ETH ZURICH

Semi-device-independent self-testing of a single quantum system based on Spekkens contextuality

by

Janek Denzler

Master thesis

Supervisors:
Dr. L. del Rio,
Prof. Dr. R. Renner

Department of Physics

April 2021

Abstract

Self-testing allows for powerful inferences about the quantum properties of an adversarial input-output device from a minimal set of assumptions, by only interacting with the device classically. The vast majority of self-testing protocols rely on Bell non-locality. This thesis explores the possibility of self-testing local input-output devices, without the need for entanglement.

Kochen-Specker contextuality, a generalization of Bell non-locality to single-partite systems, was identified as a resource for self-testing in [5]. We study their protocol in detail, derive an improved robustness to noise, and perform a numerical analysis of the constant of proportionality, for the simplest case. Carefully pointing out several problematic assumptions of the protocol, we argue for ways to adapt it.

Lastly, we propose a novel protocol based on Spekkens' operational notion of contextuality, replacing unphysical assumptions regarding noise-free and commuting measurements with approximate operational equivalences. Furthermore, the certificate of quantumness we propose is inherently robust to noise.

Contents

Chapter 1

Introduction

The term "self-testing" was coined by Mayers and Yao in 2004 [28]. Despite its adolescent age, the field is already at the forefront of quantum cryptography and the pillar on which countless protocols rest. The aim of self-testing is to acquire knowledge about the quantum properties of an adversarial input-output device from a minimal set of assumptions, by only interacting with the device classically, i.e. passing an input bitstring and receiving an output bitstring for each input-output cycle. The vast majority of self-testing protocols rely on Bell non-locality. As such, they assume that the device is a multipartite system that is split into space-like separated subsystems that are non-communicating for the duration of one input-output cycle. The key observation giving rise to the notion of self-testing is that there exist Bell non-local correlations that are compatible only with an essentially unique quantum model, comprised of quantum state and measurement operators. Bell non-locality further certifies that the input-output correlations were not simulated by a hackable classical device. While many technical tools have been developed to bound compatible quantum models for Bell-type self-testing scenarios, the task of certifying quantum properties of unipartite quantum system remains largely underexplored. A promising undertaking is to identify sets of assumptions that facilitate self-testing of a unipartite system, without using entanglement and Bell non-locality as relevant resources. Apart from providing insight into the geometry of quantum correlations for more general scenarios, the potential usefulness of such protocols lies in the fact that the no-communication assumption underlying conventional self-testing approaches is hard to implement in practice. Additionally, entanglement is often a very costly resource. The natural starting point we take on in this thesis is to study the role of Kochen-Specker contextuality as a resource for self-testing, as it can be seen as an extension of Bell non-locality to more general, unipartite scenarios.

This thesis discusses and recasts previous results [5] that identify Kochen-Specker contextuality as a resource for self-testing to a general prepare-and-measure scenario. In particular, we propose

a protocol based on Spekkens contextuality, discuss its features, and identify conditions that allow us to relax unphysical restrictions on measurement operators, imposed in [5]. Along the way, we characterize the resources in terms of memory needed by a pre-programmed classical device to simulate the optimal correlations to the protocol in [5], and present an improved proof of robustness to noise.

The thesis is organized as follows: In Section 2, we introduce the notion of Kochen-Specker contextuality, which aims to capture the way in which quantum mechanics deviates from our classical intuition, even for unipartite systems. We will derive a class of Kochen-Soecker non-contextuality inequalities in Section 2.9 that delimit classical from quantum correlations, analogously to Bell inequalities in the multipartite setting. In Section 3, we introduce graph theoretic tools, proposed in [11], that allow for a rigorous analysis of these Kochen-Specker non-contextuality inequalities. In particular, these tools enable us to identify and study optimal classical behaviours that saturate the non-contextual bounds, as well as optimal quantum models that produce maximal violations of the inequalities, relevant to self-testing. Section 4 introduces the concept of self-testing in the context of multipartite Bell scenarios, which allows for powerful inferences about the quantum properties of a correlation experiment from minimal assumptions. Additionally, Section 4 points out in what ways this concept can be modified to accommodate unipartite systems with no space-like separation. Equipped with the graph-theoretic tools presented in Section 3, we will demonstrate that the Kochen-Specker non-contextuality inequalities in 2.9 facilitate noise-robust self-testing, as was proved in [5], and improve the robustness to noise. Section 2.9 also identifies the assumptions that go into this self-testing protocol. In particular, we find that one has to bound the information carrying capacity of the device to exclude the possibility of a pre-programmed classical computer generating the correlations. In an attempt to relax some of the unphysical constraints, such as the restriction to sharp measurements, we shift to a general prepare-and-measure scenario. Spekkens revised notion of non-contextuality, which will be introduced in Section 6, is formulated within this general framework. Spekkens operational notion of contextuality will provide us with more suitable and robust tests of non-classicality. Finally, in Section 7, we submit a self-testing protocol based on Spekkens contextuality and discuss its merits and drawbacks, compared to the protocol in Section 5.

Chapter 2

Kochen-Specker contextuality

The Kochen-Specker (KS) Theorem [26] is the often overlooked brother of Bell's monumental no-go theorem [3]. In loose terms, Bell's Theorem proves that quantum mechanics (QM) cannot be described in terms of a local hidden variable model (HVM). A "Bell inequality" is a condition measurement correlations must obey, for them to permit a local HVM description. Violations of Bell inequalities have been observed in numerous experiments: The well-known Delft experiment of 2015 was the first loophole-free confirmation of such violations [21]. As will be discussed in Section 2.8, Bell locality and KS non-contextuality are closely related, locality being the weaker assumption and justified by special relativity. Nevertheless, the KS Theorem has several distinct features that deserve to be studied in their own right. Most notably, KS contextuality allows one to examine how even single quantum systems are incompatible with a seemingly orthodox classical description in terms of hidden variables [37]. In contrast, Bell tests require two or more space-like separated quantum systems, in order for the measurement statistics to exhibit Bell non-locality. The reason for this is that we observe violations of Bell inequalities for entangled quantum systems. An immediate consequence is that the notion of KS contextuality applies to a much broader class of experiments.

2.1 Introductory example

Specker's three boxes

Section 2.5 will give a formal definition of KS non-contextuality. Nevertheless, we shall first consider a simple example that illustrates the concept. It is taken from Specker's 1960 paper "The logic of propositions which are not simultaneously decidable" [39], which also gives a nice accompanying parable about an overprotective seer that teaches at the school of prophets in the fictitious place Arba'ila. Fairy tales aside, consider three boxes, labelled with 1, 2, 3, that can each be in one of two states: empty or containing a gem. The state of a box

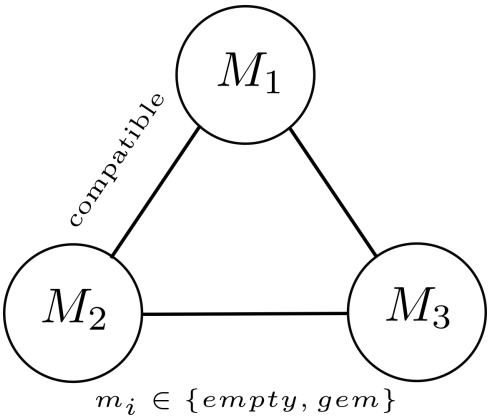


Figure 2.1: Specker's "three boxes" example. Three boxes 1,2,3 can each either hold a gem or be empty. The measurement M_i corresponds to opening the lid of the ith box and checking its contents. The outcome of measurement M_i is denoted by m_i. A solid line between measurements means that these are compatible and can be measured simultaneously. Despite all measurements being pairwise compatible, only pairs of boxes can be jointly measured. Measuring pairs of boxes always yields anti-correlated outcomes.

can be measured by opening the lid and peeking inside. Let M_1, M_2, M_3 denote the three two-outcome measurements (with outcomes "empty" or "gem") of the boxes. Further assume, rather unnaturally, that it is impossible to open all three boxes simultaneously. Rather, only pairs of boxes may be jointly measured. Figure 2.1 depicts the setting of Specker's "three boxes" example. For some repeatable preparation of the system, one curiously observes the following: opening pairs of boxes always yields anti-correlated measurement outcomes (that is, one of the boxes will be empty and the other will contain a gem). How can these statistics be reconciled with a naive model of reality that assumes the three boxes to be in some underlying state that specifies their contents? As we can only ever open two boxes simultaneously, we can never observe this hidden state of reality. Importantly, a valid assignment of gems to the three boxes must be compatible with the observed anti-correlated outcomes. It turns out, as the reader should convince themself, that any assignment of gems to the boxes is incompatible with the observed statistics. Assuming that the pairs of boxes to be measured are picked at random, in a uniform way, the probability of observing anti-correlated outcomes is bounded by $p(\text{anti-correlated}) \leqslant \frac{2}{3}$ and we conclude that the naive type of HVM we just considered is unable to account for the experimental data.

The point is that we have in fact considered a deterministic and non-contextual HVM to describe the three boxes. Deterministic just means that the underlying state of the system fixes all measurement outcomes and not just a probability distribution over possible outcomes. More interestingly, we define non-contextuality for the time being as the critical assumption that the outcome of some measurement should be the same, regardless of which other compatible[1] measurement is performed simultaneously. Allowing the outcome of a measurement to depend

[1] Two measurements are called *compatible* if they can be performed simultaneously. While this "definition" is sufficient for now, we will refine it in Section 3.1.

on what other compatible measurement is performed together with it, could indeed account for the observed correlations.

2.2 The Kochen-Specker Theorem, informally

Informally, the KS Theorem states the following:

Theorem 2.1 (Kochen-Specker, informal)**.** *A non-contextual deterministic HVM of QM for Hilbert spaces of dimension* ≥ 3 *is impossible.*

It is tempting to try to recast Specker's "three boxes" example to the realm of quantum theory. Within the theory of QM, compatible measurements are conventionally represented by commuting Hermitian operators. To implement the example in QM, we would have to find a quantum state $|\Psi\rangle$ and pairwise commuting two-outcome measurements M_1, M_2, M_3 not all simultaneously measurable, such that QM predicts $p(\text{anti-correlated}) > \frac{2}{3}$. For the same reasons, this would imply that non-contextual deterministic HVM are not compatible with the predictions of QM and would prove the KS Theorem, albeit initially only for a fixed dimension that may be greater than 3. However, if all three measurements are pairwise commuting, they necessarily have a joint eigenbasis. Thus, all three measurements could be simultaneously performed and correlations with $p(\text{success}) > \frac{2}{3}$ could not arise. Despite this, Sections 2.6, 2.7 feature examples that are similar in spirit, if a little more involved, and which can indeed be implemented by QM.

Specker hypothesized pairwise compatibility implying global compatibility to be the "fundamental theorem of QM" that shapes the set of quantum correlations [9]. This is supported by the fact that for many correlation experiments the "theorem" singles out correlations that are compatible with QM (and projective measurements) and offers an explanation for why quantum correlations cannot exceed certain mystifying bounds we will encounter in Section 3.3 [10].

Before we can discuss the KS Theorem and its proof in formal terms, as is done in Section 2.6, we will begin by establishing a general mathematical framework of HVM in Section 2.3. Section 2.4 introduces two ways of representing measurements in QM, leading to two equivalent notions of KS non-contextuality, as defined in Section 2.5. Section 2.7 presents a considerably simpler proof of the KS Theorem that holds for four-dimensional systems. The connection between KS non-contextuality and local causality will be the subject of Section 2.8, highlighting the similarities and differences between the KS Theorem and Bell's Theorem. This discussion will point out several flaws of the traditional notion of KS contextuality and addresses the need for an operational notion. A proposal of that sort by Spekkens will be introduced later in Section 6. Finally, in Section 2.9 we will discuss an important class of KS contextuality experiments

involving a qutrit, namely the KCBS scenario and generalizations of it. For these experiments one can derive KS non-contextuality inequalities on the convex set of possible correlations, analogously to Bell inequalities, that must hold for all non-contextual statistics. Interestingly, these inequalities turn out to be incompatible with the predictions of QM for some quantum states.

2.3 General framework of hidden variable models

An ontological model of QM[2], more commonly referred to as a hidden variable model (HVM), presupposes some ontic state space Λ. An ontic state $\lambda \in \Lambda$ is a state of reality or "matter of fact" description of the system. It holds all attributes of the system, regardless of whether they are known or even knowable (hence hidden variables). Importantly, a HVM posits that our system is at each time point in some ontic state with corresponding definite attributes, even when these are not being subjected to measurements. In the case of KS contextuality we will only consider **deterministic** HVM, meaning that an ontic state fixes the outcomes of all possible measurements that may be performed on the system. Uncertainty in the ontic state that was prepared results in a probabilistic distribution of measurement outcomes. It is helpful to formalize this and boil the assumptions down to a simple mathematical model. We follow the notation and terminology of [40].

With any preparations P one can associate a probability density function $\mu_P(\lambda)$ on the ontic state space Λ. One can think of a preparation as an ensemble of systems, much like in the spirit of statistical mechanics, whereby each individual member of the ensemble is in some well-defined ontic state $\lambda \in \Lambda$. The probability of picking a system from the ensemble at random with a hidden parameter contained in some region $\mathcal{A} \subset \Lambda$, is then given by $\int_{\mathcal{A}} \mu_P(\lambda) d\lambda$. By normalization, we require that $\int_\Lambda \mu_P(\lambda) d\lambda = 1$.

Measurements M can be associated with sets of indicator functions $\{\xi_{M,k}\}_k$ on Λ, one for each measurement outcome k. The value $\xi_{M,k}(\lambda) \in \{0, 1\}$ is the conditional probability of the measurement M yielding the outcome k, given that the system is in the ontic state λ. Since we are considering only deterministic HVM, meaning that the ontic state of the system fixes all measurement outcomes with certainty, this conditional probability must be either 0 or 1. Naturally, the indicator functions for the different measurement outcomes must obey $\sum_k \xi_{M,k}(\lambda) = 1$ $\forall \lambda \in \Lambda$. Section 2.4 will introduce two ways of representing measurements in QM. We will apply the framework introduced in this section to these, in order to treat quantum measurements within HVM. Lastly, for the HVM to be consistent with the testable predictions of the operational theory it should model, it must hold that: $p(k|P, M) = \int_\Lambda d\lambda\, \xi_{M,k}(\lambda) \mu_P(\lambda)$, where

[2]The following applies more generally for ontological models of arbitrary operational theories. An operational theory is one that makes testable predictions.

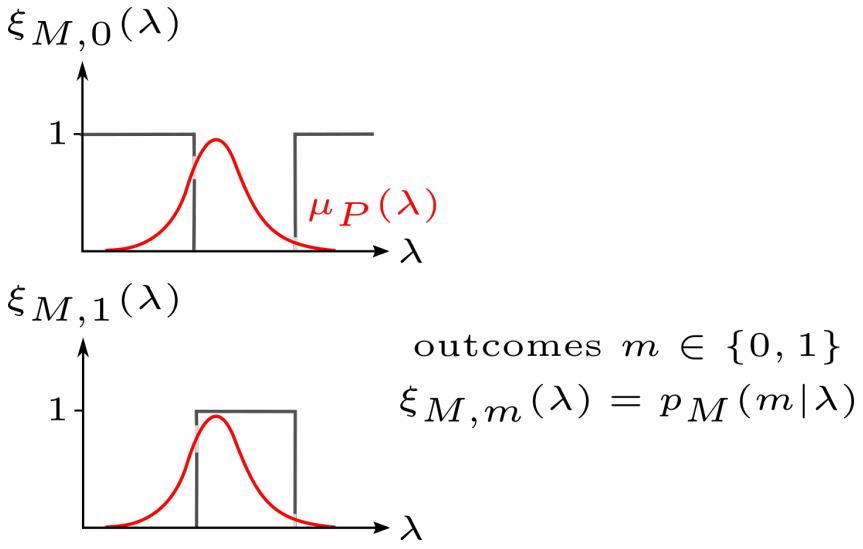


Figure 2.2: Summary of the constituents characterizing a deterministic HVM for the case of a two-outcome measurement.
A system's ontic state space Λ, here depicted as a one-dimensional axis, comprises all possible states of reality. A state preparation P can be associated with a probability density function on Λ, modelling the uncertainty in the prepared ontic state. A deterministic HVM associates with any measurement M, here a two-outcome measurement with outcomes $m \in \{0, 1\}$, a set of indicator functions $\{\xi_{M,m}\}_m$. The indicator function to measurement outcome m, $\xi_{M,m} : \Lambda \to \{0, 1\}$, maps each ontic state to the probability of observing outcome m, conditioned on the system being in that ontic state.

$p(k|P, M)$ describes the probability of observing the measurement outcome k, given measurement M and preparation P, as predicted by the operational theory. The above quantities are pictorally summarized in Figure 2.2, for the case of a two-outcome measurement M. In the following, when making reference to an arbitrary deterministic HVM of QM, we will implicitly assume an underlying ontic state space Λ and indicator functions $\{\xi_{M,k}\}_k$ for all measurements M with possible outcomes k.

2.4 Measurements in quantum mechanics

Section 2.5 will define the notion of KS non-contextuality that is at the heart of the KS Theorem. It is an assumption about the type of HVM the theorem rules out and, in loose terms, pertains to the measurements within such a HVM. To set the stage for discussing KS non-contextuality, this section introduces two ways of representing (projective) measurements in QM. As will be proved in Section 2.5, the two notions of measurements will give us two equivalent notions of KS non-contextuality, each of which will be particularly useful for certain purposes. Note that within the framework of KS contextuality we assume all quantum measurements to be projective. On the other hand, in quantum information theory, a general quantum measurement is given by a set of positive semi-definite operators that sum to the identity, a so-called positive operator valued measure (POVM) [33]. A projective measurement, or projection-valued measure (PVM),

is also a POVM, however the reverse generally does not hold. The two are related in the sense that every POVM can be seen as a PVM acting on an extended Hilbert space, followed by us discarding some degrees of freedom, a consequence of Naimark's dilation theorem [46]. Thus, assuming all quantum measurements to be projective is in principle consistent with the existence of POVMs. In Section 6, we present a revision of the traditional notion of KS contextuality due to Spekkens. Spekkens contextuality is defined largely in terms of operational primitives and is able to accommodate more general POVM quantum measurements. As physical measurements are non-ideal, we must account for these in view of our goal to certify quantum properties of a single quantum device under operational assumptions.

The conventional way of representing projective measurements in QM is in terms of Hermitian operators, with every measurement being relatable to a set of pairwise commuting and thus simultaneously measurable operators that share a common eigenbasis. Measurement outcomes are tuples of simultaneous eigenvalues. The spectrum of a Hermitian operator A is denoted by $\sigma(A)$ and the set of all Hermitian operators on a Hilbert space $\mathcal{H}$ is denoted by $\mathrm{Herm}(\mathcal{H})$.

Definition 2.2. Let $\mathcal{H}$ be the system Hilbert space.
A *projective quantum measurement* $\mathcal{M} = (X_1, X_2, \dots) \subset \mathrm{Herm}(\mathcal{H})$ is an ordered set of pairwise commuting Hermitian operators.
An *outcome of a projective quantum measurement* $\mathcal{M}$ is a tuple
$(x_1, x_2, \dots) \in \bigtimes_{X_i \in \mathcal{M}} \sigma(X_i)$ of simultaneous eigenvalues to operators in $\mathcal{M}$, meaning that there exists a common eigenvector $|\Psi\rangle \in \mathcal{H} \backslash \{0\}$ with $X_k |\Psi\rangle = x_k |\Psi\rangle \ \forall X_k \in \mathcal{M}$.

An ordered set of commuting Hermitian operators like in 2.2 is also referred to as a "measurement context". We say that X_1 is measured in the context $\mathcal{M}$. An example of a quantum measurement in this sense is given by "Pauli strings" on the composite Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$, for example

$$\mathcal{M} = (\sigma_x^1 \coloneqq \sigma_x \otimes \mathbb{1},\ \sigma_x^2 \coloneqq \mathbb{1} \otimes \sigma_x,\ \sigma_x^1 \sigma_x^2 \coloneqq \sigma_x \otimes \sigma_x)$$

with outcomes

$$\{(1,1,1), (1,\text{-}1,\text{-}1), (\text{-}1,1,\text{-}1), (\text{-}1,\text{-}1,1)\}\,.$$

Each of the four outcome tuples corresponds to one of the four joint eigenstates that together span the four-dimensional Hilbert space. Additionally, each outcome tuple contains three values, corresponding to the three pairwise commuting Hermitian operators that constitute the measurement $\mathcal{M}$. We will revisit operators of this form in Section 2.7, when discussing a simple, algebraic proof of the KS Theorem for four-dimensional Hilbert spaces.

Let us now apply the framework proposed in Section 2.3 to projective quantum measurements like in Defintion 2.2. Let A, B be two commuting Hermitian operators. If we assume the ontic state $\lambda \in \Lambda$ to be fixed, $\xi_{(A,B),(a,b)}(\lambda)$ describes a joint probability distribution

$P^{\lambda}_{\mathcal{AB}}(a,b) = \xi_{(A,B),(a,b)}(\lambda)$ for random variables $\mathcal{A}, \mathcal{B}$, that hold eigenvalues of A, B, respectively. $P^{\lambda}_{\mathcal{AB}}(a,b)$ gives the probability of obtaining outcome (a,b) when performing the joint measurement (A,B), conditioned on the system being in the ontic state λ. We can go from a joint probability distribution $P^{\lambda}_{\mathcal{AB}}$ to the "marginal" distribution $P^{\lambda}_{\mathcal{A}}$ for random variable $\mathcal{A}$, by summing over all $b \in \sigma(B)$: $\sum_{b\in\sigma(B)} P^{\lambda}_{\mathcal{AB}}(a,b) = P^{\lambda}_{\mathcal{A}}(a)$. We will revisit marginals of joint outcome probability distributions when defining the notion of KS non-contextuality in Section 2.5.

An alternative way of representing measurements in QM is motivated by the spectral decomposition of Hermitian operators. It defines projective quantum measurements in terms of rays in a projective Hilbert space $\mathcal{P}(\mathcal{H})$. The projective Hilbert space $\mathcal{P}(\mathcal{H})$ of a complex Hilbert space $\mathcal{H}$ is defined as the set of equivalence classes $\{[\,|\Psi\rangle\,]\}_{|\Psi\rangle\in\mathcal{H}\setminus\{0\}}$ with respect to the equivalence relation $|\Psi\rangle \sim |\Phi\rangle :\Leftrightarrow \exists\alpha \in \mathbb{C}\setminus\{0\} : |\Psi\rangle = \alpha\,|\Phi\rangle$. The equivalence classes of $\mathcal{P}(\mathcal{H})$ are called "rays". Intuitively, one can think of $\mathcal{P}(H)$ as the "unit sphere" of normalized vectors in $\mathcal{H}$, with every $\mathcal{H} \ni |\Psi\rangle \neq 0$ being identified with its normalized counterpart. In this picture, a measurement can be specified by an orthonormal basis (ONB) $\{|\Psi_i\rangle\}_i$ of the Hilbert space, with measurement outcomes being represented by rays.

Definition 2.3. Let $\mathcal{H}$ be the system Hilbert space.
A *projective quantum measurement with respect to a basis* $\mathcal{M} = \{|\Psi_i\rangle\}_i \subset \mathcal{P}(\mathcal{H})$ is specified by an ONB of $\mathcal{H}$. An *outcome* of such a measurement $\mathcal{M}$ is specified by a ray $|\Psi_k\rangle \in \mathcal{M}$. The Hermitian measurement operator associated with the outcome $|\Psi_k\rangle \in \mathcal{M}$ is given by the rank one projector $|\Psi_k\rangle\langle\Psi_k|$.

As will be discussed in the next section, the ONB an outcome ray is associated with is also referred to as the measurement context it belongs to. A given ray can belong to multiple measurement contexts.

2.5 Defining KS non-contextuality

We now turn to defining what it means for a HVM of QM to be KS non-contextual. There are two equivalent definitions of KS non-contextuality for deterministic HVM of QM, both commonly found in literature on the topic. As advertised in 2.4, these stem from the two notions of measurements in QM. Key to both definitions of KS non-contextuality will be interpreting a measurement as "revealing" a true property of the system. "True" means that the system has this property, regardless of whether and how the property is measured. The definitions presented here are adapted from the lecture series [41].

Let us first consider measurements $\mathcal{M}$ being represented by Hermitian operators, as in Definition 2.2. Given the ontic state of the system, a general deterministic HVM simultaneously fixes all

possible measurement outcomes via the indicator functions $\xi_{\mathcal{M},k}(\lambda)$. A priori, for a given ontic state $\lambda \in \Lambda$, the value revealed upon measuring some $A \in \mathrm{Herm}(\mathcal{H})$ may depend on what other compatible operators are measured together with A. These constitute the "measurement context". For instance, assume A, B and A, C to be pairs of commuting operators in $\mathrm{Herm}(\mathcal{H})$. Further assume B and C to be non-commuting and thus not simultaneously measurable. We can measure A in two different contexts, together with B or together with C. For a given ontic state λ, these measurements may yield a different outcome for A. However, for a HVM to be consistent with the interpretation of measurements revealing "true" properties of the system, the outcome of measuring an operator may not depend on the measurement context. The assumption of KS non-contextuality demands a functional relationship between Hermitian measurement operators and their "true value":

Definition 2.4. Let $\mathcal{H}$ be the system Hilbert space and represent measurements like in Definition 2.2. A deterministic HVM is *KS non-contextual* if it satisfies Conditions 1 and 2 for all ontic states $\lambda \in \Lambda$:

1. (context independence)

$$\forall A, B \in \mathrm{Herm}(\mathcal{H}),\, [A, B] = 0 :\, \xi_{A,a}(\lambda) = \sum_{b \in \sigma(B)} \xi_{(A,B),(a,b)}(\lambda)$$

   Thus the outcome assignment $\nu_\lambda : \mathrm{Herm}(\mathcal{H}) \to \mathbb{R}$*, defined via*

$$\nu_\lambda(A) = a \in \sigma(A) \iff \xi_{A,x}(\lambda) = \delta_{ax}$$

   is well-defined and consistent with all measurement contexts.

2. (commuting observables are assigned simultaneous eigenvalues)

$$\begin{aligned} &\forall A, B \in \mathrm{Herm}(\mathcal{H}),\, [A, B] = 0 : \\ &\nu_\lambda(A) = a,\, \nu_\lambda(B) = b \implies \exists\, |\Psi\rangle \in \mathcal{H} \backslash \{0\} \text{ with} \\ &\qquad\qquad A\,|\Psi\rangle = a\,|\Psi\rangle\,, \\ &\qquad\qquad B\,|\Psi\rangle = b\,|\Psi\rangle \end{aligned}$$

Condition 1 in Definiton 2.4 requires that, for compatible operators, taking the marginal of the joint probability distribution yields the single operator probability distribution, for all measurement contexts. This just means that $\nu_\lambda(A)$ assigns the correct outcome to measurement A for the ontic state λ, as given by the deterministic HVM, no matter what other compatible measurements are performed jointly. It is crucial that this is required to hold at the ontic level: if we take the ensemble average, as outlined in Section 2.4, this property would always hold in QM. The notion of KS non-contextuality in terms of Hermitian operators is summarized in Figure 2.3.

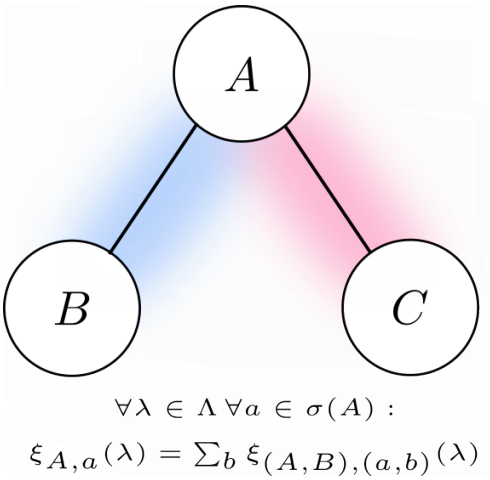


Figure 2.3: Summary of the notion of KS non-contextuality in terms of Hermitian operators. A, B and B, C are assumed to be commuting operators, whereas B and C do not commute. For every ontic state $\lambda \in \Lambda$, a deterministic non-contextual HVM assigns simultaneous eigenvalues (a, b) to the commuting operators A, B. Context independence requires that taking the marginal of the joint statistics $\xi_{(A,B),(a,b)}$ must reduce to the single operator assignment $\xi_{A,a}$.

An important consequence of Condition 2 in Definition 2.4 is often coined "functional consistency" [35]. We will only need and prove the property of "functional consistency" for functional relationships which are polynomial, however it could be extended to smooth functions with a converging (multivariate) Taylor expansion. Note that, strictly speaking, context independence implies Condition 2, as permissible outcomes of joint measurements are by definition simultaneous eigenvalues.

Lemma 2.5 (functional consistency)**.**
Let $\nu_\lambda : \mathrm{Herm}(\mathcal{H}) \to \mathbb{R}$ be a KS non-contextual value assignment, according to Definition 2.4, and let $f : \mathbb{R}^n \to \mathbb{R}$, $(x_1, x_2, \ldots, x_n) \mapsto f(x_1, x_2, \ldots, x_n)$ be a polynomial function in n real variables. Further assume that $f(X_1, X_2, \ldots, X_n) = 0$ holds as an operator identity among the observables of a pairwise commuting set $\{X_1, X_2, \ldots, X_n\} \subset \mathrm{Herm}(\mathcal{H})$. Here $f(X_1, X_2, \ldots, X_n)$ is the operator expression obtained by formally inserting operators X_i into the real variables x_i. Then it must hold that $f(\nu_\lambda(X_1), \nu_\lambda(X_2), \ldots, \nu_\lambda(X_n)) = 0$.

Proof. By definition, the value assignment ν_λ assigns simultaneous eigenvalues to the set of pairwise commuting operators. This means that there exists a joint eigenvector $\mathcal{H} \ni |\Psi\rangle \neq 0$ that corresponds to the eigenvalues assigned by ν_λ. The operator expression $f(X_1, X_2, \ldots, X_n)$ is a sum containing terms of the form $c \prod_{i=1}^n X_i^{\alpha_i}$, with $c \in \mathbb{R}$, $\alpha_i \in \mathbb{N}_0$. Acting with $f(X_1, X_2, \ldots, X_n)$ on the joint eigenvector $|\Psi\rangle$ thus "evaluates" f at $(\nu_\lambda(X_1), \nu_\lambda(X_2), \ldots, \nu_\lambda(X_n))$:

$$f(X_1, X_2, \ldots, X_n)\,|\Psi\rangle = f(\nu_\lambda(X_1), \nu_\lambda(X_2), \ldots, \nu_\lambda(X_n))\,|\Psi\rangle = 0.$$

This concludes the proof, as $|\Psi\rangle$ is non-zero. □

Corollary 2.6. *Let X_1, X_2 be arbitrary commuting observables and $\nu_\lambda : \mathrm{Herm}(\mathcal{H}) \to \mathbb{R}$ a KS non-contextual value assignment. It holds that:*

1. $\nu_\lambda(X_1 + X_2) = \nu_\lambda(X_1) + \nu_\lambda(X_2)$ [3]
2. $\nu_\lambda(X_1 X_2) = \nu_\lambda(X_1)\,\nu(X_2)$

Proof.

1. Define $X = X_1 + X_2$, which commutes with both X_1 and X_2, and apply Lemma 2.5 to the operator identity $X - X_1 - X_2 = 0$.
2. $X_1 X_2$ is a Hermitian operator that commutes with both X_1 and X_2. Apply Lemma 2.5 to the operator identity $(X_1 X_2) - (X_1)(X_2) = 0$.

□

Let us now turn to the definition of KS non-contextuality in terms of rays in a projective Hilbert space.

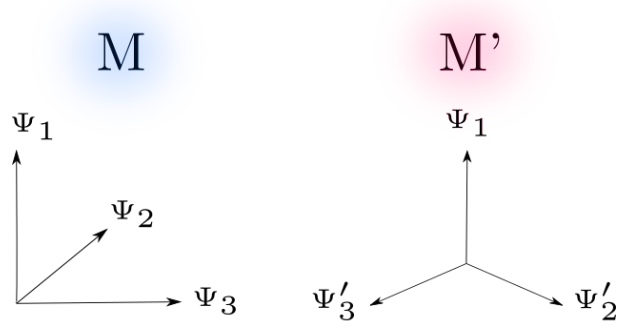
Definition 2.7. Let $\mathcal{H}$ be the system Hilbert space and represent measurements as in Definition 2.3. A deterministic HVM is *KS non-contextual* if for all ontic states $\lambda \in \Lambda$ there is a value assignment $\omega_\lambda : \mathcal{P}(\mathcal{H}) \to \{0, 1\}$ satisfying:

1. (context independence)
   $\forall\, |\Psi\rangle \in \mathcal{P}(\mathcal{H})$, $\forall$ measurements $\mathcal{M}$ with $|\Psi\rangle \in \mathcal{M}$: $\omega_\lambda(|\Psi\rangle) = \xi_{\mathcal{M},|\Psi\rangle}(\lambda)$
2. (exactly one outcome)
   $\forall$ measurements $\{|\Psi_i\rangle\}_i \subset \mathcal{P}(\mathcal{H})$ it holds that $\sum_i \omega_\lambda(|\Psi_i\rangle) = 1$

The interpretation is as follows: think of a measurement $\{|\Psi_i\rangle\}_i \subset \mathcal{P}(\mathcal{H})$ as a set of yes/no questions posed to the system. For each outcome $|\Psi_k\rangle$, a measurement "asks" if this is a true property of the system, fixed by its ontic state, and subsequently "reveals" the answer:

Given the ontic state of the system, a general deterministic HVM simultaneously fixes all possible measurement outcomes via the indicator functions $\xi_{\mathcal{M},k}(\lambda)$. This gives us an assignment similar to that above, with the crucial difference that the value assigned to an outcome $|\Psi_k\rangle$ may a priori depend on the measurement (ONB) it is regarded to be part of. This ONB is referred to as the "measurement context". For our HVM to be consistent with the interpretation of

[3] Compare this to the much stronger assumption made by Neumann in his faulty no-go theorem, namely that this relation must hold for arbitary observables. While this produces a contradiction already in two dimensional Hilbert spaces, the unfounded assumption has been mocked as silly by Mermin [31].



$$\forall \lambda \in \Lambda : \xi_{M,\Psi_1}(\lambda) = \xi_{M,\Psi_1}(\lambda)$$

Figure 2.4: Summary of the key points of KS non-contextuality, in terms of vectors in a projective Hilbert space, depicted here to be three-dimensional. The two ONB, $\{|\Psi_1\rangle, |\Psi_2\rangle, |\Psi_3\rangle\}$ and $\{|\Psi_1\rangle, |\Psi'_2\rangle, |\Psi'_3\rangle\}$ are related by a rotation in the plane. As $|\Psi_1\rangle$ appears in both measurements, a non-contextual HVM must assign the same conditional probability of observing the corresponding outcome for every ontic state $\lambda \in \Lambda$.

measurements revealing "true" properties of the system, an outcome being true or false should not depend on what measurement it is a part of. The assumption of KS non-contextuality imposes that the value assigned to any ray be independent of the measurement context, meaning that a ray appearing in multiple measurements must receive the same valuation by the indicator functions, allowing for a functional dependence on only the ray itself. Lastly, Condition 2 ensures that exactly one measurement outcome corresponds to a true property of the system, which is revealed upon performing the measurement. The key points of Definition 2.7 are summarized in Figure 2.4, which depicts two measurement bases, assuming a three dimensional Hilbert space.

