indeed, one immediately verifies the properties of the mixed model from the choices

$$
\begin{aligned}
&\mathcal{H}_{AB_A}^{(\mathrm{mix})} = \mathcal{H}_{AB_A}, \quad \mathcal{H}_{B_C C}^{(\mathrm{mix})} = \mathcal{H}_{B_C C}, \\
&\mathcal{A}^{(\mathrm{mix})} = \pi_A(\mathcal{A}), \quad \mathcal{C}^{(\mathrm{mix})} = \pi_C(\mathcal{C}), \\
&\mathcal{B}_A^{(\mathrm{mix})} = \mathcal{B}_A, \quad\quad \mathcal{B}_C^{(\mathrm{mix})} = \mathcal{B}_C, \\
&A_{\alpha|x}^{(\mathrm{mix})} = \pi_A(A_{\alpha|x}), B_{\beta|y}^{(\mathrm{mix})} = \Lambda(B_{\beta|y}), \; C_{\gamma|z}^{(\mathrm{mix})} = \pi_C(C_{\gamma|z}), \\
&\sigma_{AB_A}^{(\mathrm{mix})} = |\Omega_A\rangle\langle\Omega_A|, \sigma_{B_C C}^{(\mathrm{mix})} = |\Omega_C\rangle\langle\Omega_C|.
\end{aligned}
$$

Step 3: The mixed model obviously implies the commuting operator model. It also implies the Renou-Xu model by setting

$$
P = |\Omega_A\rangle\langle\Omega_A| \otimes \mathbb{1}_{\mathcal{B}_C C}, \quad Q = \mathbb{1}_{\mathcal{AB}_A} \otimes |\Omega_C\rangle\langle\Omega_C|
$$

(which is similar to the construction of their operators ρ, σ in the proof of Thm. 3.2

of Ref. [83]).   □

It remains to treat the finite-dimensional case, as advertised in Cor. 3.3.9. The proof combines the construction given in Thm. 3.3.10 with the well-known fact (c.f. Ref. [53, 98]) that for two-party correlations, a finite-dimensional commuting model (as in Def. 3.3.2) implies a tensor product model (as in Def. 3.3.1). Specifically, we will use the following reformulation of Theorem 1 of Ref. [53]:

Lemma 3.3.13. *If $\mathcal{F}, \mathcal{G}$ are mutually commuting C^*-algebras on a finite-dimensional Hilbert space $\mathcal{H}$, then there exist finite-dimensional Hilbert spaces $\mathcal{H}_\mathcal{F}, \mathcal{H}_\mathcal{G}$ and an isometric embedding*

$$W_{\mathcal{FG}} : \mathcal{H} \to \mathcal{H}_\mathcal{F} \otimes \mathcal{H}_\mathcal{G}$$

such that

$$W_{\mathcal{FG}}\, \mathcal{F}\, W_{\mathcal{FG}}^* \subset B(\mathcal{H}_\mathcal{F}) \otimes \mathbb{1},$$
$$W_{\mathcal{FG}}\, \mathcal{G}\, W_{\mathcal{FG}}^* \subset \mathbb{1} \otimes B(\mathcal{H}_\mathcal{G}).$$

In keeping with the notation of C^*-algebras, W^* denotes the adjoint of W, i.e. the operator that would be denoted as $W^\dagger$ in physics notation.

Proof (of the equivalences for finite-dimensional models). By the previous proof, all models imply the mixed model, where specifically $\mathcal{H}_{AB_A}, \mathcal{H}_{B_C C}$ arise from the GNS representation of the observable algebras $\mathcal{A}$ and $\mathcal{C}$ respectively. In particular, the dimensions of these Hilbert spaces are upper-bounded by the dimension of the associated algebras, which in turn can be chosen to be the ones generated by Alice's and Charlie's observables. All operators that enter the construction of the mixed model as laid out in Cor. 3.3.7 are linear maps on the tensor product of these two Hilbert spaces and therefore finite-dimensional. The third step of the previous proof then gives finite-dimensional realizations in the commuting operator model and the Renou-Xu model.

It remains to be shown that $p(\alpha\beta\gamma|xyz)$ can be realized in the Hilbert space tensor product model (Def. 3.3.4), and in particular in one involving only finite-dimensional spaces. Let a finite-dimensional mixed model realization of $p(\alpha\beta\gamma|xyz)$ with elements $\mathcal{H}_{AB_A}, \mathcal{H}_{B_C C}, \mathcal{A}, \mathcal{B}_A, \mathcal{B}_C, \mathcal{B}, A_{\alpha|x}, \sigma_{AB_A}, \ldots$ be given. Our strategy is to apply Lem. 3.3.13 separately to $\mathcal{A}, \mathcal{B}_A$ and to $\mathcal{B}_C, \mathcal{C}$. First, choosing $\mathcal{F} = \mathcal{A}$ and $\mathcal{G} = \mathcal{B}_A$ in Lem. 3.3.13 establishes the existence of two Hilbert spaces $\mathcal{H}_\mathcal{A}, \mathcal{H}_{\mathcal{B}_A}$ and an isometry

$$W_{\mathcal{AB}_A} : \mathcal{H}_{AB_A} \to \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_{\mathcal{B}_A}.$$

These allow us to choose the first set of objects that will enter the tensor product model as

$$\mathcal{H}_A^{(\text{t.p.})} = \mathcal{H}_\mathcal{A}, \qquad\qquad \mathcal{H}_{B_A}^{(\text{t.p.})} = \mathcal{H}_{\mathcal{B}_A},$$

$$A_{\alpha|x}^{(\text{t.p.})} = W_{\mathcal{A}\mathcal{B}_A} A_{\alpha|x} W_{\mathcal{A}\mathcal{B}_A}^*,$$

$$\sigma_{AB_A}^{(\text{t.p.})} = W_{\mathcal{A}\mathcal{B}_A} \sigma_{AB_A} W_{\mathcal{A}\mathcal{B}_A}^*.$$

An analogous procedure starting with $\mathcal{B}_C, \mathcal{C}$ gives

$$\mathcal{H}_{B_C}^{(\text{t.p.})} = \mathcal{H}_{\mathcal{B}_C}, \qquad\qquad \mathcal{H}_C^{(\text{t.p.})} = \mathcal{H}_\mathcal{C},$$

$$C_{\gamma|z}^{(\text{t.p.})} = W_{\mathcal{B}_C\mathcal{C}} C_{\gamma|z} W_{\mathcal{B}_C\mathcal{C}}^*,$$

$$\sigma_{B_CC}^{(\text{t.p.})} = W_{\mathcal{B}_C\mathcal{C}} \sigma_{B_CC} W_{\mathcal{B}_C\mathcal{C}}^*.$$

Finally, with

$$B_{\beta|y}^{(\text{t.p.})} = (W_{\mathcal{A}\mathcal{B}_A} \otimes W_{\mathcal{B}_C\mathcal{C}}) B_{\beta|y} (W_{\mathcal{A}\mathcal{B}_A} \otimes W_{\mathcal{B}_C\mathcal{C}})^*$$

it is straight-forward to verify the properties of the tensor product model. $\qquad\square$

It is apparent from the proof that the condition in Cor. 3.3.9 can be slightly weakened. Instead of demanding that the algebras $\mathcal{A}, \mathcal{C}$ be finite-dimensional, it is sufficient for the conclusions to hold that the GNS Hilbert space $\mathcal{H}_{AB_A} \otimes \mathcal{H}_{B_CC}$ associated with the restriction of the state to $\mathcal{A}\mathcal{C}$ is finite-dimensional.

3.3.6 Complete SDP hierarchies

In this section, we will construct complete hierarchies of relaxations for the reduced model defined in Cor. 3.3.6. Most ingredients for this construction and the completeness proof have been developed in earlier sections (based on Refs. [21, 22, 24, 40]), to which we will refer for technical details.

Let $\mathcal{D}$ be the *universal C^*-algebra* (cf. Sec. 1.1.1 and [43, Sec. II.8.3]) with generators

$$\mathcal{G} = \{\mathbb{1}, A_{\alpha|x}, B_{\beta|y}, C_{\gamma|z}\} \tag{3.58}$$

and relations

$$[A_{\alpha|x}, B_{\beta|y}] = 0, \qquad\qquad\qquad \forall \alpha, \beta, x, y, \qquad (3.59)$$

$$[B_{\beta|y}, C_{\gamma|z}] = 0, \qquad\qquad\qquad \forall \beta, \gamma, y, z, \qquad (3.60)$$

$$[C_{\gamma|z}, A_{\alpha|x}] = 0, \qquad\qquad\qquad \forall \alpha, \gamma, x, z, \qquad (3.61)$$

$$[\mathbb{1}, X] = 0 \qquad\qquad\qquad\qquad \forall X \in \mathcal{G}, \qquad (3.62)$$

$$\mathbb{1}X = X\mathbb{1} = X \qquad\qquad\qquad \forall X \in \mathcal{G}, \qquad (3.63)$$

$$X^* = X \succeq 0, \qquad\qquad\qquad \forall X \in \mathcal{G}, \qquad (3.64)$$

$$\sum_\alpha A_{\alpha|x} = \sum_\beta B_{\beta|y} = \sum_\gamma C_{\gamma|z} = \mathbb{1} \qquad\qquad \forall x, y, z. \qquad (3.65)$$

In a precise sense $\mathcal{D}$ is the direct sum of all possible realizations of this algebra as operators on Hilbert spaces. Let $K(\mathcal{D})$ be the set of all states on $\mathcal{D}$.

We aim to solve the following optimization problem:

$$f^* = \min_{\rho \in K(\mathcal{D})} \sum_{\alpha, \beta, \gamma, x, y, z} \big(\rho(A_{\alpha|x} B_{\beta|y} C_{\gamma|z}) - p(\alpha\beta\gamma|xyz) \big)^2 \qquad (3.66)$$
$$\text{s. t.} \quad \rho(ac) - \rho(a)\rho(c) = 0 \qquad \forall a \in \mathcal{A}, \ c \in \mathcal{C}$$

The objective function of this problem represents the minimal 2-norm distance between a quantum realization of the reduced model of the bilocal scenario and the observed statistics. We accept that the correlations can arise from a reduced model if $f^* = 0$ (or at most some small ε that represents numerical and statistical tolerances). By Theorem 3.3.10 this means that such correlations can also arise in the mixed model, the commuting observables model and the Renou-Xu model of the bilocal scenario. If $f^* > 0$, the correlations cannot have been produced in any of the models, including the tensor product one.

The problem (3.66) is "polynomial" in two different ways: The operators $A_{\alpha|x} B_{\beta|y} C_{\gamma|z}$ and ac are (norm limits of) non-commutative polynomials in the generators, while the objective function and the constraints are second order polynomials in the state.

We will again use two different techniques to deal with these non-linearities:

1. Non-commutative polynomial optimization (NPO) [40] provides a hierarchy of SDP relaxations for optimizing over *linear* functions on states of the universal algebra $\mathcal{D}$, subject to linear constraints. Its completeness follows from the GNS construction.

2. By passing to their *polarizations*, one can interpret the polynomial functions on $K(\mathcal{D})$ as linear functions on symmetric product states on multiple copies

of $\mathcal{D}$. Such states are constructed from symmetric extensions of states on $\mathcal{D}$ and completeness follows from a suitable quantum de Finetti theorem from Chapter 2.

In Sec. 3.3.6 below, we lay out how to use the results of the earlier sections to construct a converging hierarchy of SDP relaxations for polynomial optimization problems over algebras given in terms of generators and relations. While we focus on the bilocal scenario, the techniques can be fairly straightforwardly adapted to general algebras and polynomials, as we will see in Chapter 5. We call this approach the *polarization hierarchy*. In Sec. 3.3.6, we describe a slightly different approach more closely related to quantum inflation [22], which we also prove to be complete for the bilocal scenario.

Polarization hierarchy

To define the polarization hierarchy, choose a *level* $n \in \mathbb{N}$ and consider n copies of the generators:

$$\mathcal{G}^n = \{\mathbb{1}, A^{(i)}_{\alpha|x}, B^{(i)}_{\beta|y}, C^{(i)}_{\gamma|z}\} \qquad i \in 1, \dots n. \tag{3.67}$$

Relations analogous to those in Eqs. (3.59)-(3.65) are imposed for each i, together with relations stating that operators for different values of the superscript i commute. The resulting universal C^*-algebra $\mathcal{D}^n$ is the "largest C^*-algebra generated from n commuting copies of $\mathcal{D}$", or, more precisely, the *maximal C^*-tensor product* $\mathcal{D}^n = \mathcal{D}^{\otimes_{\max} n}$ [1, 52].

We note that this algebra is closely related to the algebra that is constructed for the most general version of the quantum inflation technique [22]. This technique works with an even larger algebra, where e.g. Bob's operators carry two indices $B^{(i,j)}_{\beta|y}$ that can be varied independently. It will turn out that our simpler model is sufficient for the bilocal scenario.

On the n-th tensor product of $\mathcal{D}$, we can linearize n-th order polynomial functions on $K(\mathcal{D})$ by passing to their polarization as follows: With every state $\sigma \in K(\mathcal{D})$ associate its n-fold symmetric product state $\Pi^n_\sigma \in K(\mathcal{D}^n)$ which is defined by its action on product operators in the obvious way:

$$\Pi^n_\sigma(x_1 \otimes \cdots \otimes x_n) = \sigma(x_1) \dots \sigma(x_n)$$

and extended to all of $\mathcal{D}^n$ by linearity and continuity. Then

$$
\sum_{\alpha,\beta,\gamma,x,y,z} \left(\sigma(A^{(1)}_{\alpha|x} B^{(1)}_{\beta|y} C^{(1)}_{\gamma|z}) - p(\alpha\beta\gamma|xyz) \right)^2
$$
$$
= \Pi_\sigma^2 \Big(\sum_{\alpha,\beta,\gamma,x,y,z} A^{(1)}_{\alpha|x} B^{(1)}_{\beta|y} C^{(1)}_{\gamma|z} A^{(2)}_{\alpha|x} B^{(2)}_{\beta|y} C^{(2)}_{\gamma|z}
$$
$$
- 2p(\alpha\beta\gamma|xyz) A^{(1)}_{\alpha|x} B^{(1)}_{\beta|y} C^{(1)}_{\gamma|z} + p(\alpha\beta\gamma|xyz)^2 \mathbb{1} \Big)
$$
$$
=: \Pi_\sigma^2(y_0),
$$

where y_0 is the element of $\mathcal{D}^2$ on which Π_σ^2 is evaluated.

Similarly, one can turn the independence constraint of (3.66) into a linear constraint on two inflation levels. However, it will turn out that for the completeness proof, it is necessary to impose constraints that are bounded from below and attain their minimal value on the feasible set of states. We will thus formulate the factorization constraints as

$$
(\sigma(ac) - \sigma(a)\sigma(c))^2 = 0,
$$

so that the polarization becomes

$$
\Pi_\sigma^4(y_{ac}) = 0,
$$

where

$$
y_{ac} := a^{(1)} c^{(1)} a^{(2)} c^{(2)} - 2a^{(1)} c^{(1)} a^{(2)} c^{(3)} + a^{(1)} c^{(2)} a^{(3)} c^{(4)}.
$$

Here, the indices indicate which copies of the POVM elements are used to generate the operator, e.g. $a^{(2)}$ can be written as (the norm limit of) a polynomial in the generators $\{\mathbb{1}, A^{(2)}_{\alpha|x}\}$. In this way, both the polynomial objective function and the polynomial constraints correspond to the linear pairing between operators $y_0, y_{ac} \in \mathcal{D}^4$ and symmetric product states in $K(\mathcal{D}^4)$.

More generally, given a degree m polynomial q whose action on states is bounded from below by 0, one can optimize over polynomial constraints of the form

$$
q(\sigma) = 0,
$$

by passing to the polarization $y_q \in \mathcal{D}^m$ of q.

Unfortunately, the set of symmetric product states is not an affine subset of state space, which means that the NPO method cannot directly optimize over it. To get around this restriction, we will combine three tricks. First, realize that NPO can

optimize over the set of *all* symmetric states. Indeed, the symmetric group S_n acts on $\mathcal{D}^n$ by permuting the indices of the generators, and a state $\rho \in K(\mathcal{D}^n)$ is *symmetric* if it satisfies the *linear* constraints $\rho(\pi(x)) = \rho(x)$ for every $x \in \mathcal{D}^n, \pi \in S_n$. Second, in both quantum and classical probability [1, 23, 24, 75, 79, 99], there is a well-known family of statements collectively known as *de Finetti theorems* that show that symmetric states on infinitely many copies are a convex combination of symmetric product states. In our particular case, "infinitely many copies" can be made rigorous as the *inductive limit of maximal C^*-tensor products*. The following de Finetti theorem, adapted to this setting, is proven in Sec. 2.1 and repeated here for easy reference.

Theorem 3.3.14 (Max tensor product Quantum de Finetti Theorem). *Let $\rho \in K(\mathcal{D}^\infty)$ be a symmetric state on an infinite maximal tensor product*

$$\mathcal{D}^\infty = \lim_{n \to \infty} \mathcal{D}^{\otimes_{\max} n}.$$

Then there exists a unique probability measure μ over states on $\mathcal{D}$ such that for all $x \in \mathcal{D}^\infty$,

$$\rho(x) = \int_{K(\mathcal{D})} \Pi_\sigma^\infty(x) \, \mathrm{d}\mu(\sigma), \tag{3.68}$$

where Π_σ^∞ is the infinite symmetric product state on $\mathcal{D}^\infty$ associated with the state σ on $\mathcal{D}$.

The third trick is to choose the polynomial constraints in such a way that they demand that point-wise non-negative polynomials are set to 0. If such an extremal condition is satisfied by a (continuous, as in Eq. (3.68)) convex combination, then it must in fact be satisfied almost surely. Applying this to the constraints and the objective function, we will see that in our case the Π_σ^∞ are almost surely a feasible solution of (3.66) that attains the minimum $f^\infty = \lim_{n \to \infty} f^n$ of the relaxation (3.69) below.

Let us now formulate the NPO hierarchy and its convergence proof more precisely. Define $\mathcal{A}^n$ to be the subalgebra of $\mathcal{D}^n$ that consists of Alice's operators and similar for Bob and Charlie. Let U be a countable basis of $\mathcal{A}^n$. Usually this basis is taken to be the set of all words in Alice's POVM elements. Define V for Bob and W for Charlie in a similar way. For $n \geq 4$ the hierarchy of NPO problems is then given

by

$$f^n = \min_{\rho \in K(\mathcal{D}^n)} \rho(y_0)$$
$$\text{s. t.} \quad \rho(\pi(abc)) = \rho(abc),$$
$$\rho(y_{ac}) = 0,$$
$$\forall \pi \in S_n, a \in U, b \in V, c \in W.$$
$$(3.69)$$

Each of these NPO problems can in turn be solved via the complete hierarchy of SDP relaxations (1.27) introduced in Ref. [40, 56], where we have used the formulation of NPO problems from Sec. 1.3.2.

The following theorem then states that (3.66) and (3.69) are equivalent in the limit.

Theorem 3.3.15. *Let $f^\infty = \lim_{n \to \infty} f^n$. It holds that $f^\infty = f^*$.*

Proof. The proof is very similar to that of Theorem 3.2.1.

It is clear that

$$f^n \leq f^* \qquad \forall n, \tag{3.70}$$

since each level of the hierarchy (3.69) is a relaxation of the optimization problem (3.66).

For the converse direction, use NPO to construct a state ω_n on $\mathcal{D}^\infty$ for each level n of the hierarchy by taking the infinite tensor product of an optimizing state of the optimization problem (3.69) at level n.

By the Banach-Alaoglu theorem applied to the state space $K(\mathcal{D}^\infty)$, this sequence admits a weak*-convergent subsequence. Let ω be its limit point. Since each ω_n obeys the constraints of Eq. (3.69), so does ω. Hence, ω is a symmetric state on the algebra $\mathcal{D}^\infty$ and Theorem 3.3.14 applies. That is, ω can be written as

$$\omega = \int d\mu(\sigma) \, \Pi_\sigma^\infty, \tag{3.71}$$

with μ a unique probability measure over states $\sigma \in K(\mathcal{D})$ and Π_σ^∞ an infinite product state on $\mathcal{D}^\infty$.

By construction, each of the y_{ac} is non-negative on the product states Π_σ^∞. Therefore, since $\omega(y_{ac}) = 0$, and μ is a probability measure, it holds that

$$\Pi_\sigma^\infty(y_{ac}) = 0 \qquad \text{almost everywhere w.r.t. } \mu.$$

That is, there exists a full measure subset $E \subset K(\mathcal{D})$ such that for all $\sigma \in E$, it holds that $\Pi_\sigma^\infty(y_{ac}) = 0$.
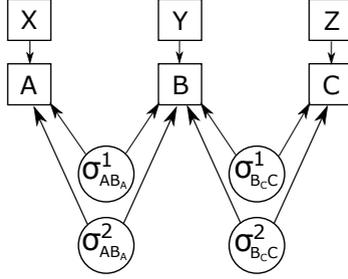
Figure 3.9: The level 2 inflation of the bilocal scenario. Each of the states σ_{AB_A} and $\sigma_{B_C C}$ has been copied. The total state of the system is permutation symmetric under the exchange of each of these copies. The inflation technique builds on this observation.

Hence, each Π_σ^∞ with $\sigma \in E$ defines a feasible state σ for the optimization problem (3.66) by restricting to one copy of the algebra $\mathcal{D}$. From this one can conclude that $\Pi_\sigma^\infty(y_0) \geq f^\infty$ for all $\sigma \in E$, for otherwise one could have taken ω to be the point measure on a state σ' such that $\Pi_{\sigma'}(y_0) < f^\infty$. This would contradict the fact that f^∞ is a minimum.

Combining this with the fact that $\omega(y_0) = f^\infty$, it must hold that

$$\Pi_\sigma^\infty(y_0) = f^\infty \qquad \text{almost everywhere w.r.t. } \mu \text{ on } E.$$

I.e., there exists a set $F \subset E$ with full measure, such that for all $\sigma \in F$ it holds that $\Pi_\sigma^\infty(y_0) = f^\infty$. Finally, we can conclude for any $\sigma \in F$

$$f^\infty = \Pi_\sigma^\infty(y_0) \geq f^*. \tag{3.72}$$

Combining Eqs. (3.70) and (3.72) yields $f^\infty = f^*$, proving the theorem. $\qquad\square$

Inflation hierarchy

There exists a second convergent hierarchy that is more closely related to the quantum inflation hierarchy of Ref. [22]. By showing convergence of such a hierarchy, we answer a question posed by Renou and Xu in Ref. [83]. The hierarchy is very similar to that of Eq. (3.69), but instead of treating the independence constraints as polynomial conditions, they are enforced by imposing additional symmetries. Loosely speaking, these new symmetry constraints posit that copies of the state σ_{AB_A} can be permuted independently of the copies of $\sigma_{B_C C}$, see Fig. 3.9 for a visualization.

The advantage of this hierarchy over the polarization hierarchy is that the symmetry constraints can already be imposed at level 2 of the hierarchy.

In the notation introduced above Eq. (3.69), the level n relaxation is given by

$$
\begin{aligned}
\tilde{f}^n = \min_{\rho \in K(\mathcal{D}^n)} \quad & \rho(y_0) \\
\text{s. t.} \quad & \rho(\pi(abc)) = \rho(abc), \quad \forall \pi \in S_n \\
& \rho(a\pi(c)) = \rho(ac), \quad \forall \pi \in S_n.
\end{aligned}
\tag{3.73}
$$

Theorem 3.3.16. *Let* $\tilde{f}^\infty = \lim_{n\to\infty} \tilde{f}^n$. *It holds that* $\tilde{f}^\infty = f^*$.

Proof. We only give a short proof sketch, since the techniques are nearly identical to the proof of Theorem 3.3.15.

Construct a state $\omega \in K(\mathcal{D}^\infty)$ as the limit of optimizing states of (3.73) (c.f. the proof of Theorem 3.3.15). By the de Finetti theorem this state has the form

$$
\omega = \int d\mu(\sigma)\Pi_\sigma^\infty.
\tag{3.74}
$$

Fix one $n \in \mathbb{N}$. Using the cycle notation, define the permutation

$$
\pi = (1, n+1)\,(2, n+2)\,\ldots\,(n, 2n),
$$

i.e. π exchanges the 1st block of n symbols with the 2nd block of n symbols. Using the additional symmetry constraints of ω, when restricted to elements of Alice and Charlie, we see that for each $a \in \mathcal{A}^n$ and $c \in \mathcal{C}^n$, and for all n

$$
\omega(ac) = \omega(a\pi(c))
\tag{3.75}
$$

$$
= \int d\mu(\sigma)\,\Pi_\sigma^\infty(a\pi(c))
\tag{3.76}
$$

$$
= \int d\mu(\sigma)\,\Pi_\sigma^\infty(a)\Pi_\sigma^\infty(\pi(c))
\tag{3.77}
$$

$$
= \int d\mu(\sigma)\,\Pi_\sigma^\infty(a)\Pi_\sigma^\infty(c),
\tag{3.78}
$$

where the symmetry of ω was used in (3.75), Eq. (3.74) was used for Eq. (3.76), and in Eqs. (3.77) and (3.78) it was used that each Π_σ^∞ is a symmetric product state over disjoint inflation levels. From this we can see that Π_σ^∞ obeys the factorization constraint almost surely with respect to μ.

The rest of the proof is now similar to that of Theorem 3.3.15. $\square$

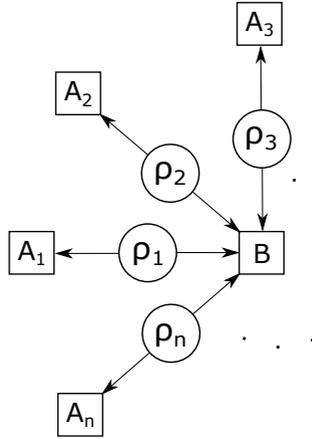We note that this result also proves convergence of the "full" quantum inflation

Figure 3.10: A star network with $n+1$ parties. The bilocal scenario is a star network with $n = 2$, where Alice is A_1 and Charlie is A_2. The settings for each of the parties have been left out in this figure.

hierarchy [22] where Bob's POVMs have two separate indices: For each n, the NPO problem that describes such a full inflation level is a relaxation of problem (3.66) that is at least as restrictive as the relaxation (3.69). Hence, its optimal value lies between f^n and f^* for every n.

3.3.7 Generalization to other causal structures

One can ask whether it is possible to use the technique of Theorem 3.3.10 to show that the quantum inflation hierarchy converges for other networks. It is not difficult to show that Theorem 3.3.10 can be adapted to the more general case of *star networks*, in which one central party shares a bipartite quantum state with n other parties, but no other connections are present. Note that the bilocal scenario is a star network with $n = 2$, where Bob acts as the central party.

Consider the star network with $n+1$ parties $A_1, \ldots, A_n, B$ depicted in Fig. 3.10. In this figure the settings for each party have been left out for simplicity, but can be added back in if necessary. The following corollary of Theorem 3.3.10 holds.