As previously stated, both definitions of KS non-contextuality, Definitions 2.4 and 2.7, are equivalent in the following sense:

Lemma 2.8. *Let $\mathcal{H}$ be a Hilbert space. A non-contextual deterministic value assignment $\omega : \mathcal{P}(\mathcal{H}) \to \{0,1\}$ according to Definition 2.7 induces a non-contextual deterministic value assignment $\nu : \mathrm{Herm}(\mathcal{H}) \to \mathbb{R}$ according to Definition 2.4, and vice versa.*

Lemma 2.8 implies that proving the impossibility of a non-contextual deterministic value assignment $\omega : \mathcal{P}(\mathcal{H}) \to \{0,1\}$ for Hilbert spaces of dimension ≥ 3, which we will demonstrate in the next section, entails the impossibility of a non-contextual deterministic HVM according to either definition, as such a HVM would come with a mapping ν and ω for every ontic state λ.

Proof. (of Lemma 2.8) We prove both implications separately.

- $\nu \Rightarrow \omega$: Define $\omega(|\Psi\rangle) \coloneqq \nu(|\Psi\rangle\langle\Psi|)$. Projectors $|\Psi\rangle\langle\Psi|$ have eigenvalues 0,1, therefore ω maps to $\{0,1\}$. Since orthogonal projectors commute, by the property of "functional

consistency", in particular Corollary 2.6, we get:

$$\{|\Psi_i\rangle\}_i \subset \mathcal{H} \text{ an ONB } \Rightarrow \sum_i \omega(|\Psi_i\rangle) = \sum_i \nu(|\Psi_i\rangle\langle\Psi_i|) = \nu(\sum_i |\Psi_i\rangle\langle\Psi_i|) = \nu(\mathbb{1}_{\mathcal{H}}) = 1.$$

- $\omega \Rightarrow \nu$: (Adapted from [37]) Define $\nu(A = \sum_{i,j} a_i |\Psi_j^{(i)}\rangle\langle\Psi_j^{(i)}|) := \sum_{i,j} a_i\, \omega(|\Psi_j^{(i)}\rangle)$, where we have used that every Hermitian operator can be spectrally decomposed. We take the eigenvalues a_i to be distinct, $\{|\Psi_j^{(i)}\rangle\}_j$ being an ONB of the eigenspace corresponding to the eigenvalue a_i. The spectral decomposition of a Hermitian operator is not unique, as we may choose different ONB of the eigenspaces, in case of degeneracies. Nonetheless the above assignment is well-defined: Take

$$A = \sum_{i,j} a_i |\Psi_j^{(i)}\rangle\langle\Psi_j^{(i)}| = \sum_{i,j} a_i |\Phi_j^{(i)}\rangle\langle\Phi_j^{(i)}|$$

to be two spectral decompositions of A. Both $\{|\Psi_j^{(i)}\rangle\}_j$ and $\{|\Phi_j^{(i)}\rangle\}_j$ span the eigenspace corresponding to the eigenvalue a_i. We can extend both subsets to ONB of $\mathcal{H}$ by adding an orthonormal set $\{|\Theta_k\rangle\}_k$ to each. By the defining property of ω we get

$$\sum_j \omega(|\Psi_j^{(i)}\rangle) = \sum_j \omega(|\Phi_j^{(i)}\rangle)$$

for all i, which means that ν is well-defined. Suppose $\{A, B, ...\} \subset \text{Herm}(\mathcal{H})$ is a set of mutually commuting Hermitian operators. There exists a shared ONB of eigenvectors $\{|\Psi_i\rangle\}_i$. Let

$$A = \sum_i a_i\, |\Psi_i\rangle\langle\Psi_i| \text{ and } B = \sum_i b_i\, |\Psi_i\rangle\langle\Psi_i|$$

be spectral decompositions of A and B, where the eigenvalues a_i, b_i can now be repeating. The assignment ν maps A and B to

$$\nu(A) = \sum_i a_i\, \omega(|\Psi_i\rangle) \text{ and } \nu(B) = \sum_i b_i\, \omega(|\Psi_i\rangle).$$

Thus, commuting operators are mapped to simultaneous eigenvalues, and ν is a KS non-contextual value assignment like in Definition 2.4.

□

2.6 Formal statement and proof of the KS Theorem

The KS Theorem was first proved by Kochen and Specker in 1968 [26]. Previously, Bell independently proved that a deterministic HVM of QM that induces a non-contextual value assignment like in Definition 2.7, must be contextual [4]. However his proof was deemed less convincing, as it used a continuum of vector directions, as opposed to an explicit finite set [31].

We are now ready to reformulate the KS Theorem in precise mathematical terms (adapted from [37]):

Theorem 2.9 (Kochen-Specker [26])**.**
For all projective Hilbert spaces $\mathcal{P}(\mathcal{H})$ of dimension ≥ 3, there exists a finite set $S \subset \mathcal{P}(\mathcal{H})$ of rays for which no valid value assignment $\omega : S \to \{0, 1\}$ exists, satisfying the conditions for a KS non-contextual model, according to Definition 2.7.

The existence of such a set $\mathcal{S}$ is in contradiction with any attempt of a KS non-contextual deterministic HVM of QM, as KS non-contextuality implies a valid value assignment for every possible ontic state of the system. Thus the KS Theorem effectively states the impossibility of a non-contextual deterministic HVM.

As Pusey notes [37], all assumptions (outcome determinism, context independence and Hilbert spaces of dimension ≥ 3) of the KS Theorem are necessary for the proof of impossibility. If we allow for context dependence, such a value assignment can be given for all sets $\mathcal{S}$, as we can assign appropriate values for each measurement contained in $\mathcal{S}$. Furthermore, relaxing the assumption of outcome determinism to an assignment of probabilities $\tilde{\omega} : \mathcal{S} \to [0, 1]$ spoils the proof of impossibility, as a valid value assignment $\tilde{\omega}$ is given for instance by the Born rule: $\tilde{\omega}(|\Psi_i\rangle) \coloneqq \mathrm{tr}(\rho\, |\Psi_i\rangle \langle\Psi_i|)$ for some non-negative (Hermitian) operator ρ with $\mathrm{tr}(\rho) = 1$. Lastly, for Hilbert spaces of dimension 2, it is possible to give an explicit form of a non-contextual deterministic HVM [31]. The deeper reason for why our proof does not extend to the two-dimensional case is that it makes heavy use of overlapping ONB (measurement contexts) in the construction of $\mathcal{S}$, meaning different measurements that share at least one ray. These will be key in establishing a contradiction for dimensions ≥ 3. However, for two-dimensional Hilbert spaces, it is impossible to construct overlapping and distinct ONB, as a single ray fixes the ONB.

The proof of the KS Theorem is by induction with respect to to the dimension. The following Lemma (induction step), once proven, reduces the proof of the KS Theorem to the three-dimensional base case. We will prove the three-dimensional base case by giving an explicit set of directions belonging to overlapping orthogonal triads, for which a KS non-contextual value assignment can be shown to be impossible. This will be demonstrated by a chain of geometric arguments.

Lemma 2.10 ([37])**.**
Assume that all projective Hilbert spaces of dimension d have a finite subset like in the KS Theorem. Let $\mathcal{P}(\mathcal{H})$ be a projective Hilbert space of dimension d+1. Then $\exists\, \mathcal{S} \in \mathcal{P}(\mathcal{H})$ with the same properties.

Proof. We follow [37]. Let $\{|0\rangle, \ldots, |d\rangle\}$ be an ONB of $\mathcal{H}$ and let $\mathcal{H}_0$, $\mathcal{H}_1 \subset \mathcal{H}$ be the d-dimensional subspaces orthogonal to $|0\rangle$, $|1\rangle$, respectively. There exist subets $\mathcal{S}_0$, $\mathcal{S}_1$ of $\mathcal{P}(\mathcal{H}_0)$, $\mathcal{P}(\mathcal{H}_1)$, for which there are no valid value assignments within the respective projective Hilbert space. Let $\mathcal{S}_0'$, $\mathcal{S}_1'$ denote the embeddings of $\mathcal{S}_0$, $\mathcal{S}_1$ into $\mathcal{H}$. We will show that $\mathcal{S}' \coloneqq \mathcal{S}_0' \cup \mathcal{S}_1' \cup \{|0\rangle, \ldots, |d\rangle\}$ then defines a subset of $\mathcal{P}(\mathcal{H})$ for which there is no valid non-contextual value assignment. As $\{|0\rangle, \ldots, |d\rangle\}$ constitutes an ONB, a valid value assignment ω' has to map $|0\rangle$ or $|1\rangle$ (or both) to 0. Both $\omega'(|0\rangle) = 0$ and $\omega'(|1\rangle) = 0$ lead to a contradiction: For instance, say $\omega'(|0\rangle) = 0$. For any ONB $\mathcal{M} \subset \mathcal{H}_0$ and its embedding $\mathcal{M}' \subset \mathcal{H}$ we have

$$\sum_{|\Psi_i\rangle \in M'} \omega'(|\Psi_i\rangle) + \omega'(|0\rangle) = \sum_{|\Psi_i\rangle \in M'} \omega'(|\Psi_i\rangle) = 1.$$

Hence, $\omega : \mathcal{S}_0 \mapsto \{0,1\}$, $\omega(|\Psi\rangle) \coloneqq \omega'(|\Psi'\rangle)$ satisfies Condition 2 of 2.7, where $|\Psi'\rangle \in \mathcal{S}_0'$ is the embedding of $|\Psi\rangle \in \mathcal{S}_0$ into $\mathcal{H}$. As per assumption, w cannot be context independent. Therefore, by extension, w' must also be context dependent. □

What remains is a proof for the existence of a finite subset $\mathcal{S}$ that does not permit a valid KS non-contextual value assignment for the case of three dimensions. The following lemma allows us to consider only the three-dimensional complex projective space $\mathcal{P}(\mathbb{C}^3)$ w.l.o.g.:

Lemma 2.11. *Let $\mathcal{H}_n$ be a n-dimensional dimensional complex Hilbert space. A valid KS non-contextual value assignment $\omega : \mathcal{P}(\mathcal{H}_n) \to \{0,1\}$ according to Definition 2.7 induces a valid value assignment $\tilde{\omega} : \mathcal{P}(\mathbb{C}^n) \to \{0,1\}$.*

Proof. Let $\{|\Psi_i\rangle\}_{i\in\{1,\ldots,n\}}$ be an ONB of $\mathcal{H}_n$. We can construct the isomorphism $\mathcal{J} : \mathcal{H}_n \to \mathbb{C}^n$, $|\Psi_k\rangle \mapsto e_k$, where $\{e_i\}_{i\in\{1,\ldots,n\}}$ is the standard basis of $\mathbb{C}^n$. This isomorphism preserves the inner product structure, as it maps an ONB of $\mathcal{H}_n$ to an ONB of $\mathbb{C}^n$. Thus, $\mathcal{J}$ preserves all orthogonality relations. A valid non-contextual value assignment according to Definition 2.7 then induces a valid non-contextual value assignment $\tilde{\omega} : \mathcal{P}(\mathbb{C}^n) \to \{0,1\}$ via $\tilde{\omega} \coloneqq \omega \circ \mathcal{J}^{-1}$. □

Having reduced the proof of the three-dimensional base case to proving the impossibility of a valid non-contextual value assignment $\tilde{\omega} : \mathcal{P}(\mathbb{C}^3) \to \{0,1\}$, we will now turn our attention to this task. Lemma 2.12 sticks to Peres' notation [34] for saving space:

Lemma 2.12. *Let* $\bar{1}$ *denote -1 and 2 denote* $\sqrt{2}$*. For the following set of 33 vectors* $\mathcal{S} \subset \mathbb{R}^3 \subset \mathbb{C}^3$ *("Peres configuration"* [4] *[34]) it is impossible to assign values* $\{0,1\}$ *to each vector such that for all orthogonal triads contained in* $\mathcal{S}$ *exactly one member receives the valuation 1:*

$$\begin{aligned}\mathcal{S} = \{&100, 010, 001, 110, 101, 011, 1\bar{1}0, \bar{1}01, 0\bar{1}1, 1\bar{1}2, \bar{1}12,\\ &\bar{2}01, 021, 102, \bar{2}11, 211, \bar{1}02, 201, \bar{1}\bar{1}2, 112, 0\bar{2}1, 012,\\ &1\bar{2}1, 121, 0\bar{1}2, 12\bar{1}, \bar{1}21, 21\bar{1}, 2\bar{1}1, \bar{2}10, 1\bar{2}0, 120, 210\}\end{aligned}$$

Here, the tuple $\bar{2}01$ *for instance corresponds to the vector* $\begin{pmatrix} -\sqrt{2} \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3$.

We consider unnormalized vectors only for convenience as this does not affect orthogonality relations.

The cubic symmetry of the set $\mathcal{S}$ greatly simplifies the proof. More precisely, the set $\mathcal{S}$ is invariant under interchange of the x, y, z axes and under a reversal of the direction of each axis $x \mapsto \bar{x}$. It is the case that

$$\omega : \mathcal{S} \to \{0,1\} \text{ is a valid value assignment } \iff \omega \circ T \text{ is a valid value assignment,}$$

T being one of the aforementioned symmetry transformations. This is straightforward to see, as the symmetry transformations of $\mathcal{S}$ are orthogonal transformations that leave orthogonality relations unchanged. The following concrete example shows how we can exploit the symmetry of $\mathcal{S}$ to make arbitrary assignments w.l.o.g.: Consider the orthogonal triad $\{001, 100, 010\}$. Assuming there exists a valid value assignment $\omega : \mathcal{S} \to \{0,1\}$, then exactly one of the three vectors in the triad is assigned a 1. Due to the symmetry of the set $\mathcal{S}$, there must be a valid value assignment ω' that assigns 001 the value 1. Analogous considerations apply to symmetry transformations $x \mapsto \bar{x}$.

On a historical note, the original KS paper [26] gives an explicit set of 117 directions in $\mathbb{R}^3$ that lead to a contradiction. The current record holders are Kochen and Conway, with a set of 31 directions, however their set is not as highly symmetric as the Peres configuration, which complicates the proof [35].

Proof. (of Lemma 2.12)
The proof of Lemma 2.12 is given by the table in Figure 2.5 and shows that any attempt to construct a valid KS non-contextual value assignment $\omega : \mathcal{S} \to \{0,1\}$ inevitably leads to a contradiction. Finding such an assignment is equivalent to a colouring problem. Identifying

[4] For a nice visualization of the "Peres configuration", see Conway and Kochen's paper "The Strong Free Will Theorem", which also makes use of it [17].

$1 \equiv$ "green", $0 \equiv$ "red", a valid value assignment corresponds to a colouring of the rays in $\mathcal{S}$, in accordance with with the rule that each orthogonal triad in $\mathcal{S}$ must have exactly one element coloured green and two elements coloured red. Such a colouring of rays is shown to be impossible, making use of the fact that $\mathcal{S}$ contains rays that are part of multiple distinct orthogonal triads. Each line of the table introduces a new orthogonal triad that is coloured, either w.l.o.g. , making use of the symmetry of $\mathcal{S}$, or according to the values of previous triads. In each line, only the first ray of the triad is coloured green. The "Other rays" in each line are orthogonal to the single green ray in that line and are consequently coloured red. The stepwise colouring of rays leaves the set $\{100, 021, 0\bar{1}2\}$ coloured red. As these rays form an orthogonal triad, this proves that a valid colouring and thus a valid value assignment are impossible. □

2.7 Algebraic proof in four dimensions

We now reap the fruits of our previous efforts that proved the equivalence of KS non-contextual value assignments to rays and such assignments to Hermitian operators. This allows us to give a far simpler and purely algebraic proof of the KS Theorem in four dimensions, which uses KS non-contextuality in terms of Hermitian operators. For the same reasons as in three dimensions, we can w.l.o.g. assume $\mathcal{H} = \mathbb{C}^4 \cong \mathbb{C}^2 \otimes \mathbb{C}^2$ [5]. Consider the set of operators depicted in Figure 2.6. They form the so-called "Mermin Square" [31] and can be written in terms of Pauli operators. The operators are arranged such that the operators in each of the three rows, columns are mutually commuting. Additionally, the product of the three observables in the rightmost column is $-\mathbb{1}_{\mathbb{C}^2\otimes\mathbb{C}^2}$. The product of the three observables in all other columns and all rows is $+\mathbb{1}_{\mathbb{C}^2\otimes\mathbb{C}^2}$. By Corollary 2.6, a KS non-contextual value assignment assigns values to the operators such that the product of values within one row, column is ± 1, in accordance with the corresponding operator identities. It follows that, by the row identities, the product of all nine values is $+1$, whereas the column identities result in a value of -1, a contradiction.

2.8 Local causality and Bell's Theorem

Locality as an instance of contextuality

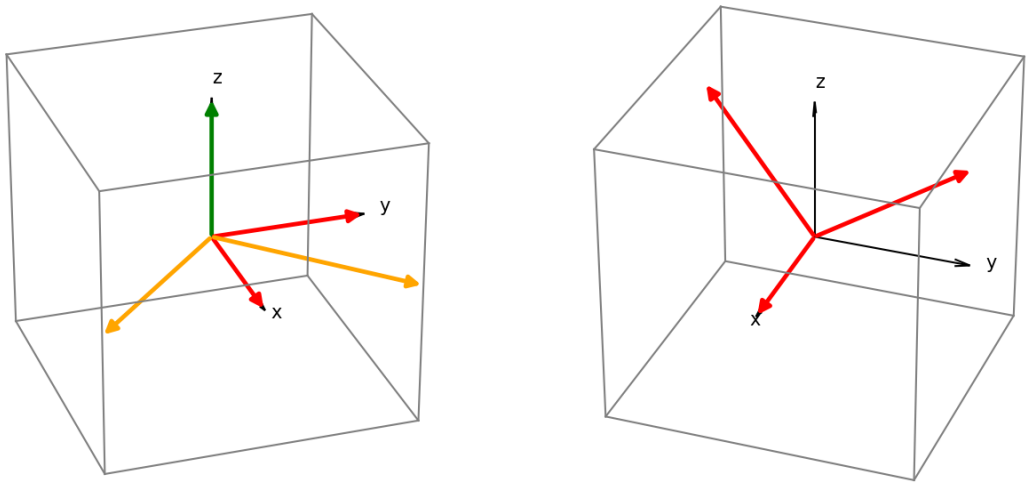
The following is meant to discuss the relationship between the KS Theorem and Bell's Theorem, in particular the connection between KS non-contextuality and Bell's notion of local causality.

Consider the well-known CHSH (thought-)experiment [15, 16, 41] that proves QM to be incompatible with a local HVM description. In particular, quantum theory predicts violations of the CHSH Bell inequality, which must be satisfied for all local HVM.

[5] This is not to say that we are now considering two remote systems, but only provides a convenient way to define the set of operators used in the proof.

Orthogonal triad			Other rays	The first ray is green because of
001	100	010	110 $1\bar{1}0$	choice of z axis
101	$\bar{1}01$	*010*		choice of x vs -x
011	$0\bar{1}1$	*100*		choice of y vs -y
$1\bar{1}2$	$\bar{1}12$	*110*	$\bar{2}01$ 021	choice of x vs y
102	*$\bar{2}$01*	*010*	$\bar{2}11$	orthogonality to second and third rays
211	*0 $\bar{1}$ 1*	*$\bar{2}$11*	$\bar{1}02$	orthogonality to second and third rays
201	*010*	*$\bar{1}$02*	$\bar{1}\bar{1}2$	orthogonality to second and third rays
112	*1 $\bar{1}$ 0*	*$\bar{1}$ $\bar{1}$ 2*	$0\bar{2}1$	orthogonality to second and third rays
012	*100*	*0 $\bar{2}$ 1*	$1\bar{2}1$	orthogonality to second and third rays
121	*$\bar{1}$01*	*1 $\bar{2}$ 1*	$0\bar{1}2$	orthogonality to second and third rays

(a) Table containing the proof of the three-dimensional base case of the KS Theorem, in terms of a colouring problem where each orthogonal triad must contain exactly one green and two red rays. Each line of the table introduces a new orthogonal triad that is coloured, either w.l.o.g. or according to the values of previous triads. In each line, only the first ray of the triad is coloured green. The "Other rays" in each line are orthogonal to the single green ray in that line and are consequently coloured red. The stepwise colouring of rays leaves the orthogonal triad {100,021,$0\bar{1}2$} coloured red, proving the impossiblity of a valid colouring. Table copied from [34].



(b) The left plot visualizes the first line of the proof-containing table, "Other rays" being coloured orange. The right plot visualizes the resulting contradiction, refer to the text.

Figure 2.5: Proof of the KS Theorem

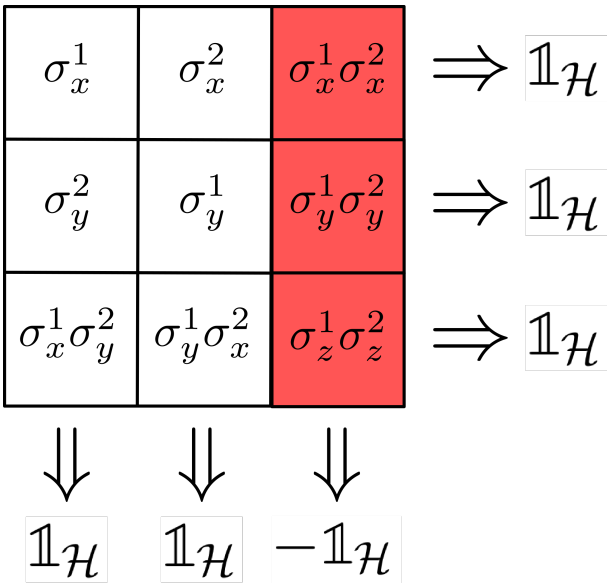


Figure 2.6: "Mermin Square": Operators on $\mathbb{C}^2 \otimes \mathbb{C}^2$ that allow for a proof of the KS Theorem in four dimensions. Tensor products between Pauli operators are omitted for readability. Within each row and each column, the operators commute. A non-contextual value assignment to these operators is contradictory, in the sense that the product of all nine values assigned to the operators would not be well-defined. For each row, the product of the three operators in that row is $\mathbb{1}_{\mathcal{H}}$. The same holds for the columns, with the notable exception that the product of the three operators in the rightmost column is $-\mathbb{1}_{\mathcal{H}}$. These identities on an operator level are key in establishing a contradiction. Figure adapted from [31].

A source S emits pairs of entangled particles in the state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)$ (EPR pairs) that split up and travel to two remote labs. In each of the labs, an experimenter can choose between two incompatible two-outcome measurements A_1, A_2 and B_1, B_2, respectively. One can for instance think of these as measuring different spin components of the particle. We will label the outcomes with ± 1. The setting of the CHSH experiment is shown in Figure 2.7a. Let X, Y be random variables that describe the measurement outcomes of A_i, B_i, respectively. For uniformly chosen detector settings, we wish to measure the probability

$$\begin{aligned} p(\text{success}) = &p(X = Y, A_1, B_1) + p(X = Y, A_2, B_1) \\ &p(X = Y, A_2, B_2) + p(X \neq Y, A_1, B_2), \end{aligned}$$

as is illustrated in Figure 2.7b. All correlations compatible with a general HVM description must be of the form

$$p(X = x, Y = y, A_i, B_j) = \int_\Lambda d\lambda\, \mu(\lambda)\, p(X = x, Y = y | A_i, B_j, \lambda)\, p(A_i, B_j | \lambda),$$

where μ is some probability density function on the ontic state space λ. Driven by special relativity, we further assume local causality to be a property of all valid classical descriptions: Rewriting $p(X = x, Y = y | A_i, B_j, \lambda)$ as $p(X = x | A_i, B_j, Y = y, \lambda) p(Y = y | A_i, B_j, \lambda)$, local causality is the assumption that, given a complete specification of the ontic state λ, the outcome

of a measurement is independent of what other measurements are performed in a space-like separated region of space-time, as well as their outcomes:

$$p(X = x|A_i, B_j, Y, \lambda) = p(X = x|A_i, \lambda)$$
$$p(Y = y|A_i, B_j, \lambda) = p(Y = y|B_j, \lambda)$$

For any ontological model of reality, local causality provides the best and most natural explanation for the fact that such superluminal influences have never been observed in experiments. Analogously, KS non-contextuality provides the most natural explanation for why, within the framework of QM, compatible observables do not disturb each other's outcome statistics. We will revisit this point in Section 6, where we introduce a revised notion of contextuality due to Spekkens. All correlations compatible with a local HVM description, i.e. an ontological model obeying local causality, must be of the form

$$p(X = x, Y = y|A_i, B_j) = \int_\Lambda d\lambda\, \mu(\lambda) p_A(X = x|A_i, \lambda) p_B(Y = y|B_j, \lambda).$$

The assumption of local causality for the space-like separated regions A and B manifests itself in the local response functions p_A and p_B. W.l.o.g. , we can assume p_A and p_B to be deterministic response functions, as we may shift all classical randomness into the ontic states and the distribution μ over these. This is sometimes referred to as Fine's Theorem [20]. Importantly, Fine's Theorem does not hold for general contextuality scenarios without space-like separated subsystems, as for two compatible local operators A, A', the set of simultaneous eigenvalues is in general not given by the Cartesian product $\sigma(A) \times \sigma(A')$. We will discuss complications that arise without space-like separated subsystems in Section 3.5.

Assuming that the experimenters can choose the detector settings freely (no super-determinism), the choice of measurement should not be correlated with the ontic state λ emitted by the source. Therefore, $p(A_i, B_j|\lambda) = p(A_i, B_j)$. All in all, for the CHSH experiment[6] to have a valid classical **HVM description** obeying **local causality** and the assumption of **freely random detector settings**, chosen according to a **uniform distribution**, all correlations must be of the form:

$$p(X = x, Y = y, A_i, B_j) = \frac{1}{4} \int_\Lambda d\lambda\, \mu(\lambda) p_A(X = x|A_i, \lambda) p_B(Y = y|B_j, \lambda), \tag{2.1}$$

where p_A and p_B are deterministic probability assignments.

Given the general expression for correlations, as prescribed by a local HVM description 2.1, it immediately follows that $p(\text{success})$ is upper-bounded by $\frac{3}{4}$ for any such HVM description. This can be verified in the same manner as in Section 2.1, where we showed deterministic and non-contextual HVM descriptions to be incompatible with Specker's three boxes: Examining Figure 2.7b, we realize that there is no deterministic assignment of measurement outcomes ± 1

[6] This can of course be extended to more general correlation experiments in a Bell-like setting.

to the four measurement operators, such that all "winning" correlations are satisfied. In fact, for any such assignment, at most three of the four "winning" correlations are satisfied. The local response functions in Equation 2.1 induce such a deterministic assignment. Therefore, for every ontic state λ, $p(\text{success})$ is upper-bounded by $\frac{3}{4}$. The ensemble average over all ontic states $\lambda \in \Lambda$ is thus upper-bounded by the same quantity. Consequently, $p(\text{success}) \leq \frac{3}{4}$ is a Bell inequality: the so-called **CHSH Bell-inequality** [15, 16, 41]:

$$p(\text{success}) = p(X = Y, A_1, B_1) + p(X = Y, A_2, B_1) + p(X = Y, A_2, B_2) + p(X \neq Y, A_1, B_2) \leq \frac{3}{4} \tag{2.2}$$

In QM, outcome probabilities are computed according to the Born rule. Assuming projective measurements on a pure quantum state, permissible quantum correlations for Bell-type experiments are of the form

$$p(X = x, Y = y | A_i, B_j) = \langle \Psi | \Pi^x_{A_i} \otimes \Pi^y_{B_j} | \Psi \rangle \,,$$

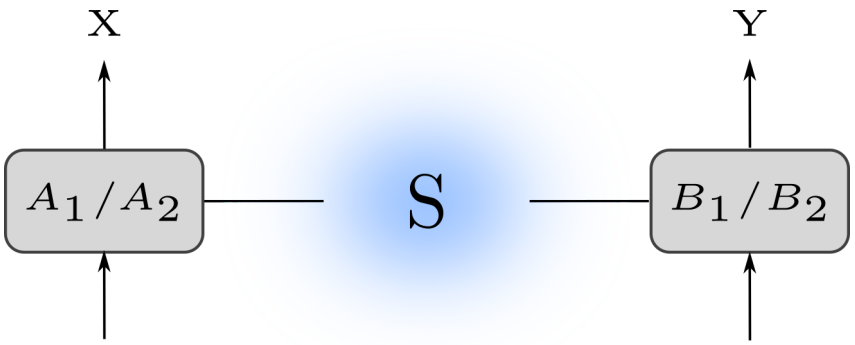
where $|\Psi\rangle \in \mathcal{H}$ is a normalized, bipartite Hilbert state vector and $\Pi^x_{A_i}$ is the projection operator corresponding to the outcome x of the measurement A_i on the subsystem A.

What is remarkable about the CHSH scenario is the fact that one can give explicit projective measurements A_1, A_2, B_1, B_2, as well as the maximally entangled pure state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)$, for which quantum theory predicts $p(\text{success}) = \cos^2(\frac{\pi}{8}) > \frac{3}{4}$: a violation of the CHSH Bell inequality. A set of Hermitian operators with this property is

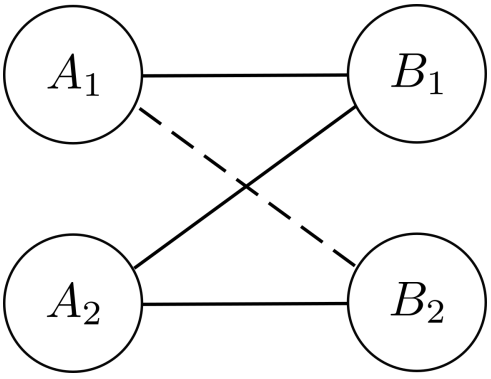
$$\begin{aligned} A_1 &= |0\rangle\langle 0| - |\pi\rangle\langle \pi| \,, \\ A_2 &= |\frac{\pi}{2}\rangle\langle\frac{\pi}{2}| - |\frac{3\pi}{2}\rangle\langle\frac{3\pi}{2}|, \\ B_1 &= |\frac{\pi}{4}\rangle\langle\frac{\pi}{4}| - |\frac{5\pi}{4}\rangle\langle\frac{5\pi}{4}|, \\ B_2 &= |\frac{3\pi}{4}\rangle\langle\frac{3\pi}{4}| - |\frac{7\pi}{4}\rangle\langle\frac{7\pi}{4}|, \end{aligned} \tag{2.3}$$

with $|\theta\rangle := \cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + \sin\left(\frac{\theta}{2}\right)|\downarrow\rangle$ [16]. Recall that we label the two distinct measurement outcomes with the eigenvalues ± 1. Each local measurement projects the state of the particle on which it acts onto one of two antipodal surface points in the xz-plane of the single qubit Bloch sphere. As the two particles are entangled, in particular maximally correlated, measuring one particle will "steer" the state of the other particle[7]. Thus, the probability of the measurements (A_1, B_2) yielding anti-correlated outcomes is given by $\frac{1}{2}\left|\langle 0|\frac{7\pi}{4}\rangle\right|^2 + \frac{1}{2}\left|\langle \pi|\frac{3\pi}{4}\rangle\right|^2 = \cos^2(\frac{\pi}{8})$. In fact, for each of the four possible measurement settings, the probability of the measurement outcomes being related as in Figure 2.7b is $\cos^2(\frac{\pi}{8})$. To see this, recall that we may associate with each local measurement outcome a point on the Bloch sphere, and by extension with every

[7] This does not violate the no-signalling principle, as can be seen by computing the reduced density operator of system B: it describes a maximally mixed state.



(a) CHSH setup leading to violations of the CHSH inequality. A source S emits entangled EPR pairs. Constituent particles are measured in space-like separated regions, with experimenters being able to choose between two incompatible measurements A_1, A_2 or B_1, B_2. The random variables X, Y hold the measurement outcome for operators A_1, A_2 or B_1, B_2, respectively.



(b) CHSH correlations for space-like separated operators A_1, A_2 and B_1, B_2. Lines indicate joint measurability. A solid line further indicates correlated outcomes, whereas a dashed line indicates anti-correlated outcomes. Let X, Y be random variables that describe the measurement outcomes of A_i, B_i, respectively. We are interested in the probability that $X = Y$ for all choices of measurements but (A_1, B_2) and $X \neq Y$ for the measurements (A_1, B_2). Adapted from [41].

Figure 2.7: CHSH Bell test

local measurement operator an axis connecting two antipodal points on the Bloch sphere, as shown in Figure 2.8. These axes are rotated by an angle of $\frac{\pi}{4}$ (in real space) relative to one another and lie within the xz-plane. For all A_i and B_j, the measurement operators are chosen in such a way that the overlap of the Bloch states corresponding to measurement outcomes that obey the correlations in Figure 2.7b is $\cos\left(\frac{\pi}{8}\right)$.

In view of Section 4, important properties of the CHSH Bell test are:

- The choice of quantum state and measurements above leads to a maximal violation of the CHSH inequality, as allowed by QM. We will show this in Section 3. The quantum supremum $B_{\mathcal{Q}}$ of a Bell inequality's left hand side is called its Tsirelson bound. The Tsirelson bound of the CHSH Bell inequality is $\frac{1}{2}(1 + \frac{1}{\sqrt{2}})$.

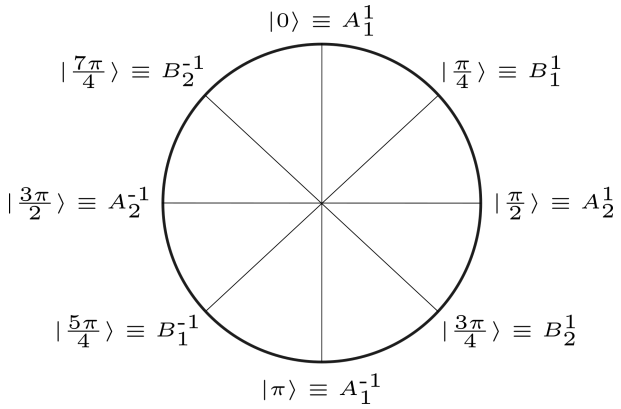


Figure 2.8: The two outcomes of each of the four local projective CHSH measurements A_1, A_2, B_1, and B_2 correspond to two antipodal surface points in the xz-plane of the Bloch sphere. For example, the projector associated with the outcome 1 of the measurement A_1 projects onto the state $|0\rangle$. The overlap between two states that are related via a rotation by $\frac{\pi}{4}$ in real space have an overlap of $\cos(\frac{\pi}{8})$. By inspection, one sees that the winning CHSH correlations in Figure 2.8 are obtained with the probability $\cos^2(\frac{\pi}{8})$.

- Assuming projective measurements, all quantum experiments $\langle\Psi|\Pi^x_{A_i} \otimes \Pi^y_{B_j}|\Psi\rangle$ producing a maximal violation of the CHSH Bell inequalitiy are essentially equivalent and related by some trivial operations. This property gives rise to the notion of self-testing and will be discussed in more detail in Section 4.

A generic Bell inequality in the bipartite scenario is of the form

$$\sum_i w_i p(\epsilon_i) \leq B_{\mathcal{L}} \ , \tag{2.4}$$

where $w_i > 0$ and ϵ_i denotes the i-th measurement event. For now, assuming bipartite Bell tests, a measurement event can be thought of as just an initial preparation P of the system (for instance in an EPR state), followed by both parties performing some local measurement a,b, and obtaining outcomes x,y: $a, b|x, y$. We have used the same notation as in [11] and will revisit the proposed graph-theoretic framework in more detail in Section 3, in order to identify the maximal quantum violation of a general KS non-contextuality inequality. Note that all linear combinations of probabilities pertaining to our correlation experiment can be expressed as a positive linear combination with $w_i > 0$, like in 2.4. For any negative w_i we may rewrite

$$w_i p(e_i) = w_i p(x_i, y_i|a_i, b_i) = w_i - w_i \sum_{(x_i', y_i') \neq (x_i, y_i)} p(x_i', y_i'|a, b) \ ,$$

obtaining a positive linear combination of probabilites.

The statistics of a Bell-type correlation experiment are fully characterized by the set $\{p(x, y|a, b)\}_{x,y,a,b}$. For the CHSH experiment, this set contains $2^4 = 16$ values between 0 and 1, which obey some

normalization conditions and can be represented by a vector in $\mathbb{R}^{16-4} = \mathbb{R}^{12}$. Identifying permissible correlations as a subset of $[0,1]^{12}$, in the case of CHSH, one can characterize the geometry of correlations that are compatible with say local causality, quantum theory, or the no-signalling principle. Each set of permissible correlations is clearly convex: While one can check that convex combination preserves the relevant constraints, there is an intuitive argument as to why this must be the case: Assume that we have two experimental procedure, each producing "allowed" statistics. As we only impose constraints on what correlations we consider to be "allowed", the actual existence of procedures generating these correlations cannot be discarded. We now consider the following new experimental procedure: we generate a random bit and then perform the "old" experimental procedure corresponding to the value of the generated bit. By changing the bias of our "coin", we can produce all statistics that are convex combinations of the initial, permissible statistics.

As discussed, all correlations $p(x,y|a,b)$ compatible with an ontological model obeying local causality are convex combinations of local, deterministic probability assignments $p_A(x|a,\lambda)p_B(y|b,\lambda)$. There are $2^4 = 16$ distinct such assignments. Thus, the set of local correlations

$$\mathcal{L} = \{\{p(x,y|a,b)\}_{x,y,a,b} \equiv \vec{v} \in \mathbb{R}^{12} : \vec{v} \text{ compatible with local causality}\}$$

is a polytope in $\mathbb{R}^{12}$ with 16 extremal vertices. A Bell inequality 2.4 defines a separative hyperplane that splits $[0,1]^{12}$ in two, such that $\mathcal{L}$ lies entirely on one side of this hyperplane (the closed half-space). The Bell hyperplane could for instance be a facet of the local polytope, as shown in Figure 2.9. A facet of $\mathcal{L}$ is one of its delimiting hyperplanes. Bell's Theorem tells us that the set of quantum correlations $\mathcal{Q}$ is strictly larger than $\mathcal{L}$. The set $\mathcal{Q}$ further cannot be characterized in terms of a finite number of delimiting hyperplanes (linear inequalities) [7]. Correlations that maximally violate the CHSH Bell inequality have a maximal distance to the Bell hyperplane.

The no-signalling principle is built into the tensor product representation of traditional QM. As such, the set $\mathcal{Q}$ is contained in the set $\mathcal{NS}$, which contains all correlations that obey no-signalling. Like $\mathcal{L}$, $\mathcal{NS}$ is a polytope: $\mathcal{NS}$ is the intersection of $[0,1]^{16}$, the hyperplanes corresponding to normalization constraints $\sum_{x,y} p(x,y|a,b) = 1$, and those defined by the no-signalling conditions

$$\sum_y p(x,y|a,b) = \sum_y p(x,y|a,b') \text{ and}$$
$$\sum_x p(x,y|a,b) = \sum_x p(x,y|a',b) \ .$$

The set $\mathcal{NS}$ is strictly bigger than $\mathcal{Q}$, as there exist correlations, such as the "PR-box", that satisfy the no-signalling principle and win the "CHSH game" with certainty [7].

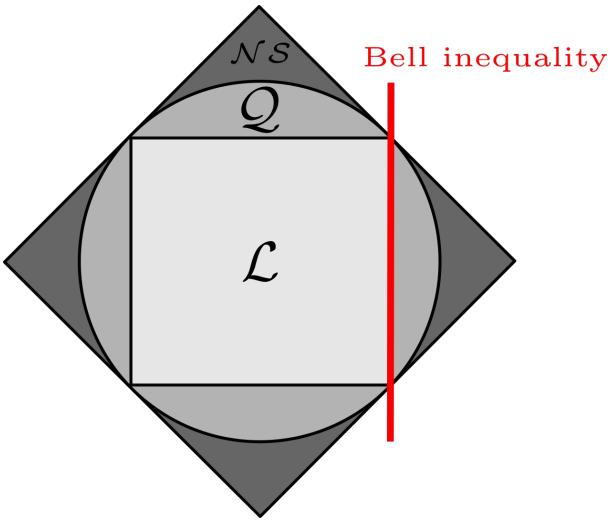


Figure 2.9: Schematic drawing depicting the geometry of the convex sets $\mathcal{L}$, $\mathcal{Q}$, and $\mathcal{NS}$, containing all local, quantum, and no-signalling correlations, respectively. The sets $\mathcal{L}$, $\mathcal{NS}$ are polytopes and can be characterized by a finite number of facet hyperplanes, whereas the set $\mathcal{Q}$ is more complicated. A Bell inequality, like the CHSH inequality, defines a separative hyperplane such that all local correlations $\mathcal{L}$ lie to one "side" of this hyperplane.

Analogous to Bell inequalities, KS non-contextuality inequalities define hyperplanes that splits the convex set of all permissible correlations in two, such that the closed half-space contains all correlations compatible with a KS non-contextual HVM description.

We have captured the essence of Bell's Theorem, cast in a quantum framework, proving that the predictions of QM are incompatible with a local HVM description. Let us now study the relationship between KS non-contextuality and Bell's notion of local causality, or equivalently local determinism (recall Fine's Theorem [20]).

The assumption of local determinism is an assumption of context independence for remote measurement contexts, like for Bell-type setting [41]. Both boil down to measurement outcomes being independent of which remote measurement is simultaneously performed (and its outcome). It is important to keep in mind however that KS non-contextuality is a more general assumption, as compatible measurements are not required to be remote. As Bell's Theorem implies the impossibility of a local HVM of QM, it in particular also proves the impossibility of a KS non-contextual deterministic HVM of QM. This, together with the fact that locality is accepted as a fundamental pillar of modern physics, is why Bell's Theorem is considered to be stronger than the KS Theorem.

Informally speaking, Bell's Theorem, as presented above, is proof of a KS Theorem in four dimensions, with the nice property that KS non-contextuality is only assumed when it can be justified by locality. A caveat that stems from only assuming KS non-contextuality for remote measurement contexts is that the CHSH proof assumes an entangled state preparation $|\Psi\rangle$.

In contrast, the KS Theorem, as formally stated in Theorem 2.9, is entirely measurement-dependent. Theorem 2.9 states that any attempt to assign outcomes to measurements in a KS non-contextual manner is inherently contradictory, no matter the state preparation. In Section 2.7 we examined a KS proof in four dimensions that made use of the full extent of KS non-contextuality. Unsurprisingly, "watering down" the assumption of KS non-contextuality, like in the CHSH proof, requires us to make an additional assumption in order to arrive at a contradiction. For the CHSH proof this assumption is that of a fixed, entangled quantum state.

Mermin gives a GHZ-like proof of an eight-dimensional KS Theorem [31], which can be recast into a proof of Bell's Theorem, further highlighting the deep connection between both theorems. Again, the information "lost" in assuming KS non-contextuality only when justified by locality is compensated for by a fixed quantum state. Thus, the fact that the KS Theorem is only measurement-dependent and in particular makes no reference to the system preparation is a consequence of the stronger assumption of KS non-contextuality imposed by the KS Theorem.

On a final note, it is worth pointing out that the statement of the KS Theorem is of a fundamentally different nature than that of Bell's Theorem. In particular, it does not propose inequality constraints on correlations that, when found to be violated, would imply the impossibility of a particular classical description. The KS Theorem goes further than that and deduces the impossibility of such an attempt altogether, only by geometric considerations that lead to a contradiction. Nevertheless, this strength is also a key weakness of the theorem. A Bell inequality is operationally testable and does not make reference to the formalism of QM. The general form of correlations compatible with a local HVM description 2.1 makes sense, independent of QM. Therefore, the CHSH Bell inequality also holds for theories beyond QM. An experimental observation of a Bell inequality violation implies that there can be no operational theory of reality (compatible with experimental observations) that admits of a locally causal hidden variable theory, quantum theory only being one possible candidate. Any such theory would be in conflict with the experimental data which violates Bell inequalities. By comparison, the KS Theorem is formulated within the framework of QM, for instance making explicit reference to the system's Hilbert space. Therefore, the scope of the KS Theorem is limited to quantum theory.

In Section 6 we revise the traditional notion of KS non-contextuality. Spekkens proposes a largely operational notion of non-contextuality that addresses these shortcomings [40]. However, for the purposes of self-testing, which assumes the validity of QM, the main asset of Spekkens contextuality will be that it allows for a consistent treatment of unsharp POVM quantum measurements.

2.9 State-dependent KS contextuality

The KCBS inequality and odd n-cycle scenarios

In Section 2.6 we found that for every projective Hilbert space of dimension ≥ 3 there exists a finite set of rays for which there exists no KS non-contextual value assignment according to Definition 2.7. Identifying these rays with projective measurements of rank one gives us a set of projective measurements for which there exists no KS non-contextual value assignment according to Definition 2.4. We call such sets of projective measurements (or rays in a projective Hilbert space) KS uncolourable, owing to the fact that one can reduce the problem of finding a valid KS non-contextual value assignment to a colouring problem, as was done in Section 2.6.

Definition 2.13. A set of projective measurements (rays in a projective Hilbert space) is called *KS colourable* if there exists a KS non-contextual assignment of measurement outcomes (values 0,1) to the measurements (rays) in the set. If there exists no such value assignment, the set is called *KS uncolourable*.

KS uncolourable sets embody the "strongest" proofs that QM is incompatible with a KS non-contextual HVM description, in the sense that the structure in the set of measurements alone is enough to arrive at a logical contradiction. Nevertheless, as alluded to in Section 2.8, also KS colourable sets of projective measurements may exhibit KS contextuality. This manifests itself in QM predicting violations of KS non-contextuality inequalities for some KS colourable sets. A KS non-contextuality inequality is, analogously to a Bell inequality, a linear inequality constraint on the convex set of correlations that is satisfied for all correlations compatible with a KS non-contextual HVM description. We will discuss the simplest KS non-contextuality inequality, the KCBS inequality, as well as an infinite class of generalizations thereof. These KS non-contextuality inequalities allow for self-testing, as will be the subject of Section 5. Before that, note that we have already come across one example of a KS non-contextuality inequality, the CHSH inequality, albeit in the context of Bell nonlocality. The set of CHSH measurements given in Section 2.8 that produce a maximal violation of the CHSH inequality when acting on an EPR pair is a KS colourable set of projective measurements. QM predicts state-dependent Bell non-locality, as witnessed by violations of the CHSH inequality. For reasons highlighted in Section 2.8, the CHSH inequality is also a KS non-contextuality inequality and the non-local correlations of the CHSH experiment exhibit state-dependent KS contextuality. Programming the source to emit separable bipartite states, as opposed to EPR pairs, will yield KS non-contextual correlations.

Note that there even exist KS colourable sets of measurements that produce state-independent KS contextuality, in the sense that QM predicts the violation of a KS non-contextuality inequality for all possible quantum state preparations. An example of such a set is one proposed by Yu and Oh [48], simply called the Yu-Oh set. It is a set of 13 rank one projectors that act on a

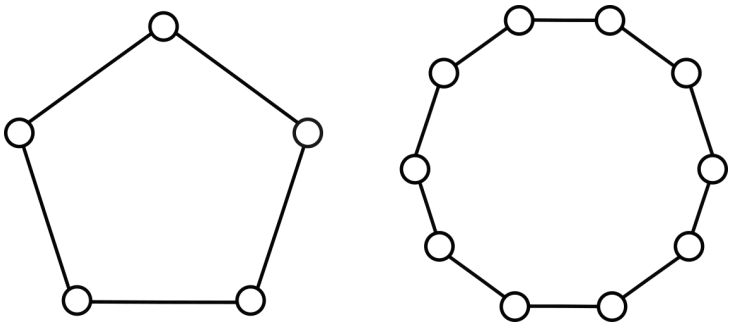


Figure 2.10: Graph showing compatibility relations for the KCBS (left) and a higher order odd n-cycle scenario (right). Each vertex corresponds to one projective measurement, adjacent measurements are compatible.

three-dimensional Hilbert space. In fact, if one identifies the rank one projectors with rays in a projective Hilbert space, the 13 Yu-Oh rays are a subset of the 33 rays of the Peres configuration that we used in proving the KS Theorem in Section 2.6. The Yu-Oh set is the minimal set of rays revealing state-independent KS contextuality, as proved in [13].

The perhaps simplest example of a quantum experiment leading to (state-dependent) KS contextuality for a single quantum system is the KCBS scenario [25]. We will present the corresponding KCBS KS non-contextuality inequality, which will become relevant when we later discuss self-testing of quantum systems in Sections 4 and 5.

For the KCBS KS contextuality scenario we assume five dichotomic projective measurements with outcomes 0 and 1. Additionally, the five measurements should obey cyclic compatibility, as shown in Figure 2.10. We freely choose a pair of compatible measurements (measurement context) at random, each with equal probability, and determine the value of the following linear sum of probabilities:

$$\sum_{i=1}^{5} p(0,1|i,i+1) \leq 2,$$

where compatibility with a KS non-contextual HVM imposes an upper bound of 2 for this expression: For any HVM description of the KCBS experiment, $\sum_{i=1}^{5} p(0,1|i,i+1)$ can be rewritten as $\int_\Lambda d\lambda\, \mu(\lambda|i,i+1) \sum_{i=1}^{5} p(0,1|i,i+1,\lambda)$, where μ is some probability density function over the ontic state space Λ. Furthermore, $\mu(\lambda|i,i+1) = \mu(\lambda)$, as we assume the choice of measurement context to be uncorrelated with the system. Analogously to Specker's three boxes or the CHSH experiment, one can easily verify that this quantity has a non-trivial upper bound for any KS non-contextual HVM description.

We now present a quantum experiment involving a KS colourable set of five projective measurements on a qutrit for which QM predicts a (maximal) violation of the KCBS KS non-contextuality inequality. The five projective measurements of this reference experiment correspond to five rank one projectors that project onto five pure qutrit states. The components of these five qutrit states with respect to the standard qutrit basis $\{|0\rangle, |1\rangle, |2\rangle\}$ are real, allowing

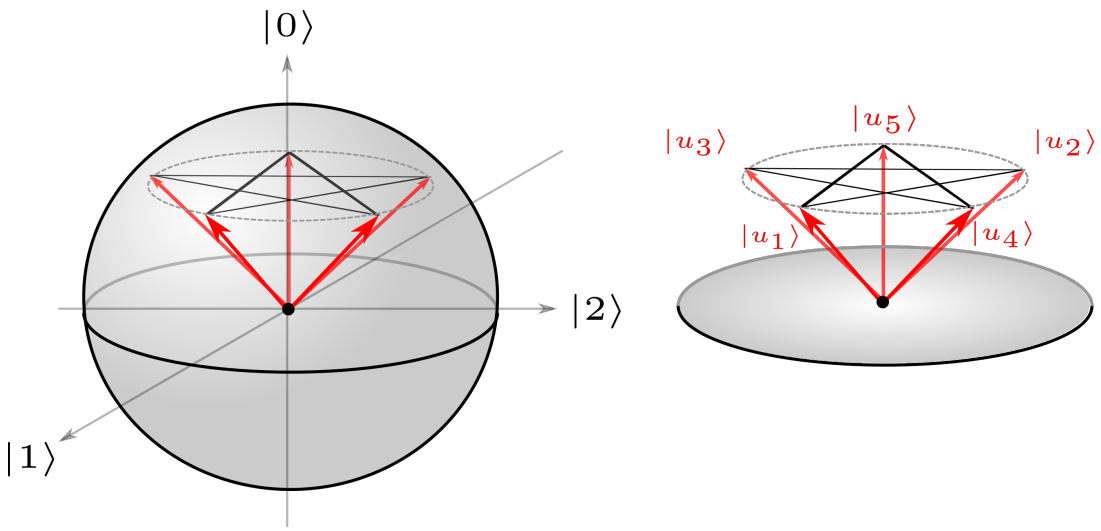


Figure 2.11: Reference KCBS experiment on a qutrit leading to a maximal violation of the KCBS KS non-contextuality inequality. An experimenter can freely choose between five rank one projective measurements $|u_i\rangle$ that satisfy cyclic compatibility: $\langle u_i|u_{i\oplus 1}\rangle = 0$. The qutrit is initially prepared in the state $|0\rangle\rangle$ and then measured in one of five measurement contexts (i,i⊕1). After the measurement the qutrit state is reset.

us to represent them as vectors in $\mathbb{R}^3$, see Figure 2.11. The angle between each of the five vectors and the $|0\rangle$ qutrit axis is $\cos^2(\theta) = \frac{1}{\sqrt{5}}$. The following is an explicit expression for five suitable qutrit states that also satisfy cyclic compatibility:

$$|u_j\rangle = (\cos(\theta),\, \sin(\theta)\sin(j\phi),\, \sin(\theta)\cos(j\phi))^T\,,$$

for $j = 1 \ldots 5$ with $\phi = \dfrac{4\pi}{5}$.

For these five projective measurements acting on the $|0\rangle$ qutrit state, QM predicts

$$\sum_{i=1}^{5} p(0,1|i,i+1) = \frac{5}{\sqrt{5}} = \sqrt{5} > 2.$$

Not only does the reference quantum experiment $|0\rangle$, $\{|u_i\rangle\}_{i=1}^5$ produce correlations that violate the KCBS non-contextuality inequality, but this violation is also maximal, as we will see in Section 3.

One can generalize the KCBS KS non-contextuality inequality to greater cycle lengths n. For the odd integers $n \geq 5$, where n parametrizes the number of dichotomic projective measurements obeying cyclic compatibility, one finds a class of KS non-contextuality inequalities [5]

$$\sum_{i=1}^{n} p(0,1|i,i+1) \leq \frac{n-1}{2}. \tag{2.5}$$

The quantum supremum for a given cycle length n is

$$B_q^{(n)} = \frac{n\cos\left(\frac{\pi}{n}\right)}{1+\cos\left(\frac{\pi}{n}\right)}\,, \tag{2.6}$$

and can be achieved by preparing the qutrit state $|0\rangle$ and performing the rank one projective measurements

$$|u_j^{(n)}\rangle = (\cos(\theta_n),\, \sin(\theta_n)\sin(j\phi_n),\, \sin(\theta_n)\cos(j\phi_n))^T\,, \tag{2.7}$$

for $j = 1 \ldots n$, $\cos^2(\theta_n) = \dfrac{\cos\left(\frac{\pi}{n}\right)}{1+\cos\left(\frac{\pi}{n}\right)}$, and $\phi = \dfrac{(n-1)}{n}\pi$.

In Section 3, we adopt the graph-theoretic framework proposed in [11] to classify correlations of an arbitrary experiment. It turns out that for a general linear combination of probabilities one can relate the KS non-contextual and quantum suprema to certain graph invariants. Computing these graph invariants corresponds to solving a mathematical optimization problem. In view of self-testing, this will enable us to identify extremal correlations and study whether these can only be realized in an essentially unique manner.

Chapter 3

CSW graph-theoretic approach to quantum correlations

In [11], the authors propose a graph-theoretic approach to classifying the statistics of an arbitrary correlation experiment that aims to determine the value of some linear combination of probabilities. In particular, [11] identifies a hierarchical structure that classifies correlations as classical (KS non-contextual), quantum, or more generally obeying the exclusivity principle [1]. This hierarchical structure is established by associating with a correlation experiment its corresponding "exclusivity graph". The exclusivity structure obeyed by the experiment imposes non-trivial constraints on the set of permissible correlations compatible with KS non-contextuality, QM, or the exclusivity principle. In particular, upper bounds of the relevant linear combination of probabilities within these three classes of theories, can be related to graph invariants of the experiment's exclusivity graph.

3.1 Generic correlation experiments in the CSW framework

The CSW framework takes an operationalist approach to correlation experiments, in the sense that any such experiment is described in terms of a sequence of (reproducible) experimental procedures. For example, one such experimental procedure may be to prepare the system in a certain way. If they provide the experimenter with an output (apart from the transformed system), such experimental procedures can be considered *tests*. For a preparation P and test M with outcomes $\{k_i\}_i$, an operational theory specifies the probabilities $p(k_i|M,P)$.

There may be multiple equivalent ways of preparing a system, in the sense that we can never tell the preparation procedures apart by experiment, as they yield identical probability distributions

[1] The exclusivity applies to pairwise exclusive events and bars their added probabilities from exceeding 1. We will come back to it shortly.

for all tests. The equivalence class of operationally indistinguishable preparation procedures P defines a state. In the same manner, the equivalence class of operationally indistinguishable tests M defines an observable[2].

We now define what it means for two observables to be compatible, or jointly measurable. While in QM two projective measurements are compatible if and only if the corresponding Hermitian operators commute, we want to define compatibility in operational terms, in order for the definition to be applicable to arbitrary operational theories.

Definition 3.1. Two observables M_1, M_2 with outcome sets $\{k_i\}_i$ and $\{k'_i\}_i$, respectively, are *jointly measurable* if there exists an observable M_{12} with outcome set $\{k_i, k'_j\}_{i,j}$ such that:

1. $\forall$ preparations P, $\forall$ outcomes $k \in \{k_i\}_i$:
   $p(k|M_1, P) = \sum_{k'} p(k, k'|M_{12}, P)$
2. $\forall$ preparations P, $\forall$ outcomes $k' \in \{k'_j\}_j$:
   $p(k'|M_2, P) = \sum_{k} p(k, k'|M_{12}, P)$

One may think of M_{12} as measuring both M_1 and M_2 simultaneously, placing the outcome of M_1 into its first and the outcome of M_2 into its second output register. By discarding one of the outcomes, we reduce M_{12} to the corresponding single-outcome measurement. At the level of probability distributions, discarding one of the outcomes corresponds to a summation over all possible outcomes held by the discarded register. For QM as operational theory and projective measurements, this definition is equivalent to operator commutativity. For general quantum measurements, Definition 3.1 implies that the POVM M_1 and M_2 are obtained by coarse graining the POVM M_{12} that jointly realizes both.

An *event* is characterized by a list of compatible tests $(M_1, M_2, \ldots, M_n)$ that yield some outcomes $(X_1, X_2, \ldots, X_n)$. Let $X_1, X_2, \ldots, X_n \,|\, M_1, M_2, \ldots, M_n$ denote such event. Some events may be operationally equivalent, in the sense that they have the identical probability of occurring, for all initial preparations of the system. Analogous to our treatment of preparations and tests, operationally equivalent test-outcome tuples define the same event. Within QM, a state is defined by a density operator, whereas equivalent events correspond to the same positive semi-definite operator or projector.

A correlation experiment, like the KCBS or CHSH experiment, consists of one or multiple experimenters performing subsets of compatible tests on an intial preparation P of the system. For instance, in the KCBS experiment, an experimenter chooses between five measurement contexts $(i, i \oplus 1)$. We assume that an experimenter can freely choose between measurement contexts, i.e. subsets of compatible tests, meaning that this choice is not correlated with say the

[2]This identification of operationally equivalent tests and preparations parallels the operational equivalence classes we will introduce in Section 6, when introducing the notion of Spekkens contextuality.

system being measured. By recording which of all could-be events occurs for many repetitions of this procedure, the experimenter(s) can determine the value of some relevant linear combination $\sum_i w_i p(\epsilon_i|P)$, where $w_i > 0$ and ϵ_i denotes a possible event i.e. a set of compatible tests that yield some outcomes. Sometimes we will omit the initial preparation P in our formulas. Let us now discuss what conditions and assumptions allow us to impose non-trivial constraints on correlations.

We define two events to be exclusive if they specify different outcomes for identical tests.

Definition 3.2 ([1])**.** For outcome-repeatable tests, two events

$$X_1, X_2, \ldots, X_n \,|\, M_1, M_2, \ldots, M_n \quad \text{and} \quad X'_1, X'_2, \ldots, X'_m \,|\, M'_1, M'_2, \ldots, M'_m$$

are *exclusive* if for some i and j

$$M_i \equiv M'_j \text{ and } X_i \neq X'_j.$$

An important assumption that underlies the CSW framework is that all measurements are assumed to be outcome-repeatable, meaning that performing two identical tests in a consecutive manner always yields identical outcomes. As such, we can, by performing a subsequent measurement M_i, perfectly distinguish between two exclusive events. The assumption of outcome repeatability compells us to restrict quantum measurements to PVM, as this is not a general feature of unsharp quantum measurements. Dropping the assumption of outcome repeatability and thereby accounting for general POVM quantum measurements will not provide us with any non-trivial constraints distinguishing the set of quantum correlations from the the set of general probabilistic assignments, at least not within the CSW framework. The reason for this that the exclusivity principle, which we invoke in Section 3.3 to obtain constraints on correlations, is in general not true for theories within which pairwise compatibility does not imply joint compatibility (Specker's "fundamental theorem", see Section 2.1). There exist POVM that are pairwise compatible, but fail to satisfy joint compatibility [22]. Section 3.4 will discuss in more detail how the CSW framework falls apart when allowing for general unsharp measurements.

For a correlation experiment with two exclusive could-be measurements events ϵ_1 and ϵ_2, the sum of their probabilities cannot exceed 1, for all initial preparations P: Take M to be the test that perfectly distinguishes between ϵ_1 and ϵ_2, like in Definition 3.2, with ϵ_1 specifying the outcome X_1 and ϵ_2 the outcome X_2. Imagine performing the test M after one input-output

round of the correlation experiment:

$$\begin{aligned} 1 &\geq p(\{X_1, X_2\}|M) \\ &\geq p(\{X_1, X_2\}|M, \epsilon_1)\, p(\epsilon_1) + p(\{X_1, X_2\}|M, \epsilon_2)\, p(\epsilon_2) \\ &= p(\epsilon_1) + p(\epsilon_2). \end{aligned}$$

The exclusivity principle extends this property to sets of pairwise exclusive events:

Principle 3.3 (Exclusivity principle)**.**
For a set of pairwise exclusive measurement events of a correlation experiment $\{\epsilon_i\}_i$*,*

$$\sum_i p(\epsilon_i) \leq 1,$$

for all initial preparations.

We denote the set of correlations that are compatible with the exclusivity principle by E_1.

Not all valid assignments of probabilities to events of a correlation experiment, i.e. probabilistic models of a correlation experiment, are compatible with the exclusivity principle. For example, consider Specker's three boxes, as introduced in Section 2.1. The relevant linear combination of probabilities that this correlation experiment tests is

$$S = \frac{1}{3}\sum_{i=1}^{3} p(0,1\,|\,i, i\oplus 1) + \frac{1}{3}\sum_{i=1}^{3} p(1,0\,|\,i, i\oplus 1).$$

For both sums, the events that appear are pairwise exclusive. Therefore, within theories obeying the exclusivity principle, S is upper-bounded by $\frac{2}{3}$. However, by assigning the probability $\frac{1}{2}$ to each event appearing in S we obtain a probability assignment that is compatible with pairwise exclusivity, but incompatible with the exclusivity principle. Furthermore, this assignment of probabilities is consistent with perfectly anti-correlated outcomes for all measurement contexts, as was considered in Section 2.1. While the exclusivity principle must not be true for arbitrary theories, it does apply to QM if all measurements are projective, and by extension to KS non-contextual correlations. This can be seen by noting that pairwise exclusive events must correspond to pairwise orthogonal projectors, as well as noting that for a set of orthogonal projectors $\{\Pi_i\}_i$, $(\,\mathbb{1} - \sum_i \Pi_i\,)$ is a projector. Therefore, in Section 3.3, we may use the exclusivity principle to impose constraints on the set of quantum (and KS non-contextual) correlations.

If we extend quantum measurements to general POVM, the exclusivity principle must no longer hold, as we just demonstrated for the Specker's three boxes correlation experiment (assign the positive operator $\frac{\mathbb{1}}{2}$ to all measurement events).

3.2 The exclusivity graph of a correlation experiment

The exclusivity graph of a correlation experiment is a convenient representation of the exclusivity structure obeyed by events, imposing constraints on what correlations are permissible within different classes of theories. The exclusivity graph of a correlation experiment $S = \sum_i w_i p(\epsilon_i)$ has a vertex for each measurement event ϵ_i appearing in S. Furthermore, exclusive measurement events are represented as adjancent vertices, as shown in Figure 3.1 for the KCBS correlation experiment, or in Figure 3.2 for Specker's three boxes.

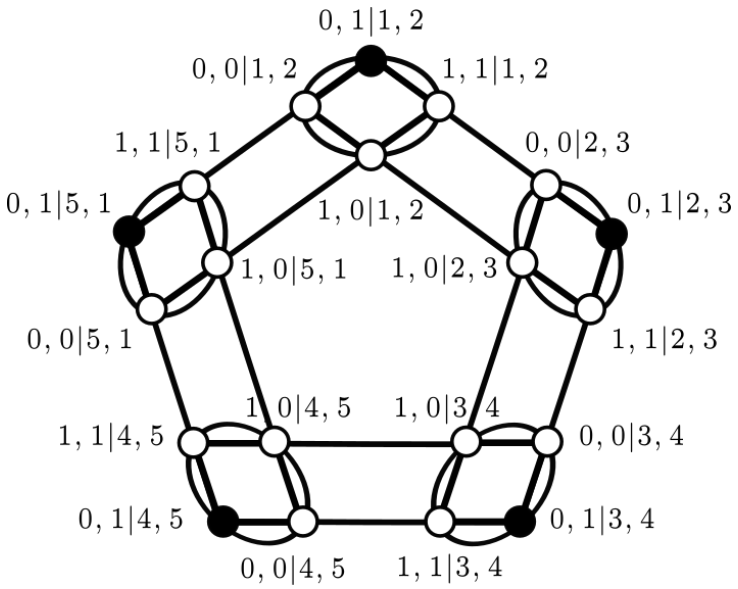
3.3 Hierarchical structure of correlations

We will now relate upper bounds for the quantity S, consistent with KS non-contextuality, QM, or the exclusivity principle, to graph invariants of the experiment's exclusivity graph. All three graph invariants can be computed by solving a corresponding mathematical optimization problem. The proofs we present are adapted from [11]:

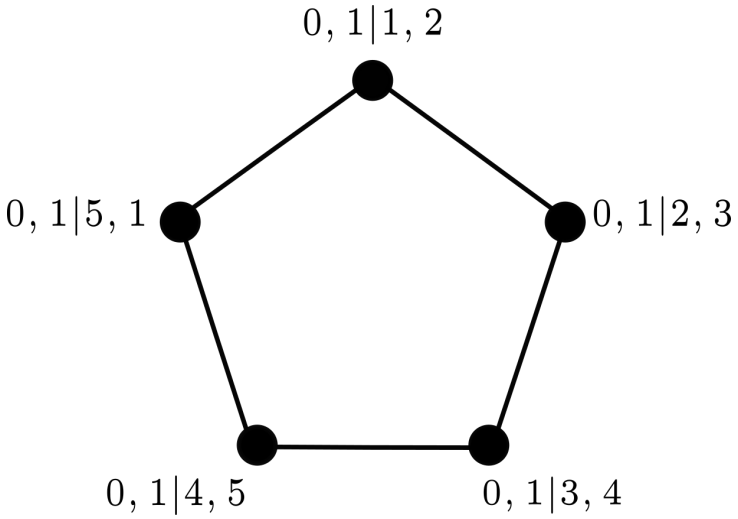
- **KS non-contextual correlations:** The CSW approach decouples KS non-contextuality from the operational theory that is QM and extends it to arbitrary operational theories. For a given exclusivity graph, classical i.e. KS non-contextual correlations are those that can be written as a convex combination of deterministic assignments $\nu : V \mapsto \{0, 1\}$ that obey the exclusivity principle, where V is the vertex set of the correlation experiment. Let us verify that this extension to arbitrary operational theories is compatible with KS non-contextuality as defined in Definition 2.4: Assuming QM and projective measurements, each event defines a (distinct) projector, as we identify operationally equivalent events. Furthermore, pairwise exclusive events correspond to pairwise orthogonal projectors and can be regarded as part of the same resolution of the identity. The correlation experiment is consistent with a KS non-contextual description according to Definition 2.4 if and only if all correlations are compatible with an assignment ν like above, up to convex combination. Therefore, for the case of QM as operational theory and perfectly sharp measurements, the generalized notion of classicality in [11] reduces to KS non-contextuality like in Definition 2.4.