Corollary 3.3.17. *The polarization and inflation hierarchies are complete for the causal compatibility problem for star networks.*

Proof. The proof relies on an extension of Theorem 3.3.10. In a star network, we
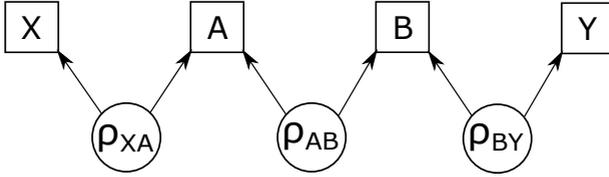
Figure 3.11: A line network of four parties without settings. The Bell scenario can be simulated by this scenario by regarding X as the setting for A and Y as the setting for B.

have a state ρ such that

$$\rho(a_1 a_2 \ldots a_n) = \rho(a_1)\rho(a_2)\ldots\rho(a_n). \tag{3.79}$$

Using this property, we can repeat the construction that leads to the proof of Theorem 3.3.10. Applying the GNS construction on the state of Eq. (3.79) yields a tensor product of n Hilbert spaces $\mathcal{H}_{A_i} \otimes \ldots \otimes \mathcal{H}_{A_n}$. Defining $\mathcal{B}_{A_i} = \pi_{A_i}(\mathcal{A}_i)'$, one can show that there is a channel $\Lambda : \mathcal{B} \to \mathcal{B}_{A_1}\bar{\otimes}\mathcal{B}_{A_2}\bar{\otimes}\ldots\bar{\otimes}\mathcal{B}_n$ that preserves the correlations, yielding a mixed model for the star network.

In this way, the causal compatibility problem becomes an instance of a state polynomial optimization problem. The claim for the polarization and scalar extension hierarchy then follows from Theorem 5.1.1, which is proven in chapter 5. For the quantum inflation hierarchy, the claim follows from Theorem 3.2.1. □

Remark. In Chapter 5 we will also treat a hierarchy of SDPs known as scalar extension in the quantum information community and as state polynomial optimization in the convex optimization community. It was shown in Ref. [39] that scalar extension is also convergent for optimization problems that are polynomials in the operators and the states. Therefore, Theorem 3.3.10 shows that it can also be used for the causal compatibility problem in the bilocal scenario. Additionally, using the same reasoning as in Corollary 3.3.17, this can be extended to the case of star networks as well.

The bilocal scenario is also a *line network*, in which the parties are arranged in a line and share a bipartite quantum state with each of their neighbors. It therefore might seem plausible that the technique can also be extended to larger line networks. This, however, turns out to not be the case. The following counter example is due to Elie Wolfe.

Example 3.3.18. Consider a line network of four parties, X, A, B, Y, as depicted

in Fig. 3.11. From the causal structure, we can deduce the following factorization constraints for all $X \in \mathcal{X}$, $A \in \mathcal{A}$, $B \in \mathcal{B}$ and $Y \in \mathcal{Y}$:

$$\rho(XAY) = \rho(XA)\rho(Y) \tag{3.80}$$
$$\rho(XBY) = \rho(X)\rho(BY). \tag{3.81}$$

These factorization conditions can be enforced in the polarization, inflation or scalar extension hierarchy. Since there are no settings in this example, we can w.l.o.g. take our POVMs to be commuting orthogonal projective measurements, which means that the conditions (3.80) and (3.81) are equivalent to

$$p(XAY) = p(XA)p(Y) \tag{3.82}$$
$$p(XBY) = p(X)p(BY). \tag{3.83}$$

However, these conditions are not enough to ensure compatibility with the four party line scenario. This can be seen by the following argument.

By letting ρ_{XA} send perfectly correlated classical bits to X and A, it is possible to simulate a network where X is a setting for A. By a similar procedure, Y can be interpreted as a setting for B. This effectively creates the Bell scenario. As we know, in the two parties, two outcomes, two settings Bell scenario, the CHSH inequality (Eq. (1)) can be violated up to $2\sqrt{2}$. However, if we only assume the factorization conditions (3.82) and (3.83), we cannot exclude correlations created by a PR box, which only requires non-signaling constraints and can reach a violation of the CSHS inequality up to 4. Indeed, consider the distribution of the PR box

$$p_{\mathrm{PR}}(x, a, b, y) = \begin{cases} \frac{1}{8} & \text{if } a \oplus b = x \cdot y, \\ 0 & \text{otherwise.} \end{cases} \tag{3.84}$$

The marginal distribution of p_{PR} on the subsets XAY and XBY leads to a uniform distribution, so that (3.82) and (3.83) hold. From this we must conclude that the factorization conditions alone are not enough to capture the correlations in general line scenarios.

Remark. We note that example 3.3.18 is again not a counter example to the convergence of the quantum inflation hierarchy. The example only shows that a similar construction as that of the proof of Theorem 3.3.10 will not work in general.

In fact, the quantum inflation hierarchy numerically refutes the PR correlations in the four party line already at inflation level 2, which can quickly be checked using the recently developed inflation toolkit [100].

3.4   Conclusion

Building on the quantum inflation hierarchy of Ref. [22], we have constructed a provably complete semidefinite programming hierarchy for the quantum causal polynomial optimization problem. Along the way, we have used the generalized Quantum de Finetti Theorem for infinite systems of Chapter 2, and the description of the NPO hierarchy as an optimization procedure over states of a universal C^*-algebra of Section 1.3.2.

We have shown the equivalence of several models of locality for the bilocal scenario. In particular, we have shown that a reduced model of bilocality, in which only Alice and Charlie are supposed to be independent, is enough to reproduce exactly the bilocal quantum distributions in the commuting observables model. Furthermore, if Alice's and Charlie's systems can be associated with a finite dimensional algebra, the correlations also coincide with the tensor product model.

Additionally, we have constructed two converging SDP hierarchies for the bilocal scenario, based on the above-mentioned classification. The polarization hierarchy makes use of the fact that certain polynomial expressions in a state can be linearized on tensor powers of that state. Here, this idea was applied to the factorization constraint between Alice and Charlie, but it can be applied to polynomials of higher order as well. The second hierarchy is a form of the quantum inflation hierarchy.

In deriving these results, we have answered two open questions of Ref. [83]:

1. whether the bilocal scenario allows for new insights into Tsirelson's problem: No.

2. and whether the quantum inflation hierarchy is complete for the bilocal scenario: Yes.

A number of follow-up questions suggest themselves.

We could not prove completeness of the original quantum inflation hierarchy for general quantum causal structures, due to the difficulty of constructing the local observable algebras ($\mathcal{A}_-, \mathcal{A}_+$, etc.). To deal with this problem in the general case, we had to manually add generators for the algebra to the NPO program, and then manually impose norm constraints. While we have argued that any constructive completeness proof will have to add elements to the algebra that is extracted from the output of the SDP hierarchy, it is not obvious that we have found the most economical way of handling the issue. We also do not know whether there are a priori finite bounds on the norm of the local operators that combine to give the POVM elements, except for the case of the bilocal scenario. Both questions merit further research.

So far, we have focused mostly on the quantum causal compatibility problem, and have not investigated objective functions beyond the 2-norm distance to mea-

sured data. While the examples given in [22, Section VII] carry over to our formulation, it would be interesting to look into further applications. For this reason, we consider quantum networks in Chapter 4, where we allow the output to be a finite-dimensional quantum state. Additionally, the more general class of optimization problems known as state polynomial optimization (SPO) will be treated in Chapter 5. In both chapters many of the concepts introduced in this chapter will reappear.

Laslty, it would be worth investigating what further properties of the universal observable algebra can be enforced with suitable constraints. For example, one could impose that some of the subalgebras are Abelian in order to model partly classical behavior.

Chapter 4

Network compatibility

Even though the inflation technique is mostly applied in the device-independent setting, where only correlations are considered, there is no theoretical reason to restrict to this case. In this section we will therefore take a closer look at how to apply the inflation and polarization techniques to analyze whether a specific quantum state can be produced in a given network scenario.

4.1 General approach

The descriptions of quantum inflation and NPO as given in sections 3.1 and 1.3.2 are not restricted to testing for compatibility with respect to a probability distribution. Instead, they allow for a wide range of optimization problems in network scenarios, as long as one can express the problem in terms of quantum states on separable C^*-algebras.

To give a simple example, one can analyze the Bell scenario using dichotomic observables, i.e. observables with a ± 1-valued spectrum, as we did in the introduction. The CHSH expression in terms of these quantities takes on the form in which it is most often found in textbooks. For dichotomic operators A_0, A_1, B_0, B_1 such that $[A_i, B_j] = 0$, we want to maximize

$$\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle. \tag{4.1}$$

It is well known [56, 101] that this expression reaches the Tsirelson bound of $2\sqrt{2}$

by choosing

$$A_0 = \sigma_Z \qquad\qquad B_0 = \frac{1}{\sqrt{2}}(\sigma_Z + \sigma_X)$$

$$A_1 = \sigma_X \qquad\qquad B_1 = \frac{1}{\sqrt{2}}(\sigma_Z - \sigma_X)$$

and measuring these observables on the maximally entangled state $|\psi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Using the NPA hierarchy [56], one can retrieve this upper bound already with a moment matrix of size 5×5, and extract the 2-qubit model above with a matrix of size 9×9.

In some sense, this example is still only testing for correlations: each dichotomic operator A is equivalent to a two-outcome measurement $\{P_0, P_1\}$ via $A = +1P_0 + (-1)P_1 = 2P_0 - \mathbb{1} = \mathbb{1} - 2P_1$. To test the compatibility of an arbitrary quantum state ρ, we will in general need a larger set of observables. Suppose one is testing the compatibility of a state in d dimensions. It is then always possible to choose a complete set of observables for that dimension, that is, a set of observables whose expectation values uniquely determine the state (up to local isometries). We will consider two special cases: the set of d-dimensional Pauli operators (which will also prepare us for Sec. 4.2), and the set of matrix units.

Let us first describe the general setup. We will restrict our attention to networks described by bigraphs, which were called correlation scenarios in Sec. 1.2.2 for the case of the causal compatibility problem. Consider a quantum network of N spatially separated parties, with some configuration of S quantum sources, given by a bipartite DAG G in a similar way as for a quantum causal structure. Each party receives a number of quantum systems from the sources, as prescribed by G, is allowed to perform local operations (LO) on these source states, and finally outputs a state of dimension d. The total Hilbert space of the outputs is thus $(\mathbb{C}^d)^{\otimes N}$ and the final state is of the form

$$\rho_{\text{LO}} = \mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_N(\sigma_1 \otimes \ldots \otimes \sigma_S), \tag{4.2}$$

where $\mathcal{E}_i$ is a local quantum channel for party i and σ_j is the state that the source j distributes. Even though the dimensions of the output states are restricted, those of the input states are not. If H is a subgraph of G, we label the reduced output state on the parties in H by ρ_H.

Alternatively, one can consider the so-called LOSR setup, in which the parties can perform local operations (LO) and possess pre-shared randomness (SR). Without loss of generality the shared randomness can be absorbed into the channels [102]. In

this case, the output state has the more general form

$$\rho_{\text{LOSR}} = \sum_{\lambda \in \Lambda} p_\lambda \mathcal{E}_1^\lambda \otimes \ldots \otimes \mathcal{E}_N^\lambda (\sigma_1 \otimes \ldots \otimes \sigma_S),$$

for a (discrete) random variable Λ.

The question that naturally arises from these setups is which (d^N-dimensional) quantum states can be produced in such a network under LO or LOSR constraints. This is known as the *network compatibility problem*.

Problem 6 (network compatibility)**.** *Given a d^N-dimensional quantum state ρ on N parties and a causal structure G with N end nodes, determine whether ρ can be produced in a LO/LOSR quantum model compatible with the causal structure.*

Remark. There is a third well-known setup of network scenarios in which one allows for local operations and classical communication (LOCC). For the discussion of network compatibility, however, this setup is less interesting: given enough rounds of communication, one can simply create any quantum state by distributing it through the network via teleportation. The analysis becomes far less trivial again if one restricts the number of rounds of communication.

To try to answer the question of Problem 6, first note that the causal compatibility problem is a special case of the network compatibility problem with LO constraints, where the output is simply a classical state. The idea behind solving the more general setup is therefore largely the same: it is still possible to perform inflation as a thought experiment. If the output state can be produced in the network, it should be possible to copy the source states and pick any of those copies to produce the output state.

This procedure is formalized in the following way. Suppose we want to test whether a state $\rho \in K((\mathbb{C}^d)^{\otimes N})$ is compatible with a given network corresponding to a DAG G. Let $\{M_i^\ell\}_i$ be a set of m observables in $B(\mathbb{C}^d)$, which act on the output space of party ℓ. Denote the algebra that is generated by all such observables on all parties of G by $\mathcal{D}_G$. Then one can calculate the expectation values $y_{M_{i_1}^1 \otimes \ldots \otimes M_{i_N}^N} = \text{tr}\big(\rho\, M_{i_1}^1 \otimes \ldots \otimes M_{i_N}^N\big)$ for all possible combinations of the observables. If the set of observables is complete, these expectation values uniquely identify the state.

To perform inflation, one now has to distinguish the LO and LOSR setups. In the case where only local operations are allowed, the setup is exactly the same as

described in section 3.2. To see this, one can perform the following calculation:

$$
\begin{aligned}
&\operatorname{tr}\!\big(\rho\, M_{i_1}^1 \otimes \ldots \otimes M_{i_N}^N\big) \\
&= \operatorname{tr}\!\big(\mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_N(\sigma_1 \otimes \ldots \otimes \sigma_S)\, M_{i_1}^1 \otimes \ldots \otimes M_{i_N}^N\big) \\
&= \operatorname{tr}\!\big((\sigma_1 \otimes \ldots \otimes \sigma_S)\, \mathcal{E}_1^* \otimes \ldots \otimes \mathcal{E}_N^*(M_{i_1}^1 \otimes \ldots \otimes M_{i_N}^N)\big) \\
&= \operatorname{tr}\!\big((\sigma_1 \otimes \ldots \otimes \sigma_S)\, U_1 \otimes \ldots \otimes U_N(M_{i_1}^1 \otimes \mathbb{1} \otimes \ldots \otimes M_{i_N}^N \otimes \mathbb{1})U_1^* \otimes \ldots \otimes U_N^*\big) \\
&= \operatorname{tr}\!\big((\sigma_1 \otimes \ldots \otimes \sigma_S)\, (U_1(M_{i_1}^1 \otimes \mathbb{1})U_1^* \otimes \ldots \otimes U_N(M_{i_N}^N \otimes \mathbb{1})U_N^*)\big),
\end{aligned}
\tag{4.3}
$$

where the definition of an LO network state was used in the first equality, the channels $\mathcal{E}_j$ were moved to the observables by taking the adjoint maps in the second equality, and the last two equalities follow from the Stinespring representation of the adjoint channel and the fact that the channels act locally on each party. From the final equation, one can see that the operators $U_j(M_{i_j}^j \otimes \mathbb{1})U_j^*$ have the same algebraic structure as the operators $M_{i_j}^j$. In other words, the problem is invariant under the local operations. This is not surprising, since we were only trying to determine the state uniquely up to local isometries.

The convergence results for the LO setup thus carry over directly from the causal compatibility problem of Chapter 3:

Corollary 4.1.1. *The modified quantum inflation technique of Section 3.2 yields a hierarchy of semidefinite programs that converges to the network compatibility problem under LO constraints. It will have the same drawbacks as in section 3.2, namely that it introduces a very large number of variables and is not monotonically convergent.*

Proof. This follows immediately from Eq. (4.3) and Theorem 3.2.1, since the construction works for general C^*-algebras. □

Corollary 4.1.2. *For the specific case of star networks, most notably the bilocal network, the original version of the quantum inflation technique is convergent for the network compatibility problem under LO constraints. The same is true for the polarization hierarchy and the scalar extension hierarchy (discussed in detail in Chapter 5).*

Proof. The construction of Sec. 3.3 holds for any C^*-algebra, and is therefore also applicable to the network compatibility problem under LO constraints, due to the calculation (4.3). □

For the LOSR setup the same procedure does not quite work, due to the preshared randomness. The shared randomness prohibits us from performing the quan-

tum inflation technique in the usual way. Instead, for the original version of quantum inflation, one needs to choose a so-called *non-fanout inflation* network [21] that is compatible with the original network.

In a non-fanout inflation G_{NF} of a DAG G, the sources and parties are copied in such a way that each copy of a party is locally indistinguishable from the original party it was copied from. That is, it receives exactly the same amount of source states, and each such source state corresponds to a copy of a source that was send to the party in the original DAG. The full level 2 inflation of the triangle scenario depicted in Fig. 3.1 does not obey these requirements, since each of the parties receives *two* copies of each of the source states that they would have received in the original triangle graph. However, if we consider the graph that corresponds to the commuting subalgebras $\mathcal{A}^{12}, \mathcal{A}^{21}, \mathcal{B}^{11}, \mathcal{B}^{22}, \mathcal{C}^{11}, \mathcal{C}^{22}$, we get an inflation of the triangle that corresponds to Fig. 4.1, known as the wheel inflation [21]. Since the algebras $\mathcal{A}^{12}$ and $\mathcal{A}^{21}$ in the full level 2 inflation are acting on different states, they can be considered as separate parties acting on spatially separated systems. The same is true for the algebras of Bob and Charlie. Note that each of the parties in the wheel inflation receives exactly the same states as their counterparts in the triangle scenario. The wheel inflation is therefore an example of a non-fanout inflation. Apart from the local correlations being the same, the wheel inflation is also indistinguishable from the triangle for each set of two neighboring parties. For example, if we consider the subalgebras $\mathcal{A}^{21}$ and $\mathcal{B}^{11}$, we see that they are acting on the states ρ_{CA}^2, ρ_{AB}^1 and ρ_{BC}^1 in exactly the same way as in the triangle scenario. Lastly, the 3-local correlations of the triangle will in general not be found in the wheel inflation: none of the combinations of the copies of the algebras $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ in the wheel inflation results from a set of three sources like in the triangle scenario.

We will now introduce some notation to treat a more general set of causal structures. Label the parties by a vector $\vec{s}_\ell$, indicating which copy of the source states they are acting on. For example, $1^{\vec{s}_1}$ is the copy of party 1, with a channel acting on the sources indicated by $\vec{s}_1$. For certain subgraphs the input states and channels will be indistinguishable from those in the original graph, as we have seen in the wheel inflation. Hence, for the non-fanout inflation network G_{NF} there should exist a set of local channels that result in a state $\omega_{G_{NF}}$, such that whenever a subgraph H_{NF} of G_{NF} is equivalent to a subgraph H of G, it holds that the reduced output state $\omega_{H_{NF}}$ of the inflation is equal to the reduced output state ρ_H of the original network. This means that all of the expectation values should correspond to those in the original graph as well. That is,

$$\mathrm{tr}\left(\omega_{H_{NF}} M_{i_1}^{\vec{s}_{\ell_1}} \otimes \ldots \otimes M_{i_K}^{\vec{s}_{\ell_K}}\right) = \mathrm{tr}\left(\rho_H M_{i_1}^{\ell_1} \otimes \ldots \otimes M_{i_K}^{\ell_K}\right)$$
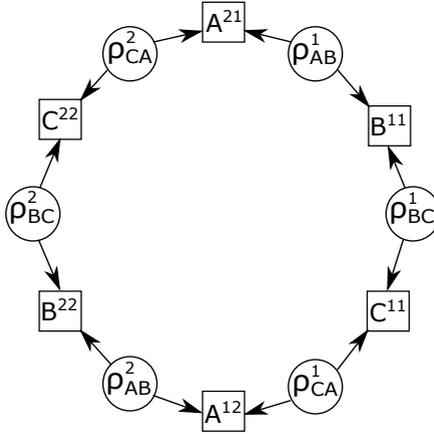
for $\ell_1, \ldots, \ell_K \in H$.

Figure 4.1: An example of a non-fanout inflation of the triangle scenario, known as the wheel inflation. Each copied party is locally indistinguishable from the party it was copied from.

Let $\mathcal{D}_{G_{NF}}$ be the algebra generated by all the observables $\{M_i^{\vec{s}_\ell}\}$ in the inflation graph G_{NF} and the relations among those variables. The network compatibility problem then takes on the form

$$
\begin{aligned}
\min_{\omega_{G_{NF}} \in K(\mathcal{D}_{G_{NF}})} \quad & 0 \\
\text{s. t.} \quad & \omega_{G_{NF}}(x') = \rho_G(x) \quad \text{for all } x' \sim x, \\
& \omega_{G_{NF}}(x') = \omega_{G_{NF}}(x) \quad \text{for all } x' \sim x,
\end{aligned}
\tag{4.4}
$$

where $x' \sim x$ if the operators x and x' are copies of the same operator, and are non-trivial only on subalgebras corresponding to isomorphic subgraphs of G_{NF}. The first constraint in Eq. (4.4) imposes that the subgraphs of G_{NF} that are isomorphic to subgraphs of the original causal structure lead to the same expectation values. The second constraint imposes that the same is true when two different subgraphs of G_{NF} are isomorphic. An example of this second constraint for the wheel inflation of Fig. 4.1, which is not implied by the first set of constraints, is

$$
\omega_{\text{Wheel}}(a^{21} b^{11} c^{11}) = \omega_{\text{Wheel}}(a^{12} b^{22} c^{22}),
$$

where $a^{21} \in \mathcal{A}^{21}$, etc.

Despite the added difficulties, there are still some convergence results for the

LOSR case as well:

Corollary 4.1.3. *There exists a hierarchy of semidefinite programs, closely related to the modified quantum inflation technique of Section 3.2 that converges to the network compatibility problem under LOSR constraints. It will have the same drawbacks as in section 3.2, namely that it introduces a very large number of variables and is not monotonically convergent.*

Proof. By treating the shared random variable as a classical unobserved source, one can also define "fanout" inflations in the LOSR setup for the modified inflation technique of section 3.2. That is, write the network state as

$$\rho_{\text{LOSR}} = \mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_N (\Lambda \otimes \sigma_1 \otimes \ldots \otimes \sigma_S),$$

where Λ is a local hidden variable distributing randomness between all parties, and each of the $\mathcal{E}_j$ now also act on a part of the variable Λ. Repeat the calculation in Eq. (4.3). Then, when defining the local algebras, add an abelian local algebra for each party which models the classical random variable received by that party. At this point we are back in the usual framework for the modified version of inflation and can claim the same convergence results[1]. $\qquad\square$

The result of Sec. 3.3 for the bilocal scenario does not obviously carry over. Though it is true that the argument only uses non-fanout inflations, the construction in Theorem 3.3.10 makes central use of the factorization condition (3.55), which does not hold for states produced in an LOSR network. One could hope to use the separability of the state instead: The GNS representation arising from the restriction of the state to Alice and Charlie is a direct integral of the form [43, III.1.6]

$$\mathcal{H}_{AC} = \int^{\oplus} \mathrm{d}\mu(\sigma) \mathcal{H}_A^{\sigma_A} \otimes \mathcal{H}_C^{\sigma_C}, \tag{4.5}$$

$$\pi_{AC}(ac) = \int^{\oplus} \mathrm{d}\mu(\sigma) \pi_A^{\sigma_A}(a) \otimes \pi_C^{\sigma_C}(c). \tag{4.6}$$

But this does not clearly lead to an algebra for Bob that is generated by a tensor product of two commuting subalgebras. We therefore leave the question of convergence of the inflation technique for the bilocal LOSR network open for future research. The scalar extension and polarization hierarchies are not suitable for this version of the network compatibility problem, since there are no factorization constraints to enforce.

---

[1] The reason why this does not work for the original version of quantum inflation is because it is not clear how to model the classical subsystem. Indeed, as pointed out in section 3.1.1 one of the biggest hurdles in proving convergence is finding a way to construct those subsystems.

In summary, the essence of the inflation technique for network scenarios is no different from that of (non-fanout) inflations as discussed in e.g. Ref. [21]. The main difference is the fact that we allow for more general algebras to optimize over. In particular, we are interested in algebras that completely determine a d^N-dimensional state via expectation values, while still having a convenient algebraic structure that is relatively easy to implement and that reduces the size of the problem.

Below, we will discuss two such examples: the algebra of generalized Pauli matrices, and the algebra of matrix units in d dimensions. We will only show the optimization problems for the original version of quantum inflation, since the optimization problem for the modified version of section 3.2 would be quite hard to parse and would not serve well as a clarifying example.

Example 4.1.4 (Pauli operators)**.** As a generalization of the well-known 2-dimensional Pauli matrices $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$, we define the *generalized Pauli matrices* in d dimensions [103]

$$X := \sum_{i=0}^{d-1} |i+1\rangle\langle i|, \qquad Z := \sum_{i=0}^{d-1} \omega^i |i\rangle\langle i|, \qquad (4.7)$$

where $\omega = e^{i\frac{2\pi}{d}}$ is a d'th root of unity and $|d\rangle$ is identified with $|0\rangle$. More abstractly, their algebra is defined through the relations

$$Z^l X^m = e^{i\frac{2\pi lm}{d}} X^m Z^l, \qquad (4.8)$$

$$(X^m)^* = X^{d-m}, \quad (Z^m)^* = Z^{d-m}, \qquad (4.9)$$

$$X^d = Z^d = \mathbb{1}. \qquad (4.10)$$

We will denote the generalized Pauli algebra by $\mathcal{P}$.

The representation (4.7) in d dimensions is particularly useful when we define the displacement operators

$$D(l, m) = e^{-i\frac{\pi lm}{d}} Z^l X^m.$$

These operators form an orthogonal set with which we can write a d-dimensional state in the Bloch representation [103]

$$\rho = \frac{1}{d} \sum_{l,m=0}^{d-1} \operatorname{tr}(\rho D(l,m)) D^*(l,m), \qquad (4.11)$$

which emphasizes the fact that the expectation values of the state on the generalized Pauli matrices uniquely determine the state. The N party state can then be charac-

terized by the expectation values on $\mathcal{P}^{\otimes N}$.

Let G_{NF} be a non-fanout inflation of a DAG G with $\tilde{N}$ end nodes. Let $S_{\text{inv}} \subset S_{\tilde{N}}$ be the set of permutations of the parties that leaves the graph G_{NF} invariant. Concretely, the optimization problem then takes the form

$$\min_{\omega \in K(\mathcal{P}^{\otimes \tilde{N}})} \quad 0$$

$$\text{s. t.} \quad \omega_{H_{NF}}(Z^{l_1}X^{m_1} \otimes \ldots \otimes Z^{l_K}X^{m_K}) = \rho_H(Z^{l_1}X^{m_1} \otimes \ldots \otimes Z^{l_K}X^{m_K})$$

$$\text{for all } H_{NF} \sim H,$$

$$\omega(Z^{l_1}X^{m_1} \otimes \ldots \otimes Z^{l_{\tilde{N}}}X^{m_{\tilde{N}}})$$

$$= \omega(Z^{\pi(l_1)}X^{\pi(m_1)} \otimes \ldots \otimes Z^{\pi(l_{\tilde{N}})}X^{\pi(m_{\tilde{N}})})$$

$$\text{for all } \vec{l} \in (\mathbb{Z}_d)^{\tilde{N}}, \vec{m} \in (\mathbb{Z}_d)^{\tilde{N}}, \pi \in S_{\text{inv}},$$

$$(4.12)$$

where $\omega_{H_{NF}}$ is the restriction of ω to the subgraph H_{NF} and the systems are ordered such that the first system in H_{NF} is equivalent to the first system in H, and so forth.