  As will be discussed in Section 6.1, lifting the notion of KS non-contextuality as relevant criterion for non-classicality to arbitrary operational theories, and assigning outcomes in a deterministic fashion to potentially unsharp measurements, gives rise to a number of inconsistencies. Therefore, this "incremental" [37] approach towards an operational notion of non-contextuality seems to be flawed. An alternative will be presented in Section 6.2.

  We now wish to determine the maximal value of S that is compatible with a classical description like above. For a set of pairwise adjacent vertices, the sum of the probabilities



(a) Graph depicting all events of the KCBS correlation experiment and the underlying exclusivity structure. An event is defined in terms of a measurement context and outcome tuple, and is represented by a vertex of the graph. Adjacent vertices correspond to "mutually exclusive" events. Figure copied from [11].



(b) Simplified exclusivity graph for the KCBS correlation experiment. Contains only those vertices in (a) whose probabilities appear in the KCBS KS non-contextuality inequality. Figure copied from [11].

Figure 3.1: KCBS exclusivity graph

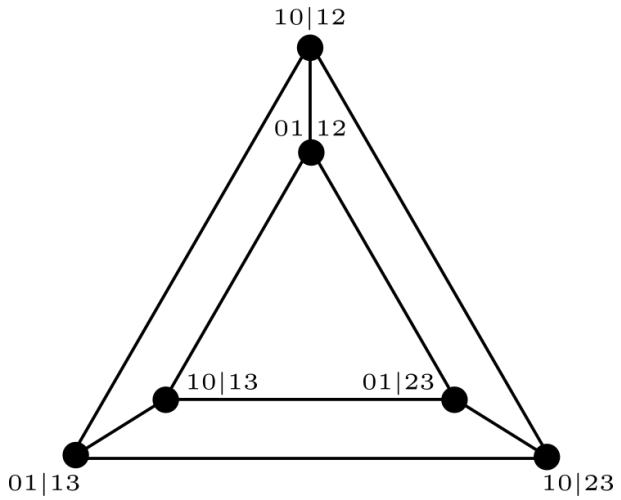


Figure 3.2: Exclusivity graph for Specker's three boxes correlations experiment, see Section 2.1. The outcome 0 corresponds to the opened box being empty, whereas the outcome 1 corresponds to the opened box containing a gem.

assigned to these cannot exceed 1. Only two independent vertices, meaning two vertices that are non-adjacent can both receive the valuation 1. It follows that the maximal value of S is

$$S_{nc} = \max_{U} \sum_{i \in U} w_i,$$

where the expression is maximized over all independent sets of the exclusivity graph. This is the independence number of the (weighted) exclusivity graph. Section 2.9 presents concrete values for the odd n-cycle scenarios, see Equation 2.5.

- **Quantum correlations:** The set of quantum correlations, restricting to projective measurements, is of the form[3]

$$p(\epsilon_i) = \langle \Psi | \Pi_i | \Psi \rangle ,$$

for some quantum state $|\Psi\rangle$ and Π_i the projector corresponding to the event ϵ_i. Adjacent events must be represented by orthogonal projectors:

$$\Pi_i \Pi_j = 0 \text{ for } \epsilon_i, \epsilon_j \text{ adjancent.}$$

For an arbitrary quantum realization $|\Psi\rangle$, $\{\Pi_i\}_i$, define the vectors

$$|u_0\rangle \coloneqq |\Psi\rangle \text{ and } |u_i\rangle \coloneqq \frac{\Pi_i |u_0\rangle}{\sqrt{\langle u_0 | \Pi_i | u_0 \rangle}}$$

[3]We can w.l.o.g. assume a pure quantum state. For mixed states ρ_A, consider the purification $|\Psi\rangle_{AE}$ and projectors $(\Pi_i \otimes \mathbb{1}_E)$.

The $(n+1)\times(n+1)$ Gram matrix[4] X of the vectors $\{\langle u_i|u_0\rangle\,|u_i\rangle\}_{i=0}^{n}$ obeys the following constraints:

$$\begin{aligned} X_{00} &= 1\,, \\ X_{0i} &= X_{ii}\,, \\ \text{and } X_{ij} &= 0 \text{ for } \epsilon_i,\, \epsilon_j \text{ adjacent} \end{aligned}$$

Additionally, $\sum_i w_i X_{ii} = \sum_i w_i |\langle\Psi|\Pi_i|\Psi\rangle|^2$, the value of the linear combination S for the quantum model $|\Psi\rangle$, $\{\Pi_i\}_i$. As we can construct a Gram matrix that satisfies these constraints and which is related to S like above for all quantum experiments , the maximum value of S, for a given exclusivity graph, that is compatible with QM is upper-bounded by the so-called Lovász semi-definite program (SDP) [5]

$$\begin{aligned} \max \sum_{i=1}^{n} w_i X_{ii} & \\ \text{subject to } X_{ii} &= X_{01}\,,\ 1\le i\le n \\ X_{ij} &= 0\,, \qquad \text{for } \epsilon_i,\ \epsilon_j \text{ adjacent} \\ X_{00} &= 1\,, \qquad X\in\mathcal{S}_+^{n+1}, \end{aligned} \tag{3.1}$$

where $\mathcal{S}_+^{1+n}$ is the set of positive semi-definite $(n+1)\times(n+1)$ matrices.

One might naively think that the Lovász bound is tight: For an optimal solution of the Lovàsz SDP 3.1, X^*, with Gram decompostion $\{|u_i\rangle\}_{i=0}^{n}$, $S=\sum_{i=1}^{n} w_i X_{ii}^*$ can be realized by the quantum experiment which consists of an experimenter performing the rank-one projective measurements $|\tilde{u_i}\rangle\langle\tilde{u_i}|$ on the state $|\tilde{u_0}\rangle$, where the vectors $\{|\tilde{u_i}\rangle\}_{i=0}^{n}$ are obtained by normalization:

$$p(\epsilon_i)=\mathrm{tr}(|\tilde{u_i}\rangle\langle\tilde{u_i}|\tilde{u_0}\rangle\langle\tilde{u_0}|)=|\langle\tilde{u_0}|\tilde{u_i}\rangle|^2=\frac{|\langle u_0|u_i\rangle|^2}{\langle u_0|u_0\rangle\langle u_i|u_i\rangle}=|\langle u_i|u_i\rangle|=X_{ii},$$

where $\epsilon_i\equiv 1\,|\,|\tilde{u_i}\rangle\langle\tilde{u_i}|$ and the last equality follows from the constraints a feasible solution of the Lovász SDP 3.1 must satisfy. However, the Lovász bound on quantum correlations arising from projective measurements is in general not tight. For Bell scenarios with multiple space-like separated subsystems, the concrete physical setting of the correlation experiment imposes additional constraints on what projectors $\{\Pi_i\}_i$ are physical.

[4]The Gram matrix X of a set of vectors $v_1,\ldots,v_n$ in an inner product space $(V,\langle\cdot,\cdot\rangle)$ is the $n\times n$ matrix with entries $X_{ij}=\langle v_i|v_j\rangle$. The Gram matrix of a set of vectors is positive semi-definite and all positive semi-definite matrices are the Gram matrix for some non-unique list of vectors. For a $n\times n$ positive semi-definite matrix X, we can always find a suitable list of vectors in $(\mathbb{C}^n,\langle\cdot,\cdot\rangle_{\mathrm{std}})$, where $\langle\cdot,\cdot\rangle_{\mathrm{std}}$ is the standard inner product, by computing the Cholesky decomposition of X. Acting with an arbitrary isometry $V:\mathbb{C}^n\to\mathbb{C}^d$, $d\ge n$, yields a Gram decomposition in $(\mathbb{C}^d,\langle\cdot,\cdot\rangle_{\mathrm{std}})$.

In particular these have to respect locality $\Pi_i = \Pi_A^{(i)} \otimes \Pi_B^{(i)}$. For contextuality scenarios the Lovász bound is generally tight, as is the case for the class of odd n-cycle scenarios presented in Section 2.9.

- **Correlations in E_1:** The fractional packing number of the weighted exclusivity graph is defined as

$$\max_{\{p_i\}_i} \sum_{i \in V} w_i \, p_i,$$

where the expression is maximized over all $\{p_i\}_i$ with $p_i \geq 0$ and $\sum_{i \in C} p_i \leq 1$, for all subsets C of pairwise exclusive events (cliques) of the exclusivity graph. This is just the the maximum value of S for correlations in E_1. For the KCBS inequality E_1 correlations can achieve $\sum_{i=1}^{5} p(0,1|i,i+1) = \frac{5}{2}$.

3.4 What about unsharp measurements?

The CSW framework falls apart when we allow for outcome unsharp measurements. In particular, as we will see in the next section, the Lovász number 3.1 no longer bounds general quantum behaviours, and even trivial POVM $\{a\mathbb{1}\}_a$ can realize the most general probabilistic models [27]. In particular, they can produce violations of KS non-contextuality inequalities that exceed those achieved by just projective measurements. One arrives at the pathological conclusion that, within the CSW framework, if we do not restrict the set of measurements, all correlations are quantum (compatible with QM), yet all correlations can be produced by trivial POVM, which are intuitively classical[5]. Thus, the hierarchical structure of correlations established in Section 3.3 breaks down for this case. Futhermore, extending our classicality criterion, namely KS non-contextuality, to unsharp measurements is conceptually problematic and raises numerous inconsistencies, as will be discussed in Section 6.1.

3.5 Complications that arise without space-like separation

Let us now comprehend in what sense trivial POVM are able to realize arbitrary probabilistic models, in particular violate KS non-contextuality inequalities, and how come one doesn't encounter the same complications for Bell-type scenarios.

To work with a concrete example, consider the KCBS correlation experiment with exclusivity graph Figure 3.1. We present an argument that parallels the one presented in [27]. Let $(i, i \oplus$

[5] A trivial POVM generates a random output, totally independent of the measured system's state. One can therefore implement an operationally equivalent measurment by discarding the system and flipping some potentially biased coins.

1) denote two compatible dichotomic measurements an experimenter can perform, which we will assume to be compatible POVM. Importantly, there generally exist many joint POVM $\{M_{00}^{i,i\oplus1}, M_{01}^{i,i\oplus1}, M_{10}^{i,i\oplus1}, M_{11}^{i,i\oplus1}\}$, satisfying Definition 3.1, and we cannot rule out any of these without making additional assumptions. For Bell-type correlation experiments, we can impose additional constraints due to locality, namely that compatible measurements act on local Hilbert spaces, and in particular that compatible POVM are commuting. In contrast to projective measurements, commutativity is sufficient but not necessary for two POVM to be compatible: If and only if all elements from both POVM commute, their joint POVM is uniquely determined as

$$M_{ab}^{i,i\oplus1} = M_a^i M_b^{i\oplus1},$$

where $i \equiv \{M_a^i\}_a$ and $i \oplus 1 \equiv \{M_b^{i\oplus1}\}_b$ [27].

Let us compare the possible correlations that may arise from commuting POVM to those that can arise from non-commuting, compatible POVM:

- **commuting POVM (Bell-type scenarios):** Due to commutativity, the probability of an event $x, y \,|\, i, i \oplus 1$ can be calculated like

  $$p(x, y \,|\, i, i \oplus 1) = \mathrm{tr}(\rho M_x^i M_y^{i\oplus1}).$$

  For Bell tests with remote measurement contexts, the joint measurement is given by a tensor product.

  If we consider all POVM i to be trivial, $M^i \equiv \{a_0^i \mathbb{1}\,,\, a_1^i \mathbb{1}\}$, then we find a joint probability distribution for all measurements M^i that is compatible with all measureable marginal correlations for the five measurement contexts,

  $$p(x, y, z, v, w) = a_x^1\, a_y^2\, a_z^3\, a_v^4\, a_w^5,$$

  where $x, \ldots, w$ denote the outcomes of the five measurements M^i. Finding such a global probability distribution implies that all correlations produced by the experiment are KS non-contextual, and in particular do not violate a KS non-contextuality inequality. To see this, assume an ontic state space that contains one ontic state for each outcome tuple (x, y, z, v, w) and fixes the outcomes of the measurements M^i accordingly. The preparation that assigns the probability $p(x, y, z, v, w)$ to the ontic state corresponding to that outcome tuple reproduces all correlations of the experiment. The ontological model is KS non-contextual, as it is deterministic and assigns outcomes to the individual operators, independent of the context they are measured in. Thus, for Bell scenarios with space-like separated parties, all compatible POVM are commuting, and trivial POVM cannot violate Bell inequalities. As alluded to in Section 2.8, we cannot carry out the same procedure for

general contextuality scenarios, as the set of simultaneous eigenvalues for two compatible operators A, B is in general not given by the Cartesian product of individual eigenvalues $\sigma(A) \times \sigma(B)$.

- **non-commuting POVM (KS-type scenarios):** General compatibile POVM, can be realized jointly in many different ways. A joint realization satisfying Definition 3.1 is of the form

$$\{M_{00}^{i,i\oplus1}, M_{01}^{i,i\oplus1}, M_{10}^{i,i\oplus1}, M_{11}^{i,i\oplus1}\}.$$

For a single system, $\{M_{xy}^{i,i\oplus1}\}_{xy}$ must no longer be the product of the individual POVM. As Kunjwal points out [27], there is one undetermined degree of freedom, i.e. one positive semi-definite operator, say $M_{01}^{i,i\oplus1}$, of the joint POVM not fixed by the marginal POVM $\{M_a^i\}_a$ and $\{M_b^{i\oplus1}\}_b$:

$$\begin{aligned} M_{00}^{i,i\oplus1} &= M_0^i - M_{01}^{i,i\oplus1} \\ M_{11}^{i,i\oplus1} &= M_1^{i\oplus1} - M_{01}^{i,i\oplus1} \\ M_{10}^{i,i\oplus1} &= \mathbb{1} - M_{01}^{i,i\oplus1} - M_{00}^{i,i\oplus1} - M_{11}^{i,i\oplus1} \\ &= \mathbb{1} - M_0^i - M_1^{i\oplus1} + M_{01}^{i,i\oplus1}. \end{aligned}$$

Let us assume $\{M_a^i\}_a = \{\frac{\mathbb{1}}{2}, \frac{\mathbb{1}}{2}\}$ and $\{M_b^{i\oplus1}\}_b = \{\frac{\mathbb{1}}{2}, \frac{\mathbb{1}}{2}\}$ to be trivial and acting on a qubit system. A possible joint POVM is

$$M_{00}^{i,i\oplus1} = 0\ ,\ M_{01}^{i,i\oplus1} = \frac{\mathbb{1}}{2}\ ,\ M_{10}^{i,i\oplus1} = \frac{\mathbb{1}}{2}\ ,\ M_{11}^{i,i\oplus1} = 0\ ,$$

which is itself trivial. Furthermore, it produces correlations that exceed the Lovász bound of the KCBS correlations experiment. This shows that, taking KS non-contextuality as the revelant criterion of non-classicality and extending it to unsharp measurements, an intuitively classical procedure can produce correlations that are highly non-classical, even post-quantum, according to the hierarchical structure established in [11]. This example can be generalized, by making use of the one undetermined positive semi-definite operator, to show that trivial POVM can actually realize arbitrary probabilistic assignments.

The above discussion is consistent with the observation that the notion of local causality is formulated entirely theory independent and makes no mention of the sharpness of measurements. Indeed, within QM, Bell inequalities apply to general quantum measurements just as they apply to projective quantum measurements, for reasons just highlighted. The aim of this Section was to motivate why, for general contextuality scenarios without additional locality constraints, extending Bell's notion of local causality, or rather local determinism, to apply to non-remote measurements contexts runs into problems.

Shifting to a general prepare-and-measure scenario, instead of extending the assumption of outcome determinism to arbitary unsharp measurements, we will follow Spekkens and adopt his proposal of a new notion of contextuality [40], which will be discussed in Section 6. Lastly, the revised tests of non-classicality on which the protocol in 7 hinges, factor in source-measurement correlations, a term coined by Kunjwal in [27], which one can think of as quantifying the sharpness of the measurements involved in the correlation experiment. Within QM, the source-measurement correlations capture the "classicality" of trivial POVM measurements, and these in fact exhibit classical correlations.

Chapter 4

Self-testing

Self-testing aims to certify properties of unknown quantum devices, only by interacting classically with these, i.e. by passing classical input strings to devices and receiving classical outputs, for instance a binary string. Ideally, we want to make a minimal number of assumptions about the devices themselves, largely treating them as black boxes[1]. This is useful for certifying the behaviour of untrusted or noisy quantum devices, even in an adversarial scenario, and for the most part does not require the experimenter to have knowledge about the exact inner workings of the devices. More concretely, many quantum key distribtion (QKD) protocols for quantum cryptography, such as Ekert91 [19], require the generation of a large number of maximally entangled states. To ensure information-theoretic security, one has to confirm the quality of the entanglement source used. Self-testing provides a means of doing this.

A self-testing protocol is essentially a correlation experiment, as introduced in Section 3.1, that allows us to make powerful inferences about the internal workings of the devices. The following example will clarify the concept. We will still assume measurements to be sharp for now:

Consider the CHSH correlation experiment, as introduced in Section 2.8, with two space-like separated systems, $\mathcal{A}$ and $\mathcal{B}$, as well as two experimenters that can perform local measurements. Assuming no-signalling, i.e. that local measurements on $\mathcal{A}$ do not disturb the outcome statistics of local measurements on $\mathcal{B}$, remote measurements are compatible according to Defintion 3.1. Recall that the CHSH correlation experiment is a test for violations of the inequality

$$\begin{aligned}\frac{1}{4}(\ & p(x=y|A_1,B_1)+p(x=y|A_2,B_1)\\ & +p(x=y|A_2,B_2)+p(x\neq y|A_1,B_2)\ )\leq\frac{3}{4}\end{aligned}$$

[1] Self-testing schemes that do not require assumptions about the underlying mechanics of the devices, and consider only the input-correlations they generate, are called "device-independent".

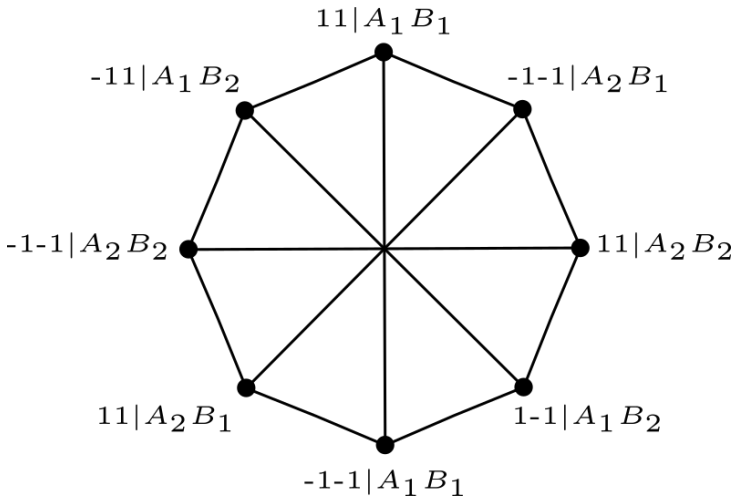


Figure 4.1: Exclusivity graph for the CHSH correlations experiment, see Section 2.8. Like in Section 2.8, the outcomes of local measurements are labelled ± 1. Vertices correspond to measurement events, exclusive measurement events are represented as adjacent.

The exclusivity graph of the CHSH correlation experiment can be seen in Figure 4.1; the equal weighting $\frac{1}{4}$ each vertex receives is omitted for readability.

The Lovász number of the CHSH exclusivity graph corresponds to the Tsirelson bound $B_q = \frac{1}{2}(1 + \frac{1}{\sqrt{2}})$. Thus, the Lovász bound is tight and the measurements 2.3, together with the maximally entangled EPR state, produce a maximal violation of the CHSH inequality 2.2.

In fact, due to space-like separation, the Lovász bound is a tight upper bound that all correlations compatible with QM have to obey, even if we allow for general POVM measurements. This follows from the Naimark dilation theorem [46]. Informally, we can extend every non-projective quantum measurement $\{M_i\}_i$ acting on the system $\mathcal{A}$ to a projective measurement $\{\Pi_i^{\mathcal{A}}\}_i$ that acts on a composite system ($\mathcal{A}$ + ancilla), where the ancillary system is prepared in some well-defined state $|0\rangle\langle 0|$. The extended projectors acting on the composite system ($\mathcal{A}$ + ancilla) produce the same output correlations as the unsharp measurements acting on the single system $\mathcal{A}$. The same can be said about local measurements on subsystem $\mathcal{B}$. Now, due to space-like separation, the projectors $\Pi_i^{\mathcal{A}} \otimes \Pi_j^{\mathcal{B}}$ corresponding to exclusive measurement events commute. As such, we can follow the same line of reasoning as in Section 3.5 to argue that the maximal violation of the CHSH inequalitiy, even with access to unsharp measurements is given by the Tsirelson bound. As we saw in Section 3.5, the same considerations do not apply to contextuality scenarios involving compatible measurements on a single system. While we can extend the individual measurements to projective measurements by adding an ancilliary system, thereby associating with each measurement event a projector, we have no grounds to assume that the projectors associated with exclusive measurement are orthogonal. Take for instance the KCBS correlation experiment with exclusivity graph Figure 3.1. Measurements 1,2 and 2,3 are compatible, however not all three are jointly measurable. We can assume the Naimark projectors associated with the measurement events $0,1\,|\,1,2$ and $0,0\,|\,1,2$ to be orthogonal, because they

correspond to distinct outcomes of a single joint measurements. This does not hold for the exclusive measurement events of the KCBS exclusivity graph: $0,1\,|\,1,2$ and $0,1\,|\,2,3$ can not be distinct outcomes of a single joint measurement, as the measurements $1,2,3$ are not even pairwise compatible.

Back to the CHSH correlation experiment: A violation of the CHSH inequality implies that the observed correlations are not compatible with a classical description of the form 2.1, but require quantum resources, in particular entanglement, to generate. What is remarkable about, but not unique to the CHSH Bell inequality[2], is that the quantum model presented in 2.3, leading to a maximal violation of the inequality, i.e. corresponding to the extremal point $\vec{v}_{max} \in \mathcal{Q}$ of the set of quantum correlations, see Figure 4.2, is "essentially unique". By this we mean that the only quantum model $|\Psi\rangle$, $\{\Pi_i\}_i$ predicting a maximal violation of the CHSH inequality is 2.3, modulo trivial changes to the reference experiment which we will specify shortly. A comprehensive proof is given in [44]. Thus, by interacting only classically with our preparation and measurement devices, we can characterize the quantum state and measurements our devices implement. We can only hope to self-test extremal correlations, as other correlations $\vec{v} \in \mathcal{Q}$ correspond to a non-unique convex combination of distinct extremal correlations.

It is apparent that we can never certify exact states and measurements, as we will always be blind to a number of trivial degrees of freedom. For instance, we cannot tell if there are any additional ancilliary systems, $\mathcal{A}'$, $\mathcal{B}'$, prepared in some joint state $\rho_{\mathcal{A}'\mathcal{B}'}$, in the remote labs that we are not probing:

$$\begin{aligned} & |\Psi\rangle_{\mathcal{A}\mathcal{B}}\,, && A_1 = \sigma_{z,\mathcal{A}} \otimes \mathbb{1}_{\mathcal{B}}\,, \ldots \\ \equiv\ & |\Psi\rangle\langle\Psi|_{\mathcal{A}\mathcal{B}} \otimes \rho_{\mathcal{A}'\mathcal{B}'}\,, && A_1 = \sigma_{z,\mathcal{A}} \otimes \mathbb{1}_{\mathcal{A}'\mathcal{B}\mathcal{B}'}\,, \ldots\,. \end{aligned}$$

Furthermore, we can never detect local unitaries $U = U_{\mathcal{A}\mathcal{A}'} \otimes U_{\mathcal{B}\mathcal{B}'}$, i.e. local basis changes:

$$\begin{aligned} & |\Psi\rangle\langle\Psi|_{\mathcal{A}\mathcal{B}} \otimes \rho_{\mathcal{A}'\mathcal{B}'}\,, && A_1 = \sigma_{z,\mathcal{A}} \otimes \mathbb{1}_{\mathcal{A}'\mathcal{B}\mathcal{B}'}\,, \ldots \\ \equiv\ & U(\,|\Psi\rangle\langle\Psi|_{\mathcal{A},\mathcal{B}} \otimes \rho_{\mathcal{A}'\mathcal{B}'})U^\dagger\,, && A_1 = U_{\mathcal{A}\mathcal{A}'}(\sigma_{z,\mathcal{A}} \otimes \mathbb{1}_{\mathcal{A}'\mathcal{B}\mathcal{B}'})U^\dagger_{\mathcal{A}\mathcal{A}'}\,, \ldots\,. \end{aligned}$$

According to Stinespring's Dilation Theorem [46], these trivial degrees of freedom can account for arbitrary local isometries $\mathcal{A} \to \mathcal{A}\mathcal{A}'$, $\mathcal{B} \to \mathcal{B}\mathcal{B}'$. Furthermore, the previous two examples suggest that in order to define a sensible notion of self-testing, we must first define some notion of equivalence for quantum models. We then aim to construct self-testing protocols that single out a unique quantum model, modulo this notion of equivalence. This notion of equivalence should account for some trivial degrees of freedom, like the ones above, but should also not be too broad, in order for self-testing to be a powerful property of select correlation experiments.

[2] Other Bell inequalities may self-test different reference experiments.

Ideally, the equivalence of quantum realizations should have an operational meaning. The following definition applies to Bell-type correlation experiments and does justice to these three criteria.

Definition 4.1 ([30])**.** A quantum experiment involving $|\Psi'\rangle_{\mathcal{A}'\mathcal{B}'\mathcal{P}}$ and local measurements $\{{P'}_x^a\}_x$, $\{{Q'}_y^b\}_y$, where $|\Psi'\rangle_{\mathcal{A}'\mathcal{B}'\mathcal{P}}$ is an arbitrary purification and ${P'}_x^a$ is the positive semi-definite operator corresponding to the outcome x of the local measurement a on system $\mathcal{A}$, is *equivalent* to a reference experiment $|\Psi\rangle_{\mathcal{A}\mathcal{B}}, \{P_x^a\}_x$, $\{Q_y^b\}_y$, if there exists a local isometry $\Phi = \Phi_{\mathcal{A}} \otimes \Phi_{\mathcal{B}}$, $\Phi_{\mathcal{A}/\mathcal{B}} \colon \mathcal{H}_{\mathcal{A}'/\mathcal{B}'} \to \mathcal{H}_{\mathcal{A}'/\mathcal{B}'} \otimes \mathcal{H}_{\mathcal{A}/\mathcal{B}}$, such that

$$(\Phi \otimes \mathbb{1}_{\mathcal{P}}) \left|\Psi'\right\rangle_{\mathcal{A}'\mathcal{B}'\mathcal{P}} = |\text{junk}\rangle_{\mathcal{A}'\mathcal{B}'\mathcal{P}} \otimes |\Psi\rangle_{\mathcal{A}\mathcal{B}}$$

$$(\Phi \otimes \mathbb{1}_{\mathcal{P}})({P'}_x^a \otimes {Q'}_y^b) \left|\Psi'\right\rangle_{\mathcal{A}'\mathcal{B}'\mathcal{P}} = |\text{junk}\rangle_{\mathcal{A}'\mathcal{B}'\mathcal{P}} \otimes (P_x^a \otimes Q_y^b) |\Psi\rangle_{\mathcal{A}\mathcal{B}} \ .$$

Operationally, this means that a physical experiment is equivalent to some reference experiment if by means of local operations alone[3], one can recover the reference experiment, up to some "junk". In particular, the experimenters can for instance extract the reference state, i.e. the maximally entangled EPR state for the CHSH correlation experiment, to their ancilliary systems, by means of local operations alone. As local operations cannot generate entanglement, their initial state must have been maximally entangled. Note that the notion of equivalence, as defined in Definition 4.1, is not an equivalence relation, as it fails to be symmetric [30]. While the experimenters can extract the reference state to ancilliary systems using only local operations, if they have access to an equivalent experiment, the converse direction must not hold: Consider the reference CHSH experiment 2.3 producing a maximal violation of the CHSH inequality 2.2. Further, consider the equivalent quantum experiment where both experimenters control an additional ancilliary system and both ancillae are prepared in an entangled state. While this extended experiment is equivalent to the reference experiment, the converse does not hold, as we cannot create entanglement by local operations alone.

Definition 4.1 accounts for all of the following statistics-preserving trivial degrees of freedom, as presented in [30]:

1. Local changes of basis
2. adding ancillae to the physical systems, prepared in any joint state (measurements do not act on these)
3. changing the action of an observable outside the support of the state
4. locally embedding the state and operators into a larger (or smaller) Hilbert space

[3]Note that the two experimenters can realize an arbitrary local isometry $\Phi = \Phi_{\mathcal{A}} \otimes \Phi_{\mathcal{B}}$ by preparing ancilliary systems in some well-defined state $|00\rangle$ and subsequently applying local unitary operations.

Claim 4.2 ([30])**.** *Any finite number of changes 1-4 transforms a quantum experiment into an equivalent quantum experiment like in Definition 4.1. Furthermore, any quantum experiment equivalent to some reference experiment can be constructed from that reference experiment by applying a finite number of changes 1-4.*

Let us constrast self-testing via Bell inequalities with self-testing via KS non-contextuality inequalities. Section 5 will examine in detail a self-testing protocol based on the class of odd n-cycle KS non-contextuality inequalities [5]. Importantly, without space-like separated subsystems, we have to give up device-independence: While relaxing the assumption of space-like separation means that we no longer require the device to be split up among remote subsystems, or that it can generate a large number of entangled states, we can no longer make use of quantum non-locality to certify that the correlations were not simulated by some pre-programmed classical computer [44]. In principle, we can never discard this possibility without imposing additional assumptions about the devices. In Section 5.3, we examine how bounding the information carrying capacity or memory of the device can be helpful in this regard. Additionally, the CSW framework in Section 3, which gives us non-trivial constraints on the set of valid quantum models, which we will be use in the self-testing proof, requires assumptions about operator compatibility and sharpness. This is illustrated in Figure 4.3. Since contextuality scenarios no longer require remote subsystems, we need to amend our notion of "equivalence" for quantum experiments, see Definition 4.1, accordingly. We will come back to this in Section 5, but already note that contextuality-based self-testing protocols aim to single out a quantum realization, up to a **global** isometry.

Finally, self-testing results should be robust to noise. Informally, this means that a near-optimal violation of the Bell or KS non-contextuality inequality should imply that all compatible quantum models are, up to equivalence, "close" to the reference quantum experiment. The self-testing protocol we will examine in Section 5 is robust, see Theorem 5.4.

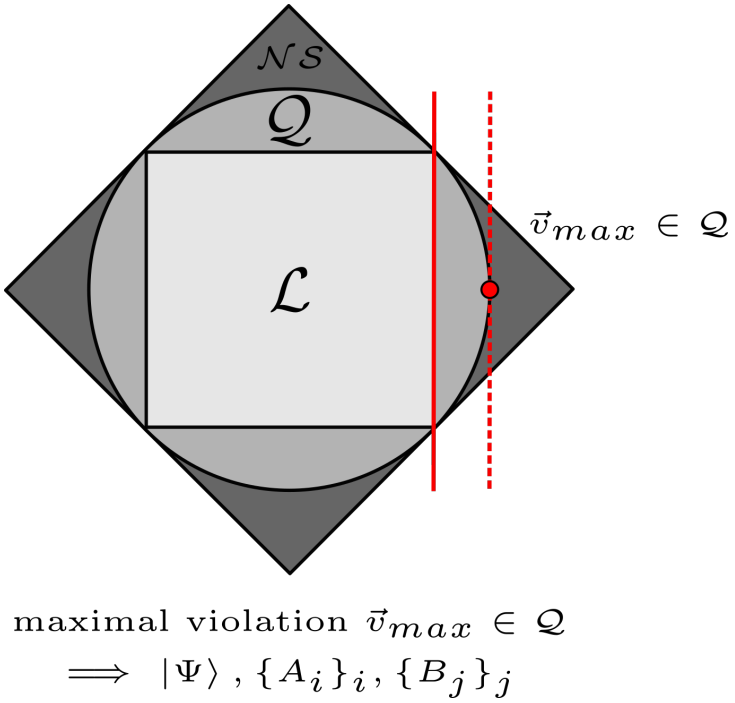


Figure 4.2: Schematic drawing depicting the geometry of the convex sets $\mathcal{L}$, $\mathcal{Q}$, and $\mathcal{NS}$, containing all local, quantum, and no-signalling correlations, respectively. A Bell inequality, like the CHSH inequality, defines a separative hyperplane such that all local correlations $\mathcal{L}$ lie to one "side" of this hyperplane, here represented by the unbroken red line. The extremal point(s) $\vec{v}_{max} \in \mathcal{Q}$ of the set of quantum correlations $\mathcal{Q}$ corresponds to a maximal violation of the CHSH inequality. The CHSH inequality can be used for self-testing, as the quantum model $|\Psi\rangle$,$\{\Pi_i\}_i$ compatible with a maximal violation of the inequality is unique, up to trivial degrees of freedom.

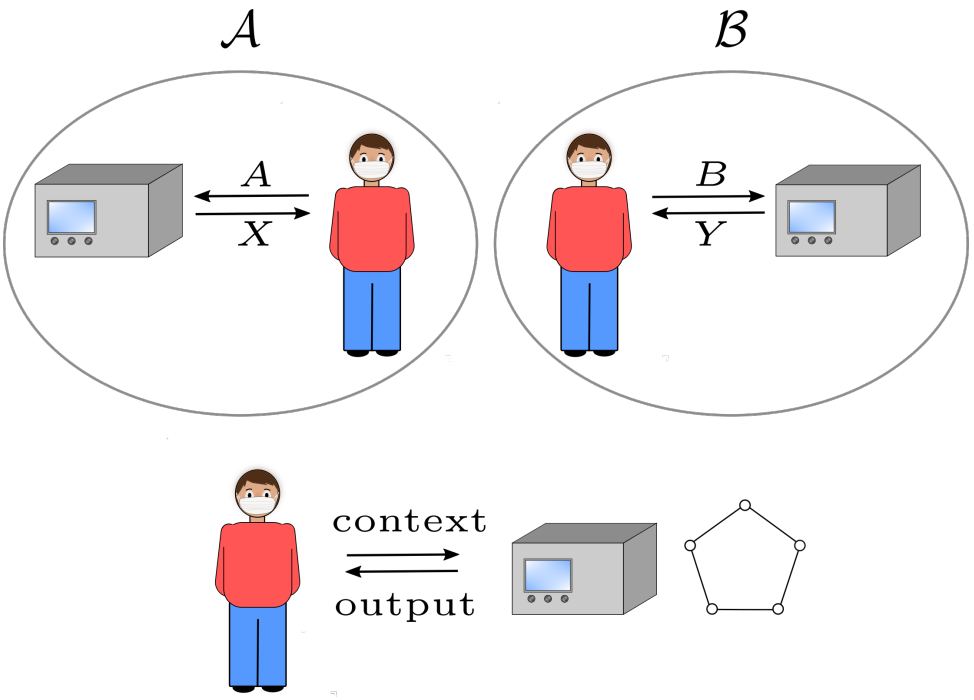


Figure 4.3: Top: Typical device-independent Bell-type self-testing scenario. The two subsystems $\mathcal{A}$ and $\mathcal{B}$ are space-like separated. Accordingly, the black box device, which can be interpreted as a correlation experiment testing a Bell inequality, is split into two non-communicating parts. An experimenter can interact classically with a local sub-device, by passing an input value and recording the outcome returned by the device. A violation of the tested Bell inequality implies that the corresponding correlations were obtained by quantum means, i.e. joint measurements on an entangled state shared between $\mathcal{A}$, $\mathcal{B}$. For some Bell inequalities, a maximal violation is only compatible with an essentially unique quantum realization, i.e. quantum state and set of measurements. **Bottom:** In contrast, self-testing protocols based on KS non-contextuality inequality violations do not assume the device to be split into multiple non-communicating parts. Instead, a single experimenter interacts classically with the device by freely choosing between measurement contexts and determining the corresponding outcome probabilities. Dropping the unnatural assumption of space-like separation complicates the problem of self-testing considerably and we are required to make additional assumptions about the device, for example regarding compatibility relations and information carrying capacity. Thus, contextuality-based self-testing protocols are only semi-device-independent.

Chapter 5

Self-testing via KS non-contextuality inequalities

5.1 Preceding results

In Section 3, we considered general correlation experiments testing some linear combination of probabilities S and found that any quantum realization $|\Psi\rangle$, Π_i can be translated into a feasible solution X of the Lovász SDP 3.1, such that the value of S, as predicted by QM, corresponds to the objective function the SDP aims to maximize, evaluated at the feasible solution X. Conversely, if there are no additional physical constraints, for example due to space-like separated subsystems, one can construct a quantum realization for every feasible solution X that achieves $S = \sum_{i=1}^{n} X_{ii}$. Such quantum realization was constructed from an arbitrary Gram decomposition of X. This connection between feasible solutions to the Lovász SDP and quantum models of the correlation experiment will be central to proving robust self-testing based on KS non-contextuality inequalities.

As outlined in Section 4, the essence of self-testing is that a maximal violation of suitable Bell or KS non-contextuality inequalities is only compatible with an essentially unique quantum realization $|\Psi\rangle$, $\{\Pi_i\}_i$, up to statistics-preserving trivial degrees of freedom. Furthermore, the inequalities are tests of non-classicality. If our experiment produces correlations that violate these, this certifies that the correlations were not simulated by some classical mechanism, which would render all proofs of information-theoretic security superfluous. Extra care has to be taken in the case of contextuality scenarios, as will be the topic of Section 5.3. Ideally, such self-tests should be robust to noise, i.e. a ϵ-suboptimal violation should imply that the only compatible quantum realizations are "close" to the ideal reference implementation, up to trivial degrees of freedom.

We will now sketch the proof in [5], which demonstrates that the class of odd n-cycle contextuality scenarios allows for robust self-testing. Recall that an odd n-cycle correlation experiment assumes cyclic compatibility relations, like in Figure 2.10. The self-testing protocol consists of an experimenter choosing freely between the n measurement contexts $(i, i \oplus 1)$ and determining the outcome correlations $p(0,1 \,|\, i, i \oplus 1)$ that constitute the test S. The main "ingredients" of the proof are the following three lemmas:

Lemma 5.1 ([5])**.** *For the class of odd n-cycle exclusivity graphs, $n \geq 5$, the Lovász SDP 3.1 has a unique optimal solution X_n^*.*

From Section 3.3 we know that the reference quantum experiment $\{|u_j^{(n)}\rangle\}_{j=0}^n$ presented in Section 2.9 produces a maximal violation of the odd n-cycle KS non-contextuality inequality, because it attains the Lovász bound. The unique optimal solution X_n^* is obtained from this reference quantum realization by the usual procedure:

$$X_n^* = \text{Gram}(\, |u_0\rangle \,,\, \langle u_1^{(n)}|u_0\rangle |u_1^{(n)}\rangle \,,\, \ldots \,,\, \langle u_n^{(n)}|u_0\rangle |u_n^{(n)}\rangle \,) \,.$$

The following two lemmas will be important to prove robustness:

Lemma 5.2 ([5])**.** *Let $(\mathcal{G}_{ex}, w)$ be a weighted exclusivity graph with n vertices and assume that 3.1 has a unique optimal solution $X^* \in \mathcal{S}_+^{n+1}$. Further, let $\tilde{X}$ be a feasible solution of 3.1 that is ϵ-suboptimal, i.e.*

$$\sum_{i=1}^{n} w_i \, \tilde{X}_{ii} \geq B_q(\mathcal{G}_{ex}, w) - \epsilon,$$

where $B_q(\mathcal{G}_{ex}, w) = \sum_{i=1}^n w_i X_{ii}^*$ is the optimal solution of the Lovász SDP for the weighted exclusivity graph $(\mathcal{G}_{ex}, w)$. Then

$$\|\tilde{X} - X^*\|_F \leq \mathcal{O}(\epsilon) \,,$$

where $\|A\|_F = \text{tr}(A^\dagger A)$ is the Frobenius norm.

Lemma 5.3. *Let $\tilde{X} \in S_+^{n+1}$ be a positive semi-definite matrix, ϵ-close to $X^* \in S_+^{n+1}$ like in 5.1*

$$\|\tilde{X} - X^*\|_F \leq \epsilon.$$

Let $\{|\tilde{u}_i\rangle\}_{i=0}^n \subset \mathbb{C}^d$ and $\{|u_i\rangle\}_{i=0}^n \subset \mathbb{C}^d$ be Gram decompositions of $\tilde{X}$ and X^*, respectively. Then there exists a $d \times d$ unitary matrix U such that for all $i \in \{0, \ldots, n\}$

$$\||\tilde{u}_i\rangle - U|u_i\rangle\|_2 \leq \kappa(n)\epsilon + \mathcal{O}(\kappa(n)^2\epsilon^2) \,, \,,$$

where $\kappa(n) = \frac{1}{2}\sqrt{n+4}$ for all $n > 7$.

We will omit the proofs of Lemmas 5.1 and 5.2, as they do not provide much conceptual insight, and offer only a few comments. We include the proof of Lemma 5.3, as [5] proves only a weaker

version, in the sense that $d = n$. Furthermore, the proof we present establishes a tighter error bound of the order $\mathcal{O}(\kappa(n)\epsilon)$, compared to $\mathcal{O}(\sqrt{(n+1)\epsilon})$ in [5]. As opposed to the result in [5], the proof of Lemma 5.3 makes use of the fact that the verifier has knowledge about the ideal Gram matrix X^*.

The proof of Lemma 5.1 utilizes SDP duality theory and shows a sufficient condition for the unicity of X_n^* to be satisfied by explicit construction. It applies only to the class of odd n-cycle scenarios, as it makes use of the cyclic exclusivity constraints that enter the Lovász SDP 3.1. Lemma 5.2 is a general result for the optimization problem 3.1, and as such applies to arbitrary exclusivity graphs. Like Lemma 5.1, the proof of Lemma 5.2 makes heavy use of SDP duality theory.

Proof. (of Lemma 5.3)
The proof is a generalization of the proof of Theorem 7.3.11 in [23], which only considers the noise-less case $\epsilon = 0$, and we for the most part will adopt their notation.

Define A, $B \in \mathbb{C}^{n+1,d}$ as the matrices with rows $\{|u_i\rangle\}_{i=0}^n \subset \mathbb{C}^d$ and $\{|\tilde{u}_i\rangle\}_{i=0}^n \subset \mathbb{C}^d$, respectively. The Gram matrices $\tilde{X}$ and X^* are given by $X^* = AA^\dagger$ and $\tilde{X} = BB^\dagger$.

Let $A^\dagger = V\Sigma W^\dagger$ be a singular value decomposition (SVD) of $A^\dagger$, $V \in \mathbb{C}^{d,d}$ and $W \in \mathbb{C}^{n+1,n+1}$ being unitary matrices, and Σ a rectangular diagonal matrix containing the singular values of $A^\dagger$. By definition,

$$X^* = AA^\dagger = (A^\dagger)^\dagger A^\dagger = W\Sigma^\dagger\Sigma W^\dagger = W\Lambda W^\dagger,$$

where we have defined $\mathbb{C}^{n+1,n+1} \ni \Lambda \coloneqq \Sigma^\dagger\Sigma$. Denote the non-zero singular values of $A^\dagger$, ordered according to magnitude, by $\sigma_1 \geq \cdots \geq \sigma_r > 0$, where r is the rank of $A^\dagger$. Note that A and $A^\dagger$ have identical singular values. As acting with permutation matrices on V, $W^\dagger$ preserves unitarity, we can w.l.o.g assume

$$\Lambda = \begin{pmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_r^2 & \\ & & & 0_{n+1-r} \end{pmatrix},$$

where empty spaces indicate zero entries. Further, let

$$\Sigma_r = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \in \mathbb{C}^{r,r}$$

and partition $W \in \mathbb{C}^{n+1,n+1}$ like $\begin{pmatrix} W_1 & W_2 \end{pmatrix}$, with $W_1 \in \mathbb{C}^{n+1,r}$ and $W_2 \in \mathbb{C}^{n+1,n+1-r}$. Theorem 7.3.2 in [23] implies that there exists an isometry $V_1 \in \mathbb{C}^{d,r}$ such that

$$A^\dagger = V_1 \Sigma_r W_1^\dagger.$$

The first step towards proving Lemma 5.3 is to generalize the result of Theorem 7.3.2 in [23]. Concretely, we want to show that $B^\dagger$, obeying $(B^\dagger)^\dagger B^\dagger \approx X^* = W \Lambda W^\dagger$, can be approximately written like $B^\dagger \approx V_2 \Sigma_r W_1^\dagger$, where $V_2 \in \mathbb{C}^{d,r}$ is an isometry.

By definition,

$$(B^\dagger)^\dagger B^\dagger = \tilde{X} \overset{\epsilon}{\approx} X^* = W \Lambda W^\dagger.$$

Let

$$D \coloneqq \Sigma_r \oplus \mathbb{1}_{n+1-r} \in \mathbb{C}^{n+1,n+1}$$

be the $(n+1) \times (n+1)$ diagonal matrix with diagonal entries corresponding to the non-zero singular values of $A^\dagger$. The remaining $n+1-r$ diagonal entries are set to 1, such that D is invertible. Further, define

$$X \coloneqq B^\dagger W D^{-1} \in \mathbb{C}^{d,n+1}.$$

We partition X like $X = \begin{pmatrix} \tilde{V}_2 & Z \end{pmatrix}$, where $\tilde{V}_2 \in \mathbb{C}^{d,r}$. It follows that

$$\begin{aligned} X^\dagger X &= D^{-1} W^\dagger B B^\dagger W D^{-1} \\ &= D^{-1} W^\dagger A A^\dagger W D^{-1} &&+ D^{-1} W^\dagger (B B^\dagger - A A^\dagger) W D^{-1} \\ &= D^{-1} \Lambda D^{-1} &&+ D^{-1} W^\dagger (\tilde{X} - X^*) W D^{-1} \\ &= \begin{pmatrix} \mathbb{1}_r & \\ & 0_{n-r} \end{pmatrix} &&+ D^{-1} W^\dagger (\tilde{X} - X^*) W D^{-1}. \end{aligned}$$

Therefore,

$$X^\dagger X = \begin{pmatrix} \tilde{V}_2^\dagger \\ Z^\dagger \end{pmatrix} \begin{pmatrix} \tilde{V}_2 & Z \end{pmatrix} = \begin{pmatrix} \tilde{V}_2^\dagger \tilde{V}_2 & \tilde{V}_2^\dagger Z \\ Z^\dagger \tilde{V}_2 & Z^\dagger Z \end{pmatrix} = \begin{pmatrix} \mathbb{1}_r & \\ & 0_{n+1-r} \end{pmatrix} + D^{-1} W^\dagger (\tilde{X} - X^*) W D^{-1}.$$

Next, we aim to bound the Frobenius distance

$$\left\| \begin{pmatrix} \tilde{V}_2^\dagger \tilde{V}_2 & \tilde{V}_2^\dagger Z \\ Z^\dagger \tilde{V}_2 & Z^\dagger Z \end{pmatrix} - \begin{pmatrix} \mathbb{1}_r & \\ & 0_{n+1-r} \end{pmatrix} \right\|_F.$$

Unitaries leave the Frobenius norm invariant, as $\|A\|_F^2 = \sum_k \|Ae_k\|_2^2 = \sum_k \|UAe_k\|_2^2 = \|UA\|_F^2$, where U is a unitary. We find:

$$\left\| \begin{pmatrix} \tilde{V}_2^\dagger \tilde{V}_2 & \tilde{V}_2^\dagger Z \\ Z^\dagger \tilde{V}_2 & Z^\dagger Z \end{pmatrix} - \begin{pmatrix} \mathbb{1}_r & \\ & 0_{n-r} \end{pmatrix} \right\|_F = \|WD^{-1}W^\dagger(\tilde{X} - X^*)WD^{-1}W^\dagger\|_F \leq \epsilon \max\left\{ {\sigma_{\min}^{(n)}}^{-2}, 1 \right\},$$

where $\sigma_{\min}^{(n)}$ is the non-zero singular value of A with smallest magnitude, the square of which is the smallest non-zero eigenvalue of X^*, $\lambda_{\min}^{(n)}$. For the KCBS scenario, $n = 5$, we find $\lambda_{\min}^{(n)} = \frac{1}{2}(\sqrt{5} - 1) \approx 0.618$.

Therefore,

$$\|\tilde{V}_2^\dagger \tilde{V}_2 - \mathbb{1}_r\|_F \leq \epsilon \max\left\{ {\lambda_{\min}^{(n)}}^{-1}, 1 \right\} \quad \text{and} \tag{5.1}$$

$$\|Z\|_F \leq \epsilon \max\left\{ {\lambda_{\min}^{(n)}}^{-1}, 1 \right\} + \mathcal{O}\left(\epsilon^2 \max\left\{ {\lambda_{\min}^{(n)}}^{-1}, 1 \right\}^2 \right), \tag{5.2}$$

where 5.2 follows from 5.1, as we will show. For now, let us assume 5.2 to be true.

By definition, $X = B^\dagger W D^{-1}$, and hence

$$B^\dagger = XDW^\dagger = \begin{pmatrix} \tilde{V}_2 & Z \end{pmatrix} \begin{pmatrix} \Sigma_r & \\ & \mathbb{1}_{n-r} \end{pmatrix} \begin{pmatrix} W_1^\dagger \\ W_2^\dagger \end{pmatrix} = \begin{pmatrix} \tilde{V}_2 \Sigma_r & Z \end{pmatrix} \begin{pmatrix} W_1^\dagger \\ W_2^\dagger \end{pmatrix} = \tilde{V}_2 \Sigma_r W_1^\dagger + Z W_2^\dagger$$

Therefore,

$$\|B^\dagger - \tilde{V}_2 \Sigma_r W_1^\dagger\|_F = \|ZW_2^\dagger\|_F = \|Z\|_F \leq \epsilon \max\left\{ {\lambda_{\min}^{(n)}}^{-1}, 1 \right\} + \mathcal{O}\left(\epsilon^2 \max\left\{ {\lambda_{\min}^{(n)}}^{-1}, 1 \right\}^2 \right).$$

Crucially, $\tilde{V}_2$ is approximately isometric, which follows from the previously derived bound on $\|\tilde{V}_2^\dagger \tilde{V}_2 - \mathbb{1}_r\|_F$, as we will now prove:

Consider a SVD of $\tilde{V}_2$,

$$\tilde{V}_2 = S \tilde{\Sigma} T^\dagger,$$

and define V_2 as the isometry

$$V_2 := S \begin{pmatrix} \mathbb{1}_r \\ 0_{d-r,r} \end{pmatrix} T^\dagger.$$

V_2 is isometric because

$$V_2^\dagger V_2 = T \begin{pmatrix} \mathbb{1}_r & 0_{r,d-r} \end{pmatrix} \begin{pmatrix} \mathbb{1}_r \\ 0_{d-r,r} \end{pmatrix} T^\dagger = T \, \mathbb{1}_r \, T^\dagger = \mathbb{1}_r,$$

and its Frobenius distance to $\tilde{V}_2$ is

$$\|\tilde{V}_2 - V_2\|_F^2 = \left\| S \left[\tilde{\Sigma} - \begin{pmatrix} \mathbb{1}_r \\ 0_{d-r,r} \end{pmatrix} \right] T^\dagger \right\|_F^2 = \mathrm{tr} \left(\left[\tilde{\Sigma}^\dagger - \begin{pmatrix} \mathbb{1}_r & 0_{r,d-r} \end{pmatrix} \right] \left[\tilde{\Sigma} - \begin{pmatrix} \mathbb{1}_r \\ 0_{d-r,r} \end{pmatrix} \right] \right)$$

Let r' be the rank of $\tilde{V}_2$, and $\tilde{\sigma_1} \geq \ldots \tilde{\sigma}_{r'} > 0$ the non-zero singular values of $\tilde{V}_2$. Again, w.l.o.g.

$$\tilde{\Sigma}^\dagger = \begin{pmatrix} \tilde{\sigma}_1 & & & & \\ & \ddots & & & \\ & & \tilde{\sigma}_{r'} & & \\ & & & 0_{r-r'} & \\ & & 0_{d-r,r} & & \end{pmatrix}.$$

Hence, $\|\tilde{V}_2 - V_2\|_F$ can be re-written like

$$\|\tilde{V}_2 - V_2\|_F^2 = \mathrm{tr} \left(\left(\begin{pmatrix} \tilde{\sigma}_1^2 & & & \\ & \ddots & & \\ & & \tilde{\sigma}_{r'}^2 & \\ & & & 0_{r-r'} \end{pmatrix} - 2 \begin{pmatrix} \tilde{\sigma}_1 & & & \\ & \ddots & & \\ & & \tilde{\sigma}_{r'} & \\ & & & 0_{r-r'} \end{pmatrix} + \mathbb{1}_r \right)$$

$$= \sum_{i=1}^{r'} (\tilde{\sigma}_i^2 - 2\tilde{\sigma}_i + 1) + (r - r') = \sum_{i=1}^{r'} (\tilde{\sigma}_i - 1)^2 + (r - r').$$

The condition $\|\tilde{V}_2^\dagger \tilde{V}_2 - \mathbb{1}_r\|_F \leq \epsilon \max\left\{ \lambda_{\min}^{(n)\,-1}, 1 \right\}$ implies

$$\left\| T \begin{pmatrix} \tilde{\sigma}_1^2 & & & \\ & \ddots & & \\ & & \tilde{\sigma}_{r'}^2 & \\ & & & 0_{r-r'} \end{pmatrix} T^\dagger - \mathbb{1}_r \right\|_F^2 = \sum_{i=1}^{r'} (\tilde{\sigma}_i^4 - 2\tilde{\sigma}_i^2 + 1) + (r - r')$$

$$= \sum_{i=1}^{r'} (\tilde{\sigma}_i - 1)^2 (\tilde{\sigma}_i + 1)^2 + (r - r') \leq \epsilon^2 \max\left\{ \lambda_{\min}^{(n)\,-1}, 1 \right\}^2 ,$$

which in particular implies that, for $\epsilon < \min\left\{ \lambda_{\min}^{(n)}, 1 \right\}$, the rank of $\tilde{V}_2$ must match that of $A^\dagger$, i.e. $r = r'$. Additionally, $\tilde{V}_2$ is approximately isometric:

$$\|\tilde{V}_2 - V_2\|_F^2 = \sum_{i=1}^{r} (\tilde{\sigma}_i - 1)^2 \leq \epsilon^2 \max\left\{ \lambda_{\min}^{(n)\,-1}, 1 \right\}^2 .$$

Together, this gives us the desired intermediate result. Namely, up to leading order,

$$\|B^\dagger - V_2\Sigma_r W_1^\dagger\|_F \leq \|B^\dagger - \tilde{V}_2\Sigma_r W_1^\dagger\|_F + \|(\tilde{V}_2 - V_2)\Sigma_r W_1^\dagger\|_F \leq \epsilon \max\left\{{\lambda_{\min}^{(n)}}^{-1}, 1\right\}\left(1 + \sigma_{\max}^{(n)}\right),$$

where $\sigma_{\max}^{(n)}$ is the maximum singular value of $A^\dagger$. We take the second term to be negligible. For the KCBS scenario, $n = 5$, $\sigma_{\max}^{(n)} = \sqrt{2}$, and consequently $\|B^\dagger - V_2\Sigma_r W_1^\dagger\|_F \leq \kappa\epsilon$, where $\kappa \approx 3.55$. For $n > 7$, one finds numerically that $\max\left\{{\lambda_{\min}^{(n)}}^{-1}, 1\right\} = 1$. Additionally, $\lambda_{\max}^{(n)} = {\sigma_{\max}^{(n)}}^2$ scales linearly in n like $\lambda_{\max}^{(n)} \approx \frac{1}{4}n + 1$.

To recap, we have just shown that

$$A^\dagger = V_1\Sigma_r W_1^\dagger$$
$$B^\dagger \overset{\mathcal{O}(\kappa(n)\epsilon)}{\approx} V_2\Sigma_r W_1^\dagger,$$

where V_2 is an isometry. The next proof step is to show that there exists a unitary $U \in \mathbb{C}^{d,d}$ such that $B^\dagger \approx UA^\dagger$.

According to Theorem 2.1.18 in [23], there exists a unitary $U \in \mathbb{C}^{d,d}$ such that $V_2 = UV_1$. This unitary U also relates the matrices $B^\dagger$ and $A^\dagger$, as

$$B^\dagger \overset{\mathcal{O}(\kappa(n)\epsilon)}{\approx} V_2\Sigma_r W_1^\dagger = UV_1\Sigma_r W_1^\dagger = UA^\dagger.$$

Finally, we can use this result to bound the distance between the two Gram decompositions $\{u_i\}_{i=0}^n$ and $\{\tilde{u}_i\}_{i=0}^n$, and thereby conclude the proof:

$$\begin{aligned}\|B^\dagger - UA^\dagger\|_F &= \|\left(\tilde{u}_1 \mid \ \ldots \ \mid \tilde{u}_n\right) - U\left(u_1 \mid \ \ldots \ \mid u_n\right)\|_F \\ &= \|\left(\tilde{u}_1 - Uu_1 \mid \ \ldots \ \mid \tilde{u}_n - Uu_n\right)\|_F \leq \kappa(n)\epsilon + \mathcal{O}\left(\kappa(n)^2\epsilon^2\right),\end{aligned}$$

where vertical lines delimit columns. Since the Frobenius norm of an $n \times m$ matrix is just the two norm of the $(n \cdot m)$-long vector containing all matrix entries, we have

$$\forall i \in \{0, \ldots, n\} : \||\tilde{u}_i\rangle - U|u_i\rangle\|_2 \leq \kappa(n)\epsilon + \mathcal{O}\left(\kappa(n)^2\epsilon^2\right).$$

The only thing that remains is proving the inequality 5.2. By considering the off-diagonal entries of $X^\dagger X$, we find that

$$\epsilon\,{\lambda_{\min}^{(n)}}^{-1} \geq \|Z^\dagger\tilde{V}_2\|_F = \|Z^\dagger V_2 - Z^\dagger(V_2 - \tilde{V}_2)\|_F.$$

The reverse triangle inequality and submultiplicativity of the Frobenius norm yields

$$\epsilon\,{\lambda_{\min}^{(n)}}^{-1} \geq \|Z^\dagger V_2\|_F - \|Z^\dagger(V_2 - \tilde{V}_2)\|_F \geq \|Z\|_F\,(1 - \|V_2 - \tilde{V}_2\|_F),$$

which implies

$$\|Z\|_F \leq \epsilon\, \lambda_{\min}^{(n)\,-1} + \mathcal{O}\left(\epsilon^2 \lambda_{\min}^{(n)\,-2}\right) .$$

□

Note that Lemma 5.3 is of no use for Bell-type self-testing scenarios, as it relates two Gram decompositions only by a global unitary U.

Having distilled the main proof components into three auxiliary theorems, Lemmas 5.1-5.3, we will now demonstrate that these imply the desired self-testing result. [5] for the most part omits these final proof steps. Section 5.2 will provide an in-depth discussion about all of the assumptions that go into the self-testing protocol.

Theorem 5.4 (Odd n-cycle KS non-contextuality inequalities facilitate self-testing [5])**.**
Let $n \geq 5$ be an odd integer. The only quantum model $|\Psi\rangle$, $\{\Pi_i\}_{i=1}^n$ compatible with a maximal violation of the odd n-cycle KS non-contextuality inequality 2.5 with exclusivity graph $\mathcal{G}_{ex}^{(n)}$ is 2.7, up to a global isometry:

$$\exists \text{ isometry } V : |\Psi\rangle = V|u_0\rangle$$

$$\forall i \in \{1, \ldots, n\} : \Pi_i\, |\Psi\rangle = V|u_i^{(n)}\rangle\langle u_i^{(n)}|\, |u_0\rangle$$

Furthermore, all quantum models $|\Psi\rangle$, $\{\Pi_i\}_{i=1}^n$ compatible with an ϵ-suboptimal violation of 2.5, i.e.

$$\sum_{i=0}^{n} |\langle\Psi|\Pi_i|\Psi\rangle|^2 > B_q(\mathcal{G}_{ex}^{(n)}) - \epsilon\,,$$

where $B_q(\mathcal{G}_{ex}^{(n)})$ is the quantum supremum 2.6, are $\mathcal{O}(\epsilon)$ close in 2-norm to 2.7, up to a global isometry:

$$\exists \text{ isometry } V : \|\, |\Psi\rangle - V|u_0\rangle\|_2 \;\leq\; \mathcal{O}(\epsilon)$$

$$\forall i \in \{1, \ldots, n\} : \|\Pi_i\, |\Psi\rangle - V|u_i^{(n)}\rangle\langle u_i^{(n)}|\, |u_0\rangle\| \leq \mathcal{O}(\epsilon)$$

Here, the global isometry V expresses the fact that the statistics of a correlation experiment are always blind to embeddings into a higher-dimensional space and a global basis change. It is too optimistic to hope to determine these trivial degrees of freedom.

Proof. We denote the unique optimal solution of the Lovász SDP, assuming an underlying cyclic compatibility graph with n vertices, by X_n^*. One valid Gram decomposition of X_n^* is

$$X_n^* = \mathrm{Gram}(\, |\mathrm{u}_0\rangle\,,\, \langle \mathrm{u}_1^{(n)}|\mathrm{u}_0\rangle|\mathrm{u}_1^{(n)}\rangle\,,\, \ldots\,,\, \langle \mathrm{u}_n^{(n)}|\mathrm{u}_0\rangle|\mathrm{u}_n^{(n)}\rangle\,)\,,$$

with $\{|u_i^{(n)}\rangle\}_{i=0}^n$ the set of vectors defined in 2.7.

Say the odd n-cycle correlation experiment produces an ϵ-suboptimal violation of the KS non-contextuality inequality 2.5. Any quantum model $|\Psi\rangle$, $\{\Pi_i\}_{i=1}^n$ consistent with these statistics can be translated into a Gram matrix $\tilde{X}$ that satisfies the assumptions of Lemma 5.2:

$$\tilde{X}_n = \text{Gram}(\, |\tilde{u}_0\rangle \,,\, \langle \tilde{u}_1|\tilde{u}_0\rangle|\tilde{u}_1\rangle \,,\, \ldots \,,\, \langle \tilde{u}_n|\tilde{u}_0\rangle|\tilde{u}_n\rangle \,)\,,$$

where

$$\begin{aligned} |\tilde{u}_0\rangle &\coloneqq |\Psi\rangle \\ |\tilde{u}_i\rangle &\coloneqq \frac{\Pi_i|\Psi\rangle}{\sqrt{\langle\Psi|\Pi_i|\Psi\rangle}}\,. \end{aligned}$$

If ϵ is sufficiently small, and the quantum model violates the KS non-contextuality inequality 2.5, then the underlying Hilbert space $\mathcal{H}$ must be at least three-dimensional, since there exists an explicit KS non-contextual hidden variable model for a qubit [31]. Let $d \geq 3$ denote the dimension of $\mathcal{H}$. If $\mathcal{J}$ is an inner product isomorphism $\mathcal{H} \mapsto \mathbb{C}^d$, define $|\tilde{\mathbf{u}}_i^{(n)}\rangle \coloneqq \mathcal{J}|\tilde{u}_i^{(n)}\rangle$. For the vectors $\{|u_i^{(n)}\rangle\}_{i=0}^n \subset \mathbb{C}^3$, define $\{|\mathbf{u}_i^{(n)}\rangle\}_{i=0}^n$ to be their embedding into the larger space $\mathbb{C}^d$. As such, the vectors $|u_i^{(n)}\rangle$ and $|\mathbf{u}_i^{(n)}\rangle$ are related by the isometric embedding

$$\begin{aligned} V' : \mathbb{C}^3 &\to \mathbb{C}^d \\ e_i^{\mathbb{C}^3} &\mapsto e_i^{\mathbb{C}^d}, \end{aligned}$$

where $\{e_i^{\mathbb{C}^d}\}$ is the standard basis of $\mathbb{C}^d$.

Lemmas 5.2 and 5.3 imply that

$$\begin{aligned} \forall\, i \in \{0,\ldots,n\},\ \exists \text{ unitary } U \in \mathbb{C}^{d,d} : \|\langle\tilde{\mathbf{u}}_i^{(n)}|\tilde{\mathbf{u}}_0^{(n)}\rangle|\tilde{\mathbf{u}}_i^{(n)}\rangle - \langle\mathbf{u}_i^{(n)}|\mathbf{u}_0^{(n)}\rangle\, U|\mathbf{u}_i^{(n)}\rangle\|_2 \\ = \|\langle\tilde{u}_i^{(n)}|\tilde{u}_0^{(n)}\rangle|\tilde{u}_i^{(n)}\rangle - \langle u_i^{(n)}|u_0\rangle\, \mathcal{J}^{-1}UV'|u_i^{(n)}\rangle\|_2 \leq \mathcal{O}(\epsilon) \end{aligned} \tag{5.3}$$

For convenience, we define $V \coloneqq \mathcal{J}^{-1}UV'$. V is an isometry because the inner product isomorphism $\mathcal{J}$ is unitary.

We want 5.3 to hold true for the normalized vectors $|\tilde{u}_i^{(n)}\rangle$, $|u_i^{(n)}\rangle$, without pre-factors, as this would immediately imply Theorem 5.4. This is a direct consequence of the following, final Lemma. □

Lemma 5.5 ([5])**.** *Let* $\| \,|a\rangle - |b\rangle\, \|_2 \leq \delta$ *for vectors* $|a\rangle$ *and* $|b\rangle$*, such that* $\| \,|a\rangle\, \|_2 \geq 2\delta$*. Let* $|\hat{a}\rangle$ *and* $|\hat{b}\rangle$ *be* $|a\rangle$ *and* $|b\rangle$*, normalized to have unit norm. Then, we have that* $\||\hat{a}\rangle - |\hat{b}\rangle\|_2 \leq \frac{2\delta}{\||a\rangle\|_2}$.

The proof is found in [5] (Lemma 9). As $|\langle u_i^{(n)}|u_0\rangle| = 5^{\frac{1}{4}}$, the assumption in Lemma 5.5 is true for sufficiently small ϵ.

Having proved that the only quantum model $|\Psi\rangle$, Π_i compatible with a maximal violation of 2.5 is 2.7, up to a global isometry, and that this self-testing result is robust, we will now analyze what assumptions the protocol in [5] requires for practical purposes.

5.1.1 Proportionality constant in Lemma 5.2

In order to get a feeling for the constant of proportionality in Lemma 5.2, we perform a short numerical analysis for the simplest case, $n = 5$. An analytic approach would most likely involve deriving a closed-form expression for the Hölderian error bounds used in [5].

Our strategy to obtain an estimate for the constant of proportionality is to generate a large number of random matrices that are both in the feasible set of the Lovász SDP 3.1 and ϵ-suboptimal. The sketch in Figure 5.1 helps to visualize the set of matrices from which we sample, although it does not do justice to the large affine dimension of the feasible set.