Example 4.1.5. (Matrix units) A second example is given by the algebra of matrix units [43, Example II.8.3.2.iv]. This algebra is defined as the universal algebra of a set of generators $\mathcal{G} = \{e_{i,j}\}$, $i, j \in [0, d-1]$ and the relations

$$\mathcal{R} = \{e_{ij}e_{kl} = \delta_{jk}e_{il}, e_{ij}^* = e_{ji}\}. \tag{4.13}$$

The universal C^*-algebra $C^*(\mathcal{G}|\mathcal{R})$ is isomorphic to $\mathbb{M}_d$, the set of $d \times d$ matrices, where each of the generators e_{ij} can be thought of as the $d \times d$ matrix with a 1 on position (i, j) and 0 everywhere else.

It is clear that such an algebra also uniquely determines a d-dimensional state (up to local isometries), simply by writing down the $d \times d$ density matrix and requiring $\text{tr}(\rho e_{ij}^*) = \rho_{ij}$. More generally, one can determine a state on N parties by taking expectation values of $(C^*(\mathcal{G}|\mathcal{R}))^{\otimes N}$ so that the optimization problem becomes

$$\min_{\omega \in K((C^*(\mathcal{G}|\mathcal{R}))^{\otimes \tilde{N}})} \quad 0$$

$$\text{s. t.} \quad \omega_{H_{NF}}(e_{ij} \otimes \ldots \otimes e_{i'j'}) = \rho_H(e_{ij} \otimes \ldots \otimes e_{i'j'})$$

$$\text{for all } H_{NF} \sim H,$$

$$\omega_{H_{NF}^1}(e_{ij} \otimes \ldots \otimes e_{i'j'}) = \omega_{H_{NF}^2}(e_{ij} \otimes \ldots \otimes e_{i'j'})$$

$$\text{for all } H_{NF}^1 \sim H_{NF}^2,$$

$$(4.14)$$

where again the systems are ordered in the obvious way.

4.2 Analytic proof of incompatibility for graph states

The following section is largely taken from Paper [3]. Its wording and notation have been slightly modified to be consistent with the rest of the text. Furthermore, the supplementary material has been inserted in the main text in order to make the argument more linear.

Out of all proposed real-life implementations of quantum information, quantum networks stand as one of the most promising ones. They appear as key components for several proposals of possible applications of quantum information theory, such as quantum key distribution [26, 27, 28], clock synchronisation [29], parallel computing [30] or even a quantum internet [31, 32, 33, 34]. What is more, their relative simplicity as compared to other quantum technologies, makes them perhaps the closest ones to commercial implementation. This sentiment is also supported by the significant progress in the experimental implementation of quantum networks that has recently been made [104, 105, 106, 107, 108].

At their core, quantum networks are simply collections of parties and of sources of multipartite quantum states. Their most natural model (referred to as LOCC networks) is one that allows the parties to act with the most general local operations on their shares of the distributed states and to coordinate their actions by using classical communication. However, while connected LOCC networks enable preparing any multipartite state, the use of classical communication might be problematic for their commercial implementations.

When considering the possible future applications of quantum networks one has to take into account that the distances between parties will substantially increase as compared to the current state-of-the-art experiments. Therefore, while classical communication between parties can be considered almost instantaneous in a lab setup, this will not be the case for quantum networks spanning many different countries or even continents. From that point of view, it would be beneficial, e.g. for quantum key distribution protocols, to use quantum states that require as little classical communication as possible to be generated. This has not only the potential to decrease the latency of quantum networks but also to reduce the noise therein, as the longer a quantum state has to be stored, the noisier the state gets.

One is thus forced to consider quantum networks in which the amount of classical communication between the parties is limited or even no communication is allowed (see Ref. [109]). A possible model of quantum networks that fulfills this requirement is one in which the parties can apply arbitrary quantum channels to their particles, however, they cannot use classical communication. Instead, they are allowed to orchestrate their actions by using some pre-shared classical information. We call such networks LOSR (local operations and shared randomness) quantum networks.

While the above no-communication assumption severely reduces the capability of generating multipartite states in LOSR quantum networks as compared to the LOCC ones, they are still more general than those in which the parties can only implement unitary operations and no randomness is shared. While the LOSR networks have become an object of intensive studies [102, 110, 111, 112, 113], it remains unclear what multipartite entangled states can actually be prepared in them.

In quantum information graph states stand as one of the most representative classes of multipartite states, including the Greenberger-Horne-Zeilinger [114], cluster [115] or the absolutely maximally entangled [116] states. Moreover, they are key resources for many applications, such as quantum computing [117, 118, 119], multipartite secret sharing [120], or quantum metrology [121]. It is thus a question of utmost importance how difficult it is to prepare such states in quantum networks.

In this chapter we will prove a negative result to this question in the sense that no graph states of arbitrary prime local dimension (or any state sufficiently close to them) can be prepared in LOSR networks with bipartite sources. We thus generalize the recent results that the three-partite GHZ states [102, 122] or any N-qubit graph states with N up to 12 cannot be prepared in such networks [112]. Simultaneously, our work is complementary to Ref. [113] showing that no qubit or qutrit graph states of an arbitrary number of parties can be prepared in LOSR networks even with $(N - 1)$-partite sources. Our proof employs the quantum inflation method [21, 22] which is perfectly suited to tackle these types of questions [22, 102].

While finishing the original manuscript [3] of this result, we became aware of a very similar result by Y.-X. Wang *et al* [123], and coordinated the publication of the preprints of the manuscripts.

4.2.1   Preliminaries

(1) Graph states. Consider a multigraph G, which is a graph in which any two vertices can be connected by more than one edge, but no edge can connect a vertex to itself. This can be seen as a special case of a weighted graph, in which the weights are integers. Let $\Gamma_{i,j}$ denote the number of edges connecting vertices i and j, and let $\mathcal{N}_i$ be the neighbourhood of vertex i, i.e. the set of vertices that are connected to i by at least one edge (see Fig. 4.2 for an example). To associate a quantum state to an N-vertex multigraph G, we consider a Hilbert space $\mathcal{H} = \mathbb{C}_d^{\otimes N}$, where each qudit space $\mathbb{C}_d$ corresponds to one of the vertices of G. We assume that d is prime and will consider multigraphs such that $d \geq \max_{i,j} \Gamma_{i,j}$. To each vertex i we associate the operator

$$g_i = X_i \prod_{j \in \mathcal{N}_i} (Z_j)^{\Gamma_{i,j}}, \tag{4.15}$$

$$\Gamma_{1,2} = 2$$
$$\Gamma_{1,3} = 1$$
$$\Gamma_{2,3} = 0$$

$$\mathcal{N}_1 = \{2,3\}$$
$$\mathcal{N}_2 = \{1\}$$
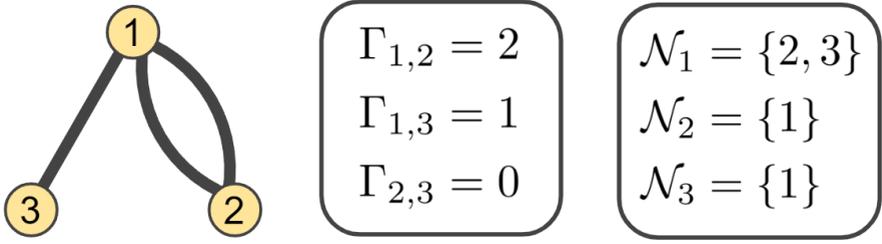$$\mathcal{N}_3 = \{1\}$$

Figure 4.2: An example of a multigraph with three vertices.

where X and Z are the generalized Pauli matrices of Eq. (4.7). The subscripts in (4.15) label the subsystems on which these operators act. One defines a graph state $|G\rangle$ associated to G to be the unique state in $\mathcal{H}$ obeying $g_i |G\rangle = |G\rangle$ ($i = 1, \ldots, N$) (for a review on graph states see e.g. [35]).

(2) Quantum networks. Let us consider a scenario in which N parties, labelled $1, \ldots, N$, receive quantum states distributed by independent sources. Each party i can perform an arbitrary local operation represented by a quantum channel $\mathcal{E}_i$, on their shares of these states. We also assume that parties cannot communicate with each other, but instead all have access to some shared randomness, which is a random variable Λ with a distribution $\{p_\lambda\}_{\lambda \in \Lambda}$. These assumptions describe the LOSR scenario.

There is one more assumption, independent of LOSR, that we make: the sources distributing quantum states are bipartite, that is, every source distributes a quantum state to only two parties. We say that two parties are connected if they share a source state.

The most general state that can be produced in such an LOSR network is given by [112]

$$\rho = \sum_\lambda p_\lambda \mathcal{E}_1^\lambda \otimes \cdots \otimes \mathcal{E}_N^\lambda \left[\sigma_{1,2} \otimes \sigma_{1,3} \otimes \cdots \otimes \sigma_{N-1,N} \right], \qquad (4.16)$$

where $\sigma_{i,j}$ denotes a source state that is distributed between parties i and j, $\sum_\lambda p_\lambda = 1$, and the superscript λ denotes the dependence of local operations on the shared random variable.

Here, three remarks are in order. First, tensor products of $\mathcal{E}_i^\lambda$ and of $\sigma_{i,j}$ are taken with respect to different sets of subsystems; while the former is taken with respect to different parties, the latter separates states from different sources.

Second, in general we could expect the distributed states $\sigma_{i,j}$ to also depend on Λ, since the sources can be classically correlated as well. However, because we

do not impose any restriction on the dimension of the $\sigma_{i,j}$, one can get rid of this dependency by considering a Hilbert space of sufficiently high dimension [102].

Third, in this work, we will assume that every network we consider (not including inflations) is fully-connected, i.e., each party shares a bipartite state with every other party. We can make this assumption without loss of generality because the behaviour of each quantum network can always be simulated by the fully connected one: taking $\sigma_{i,j}$ to be the maximally mixed state produces the same outcomes as removing the connection between the nodes i and j. Thus, since we are trying to show *in*compatibility of graph states, finding that they cannot be produced in this more powerful setup where all parties share a source state is sufficient to show that they cannot be produced in any bipartite network with LOSR.

(3) Network inflation. Let us briefly recall the relevant tools from the network inflation method [21, 22] which we use to derive the main results. Given some network $\mathcal{O}$, an inflation network $\mathcal{I}$ of $\mathcal{O}$ is a network that consists of multiple copies of parties and sources from the original network. Whether two parties are connected in $\mathcal{I}$ depends on the choice of inflation, with the only restriction that they are connected via a copy of a source from $\mathcal{O}$.

This construction is very general as many different inflations can be considered for a given network $\mathcal{O}$. Here we focus on a certain class of inflations that are tailored to our proof. Consider an N-partite network $\mathcal{O}$ that we want to analyse; as mentioned before, we will assume it to be fully connected. In our approach, every inflation $\mathcal{I}$ of $\mathcal{O}$ that we consider consists of two copies of the parties from $\mathcal{O}$ labeled $1, \ldots, N$ and $1', \ldots, N'$. We assume that parties i and i' apply the same local operation as the original parties in $\mathcal{O}$: $\mathcal{E}_i^{\mathcal{I}} = \mathcal{E}_{i'}^{\mathcal{I}} = \mathcal{E}_i^{\mathcal{O}}$, where the superscripts indicate the network. We also assume that each party i in $\mathcal{I}$ is connected to either j or j' (but never to both). Furthermore, if two copies of parties share a source state, this state is a copy of the state that is shared between the original parties in $\mathcal{O}$. These last two assumptions imply that for every pair of parties $i, j \in \mathcal{O}$ and any inflation $\mathcal{I}$, exactly one of the following statements is true:

$$\sigma_{i,j}^{\mathcal{I}} = \sigma_{i',j'}^{\mathcal{I}} = \sigma_{i,j}^{\mathcal{O}}, \quad \text{or} \quad \sigma_{i,j'}^{\mathcal{I}} = \sigma_{i',j}^{\mathcal{I}} = \sigma_{i,j}^{\mathcal{O}}. \tag{4.17}$$

Lastly, it is assumed that the shared randomness in $\mathcal{I}$ is distributed between all copies of the parties, such that the total output state is again described by Eq. (4.16).

Two networks are said to be isomorphic if one can be transformed into the other by switching a subset of the non-primed parties with their primed counterparts.

The above assumptions allow us to establish very useful relations between expectation values $\langle \cdot \rangle_{\mathcal{I}_1}$ and $\langle \cdot \rangle_{\mathcal{I}_2}$ calculated over states from two different inflations $\mathcal{I}_1, \mathcal{I}_2$.

Let $\mathcal{O}$ be a network and $B = \bigotimes_{i \in \mathcal{O}} B_i$. We call a subnetwork of $\mathcal{O}$ the B sub-

network if it consists only of parties i for which $B_i \neq \mathbb{1}$ and sources that distribute states to said parties. For example, if $\mathcal{O}$ is a tripartite, fully-connected network and $B = X \otimes X \otimes \mathbb{1}$, then the B subnetwork of $\mathcal{O}$ is a network consisting of parties 1 and 2 and the sources distributing a state to those parties.

Fact 4.2.1. *Consider a network $\mathcal{O}$ and two different inflations of it, $\mathcal{I}_1$ and $\mathcal{I}_2$. Consider also two matrices $B = \bigotimes_{i \in \mathcal{I}_1} B_i$ and $C = \bigotimes_{i \in \mathcal{I}_2} C_i$ that act nontrivially on some subnetworks $\mathcal{I}'_i \subseteq \mathcal{I}_i$. Then, $\langle B \rangle_{\mathcal{I}_1} = \langle C \rangle_{\mathcal{I}_2}$ if*

1. *The subnetworks $\mathcal{I}'_1$ and $\mathcal{I}'_2$ are isomorphic.*

2. *For every party i we have that either $B_i = C_i$ and $B_{i'} = C_{i'}$, or $B_{i'} = C_i$ and $B_i = C_{i'}$.*

The above fact can be proven with the use of (4.16) by tracing out the parties where B and C act trivially. We will see several examples and applications of this fact below.

Note that since the original network $\mathcal{O}$ can be considered a trivial inflation of itself, Fact 4.2.1 also enables establishing relations between $\langle \cdot \rangle_{\mathcal{O}}$ and $\langle \cdot \rangle_{\mathcal{I}}$ for some inflation $\mathcal{I}$.

4.2.2 Incompatibility of graph states

Let us now move on to the main results of this section, namely that no graph state can be generated in quantum networks with bipartite sources. We begin by presenting the key ingredients of our approach, which is inspired by the recent work [112]. The main idea of the proof is to show that the assumption that a graph state can be generated in a network leads to the violation of a certain inequality that follows from the lemma below.

Lemma 4.2.2. *Consider two unitary matrices A_1, A_2 acting on some Hilbert space $\mathbb{C}_D$ with D being a multiple of some prime number $d \geq 2$. Assume moreover that A_i satisfy $A_1^d = A_2^d = \mathbb{1}$ and $A_1 A_2 = \omega^q A_2 A_1$ for $\omega = e^{i\frac{2\pi}{d}}$ and for some $q \in \{1, \ldots, d-1\}$. Then*

$$\sum_{k=0}^{d-1} \langle A_1^k + A_2^k \rangle \leq d + \sqrt{d} \tag{4.18}$$

where $\langle \cdot \rangle \equiv \mathrm{Tr}[\rho(\cdot)]$ and the above holds true for any state ρ acting on $\mathbb{C}_D$.

Proof. We start from rewriting the sum in Eq. (4.18) in the following form

$$\sum_{k=0}^{d-1}\langle A_1^k + A_2^k\rangle_\rho = \sum_{k=0}^{d-1}\mathrm{tr}\left[(A_1^k + A_2^k)\rho\right], \tag{4.19}$$

where ρ is an arbitrary state acting on $\mathbb{C}^D$. Let us then assume that $A_1 A_2 = \omega^q A_2 A_1$ for some $q \in \{1, \ldots, d-1\}$ and that $A_1^d = A_2^d = \mathbb{1}$. As proven in Ref. [124, Proposition B.1] these two conditions imply the existence of a unitary $U : \mathbb{C}^D \rightarrow \mathbb{C}^d \otimes \mathcal{H}'$ such that

$$UA_1U^\dagger = X \otimes \mathbb{1}, \quad UA_2U^\dagger = Z^q \otimes \mathbb{1}, \tag{4.20}$$

where $\mathbb{1}$ is the identity acting on $\mathcal{H}'$. Denoting $\rho' = U\rho U^\dagger$, Eq. (4.19) can be rewritten as

$$\sum_{k=0}^{d-1}\mathrm{tr}\left[((X^k + Z^{qk}) \otimes \mathbb{1})\rho'\right]. \tag{4.21}$$

We can trace out the subsystems corresponding to the Hilbert space $\mathcal{H}'$ which leads us to

$$\sum_{k=0}^{d-1}\mathrm{tr}\left[(X^k + Z^{qk})\tilde{\rho}\right] = \sum_{k=0}^{d-1}\mathrm{tr}\left[(X^k + Z^{qk})\tilde{\rho}\right], \tag{4.22}$$

where $\tilde{\rho} = \mathrm{tr}_{\mathcal{H}'}\rho'$.

Let us now consider the eigendecompositions of both X and Z operators,

$$X = \sum_{i=0}^{d-1}\omega^i\,|\phi_i\rangle\langle\phi_i|, \quad Z = \sum_{i=0}^{d-1}\omega^i\,|i\rangle\langle i|, \tag{4.23}$$

where $|\phi_i\rangle$ and $|i\rangle$ are the eigenvectors of X and Z, respectively. Exploiting the well-known fact that

$$\sum_{k=0}^{d-1}\omega^{ki} = d\delta_{i,0}, \tag{4.24}$$

it is not difficult to observe that

$$\sum_{k=0}^{d-1}X^k = d\,|\phi_0\rangle\langle\phi_0|, \quad \sum_{k=0}^{d-1}Z^{kq} = d\,|0\rangle\langle0|, \tag{4.25}$$

where $|\phi_0\rangle$ and $|0\rangle$ are eigenstates corresponding to the eigenvalue 1 of X and Z^q,

respectively. This, taking into account, Eq. (4.22), allows us to rewrite Eq. (4.19) as

$$\sum_{k=0}^{d-1} \langle A_1^k + A_2^k \rangle_\rho = d \operatorname{tr} \left[(|\phi_0\rangle\langle\phi_0| + |0\rangle\langle0|)\tilde{\rho} \right]. \tag{4.26}$$

Let us now notice that for any pair of normalized vectors $|\psi_i\rangle$ $(i = 1, 2)$, the sum of projectors $P = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|$ is a rank-two matrix whose eigenvalues are $1 \pm |\langle\psi_1|\psi_2\rangle|$ and therefore $\operatorname{tr}[P\sigma] \le 1 + |\langle\psi_1|\psi_2\rangle|$ for any σ. Taking into account that the eigenbasis of X and Z^q is mutually unbiased, meaning that $|\langle0|\phi_0\rangle| = 1/\sqrt{d}$, the above bound implies that

$$\sum_{k=0}^{d-1} \langle A_1^k + A_2^k \rangle_\rho \le d + \sqrt{d}, \tag{4.27}$$

which completes the proof.   □

While the inequality of Lemma 4.2.4 is enough to prove the main result of this chapter, we have also found another way allowing to reach the same result that exploits an equality instead. This is shown in the following Lemma.

Lemma 4.2.3. *Given two positive real numbers λ_1, λ_2 and two matrices A_1, A_2 such that $A_1^d = A_2^d = \mathbb{1}$ and $A_1 A_2 = \omega^q A_2 A_1$ for $q \in \{1, \ldots, d-1\}$, if q and d are coprime, then it holds that*

$$\left\langle (\lambda_1 A_1 + \lambda_2 A_2)^{nd} \right\rangle_\rho = (\lambda_1^d + \lambda_2^d)^n \qquad \forall n \in \mathbb{N}, \tag{4.28}$$

where ρ is an arbitrary state acting on $\mathbb{C}^D$.

Proof. Exploiting the commutation relation $A_2 A_1 = \eta A_1 A_2$, where $\eta = \omega^q$ and the fact that $A_i^d = \mathbb{1}$ we can rewrite

$$(\lambda_1 A_1 + \lambda_2 A_2)^d$$
$$= \sum_{k=1}^{d-1} A_1^{d-k} A_2^k \lambda_1^{d-k} \lambda_2^k \sum_{i_1=0}^{d-k} \cdots \sum_{i_k=i_{k-1}}^{d-k} \eta^{\sum_{j=1}^k i_j}$$
$$+ A_1^d \lambda_1^d + A_2^d \lambda_2^d. \tag{4.29}$$

Using Eq. (4.31) proven below in Lemma 4.2.4, the first term on the right vanishes and the above considerably simplifies to

$$(\lambda_1 A_1 + \lambda_2 A_2)^d = A_1^d \lambda_1^d + A_2^d \lambda_2^d, \tag{4.30}$$

which due to the fact that $A_i^d = \mathbb{1}$ leads us directly to Eq. (4.28), completing the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 4.2.4. *Given $\eta = \omega^{-q}$ for $q \in \{1, \ldots, d-1\}$, if q and d are coprime, the following holds true*

$$\sum_{i_1=0}^{d-k}\sum_{i_2=i_1}^{d-k}\cdots\sum_{i_k=i_{k-1}}^{d-k}\eta^{\sum_{j=1}^{k}i_j} = 0, \tag{4.31}$$

where $k \in \{1, \ldots, d-1\}$.

Proof. Let us consider the expression on the left-hand side of Eq. (4.31) and shift the summation index in the last sum by k. That is, we apply the transformation $i_k \to i_k - k$, which leads us to

$$\sum_{i_1=0}^{d-k}\cdots\sum_{i_k=i_{k-1}}^{d-k}\eta^{\sum_{j=1}^{k}i_j} = \eta^{-k}\sum_{i_1=0}^{d-k}\cdots\sum_{i_k=i_{k-1}+k}^{d}\eta^{\sum_{j=1}^{k}i_j}. \tag{4.32}$$

We then implement a similar transformation to the second to last, where this time we shift the index by $k-1$ $[i_{k-1} \to i_{k-1} - (k-1)]$, which allows us to rewrite the above as

$$\eta^{-k}\sum_{i_1=0}^{d-k}\cdots\sum_{i_{k-1}=i_{k-2}}^{d-k}\sum_{i_k=i_{k-1}+k}^{d}\eta^{\sum_{j=1}^{k}i_j}$$

$$= \eta^{-k-(k-1)}\sum_{i_1=0}^{d-k}\cdots\sum_{i_{k-1}=i_{k-2}+k-1}^{d-1}\sum_{i_k=i_{k-1}+1}^{d}\eta^{\sum_{j=1}^{k}i_j}. \tag{4.33}$$

Notice that this action changes also the summation range of the last sum.

We then recursively shift the remaining summation indices as $i_j \to i_j - j$ for $j = 1, \ldots, k-2$, which yields the following expression

$$\sum_{i_1=0}^{d-k}\sum_{i_2=i_1}^{d-k}\cdots\sum_{i_k=i_{k-1}}^{d-k}\eta^{\sum_{j=1}^{k}i_j}$$

$$= \eta^{-k(k+1)/2}\sum_{i_1=1}^{d-k+1}\sum_{i_2=i_1+1}^{d-k+2}\cdots\sum_{i_k=i_{k-1}+1}^{d}\eta^{\sum_{j=1}^{k}i_j}, \tag{4.34}$$

which can also be stated as

$$\sum_{i_1=0}^{d-k} \sum_{i_2=i_1}^{d-k} \cdots \sum_{i_k=i_{k-1}}^{d-k} \eta^{\sum_{j=1}^{k} i_j} = \eta^{-k(k+1)/2} \sum_{1 \leq i_i < i_2 < \cdots < i_k \leq d} \eta^{\sum_{j=1}^{k} i_j}. \quad (4.35)$$

Our aim now is to prove that the above sum vanishes. To this end, we employ the so-called Vieta's formulas which relate roots x_j of any complex polynomial of degree n

$$P_n(x) = \sum_{i=0}^{n} a_i x^i, \quad (4.36)$$

to the coefficients a_i defining it,

$$\sum_{1 \leq i_i < i_2 < \cdots < i_k \leq n} \left(\prod_{j=1}^{k} x_{i_j} \right) = (-1)^k \frac{a_{n-k}}{a_n}. \quad (4.37)$$

Let us then consider a particular polynomial of degree d,

$$P_d^\omega(x) = x^d - 1, \quad (4.38)$$

whose roots are obviously ω^i. However instead of enumerating these roots by powers of ω, i.e., $x_i = \omega^i$, we can use the fact that q and d are coprime and enumerate them by powers of η, that is,

$$x_i = \eta^i. \quad (4.39)$$

Since all coefficients of this polynomial except a_d and a_0 are zero, Eq. (4.37) implies that for any $k = 1, \ldots, d-1$,

$$\sum_{1 \leq i_i < i_2 < \cdots < i_k \leq d} \prod_{j=1}^{k} \eta^{i_j} = \sum_{1 \leq i_i < i_2 < \cdots < i_k \leq d} \eta^{\sum_{j=1}^{k} i_j} = 0. \quad (4.40)$$

Substitution of the above equation to Eq. (4.35) ends the proof. $\qquad \square$

Remark. Though the proof of Lemma 4.2.4 seems rather technical, the intuition behind it is easier to understand: Due to the commutation relation between A_1 and A_2, each of the terms $A_1^k A_2^{d-k}$ in the binomial expansion of $(A_1 + A_2)^d$ appears equally often with each of the phases $e_q^{i\frac{2\pi q}{d}}$. Hence, summing over all of these terms exactly cancels out this sum.

In order to show a violation of (4.18) we will also use the following fact.

Fact 4.2.5. *Consider three mutually commuting unitary matrices B_i that obey $B_i^d = \mathbb{1}$. If $\langle B_1 B_3 \rangle = \langle B_2 B_3^\dagger \rangle = 1$, then $\langle B_1 B_2 \rangle = 1$.*

This fact follows from an observation that $B_1 B_3$ and $B_2 B_3^\dagger$ are unitary and therefore the fact that $\langle B_1 B_3 \rangle = \langle B_2 B_3^\dagger \rangle = 1$ holds true for some $|\psi\rangle$ implies that $B_1 B_3 |\psi\rangle = |\psi\rangle$ and $B_2 B_3^\dagger |\psi\rangle = |\psi\rangle$. Since B_i mutually commute, one concludes that $B_1 B_2 |\psi\rangle = |\psi\rangle$ which yields the above implication.

Facts 4.2.1 and 4.2.5 together with Lemma 4.2.2 are the key elements of the proof technique. However, before showing how they are combined, let us present an illustrative example.

Example 4.2.6. Consider the triangle network $\mathcal{O}^\triangle$ presented in Fig. 4.3. The most general form of a state that can be generated in such a network under the LOSR assumption is given by

$$\rho^{\mathcal{O}^\triangle} = \sum_\lambda p_\lambda \mathcal{E}_1^{(\lambda)} \otimes \mathcal{E}_2^{(\lambda)} \otimes \mathcal{E}_3^{(\lambda)} (\sigma_{1,2} \otimes \sigma_{1,3} \otimes \sigma_{2,3}), \qquad (4.41)$$

where $\sigma_{i,j}$ is a state shared by parties i and j and $\mathcal{E}_i^{(\lambda)}$ is a local operation performed by the party i.

Next, let us fix $d = 3$ and consider the graph state $|G_\triangle\rangle$ corresponding to the graph[2] of Fig. 4.2, i.e., the state that satisfies $g_i |G_\triangle\rangle = |G_\triangle\rangle$ for $i \in \{1, 2, 3\}$, where the stabilizing operators are given by

$$g_1 = X_1 Z_2^2 Z_3, \qquad g_2 = Z_1^2 X_2, \qquad g_3 = Z_1 X_3. \qquad (4.42)$$

In what follows we prove that this state cannot be generated in the network $\mathcal{O}^\triangle$. In other words, we demonstrate that $|G_\triangle\rangle$ does not admit the form given in Eq. (4.41). This goal is achieved via a proof by contradiction: by assuming that the above graph state can be generated in $\mathcal{O}^\triangle$, and then showing that this assumption leads to the violation of an inequality that follows from Lemma 4.2.2, we reach a contradiction.

To this end, let us consider an inflation of $\mathcal{O}^\triangle$, denoted $\mathcal{I}_0^\triangle$ (cf. Fig. 4.3), and two operators defined on it, $g_1 = X_1 Z_2^2 Z_3$ and $Z_1^2 X_{2'}$. Notice, that while the former acts only on the non-primed parties, the latter acts both on party 1, as well as on the primed party $2'$. So, importantly, even though g_1 and g_2, as defined in Eq. (4.42) commute, this is not the case for g_1 and $Z_1^2 X_{2'}$. Instead, it holds that $g_1 Z_1^2 X_{2'} = \omega Z_1^2 X_{2'} g_1$. Hence, these two operators fulfill the conditions of Lemma 4.2.2 with

---

[2] Note that there are two notions of graphs here that are not to be confused: the graph $\mathcal{O}^\triangle$ corresponds to the network, while the graph $G_\triangle$ corresponds to the graph state $|G_\triangle\rangle$.
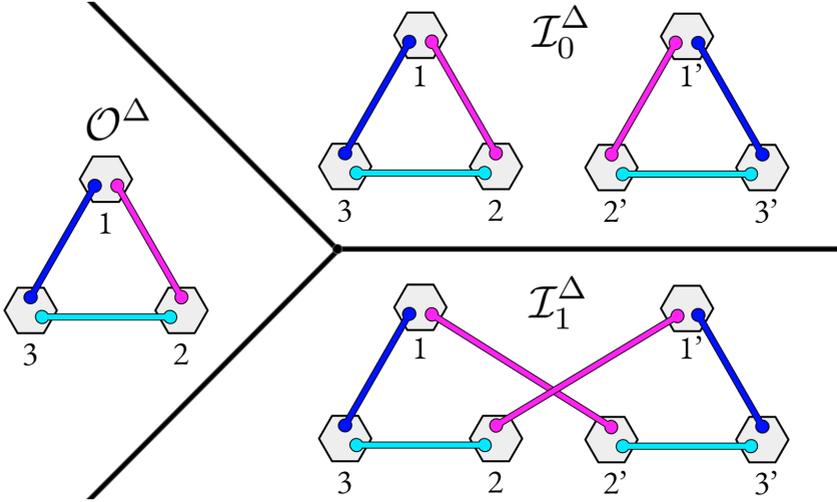
Figure 4.3: A fully connected tripartite quantum network $\mathcal{O}^\Delta$ with its two inflations $\mathcal{I}_0^\Delta$ and $\mathcal{I}_1^\Delta$. The "edges" represent bipartite states shared by the parties.

$q = 1$, so that we have

$$\sum_{k=0}^{2} \left(\langle g_1^k \rangle_{\mathcal{I}_0^\Delta} + \langle (Z_1^2 X_{2'})^k \rangle_{\mathcal{I}_0^\Delta} \right) \leq 3 + \sqrt{3}, \tag{4.43}$$

where $\langle \cdot \rangle_{\mathcal{I}_0^\Delta}$ denotes the expected value calculated with respect to any state that can be generated in $\mathcal{I}_0^\Delta$. For completeness, let us notice that such a state has the following form

$$\rho^{\mathcal{I}_0^\Delta} = \sum_\lambda p_\lambda \mathcal{E}_1^{(\lambda)} \otimes \mathcal{E}_2^{(\lambda)} \otimes \mathcal{E}_3^{(\lambda)} \otimes \mathcal{E}_{1'}^{(\lambda)} \otimes \mathcal{E}_{2'}^{(\lambda)} \otimes \mathcal{E}_{3'}^{(\lambda)} (\sigma^{\mathcal{I}_0^\Delta}), \tag{4.44}$$

where

$$\sigma^{\mathcal{I}_0^\Delta} = \sigma_{1,2} \otimes \sigma_{1,3} \otimes \sigma_{2,3} \otimes \sigma_{1',2'} \otimes \sigma_{1',3'} \otimes \sigma_{2',3'}. \tag{4.45}$$

Our goal now is to show that the assumption that the graph state $|G_\Delta\rangle$ can be generated in $\mathcal{O}^\Delta$ leads to a contradiction with inequality (4.43). We achieve it by proving that, under this assumption, every expected value in (4.43) must equal one.

Let us start with the first expected value in (4.43). Given that $|G_\Delta\rangle$ can be pre-

pared in $\mathcal{O}^{\triangle}$, it follows from Fact 4.2.1 that

$$\langle g_1^k \rangle_{\mathcal{I}_0^{\triangle}} = \langle g_1^k \rangle_{\mathcal{O}^{\triangle}} = 1. \tag{4.46}$$

This can also be seen by directly calculating the reduced state of $\rho^{\mathcal{I}_0^{\triangle}}$ corresponding to the non-primed parties, which is exactly the state that can be prepared in the original network. That is,

$$
\begin{aligned}
\rho_{1,2,3}^{\mathcal{I}_0^{\triangle}} &= \operatorname{Tr}_{1',2',3'}\left(\rho^{\mathcal{I}_0^{\triangle}}\right) \\
&= \sum_{\lambda} p_\lambda \mathcal{E}_1^{(\lambda)} \otimes \mathcal{E}_2^{(\lambda)} \otimes \mathcal{E}_3^{(\lambda)}(\sigma_{1,2} \otimes \sigma_{1,3} \otimes \sigma_{2,3}) = \rho^{\mathcal{O}^{\triangle}}. \tag{4.47}
\end{aligned}
$$

From this it follows that

$$\langle g_1^k \rangle_{\mathcal{I}_0^{\triangle}} = \operatorname{Tr}\left(g_1^k \rho^{\mathcal{I}_0^{\triangle}}\right) = \operatorname{Tr}\left(g_1^k \rho_{1,2,3}^{\mathcal{I}_0^{\triangle}}\right) = \operatorname{Tr}\left(g_1^k \rho^{\mathcal{O}^{\triangle}}\right), \tag{4.48}$$

where the second equality follows from the fact that g_1 acts only on the non-primed parties, whereas the third equality is a consequence of (4.47). We can then employ the assumption that the state generated by $\mathcal{O}^{\triangle}$ is our graph state, i.e., $\rho^{\mathcal{O}^{\triangle}} = |G_{\triangle}\rangle\langle G_{\triangle}|$, and therefore

$$\langle g_1^k \rangle_{\mathcal{I}_0^{\triangle}} = \operatorname{Tr}\left(g_1^k |G_{\triangle}\rangle\langle G_{\triangle}|\right) = \operatorname{Tr}\left(|G_{\triangle}\rangle\langle G_{\triangle}|\right) = 1, \tag{4.49}$$

As for the second expected value in (4.43), we consider another inflation $\mathcal{I}_1^{\triangle}$ of $\mathcal{O}^{\triangle}$ (cf. Fig. 4.3) and use Fact 4.2.1 to show that

$$\left\langle Z_1^2 X_{2'} \right\rangle_{\mathcal{I}_0^{\triangle}} = \langle g_2 \rangle_{\mathcal{I}_1^{\triangle}}, \tag{4.50}$$

where $\langle \cdot \rangle_{\mathcal{I}_1^{\triangle}}$ stands for an expected value calculated on the state that can be prepared in $\mathcal{I}_1^{\triangle}$. This relation follows from noticing that the two party reduced density matrix corresponding to nodes 1 and 2' of the state preparable in $\mathcal{I}_0^{\triangle}$ is the same as the one corresponding to the nodes 1 and 2 of the state that can be generated in $\mathcal{I}_1^{\triangle}$.

Let us now prove that $\langle g_2 \rangle_{\mathcal{I}_1^{\triangle}} = 1$. Since the nodes 1 and 2 are disconnected in $\mathcal{I}_1^{\triangle}$, but are connected in $\mathcal{O}^{\triangle}$ we cannot directly obtain this expected value from the original network $\mathcal{O}^{\triangle}$ and the state $|G_{\triangle}\rangle$. However, we can compute it indirectly by employing Fact 4.2.5. The idea is to link $\langle g_2 \rangle_{\mathcal{I}_1^{\triangle}}$ to two other expected values. If the latter are chosen appropriately then we will be able to calculate them in the same way as we calculated the value of $\langle g_1 \rangle_{\mathcal{I}_0^{\triangle}}$.

Notice first that g_2 and g_3 commute and give the identity when raised to the third power. Hence, the assumptions of Fact 4.2.5 are satisfied.

Furthermore, it follows from Fact 4.2.5 that if

$$\langle g_2 g_3 \rangle_{\mathcal{I}_1^\triangle} = \langle g_3 \rangle_{\mathcal{I}_1^\triangle} = 1, \tag{4.51}$$

it holds that $\langle g_2 \rangle_{\mathcal{I}_1^\triangle} = 1$.

The above observation allows us to shift the focus from $\langle g_2 \rangle_{\mathcal{I}_1^\triangle}$ to the expected values appearing in Eq. (4.51), that is, $\langle g_2 g_3 \rangle_{\mathcal{I}_1^\triangle}$ and $\langle g_3 \rangle_{\mathcal{I}_1^\triangle}$.

Now, it follows from Fact 4.2.1 that

$$\langle g_2 g_3 \rangle_{\mathcal{I}_1^\triangle} = \langle g_2 g_3 \rangle_{\mathcal{O}^\triangle} = 1, \tag{4.52}$$

as well as that

$$\langle g_3 \rangle_{\mathcal{I}_1^\triangle} = \langle g_3 \rangle_{\mathcal{O}^\triangle} = 1. \tag{4.53}$$

From this, we conclude $\langle g_2 \rangle_{\mathcal{I}_1^\triangle} = 1$, which also implies that $\left\langle Z_1^2 X_{2'} \right\rangle_{\mathcal{I}_0^\triangle} = 1$ by Eq. (4.50).

By the same argument it also holds that $\left\langle Z_1 X_{2'}^2 \right\rangle_{\mathcal{I}_0^\triangle} = 1$. Combining this with Eq. (4.46) implies that the left-hand side of (4.43) equals 6, leading to a contradiction. Thus, the graph state $|G_\triangle\rangle$ cannot be prepared in the network $\mathcal{O}^\triangle$.

We are now ready to present our main result that no graph states of arbitrary local prime dimension can be produced in LOSR quantum networks with bipartite sources, generalizing the results of Refs. [112, 113].

Theorem 4.2.7. *Consider a graph G with $N \geq 3$ vertices and where at least one vertex i has a neighbourhood $|\mathcal{N}_i| \geq 2$. The graph state $|G\rangle \in \mathbb{C}_d^{\otimes N}$, where d is prime, corresponding to a graph G cannot be generated in an LOSR N-partite quantum network with bipartite sources.*

For the proof of Theorem 4.2.7, it will be convenient to consider certain sets of graph states separately. This division into sets is done with the help of Lemma 4.2.8 below.

For the proof of this lemma we will recall the notion of local complementation. First, if a graph G can be transformed into a graph G' using a set of transformations called local complementations, then there exists a set of local unitaries U_i such that

$$|G\rangle = \bigotimes_{i=1}^N U_i |G'\rangle \tag{4.54}$$

(see e.g. Theorem 5 in [125]). The local complementations on the vertex n consist

of the transformations of a graph G to a graph G' such that

$$\Gamma'_{i,j} = \Gamma_{i,j} + a_n \Gamma_{i,n} \Gamma_{j,n}, \tag{4.55}$$

where $a_n \in \{0, \ldots, d-1\}$.

Lemma 4.2.8. *Every graph that has at least 3 vertices and has at least one vertex with a neighbourhood $|\mathcal{N}_i| \geq 2$ can be transformed using local complementations and relabelling into a graph G that fulfils $|\mathcal{N}_1 \setminus \mathcal{N}_2| \geq 2$, $\Gamma_{1,2} \neq 0$ and one of the four following sets of conditions:*

1. *$\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$,*

2. *$\mathcal{N}_1 \cap \mathcal{N}_2 \neq \emptyset$ and for all $n \in \mathcal{N}_1 \cap \mathcal{N}_2$ we have $\mathcal{N}_2 \setminus \{n\} \neq \mathcal{N}_n \setminus \{2\}$,*

3. *$|\mathcal{N}_1 \cap \mathcal{N}_2| = 1$ and there exists $n \in \mathcal{N}_1 \cap \mathcal{N}_2$ such that $\mathcal{N}_2 \setminus \{n\} = \mathcal{N}_n \setminus \{2\}$,*

4. *$|\mathcal{N}_1 \cap \mathcal{N}_2| \geq 2$, there exists $n \in \mathcal{N}_1 \cap \mathcal{N}_2$ such that $\mathcal{N}_2 \setminus \{n\} = \mathcal{N}_n \setminus \{2\}$, and there exists $a_n \in \{1, \ldots, d-1\}$ such that for all $i \in \mathcal{N}_n \setminus \{2\}$ we have $\Gamma_{2,i} + a_n \Gamma_{2,n} \Gamma_{n,i} = 0$.*

Proof. We assume that a graph G has at least 3 vertices and that at least one vertex has a neighbourhood $|\mathcal{N}_i| \geq 2$, and so, without loss of generality, we can take $|\mathcal{N}_1| \geq 2$ and $\Gamma_{1,2} \neq 0$. Then we act a_1 times with a local complementation on vertex 1, which gives us

$$\Gamma'_{2,n} = \Gamma_{2,n} + a_1 \Gamma_{1,2} \Gamma_{1,n}, \tag{4.56}$$

where $n \in \mathcal{N}_1$. We choose a_1 such that $\Gamma'_{2,n} = 0$, and since local complementation on 1 does not change $\Gamma_{1,2}$, it implies $|\mathcal{N}_1 \setminus \mathcal{N}_2| \geq 2$. If after this operation we have $\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$ then the first set of conditions in the lemma is fulfilled.

As for the case when $\mathcal{N}_1 \cap \mathcal{N}_2 \neq \emptyset$, the graph G fulfils the second set of conditions in the lemma if for all $n \in \mathcal{N}_1 \cap \mathcal{N}_2$ we have

$$\mathcal{N}_2 \setminus \{n\} \neq \mathcal{N}_n \setminus \{2\}. \tag{4.57}$$

Let us now consider a case where for some $n \in \mathcal{N}_1 \cap \mathcal{N}_2$ the above equation does not hold true. If $|\mathcal{N}_1 \cap \mathcal{N}_2| = 1$ then the third set of conditions in the lemma is fulfilled. Otherwise, let us consider a local complementation on the vertex n

$$\Gamma'_{2,i} = \Gamma_{2,i} + a_n \Gamma_{2,n} \Gamma_{n,i}, \tag{4.58}$$

where $i \in \mathcal{N}_n$. We assume that (4.57) does not hold for n, hence for every i we also have $i \in \mathcal{N}_2$. If one can find a_n such that $\Gamma'_{2,i} = 0$ for all $i \in \mathcal{N}_n$, then the fourth set of assumptions from the lemma is fulfilled.

If choosing such an a_n is not possible, then we can take a_n such that $\Gamma'_{1,2} \neq 0$ and such that there exists $i \in \mathcal{N}_2$ for which $\Gamma'_{2,i} = 0$. Clearly, after this transformation we get

$$|\mathcal{N}'_2| < |\mathcal{N}_2|. \tag{4.59}$$

Let us investigate the properties of this transformed graph. First of all $\Gamma_{1,2} \neq 0$, since that was the condition on a_n. Next, the local complementation on the vertex n by definition does not change $\Gamma_{i,n}$ for all i, so in particular we still have $\Gamma_{1,n} \neq 0$ and $\Gamma_{2,n} \neq 0$ which implies $\mathcal{N}_1 \cap \mathcal{N}_2 \neq \emptyset$. Last but not least, the action of a local complementation on vertex $n \in \mathcal{N}_1 \cap \mathcal{N}_2$ cannot decrease the size of a set $\mathcal{N}_1 \setminus \mathcal{N}_2$, which means that the assumption $|\mathcal{N}_1 \setminus \mathcal{N}_2| \geq 2$ still holds.

To see why the last claim holds true, notice that to decrease the size of a set $\mathcal{N}_1 \setminus \mathcal{N}_2$ we would have to have $\Gamma_{1,j} \neq 0$ and $\Gamma_{2,j} = 0$ for some j before the action of the local complementation on n, and after said action either $\Gamma'_{1,j} = 0$ or $\Gamma'_{2,j} = 0$. However, from the assumption $\mathcal{N}_2 \setminus \{n\} \neq \mathcal{N}_n \setminus \{2\}$ it follows that $\Gamma_{n,j} = 0$. Crucially, this implies that under the action of local complementation on n we have

$$\Gamma'_{j,k} = \Gamma_{j,k} \tag{4.60}$$

for all k, hence $\Gamma'_{1,j} = \Gamma_{1,j}$ and $\Gamma'_{2,j} = \Gamma_{2,j}$.

The key observation here is that the new graph created after action of local complementation fulfils $|\mathcal{N}_1 \setminus \mathcal{N}_2| \geq 2$, $\Gamma_{1,2} \neq 0$ and $\mathcal{N}_1 \cap \mathcal{N}_2 \neq \emptyset$, and so we can check if the new graph fulfils conditions 2, 3 or 4 from the lemma and if they are not fulfilled, then we again create a new graph by action of local complementation on some $\tilde{n} \in \mathcal{N}_1 \cap \mathcal{N}_2$. Every time we act with this local complementation we get (4.59), which implies that after a finite amount of repetition of this procedure, we will get $|\mathcal{N}_1 \cap \mathcal{N}_2| = 1$ which fulfils either the second or third condition from the lemma. □

With that, we are finally ready to prove Theorem 4.2.7. The proof is rather technical, so let us briefly describe the key idea. Our goal is to find a contradiction to Lemma 4.2.2. In order to do that, we construct a series of implications between expectation values of different inflations, using Facts 4.2.1 and 4.2.5 repeatedly. In one of those inflations we will be able to directly link the expectation value of an operator to the expectation value of a stabilizer in the original network. When we assume that we have a graph state, this leads us to conclude that the expectation value equals 1. Applying this trick to both operators in the statement of Lemma 4.2.2, we reach a total of $2d$ on the left hand side, which is a contradiction. This leads us to conclude that the state we started with could not have been a graph state, completing the proof.

Proof (of Theorem 4.2.7). We assume that the graph G consists of at least three ver-
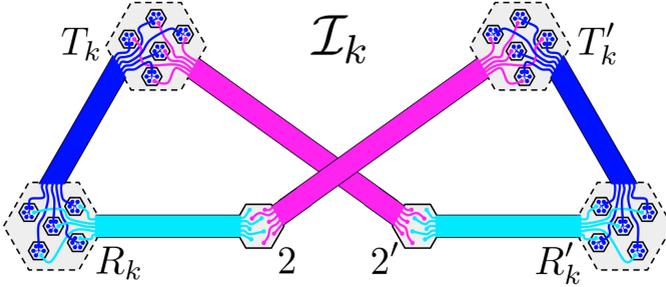
Figure 4.4: Inflations $\mathcal{I}_k$. Here T_k, T'_k, R_k, R'_k are not individual parties, but rather sets of parties. Every party from a set is connected to all other parties from that set and if two sets are connected then every party from one set is connected to every party from the other set.

tices and has at least one vertex i such that $|\mathcal{N}_i| \geq 2$. So by virtue of Lemma 4.2.8 by using local complementations and relabeling we can transform G into a graph that fulfills one of the four sets of conditions. Acting with a local complementation on a graph G corresponds to acting with local unitaries on a graph state $|G\rangle$, and so it is sufficient to only consider graphs that fulfill one of the four sets of conditions in Lemma 4.2.8.

Since every inflation in this proof has a very similar structure, let us begin by discussing these similarities in order to make the proof easier to follow. All inflations $\mathcal{I}_k^l$ that we consider in this proof share two assumptions. First, every $\mathcal{I}_k^l$ consists of exactly two copies of every party, denoted i and i', and two copies of every source from $\mathcal{O}$. Second, a non-primed party i (for $i \neq 2$) is connected to every other non-primed party j (for $j \neq 2$) and to either 2 or $2'$. Likewise, every primed party i' (for $i' \neq 2'$) has to be connected to every other primed party j' (for $j' \neq 2'$) and to either 2 or $2'$. These two assumptions allow us to precisely describe an inflation just by defining the set of non-primed parties connected to $2'$, which we denote by T_k^l. For a graphical description, see Fig. 4.4.

We structure this proof in the following way: in *Part 1* we consider graphs G fulfilling the first or the second set of conditions from Lemma 4.2.8, in *Part 2* we consider graphs G fulfilling the third set of conditions and in *Part 3* we consider graphs fulfilling the fourth set of conditions.

Part 1. Here, we assume that a graph G fulfils the first or second set of conditions from Lemma 4.2.8. Both these sets of conditions imply the same structure of the first

two stabilizing operators, namely:

$$
\begin{aligned}
g_1 &= X_1 Z_2^{\Gamma_{1,2}} Z_{\mathcal{N}_1 \setminus \{2\}}, \\
g_2 &= Z_1^{\Gamma_{1,2}} X_2 Z_{\mathcal{N}_2 \setminus \{1\}},
\end{aligned}
\tag{4.61}
$$

where

$$
Z_{\mathcal{N}_i \setminus \{j\}} = \prod_{k \in \mathcal{N}_i \setminus \{j\}} Z_k^{\Gamma_{i,k}}.
\tag{4.62}
$$

Let us consider an inflation $\mathcal{I}_0^1$ for which

$$
T_0^1 = \left(\mathcal{N}_1 \setminus (\mathcal{N}_2 \cup \{2\}) \right) \cup \{1\}.
\tag{4.63}
$$

By virtue of Lemma 4.2.2, the following holds true:

$$
\sum_{k=0}^{d-1} \langle \tilde{g}_1^k \rangle_{\mathcal{I}_0^1} + \langle g_2^k \rangle_{\mathcal{I}_0^1} \leq d + \sqrt{d},
\tag{4.64}
$$

where $\tilde{g}_1 = X_1 Z_{2'}^{\Gamma_{1,2}} Z_{\mathcal{N}_1 \setminus \{2\}}$. To complete the proof we will show that the above inequality is violated, since if one could generate the aforementioned graph state in the network $\mathcal{O}$, then the above operators stabilize the state generated in the inflation $\mathcal{I}_0^1$. The strategy is the same as the one we applied in Example 4.2.6.

We begin our analysis with the operator $\tilde{g}_1$. To show that this operator is a stabilizing operator, let us consider another inflation $\mathcal{I}_1^1$ for which

$$
T_1^1 = \mathcal{N}_1 \cap \mathcal{N}_2.
\tag{4.65}
$$

Notice that the union

$$
T_0^1 \cup T_1^1 \cup \{2'\} = (\mathcal{N}_1 \setminus \{2\}) \cup \{1, 2'\}
\tag{4.66}
$$

describes a set of parties in a $\tilde{g}_1$ subnetwork of $\mathcal{I}_0^1$. In this subnetwork, every party from the set T_0^1 is connected to $2'$ and to T_1^1, but $2'$ and T_1^1 are disconnected. Similarly, in the inflation $\mathcal{I}_1^1$, every party from the set T_0^1 is connected to 2 and to T_1^1, but 2 and T_1^1 are disconnected, hence by Fact 4.2.1 we have

$$
\langle \tilde{g}_1 \rangle_{\mathcal{I}_0^1} = \langle g_1 \rangle_{\mathcal{I}_1^1}.
\tag{4.67}
$$

The first set of assumptions from Lemma 4.2.8 implies $T_1^1 = \emptyset$ which gives us $\langle g_1 \rangle_{\mathcal{I}_1^1} = \langle g_1 \rangle_{\mathcal{O}} = 1$, since we assume that we generate a graph state $|G\rangle$ in the

original network $\mathcal{O}$. Of course, for the second set of assumptions we have $T_1^1 \neq \emptyset$, and so the same implication does not hold. Therefore, for the next few paragraphs we only consider the second set of assumptions from Lemma 4.2.8 and we will return to considering both after showing that $\langle g_1 \rangle_{\mathcal{I}_1^1} = \langle g_1 \rangle_{\mathcal{O}} = 1$ holds also for the second set of assumptions.

One of the assumptions from the second set of assumptions from Lemma 4.2.8 states that for all $n \in \mathcal{N}_1 \cap \mathcal{N}_2$ we have

$$\mathcal{N}_2 \setminus \{n\} \neq \mathcal{N}_n \setminus \{2\}. \tag{4.68}$$

Consequently, given two stabilizing operators g_2 and g_n there has to exist a qubit i such that $i \neq 2, n$ and $g_2^{(i)} \neq g_n^{(i)}$, where $g_k^{(i)}$ denotes a matrix of g_k acting on the subsystem corresponding to the party i. Utilising the symmetry: $g_i^{(j)} = g_j^{(i)}$, we get two possibilities: either

$$g_i^{(2)} = \mathbb{1}, \qquad g_i^{(n)} = Z^{\Gamma_{i,n}} \tag{4.69}$$

or

$$g_i^{(2)} = Z^{\Gamma_{2,i}}, \qquad g_i^{(n)} = \mathbb{1}, \tag{4.70}$$

where $\Gamma_{i,n} \neq 0$ and $\Gamma_{2,i} \neq 0$ respectively. In what follows, we assume that Eq. (4.69) is true, and we will later come back to the case of Eq. (4.70).

Using Fact 4.2.5 we have the implication

$$\langle g_i^{-l} \rangle_{\mathcal{I}_1^1} = \langle g_1 g_i^l \rangle_{\mathcal{I}_1^1} = 1 \quad \Rightarrow \quad \langle g_1 \rangle_{\mathcal{I}_1^1} = 1, \tag{4.71}$$

for all $l \in \{1, \ldots, d-1\}$.