We generate $\sim 10^4$ feasible, ϵ-suboptimal matrices $\{X_i^\epsilon\}_i$ and compute the maximum ratio

$$\nu_\epsilon(n=5) \coloneqq \max_i \frac{\|X_i^\epsilon - X^*\|_F}{\sum_i X_{ii}^* - X_{ii}},$$

where X^* denotes the optimal solution of the SDP 3.1 for $n = 5$. Furthermore, we perform the analysis for different values of ϵ, which range over many orders of magnitude, since we aim to find a universal constant that upper-bounds the ratio between deviation in Frobenius norm and suboptimality. Our estimate for the universal constant is then

$$\nu(n=5) \coloneqq \max_\epsilon \nu_\epsilon(n=5).$$

The set of Hermitian matrices on $\mathbb{C}^{n+1}$, $\mathrm{Herm}(\mathbb{C}^{n+1})$, is a $(n+1)^2$-dimensional real vector space. Define the (n^2-2n)-dimensional subspace

$$\bar{\mathcal{A}} \coloneqq \{X \in \mathrm{Herm}(\mathbb{C}^{n+1}) \,|\, X_{ii} = X_{0i} \text{ for } 1 \le i \le n, \text{ and } X_{ij} = 0,\ i \sim j\} \subset \mathrm{Herm}(\mathbb{C}^{n+1}).$$

The affine space

$$\mathcal{A} \coloneqq \bar{\mathcal{A}} + e_{00},$$

is obtained by translating the subspace $\bar{\mathcal{A}}$ along the vector e_{00}, where e_{ij} is the matrix that has a 1 at (i,j), and 0 elsewhere. The feasible set of the Lovász SDP 3.1 is then $\mathcal{A} \cap S_+^{n+1}$.

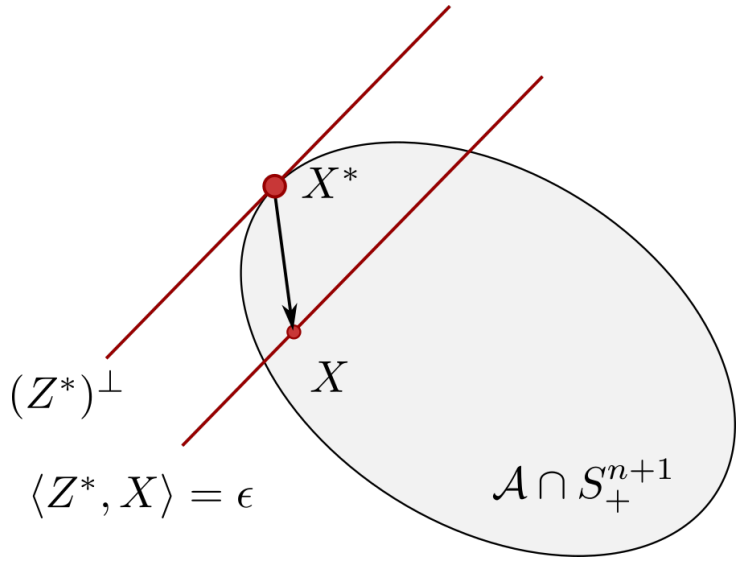


Figure 5.1: The feasible set of the Lovász SDP 3.1 is a bounded convex set. It is given by the intersection of the set of positive semi-definite matrices S_+^{n+1} and the affine space $\mathcal{A}$, as defined in the the text. Z^* is the optimal solution of the SDP that is dual to 3.1. The primal optimal solution of 3.1 is the unique point of intersection of the feasible set $\mathcal{A} \cap S_+^{n+1}$ and the subspace $(Z^*)^\perp$ [5]. As such, the subspace $(Z^*)^\perp$ is a supporting hyperplane of the feasible set, which lies entirely on one side of this hyperplane. Furthermore, the set of ϵ-suboptimal, feasible matrices is given by the intersection of $\mathcal{A} \cap S_+^{n+1}$ and $\{X \in \mathrm{Herm}(\mathbb{C}^{n+1}) \,|\, \langle Z^*, X\rangle = \epsilon\}$ [5].

To generate appropriate random matrices, we note that the vector connecting X^* and any X_i must be in $\bar{\mathcal{A}}$. An ONB of $\bar{\mathcal{A}}$, with respect to the inner product $\langle X, Y\rangle = \mathrm{tr}(X^\dagger Y)$, is given by the $(n^2 - 2n)$ matrices

$$
\begin{aligned}
D_i &:= \frac{e_{ii} + e_{0i} + e_{i0}}{\sqrt{3}} \quad \text{, for } 1 \leq i \leq n, \\
T_{ij} &:= \frac{e_{ij} + e_{ji}}{\sqrt{2}} \quad \text{, for } i \not\sim j, i > j, \\
R_{ij} &:= i\frac{e_{ij} - e_{ji}}{\sqrt{2}} \quad \text{, for } i \not\sim j, i > j.
\end{aligned}
$$

By sampling $(n^2 - 2n)$ real parameters from a Gaussian distribution, we can generate random, normalized shifts that are uniformly distributed on the $(n^2 - 2n)$-dimensional unit sphere. To get an ϵ-suboptimal matrix, we simply multiply this shift by ϵ. To check whether the shift preserves positive semi-definiteness, we attempt to perform a Cholesky decomposition of the matrix $X^* + \epsilon V + 10^{-12}\mathbb{1}$. The regularizing term $10^{-12}\mathbb{1}$ accounts for the fact that the Cholesky decomposition can only be computed for positive definite matrices, whereas we want to check for positive semi-definiteness.

Figure 5.2 captures the results of the numerical analysis for 5 independent trials. We estimate $\nu(n = 5) \sim 1.5$.

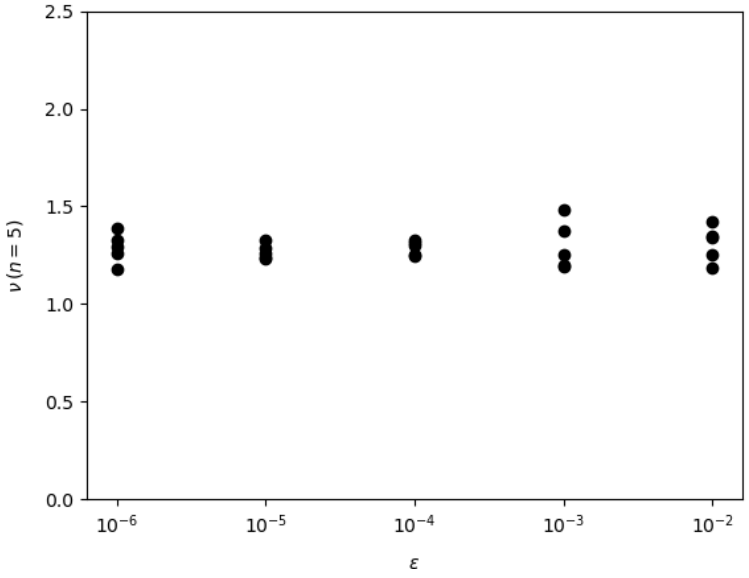


Figure 5.2: Numerical analysis of the proportionality constant $\nu(n)$ in Lemma 5.2, for $n = 5$. These values are computed by randomly generating $\sim 10^4$ feasible, ϵ-suboptimal matrices, where ϵ ranges over several orders of magnitude. Plotted is the maximal ratio between deviation in Frobenius norm and suboptimality, $\nu_\epsilon(n = 5)$, for 5 independent trials. Each black dot corresponds to $\sim 10^4$ feasible, ϵ-suboptimal matrices. We estimate the constant of proportionality to be $\max_\epsilon \nu_\epsilon(n = 5) \sim 1.5$.

5.2 Assumptions

The following is an as comprehensive as possible list of explicit or implicit assumptions that are required to self-test the reference quantum experiment 2.7, on the basis of Theorem 5.4:

1. i.i.d. rounds $x, y \,|\, i, i \oplus 1$
2. uncorrelated preparation and measurement devices,
   freely random choice of measurement context $i, i \oplus 1$
3. sharp (outcome-repeatable and minimally disturbing) measurements
4. cyclic compatibility
5. to certify quantumness:
   (a) sequential measurements
      - system not disturbed between measurements

   (b) information carrying capacity of the unknown system does not exceed "simulable" bound (each measurement device is used only once)

Let us go through Assumptions 1-5 in order:

Assumptions 1, namely that the devices behave in the same way for every input-output round of the correlation experiment, is made in a large majority of current self-testing works [44]. It allows us to lift relative frequencies observed in a correlation experiment to well-defined probability distributions $p(x, y \,|\, i, i \oplus 1)$ that are valid for all rounds of the protocol, a round being characterized by one input-output cycle for a given measurement context.

For a second, assume that Assumption 2 is not true, and consider the preparation and measurement devices to be correlated. We express this in terms of some shared classical randomness λ. As noted in [32], such shared randomness can render the task of self-testing futile. Consider a quantum experiment where, based on the value of λ, a device prepares $\rho_\lambda = U_\lambda \rho U_\lambda^\dagger$, with U_λ unitary, and the measurement event $0, 1 \,|\, i, i \oplus 1$ is characterized by a positive semi-definite operator $F_\lambda^{i,i\oplus 1} = U_\lambda F^{i,i\oplus 1} U_\lambda^\dagger$. The probability $p(0, 1 \,|\, i, i \oplus 1) = \int_\Lambda d\lambda \, \mu(\lambda) \operatorname{tr}(\rho_\lambda F_\lambda^{i,i\oplus 1}) = \operatorname{tr}(\rho F^{i,i\oplus 1})$. This correlation experiment with shared randomness is therefore operationally indistinguishable from one without, and we cannot hope to relate the physical experiment to the reference experiment via a single unitary operator, based on observed correlations alone.

The necessity of Assumption 3 for the soundness of the self-testing protocol was already highlighted in Sections 3.4 and 3.5. For unsharp quantum measurements, the class of odd n-cycle inequalities can be super-maximally violated by trivial POVM. Therefore, the CSW hierarchy falls apart and the Lovász SDP no longer characterizes optimal quantum violations of KS non-contextuality inequalities. Assumption 3 was critical in the above approach. Nevertheless, we would like to relax it - in a realistic setting, measurements will never be truly projective.

Assumption 4 is required by the CSW framework. If the measurements $i, i \oplus 1$ were not compatible, the measurement events in Figure 3.1 could not be related to well-defined projectors Π_i, even for sequential measurements, as the product of two non-commuting projectors is not necessarily a projector.

As Assumption 5 is extensive, we devote an entire section to it. We assume sequential measurements, as they allow us to bound classical simulations reproducing the statistics in a non-trivial manner, something we will examine in Section 5.3. We assume that the system is not disturbed in between sequential measurements, so that the measurement sequence $i, i \oplus 1$ corresponds to a well-defined projector like before.

5.3 Sequential measurements and the memory assumption

As described in Section 4, self-testing allows us to make powerful inferences about the quantum properties of a system. For applications of self-testing, the prime example being quantum key distribution (QKD), it is of great importance that the certificate not only encompasses the quantum state and measurements consistent with the maximal violation, but that the optimal input-output correlations also bear witness to the quantumness of the system. This means that we must exclude the possibility of the correlations being simulated by some classical mechanism,

as this would for instance allow the copying of information and render the protocol information-theoretically insecure. To do so, we wish to study the resources a classical device requires, in order to simulate the ideal reference quantum realization. The odd n-cycle contextuality scenarios will be of particular interest to us. In the case of Bell inequalities, the relevant resources have been identified and studied extensively [6, 45]. As highlighted in Section 2.8, assuming local causality and freely chosen detector settings, correlations that violate a Bell inequality are incompatible with a local hidden variable model description. Thus, classically simulating correlations that exhibit Bell non-locality requires superluminal communication ("communication cost") [45] and would violate special relativity. For self-testing via Bell inequalities this means that, assuming any adversary is bounded by the no-signalling principle, the violation of a Bell inequality certifies quantumness, provided that both subsystems are indeed space-like separated. Analogously to Bell non-locality and the no-signalling principle, we wish to identify a physical principle that enables us to differentiate between a quantum system exhibiting KS non-contextuality and a classical mechanism simulating correlations that violate a KS non-contextuality inequality. One strength of contextuality as a competing notion of nonclassicality is that contextuality experiments comprise a much broader class of experiments. In particular, generic contextuality experiments don't presuppose multiple space-like separated systems, but can consist of sequences of state preparations and measurements of a single system. However, this generality also complicates the analysis of the relevant resources required for simulating KS contextuality. Moreover, we can in principle never distinguish between a single-system quantum experiment exhibiting some input-output correlations and a classical pre-programmed computer simulating these [44]. For Bell-type scenarios this possibility is circumvented by extending the experiment to multiple space-like separated subsystems. In Section 5.3 we examine the "memory cost" of KS contextuality [14, 24]: As we will see, the minimal internal entropy of a classical system emulating KS contextuality in many cases exceeds the information-carrying capacity of the quantum system it aims to emulate. To weed out the possibility of a classical mechanism producing the correlations, in [5] it is assumed that "[...] the measured system has no more memory than its information carrying capacity". However, for self-testing scenarios the information carrying capacity of the physical device is unknown. One therefore has to justify imposing an upper bound on the unknown device's information carrying capacity to conclude quantumness. As we will see, these upper bounds have to be of the order of a few bits, which means that the required memory assumption is a very strong one indeed.

In Section 7, where we propose a self-testing protocol based on Spekkens contextuality, the cyclic compatibility and measurement sharpness assumptions will be replaced by assumptions about (approximate) operational equivalences of certain experimental procedures. The protocol will also assume an upper bound on the information carrying capacity of the unknown system. We will discuss its features in Section 7.3.

Alternatively, Cabello proposes using Landauer's principle [14] to translate the memory cost of KS contextuality into a "thermodynamical cost" [12, 47], in terms of heat the system dissipates when measured. We briefly sketch the general idea, however, as the minimal unit of Landauer

heat a classical simulation must dissipate per measurement will be of the order $\sim k_B T$, this proposal is hardly viable in a realistic experimental setting: For sufficiently long measurement sequences, a classical simulation with finite memory must "forget" information about the past input-output sequence. It is as such a logically irreversible computation [47]. Landauer's principle states that whenever classical information is processed in a logically irreversible manner, for example when erasing a classical bit of information, the entropy of the environment increases by a corresponding amount. For the erasure (setting to zero) of a classical bit this entropy increase is at least $\Delta S_{env} \geq -\Delta S_{sys} = -(0 - k_B \log(2)) = k_B \log(2)$. This entropy increase is related to the heat dissipated to the environment reservoir via $\Delta Q = (\Delta S)\, T$, meaning that the minimal amount of dissipated heat for the erasure of a single classical bit is given by $\Delta Q = k_B T \log(2)$.

5.3.1 Setting and preliminary considerations

We consider correlations experiments that consist of a series of sequential measurements for different measurement contexts $i, i \oplus 1$. Within quantum theory it does not matter whether we consider joint or sequential measurements of compatible observables, as both produce identical correlations. To that effect we will assume that our system is not perturbed between successive measurements. If one measurement is performed at time t and another at $t' > t$, then we want the system's state to remain unchanged in the time interval (t, t').

Let us now examine how, in an adversarial setting where correlations may not capture the outcome statistics of compatible and ideal quantum measurements, a classical device with an internal memory can simulate the input-output process characterized by the correlation experiment.

5.3.2 Input-output processes as information transducers

In most general terms, a KS contextuality experiment is an input-output process. To verify that an experiment producing a maximal violation of say the KCBS KS non-contextuality inequality is compatible with the ideal reference experiment, as mandated by self-testing, we may in principle perform an infinite sequence of measurements $\overleftrightarrow{X} = \{\ldots, X_{-1}, X_0, X_1, \ldots\}$, yielding an infinite sequence of outputs $\overleftrightarrow{Y} = \{\ldots, Y_{-1}, Y_0, Y_1, \ldots\}$ [14]. Here, $\overleftrightarrow{X}$ and $\overleftrightarrow{Y}$ are bi-infinite sequences of random variables and define stochastic processes. The random variable X_i represents the freely chosen random measurement at time $t = i \in \mathbb{Z}$, whereas the random variable Y_i represents the output obtained when performing the measurement X_i at time i. We assume that all X_i and Y_i take values from finite alphabets $\mathcal{X}$ and $\mathcal{Y}$; most of the measurements we will consider are dichotomic, i.e. $|\mathcal{Y}| = 2$, as is the case for the KCBS KS contextuality scenario. Furthermore, we can distinguish the present $t = 0$ and partition the processes $\overleftrightarrow{X}$, $\overleftrightarrow{Y}$ into past and future: $\overleftarrow{X} \coloneqq \ldots, X_{t-2}, X_{t-1}$, $\overrightarrow{X} \coloneqq X_t, X_{t+1}, \ldots$. The reference time is in fact arbitrary, as the stochastic processes we will consider turn out to be stationary. Any classical

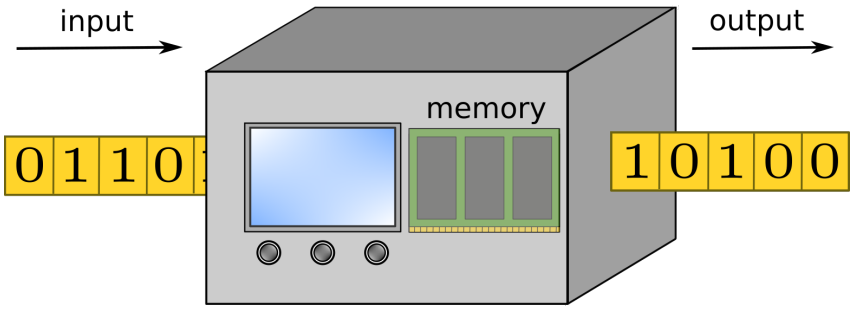


Figure 5.3: A classical device simulating the ideal reference statistics of some KS-type experiment, consisting of sequential measurements. The device is fed a sequence of inputs (measurement choices) and produces outputs (measurement outcomes), based on its internal state prior to an input. The internal state of the device, which one can equate with a memory, stores relevant historical information about the past input-output process, on the basis of which the device generates the correct input-output statistics. After one input-output cycle, the device can update its memory state.

input-output mechanism producing a maximal violation that we cannot "unmask" on the basis of correlations alone, must be consistent with the ideal reference input-output process.

Any classical apparatus simulating the reference correlations has the form of a general information transducer, transforming an input stochastic process into an output stochastic process [14]. A classical simulation is at all times i in some well-defined internal state S_i that holds information about the past input-output sequence $\overleftarrow{Z} := (\overleftarrow{X}, \overleftarrow{Y})$, stored inside a memory. As the future process may be correlated with the past, the internal memory of the device allows it to produce statistically correct predictions. If we pass a valid input symbol from $\mathcal{X}$ (think of $\mathcal{X}$ as the n measurement settings a verifier can choose from for the odd n-cycle scenario) to the transducer it will generate an output symbol in $\mathcal{Y}$, based on the information about the past encoded in the internal state prior to the input. We assume that the choice of measurement setting is freely random and not correlated with the internal state of the machine. The output may be accompanied by an update of the transducer's internal memory. One can think of this input-output process as feeding the transducer an infinite tape, whereby each tape cell contains a valid input symbol (choice of measurement). The transducer then generates an output, overwrites the corresponding tape cell with said output, and updates its internal state, see Figure 5.3. Note that we will only consider the RAM memory cost of such a simulation, which relates to the number of memory bits the device needs to store the relevant historical information about the past process[1], and in particular not to the memory required to store say output and transition probabilities.

Definition 5.6 ([2])**.** An *information transducer* is a tuple $(\mathcal{X}, \mathcal{Y}, \mathcal{S}, \mathcal{T})$ where

- $\mathcal{X}$, $\mathcal{Y}$ are the transducer's (finite) input and output alphabets,
- $\mathcal{S}$ is a (finite or countably infinite) set of internal memory states, and

[1] alternatively, the RAM memory cost can be regarded as the uncertainty (in bits) in the internal state of the transducer.

- $\mathcal{T} \equiv \{p(S_{i+1} = s', Y_i = y | S_i = s, X_i = x)\}_{x,y,s,s'}$ is the set of conditional ouput-transition probabilities.

More concretely, $p(S_{i+1} = s', Y_i = y | S_i = s, X_i = x)$ denotes the probability of the transducer outputting y and transitioning to the internal state s' after the ith input-output cycle, conditioned on the input x and internal state s.

We require a valid classical model to produce statistics that are consistent with the infinite family of stochastic processes describing the quantum probabilities of the ideal reference experiment: $\{P(\overleftrightarrow{Y} \mid \overleftrightarrow{x})\}_{\overleftrightarrow{x}}$, $\overleftrightarrow{x}$ being an instantiation of the input sequence $\overleftrightarrow{X}$.

Input-output transducers have been extensively studied within the field of computational mechanics. Of interest to us are memory-optimal classical models that produce the correct statistics $\{P(\overleftrightarrow{Y} \mid \overleftrightarrow{x})\}_{\overleftrightarrow{x}}$. A transducer is memory-optimal if it stores the minimal amount of information about the past (in terms of number of bits) to make statistically correct future predictions. It has been shown that for stationary input-output processes there exists a unique memory-optimal transducer producing the correct statistics: the process' ϵ-transducer [2]. The memory-optimal transducer does not store all information about the past $\overleftarrow{Z}$, as this may be highly redundant. In particular, the memory-optimal transducer does not differentiate between pasts that predict statistically identical futures. As such, the internal states of an ϵ-transducer are given by the causal states of the process.

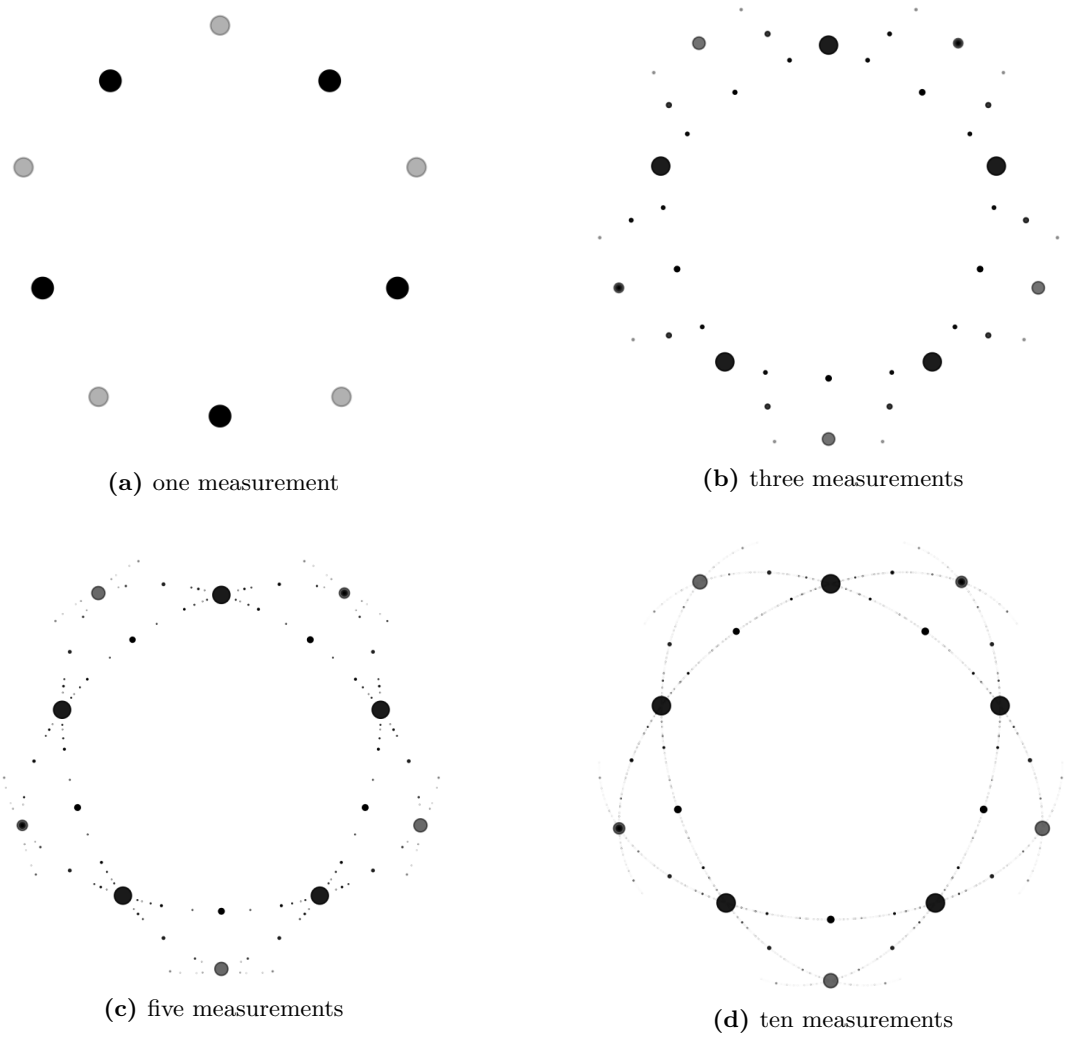
Definition 5.7. Let $\overleftarrow{X}$, $\overrightarrow{X}$ denote the past and future stochastic input processes, respectively. The *causal states* $[\overleftarrow{z}]$ of an input-output process $\overleftrightarrow{Y} \mid \overleftrightarrow{X}$ are the equivalence classes on the set of pasts $\overleftarrow{z}$ with respect to the equivalence relation

$$\overleftarrow{z} \sim_\epsilon \overleftarrow{z}' :\Leftrightarrow \; P(\overrightarrow{Y} \mid \overrightarrow{X}, \overleftarrow{Z}{=}\overleftarrow{z}) = P(\overrightarrow{Y} \mid \overrightarrow{X}, \overleftarrow{Z}{=}\overleftarrow{z}').$$

5.3.3 Memory-optimal classical simulation of quantum correlations

In [14] it is shown that, for a "causally complete" KS contextuality reference experiment, the causal states of the process' ϵ-transducer correspond one-to-one to the possible quantum states $\{|\Phi_{\overleftarrow{z}}\rangle\}_{\overleftarrow{z}}$ the system can occupy after all past measurements $\overleftarrow{x}$. The assumption of causal completeness holds for many KS contextuality scenarios and in the following we will examine the odd n-cycle KS contextuality scenarios, in particular the KCBS scenario for n = 5. In [14] the same analysis is performed for the Peres-Mermin (see Section 2.7) and Yu-Oh (see Section 2.9) sets of measurements.

For all ideal odd n-cycle reference experiments, the number of possible post-measurement states tends to infinity when we increase the number of measurements in our sequence. However, the post-measurement states are not all equally probable. Figure 5.4 illustrates this behaviour for the KCBS scenario n = 5.



(a) one measurement

(b) three measurements

(c) five measurements

(d) ten measurements

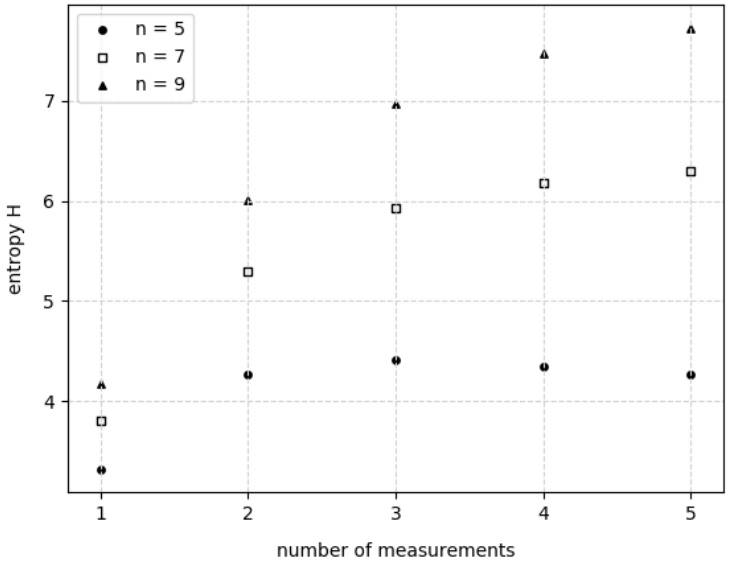
Figure 5.4: Possible post-measurement quantum states after one, three, five, and ten measurements for the ideal KCBS experiment. The components of the post-measurement quantum states w.r.t. the standard qutrit basis are all real. The phases (signs) of the real qutrit states are chosen such that the corresponding vectors lie on a joint semisphere of $\mathbb{R}^3$. What is plotted are centered views of the three-dimensional plots along the negative x direction. As can be seen, the number of possible post-measurement states increases with the number of sequential measurements. The volume of a scattered point is proportional to the probability of the system being in that state after the measurements. The probability distribution over possible post-measurement quantum states becomes approximately stationary for large measurement sequences. The fact that all states lie on one of five semi-circles is a consequence of the compatibility structure of the five KCBS measurements. Analogous plots are obtained in [14] for the Peres Mermin and Yu-Oh KS contextuality scenarios.

Problematically, the process describing the ideal quantum experiment is not stationary, i.e. not invariant under translations in time, as can be seen from Figure 5.4. However, if we "truncate" the process and consider only input-output cycles after sufficiently many initial measurements, the probability distribution over quantum states and by extension the stochastic process describing the experiment becomes stationary for all practical purposes. We only have a notion of the memory-optimal statistics-emulating classical model for the case of stationary processes. For this reason we will always consider the truncated input-output process.

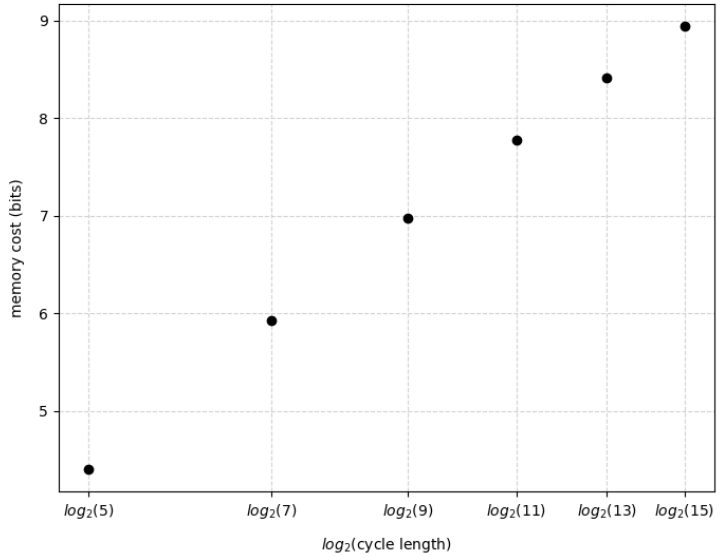
To numerically compute the memory cost, we consider all possible measurement sequences of a given length k. The statistical complexity of the process (RAM memory cost) is then the Shannon entropy $H = -\sum_i p_i \log_2 p_i$ of the probability distribution over all achievable post-measurement quantum states after k measurements, in the limit of $k \to \infty$ [2, 14]. Recall that this probability distribution becomes approximately stationary, reflected by the above expression converging. We find the number of bits (RAM) needed by a memory-optimal simulation of the ideal reference KCBS experiment to be just over 4. The statistical complexity thus exceeds the information carrying capacity $\mathcal{C} = \log_2(3)$ of the reference qutrit. In [14] the same qualitative behaviour is observed for the Peres-Mermin and Yu-Oh sets of measurements, both of which exhibit state-independent KS contextuality.

An interesting feature we observe is that the memory cost of simulating the ideal reference experiment for odd n-cycle contextuality scenarios with n $\geq$ 5 odd increases with the cycle length n, as is shown in Figure 5.5. However, this increase is only logarithmic. Furthermore, what does not change is the fact that, no matter the cycle length, we need to assume an upper bound on the unknown device's information carrying capacity in order to conclude quantumness. Additionally, we need to assume that each measurement device is used only once, requiring a potentially infinite supply, as otherwise the measurement device, and not the system whose information carrying capacity we bound, could retain memory about the past stochastic process.

Note that in Figure 5.5 only the asymptotic values of H are physically relevant, characterizing the stationary regime. We have in fact adapted the protocol in [5], by allowing the verifier to choose between n measurements for every input. In particular, the verifier is not limited to just two sequential measurements within one context $i, i \oplus 1$, but can perform any measurement sequence in $\{1, \ldots, n\}^m$, where m is the length of the sequence. This comes with an increased measurement complexity, and the verifier must check for compatibility with the ideal reference experiment for n^m measurement sequences, by comparing $n^m(2^m - 1)$ correlations to the expected statistics. The measurement sequences have to be sufficiently long to approximate the stationary regime. In fact, for just two sequential measurements $i, i \oplus 1$, as proposed in [5], one can find a classical model that reproduces the correlations of the ideal reference experiment, and requires less than one bit of memory. An information transducer with these properties is given in Figure 5.6. As such, the simulable bound on the information carrying capacity no longer exceeds the information carrying capacity of a qutrit. This highlights the need for an increased measurement complexity.



(a) Plotted is the Shannon entropy of the probability distribution over all achievable post-measurement quantum states after a given number of measurements for the 5, 7, and 9-cycle KS contextuality scenarios. The probability distribution over possible post-measurement quantum states becomes approximately stationary for a sufficiently large number of measurements and the entropies thus saturate. This stationary limit of H is the process' RAM memory cost, and it increases with an increasing cycle length



(b) The RAM memory cost of classically simulating the ideal reference experiment for n-cycle KS contextuality scenarios increases logarithmically with the number of ideal measurements obeying cyclic compatibility that have to be implemented.

Figure 5.5: The RAM memory cost of classically simulating the ideal reference experiment for odd n-cycle KS contextuality scenarios in terms of the cycle length n.