From (4.69) it follows that the g_i^{-l} subnetwork of $\mathcal{I}_1^1$ does not contain party 2. Recall that we choose inflations such that every non-primed party is connected to every other party with an exception of party 2, meaning that if a subnetwork does not contain the party 2 or $2'$ then it is fully connected, which is exactly the case for the g_i^{-l} subnetwork of $\mathcal{I}_1^1$. Consequently, from Fact 4.2.1 it follows that

$$\langle g_i^{-l} \rangle_{\mathcal{I}_1^1} = \langle g_i^{-l} \rangle_{\mathcal{O}} = 1, \tag{4.72}$$

where in the second equality we used the assumption that we generate a graph state on the network $\mathcal{O}$. As for the expected value $\langle g_1 g_i^l \rangle_{\mathcal{I}_1^1}$, in order to calculate it, we first have to introduce another inflation $\mathcal{I}_2^1$ for which

$$T_2^1 = (\mathcal{N}_1 \cap \mathcal{N}_2) \setminus \{n\}. \tag{4.73}$$

Notice, that the only difference between inflations $\mathcal{I}_1^1$ and $\mathcal{I}_2^1$ is that in the former,

parties $2'$ and n are connected and in the latter 2 and n. Crucially, since d is prime and $\Gamma_{i,n} \neq 0$, we can always take

$$l = -\Gamma_{1,n}\Gamma_{i,n}^{-1}, \tag{4.74}$$

where we take the inverse and the negation to be operations in $\mathbb{Z}_d$, i.e., $l \in \{1, \ldots, d-1\}$. This implies that the $g_1 g_i^l$ subnetwork of $\mathcal{I}_1^1$ does not contain party n, therefore by Fact 4.2.1 we have

$$\langle g_1 g_i^l \rangle_{\mathcal{I}_1^1} = \langle g_1 g_i^l \rangle_{\mathcal{I}_2^1}. \tag{4.75}$$

Lastly, by using Fact 4.2.5 once more we can show that if $\langle g_1 \rangle_{\mathcal{I}_2^1} = \langle g_i^l \rangle_{\mathcal{I}_2^1} = 1$ then $\langle g_1 g_i^l \rangle_{\mathcal{I}_2^1} = 1$. Furthermore, since g_i^l subnetwork of $\mathcal{I}_2^1$ does not contain party 2, we have $\langle g_i^l \rangle_{\mathcal{I}_2^1} = \langle g_i^l \rangle_{\mathcal{O}} = 1$, and so:

$$\langle g_1 \rangle_{\mathcal{I}_2^1} = 1 \;\Rightarrow\; \langle g_1 g_i^l \rangle_{\mathcal{I}_2^1} = 1 \;\Rightarrow\; \langle g_1 \rangle_{\mathcal{I}_1^1} = 1. \tag{4.76}$$

As for the case when Eq. (4.70) is true, the proof has mostly the same structure, but with g_i^{-l} and $g_1 g_i^l$ swapped. The only other difference is that, using the fact that d is prime and $\Gamma_{2,i} \neq 0$, we take $l \in \mathbb{Z}_d$ such that

$$l = -\Gamma_{1,2}\Gamma_{2,i}^{-1}. \tag{4.77}$$

As a result, we have the same implication as in the first case:

$$\langle g_1 \rangle_{\mathcal{I}_2^1} = 1 \;\Rightarrow\; \langle g_1 \rangle_{\mathcal{I}_1^1} = 1. \tag{4.78}$$

Since both cases result in the above implication, we can now take a different n and apply the same procedure to produce an implication

$$\langle g_1 \rangle_{\mathcal{I}_3^1} = 1 \;\Rightarrow\; \langle g_1 \rangle_{\mathcal{I}_2^1} = 1, \tag{4.79}$$

where $T_3^1 = (\mathcal{N}_1 \cap \mathcal{N}_2) \setminus \{n, n'\}$. Repeating this procedure for all $n \in T_1$ results in a chain of implications

$$\langle g_1 \rangle_{\mathcal{I}_q^1} = 1 \;\Rightarrow\; \langle g_1 \rangle_{\mathcal{I}_{q-1}^1} = 1 \;\Rightarrow\; \cdots \;\Rightarrow\; \langle g_1 \rangle_{\mathcal{I}_1^1} = 1, \tag{4.80}$$

where $q = |\mathcal{N}_1 \cap \mathcal{N}_2| + 1$. Finally, for $\mathcal{I}_q^1$ we have $T_q^1 = \emptyset$, and so

$$\langle g_1 \rangle_{\mathcal{I}_q^1} = \langle g_1 \rangle_{\mathcal{O}} = 1. \tag{4.81}$$

This is the first statement in the chain of implications (4.80), and so

$$\langle g_1 \rangle_{\mathcal{I}_1^1} = 1, \tag{4.82}$$

which, by the virtue of Eq. (4.67) implies

$$\langle \tilde{g}_1 \rangle_{\mathcal{I}_0^1} = 1. \tag{4.83}$$

As a direct implication of the above, we have that $\tilde{g}_1$ is a stabilizing operator of the state on $\mathcal{I}_0^1$.

We now come back to the inequality (4.64) and show that g_2 is also a stabilizing operator of a state generated in the inflation $\mathcal{I}_0^1$. We again consider both the first and second set of assumption from Lemma 4.2.8. The g_2 subnetwork of $\mathcal{I}_0^1$ is a fully connected subnetwork, apart from the lacking connection between parties 1 and 2. Therefore, from Fact 4.2.1 we have that

$$\langle g_2 \rangle_{\mathcal{I}_0^1} = \langle g_2 \rangle_{\mathcal{I}_{-1}^1}, \tag{4.84}$$

where $\mathcal{I}_{-1}^1$ is an inflation with

$$T_{-1}^1 = \{1\}. \tag{4.85}$$

From $|\mathcal{N}_1 \setminus \mathcal{N}_2| \geq 2$ it follows that there exists $i \in \mathcal{N}_1$ for which

$$g_i^{(1)} = Z^{\Gamma_{1,i}}, \qquad g_i^{(2)} = \mathbb{1}, \tag{4.86}$$

where we used the symmetry $g_i^{(j)} = g_j^{(i)}$. Next, using Fact 4.2.5 once again, we have

$$\langle g_i^{-l} \rangle_{\mathcal{I}_{-1}^1} = \langle g_2 g_i^l \rangle_{\mathcal{I}_{-1}^1} = 1 \;\Rightarrow\; \langle g_2 \rangle_{\mathcal{I}_{-1}^1} = 1, \tag{4.87}$$

where we take $l \in \mathbb{Z}_d$ such that

$$l = -\Gamma_{1,2}\Gamma_{1,i}^{-1}. \tag{4.88}$$

Notice, that $\langle g_2 g_i^l \rangle_{\mathcal{I}_{-1}^1}$ has a trivial action on the first party and $\langle g_i^{-l} \rangle_{\mathcal{I}_{-1}^1}$ has a trivial action on the second party meaning that the $g_2 g_i^l$ and g_i^{-l} subnetworks of $\mathcal{I}_{-1}^1$ are fully connected which by Fact 4.2.1 implies

$$\langle g_i^{-l} \rangle_{\mathcal{I}_{-1}^1} = \langle g_i^{-l} \rangle_{\mathcal{O}} = 1, \qquad \langle g_2 g_i^l \rangle_{\mathcal{I}_{-1}^1} = \langle g_2 g_i^l \rangle_{\mathcal{O}} = 1, \tag{4.89}$$

and so by virtue of Eq. (4.84) and Eq. (4.87) we have

$$\langle g_2 \rangle_{\mathcal{I}_0^1} = 1, \tag{4.90}$$

i.e., g_2 is a stabilizing operator of the state generated in $\mathcal{I}_0^1$.

As a consequence of the above result and (4.82), the inequality (4.64) is violated, meaning that our assumption must have been wrong: graph states corresponding to a graph that admit the first or the second set of assumptions from Lemma 4.2.8 cannot be generated in a network with bipartite sources.

Part 2. We now assume that the graph G fulfills the third set of conditions from Lemma 4.2.8. In this part we will only use three stabilizing operators, namely

$$
\begin{aligned}
g_1 &= X_1 Z_2^{\Gamma_{1,2}} Z_n^{\Gamma_{1,n}} Z_{\mathcal{N}_1 \setminus \{2,n\}}, \\
g_2 &= Z_1^{\Gamma_{1,2}} X_2 Z_n^{\Gamma_{2,n}} Z_{\mathcal{N}_2 \setminus \{1,n\}}, \\
g_n &= Z_1^{\Gamma_{1,n}} Z_2^{\Gamma_{2,n}} X_n Z_{\mathcal{N}_n \setminus \{1,2\}},
\end{aligned}
\tag{4.91}
$$

where $\mathcal{N}_1 \cap \mathcal{N}_2 = \{n\}$. Let us consider an inflation $\mathcal{I}_0^2$ for which $T_0^2 = \mathcal{N}_2 \setminus \{1,n\}$. We can use Lemma 4.2.2 to construct an inequality:

$$
\sum_{k=0}^{d-1} \left\langle s^k \right\rangle_{\mathcal{I}_0^2} + \left\langle g_1^k \right\rangle_{\mathcal{I}_0^2} \le d + \sqrt{d},
\tag{4.92}
$$

where

$$
s = X_{2'} Z_{2'}^{l\Gamma_{2,n}} Z_n^{\Gamma_{2,n}} X_n^l Z_{\mathcal{N}_2 \setminus \{1,n\}}^l Z_{\mathcal{N}_n \setminus \{1,2\}}^l,
\tag{4.93}
$$

and we take $l \in \mathbb{Z}_d$ such that

$$
l = -\Gamma_{1,2} \Gamma_{1,n}^{-1}.
\tag{4.94}
$$

From the assumption $|\mathcal{N}_1 \cap \mathcal{N}_2| = 1$ we can infer that the g_1^k subnetwork of $\mathcal{I}_0^2$ does not contain any party from the set T_0^2, hence we can use Fact 4.2.1 and our assumption about the generation of a graph state in the network $\mathcal{O}$ to get

$$
\langle g_1^k \rangle_{\mathcal{I}_0^2} = \langle g_1^k \rangle_{\mathcal{O}} = 1.
\tag{4.95}
$$

Thus, for all k, g_1^k is a stabilizing operator of a state generated in $\mathcal{I}_0^2$.

In order to tackle the second operator from Eq. (4.92), introduce another inflation $\mathcal{I}_1^2$ with $T_1^2 = \{n\}$. From Fact 4.2.1 it follows that

$$
\langle s \rangle_{\mathcal{I}_0^2} = \left\langle g_2 g_n^l \right\rangle_{\mathcal{I}_1^2}.
\tag{4.96}
$$

The above can be shown by comparing the two operators and the corresponding subnetworks of $\mathcal{I}_0^2$ and $\mathcal{I}_1^2$: the former is fully-connected apart from the lacking connection between $2'$ and n, and the latter is fully-connected apart from the lacking connection between 2 and n.

By virtue of Fact 4.2.5

$$\langle g_1^{-m} g_2 \rangle_{\mathcal{I}_1^2} = \langle g_1^m g_n^l \rangle_{\mathcal{I}_1^2} = 1 \quad \Rightarrow \quad \langle g_2 g_n^l \rangle_{\mathcal{I}_1^2} = 1, \tag{4.97}$$

where we take $m \in \mathbb{Z}_d$ to be

$$m = \Gamma_{2,n} \Gamma_{1,n}^{-1}. \tag{4.98}$$

A quick look at the $g_1^{-m} g_2$ and $g_1^m g_n^l$ subnetworks of $\mathcal{I}_1^2$ reveals that both of them are fully-connected, since they do not contain party n and party 2 respectively, hence from Fact 4.2.1 we have

$$\langle g_1^{-m} g_2 \rangle_{\mathcal{I}_1^2} = \langle g_1^{-m} g_2 \rangle_{\mathcal{O}} = 1, \langle g_1^m g_n^l \rangle_{\mathcal{I}_1^2} = \langle g_1^m g_n^l \rangle_{\mathcal{O}} = 1, \tag{4.99}$$

However, the above implies

$$\langle s \rangle_{\mathcal{I}_0^2} = 1, \tag{4.100}$$

and so, as was the case with g_1, the above operator is a stabilizing operator of a state generated in $\mathcal{I}_0^2$. Therefore, the inequality (4.92) is violated, which implies that the graph state corresponding to a graph that fulfills the third set of conditions from Lemma 4.2.8 cannot be generated in the considered quantum network.

Part 3. In this part, we consider graphs G that fulfill the fourth set of conditions from Lemma 4.2.8. The stabilizing operators used in this part of the proof have the same general structure as the ones we have used in the previous part (see Eq. (4.91)).

Let us consider an inflation $\mathcal{I}_0^3$ for which $T_0^3 = \emptyset$. Using Lemma 4.2.2 we can construct the following inequality:

$$\sum_{k=0}^{d-1} \left\langle \left(X_{2'} Z_{2'}^{l\Gamma_{2,n}} Z_n^{\Gamma_{2,n}} X_n^l \right)^k + g_1^k \right\rangle_{\mathcal{I}_0^3} \le d + \sqrt{d}, \tag{4.101}$$

where $l \in \mathbb{Z}_d$. From Fact 4.2.1 and from the assumption that we can generate a graph state $|G\rangle$ in the network $\mathcal{O}$ it follows that

$$\langle g_1 \rangle_{\mathcal{I}_0^3} = \langle g_1 \rangle_{\mathcal{O}} = 1, \tag{4.102}$$

and so g_1 is a stabilizing operator of a state generated in $\mathcal{I}_0^3$. Next, let us consider an inflation $\mathcal{I}_1^3$ for which $T_1^3 = \{n\}$. It quickly follows from Fact 4.2.1 that

$$\left\langle X_{2'} Z_{2'}^{l\Gamma_{2,n}} Z_n^{\Gamma_{2,n}} X_n^l \right\rangle_{\mathcal{I}_0^3} = \left\langle X_2 Z_2^{l\Gamma_{2,n}} Z_n^{\Gamma_{2,n}} X_n^l \right\rangle_{\mathcal{I}_1^3}. \tag{4.103}$$

Furthermore, one of the assumptions states that there exists $a_n \in \{1, \ldots, d-1\}$

such that for all $i \in \mathcal{N}_n \setminus \{2\}$ we have

$$\Gamma_{2,i} + a_n \Gamma_{2,n} \Gamma_{n,i} = 0. \tag{4.104}$$

Taking $l = a_n \Gamma_{2,n}$ we get

$$\left\langle X_2 Z_2^{l\Gamma_{2,n}} Z_n^{\Gamma_{2,n}} X_n^l \right\rangle_{\mathcal{I}_1^3} = \left\langle g_2 g_n^l \right\rangle_{\mathcal{I}_1^3}. \tag{4.105}$$

Now, we can use Fact 4.2.5 to construct an implication

$$\langle g_1^m g_2 \rangle_{\mathcal{I}_1^3} = \left\langle g_1^{-m} g_n^l \right\rangle_{\mathcal{I}_1^3} = 1 \quad \Rightarrow \quad \left\langle g_2 g_n^l \right\rangle_{\mathcal{I}_1^3} = 1, \tag{4.106}$$

where we take $m \in \mathbb{Z}_d$ such that

$$m = -\Gamma_{2,n} \Gamma_{1,n}. \tag{4.107}$$

It follows that the $g_1^m g_2$ subnetwork of $\mathcal{I}_1^3$ does not contain party n and that the $g_1^{-m} g_n^l$ subnetwork of $\mathcal{I}_1^3$ does not contain party 2. Therefore, both of these subnetworks are connected, which by the virtue of Fact 4.2.1 implies

$$\langle g_1^m g_2 \rangle_{\mathcal{I}_1^3} = \langle g_1^m g_2 \rangle_{\mathcal{O}} = 1, \tag{4.108}$$

$$\left\langle g_1^{-m} g_n^l \right\rangle_{\mathcal{I}_1^3} = \left\langle g_1^{-m} g_n^l \right\rangle_{\mathcal{O}} = 1. \tag{4.109}$$

Consequently, both operators in (4.101) stabilize the state generated in $\mathcal{I}_0^3$ and so the inequality (4.101) is violated which ends the proof. $\qquad\square$

4.2.3 Robustness

Using a standard continuity argument one can extend the above result to any state that is sufficiently close to a graph state. Indeed, recalling the definition of the fidelity $F(\rho, \sigma)$ between two states, where one of the states is a pure state $\sigma = |\psi\rangle\langle\psi|$:

$$F(\rho, |\psi\rangle\langle\psi|) = \langle\psi|\rho|\psi\rangle, \tag{4.110}$$

we can formulate the following Theorem.

Theorem 4.2.9. *Let us consider a state ρ and a graph state $|G\rangle$. Moreover, let $q = |\mathcal{N}_1 \cap \mathcal{N}_2| + 1$ for graphs G that fulfill (4.68) and $q = 1$ in other cases. If*

$$F(\rho, |G\rangle\langle G|) > 1 - \frac{1}{8}\left(\beta^2 + 2\gamma - \beta\sqrt{\beta^2 + 4\gamma}\right), \tag{4.111}$$

where $\beta = 2q - 1$ and $\gamma = (d - \sqrt{d})/(d - 1)$, *then ρ cannot be generated in an LOSR network with bipartite sources.*

Proof. Rewriting a state ρ in an orthogonal basis that contains $|\psi\rangle$ and diagonalising the subspace orthogonal to $|\psi\rangle$, one gets

$$\rho = (1 - \delta)\,|\psi\rangle\langle\psi| + \sum_{i=1}^{d} c_i\,|f_i\rangle\langle f_i| + \sum_{j=1}^{d}(a_{0,j}\,|\psi\rangle\langle f_j| + a_{0,j}^*\,|f_j\rangle\langle\psi|), \quad (4.112)$$

where $\sum_{i=1}^{d} c_i = \delta$, so that $\mathrm{tr}(\rho) = 1$. Note that this is always possible, since ρ is Hermitian and thus so is every principal submatrix of ρ. This then yields

$$F(\rho, |\psi\rangle\langle\psi|) = 1 - \delta. \quad (4.113)$$

We can express every state with fidelity $1 - \delta$ with respect to a graph state $|\psi\rangle = |G\rangle$ in the form of Eq. (4.112).

In order to find our desired bound on the fidelity of ρ we will need two facts. First, given a stabilizing operator s of the graph state $|G\rangle$, we have

$$\left|\langle\mathbb{1} - s\rangle_\rho\right| = \left|1 - (1 - \delta)\,\langle s\rangle_{|G\rangle} - \sum_i \mathrm{tr}(c_i s\,|f_i\rangle\langle f_i|)\right.$$
$$\left. - \sum_j \mathrm{tr}(a_{0,j} s\,|G\rangle\langle f_i| + a_{0,j}^*\,|f_i\rangle\langle G|\,s)\right| \quad (4.114)$$
$$\leq \delta + \sum_i c_i = 2\delta.$$

Second, given two unitary operators s_1 and s_2 such that $|\langle\mathbb{1} - s_1\rangle| \leqslant \mu, |\langle\mathbb{1} - s_2\rangle| \leqslant \nu$, and $\mu \leqslant \nu$ the following holds true:

$$|\langle\mathbb{1} - s_1 s_2\rangle| = |\langle(\mathbb{1} - s_1)s_2 + \mathbb{1} - s_2\rangle|$$
$$\leq \left|\sqrt{\langle(\mathbb{1} - s_1)(\mathbb{1} - s_1^\dagger)\rangle\langle s_2^\dagger s_2\rangle}\right| + |\langle\mathbb{1} - s_2\rangle|$$
$$\leq \left|\sqrt{\langle 2\mathbb{1} - s_1 - s_1^\dagger\rangle}\right| + \nu \quad (4.115)$$
$$\leq \sqrt{2\mu} + \nu,$$

where we have used the triangle inequality and the Cauchy-Schwarz inequality.

The next step is to assess how such a deviation from a unit expected value prop-

agates in the reasoning of the proof of Theorem 4.2.7. As that proof is divided into three parts, we have to consider each of them separately.

We begin with *Part 1* of the proof of Theorem 4.2.7. Unlike in the proof of Theorem 4.2.7, here we are forced to start from the "end" and consider the inflation $\mathcal{I}_q^1$ first. From Eq. (4.114) it follows that

$$|\langle \mathbb{1} - g_1 \rangle_{\mathcal{I}_q^1}| \leqslant 2\delta, \qquad |\langle \mathbb{1} - g_i^l \rangle_{\mathcal{I}_q^1}| \leqslant 2\delta, \tag{4.116}$$

where we assume that Eq. (4.69) holds true for $n \in T_{q-1}$. The case when Eq. (4.70) is true gives the same result, and so for simplicity we will only focus on Eq. (4.69).

From Eq. (4.115) it follows that

$$|\langle \mathbb{1} - g_1 g_i^l \rangle_{\mathcal{I}_q^1}| \leqslant 2\sqrt{\delta} + 2\delta. \tag{4.117}$$

From the proof of Theorem 4.2.7 we have that

$$\langle \mathbb{1} - g_1 g_i^l \rangle_{\mathcal{I}_q^1} = \langle \mathbb{1} - g_1 g_i^l \rangle_{\mathcal{I}_{q-1}^1}, \tag{4.118}$$

and so we also have

$$|\langle \mathbb{1} - g_1 g_i^l \rangle_{\mathcal{I}_{q-1}^1}| \leqslant 2\sqrt{\delta} + 2\delta. \tag{4.119}$$

For inflation $\mathcal{I}_{q-1}^1$ we also have

$$|\langle \mathbb{1} - g_i^{-l} \rangle_{\mathcal{I}_{q-1}^1}| \leqslant 2\delta. \tag{4.120}$$

We can again use Eq. (4.115) for $s_1 = g_i^{-l}$ and $s_2 = g_1 g_i^l$ which yields

$$|\langle \mathbb{1} - g_1 \rangle_{\mathcal{I}_{q-1}^1}| \leqslant 4\sqrt{\delta} + 2\delta. \tag{4.121}$$

Applying this procedure $q - 1$ times, where $q = |\mathcal{N}_1 \cap \mathcal{N}_2| + 1$, gives us

$$\left|\langle \mathbb{1} - \tilde{g}_1 \rangle_{\mathcal{I}_0^1}\right| \leqslant 4(q-1)\sqrt{\delta} + 2\delta, \tag{4.122}$$

where $\tilde{g}_1 = X_1 Z_{2'} Z_{\mathcal{N}_1 \setminus \{2\}}$. As this proof can be performed for any power $k \in \{1, \ldots, d-1\}$ it holds more generally that

$$\left|\langle \mathbb{1} - \tilde{g}_1^k \rangle_{\mathcal{I}_0^1}\right| \leqslant 4(q-1)\sqrt{\delta} + 2\delta. \tag{4.123}$$

Similarly, one can show that

$$\left|\langle \mathbb{1} - g_2^k \rangle_{\mathcal{I}_0^1}\right| \leqslant 2\sqrt{\delta} + 2\delta. \tag{4.124}$$

We can now use the above inequalities to find for which values of δ one still breaks the bounds of Lemma 4.2.2:

$$2(d-1) - \left| \sum_{k=1}^{d-1} \langle \tilde{g}_1^k \rangle_{\mathcal{I}_0^1} + \langle g_2^k \rangle_{\mathcal{I}_0^1} \right| \leq \left| \sum_{k=1}^{d-1} \langle \mathbb{1} - \tilde{g}_1^k \rangle_{\mathcal{I}_0^1} + \langle \mathbb{1} - g_2^k \rangle_{\mathcal{I}_0^1} \right| \quad (4.125)$$

$$\leq 2(d-1) \left[(2q-1)\sqrt{\delta} + 2\delta \right], \quad (4.126)$$

so that

$$\left| \sum_{k=1}^{d-1} \langle \tilde{g}_1^k \rangle_{\mathcal{I}_0^1} + \langle g_2^k \rangle_{\mathcal{I}_0^1} \right| \geq 2(d-1) \left[1 - (2q-1)\sqrt{\delta} - 2\delta \right]. \quad (4.127)$$

Notice, that the $k = 0$ terms in the inequality of Lemma 4.2.2 always give exactly one which can be subtracted from both sides of an inequality. Hence, the bound of $d + \sqrt{d} - 2$ is violated if

$$(2q-1)\sqrt{\delta} + 2\delta < \frac{d - \sqrt{d}}{2(d-1)}. \quad (4.128)$$

Solving this inequality yields

$$\delta < \frac{1}{8} \left(\beta^2 + 2\gamma - \beta\sqrt{\beta^2 + 4\gamma} \right), \quad (4.129)$$

where $\beta = 2q - 1$ and $\gamma = (d - \sqrt{d})/(d - 1)$.

As for *Part 2* and *Part 3* of the proof of Theorem 4.2.7, the derivation of the fidelity bound is relatively simple as compared to the case of *Part 1*, and so we will only show the result. For both *Part 2* and *Part 3* the inequality is violated if

$$(\sqrt{\delta} + 2\delta) < \frac{d - \sqrt{d}}{2(d-1)}, \quad (4.130)$$

which yields

$$\delta < \frac{1}{8}(1 + 2\gamma - \sqrt{1 + 4\gamma}), \quad (4.131)$$

where $\gamma = (d - \sqrt{d})/(d - 1)$. Notice that this bound corresponds to the bound (4.129) for $q = 1$, and so the case where $\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$. Coincidentally, the left hand side of Eq. (4.129) decreases with increasing q, and so Eq. (4.131) is also an optimal case of Eq. (4.129). $\qquad \square$

As an example, let us consider the graph state corresponding to the graph presented on Fig. 4.2. Here $q = 1$, $d = 3$, and so (4.111) simplifies to $F(\rho, |G\rangle) > 0.952$. Interestingly, the bound (4.111) can be relaxed by increasing d; in the limit $d \to \infty$ for $q = 1$ we obtain $F(\rho, |G\rangle) > 0.905$.

4.2.4   Conclusion

We showed classical communication between parties is necessary for the generation of qudit graph states of prime local dimension (and all states that are in their vicinity) in quantum networks with bipartite sources. We achieve this goal by employing the quantum inflation method. In fact, we demonstrate that the use of many different inflations of the same network might be beneficial over using just two inflations as done before in the literature, and hence our work might inspire future research involving quantum networks. Our results serve as a guide to experimental physicists who wish to implement protocols on quantum networks that involve graph states. On the other hand, they hint at a possible gain from the construction of protocols based on other states than the graph states.

Still, many questions concerning LOSR networks remain unexplored. The most obvious one is whether graph states are preparable with k-partite sources for $k \geqslant 3$. Even if from the application viewpoint this case seems less important than that of $k = 2$, answering this question would allow us to understand quantum networks on a deeper level. One can also ask whether other classes of multipartite states can be obtained in LOSR networks. Apart from the graph states, this question was answered negatively for symmetric or antisymmetric states of any local dimension [112] and pure genuinely entangled states of local dimension 2 and 3 in Ref. [113]. On the other hand, it would be interesting to determine the minimal amount of classical communication required to generate graph states in LOCC networks and to identify other classes of states that are efficiently preparable in this sense.

Chapter 5

State Polynomial Optimization

The semidefinite programming hierarchies in Chapter 3 can be phrased for a more general type of problem known as state polynomial optimization (SPO). In SPO we aim to optimize a polynomial expression in a state of a C^*-algebra $\mathcal{A}$ under polynomial (equality) constraints. An optimizing state then also induces a representation of the C^*-algebra through the GNS construction.