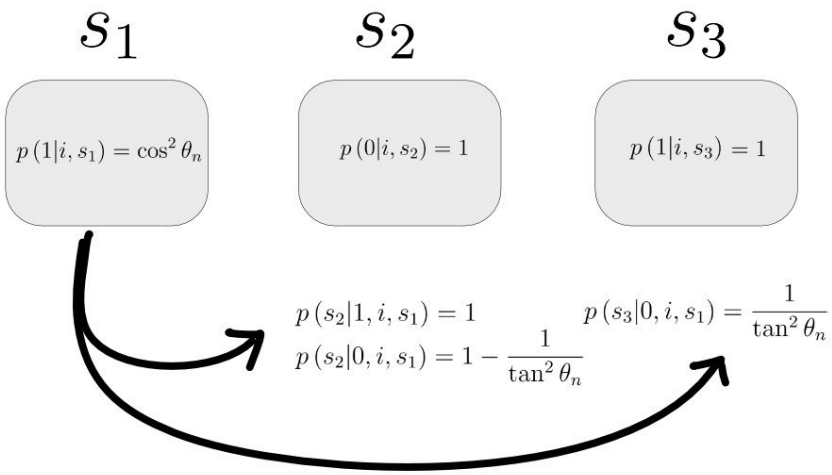


Figure 5.6: Simple classical information transducer with three internal memory states, s_1, s_2, s_3, which reproduces all correlations of the ideal odd n-cycle experiment, provided that the experimenter performs only two sequential measurements $i, i \oplus 1$. The rectangles with grey border give the outcome probabilities, conditioned on the respective memory state s_j and measurement i. The arrows indicate transitions in the internal state, and are labeled by the respective conditional transition probabilities. Prior to each input-output cycle $i, i \oplus 1$, the transducer's memory is reset to the state s_1. The memory cost of this simulation, i.e. the maximal uncertainty in the internal memory state of the machine, is less than one bit (after one measurement, the system is in one of two states with unequal probability), and in particular does not exceed the qutrit information carrying capacity. This simple example is meant to motivate an increased measurement complexity, in order to establish a clear separation in terms of resources between the ideal qutrit and any classical simulation.

Chapter 6

A revised notion of non-classicality
Spekkens contextuality

6.1 Outcome determinism for unsharp measurements (ODUM)

We want a notion of classicality that accomodates unsharp measurements and is robust to noise. A first approach might entail simply extending KS non-contextuality to general POVM measurements, in the sense that a classical HVM description assigns a conditional probability $p \in \{0, 1\}$ to all positive semi-definite operators, such that for every resolution of the identity only one operator receives the valuation 1. This incremental approach has been given the name "ODUM", which stands for "outcome determinism for unsharp measurements". The problem with ODUM is that it runs into numerous inconsistencies, as convincingly argued in [42]. We confine ourselves to the most immediate inconsistency: Consider the quantum experiment consisting of an experimenter repeatedly performing the trivial POVM $\{\frac{\mathbb{1}}{2}, \frac{\mathbb{1}}{2}\}$ on some arbitrary preparation ρ. Any classical value assignment to positive semi-definite operators in accordance with ODUM is in conflict with the normalization conditions for general ontological models in Section 2.3. In particular, for a given ontic state λ, either the added probabilities of the two outcomes is zero or exceeds one. Thus, either this trivial POVM is non-classical according to the extended notion just presented, or the new notion is not meaningful. The trivial POVM corresponds to a fair coin flip - it can be implemented by simply throwing away the system and generating a random bit. Such a "measurement" should by all means be considered classical. For this implementation, the only sensible choice is to assign to every every ontic state λ the conditional probability $\frac{1}{2}$, as measurement and system are completely decoupled. The same line of reasoning leads to the conclusion that the approach in Section 3.3 cannot be extended to general unsharp measurements, as this produces inconsistencies when treating QM.

6.2 Operational approach due to Spekkens

The previous examples highlight the need for a new notion of contextuality that captures the spirit of KS non-contextuality, but is applicable to realistic experimental scenarios. Let us take a step back and identify common features of the contextuality scenarios discussed so far, which led to behaviour we deemed non-classical. The setting was in all cases for the most part identical: we considered sets of projective measurements, with the key property that a given projector was shared amongst multiple measurements. Identical projectors appearing in multiple measurement contexts were crucial for proving a contradiction. Spekkens' approach captures this key feature and operationalizes the underlying assumptions. In QM, a positive semi-definite operator defines an operational equivalence class, in the sense that measurement events which are operationally indistinguishable for all state preparations must be represented by the same operator. An analogous observation applies to density operators describing the preparation of a system. On that account, instead of assuming a projector to be part of several measurement contexts, Spekkens assumes certain measurement events to be operationally equivalent; we will formalize this shortly. Assuming KS non-contextuality, the implication of a projector appearing as part of multiple measurements is that it must be assigned the same conditional probability $\{0, 1\}$ by the ontological model, for all ontic states λ. In other words, operationally equivalence implies an identical ontological representation. We will find that Spekkens' revised notion of non-contextuality follows exactly this prescription.

In [40], Spekkens introduces a largely operational notion of contextuality for preparations, measurements, and transformations. It is a revision of KS contextuality that addresses several shortcomings of the traditional definition. Contextuality is generalized to apply to arbitrary operational theories. In comparison, KS contextuality is limited to the scope of QM, as it is defined within the framework of quantum theory. Furthermore, the revised definiton will not assume outcome determinism at an ontic level, accomodating HVM whose indicator functions are more general mappings $\Lambda \mapsto [0, 1]$. Taking QM as operational theory, Spekkens contextuality extends to arbitrary measurements, including physical, non-projective ones. In fact, within the Spekkens framework, measurements, preparations, and transformations can be considered black box devices, i.e. primitives of which one can compose an experiment. An operational theory makes predictions about these operational primitives, i.e. assigns probabilities to measurement outcomes that are consistent with experimental observations. In QM, for instance, outcome probabilities are given by the Born rule, where we posit that every preparation is operationally fully specified by a density operator and every measurement by a POVM. A Spekkens non-contextual operational theory is one that is compatible with a Spekkens non-contextual ontological model, which we will define shortly. Spekkens' revised concept of non-contextuality will serve as a new, more suitable notion of classicality that is in particular compatible with unsharp measurements.

This section only discusses Spekkens non-contextuality for preparations and measurements, as

transformations will not be relevant to our discussion[1]. Preparations and measurements will be represented as outlined in Section 2.3. In the following, if not stated otherwise, the term "contextuality" and all variant forms will refer to Spekkens contextuality.

The essence of what it means for an ontological model to be non-contextual can be summarized as follows: An ontological model of an operational theory is *non-contextual* if the operational equivalence of two experimental procedures, i.e. two preparations or two measurements, implies that they have an equivalent representation in the ontological model. Two experimental procedures are operationally equivalent if they cannot be distinguished by any measurement statistics.

Let us formalize the above, starting with preparation non-contextuality. We can define an equivalence relation on the set of all preparation procedures P that partions the set into operational equivalence classes $[P]$. By the guiding principle above, we require equivalent preparations to have the same ontological representation in a non-contextual HVM.

Definition 6.1 ([40])**.**
Two preparation procedures P, P' are *equivalent*, denoted $P \sim P'$, if

$$\forall \text{ measurements } M, \forall \text{ measurement outcomes } k \,:\, p(k|P, M) = p(k|P', M)\,.$$

An ontological model is *preparation non-contextual* if equivalent preparations have identical associated probability density functions on the ontic state space:

$$\forall \text{ preparations } P \,:\, \mu_P = \mu_{[P]}\,.$$

Measurement non-contextuality is defined in an analogous manner:

Definition 6.2 ([40])**.**
Two measurement procedures M, M' are *equivalent*, denoted $M \sim M'$, if their outcomes can be associated one-to-one and

$$\forall \text{ outcomes } k\,, \forall \text{ preparations } P \,:\, p(k|P, M) = p(k|P, M')\,.$$

An ontological model is *measurement non-contextual* if equivalent measurements have identical associated indicator functions:

$$\forall \text{ measurements } M\,, \forall \text{ outcomes } k \,:\, \xi_{M,k} = \xi_{[M],k}\,.$$

[1] Consider a prepare-and-measure type experiment that consists of an initial state preparation, followed by some transformation of the system, and finally a measurement. We may consider the transformation as part of the initial preparation procedure, or as part of the final measurement, and describe the experiment solely in terms of preparation procedures and measurements.

Spekkens proves three no-go theorems for ontological models of QM [40]. These rule out preparation non-contextual, measurement non-contextual, and transformation non-contextual models, respectively. All three proofs apply to two-dimensional Hilbert spaces, making them stronger than traditional proofs of contextuality.

Whenever referring to a non-contextual ontological model, without specifying preparation or measurement contextual, we will assume the ontological model to be both preparation and measurement non-contextual. As the reasons for assuming classical correlations to be compatible with a measurement non-contextual HVM description are the same as those for preparation non-contextuality, the combination of both is the only natural assumption of non-contextuality.

We now discuss to what extent non-contextuality is a sensible notion of classicality. A contextual ontological model implies a difference in reality that cannot be observed. Such a model is in conflict with Leibniz' "Identity of Indiscernibles" [8, 43]. Leibniz's principle states that two empirically indistinguishable scenarios are to be ontologically equivalent. As such, it rejects ontological models for which there exist ontologically distinct, but empirically indistinguishable scenarios. Recall that, within QM, KS non-contextuality can be seen as a generalization of Bell's notion of local determinism to non-remote measurement contexts. One can also link Leibniz' principle, which is at the heart of Spekkens' revised notion of contextuality to Bell's notion of local causality. Like before, local causality is the weaker assumption, which is owed to the fact that Bell-type experiments impose additional constraints on the experiment's causal structure:

In principle, it might well be the case that an agent's free choice of measurement can have a causal influence on a remote system and alter the ontic or "matter of fact" state of that system. However, physicists have yet to devise an experiment producing correlations that are in conflict with the no-signalling principle, which is reinforced by special relativity. Therefore, to the best of our knowledge, such causal influence, if it exists, is not detectable. This in turn means that an agent in the remote frame of reference can detect no difference in the physical properties of his system for different free choices of measurement. Local causality is the assumption that there is in fact no difference at the ontic level, and thus provides the most natural explanation for our inability to detect a causal influence. Another, more hair-raising explanation could be that all possible ensemble preparations accessible to an experimenter are too spread out to resolve minor differences in the conditional probabilities. The assumption of local causality just corresponds to Leibniz' principle applied to Bell-type experiments with space-like separated measurements. The assumption of Spekkens non-contextuality simply extends the applicability of Leibniz's principle to arbitrary prepare-and-measure type experiments, in particular those without remote subsystems.

Just as local causality is motivated by the no-signalling principle, local causality providing the most natural explanation for our apparent inability to communicate superluminally, a point can be made that non-contextuality is perhaps equally well motivated by Leibniz' Identity of indiscernibles. Analogously to local causality, non-contextuality is the most natural explanation

for operational equivalences. Spekkens points out that the credentials of Leibniz' principle parallel those of no-signalling and that it was critical in Einstein's conception of relativity [8, 43]. Take for instance the equivalence principle which states that one cannot distinguish between being at rest in a uniform gravitational field and accelerating uniformly through free space. Within Newtonian mechanics the two scenarios are ontologically distinct, as one's absolute acceleration differs in both cases. On his way to general relativity, Einstein reasoned that the empirical indistinguishablitiy of both scenarios implies that they should be treated as ontologically equivalent by a physical theory, a powerful invocation of Leibniz's principle that reshaped our perception of reality. An almost identical line of reasoning is found in Einstein's 1905 paper "On the electrodynamics of moving bodies" [18]. He notes that the predictions of Maxwell's equations depend only on relative motion. For example, the current induced in a coil will have the same magnitude and direction, no matter if we consider a magnet to be moving through the coil or the coil moving through the field of the magnet with equal but opposite velocity. To Einstein, this empirical indistinguishability was in conflict with the prevailing aether theories of the time. An aether would define a distinguished frame of reference, rendering the two empirically equivalent induction experiments ontologically distinct. Once again, Einstein's invocation of Lebniz' principle lead him to abandon aether theory and conceive relativity theory, in which both scenarios are equivalent.

One can lift KS non-contextuality inequalities to Spekkens non-contextuality inequalities by substituting assumptions about operator compatibility and sharpness for assumptions about operational equivalences [27]. Like before, these lifted non-contextuality inequalities, when violated, bear witness to the quantumness of the system, in the Spekkens sense. Violations of non-contextuality inequalities that prove nature's incompatibility with a non-contextual HVM description have been observed in experiments involving photonic qubit systems [29]. While the lifted non-contextuality inequalities no longer make unphysical assumptions about the measurements themselves, they introduce two new practical complications: the issue of tomographic completeness, and that of exact operational equivalence. Experimental tests of non-contextuality inequalities assume that we can implement a tomographically complete (TC) set of measurements, relative to the accessible[2] preparations, and a TC set of preparations, relative to the accessible measurements. The reason for this is that, according to Definitons 6.1, 6.2, two experimental procedures can only be asserted operationally equivalent if the two procedures are operationally indistinguishable, for all possible prepare-and-measure-type experiment that utilize them. For instance, two preparations are only operationally equivalent if they produce the same statistics for all possible measurements, requiring in principle an infinite number of tests. In practice, we aim to identify a TC set of measurements with respect to the two preparations, as defined in REF, such that there is a functional relationship between the statistics of an arbitrary measurement performed on one of the available preparations and the statistics of only the measurements of the TC set. As such, if two preparations are operationally equivalent with respect to a TC set of measurements, they are operationally equivalent with respect to all

[2] An "accessible" preparation procedure is one that is implemented by the experimental apparatus at hand, and whose statistics are recorded as part of the experiment.

measurements. Within QM, a TC set of measurements for a qubit system is given by the three Pauli measurements σ_x, σ_y, and σ_z, that together determine the three Bloch vector components of an arbitrary state.

Definition 6.3 ([38])**.** A set of binary measurements (with outcomes 0,1) $\mathcal{M}_C$ is tomographically complete (TC), relative to a set of preparations $\mathcal{P}$, if for any measurements M (not necessarily contained in $\mathcal{M}$) with outcome set $\mathcal{K}_M$ there exists a deterministic function $f_M : \mathcal{K}_M \times [0,1]^{|\mathcal{M}_C|} \to [0,1]$ with the properties that

1. $\forall k \in \mathcal{K}_M :\ f(k\,,\,\cdot) : [0,1]^{|\mathcal{M}_C|} \to [0,1]$ is a linear mapping, and
2. $\forall k \in \mathcal{K}_M\, \forall P \in \mathcal{P}$: $p(k \,|\, P, M) = f_M(k, \{p(0 \,|\, P, M')\}_{M' \in \mathcal{M}_C})$,

where $p(k \,|\, P, M)$ is the empirical probability of obtaining the outcome k when performing the measurement M on a system prepared like P.

Note that Condition 1 ensures that the mapping f_M is compatible with convex mixtures of preparation procedures.

Problematically, without additional assumptions, we can in practice never declare a set of measurements to be TC with respect to some set of preparations. What we can do is try our best to refute such a claim. Additionally, there are tests of contextuality that account for some number of unknown degrees of freedom, in particular some number of unknown measurements that complete the accessible measurements to a TC set [38]. These strengthened tests of contextuality require knowledge about an upper bound on the information carrying capacity of the system, i.e. the amount of information (in bits) that one can encode in the state of the system in a retrievable manner. By implementing more preparations and measurements one can in principle account for an arbitrary number of unknown degrees of freedom, although the trade-off is exponential. Section 7 will make use of the results in [38], which will ultimately allow us to certify the quantumness of the unknown device in a robust manner, assuming an upper bound on its information carrying capacity. Just like in the case of KS contextuality, where one had to assume operator compatibility and sharp measurements, we cannot expect to devise a fully device-independent self-testing protocol based on non-contextuality.

Another issue with experimental tests of contextuality is that Definitions 6.1, 6.2 assume exact operational equivalences, requiring infinite precision. This problem has been tackled in [29, 36]. Let us for the moment consider operational equivalences amongst preparations. The key is noticing that performing measurements on some set of preparations defines the statistics for all preparations in the convex hull. As part of post-processing, we can define new preparations in the convex hull, whose statistics we know, and that are exactly operationally equivalent. The price to pay is that these new preparations may be less optimal than the original ones. This "fitting" process that establishes exact operational equivalences is done within the framework of generalized probabilistic theories (GPT). Operational equivalences amongst measurements

are treated in an analogous manner. The protocol we propose in Section 7 in fact only assumes approximate operational equivalence of certain preparations.

Finally, we wish to understand how KS contextuality and Spekkens contextuality are related. It turns out that a Spekkens non-contextual ontological model of QM assign deterministic outcomes to projective quantum measurements and induces a KS non-contextual outcome assignment, like in Definition 2.4. KS non-contextuality is an assumption that restricts the ontological representation of compatible (commuting) projective measurements. Let us therefore examine what conditions Spekkens non-contextuality imposes on the ontological representation of compatible measurements. Spekkens non-contextuality covers arbitrary, not necessarily sharp, measurements and is applicable to arbitrary operational theories. Therefore, recall Definition 3.1, which defines what it means for two measurements to be compatible, purely in operational terms.

Let M_1 and M_2 be compatible measurements according to Definition 3.1, and M_{12} a dual outcome measurement that jointly realizes M_1 and M_2. By definition, the measurement procedure M_1 is operationally equivalent to measuring M_{12} and discarding register two. Analogously, the measurement procedure M_2 is operationally equivalent to measuring M_{12} and discarding register one. What can be said about the ontological representation of compatible measurements within a Spekkens non-contextual model? In a Spekkens non-contextual ontological model, operationally equivalent measurement procedures have identical associated indicator functions. For the measurements M_1, M_2, and M_{12} like above, this implies:

$$\forall\, \lambda \in \Lambda,\, \forall\, P,\, \forall\, a_k : p(a_k \,|\, M_1, P, \lambda) = \sum_{b_j} p(a_k, b_j \,|\, M_{12}, P, \lambda)$$
$$\forall\, \lambda \in \Lambda, \forall\, P,\, \forall\, b_k : p(b_k \,|\, M_2, P, \lambda) = \sum_{a_i} p(a_i, b_k \,|\, M_{12}, P, \lambda)$$

Consequently, a Spekkens measurement non-contextual ontological model is one for which the probability of measuring say M_1 and obtaining the outcome a_k, conditioned on the system being in any ontic state, is independent of what other compatible measurements are simultaneously performed. If we consider QM as operational theory, this is the essence of KS non-contextuality, as discussed in Section 2.5, with the notable difference that we have decoupled outcome determinism from context independence.

Interestingly, it can proven that a Spekkens non-contextual ontological model must fix the outcomes of sharp measurements deterministically [42]. This result is robust, in the sense that "almost sharp" measurements are assigned outcomes "almost deterministically". Thus, for projective measurements, the notion of Spekkens non-contextuality reduces to KS non-contextuality. For Spekkens non-contextual ontological models, outcome determinism for sharp measurements can thus be derived from within the framework itself, and is not an assumption introduced ad-hoc, as was the case for KS non-contextuality.

Chapter 7

Protocol based on Spekkens contextuality

We now generalize the self-testing result in [5], which was discussed in Section 5, to the framework of Spekkens contextuality. In particular, this will lift the usual restriction to noise-free projective measurements. Additional features will be discussed in Section 7.3. The protocol we will consider assumes approximate operational equivalences for some preparations, as well as an upper bound on the unknown system's information carrying capacity, much like the protocol in Section 5.

Let $n \geq 5$ be an odd integer. The protocol we consider is again reminiscent of the ideal n-cycle reference experiment 2.7. As such, we assume an experimenter to be able to freely choose between n three-outcome measurements $\{M_i\}_{i=1}^n$ with outcomes m_1, m_2, m_3. We assume measurements to be chosen according to a uniform distribution and denote measurement events by $m_k \,|\, M_i$, or $m_k, m_l \,|\, M_i, M_j$ for two sequential measurements. As shown in Figure 7.1, for ideally operating devices, the measurement events $m_1 \,|\, M_i$ correspond to the rank-one projectors $|u_i^{(n)}\rangle\langle u_i^{(n)}|$ defined by 2.7. Further, in the ideal case, the measurement events

$$m_3 \,|\, M_i \equiv m_1 \,|\, M_{i\oplus 1} \equiv |u_{i\oplus 1}^{(n)}\rangle\langle u_{i\oplus 1}^{(n)}|$$

and the events $m_2 \,|\, M_i$ correspond to the projectors

$$m_2 \,|\, M_i \equiv \mathbb{1} - |u_i^{(n)}\rangle\langle u_i^{(n)}| - |u_{i\oplus 1}^{(n)}\rangle\langle u_{i\oplus 1}^{(n)}|$$

onto the pure qutrit state completing the orthogonal triad $\subset \mathbb{C}^3$. In Figure 7.1, the events $m_3 \,|\, M_i$ and $m_1 \,|\, M_{i\oplus 1}$ are "non-overlapping", despite corresponding to the same ideal projectors, to highlight that we do not need to assume operational equivalence, contrary to the assumptions made in [27].

Like in Section 5, we assume i.i.d. rounds, i.e. the devices to operate in an identical manner for all rounds of the protocol, in particular independently of previous input-output cycles. This

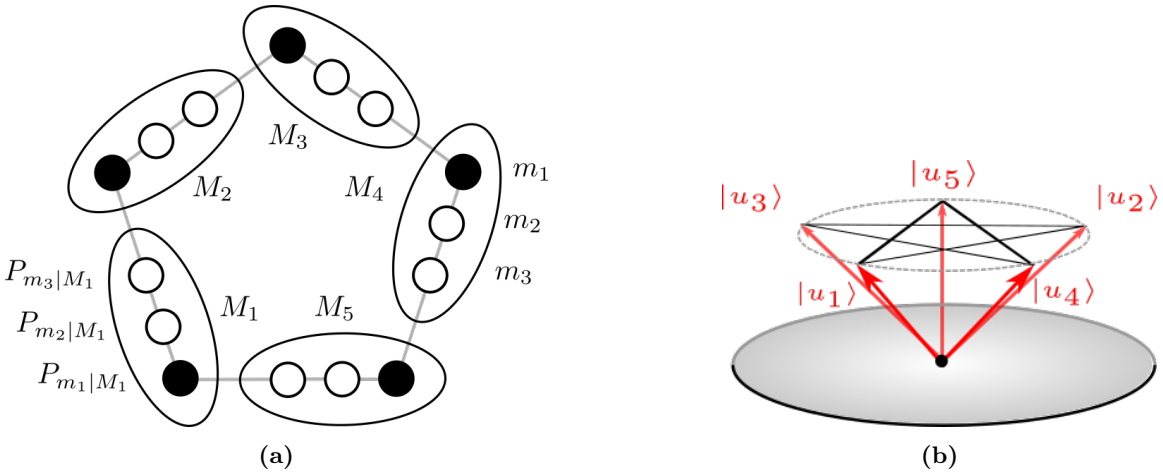


Figure 7.1: (a): Measurement events and retrospective preparations corresponding to a self-testing scenario consistent with cycle length $n = 5$. The experimenter can choose from five three-outcome measurements M_i, with outcomes $\{m_1, m_2, m_3\}$. In post-processing, the experimenter defines $2n$ preparations

$$\{P_{m_k|M_i}\}_{k,i} \quad \text{, for } k \in \{1,2\} \text{ and } i \in \{1,\ldots,n\}.$$

We define $P_{m_3|M_i} \coloneqq P_{m_1|M_{i\oplus 1}}$. In the ideal case, the measurement events $\{m_k|M_i\}_{k\in\{1,2,3\}}$ correspond to rank-one projectors onto the orthogonal triad $\subset \mathbb{C}^3$ containing the cycle states $|u_i\rangle$ and $|u_{i\oplus 1}\rangle$. Furthermore, in the ideal case, the preparations $P_{m_k|M_i}$ are pure states that coincide with the vectors in this triad. The measurement events $m_3|M_i$ and $m_1|M_{i\oplus 1}$ are non-overlapping to illustrate that the self-testing protocol does not assume operational equivalence of these, despite the fact that they correspond to the same rank-one projector onto $|u_{i\oplus 1}\rangle$ in the ideal case. **(b):** Juxtaposition with the cycle states belonging to the odd n-cycle scenario for $n = 5$. Figure (a) can be seen as a top view of (b), with every black dot corresponding to a cycle state $|u_i\rangle$.

allows us to lift relative frequencies to probabilities. We will discuss all other assumptions in due time. The protocol consists of the following steps:

1. The experimenter prepares the system in a distinguished preparation P_0 by selecting the appropriate setting on the corresponding device.

2. The experimenter samples from a uniform distribution and selects two integers $(i, j) \in \{1,\ldots,n\}^2$ at random.

3. The experimenter performs the measurements M_i, M_j in sequence, first M_i, then M_j. He does so by selecting the respective settings on the corresponding devices and records the outcomes.

4. Steps 1-3 are repeated to obtain estimates for the probability distributions $p((m_k, m_l) \,|\, (M_i, M_j), P_0)$.

5. Additionally, the experimenter performs single measurements on two auxilliary preparations P_1 and P_2 he can choose from and obtains estimates for the probability distributions $p(m_k \,|\, M, i, P_j)$, for $j \in 1, 2$.

6. Post-processing

(a) Certificate of quantumness

(b) Bounding compatible quantum models

We write the free choice of measurements (M_i, M_j) and subsequent outcomes (m_k, m_l) as ordered lists to indicate that the distributions $p((m_k, m_l) \,|\, (M_i, M_j), P_0)$ are in general not independent of the order, since we do not assume any underlying compatibility relations like we did in Section 5. In the ideal case, the distinguished preparation P_0 corresponds to the $|0\rangle$ qutrit state, which produces a maximal violation of the KCBS inequality. Additionally, we assume access to two additional auxilliary preparations, P_1 and P_2, which we require to establish approximate operational equivalences during post-processing. In the ideal case the preparations P_1 and P_2 correspond to the $|1\rangle$ and $|2\rangle$ qutrit states, forming an orthogonal triad with $|0\rangle$. The convex combination $\frac{1}{3}\sum_{i=0}^{2} P_i \equiv \frac{1}{3}\sum_{i=0}^{2} |i\rangle\langle i| = \frac{\mathbb{1}}{3}$ corresponds to the fully mixed state $\frac{\mathbb{1}}{3}$, as does the convex combination with equal weights of the ideal projectors corresponding to the measurement events $\{m_k|M_i\}_{k=1}^{3}$, for all $i \in \{1, \ldots, n\}$. Here, we implicitly defined the convex combination of two preparations to be the preparation that is implemented by generating a random number and performing the preparation procedure associated with that outcome. We note that Steps 1-4 of the protocol correspond almost one-to-one to the steps of the protocol in [5], which was discussed in Section 5, the only difference being that the experimenter now freely chooses between n^2, as opposed to n measurement settings.

We now elaborate on Step 6 of the protocol, detailing how an agent may post-process the acquired data, with the goal of self-testing the apparatus. The experimenter retrospectively defines $2n$ preparations, which we denote $P_{m_k|M_i}$, for $i \in \{1, \ldots, n\}$ and $k \in \{1, 2\}$. As suggested by the notation, $P_{m_k|M_i}$ is prepared by performing the measurement M_i on a system initially prepared like P_0, and conditioning on the outcome m_k. We refer to these preparations as retrospective, since they are not directly implemented by the experimenter. However, one can infer the relevant statistics $p(m_k \,|\, M_i, P_{m_l|M_j})$ for these preparations from the distributions $p((m_k, m_l) \,|\, (M_i, M_j))$:

$$p(m_k \,|\, M_i, P_{m_l|M_j}) = \frac{p((m_k, m_l) \,|\, (M_i, M_j))}{\sum_{k' \in \{1,2\}} p((m_{k'}, m_l) \,|\, (M_i, M_j))}.$$

These retrospective preparations will be used in Section 7.1, where we introduce a criterion to certify the quantumness of the apparatus.

The experimenter can also infer the distribution $p(m_k \,|\, M_i, P_0)$ from the correlations $p((m_k, m_l) \,|\, (M_i, M_j), P_0)$:

$$p(m_k \,|\, M_i, P_0) = \sum_{l,j} \frac{1}{n} p((m_k, m_l) \,|\, (M_i, M_j), P_0).$$

These will become relevant in Section 7.2, when bounding the compatible quantum models.

7.1 Step 6a: Certificate of quantumness

By coarse-graining, we obtain $2n$ binary measurements from the n three-outcome measurements M_i: for each M_i, we "lump together" the outcomes m_2, m_3, and m_1, m_3. In total, we define $2n$ binary measurements and $2n$ preparations during post-processing. We denote the outcomes of the $2n$ binary measurements as 0, 1, where the outcome 0 corresponds to the two coarse-grained measurement events. If we consider the ideal reference experiment 2.7, the outcomes 0, 1 correspond to the eigenvalues of the rank-one projectors we can associate with each of the binary measurments. There are n ideal binary measurements that probe the probability of finding the system in one of the n one-dimensional subspaces spanned by the n cycle states. The other n ideal binary measurements correspond to rank-one projectors onto states in an orthogonal triad containing two cycle states. For the extent of this subsection we will for simplicity index the $2n$ binary measurements and preparations like $\{M_i\}_{i\in\{1,\dots,2n\}}$ and $\{P_i\}_{i\in\{1,\dots,2n\}}$. For ideal measurements and preparations, we can choose the order such that the ideal measurement M_i is the rank-one projector onto the ideal pure state preparation P_i. Therefore, for ideal devices, the retrospective preparations and binary measurements, if ordered in this manner, satisfy

$$\epsilon := \max_i p(0|M_i, P_i) \overset{\text{ideal}}{=} 0. \tag{7.1}$$

For noisy devices, we expect to find an ordering for which 7.1 is approximately satisfied, with the parameter ϵ characterizing the amount of noise. The parameter ϵ can be determined from the observed statistics during post-processing.

Apart from the noise-characterizing parameter ϵ, the other important figure of merit is

$$\eta := \min_{\substack{i,j \\ i\neq j}} p(0|M_i, P_j).$$

The parameter η can be thought of as a measuring the maximum overlap or closeness of distinct states. Roughly speaking, the smaller η, the closer distinct states are. As such, we expect η to decrease as the cycle length n increases.

Let $m = 2^k$. In Appendix B of [38] it is proven that for a set of preparations $\{P_i\}_{i=1}^m$ and measurements $\{M_i\}_{i=1}^m$, producing statistics obeying $\epsilon < \frac{1}{4}\eta^2$, to be compatible with a (preparation) non-contextual ontological description, requires at least k binary measurements in any set that is TC with respect to the preparations $\{P_i\}_{i=1}^m$. The intuition behind this is the following: From the statistics of the m preparations $\{P_i\}_{i=1}^m$ one can infer the statistics of any preparation in the convex hull of the $\{P_i\}_{i=1}^m$. If the m preparations are not very "close", as measured by η, relative to their sharpness, as measured by ϵ (this is enforced by the condition $\epsilon < \frac{1}{4}\eta^2$), then the associated probability density functions $\{\mu_i\}_{i=1}^m$ will have non-significant overlap. In particular, if we regard the m probability density functions on the ontic state space Λ as vectors $\{v_i\}_{i=1}^m$ in $\mathbb{R}^{|\Lambda|}$, then the proof in Appendix B of [38] establishes that the $m = 2^k$ vectors vary in k linearly independent directions. Demanding preparation non-contextuality, every distinct

preparation in the convex hull of the $\{P\}_{i=1}^{n}$ must produce different predictions for at least one of the measurements in a TC set. Finally, [38] notes that the function mapping a probability density function μ_j, or rather the corresponding vector in $\mathbb{R}^{|\Lambda|}$ to the probability $p(m_k \,|\, M_i, \mu_j)$ of obtaining the outcome m_k when performing the measurement M_i for the initial preparation μ_j is linear. Hence, such mapping $\mathbb{R}^{|\Lambda|} \to \mathbb{R}$ is of the form $(x_1, \ldots, x_{|\Lambda|}) \mapsto a_1 x_1 + \cdots + a_n x_n = \vec{a} \cdot \vec{x}$ with $0 \leq a_i \leq 1$, and x_i the components w.r.t the standard basis of $\mathbb{R}^{|\Lambda|}$. Binary measurements can therefore only distinguish between preparations along a single direction, $\vec{a}$. Preparations whose ontological representations differ w.r.t. directions orthogonal to $\vec{a}$, are assigned the same probabilities by the binary measurement characterized by $\vec{a}$. A TC set of measurement must contain at least k binary measurements for k "linearly independent"‘ preparations, if we demand compatiblity with a preparation non-contextual ontological model.

We now turn to the bounded memory assumption required by the protocol, which relies on the results of [38], the relevant ones of which were sketched above. Assume that the experimenter has access to $n \geq 5$ three-outcome measurement settings, like we considered at the beginning of Section 7, where n is odd (think of n as the cycle-length). Further, let k be the largest integer satisfying $2n \geq 2^{k+1}$. To certify quantumness, we assume that there exists a TC set of measurements, relative to the $2n$ retrospective preparations, with k binary measurements, i.e. that k binary measurements are sufficient to fully characterize the statistics of the accessible preparations. If we find the condition $\epsilon < \frac{1}{4}\eta^2$ to be satisfied, we consider the system to be quantum. How does this assumption relate to the assumption of bounded memory in Section 5.3?

In Section 5.3, we studied the minimal memory (in bits) a classical simulation requires to reproduce the ideal statistics. For $m \geq 1$ sequential (binary) measurements, consider the 2^m accessible preparation procedures that consist of subjecting the system to m binary measurements and conditioning on one of the 2^m outcome strings. Some of these preparations can be identified, as they are operationally indistinguishable and result in the same causal state. To ensure correct operation, an optimal classical device retains minimal information about the past input-output process, and retrieves that information when confronted with a subsequent input. In this sense, the above preparation procedures can serve to encode some amount of information in the state of the device. To certify quantumness, we have to upper bound the information content one can encode in the state of the device, given the set of accessible preparations. We do so by bounding the information carrying capacity of the unknown system, i.e. the number of bits one can encode in the state of the system in a retrievable manner, using an optimal encoding scheme, and that the retained memory can never exceed the information carrying capacity.