Let us formalize the notion of an SPO problem. Let $\mathcal{A}$ be a C^*-algebra with state space $K(\mathcal{A})$. Following the notation of Ref. [39] we define a (commutative) state polynomial p as a formal polynomial in state symbols $\sigma(w)$, where w is a word in symbols indexed in the generators of $\mathcal{A}$. The degree of a state monomial $\sigma(w_1)\sigma(w_2)\ldots$ is the sum of the degrees of the words $w_1, w_2, \ldots$, i.e. $|\prod_i \sigma(w_i)| = \sum_i |w_i|$. The degree $|p|$ of a state polynomial is the highest degree of the state monomials that it consists of.

One can evaluate a state polynomial on a state $\omega \in K(\mathcal{A})$ and its GNS representation $(\mathcal{H}_\omega, \pi_\omega(\mathcal{A}), |\Omega\rangle)$. That is

$$
\begin{aligned}
&p : (K(\mathcal{A}), \mathcal{A}) \to \mathbb{C} \\
&p : (\omega, \mathcal{A}) \mapsto p(\langle\Omega|\cdot|\Omega\rangle, \pi_\omega(\mathcal{A})).
\end{aligned}
\tag{5.1}
$$

When there is no confusion we will often write ω instead of the expectation value with respect to the GNS vector state $|\Omega\rangle$, to reduce clutter.

Example 5.0.1. As a simple example, consider the state polynomial that represents the factorization of a state with respect to two operators that we have encountered several times in Chapter 3:

$$
p = \sigma(AB) - \sigma(A)\sigma(B).
\tag{5.2}
$$

At the moment this is only a symbolic expression and can be written down regardless of the existence of states and representations that obey such an equation.

This expression could, for example, be evaluated on realizations of the operators A and B as

$$\pi(A) = X \otimes \mathbb{1} \in \pi(\mathcal{A}), \qquad \pi(B) = \mathbb{1} \otimes Z \in \pi(\mathcal{A})$$

and the functional ρ that corresponds to the state $|+\rangle \otimes |+\rangle$, in which case

$$\begin{aligned}
p(\rho, \pi(\mathcal{A})) &= \rho(\pi(AB)) - \rho(\pi(A))\rho(\pi(B)) \\
&= \langle++|X \otimes Z|++\rangle - \langle+|X|+\rangle \, \langle+|Z|+\rangle = 0,
\end{aligned}$$

which confirms that this state factorizes over these operators.

Using the notation of state polynomials, we can formalize the optimization problem as follows.

Problem 7 (State polynomial optimization). *Given a C^*-algebra $\mathcal{A}$ with state space $K(\mathcal{A})$, and a countable set of state polynomials p_i, $i \in \{0, 1, 2, \ldots\}$, find*

$$\begin{aligned}
f^* = \min_{\rho \in K(\mathcal{A})} \quad & p_0(\rho, \pi_\rho(\mathcal{A})) \\
\text{s. t.} \quad & p_i(\rho, \pi_\rho(\mathcal{A})) = 0.
\end{aligned} \tag{5.3}$$

The semidefinite programming hierarchies discussed and developed in Chapter 3 can be used to tackle this problem. The key ingredients to the proof of this statement are the Theorems derived in Chapter 2. They show that there is a way of using linear constraints in the state to enforce that the state is a convex combination of product states.

There are roughly two known ways of writing down a convergent hierarchy of relaxations to the SPO problem. The first hierarchy, which we have come to call the *polarization hierarchy*, will be discussed in Sec. 5.1. The proof of convergence is mostly a repetition of Theorem 3.3.15, phrased in the more general language of SPO. The relation to the quantum inflation hierarchy will also be discussed.

The second hierarchy is known as the scalar extension hierarchy in the quantum information community [41]. A more involved version of this hierarchy was shown to be convergent for SPO problems in [39]. The proof uses the observation that any evaluation of a state polynomial simply consists of the multiplication and addition of scalar variables, which allows one to use the *commutative* polynomial optimization hierarchy of Lasserre [73]. The algebra is then 'build from the inside' through a large amount of localizing matrices that restrict the values these scalar operators can take.

Here we give a slightly different proof of convergence of the scalar extension

hierarchy, again making use of polarizations. This sometimes allows us to remove a large number of localizing matrices compared to the formulation of Ref.[39], which are then replaced by polarizations of the constraints they represent. The details can be found below in Sec. 5.2.

Lastly, in Sec. 5.3, we compare the two approaches, and see that they are in fact not so different.

5.1   Polarization

Both the polarization hierarchy and the quantum inflation hierarchy use the observation that a state polynomial on an algebra $\mathcal{A}$ of degree n can be written as an expression that is linear in a symmetric product state that acts on n copies of the algebra $\mathcal{A}^{\otimes n}$. Using again the example of Eq. (5.2), and the evaluation with respect to a state ρ and its GNS representation π_ρ, we get

$$
\begin{aligned}
p(\rho, \pi_\rho(\mathcal{A})) &= \rho(\pi_\rho(AB)) - \rho(\pi_\rho(A))\rho(\pi_\rho(B)) \\
&= \rho^{\otimes 2}(\pi_\rho(AB) \otimes \mathbb{1}) - \rho^{\otimes 2}(\pi_\rho(A) \otimes \pi_\rho(B)) \\
&= \rho^{\otimes 2}(\pi_\rho(AB) \otimes \mathbb{1} - \pi_\rho(A) \otimes \pi_\rho(B)) =: \rho^{\otimes 2}(y_p).
\end{aligned}
$$

Therefore, if we can guarantee that our algebra is a tensor product of n copies of the algebra $\mathcal{A}$, and the state is a symmetric product state on this algebra, we can rewrite all the state polynomials to linear expressions and write the SPO problem as an NPO problem instead. The procedure of turning a state polynomial into a linear expression on the tensor product of copies of the algebra is called *polarization*. The operator y_p is also called the polarization of $p(\rho, \pi_\rho(\mathcal{A}))$.

Up to this point we have been vague about the type of tensor product between the copies of the algebras. For the purpose of linearizing the polynomial any C^*-tensor product would do. But since our goal is to find a convergent hierarchy of SDP relaxations, we once again take the pragmatic approach and use the maximal tensor product in order to show convergence for any C^*-tensor product.

The first hierarchy of relaxations for Problem 7 uses a combination of Theorem 2.1.1 and the polarization technique described above. Let y_0 be the polarization of the state polynomial p_0 and let y_i be the polarizations that corresponds to the state polynomials p_i^2 for all $i \geq 1$. Then we get the following hierarchy of NPO problems

as relaxations of Problem 7.

$$f_{\text{pol}}^n = \min_{\rho \in K(\mathcal{A}^{\otimes n})} \rho(y_0)$$

$$\text{s. t.} \quad \rho(\pi(a_1 \otimes \ldots \otimes a_n)) = \rho(a_1 \otimes \ldots \otimes a_n), \quad \forall \pi \in S_n, a_i \in \mathcal{A}$$

$$\rho(y_i) = 0, \tag{5.4}$$

where π permutes the n tensor factors of $\mathcal{A}^{\otimes n}$.

Theorem 5.1.1. *It holds that* $\lim_{n \to \infty} f_{pol}^n = f^*$.

Proof. The proof once again follows the structure of the proofs of Theorems 3.2.1 and 3.3.15.

First note that $f_{\text{pol}}^n \leq f^*$ for all n, since each of the NPO problems in (5.4) is a relaxation of Problem 7.

For the converse direction, we realize that the states we optimize over in the relaxations (5.4) are symmetric, such that in the limit $n \to \infty$ Theorem 2.1.1 holds.

Let $\omega_n \in K(\mathcal{A}^\infty)$ be the infinite tensor product of the optimizing state of relaxation (5.4) at level n. By the Banach-Alaoglu theorem, there exists a weak*-convergent subsequence in $K(\mathcal{A}^\infty)$. If we construct the state $\omega \in K(\mathcal{A}^\infty)$ as the limit point it must be a separable state by Theorem 2.1.1. That is, it is of the form

$$\omega = \int \mathrm{d}\mu(\sigma)\Pi_\sigma, \tag{5.5}$$

where Π_σ is a symmetric product state on $\mathcal{A}^\infty$, and μ is a probability measure on states of $\mathcal{A}$.

On each of the states Π_σ it holds that $\Pi_\sigma(y_i) = p_i^2(\sigma, \mathcal{A}) \geq 0$ by definition of polarization. Since μ is a probability measure and for all $i \geq 1$ we have $\omega(y_i) = 0$ it must therefore hold that $\Pi_\sigma(y_i) = 0$ almost surely with respect to μ, which also implies that $p_i(\sigma, \mathcal{A}) = 0$ a.e. w.r.t. μ. Let $E \subset K(\mathcal{A})$ be the full measure subset where the constraints hold. Then each $\sigma \in E$ is a feasible state for the SPO problem.

From this it follows that $\Pi_\sigma(y_0) \geq f_{\text{pol}}^\infty$ a.e. w.r.t. μ, for otherwise we could have chosen the point measure on a feasible state $\sigma' \in E$ such that $\Pi_{\sigma'}(y_0) < f_{\text{pol}}^\infty$, which contradicts that f_{pol}^∞ is a minimum. Combining this with the fact that $\omega(y_0) = f_{\text{pol}}^\infty$, it holds that

$$\Pi_\sigma(y_0) = f_{\text{pol}}^\infty \quad \text{a.e. w.r.t. } \mu \text{ on } E.$$

Hence, there exists a feasible state σ such that

$$f_{\text{pol}}^\infty = \Pi_\sigma(y_0) \geq f^*, \tag{5.6}$$

which concludes the proof. $\square$

It should again be noted that it is in general not possible to extract an optimal state, due to the form of Eq. (5.5).

Furthermore, a drawback of the polarization hierarchy might seem to be that the degree of each of the polynomial constraints must be doubled to impose that constraint. However, also adding the polarization of the non-doubled constraints retains the convergence result, and allows to impose a relaxation of the constraint already at lower levels of the hierarchy.

5.1.1   Relation to quantum inflation hierarchy

The quantum inflation hierarchy is closely related to the polarization hierarchy, but instead of using the polarizations y_i for $i \geq 1$, it uses the fact that if a state factorizes with respect to two operators, it is possible to permute these two operators individually. That is, if we once again look at our example (5.2), it must hold that for a product state on $\mathcal{A}^{\otimes n}$

$$\rho^{\otimes n}(AB) = \rho^{\otimes n}(A)\rho^{\otimes n}(B)$$
$$= \rho^{\otimes n}(A\pi(B)) =: \rho^{\otimes n}(\tilde{y})$$

for all $\pi \in S_n$. We call $\tilde{y}$ an inflation of p if A and $\pi(B)$ act on different tensor factors. Hence, this trick also allows us to write polynomials in the state as linear expressions in a symmetric tensor product of that state, by replacing a k-th order polynomial in the state with a permuted version of the operator that acts non-trivially on k copies of the state.

We can then write down the following hierarchy of NPO relaxations to the SPO problem.

$$f_{\text{infl}}^n = \min_{\rho \in K(\mathcal{A}^{\otimes n})} \rho(\tilde{y}_0)$$
$$\text{s. t.} \quad \rho(\pi(a_1 \otimes \ldots \otimes a_n)) = \rho(a_1 \otimes \ldots \otimes a_n), \quad \forall \pi \in S_n, a_i \in \mathcal{A}$$
$$\rho(\tilde{y}_i) = 0, \tag{5.7}$$

where $\tilde{y}_i$ are the inflations of the polynomials p_i.

For the inflation hierarchy, convergence is less clear. It seems to be tailored to the specific problem of causal compatibility, and seems difficult to extend beyond the result of Theorem 3.2.1. The main difficulty stems from the identification of Eq. (3.35) with Eq. (3.38): it is a crucial part of the proof that the commuting subalgebras generate the global algebra, and that the separability of the state can be captured in this one expression by identifying each state in the convex combination with a product state. In a general problem of the form of Problem (7), this is not the case.

For the broader context of SPO, the hierarchy (5.7) can of course still be used as a hierarchy of relaxations.

5.2   Scalar Extension

In addition to the state polynomials for a C^*-algebra $\mathcal{A}$ defined at the beginning of this chapter, one can create formal *non-commutative state polynomials* in state symbols $\sigma(v)$ and words w. Such expressions are linear combinations of objects of the form pw, where p is a (commutative) state polynomial and w is a word in the generators of $\mathcal{A}$. Note that if q is a non-commutative state polynomial, $\sigma(q)$ is a commutative state polynomial, where each word w is replaced by its state symbol $\sigma(w)$.

The authors of Ref. [41] made the observation that the evaluation of a monomial in state symbols consists simply of scalar variables. If we treat those scalars as elements of our algebra, it is once again possible to write a polynomial expression in the state as a linear expression on a larger algebra. For example, looking at our standard factorization example (5.2), we see

$$\sigma(AB) - \sigma(A)\sigma(B) = \sigma(AB) - \sigma(A\sigma(B))$$
$$= \sigma(AB - A\sigma(B)).$$

Treating $\sigma(B)$ as a *scalar extension* of the algebra has linearized the expression. However, since the value of $\sigma(B)$ is also subject to optimization, it is not possible to truly treat $\sigma(B)A$ as a scalar multiple of A in our SDP. Instead, we add $\sigma(B)$ as a generator to our algebra and require it to lie in the center. This yields an algebra of non-commutative state polynomials. In this formulation the variable $\sigma(B)$ is not guaranteed to be a scalar in the GNS representation of the optimal state.

The authors of Ref. [39] made this procedure more systematic and proved that there is a hierarchy of semidefinite programming relaxations to Problem 7 that uses this technique and converges to the optimal value f^*, despite the above-mentioned objection. Here we prove the same statement, with a slightly different hierarchy and proof technique. This new hierarchy, inspired by the approach of Ref. [39], possibly

avoids a large number of localizing matrices that are present in their hierarchy. As we will see, it also bears some resemblance to the polarization hierarchy.

We start by defining the notion of a *quadratic module*. A quadratic module can be defined in several contexts, such as commutative polynomial optimization [73, 126], non-commutative polynomial optimization [40, 127], and now also in the context of (non-commutative) state polynomial optimization [39]. We will adopt this last definition from Ref. [39], as it applies to the SPO problem. Quadratic modules are useful for defining positivity and boundedness in an algebraic way.

Let $\mathcal{A}$ be a $*$-algebra and let $C \subset \mathcal{A}$. We will assume that the set C is *balanced*, that is, it is closed under involution and if $c \in C$ is not self-adjoint, then $-c \in C$. We then define the quadratic module $QM(C)$ generated by C as the smallest set in $\mathcal{A}$ such that

$$\mathbb{1} \in QM(C), \quad QM(C) + QM(C) \subseteq QM(C), \quad x^* QM(C) x \subseteq QM(C) \,\forall x \in \mathcal{A}.$$

A quadratic module is *archimedean* if for each $a \in \mathcal{A}$, there is an $m_a > 0$, such that $m_a \pm a \in QM(C)$.

Using this formalism, we can write down two hierarchies of relaxations to Problem 7 that can be proven to be complete. In fact, we can solve an even more general problem that also involves inequality constraints.

Problem 8 (State polynomial optimization with inequalities). *Given a C^*-algebra $\mathcal{A}$ with state space $K(\mathcal{A})$, and a countable set of state polynomials p_i, $i \in \{0, 1, 2, \ldots\}$, find*

$$\begin{aligned} f^\star = \min_{\rho \in K(\mathcal{A})} \quad & p_0(\rho, \pi_\rho(\mathcal{A})) \\ s.\ t. \quad & p_i(\rho, \pi_\rho(\mathcal{A})) \geq 0. \end{aligned} \tag{5.8}$$

For an optimization problem of the form (5.8) over a universal C^*-algebra $\mathcal{A} = C^*(\mathcal{G}|\mathcal{R})$, with basis $B = \{b\}$[1] and quadratic module $QM(\mathcal{R})$, we define the scalar extension C^*-algebra $\mathcal{C}_{\text{SE}}$ as $C^*(\mathcal{G}_{\text{SE}}|\mathcal{R}_{\text{SE}})$, where

$$\mathcal{G}_{\text{SE}} = \{\sigma(b)\}_{b \in B}, \tag{5.9}$$

$$\mathcal{R}_{\text{SE}} = \{\sigma(r)\}_{r \in QM(\mathcal{R})} \cup \{s_i\} \cup \{[\sigma(b), \sigma(b')]\}_{b, b' \in B} \tag{5.10}$$

where each element $s_i \in \mathcal{F}(\mathcal{G}_{\text{SE}})$ is the scalar extension that corresponds to the the state polynomial p_i. If the relations $\mathcal{R}$ are sufficient to make each element of the algebra $\mathcal{A}$ bounded, then the relations $\mathcal{R}_{\text{SE}}$ will ensure that each of the scalar extensions is also bounded, as we will see below in Lemma 5.2.3.

---

[1] This basis can be chosen to be overcomplete, e.g. by taking all words in the generators.

The first hierarchy (Eq. (5.12)) is equivalent to the hierarchy of Ref. [39], rephrased in the language of this thesis. It expresses the optimization problem fully in terms of the scalar extension variables, essentially 'rebuilding the algebra from within' by requiring the relations $\mathcal{R}_{SE}$ to hold. Since the optimization is then only over commutative variables, it is possible to use the polynomial optimization hierarchy of Lasserre [73] to enforce the polynomial constraints in the state. Using the GNS construction, it is shown that there exists a state that reaches the optimal value $f^\star$. A proof of these facts can be found in section 5 of Ref. [39]. Below, we will give an alternative, but equivalent proof for the hierarchy (5.14).

Problem 8 can be relaxed to the following commutative polynomial optimization problem. Let s_0 be the element in $\mathcal{C}_{SE}$ that corresponds to the state polynomial p_0, fully expressed in scalar extension variables.

$$f^*_{SE} = \min_{\rho \in K(\mathcal{C}_{SE})} \rho(s_0), \tag{5.11}$$

which can be solved via the hierarchy of relaxations (cf. Eq. (1.27))

$$
\begin{aligned}
f^k_{SE} = \min_{L \in (\mathcal{F}^{2k})^*} \quad & L(s_0) \\
\text{s. t.} \quad & \Gamma^k_{\sigma(\mathbb{1})} = 1 \\
& \Gamma^k \succeq 0 \\
& \Lambda^k_q \succeq 0 \qquad \forall q \in \mathcal{R}_{SE} \cap \mathrm{span}(\mathcal{F}^{2k})
\end{aligned}
\tag{5.12}
$$

where $\mathcal{F}^{2k}$ is the set of words in the generators of $\mathcal{C}_{SE}$ of length $\leq 2k$. Recall that the definition of the moment matrix Γ^k includes the fact that one should put $\Gamma_{x,y} = \Gamma_{x',y'}$ if $\rho(x^*y) = \rho((x')^*y')$ (and similar for the localizing matrices).

Theorem 5.2.1 (Klep *et al.* [39])**.** *If $QM(\mathcal{R})$ is an archimedean quadratic module, it holds that $\lim_{k \to \infty} f^k_{SE} = f^*_{SE} = f^\star$.*

The second hierarchy, a version of which will appear in the revised version of Ref. [83][2], is very similar, but in some cases improves over the hierarchy (5.12), in the sense that a fixed level k of the hierarchy is less expensive to run. Instead of having all the constraints $\mathcal{R}_{SE}$ represented by moment matrices, which are computationally expensive, it is possible to impose each equality constraint in $\{p_i\}$ via a polarization.

To this end, let us split up the relations $\mathcal{R}_{SE}$ into the sets $Q = \{q_i\} \subset \{s_i\}$ of equality constraints stemming from $\{p_i\}$ and all other constraints $T = \{t_j\}$. This allows us to define the algebra $\tilde{\mathcal{C}}_{SE} = C^*(\mathcal{G}_{SE}|T)$.

---

[2] Will appear as Ref. [128].

For each equality constraint q_i, we define the polarization y_{q_i} simply by the square of that constraint, i.e. $y_{q_i} = q_i^2$. Note that there is no need to do any permutations, like in the case of the hierarchy of Eq. (5.4), as we will see in the proof of Theorem 5.2.2 below.

We then get the following optimization problem.

$$
\begin{aligned}
\tilde{f}_{\mathrm{SE}}^* = \min_{\rho \in K(\tilde{C}_{\mathrm{SE}})} \quad & \rho(s_0) \\
\text{s. t.} \quad & t_j \succeq 0 \qquad \forall t_j \in T, \\
& \rho(y_{q_i}) = 0 \qquad \forall q_i \in Q,
\end{aligned}
\tag{5.13}
$$

which has relaxations of the form

$$
\begin{aligned}
\tilde{f}_{\mathrm{SE}}^k = \min_{L \in (\mathcal{F}^{2k})^*} \quad & L(s_0) \\
\text{s. t.} \quad & \Gamma_{\sigma(\mathbb{1})}^k = 1 \\
& \Gamma^k \succeq 0 \\
& \Lambda_{t_j}^k \succeq 0 \qquad \forall t_j \in T \cap \mathrm{span}(\mathcal{F}^{2k}) \\
& L(y_{q_i}) = 0 \qquad \forall q_i \in Q \text{ s. t. } y_{q_i} \in \mathrm{span}(\mathcal{F}^{2k}),
\end{aligned}
\tag{5.14}
$$

Theorem 5.2.2. *If $QM(\mathcal{R})$ is an archimedean quadratic module, it holds that $\lim_{k \to \infty} \tilde{f}_{SE}^k = \tilde{f}_{SE}^* = f^\star$.*

The proof of the Theorem consists of many similar elements as the proof of Theorem 5.2.1 proven in Ref. [39]. However, we try to phrase it using the language developed in Chapters 1 and 2. It relies roughly on two observations:

1. The extremal states, i.e. the pure states, of a commutative C^*-algebra are $*$-homomorphisms. For the GNS representation induced by such states the scalar extension variables actually become scalars, justifying the terminology.

2. A form of the GNS construction applied to the scalar extension variables under the constraints $\mathcal{R}_{\mathrm{SE}}$ allows us to build a quantum model that corresponds to a representation of the C^*-algebra $\mathcal{A}$ and a state that obeys the polynomial constraints $\{p_i\}$.

Lemma 5.2.3. *Let $\mathcal{A} = C^*(\mathcal{G}|\mathcal{R})$ be a universal C^*-algebra. Then the generators $\mathcal{G}_{SE}$ of the induced scalar extension algebra $\tilde{C}_{SE}$ are bounded by the relations T, such that $C^*(\mathcal{G}_{SE}|T)$ is a well-defined abelian universal C^*-algebra.*

Proof. Assume w.l.o.g. that the generators are self-adjoint. It follows from the assumptions in the definition of a universal C^*-algebra that the relations $\mathcal{R}$ imply that

the generators $g \in \mathcal{G}$ are bounded. That is, the relations ensure that there exists an $N_g \geq 0$ such that

$$N_g \mathbb{1} - g^* g \succeq 0.$$

It follows that for every word w

$$w^*(N_g \mathbb{1} - g^* g)w \in QM(\mathcal{R}) \subset T, \tag{5.15}$$

such that Lemma 1.1.5 implies

$$N_g \sigma(w^* w) - \sigma(w^* g^* g w) = \sigma(w^*(N_g \mathbb{1} - g^* g)w) \succeq 0. \tag{5.16}$$

By inductively reducing the length of the word w using Eq. (5.16), it follows that there always exist an $N_w \geq 0$ such that

$$N_w - \sigma(w^* w) \succeq 0. \tag{5.17}$$

Following a trick that is often used for Positivstellensätze (see e.g. also [39, Thm. 5.5]), we get that

$$\frac{1}{4} + N_w \pm \sigma(w) = \sigma\left(\left(\frac{1}{2} \pm w\right)^*\left(\frac{1}{2} \pm w\right)\right) + N_w - \sigma(w^* w) \succeq 0, \tag{5.18}$$

which implies

$$\begin{aligned}
0 &\preceq \left(\frac{1}{4} + N_w + \sigma(w)\right)^*\left(\frac{1}{4} + N_w - \sigma(w)\right) \\
&= \left(\frac{1}{4} + N_w\right)^2 - \sigma(w)^* \sigma(w) + \left(\frac{1}{4} + N_w\right)(\sigma(w^*) - \sigma(w)) \\
&\preceq 3\left(\frac{1}{4} + N_w\right)^2 - \sigma(w)^* \sigma(w),
\end{aligned}$$

where the last inequality follows from applying Eq. (5.18) twice more.

That is, all generators of $\tilde{\mathcal{C}}_{\text{SE}}$ are bounded. □

Proof. (of Theorem 5.2.2) By lemma 5.2.3, we see that we are indeed optimizing over a universal algebra given by $\mathcal{C}_{\text{SE}} = C^*(\mathcal{G}_{\text{SE}}|T)$. This turns (5.13) into a standard NPO problem as defined in Sec. 1.3.2[3], with relaxations given by (5.14). At each finite level of the hierarchy, we are optimizing over a subset of the scalar extension generators and their relations. A functional $L \in (\mathcal{F}^{2k})^*$ can then be extended to all

---

[3] Actually it is a *commutative* polynomial optimization problem, but this still fits into the NPO framework as well.

of $\mathcal{C}_{\text{SE}}$, e.g. by putting $L(x) = 0$ for $x \notin \mathcal{F}^{2k}$. By the convergence of NPO, there exists a limiting functional that is a state on $\mathcal{C}_{\text{SE}}$ that obeys all constraints. We let ρ be this limiting optimizing state.

Since $\tilde{\mathcal{C}}_{\text{SE}}$ is an abelian C^*-algebra, we can apply the result of Corollary 2.2.1 to show that ρ has a weak*-integral decomposition of the form

$$\rho = \int \mathrm{d}\mu(\phi)\phi, \tag{5.19}$$

where each ϕ is a $*$-homomorphism and μ is a probability measure over pure states. In Ref. [39] this is referred to as the Kadison-Dubois representation theorem.

Each of the states ϕ induces a one-dimensional GNS representation, where every scalar extension variable is in fact simply a scalar.

The constraints T were imposed algebraically, and thus apply to any representation of $\tilde{\mathcal{C}}_{\text{SE}}$. Additionally, by once again applying the same reasoning as in the proof of Theorem 5.1.1, the polarization constraints enforce that $\phi(y_{q_i}) = 0$ and $\phi(s_0) = \tilde{f}_{\text{SE}}^*$ almost surely with respect to μ. Choosing a ϕ^* that is in the full measure subset where these equalities hold thus gives an optimal solution to the NPO problem (5.13), which yields the same optimal value as problem (5.11), by virtue of the polarizations $\{y_{q_i}\}$.

What remains to be shown is that this solution also provides an optimal solution to the original problem, namely problem 8. First note that (5.13) is a relaxation to Problem 8, so that $f^\star \geq \tilde{f}_{\text{SE}}^*$.

For the converse, we will use the optimizing state ϕ^* to build a state $\psi \in K(\mathcal{A})$ and its corresponding GNS representation, and show that it obeys all the constraints. We define $\psi \in \mathcal{A}^*$ as

$$\psi(x) = \phi^*(\sigma(x)) \quad \forall x \in \mathcal{A}. \tag{5.20}$$

Here $\sigma(x)$ should be understood as $\sum_b \lambda_b \sigma(b)$, where $x = \sum_b \lambda_b b$ is the expansion of x in the basis $\{b\}$.

To show that ψ is a valid state and we can use it to build a representation of $\mathcal{A}$ on a Hilbert space, we need to show that it is positive on $QM(\mathcal{R})$ and that $\psi(\mathbb{1}) = 1$ (see e.g. [129, Theorem 1.27] or [40], which largely boils down to an application of the GNS construction). Additionally, we need to check that all the polynomial constraints $\{p_i\}$ are satisfied.

The normalization requirement is easily verified, since

$$\psi(\mathbb{1}_{\mathcal{A}}) = \phi^*(\sigma(\mathbb{1}_{\mathcal{A}})) = \phi^*(\mathbb{1}_{\tilde{\mathcal{C}}_{\text{SE}}}) = 1.$$

The other requirements follow from the constraints $\sigma(QM(\mathcal{R})) \in \mathcal{R}_{\text{SE}}$ and $\{s_i\}$:

$$\psi(r) = \phi^*(\sigma(r)) \geq 0, \qquad \forall r \in QM(\mathcal{R}),$$
$$p_i(\psi, \pi_\psi(\mathcal{A})) = \phi^*(\sigma(s_i)) = \phi^*(s_i) \geq 0, \qquad \forall s_i.$$

By the definition of Eq. (5.20), it then holds that $p_0(\psi, \pi_\psi(\mathcal{A})) = \tilde{f}^*_{\text{SE}}$. Furthermore, since this yields a feasible point for Problem 8, we have

$$p_0(\psi, \pi_\psi(\mathcal{A})) \geq f^\star,$$

from which we conclude $f^\star = \tilde{f}^*_{\text{SE}} = \lim_{k \to \infty} \tilde{f}^k_{\text{SE}}$. $\qquad\square$

As a result of [39, Lemma 6.4 and 6.5], both versions of the hierarchy, Eqs. (5.12) and (5.14), can be expressed with fewer localizing matrices, by relating them to the non-commutative problem they derive from. To be more precise, we include all the non-commutative generators and their relations, and require entries of the moment matrix to obey

$$(\Gamma^k)_{v,w} = L(\sigma(v^*w)), \tag{5.21}$$

so that the non-commutative variables and their scalar extensions evaluate to the same number. The consequence is that the relations for $\{\sigma(r)\}_{r \in QM(\mathcal{R})}$ are automatically obeyed due to the relations imposed on the non-commutative variables. That is, we will effectively optimize over symmetric states on $C^*(\mathcal{G}|\mathcal{R}) \otimes C^*(\mathcal{G}_{\text{SE}}|\mathcal{R}_{\text{SE}})$. We thus get the following alternative SDP hierarchy for the hierarchy of Eq. (5.12)

$$
\begin{aligned}
g^k_{\text{SE}} = \min_{L \in (\tilde{\mathcal{F}}^{2k})^*} \quad & L(s_0) \\
\text{s. t.} \quad & \Gamma^k_{\mathbb{1}} = 1 \\
& \Gamma^k \succeq 0 \\
& \Lambda^k_q \succeq 0 \qquad \forall q \in (\mathcal{R} \cup \{s_i\}_i) \cap \text{span}(\tilde{\mathcal{F}}^{2k}),
\end{aligned}
\tag{5.22}
$$

where $\tilde{\mathcal{F}}^{2k}$ consists of all the words in the commutative and non-commutative generators up to length $2k$. Note that it is now implicitly assumed that Eq. (5.21) is imposed.

It is still possible to use the polarization trick as well. We distinguish between the equality constraints $S_{eq} \subset \{s_i\}$ and inequality constraints $S_{ineq} \subset \{s_i\}$ that arise from the polynomial constraints p_i. The alternative hierarchy for Eq. (5.14)

then becomes

$$\tilde{g}_{\text{SE}}^k = \min_{L \in (\bar{\mathcal{F}}^{2k})^*} \quad L(s_0)$$

$$\text{s. t.} \quad \Gamma_{\mathbb{1}}^k = 1$$
$$\Gamma^k \succeq 0 \tag{5.23}$$
$$\Lambda_{t_j}^k \succeq 0 \qquad \forall t_j \in (\mathcal{R} \cup S_{ineq}) \cap \text{span}(\tilde{\mathcal{F}}^{2k})$$
$$L(y_{q_i}) = 0 \qquad \forall q_i \in S_{eq} \text{ s. t. } y_{q_i} \in \text{span}(\tilde{\mathcal{F}}^{2k}).$$

The following corollary then follows immediately (see e.g. [39, Lemma 6.5]).

Corollary 5.2.4. *If $QM(\mathcal{R})$ is an archimedean quadratic module, it holds that* $\lim_{k \to \infty} g_{SE}^k = g_{SE}^* = f^\star$ *and* $\lim_{k \to \infty} \tilde{g}_{SE}^k = \tilde{g}_{SE}^* = f^\star$.

Some remarks are in order.
Remarks:

- In practice the size of each relaxation in both scalar extension hierarchies can often be greatly reduced by making use of the equality constraints to reduce the set of basis elements of the algebra. For example, in optimization over algebras on a causal structure, the commutation relations and P(O)VM conditions can be used to this end. The use of such symmetries is one of the main reasons why it is possible to use the scalar extension hierarchy to get non-trivial results for optimization problems.

- In the hierarchies (5.14) and (5.23), we chose to enforce some equality constraints on the level of the state through polarizations, as opposed to on the level of the algebra via localizing matrices. In many cases, a larger set of constraints can be enforced through polarizations. The only requirement is that the algebraic constraints should be sufficient to bound each generator. If an upper bound on the norm of basis elements is known (e.g. because they are POVM elements), it is possible to add constraints of the form

$$N - \sum_{\{x\} \subset \{b\}} \sigma(x)^* \sigma(x) \geq 0, \tag{5.24}$$

where $N \geq \sum_x \|x\|$, which ensures boundedness of the operators in $\{x\}$. Any polynomial equality constraints involving only elements in $\{x\}$ can then be imposed through polarizations.

As an example, for the case of causal optimization, at any level of the hierarchy it is sufficient to have only one constraint of the form (5.24), since we know

each basis element, as a product of POVM elements, is bounded in norm by 1. All other polynomial equality constraints can then be imposed through polarizations, greatly reducing the complexity of the problem at higher levels of the hierarchy.

- It is in general not possible to extract an optimal solution, since it will be of the form (5.19), except when the optimal state turns out to be pure.

5.3   Comparison of the two approaches

The polarization hierarchy and the scalar extension hierarchy a priori seem fairly different. However, here we will try to argue that both approaches are actually quite similar, but they treat the two kinds of polynomials that appear in state polynomial optimization in a different order. For identification purposes, let us call polynomials in the operators of the algebra *operator polynomials*, while polynomials in the state are still referred to as *state polynomials*.

- The polarization hierarchy first copies the algebra $\mathcal{A}$, which is easy to do. The resulting algebra is still a non-commutative algebra, for which we know that the NPO hierarchy is convergent. This allows us to build operator polynomials.

  As a second step, the goal is to construct a symmetric state on this algebra. In the limit, such a state is a convex mixture of symmetric product states due to the quantum de Finetti theorem. Letting the product states act on the copied algebra allows us to build state polynomials.

- Scalar extension, on the other hand, first constructs a set of commuting operators by mimicking the action of a state on a basis of the algebra $\mathcal{A}$. Then these commuting operators generate an abelian algebra, for which the Lasserre hierarchy is known to be convergent. This convergence relies on the Kadison-Dubois representation theorem. Using this abelian algebra, one can construct state polynomials.

  As a second step, it is shown that a representation of the algebra $\mathcal{A}$ can be recovered if a certain set of constraints is imposed on the commuting variables. It then requires some thought which set of constraints is necessary and sufficient for this purpose. By choosing the right set of constraints, one recovers the set of operator polynomials.

The Kadison-Dubois representation theorem and the quantum de Finetti theorem essentially stem from the same observation, namely that the extremal points of the unit ball in the dual of an abelian C^*-algebra are $*$-homomorphisms. In the case

of the polarization hierarchy, one has to work quite hard to use this statement in the form of a de Finetti theorem, while for abelian algebras, the statement can be applied directly. On the other hand, the fact that the optimization problem of polarization yields a state and representation for the algebra $\mathcal{A}$ is not hard to see: it is a standard NPO problem. In contrast, for the scalar extension hierarchy the difficulty now lies in choosing an appropriate set of constraints, such that the solution also yields a way to construct a state and representation for $\mathcal{A}$.

In the end, the difference in proof techniques seems to stem mostly from familiarity with each of the concepts: as researchers in quantum information theory, non-commutative operators and product states seem like a natural approach to the problem, while for researchers in classical optimization it would be reasonable to use the Lasserre hierarchy.

Which hierarchy is more useful in practice is as of yet unclear, due to the fact that the polarization hierarchy has not been implemented yet. For the particular problem of causal optimization, the recently developed quantum inflation toolbox of Ref. [100] can be used to compare the closely related quantum inflation hierarchy to the results of the scalar extension hierarchy in Ref. [39]. The authors of Ref. [39], however, were able to greatly reduce the size of their test cases by using the symmetry and sparsity of the problem. For a fair comparison to the quantum inflation hierarchy, similar reductions should be applied to the problems constructed by the toolbox of Ref. [100]. As of yet such reductions are not part of the toolbox and would require a significant amount of work to implement. Writing a program that implements the polarization hierarchy is left for later research.

Conclusion

In this final section, I will summarize the main results presented in this thesis, and discuss some of the open problems that I hope to see solved in the near future.

The main focus of the thesis has been on the quantum causal compatibility problem, in which one tries to determine whether a certain observed probability distribution can be obtained from a quantum model that fits a particular causal hypothesis. Though the problem itself is easy to understand, finding a systematic way to answer the question posed in the problem has been considered extremely challenging, even in the classical case.

Recently, a new method to analyze the causal compatibility has been proposed in the form of the inflation technique. This technique has been developed for the classical, as well as for the quantum causal compatibility problem. Its merit is that it relaxes independence constraints, which are expressions that are at least quadratic in the state, to symmetry constraints, which can be formulated as linear constraints on a larger algebra. By increasing the size of the algebra and the number of symmetry constraints, the inflation technique gives a hierarchy of linear programs in the classical case, and semidefinite programs in the quantum case, which provide ever tighter compatibility tests. If one of the tests fails, the distribution cannot have been the result of the causal hypothesis. An important question then becomes whether the converse also holds, i.e. whether these hierarchies are complete in the sense that any incompatible distribution is eventually caught at some level of the hierarchy. In the classical case, this question has been answered in the affirmative. This thesis addressed the same question for the quantum inflation hierarchy.

In Chapter 3 we partially answered this question, based on the results from the publications [1] and [2]. Before attempting to give such an answer, however, one needs to properly define what a quantum model with a causal hypothesis entails. In particular, there is a choice to be made in how to model local subsystems. Throughout the thesis, we have made the choice to use the commuting operators model, in which subsystems are defined through commuting subalgebras, as opposed to the

more commonly used Hilbert space tensor product model. The mathematical tools necessary to treat this way of modeling quantum physics have been discussed in Chapter 1. This Chapter also included an introduction to semidefinite programming, and concludes with the observation that non-commutative polynomial optimization (NPO) problems can be phrased as optimization problems over C^*-algebras.

In Sections 3.1 and 3.2 we have developed an alternative quantum inflation hierarchy that is provably convergent. It relied roughly on three observations. Firstly, it is noted that symmetry conditions are expressed as constraints that are linear in the state and elements of the algebra. It is possible to optimize over such expressions using the NPO hierarchy. Secondly, symmetric states are closely related to separable states, a statement that is made precise in the form of a quantum de Finetti theorem. This allows use to use linear expressions on symmetric states as a relaxation of independence constraints that are implied by the causal structure. Lastly, and perhaps most importantly for a convergence argument, it is observed that the local algebras of each of the parties must consist of a number of subalgebras corresponding to the number of quantum systems that that party receives. Additionally, the independence constraints are to be formulated with respect to these subalgebras. Any constructive proof of convergence must therefore be able to certify the existence of these local subalgebras and a state obeying the independence constraints, solely from the data provided by the inflation hierarchy.

Building on this third observation, we have constructed an inflation hierarchy in which the local subalgebras are put in 'by hand'. That is, we have introduced a number of generators corresponding to each subalgebra, and required the POVM elements of each party to be build up from these generators. This has allowed us to formulate the independence constraints of the causal structure fully in terms of factorization conditions on these subalgebras. This construction furthermore suggests more clearly that a quantum de Finetti theorem might be applicable. Indeed, in Chapter 2, we have proven such a theorem (Theorem 2.1.1) in the context of general tensor products of C^*-algebras, extending a previous result that was restricted to minimal tensor products. Using this theorem, it could be shown that, in the limit, the symmetry constraints force the state to be a separable state, i.e. a convex mixture of product states. Having used the second and third observations, it is a straightforward application of NPO theory to formulate this new inflation hierarchy as a hierarchy of semidefinite programs.

The final step in proving convergence of this new inflation hierarchy involved singling out a state in the convex mixture of product states that also obeys all the constraints. For this we used a trick known as polarization. It used the observation that positive operators that evaluate to zero on a separable state, must evaluate to zero almost surely for the product states in the convex mixture that defines the separable state. In this way, it was shown that if all of the levels of the hierarchy are

passed, there must exist a state that is compatible with the causal explanation, and that produces the probability distribution up to arbitrary precision.

The new quantum inflation hierarchy has some significant drawbacks: it introduced a large number of additional variables and was therefore even more impractical than the original quantum inflation hierarchy. Additionally, the convergence was not monotonic: the number of generators for each local subalgebra became a parameter of the hierarchy, and thus restricting the number of generators was also restricting the quantum model. The question whether the original quantum inflation hierarchy is convergent therefore remains very relevant.

In Section 3.3 we have shown that for the specific case of the bilocal scenario, the original quantum inflation hierarchy is indeed convergent. As noted above, this required constructing local subalgebras for Bob, the middle party in the bilocal scenario, and constructing a global state that obeys the correct independence constraints. This was done by defining Bob's algebra to be the tensor product of the commutants of the GNS representations of Alice's and Charlie's algebras. It was then shown that it is sufficient to require independence of the state with respect to the algebras of Alice and Charlie. This section also developed a new type of SDP hierarchy, which we called the polarization hierarchy. Though it was inspired by the quantum inflation hierarchy, it is distinct in some relevant ways. Both the polarization hierarchy and the inflation hierarchy were shown to be convergent for the quantum causal compatibility problem in the bilocal scenario. The construction also worked for star networks, but failed to work for general causal structures. It remains an interesting open question whether the symmetries that follow from the inflation technique can be combined with this technique of (re)constructing local algebras to build a quantum model from the inflation or polarization technique for more general causal structures.

Chapter 4 treated a more general setup than the quantum causal compatibility problem. Instead of restricting the end nodes in a causal network to be classical random variables, the end result could now also be a quantum state. From this setup, the quantum network compatibility problem naturally arose, in which it is asked which quantum states can be produced in certain causal networks. The goal of Section 4.1 was to show that there exist similar quantum inflation and polarization hierarchies for these types of problems. In particular, the distinction was made between the case where only local operations (LO) were allowed, and the case where local operations and shared randomness (LOSR) were allowed. After realizing that the techniques of the previous chapters apply generally to C^*-algebras, it was quickly seen that the LO network compatibility problem was essentially the same as the causal compatibility problem. Therefore, the results from previous sections largely carried over. For the LOSR network compatibility problem, the results from Secs. 3.1 and 3.2 could be adapted to this case as well.

Apart from treating the inflation technique numerically with SDPs, it is also possible to derive certain analytic results. This was the goal of Sec. 4.2. Here, we showed that non-trivial graph states cannot be produced in bipartite LOSR networks. Since graph states form an important class of quantum states, this negative result can be seen as an argument against using LOSR networks. The argument involved a proof by contradiction, in which it was shown that an inequality was violated in an inflated network if the state was indeed assumed to be a graph state. In deriving the inequality, it was used that graph states are eigenstates of certain combinations of (generalized) Pauli matrices with eigenvalue 1, and that Pauli matrices obey certain commutation relations. To arrive at a contradiction, it was then used repeatedly that unit expectation values with respect to Pauli matrices have implications on the expectation values of other Pauli matrices. Linking many such expectation values together using the symmetries of the inflation technique eventually allowed us to exceed the bound of the derived inequality, thus reaching a contradiction.

Section 4.2 furthermore included a robustness result, in which it was shown that states that are sufficiently close to a graph state also cannot be produced in bipartite LOSR networks. Such robustness results are relevant in practice, since current quantum devices are not yet fault-tolerant.

In future work, it might be interesting to consider a more realistic setup, such as a setup that allows a restricted amount of classical communication, or quantum sources with bounded dimensions.

Lastly, Chapter 5 treated a more general type of optimization problem, which is known as state polynomial optimization. In such optimization problems, one optimizes over states on a C^*-algebra, under state polynomial (in)equality constraints. The polarization hierarchy was extended to this more general setting and compared to the recently developed scalar extension hierarchy. Additionally, an alternative proof of the convergence of the scalar extension hierarchy was given, and the polarization trick was included in this setting to possibly reduce the size of the semidefinite program. Though both hierarchies achieve convergence in the limit, which of them has faster convergence in practice is an interesting and relevant question, left for future research. In this context, it is also valuable to investigate which subsets of generators and constraints provide the strongest restrictions to the problem, such that the numerical optimization can be adjusted to prioritize them.

It would be interesting to develop methods to analyze causal structures of other GPTs and compare them to quantum causal structures. For example, it is known that a triangle causal structure that adheres to the limitations of boxworld (i.e. consisting of PR boxes) is in some respects less powerful than the equivalent quantum causal structure [130]. This is in some sense surprising, since PR boxes are strictly more

powerful than quantum theory in the Bell scenario [131]. Creating a version of the inflation method that works for GPTs, like boxworld, is therefore of interest, since it allows us to create experiments that can disprove such theories.

In addition, the development of other methods to tackle the causal compatibility problem that do not rely on the inflation technique, and are perhaps less computationally expensive, remains a relevant line of inquiry.

Acknowledgments

I would like to express my sincerest gratitude toward all of the wonderful people that supported and encouraged me during the completion of the PhD. First and foremost, I greatly appreciate the guidance and expertise provided by my advisor, David Gross. He has helped me grow both on an academic and on a personal level. His ideas, knowledge and questions have propelled the research forward, and his kindness and sense of humor have made it a pleasure to work with him.

I also extend my thanks to the other members of my defense committee and thesis advisory committee, Frank Vallentin, Michael Walter and Paul van Loosdrecht, for providing valuable feedback and for evaluating this manuscript.

I am very grateful for all the members of David's research group that I had the privilege to meet and work with. In particular, my office mate Paulina Goedicke, who gives a completely different meaning to the expression 'another day in the office', my good friend David Wierichs, who has been around to support me throughout the entire PhD project, my friend-on-the-other-side-of-the-world Felipe Montealegre Mora, who was always up for a chat, and my-friend-who-is-just-a-train-ride-away, Mariami Gachechiladze, who helped me through the first project during the additional difficulties provided by the pandemic. Additionally I would like to thank Julius Zeiss, Mariela Boevska, Johan Åberg, Vahideh Eshaghian, Fabian Henze, Christian Gorjaew, Taylor Garnowski, Lionel Dmello, Tanmay Singal, Arne Heimendahl, Yaiza Aragonés, Markus Heinrich and Mark Goh for making the office such a nice environment. Both the serious and the less serious conversations have contributed significantly to the completion of the PhD project.

Furthermore, I want to thank my collaborators and colleagues for all the useful discussions and many insights that they provided. This is a long list that includes, but is not limited to, David Gross, Mariami Gachechiladze, Nikolai Miklin, Reinhard Werner, Owidiusz Makuta, Remigiusz Augusiak, Julius Zeiss, Marc-Olivier Renou, Xiangling Xu, Elie Wolfe and Alex Pozas-Kerstjens.

In addition, I would like to thank David and Paulina for giving their valuable feedback on this manuscript.

Lastly, I want to thank my friends and family: my dad, mom, step-mom, brother and grandmother, who have always been extremely supportive of all the decisions I have made throughout my life; my friends from university, in particular Deanna Eerbeek who has always had more confidence in my academic capabilities than I did; and my friends from Alkmaar, most notably Sam Cornelissen, my best and longest friend.

Thank you all for being such an integral part of my life. This thesis would not have been possible without your support.

Bibliography

[1] Laurens T Ligthart, Mariami Gachechiladze, and David Gross. "A convergent inflation hierarchy for quantum causal structures". Communications in Mathematical Physics **401**, 2673–2714 (2023). Appearances: iii, xv, 4, 5, 6, 12, 20, 25, 31, 40, 43, 46, 47, 49, 51, 69, 70, 85, 87, 147

[2] Laurens T Ligthart and David Gross. "The inflation hierarchy and the polarization hierarchy are complete for the quantum bilocal scenario". Journal of Mathematical Physics **64** (2023). Appearances: iii, xv, 31, 47, 69, 147

[3] Owidiusz Makuta, Laurens T Ligthart, and Remigiusz Augusiak. "No graph state is preparable in quantum networks with bipartite sources and no classical communication". npj Quantum Information **9**, 117 (2023). Appearances: iii, xv, 105, 106

[4] Alain Aspect, Philippe Grangier, and Gérard Roger. "Experimental realization of einstein-podolsky-rosen-bohm gedankenexperiment: A new violation of bell's inequalities". Physical review letters **49**, 91 (1982). Appearances: x

[5] Alain Aspect, Jean Dalibard, and Gérard Roger. "Experimental test of bell's inequalities using time-varying analyzers". Physical review letters **49**, 1804 (1982). Appearances: x

[6] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. "Bell nonlocality". Reviews of Modern Physics **86**, 419 (2014). Appearances: x

[7] Jian-Wei Pan, Zeng-Bing Chen, Chao-Yang Lu, Harald Weinfurter, Anton Zeilinger, and Marek Żukowski. "Multiphoton entanglement and interferometry". Rev. Mod. Phys. **84**, 777–838 (2012). Appearances: x

[8] Albert Einstein, Boris Podolsky, and Nathan Rosen. "Can quantum-mechanical description of physical reality be considered complete?". Physical review **47**, 777 (1935). Appearances: x

[9] John S Bell. "On the einstein podolsky rosen paradox". Physics Physique Fizika **1**, 195 (1964). Appearances: x

[10] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. "Proposed experiment to test local hidden-variable theories". Phys. Rev. Lett. **23**, 880–884 (1969). Appearances: x

[11] Rafael Chaves, Christian Majenz, and David Gross. "Information–theoretic implications of quantum causal structures". Nature communications **6**, 1–8 (2015). Appearances: xi, 12, 60, 69

[12] J v. Neumann. "Zur algebra der funktionaloperationen und theorie der normalen operatoren". Mathematische Annalen **102**, 370–427 (1930). Appearances: xii, 2

[13] Francis J Murray and J v Neumann. "On rings of operators". Annals of MathematicsPages 116–229 (1936). Appearances: xii, 2

[14] Francis J Murray and John Von Neumann. "On rings of operators. ii". Transactions of the American Mathematical Society **41**, 208–248 (1937). Appearances: xii, 2

[15] Francis J Murray and John von Neumann. "On rings of operators. iv". Annals of MathematicsPages 716–808 (1943). Appearances: xii, 2

[16] J v Neumann. "On rings of operators. iii". Annals of MathematicsPages 94–161 (1940). Appearances: xii, 2

[17] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. "MIP* = RE" (2020). arXiv:2001.04383. Appearances: xii, 15, 17, 42, 46, 71, 72

[18] Judea Pearl. "Causality". Cambridge university press. Cambridge (2009). Appearances: xii, 10, 11, 12

[19] Alexander Balke and Judea Pearl. "Bounds on treatment effects from studies with imperfect compliance". Journal of the American Statistical Association **92**, 1171–1176 (1997). Appearances: xii

[20] Joshua D Angrist, Guido W Imbens, and Donald B Rubin. "Identification of causal effects using instrumental variables". Journal of the American statistical Association **91**, 444–455 (1996). Appearances: xii

[21] Elie Wolfe, Robert W Spekkens, and Tobias Fritz. "The inflation technique for causal inference with latent variables". Journal of Causal Inference**7** (2019). Appearances: xii, xiv, 43, 69, 83, 100, 103, 106, 108

[22] Elie Wolfe, Alejandro Pozas-Kerstjens, Matan Grinberg, Denis Rosset, Antonio Acín, and Miguel Navascués. "Quantum inflation: A general approach to quantum causal compatibility". Physical Review X **11**, 021043 (2021). Appearances: xii, 12, 33, 40, 41, 42, 43, 44, 45, 46, 47, 48, 60, 62, 64, 65, 69, 75, 83, 85, 89, 91, 94, 95, 106, 108

[23] Miguel Navascués and Elie Wolfe. "The inflation technique completely solves the causal compatibility problem". Journal of Causal Inference **8**, 70–91 (2020). Appearances: xii, 40, 41, 43, 46, 47, 69, 87

[24] G A Raggio and Reinhard F Werner. "Quantum statistical mechanics of general mean field systems". Helv. Phys. Acta**62** (1989). url: dair.dias.ie/id/eprint/902. Appearances: xiii, 15, 24, 25, 32, 33, 47, 48, 83, 87

[25] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert. ""event-ready-detectors" bell experiment via entanglement swapping". Phys. Rev. Lett. **71**, 4287–4290 (1993). Appearances: xiii, 69

[26] Hoi-Kwong Lo, Marcos Curty, and Kiyoshi Tamaki. "Secure quantum key distribution". Nature Photonics **8**, 595–604 (2014). Appearances: xiii, 105

[27] Sheng-Kai Liao, Wen-Qi Cai, Wei-Yue Liu, Liang Zhang, Yang Li, Ji-Gang Ren, et al. "Satellite-to-ground quantum key distribution". Nature **549**, 43–47 (2017). Appearances: xiii, 105

[28] Yan-Lin Tang, Hua-Lei Yin, Si-Jing Chen, Yang Liu, Wei-Jun Zhang, Xiao Jiang, et al. "Measurement-device-independent quantum key distribution over 200 km". Physical Review Letters **113**, 190501 (2014). Appearances: xiii, 105

[29] P. Kómár, E. M. Kessler, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin. "A quantum network of clocks". Nature Physics **10**, 582–587 (2014). Appearances: xiii, 105

[30] Rhea Parekh, Andrea Ricciardi, Ahmed Darwish, and Stephen DiAdamo. "Quantum algorithms and simulation for parallel and distributed quantum computing" (2021). arXiv:2106.06841. Appearances: xiii, 105

[31] H J Kimble. "The quantum internet". Nature **453**, 1023–1030 (2008). Appearances: xiii, 105

[32] Stephanie Wehner, David Elkouss, and Ronald Hanson. "Quantum internet: A vision for the road ahead". Science **362**, 303 (2018). Appearances: xiii, 105

[33] Angela Sara Cacciapuoti, Marcello Caleffi, Francesco Tafuri, Francesco Saverio Cataliotti, Stefano Gherardini, and Giuseppe Bianchi. "Quantum internet: Networking challenges in distributed quantum computing". IEEE Network **34**, 137–143 (2020). Appearances: xiii, 105

[34] Jessica Illiano, Marcello Caleffi, Antonio Manzalini, and Angela Sara Cacciapuoti. "Quantum internet protocol stack: A comprehensive survey". Computer Networks **213**, 109092 (2022). Appearances: xiii, 105

[35] Marc Hein, Wolfgang Dür, Jens Eisert, Robert Raussendorf, M Nest, and H-J Briegel. "Entanglement in graph states and its applications" (2006). arXiv:quant-ph/0602096. Appearances: xiv, 107

[36] Shiang Yong Looi, Li Yu, Vlad Gheorghiu, and Robert B Griffiths. "Quantum-error-correcting codes using qudit graph states". Physical Review A **78**, 042303 (2008). Appearances: xiv

[37] Denis Rosset, Cyril Branciard, Tomer Jack Barnea, Gilles Pütz, Nicolas Brunner, and Nicolas Gisin. "Nonlinear bell inequalities tailored for quantum networks". Physical review letters **116**, 010403 (2016). Appearances: xiv

[38] Victor Pozsgay, Flavien Hirsch, Cyril Branciard, and Nicolas Brunner. "Covariance bell inequalities". Physical Review A **96**, 062128 (2017). Appearances: xiv

[39] Igor Klep, Victor Magron, Jurij Volčič, and Jie Wang. "State polynomials: positivity, optimization and nonlinear bell inequalities" (2023). arXiv:2301.12513. Appearances: xiv, xv, 31, 92, 132, 133, 134, 137, 138, 139, 140, 141, 142, 143, 144, 146

[40] Stefano Pironio, Miguel Navascués, and Antonio Acín. "Convergent relaxations of polynomial optimization problems with noncommuting variables". SIAM Journal on Optimization **20**, 2157–2180 (2010). Appearances: xiv, 15, 20, 21, 22, 43, 47, 70, 75, 83, 84, 88, 138, 142

[41] Alejandro Pozas-Kerstjens, Rafael Rabelo, Łukasz Rudnicki, Rafael Chaves, Daniel Cavalcanti, Miguel Navascués, and Antonio Acín. "Bounding the sets of classical and quantum correlations in networks". Phys. Rev. Lett. **123**, 140503 (2019). Appearances: xiv, 31, 69, 133, 137

[42] Masamichi Takesaki. "Theory of operator algebras i". Volume 124. Springer Science & Business Media. Berlin (2001). Appearances: 2, 3, 6, 71

[43] Bruce Blackadar. "Operator algebras: theory of C^*-algebras and von neumann algebras". Volume 122. Springer Science & Business Media. Berlin (2006). Appearances: 3, 5, 6, 7, 8, 25, 30, 32, 47, 70, 73, 78, 79, 80, 83, 102, 104

[44] Richard V Kadison and John R Ringrose. "Fundamentals of the theory of operator algebras. vol. i". Oxford University Press. (1983). Appearances: 8, 79

[45] Walter Rudin. "Functional analysis". International series in pure and applied mathematics. McGraw-Hill. (1991). url: books.google.de/books?id=Sh_vAAAAMAAJ. Appearances: 9

[46] Joe Henson, Raymond Lal, and Matthew F Pusey. "Theory-independent limits on correlations from generalized bayesian networks". New Journal of Physics **16**, 113043 (2014). Appearances: 10

[47] Tobias Fritz. "Beyond bell's theorem: correlation scenarios". New Journal of Physics **14**, 103001 (2012). Appearances: 12, 13, 17

[48] Fabio Costa and Sally Shrapnel. "Quantum causal modelling". New Journal of Physics **18**, 063032 (2016). Appearances: 12

[49] Jonathan Barrett, Robin Lorenz, and Ognyan Oreshkov. "Quantum causal models" (2019). arXiv:1906.10726. Appearances: 12

[50] John-Mark A Allen, Jonathan Barrett, Dominic C Horsman, Ciarán M Lee, and Robert W Spekkens. "Quantum common causes and quantum causal models". Physical Review X **7**, 031021 (2017). Appearances: 12

[51] Franco Strocchi. "An introduction to the mathematical structure of quantum mechanics: a short course for mathematicians". Volume 28. World Scientific. London (2008). url: books.google.de/books?id=Bn7MaT3X8fkC. Appearances: 14

[52] Ola Bratteli and Derek William Robinson. "Operator algebras and quantum statistical mechanics: Volume 1: C^*-and W^*-algebras. symmetry groups. decomposition of states". Springer. Berlin (2012). Appearances: 14, 27, 30, 33, 70, 85

[53] Volkher B Scholz and Reinhard F Werner. "Tsirelson's problem" (2008). arXiv:0812.4305. Appearances: 15, 42, 71, 72, 82

[54] Marius Junge, Miguel Navascues, Carlos Palazuelos, David Perez-Garcia, Volkher B Scholz, and Reinhard F Werner. "Connes' embedding problem and Tsirelson's problem". Journal of Mathematical Physics 52, 012102 (2011). Appearances: 15, 42, 71, 72

[55] Tobias Fritz. "Tsirelson's problem and Kirchberg's conjecture". Reviews in Mathematical Physics 24, 1250012 (2012). Appearances: 15, 42, 71, 72

[56] Miguel Navascués, Stefano Pironio, and Antonio Acín. "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations". New Journal of Physics 10, 073013 (2008). Appearances: 15, 18, 20, 70, 88, 96, 97

[57] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. "Linear matrix inequalities in system and control theory". SIAM. (1994). Appearances: 17

[58] Lieven Vandenberghe and Stephen Boyd. "Applications of semidefinite programming". Applied Numerical Mathematics 29, 283–299 (1999). Appearances: 17

[59] Friedrich Pukelsheim. "Optimal design of experiments". SIAM. (2006). Appearances: 17

[60] Farid Alizadeh. "Interior point methods in semidefinite programming with applications to combinatorial optimization". SIAM journal on Optimization 5, 13–51 (1995). Appearances: 17

[61] Martin Grötschel, László Lovász, and Alexander Schrijver. "Geometric algorithms and combinatorial optimization". Volume 2. Springer Science & Business Media. (2012). url: books.google.de/books?id=x1zmCAAAQBAJ. Appearances: 17

[62] Michel X Goemans and David P Williamson. ". 879-approximation algorithms for max cut and max 2sat". In Proceedings of the twenty-sixth annual ACM symposium on Theory of computing. Pages 422–431. (1994). Appearances: 17

[63] Etienne De Klerk, Dmitrii V Pasechnik, and Renata Sotirov. "On semidefinite programming relaxations of the traveling salesman problem". SIAM Journal on Optimization **19**, 1559–1573 (2009). Appearances: 18

[64] Andrew C Doherty, Pablo A Parrilo, and Federico M Spedalieri. "Complete family of separability criteria". Physical Review A **69**, 022308 (2004). Appearances: 18

[65] Felix Huber and Nikolai Wyderka. "Refuting spectral compatibility of quantum marginals" (2022). arXiv:2211.06349. Appearances: 18

[66] Armin Tavakoli, Alejandro Pozas-Kerstjens, Peter Brown, and Mateus Araújo. "Semidefinite programming relaxations for quantum correlations" (2023). arXiv:2307.02551. Appearances: 18

[67] Mirjam Dür and Franz Rendl. "Conic optimization: a survey with special focus on copositive optimization and binary quadratic problems". EURO Journal on Computational Optimization **9**, 100021 (2021). Appearances: 18

[68] Etienne de Klerk and Frank Vallentin. "On the turing model complexity of interior point methods for semidefinite programming". SIAM Journal on Optimization **26**, 1944–1961 (2016). Appearances: 18, 20

[69] Frank Vallentin. "Lecture notes: Semidefinite programs and harmonic analysis" (2008). arXiv:0809.2017. Appearances: 18, 19

[70] Martin Grötschel, László Lovász, and Alexander Schrijver. "The ellipsoid method and its consequences in combinatorial optimization". Combinatorica **1**, 169–197 (1981). Appearances: 20

[71] Monique Laurent. "Sums of squares, moment matrices and optimization over polynomials". Emerging applications of algebraic geometryPages 157–270 (2009). Appearances: 20

[72] Miguel Navascués, Stefano Pironio, and Antonio Acín. "Bounding the set of quantum correlations". Physical Review Letters **98**, 010401 (2007). Appearances: 20

[73] Jean B Lasserre. "Global optimization with polynomials and the problem of moments". SIAM Journal on optimization **11**, 796–817 (2001). Appearances: 20, 133, 138, 139

[74] Erling Størmer. "Symmetric states of infinite tensor products of C^*-algebras". Preprint series: Pure mathematics (1967). url: http://urn.nb.no/URN:NBN:no-8076. Appearances: 24, 25, 47

[75] Carlton M Caves, Christopher A Fuchs, and Rüdiger Schack. "Unknown quantum states: the quantum de finetti representation". Journal of Mathematical Physics **43**, 4537–4559 (2002). Appearances: 24, 47, 87

[76] Renato Renner. "Symmetry of large physical systems implies independence of subsystems". Nature Physics **3**, 645–649 (2007). Appearances: 24, 47

[77] Matthias Christandl, Robert König, Graeme Mitchison, and Renato Renner. "One-and-a-half quantum de finetti theorems". Communications in mathematical physics **273**, 473–498 (2007). Appearances: 24

[78] Bruno De Finetti. "Funzione caratteristica di un fenomeno aleatorio". Atti del Congresso Internazionale dei Matematici: Bologna del 3 al 10 de settembre di 1928Pages 179–190 (1929). url: eudml.org/doc/79004. Appearances: 24, 47

[79] Persi Diaconis and David Freedman. "Finite exchangeable sequences". The Annals of ProbabilityPages 745–764 (1980). url: www.jstor.org/stable/2242823. Appearances: 24, 47, 87

[80] Guillaume Aubrun, Alexander Müller-Hermes, and Martin Plávala. "Monogamy of entanglement between cones" (2022). arXiv:2206.11805. Appearances: 24, 47

[81] G.J. Murphy. "C^*-algebras and operator theory". Academic Press. San Diego (2014). Appearances: 27, 29

[82] Alejandro Pozas Kerstjens. "Quantum information outside quantum information" (2019). Appearances: 31, 69

[83] Marc-Olivier Renou and Xiangling Xu. "Two convergent npa-like hierarchies for the quantum bilocal scenario" (2022). arXiv:2210.09065v2. Appearances: 31, 69, 76, 77, 78, 80, 81, 82, 89, 94, 139

[84] Stephen Curran. "Quantum exchangeable sequences of algebras". Indiana University mathematics journalPages 1097–1125 (2009). url: www.jstor.org/stable/24903244. Appearances: 34, 47

[85] William Slofstra. "The set of quantum correlations is not closed". In Forum of Mathematics, Pi. Volume 7. Cambridge University Press (2019). Appearances: 47

[86] Thomas S Richardson. "A factorization criterion for acyclic directed mixed graphs" (2014). arXiv:1406.6764. Appearances: 63

[87] Mariami Gachechiladze, Nikolai Miklin, and Rafael Chaves. "Quantifying causal influences in the presence of a quantum common cause". Physical Review Letters 125, 230401 (2020). Appearances: 63

[88] Thomas Van Himbeeck, Jonatan Bohr Brask, Stefano Pironio, Ravishankar Ramanathan, Ana Belén Sainz, and Elie Wolfe. "Quantum violations in the instrumental scenario and their relations to the bell scenario". Quantum 3, 186 (2019). Appearances: 63

[89] Iris Agresti, Gonzalo Carvacho, Davide Poderini, Leandro Aolita, Rafael Chaves, and Fabio Sciarrino. "Experimental connection between the instrumental and bell inequalities". Multidisciplinary Digital Publishing Institute Proceedings 12, 27 (2019). Appearances: 63

[90] Iris Agresti, Davide Poderini, Beatrice Polacchi, Nikolai Miklin, Mariami Gachechiladze, Alessia Suprano, Emanuele Polino, Giorgio Milani, Gonzalo Carvacho, Rafael Chaves, et al. "Experimental test of quantum causal influences". Science advances 8, eabm1515 (2022). Appearances: 63

[91] Robin J Evans. "Graphs for margins of bayesian networks". Scandinavian Journal of Statistics 43, 625–648 (2016). Appearances: 64

[92] Masato Koashi and Andreas Winter. "Monogamy of quantum entanglement and other correlations". Physical Review A 69, 022309 (2004). Appearances: 64

[93] Cyril Branciard, Denis Rosset, Nicolas Gisin, and Stefano Pironio. "Bilocal versus nonbilocal correlations in entanglement-swapping experiments". Physical Review A 85, 032119 (2012). Appearances: 69

[94] Askery Canabarro, Samuraí Brito, and Rafael Chaves. "Machine learning nonlocal correlations". Physical review letters 122, 200401 (2019). Appearances: 69

[95] Armin Tavakoli, Alejandro Pozas-Kerstjens, Ming-Xing Luo, and Marc-Olivier Renou. "Bell nonlocality in networks". Reports on Progress in Physics 85, 056001 (2022). Appearances: 69

[96] Klaas Landsman. "Foundations of quantum theory: From classical concepts to operator algebras". Springer Nature. (2017). Appearances: 70

[97] V. Moretti. "Fundamental mathematical structures of quantum theory: Spectral theory, foundational issues, symmetries, algebraic formulation". Springer. (2019). Appearances: 70

[98]   B.S. Tsirelson. "Bell inequalities and operator algebras (open problem 33 in quantum information)" (2007). Appearances: 72, 82

[99]   Persi Diaconis and David Freedman. "A dozen de Finetti-style results in search of a theory". Annales de l'IHP Probabilités et statistiques **23**, 397–423 (1987). url: http://www.brunodefinetti.it/Bibliografia/adozen.pdf. Appearances: 87

[100] Emanuel-Cristian Boghiu, Elie Wolfe, and Alejandro Pozas-Kerstjens. "Inflation: a python library for classical and quantum causal compatibility". Quantum **7**, 996 (2023). Appearances: 93, 146

[101] Boris S Cirel'son. "Quantum generalizations of bell's inequality". Letters in Mathematical Physics **4**, 93–100 (1980). Appearances: 96

[102] Miguel Navascués, Elie Wolfe, Denis Rosset, and Alejandro Pozas-Kerstjens. "Genuine network multipartite entanglement". Physical Review Letters **125**, 240505 (2020). Appearances: 97, 106, 108

[103] Ali Asadian, Paul Erker, Marcus Huber, and Claude Klöckl. "Heisenberg-weyl observables: Bloch vectors in phase space". Physical Review A **94**, 010301 (2016). Appearances: 103

[104] Shuang Wang, Wei Chen, Zhen-Qiang Yin, Yang Zhang, Tao Zhang, Hong-Wei Li, et al. "Field test of wavelength-saving quantum key distribution network". Optics Letters **35**, 2454 (2010). Appearances: 105

[105] M. Sasaki, M. Fujiwara, H. Ishizuka, W. Klaus, K. Wakui, M. Takeoka, et al. "Field test of quantum key distribution in the tokyo QKD network". Optics Express **19**, 10387 (2011). Appearances: 105

[106] Bernd Fröhlich, James F. Dynes, Marco Lucamarini, Andrew W. Sharpe, Zhiliang Yuan, and Andrew J. Shields. "A quantum access network". Nature **501**, 69–72 (2013). Appearances: 105

[107] Sheng-Kai Liao, Wen-Qi Cai, Johannes Handsteiner, Bo Liu, Juan Yin, Liang Zhang, et al. "Satellite-relayed intercontinental quantum network". Physical Review Letters **120**, 030501 (2018). Appearances: 105

[108] Juan Yin, Yu-Huai Li, Sheng-Kai Liao, Meng Yang, Yuan Cao, Liang Zhang, et al. "Entanglement-based secure quantum cryptography over 1,120 kilometres". Nature **582**, 501–505 (2020). Appearances: 105

[109] Cornelia Spee and Tristan Kraft. "Transformations in quantum networks via local operations assisted by finitely many rounds of classical communication" (2021). arXiv:2105.01090. Appearances: 105

[110] Ya-Li Mao, Zheng-Da Li, Sixia Yu, and Jingyun Fan. "Test of genuine multipartite nonlocality in network" (2022). arXiv:2201.12753. Appearances: 106

[111] Xavier Coiteux-Roy, Elie Wolfe, and Marc-Olivier Renou. "Any physical theory of nature must be boundlessly multipartite nonlocal". Physical Review A 104, 052207 (2021). Appearances: 106

[112] Kiara Hansenne, Zhen-Peng Xu, Tristan Kraft, and Otfried Gühne. "Symmetries in quantum networks lead to no-go theorems for entanglement distribution and to verification techniques". Nature Communications 13, 496 (2022). Appearances: 106, 107, 109, 117, 131

[113] Ming-Xing Luo. "New genuinely multipartite entanglement". Advanced Quantum Technologies 4, 2000123 (2021). Appearances: 106, 117, 131

[114] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. "Going beyond Bell's theorem". Pages 69–72. Springer Netherlands. Dordrecht (1989). Appearances: 106

[115] Hans J. Briegel and Robert Raussendorf. "Persistent entanglement in arrays of interacting particles". Physical Review Letters 86, 910–913 (2001). Appearances: 106

[116] Wolfram Helwig. "Absolutely maximally entangled qudit graph states" (2013). arXiv:1306.2879. Appearances: 106

[117] Robert Raussendorf and Hans J. Briegel. "A one-way quantum computer". Physical Review Letters 86, 5188–5191 (2001). Appearances: 106

[118] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest. "Measurement-based quantum computation". Nature Physics 5, 19–26 (2009). Appearances: 106

[119] Xing-Can Yao, Tian-Xiong Wang, Hao-Ze Chen, Wei-Bo Gao, Austin G Fowler, Robert Raussendorf, et al. "Experimental demonstration of topological error correction". Nature 482, 489–494 (2012). Appearances: 106

[120] Mark Hillery, Vladimír Bužek, and André Berthiaume. "Quantum secret sharing". Physical Review A 59, 1829–1834 (1999). Appearances: 106

[121] Géza Tóth and Iagoba Apellaniz. "Quantum metrology from a quantum information science perspective". Journal of Physics A: Mathematical and Theoretical **47**, 424006 (2014). Appearances: 106

[122] Tristan Kraft, Sébastien Designolle, Christina Ritz, Nicolas Brunner, Otfried Gühne, and Marcus Huber. "Quantum entanglement in the triangle network". Physical Review A **103**, L060401 (2021). Appearances: 106

[123] Yi-Xuan Wang, Zhen-Peng Xu, and Otfried Gühne. "Quantum networks cannot generate graph states with high fidelity" (2022). arXiv:2208.12100. Appearances: 106

[124] Jędrzej Kaniewski, Ivan Šupić, Jordi Tura, Flavio Baccari, Alexia Salavrakos, and Remigiusz Augusiak. "Maximal nonlocality from maximal entanglement and mutually unbiased bases, and self-testing of two-qutrit quantum systems". Quantum **3**, 198 (2019). Appearances: 110

[125] Mohsen Bahramgiri and Salman Beigi. "Graph states under the action of local clifford group in non-binary case" (2007). arXiv:quant-ph/0610267. Appearances: 117

[126] Mihai Putinar. "Positive polynomials on compact semi-algebraic sets". Indiana University Mathematics Journal **42**, 969–984 (1993). url: www.jstor.org/stable/24897130. Appearances: 138

[127] J Helton and Scott McCullough. "A positivstellensatz for non-commutative polynomials". Transactions of the American Mathematical Society **356**, 3721–3737 (2004). Appearances: 138

[128] Marc-Olivier Renou, Xiangling Xu, and Laurens T Ligthart. "Two convergent npa-like hierarchies for the quantum bilocal scenario" (2022). arXiv:2210.09065. Appearances: 139

[129] Sabine Burgdorf, Igor Klep, Janez Povh, et al. "Optimization of polynomials in non-commuting variables". Volume 2. Springer. (2016). Appearances: 142

[130] Giorgos Eftaxias, Mirjam Weilenmann, and Roger Colbeck. "Joint measurements in boxworld and their role in information processing" (2022). arXiv:2209.04474. Appearances: 150

[131] Sandu Popescu and Daniel Rohrlich. "Quantum nonlocality as an axiom". Foundations of Physics **24**, 379–385 (1994). Appearances: 151

Index

Zusammenfassung in deutscher Sprache

Die Charakterisierung von Korrelationen, die aus der Durchführung von quantenmechanischen Messungen entstehen können, ist eine relevante, aber auch anspruchsvolle Aufgabe. Eine solche Charakterisierung, bekannt als das Problem der quantenkausalen Kompatibilität, bietet uns zum einen Einblicke in die Vorteile der Quantentheorie gegenüber einer klassischen Theorie, kann aber auch zum anderen deren Einschränkungen aufzeigen. Die Lösung dieses Problems wird besonders herausfordernd, wenn die quantenmechanischen Zustände und Messungen mit einer gegebenen kausalen Struktur kompatibel sein müssen. Eine kausale Struktur bestimmt die kausalen Abhängigkeiten der Parteien und Systeme, die an einem Experiment beteiligt sind. Man kann zum Beispiel das Bell-Szenario als eine der einfachsten kausalen Strukturen betrachten, in der zwei räumlich voneinander getrennte Parteien, Alice und Bob, Messungen an einer gemeinsamen Quelle durchführen. In allgemeineren kausalen Strukturen kann es mehr Parteien und mehr unabhängige Quellen geben.

Kürzlich wurde eine systematische Methode zur Analyse der Korrelationen in klassischen und quantenkausalen Strukturen in Form der sogenannten Inflationstechnik entwickelt. In der Inflationstechnik werden die kausalen Abhängigkeiten, die algorithmisch schwer zu codieren sind, zu leicht zu codierenden Symmetrie-Nebenbedingungen für eine größere Anzahl von Parteien relaxiert. Im Fall der klassischen Physik liefert dies eine konvergierende Hierarchie von linearen Programmen für das Kompatibilitätsproblem. Im Fall der Quantenphysik hingegen führt dies zu einer Hierarchie von immer restriktiver werdenden semidefiniten Programmierungs-Relaxationen. Ob diese Hierarchie auch vollständig ist, bleibt eine offene Frage.

Eines der Hauptergebnisse dieser Dissertation besteht darin zu zeigen, dass eine modifizierte Version der Quanten-Inflationstechnik für das Problem der quantenkausalen Kompatibilität konvergent ist. Diese modifizierte Hierarchie führt einen zusätzlichen Parameter r ein, der den Schmidt-Rang der Observablen einschränkt. Für jeden Wert von r wird eine Hierarchie von Kompatibilitätstests bereitgestellt, die in dem Sinne vollständig ist, dass sie jede Wahrscheinlichkeitsverteilung auf irgendeinem endlichen Niveau der Hierarchie detektiert, die mit dem durch r modifizierten kausalen Modell unvereinbar ist. Solche Kompatibilitätstests werden im Allgemeinen als nicht-kommutative polynomiale Optimierungsprobleme formuliert für die in dieser Arbeit eine C^*-algebraische Beschreibung geliefert wird.

Zusätzlich wird eine separate Hierarchie von semidefiniten Programmen, die Polarisation-Hierarchie, entwickelt. Es wird gezeigt, dass sowohl die Polarisation-Hierarchie als auch die ursprüngliche Quanten-Inflation-Hierarchie für die kausale Struktur, die als das bilokale Szenario bekannt ist, vollständig ist. Im bilokalen Szenario gibt es drei räumlich weit voneinander entfernte Parteien, Alice, Bob und Charlie, die Messungen an zwei unabhängigen Quellen durchführen: eine, die von

Alice und Bob geteilt wird, und eine andere, die von Bob und Charlie geteilt wird. Es wird gezeigt, dass ein Modell für Bobs Algebra, welche aus zwei kommutierenden Teilalgebren besteht, aus den Kommutanten von Darstellungen von Alices und Charlies Algebren konstruiert werden kann. Diese Konstruktion gibt auch Einblick in eine bilokale Version von Tsirelsons Problem. Außerdem wird gezeigt, dass wenn Alices und Charlies Systeme mit einer endlichdimensionalen Darstellung modelliert werden können, das Modell der kommutierenden Observablen und das Tensorproduktmodell der Lokalität in der Quantentheorie übereinstimmen. Diese Konvergenzergebnisse stützen sich auf die Tatsache, dass bestimmte Symmetrien im Grenzfall die Unabhängigkeit von Zufallsvariablen oder Quantenzuständen implizieren. Solche Aussagen sind im Allgemeinen bekannt als de Finetti-Theoreme. Für den Spezialfall der C^*-algebraischen Beschreibung der Quantenmechanik wie er in dieser Dissertation untersucht wird, war ein de Finetti-Theorem über den Sonderfall des minimalen Tensorprodukts hinaus noch nicht bewiesen worden. Ein weiteres Ergebnis der Dissertation ist der Beweis, dass ein quantenmechanisches de Finetti-Theorem auch für allgemeine Tensorprodukte von C^*-Algebren gilt.

Das Problem der quantenkausalen Kompatibilität kann als eine Version des Kompatibilitätsproblems in Quantennetzwerken angesehen werden, welches sich mit der Frage beschäftigt, welche Quantenzustände in einer bestimmten kausalen Struktur erzeugt werden können unter der Annahme, dass das Endprodukt ein klassischer Zustand ist. In dieser Arbeit wird gezeigt, dass die für das kausale Kompatibilitätsproblem entwickelten Techniken auch auf den allgemeineren Fall von Quantennetzwerken angewendet werden können. Darüber hinaus wird ein analytischer Beweis dafür erbracht, dass Graphzustände nicht in bipartiten Quantennetzwerken erzeugt werden können. Dieser Beweis stützt sich erneut auf die Inflationstechnik, indem Korrelationen unterschiedlicher Inflationen des Netzwerks miteinander verknüpft werden. Nimmt man nun an, dass die Korrelationen aus einem Graphzustand resultieren, kann gezeigt werden, dass dies die Schranke einer bestimmten Ungleichung verletzt, was zu einem Widerspruch führt.

Zu guter Letzt wird gezeigt, dass die Polarisation-Hierarchie dazu verwendet werden kann, eine große Klasse von Optimierungsproblemen zu lösen, die als Zustandspolynom Optimierungsprobleme bekannt sind. Bei Zustandspolynom Optimierungsproblemen besteht das Ziel darin, die Ziel- und Nebenbedingungsfunktionen zu optimieren, welche polynomiale Funktionen der Erwartungswerte von Observablen sind. Dies ermöglicht es einem beispielsweise, die Kovarianzen oder nichtlineare Bell-Ungleichungen zu optimieren. Darüber hinaus wird auch eine alternative Version einer kürzlich entwickelten Hierarchie dargestellt, die dieses Problem ebenfalls löst und in welche der Polarisationstrick integriert ist.