We shall now link these observations regarding the memory assumption in Section 5.3 to the assumptions made by the protocol in Section 7. We define the information carrying capacity of a system S as

$$\mathcal{I} := \max_{\substack{\mathcal{R} \subset \mathcal{P} \\ \text{perfectly distinguishable}}} \log_2(|\mathcal{R}|),$$

where the preparations $\mathcal{R} \subset \mathcal{P}$ are perfectly distinguishable if there exists a measurement procedure that can with certainty identify which $P \in \mathcal{R}$ was prepared. The superset $\mathcal{P}$ contains all possible preparations of S. The existence of a TC set with respect to $\mathcal{P}$ consisting of k binary measurements implies $\mathcal{I} \leq \log_2(k)$ bits, and vice versa. We will prove both implications at the end of this section. In this sense, an upper bound on the minimal size of TC sets of binary measurements with respect to the accessible preparations directly relates to an upper bound on the information content we can encode in the system, given access to that same set of preparation procedures.

If we wish to allow for k measurements in a TC set, we require an apparatus consistent with a cycle length n such that

$$\begin{aligned} & 2n \geq 2^{k+1} \\ \iff & \log_2(n) \geq k. \end{aligned}$$

The relationship between the cycle length and the number of binary measurements in a TC set that this scheme accounts for is logarithmic.

The price one has to pay to account for a greater information carrying capacity is two-fold:

- the apparatus must include additional measurement settings to be consistent with a greater cycle length n
- for a greater cycle-length n, a greater fidelity to the reference experiment is required, since for the ideal odd n-cycle scenario the relevant parameter $\frac{1}{4}\eta^2$, which bounds how sharp the measurement implementations must be for self-testing, decreases with an increasing cycle length n:

  $$\frac{1}{4}\eta^2 = \mathcal{O}\left(\frac{1}{n^4}\right).$$

  This can be seen by computing the Laurent series at $n = \infty$ of the quantities $(1-|\langle u_i|u_j\rangle|^2)$, $(1-|\langle u_i|v_j\rangle|^2)$, and $(1-|\langle v_i|v_j\rangle|^2)$, where $|v_i\rangle \in \mathbb{R}^3$ is the ray orthogonal to $|u_i\rangle$ and $|u_{i\oplus 1}\rangle$, and $i \neq j$.

As announced, we now study more formally how the number of binary measurements in a TC set relates to the information carrying capacity of the system:

Lemma 7.1. *Let S be a classical system with ontic state space Λ. The information carrying capacity $\mathcal{I}$ of S is directly related to the minimal number of binary measurements that constitute a TC complete set like*

$$\exists \text{ TC set of binary measurements, relative to the set of all preparations on } S, \mathcal{P}, \text{ that is of cardinality } k \iff \mathcal{I} \leq \log_2(k).$$

Proof. We prove both implications separately:

"$\Longrightarrow$": Let $\mathcal{M}_C = \{M_i\}_{i=1}^k$ be a TC set of binary measurements of cardinality k. Further, for some subset $\mathcal{R} = \{P_i\}_i \subset \mathcal{P}$, assume that there exists a $|\mathcal{R}|$-outcome measurement M_{dist} that can with certainty distinguish between the preparations $\mathcal{R}$. Since $\mathcal{M}_C$ is TC, there is a mapping $f_{M_{\text{dist}}}$ like in Definition 6.3 that relates the statistics of the distinguishing measurement M_{dist} and the statistics of the measurements in $\mathcal{M}_C$. If we fix an outcome $j \in \mathcal{K}_{M_{\text{dist}}}$, $f_{M_{\text{dist}}}(j,\cdot) : [0,1]^k \to [0,1]$ takes a k-component real vector, whose ith entry we associate with the probability $p(0 \,|\, P, M_i)$ for a suitable preparation P of the system, and maps it to $p(j \,|\, P, M_{\text{dist}})$. The k-component vector belonging to the preparation $P_i \in \mathcal{R}$ is mapped to δ_{jl}. Therefore, due to the linearity of the mappings $f_{M_{\text{dist}}}(j,\cdot)$, the k-component vectors belonging to preparations in $\mathcal{R}$ are all linearly independent. Since any linearly independent set $\subset \mathbb{R}^k$ has cardinality $\leq k$, the same holds true for $\mathcal{R}$, implying that $\mathcal{I} \leq \log_2(k)$.

"$\Longleftarrow$": $\mathcal{I} \leq \log_2(k)$ implies $|\Lambda| \leq k$, and we can label the ontic states like $\Lambda = \{\lambda_i\}_{i=1}^k$ (in case $|\Lambda| < k$, we can add appropriately many "null" ontic states to Λ, such that these do not lie in the support of all preparations $\mathcal{P}$). A TC complete set of binary measurements $\{M_i\}_{i=1}^k$ with outcomes 0,1 is defined by the indicator functions $\xi_{M_i,1}(\lambda_j) = \delta_{ij}$. Clearly, $|\{M_i\}_{i=1}^k| = k$. □

We end this subsection with two important comments. Firstly, the exponentially large number of preparations and binary measurements satisfying $\epsilon < \frac{1}{4}\eta$, used to demonstrate incompatibility with a non-contextual classical model, up to a threshold information carrying capacity, is only a sufficient but not necessary condition. In fact, [38] gives a set of two binary measurements and four preparations that lie in a single slice of the Bloch sphere and exhibit preparation contextuality, where in order to account for one unknown measurement it suffices to consider only one additional preparation. The initial experiment only requires $4 < 2^{2+1} = 8$ preparations to demonstrate contextuality, assuming that the two binary measurements are TC. Furthermore, even if we assume three binary measurements to constitute a TC set, the number of required preparations increases linearly, and not exponentially. Further research could study the family of odd n-cycle scenarios in-depth, with the goal of identifying a sufficient and necessary condition relating the threshold information carrying capacity and the required cycle length n.

The second comment, which is relevant to the next subsection, is that the assumption $\mathcal{I} \leq \log_2(k)$, which for a classical system is equivalent to the existence of a TC set of binary measurements of cardinality k, implies for a quantum system with Hilbert space $\mathcal{H}$ that $\dim(\mathcal{H}) \leq k$.

7.2 Step 6b: Bounding compatible quantum models

As we will see, the key quantity which determines how close the implemented measurements and preparations are to their ideal counterparts is the noise-characterizing parameter ϵ, as defined in Section 7.1. Loosely speaking, it quantifies "how sharp" the implemented measurements are.

As discussed in Section 7.1, we assume a Hilbert space of bounded dimension, say $\dim(\mathcal{H}) \leq d$ (we will often use k as an index of summation). For notational simplicity we define $P_{m_3|M_i} \coloneqq P_{m_1|M_{i\oplus 1}}$.

We will now introduce the second key assumption of the protocol, which allows us to bound the quantum models compatible with the observed correlations: We assume that for all $i \in \{1, \ldots, n\}$ there exists a convex combination $\{p_j^{(i)}\}_j$ of the preparations $\{P_{m_j|M_i}\}_j$ and a convex combination $\{q_j\}_j$ of the preparations $\{P_0, P_1, P_2\}$ such that these are (approximately) operationally equivalent:

$$\forall i \in \{1, \ldots, n\} : S_i \coloneqq \sum_{j=1}^{3} p_j^{(i)} P_{m_j|M_i} \sim S_* \coloneqq \sum_{j=0}^{2} q_j P_j. \tag{7.2}$$

Although assuming approximate operational equivalences suffices, we will for now take the equivalences in 7.2 to be exact. We consider two quantum preparations to be approximately operationally equivalent if the corresponding density operators are "close" with respect to the Frobenius norm. Recall that for ideal devices and equal weights $p_j^{(i)} = q_j = \frac{1}{3}$, both S_i and S_* corresponding to the fully mixed state $\frac{\mathbb{1}}{3}$. Note that the assumption of (approximate) operational equivalence cannot be verified on the basis of statistics alone, and requires us to impose additional constraints, rendering the protocol semi-device-independent. What one can do is check whether the claim holds with respect to the available measurement settings.

The reason we assume the convex combinations S_i and S_* to be (approximately) operationally equivalent is to relate the action of the measurements M_i on the preparations $P_{m_j|M_i}$ to how the measurements act on the initial preparation P_0. MAYBE INCLUDE SOS PAPER

We define $|u_0\rangle$ as an arbitrary purification of ρ_0, where ρ_0 is the density operator corresponding to the preparation P_0. In addition to $|u_0\rangle$, we define the vectors $|u_i\rangle$, $i \in \{1, \ldots, n\}$, like

$$|u_i\rangle \coloneqq (F_{m_1|M_i} \otimes \mathbb{1})|u_0\rangle. \tag{7.3}$$

In the ideal case, these just correspond to the n cycle states of the odd n-cycle scenario. In general, the $(n+1) \times (n+1)$ Gram matrix $X_{ij} = \langle u_i | u_j \rangle$ is not a feasible solution of the Lovász SDP 3.1. While X satisfies $X_{00} = 1$ and is positive semi-definite by construction, the remaining linear constraints are not necessarily satisfied if the $F_{m_1|M_i}$ are not cyclically orthogonal projectors. Note that the entries $X_{0i} = X_{i0} = \mathrm{tr}(F_{m_1|M_i}\rho_0)$ are operationally accessible.

We briefly outline the strategy by which we bound the quantum models that are compatible with the observed correlations: We construct a feasible Gram matrix $\tilde{X}_{ij} = \langle \tilde{u}_i | \tilde{u}_j \rangle$ satisfying all constraints of the SDP 3.1, such that the $|u_i\rangle$ and $|\tilde{u}_i\rangle$ are "close" w.r.t. the 2-norm. Additionally, we require that we can bound the optimality of the feasible $\tilde{X}$ in terms of terms of the operationally accessible quantity $\sum_i \mathrm{tr}(F_{m_1|M_i}\rho_0)$. This, together with Lemmas 5.2 and 5.3, allows us to relate the noisy measurements acting on noisy preparations, $(F_{m_1|M_i} \otimes \mathbb{1})|u_0\rangle$, to the ideal ones, $\Pi_i^{\mathrm{id}}|u_0^{\mathrm{id}}\rangle$.

The following Theorem is consistent with this proof strategy:

Theorem 7.2. *Define* $\{|u_i\rangle\}_{i=0}^n$ *like in 7.3. There exist vectors* $\{|\tilde{u}_i\rangle\}_{i=0}^n$ *such that the Gram matrix* $\tilde{X}_{ij} = \langle\tilde{u}_i|\tilde{u}_j\rangle$ *is a feasible solution of the Lovász SDP 3.1 and*

$$\||\tilde{u}_i\rangle - |u_i\rangle\|_2 \leq \mathcal{O}\left(d^{13/12}\,\epsilon^{1/3}\,\eta^{-1/4}\right).$$

Proof. See Appendix A. □

Let $X^* \in \mathbb{C}^{n+1,n+1}$ be the optimal Gram matrix, as defined in Section 5. $\tilde{X}$ satisfies

$$\left|\sum_i \left(X^*_{ii} - \tilde{X}_{ii}\right)\right| \leq \left|\sum_i \left(\mathrm{tr}(\rho_0 F_{m_1|M_i}) - X^*_{ii}\right)\right| + \mathcal{O}\left(d^{13/12}\,\epsilon^{1/3}\,n\,\eta^{-1/4}\right).$$

As such, by Lemma 5.3, as long as the operationally accessible deviation from the primal optimal value in the above inequality is of the order $\mathcal{O}\left(d^{13/12}\,\epsilon^{1/3}\,\eta^{-1/4}\right)$, all compatible quantum models must satisfy

$$\left\|(F_{m_1|M_i} \otimes \mathbb{1})|u_0\rangle - \Pi_i^{\mathrm{id}}|u_0^{\mathrm{id}}\rangle\right\|_2 \leq \mathcal{O}\left(\nu(n)\,n^{3/2}\eta^{-1/4}\,d^{13/12}\,\epsilon^{1/3}\right).$$

Here, $\nu(n)$ denotes the constant of proportionality in Lemma 5.2, which depends on the cycle length n. The constant $\nu(n)$ is discussed in more detail in Section 5.1.1.

7.3 Discussion

TODO

Appendix A

Proof of Theorem 7.2

A.1 Notation and preliminary considerations

Denote the density operator corresponding to the preparation $P_{m_k|M_i}$ by

$$\rho_{m_k|M_i} = \sum_m p_m^{m_k|M_i} |\Psi_m^{m_k|M_i}\rangle\langle\Psi_m^{m_k|M_i}| \,, \tag{A.1}$$

where $|\Psi_m^{m_k|M_i}\rangle \in \mathbb{C}^k$, and the positive semi-definite operator corresponding to the measurement event $m_k|M_i$ by

$$F_{m_k|M_i} = \sum_l \lambda_l^{m_k|M_i} |\alpha_l^{m_k|M_i}\rangle\langle\alpha_l^{m_k|M_i}| \,, \tag{A.2}$$

where $|\alpha_l^{m_k|M_i}\rangle \in \mathbb{C}^k$. By 7.1,

$$\begin{aligned} p(m_k|M_i, P_{m_k|M_i}) &\geq 1 - 2\epsilon \\ p(m_k|M_i, P_{m_{k'}|M_i}) &\leq \epsilon \,, \end{aligned}$$

for $k' \neq k$. Therefore,

$$\sum_{l,m} \lambda_l^{m_k|M_i} p_m^{m_k|M_i} |\langle\alpha_l^{m_k|M_i}|\Psi_m^{m_k|M_i}\rangle|^2 \geq 1 - 2\epsilon \tag{A.3}$$

and analogously, for $k' \neq k$,

$$\sum_{l,m} \lambda_l^{m_k|M_i} p_m^{m_{k'}|M_i} |\langle\alpha_l^{m_k|M_i}|\Psi_m^{m_{k'}|M_i}\rangle|^2 \leq \epsilon. \tag{A.4}$$

Consider the density operator $\rho_{m_k|M_i}$, like in A.1. For eigenvalues $p_m^{m_k|M_i}$ of the noisy density operator that are very small, say of the order ϵ, we cannot infer from the statistics

$\{p(m_j \,|\, M_i, P_{m_k|M_i})\}_{j,k}$ alone much about how the accessible measurements act on the corresponding subspace $\mathbb{C}|\Psi_m^{m_k|M_i}\rangle$, since this subspace is "insignificant" in terms of statistics. To sidestep this issue, we consider only a "statistically relevant" subspace of $\mathbb{C}^k$: Define $\Pi_{\text{relev}}^{(i)}$ as the projector onto the subspace

$$V_{\text{relev}}^{(i)} = \text{span}\left(\bigcup_{k=1}^{3} \left\{ |\Psi_m^{m_k|M_i}\rangle \right\}_{m:p_m^{m_k|M_i} \geq \mathcal{X}} \right). \tag{A.5}$$

For now, $\mathcal{X}$ is an arbitrary cutoff, characterizing the minimum magnitude of eigenvalues for which the corresponding eigenvector in A.1 spans a statistically relevant one-dimensional subspace. We will later set $\mathcal{X}$ to an appropriate value in terms of ϵ.

The conditions A.3 and A.4 become

$$\sum_{m:p_m^{m_k|M_i} \geq \mathcal{X}} \sum_l (1 - \lambda_l^{m_k|M_i}) |\langle \alpha_l^{m_k|M_i} | \Psi_m^{m_k|M_i} \rangle|^2 \leq \frac{2\epsilon}{\mathcal{X}} \tag{A.6}$$

and

$$\sum_{l,m:p_m^{m_{k'}|M_i} \geq \mathcal{X}} \lambda_l^{m_k|M_i} |\langle \alpha_l^{m_k|M_i} | \Psi_m^{m_{k'}|M_i} \rangle|^2 \leq \frac{\epsilon}{\mathcal{X}}, \tag{A.7}$$

for $k' \neq k$.

For notational simplicity, we will write $\sum_{m:p_m^{m_k|M_i} \geq \mathcal{X}}$ as $\sum_m'$, whenever it is clear what measurement event $m_k \,|\, M_i$ we are referring to, and analogously for two summation indices.

A.2 Relating the retrospective preparations $\rho_{m_k|M_i}$ to ρ_0

Without further assumptions, it is impossible to infer anything about how the measurements M_i act on the preparation P_0 from the correlations $p(m_k \,|\, M_i, P_{m_l|M_j})$ alone. For instance, we cannot rule out that the density operator ρ_0 is in fact orthogonal to all $\rho_{m_l|M_j}$. Recall that $P_{m_l|M_j}$ is defined as the preparation procedure that involves performing the measurement M_j on the preparation P_0, and conditioning on the outcome m_l. If we take $\left\{ E_l^{(j)} \right\}_l$ to be a set of Kraus operators consistent with the measurement M_j, i.e. $E_l^{(j)\dagger} E_l^{(j)} = F_{m_l|M_j}$, one immediately sees that there in fact exists an infinite family of Kraus operators $UE_l^{(j)}$ satisfying this exact condition, where U is an arbitary unitary operator. As such, the post-measurement states are completely undetermined. Note that for PVM we are not plagued with this unitary indeterminacy. In order to get ahead, we assume the operational equivalence of some preparations, as detailed at the beginning of Section 7.2.

For the convex combinations S_i, as defined in 7.2, we find:

$$\left\| \sum_j p_j^{(i)} \left(\rho_{m_j|M_i} - \Pi_{\text{relev}}^{(i)} \rho_{m_j|M_i} \Pi_{\text{relev}}^{(i)} \right) \right\|_F \leq \sum_j p_j^{(i)} \left\| \rho_{m_j|M_i} - \Pi_{\text{relev}}^{(i)} \rho_{m_j|M_i} \Pi_{\text{relev}}^{(i)} \right\|_F \leq d^{1/2} \mathcal{X}.$$

Due to operational equivalence $S_i \sim S_*$, we can write:

$$\left\| \sum_j q_j (\rho_j - \Pi_{\text{relev}}^{(i)} \rho_j \Pi_{\text{relev}}^{(i)}) \right\|_F \leq d^{1/2} \mathcal{X},$$

where ρ_j is the density operator corresponding to the preparation P_j. Therefore,

$$q_0 \left\| (\rho_0 - \Pi_{\text{relev}}^{(i)} \rho_0 \Pi_{\text{relev}}^{(i)}) \right\|_F \leq d^{1/2} \mathcal{X}, \tag{A.8}$$

where we have used Lemma A.1:

Lemma A.1. *Let A, B be two positive semi-definite operators. Then*

$$\|A + B\|_F \geq \|A\|_F.$$

Proof. $\|A + B\|_F^2 = \operatorname{tr}(A^2 + B^2 + 2AB) \geq \operatorname{tr}(A^2) = \|A\|_F^2$, as $\operatorname{tr}(B^2)$, $\operatorname{tr}(AB) \geq 0$ due to semi positive-definiteness. This can be seen by simply inserting an arbitrary spectral decomposition for A, B, and noting that all eigenvalues are non-negative. □

The operators in A.8 are indeed positive semi-definite: they are Hermitian and satisfy

$$\langle \phi | \rho_j - \Pi_{\text{relev}}^{(i)} \rho_j \Pi_{\text{relev}}^{(i)} | \phi \rangle \geq 0,$$

for all $|\phi\rangle \in \mathbb{C}^k$, as can be verified by expanding an arbitrary $|\phi\rangle$ in terms of an ONB with respect to which the projector $\Pi_{\text{relev}}^{(i)}$ is diagonal. Note that we expect $q_0 \approx \frac{1}{3}$ for noisy devices.

Loosely speaking, the above implies that, under the assumption of operational equivalence, ρ_0 has no statistically significant spectral components that are not in $V_{\text{relev}}^{(i)}$. This allows us to infer how the measurements M_i act on P_0 from the correlations $p(m_k \,|\, M_i, P_{m_l|M_j})$.

A.3 Constructing a feasible Gram matrix

In this section, we construct a feasible Gram matrix $\tilde{X}_{ij} = \langle \tilde{u}_i | \tilde{u}_j \rangle$, starting with the vectors $\{|u_i\rangle\}_{i=0}^n$, as defined in Section 7.2. Recall that the matrix $X_{ij} = \langle u_i | u_j \rangle$ in general does not satisfy the $n + (n+1) = 2n+1$ independent linear constraints of the Lovász SDP. In order to estimate how close the matrix X is to the feasible region, we bound the $2n+1$ problematic

entries

$$|X_{0i} - X_{ii}| \text{ , } 1 \leq i \leq n, \text{ and} \tag{A.9}$$

$$|X_{ij}| \text{ , for } j > i,\ i \sim j. \tag{A.10}$$

Starting with $|X_{0i} - X_{ii}|$, we find that

$$|X_{0i} - X_{ii}| = \operatorname{tr}\left((F_{m_1|M_i} - F^2_{m_1|M_i})\rho_0\right) \tag{A.11}$$

$$\leq \operatorname{tr}\left((F_{m_1|M_i} - F^2_{m_1|M_i})\,\Pi^{(i)}_{\text{relev}}\rho_0\Pi^{(i)}_{\text{relev}}\right) + d^{1/2}\left\|\rho_0 - \Pi^{(i)}_{\text{relev}}\rho_0\Pi^{(i)}_{\text{relev}}\right\|_F \tag{A.12}$$

$$\leq d\frac{\mathcal{X}}{q_0} + \frac{1}{q_0}\sum_k p_{m_k|M_i} \operatorname{tr}\left((F_{m_1|M_i} - F^2_{m_1|M_i})\,{\sum_l}' p_l^{m_k|M_i}|\Psi_l^{m_k|M_i}\rangle\langle\Psi_l^{m_k|M_i}|\right) \tag{A.13}$$

$$\leq d\frac{\mathcal{X}}{q_0} + \frac{1}{q_0}\left(\epsilon + {\sum_l}' p_l^{m_1|M_i} \operatorname{tr}\left((F_{m_1|M_i} - F^2_{m_1|M_i})\,|\Psi_l^{m_1|M_i}\rangle\langle\Psi_l^{m_1|M_i}|\right)\right) \tag{A.14}$$

$$\leq d\frac{\mathcal{X}}{q_0} + \frac{2\epsilon}{q_0}. \tag{A.15}$$

From A.11 to A.12 we used the Cauchy Schwarz inequality. To obtain A.13, we make use of the operational equivalence $S_* \sim S_i$. The final upper bound A.15 follows from the inequalities A.3 and A.4.

We can choose the cutoff $\mathcal{X} > 0$ to be arbitrarily small, therefore

$$|X_{0i} - X_{ii}| \leq \frac{2\epsilon}{q_0}\,. \tag{A.16}$$

Bounding the off-diagonal matrix elements $|X_{ij}|$, $j > i$, $i \sim j$, is more involved. The overlap $|X_{ij}| = |\langle u_i|u_j\rangle|$ is by definition equal to

$$\left|\langle u_i|u_j\rangle\right| = \left|\operatorname{tr}\left(\rho_0\, F_{m_1|M_i} F_{m_1|M_{i\oplus 1}}\right)\right|. \tag{A.17}$$

Making use of the triangle and Cauchy-Schwarz inequalities, we obtain

$$\begin{aligned}\left|\operatorname{tr}\left(\rho_0\, F_{m_1|M_i} F_{m_1|M_{i\oplus 1}}\right)\right| \leq{}& \left|\operatorname{tr}\left(F_{m_1|M_i} F_{m_1|M_{i\oplus 1}}\Pi^{(i)}_{\text{relev}}\rho_0\Pi^{(i)}_{\text{relev}}\right)\right| \\ &+ \left\|\left(\rho_0 - \Pi^{(i)}_{\text{relev}}\rho_0\Pi^{(i)}_{\text{relev}}\right)^{1/2}\right\|_F \left\|\left(\rho_0 - \Pi^{(i)}_{\text{relev}}\rho_0\Pi^{(i)}_{\text{relev}}\right)^{1/2} F_{m_1|M_i} F_{m_1|M_{i\oplus 1}}\right\|_F,\end{aligned} \tag{A.18}$$

where the square root of a positive semi-definite operator is well-defined as the Hermitian operator obtained by replacing all eigenvalues with their respective square roots in some spectral decomposition. In the following, all matrix norms are taken with respect to the Frobenius norm, allowing us to omit the 'F' subscript for simplicity.

The second term in A.18 is less or equal to $\mathrm{tr}\left(\rho_0 - \Pi^{(i)}_{\mathrm{relev}}\rho_0\Pi^{(i)}_{\mathrm{relev}}\right) \leq d\mathcal{X}$. We can bound the first term like

$$\begin{aligned}\left|\mathrm{tr}\left(F_{m_1|M_i}F_{m_1|M_{i\oplus1}}\Pi^{(i)}_{\mathrm{relev}}\rho_0\Pi^{(i)}_{\mathrm{relev}}\right)\right| \leq &\left\|{\sum_{j,k}}' |\Psi_k^{m_j|M_i}\rangle\langle\Psi_k^{m_j|M_i}|F_{m_1|M_i}F_{m_1|M_{i\oplus1}}\right\| \\ &+ \left\|\Pi^{(i)}_{\mathrm{relev}} - {\sum_{j,k}}' |\Psi_k^{m_j|M_i}\rangle\langle\Psi_k^{m_j|M_i}|\right\| \left[\mathrm{tr}(F_{m_1|M_i}\rho_0)\,\mathrm{tr}(F_{m_1|M_{i\oplus1}}\rho_0)\right]^{1/2}.\end{aligned} \tag{A.19}$$

The first term on the right hand side of A.19 is to leading order $\leq 2\left(\frac{d\epsilon}{\mathcal{X}}\right)^{1/2}$, which follows from A.7. The second term is less or equal to

$$\left[\mathrm{tr}(F_{m_1|M_i}\rho_0)\,\mathrm{tr}(F_{m_1|M_{i\oplus1}}\rho_0)\right]^{1/2}\left({\sum_{j\neq j'}}{\sum_{k,k'}}' |\langle\Psi_k^{m_j|M_i}|\Psi_{k'}^{m_{j'}|M_i}\rangle|^2\right)^{1/2}. \tag{A.20}$$

We thus need to bound the overlap between statistically significant eigenvectors belonging to distinct measurement outcomes.

In the ideal case statistically significant eigenvectors belonging to distinct measurement outcomes are orthogonal. As we will prove now, this orthogonality still holds approximately in a noisy setting, if the noise-characterizing parameter ϵ is sufficiently small.

Using the fact that $\sum_l F_{m_l|M_i} = \mathbb{1}$, we can write

$$|\langle\Psi_k^{m_j|M_i}|\Psi_{k'}^{m_{j'}|M_i}\rangle| \leq \sum_l |\langle\Psi_k^{m_j|M_i}|F_{m_l|M_i}|\Psi_{k'}^{m_{j'}|M_i}\rangle| \tag{A.21}$$

$$\leq \sum_{l,r} \lambda_r^{m_l|M_i}|\langle\Psi_k^{m_j|M_i}|\alpha_r^{m_l|M_i}\rangle|\,|\langle\alpha_r^{m_l|M_i}|\Psi_{k'}^{m_{j'}|M_i}\rangle|. \tag{A.22}$$

If we get rid of the pesky square root in A.20 by using the fact that $\|\cdot\|_2 \leq \|\cdot\|_1$, insert A.22, and use the inequality A.7, we find that the expression A.20 is to leading order less or equal to

$$2\cdot 3\left[\mathrm{tr}(F_{m_1|M_i}\rho_0)\,\mathrm{tr}(F_{m_1|M_{i\oplus1}}\rho_0)\right]^{1/2} d^{3/2}\left(\frac{\epsilon}{\mathcal{X}}\right)^{1/2}. \tag{A.23}$$

If we consider all terms that contribute to the upper bound on $|X_{ij}|$, we have one contribution $\propto \mathcal{X}$, as well as two contribution $\propto \mathcal{X}^{-1/2}$. This means that we can no longer choose $\mathcal{X}$ to be arbitrarily small or large, as then one of the two contributions would blow up. To find the

optimal value of $\mathcal{X}$, we compute the partial derivative of $|X_{ij}|$ with respect to $\mathcal{X}$ and set this expression to zero. Defining $c_i \coloneqq \left[\operatorname{tr}(F_{m_1|M_i}\rho_0)\operatorname{tr}(F_{m_1|M_{i\oplus 1}}\rho_0)\right]^{1/2}$, the optimal value of $\mathcal{X}$ turns out to be $\mathcal{X} = \left[(d\epsilon)^{1/2}(3c_i + \frac{1}{d})\right]^{2/3}$. For noisy devices we expect $c_i \approx \cos^2(\theta_n)$, as defined in Section 2.9.

Piecing everything together, we can bound $|X_{i\,i\oplus 1}|$ like

$$|X_{i\,i\oplus 1}| \leq \mathcal{O}\left(d^{4/3}\epsilon^{1/3}\right). \tag{A.24}$$

Importantly, the numerical constant is operationally accessible.

Having derived upper bounds for both $|X_{ij}|$ and $|X_{0i} - X_{ii}|$, we can now enforce the linear constraints of the Lovász SDP 3.1, beginning with cyclic orthogonality.

Lemma A.2. *Consider the set of vectors $\{|u_i\rangle\}_{i=1}^n \subset \mathbb{C}^d$, as defined in Section 7.2, with the property that $|\langle u_i, u_{i\oplus 1}\rangle| \leq \delta_{\epsilon,d}$ for all $1 \leq i \leq n$. If we use the standard Gram-Schmidt method to orthogonalize adjacent vectors, the new set of vectors $\{|u_i'\rangle\}_{i=1}^n$ satisfies*

$|\langle u_i'|u_{i\oplus 1}'\rangle| = 0$, *and to leading order*

$$\||u_i'\rangle - |u_i\rangle\|_2 \leq \delta \operatorname{tr}(\rho_0 F_{m_1|M_{i\ominus 1}})^{-1/2}\left(1 + \frac{\operatorname{tr}(\rho_0 F_{m_1|M_i})^{1/2}}{\operatorname{tr}(\rho_0 F_{m_1|M_{i\ominus 2}})^{1/2}}\right) =: \Theta_i^\delta, \text{ in case } 1 \leq i < n, \text{ and}$$

$$\||u_n'\rangle - |u_n\rangle\|_2 \leq \Theta_n^\delta + \left[\kappa\, d^{1/2}\left(1 - \operatorname{tr}(\rho_{m_1|M_i}F_{m_1|M_j})\right)^{1/2} + \left(\operatorname{tr}(\rho_0 F_{m_1|M_1}) - \frac{\operatorname{tr}(\rho_0 F_{m_1|M_i})^2}{\operatorname{tr}(\rho_0 F_{m_1|M_{n-1}})}\right)\right]^{-1/2}$$
$$\cdot\left[\delta + \left(\frac{\operatorname{tr}(F_{m_1|M_1}\rho_0)}{\operatorname{tr}(F_{m_1|M_{n-1}}\rho_0)}\right)^{1/2}\left(\delta + \operatorname{tr}(F_{m_1|M_n}\rho_0)^{1/2}\,\Theta_{n-1}^\delta\right)\right],$$

where

$$\kappa = 2\,\frac{\operatorname{tr}(\rho_0 F_{m_1|M_i})}{\operatorname{tr}(\rho_0 F_{m_1|M_{n-1}})}.$$

Here, $i \neq j \in \{1, n-1\}$ such that $\operatorname{tr}(\rho_0 F_{m_1|M_i}) \geq \operatorname{tr}(\rho_0 F_{m_1|M_j})$.

Proof. The standard Gram-Schmidt method defines $|u_1'\rangle \coloneqq |u_1\rangle$. Subsequent $|u_i'\rangle$, $1 < i < n$, are found by subtracting from $|u_i\rangle$ its projection onto $|u_{i-1}'\rangle$:

$$|u_i\rangle = \frac{\langle u_{i-1}'|u_i\rangle}{\langle u_{i-1}'|u_{i-1}'\rangle}|u_{i-1}'\rangle. \tag{A.25}$$

As such, the statement of Lemma A.2 holds trivially for $i = 1, 2$.

Let $2 < i < n$. By the above method, we find that, to leading order,

$$\||u_i'\rangle - |u_i\rangle\|_2 \leq \delta\,\||u_{i-1}\rangle\|_2^{-1}\left(1 + \frac{\||u_i\rangle\|_2}{\||u_{i-2}\rangle\|_2}\right), \tag{A.26}$$

and A.2 follows by definition of the $|u_i\rangle$.

The case where $i = n$ is more tedious, since we require that $|u'_n\rangle$ be orthogonal to both $|u'_1\rangle = |u_1\rangle$ and $|u'_{n-1}\rangle$. To achieve this, we subtract from $|u_n\rangle$ its component parallel to $|u'_{n-1}\rangle$, as well as its component parallel to

$$|\bar{u}_1\rangle = |u'_1\rangle - \frac{\langle u'_{n-1}|u'_1\rangle}{\langle u'_{n-1}|u'_{n-1}\rangle}|u'_{n-1}\rangle.$$

The computation of the upper bound on $\||u'_n\rangle - |u_n\rangle\|_2$ is analogous to the case where $i < n$, with the notable exception that, in order to get a meaningful upper bound, we must prove that $\||\bar{u}_1\rangle\|_2$ is non-zero. In particular, we want that $\delta\||\bar{u}_1\rangle\|_2^{-1}$ remains small, since $|u'_n\rangle$ includes correction terms $\sim \delta\||\bar{u}_1\rangle\|_2^{-1}$. Fortunately, it turns out that $\||\bar{u}_1\rangle\|_2 \geq \mathcal{O}(\eta)$, where η is defined in Section 7.1. We will later set δ to $\max_i |X_{i\,i\oplus 1}| \leq \mathcal{O}\left(d^{4/3}\epsilon^{1/3}\right)$, and assume that ϵ is much smaller than $\eta = \mathcal{O}(\frac{1}{n})$.

En route to lower bounding $\||\bar{u}_1\rangle\|_2^2 = \langle\bar{u}_1|\bar{u}_1\rangle \approx \langle u_1|u_1\rangle - \frac{|\langle u_1|u_{n-1}\rangle|^2}{\langle u_{n-1}|u_{n-1}\rangle}$, we show that

$$\left\|\Pi^{(1)}_{\text{relev}} F_{m_1|M_{n-1}} - {\sum_k}' |\Psi_k^{m_1|M_1}\rangle\langle\Psi_k^{m_1|M_1}|\right\| \leq 2\left(\frac{1-\text{tr}(\rho_{m_1|M_1}F_{m_1|M_{n-1}})}{\mathcal{X}}\right).$$

This is a direct consequence of Lemma A.3. From this intermediate result, setting

$$\mathcal{X} \propto \left(\frac{1-\text{tr}(\rho_{m_1|M_1}F_{m_1|M_{n-1}})}{d}\right)^{1/2},$$

it follows that

$$\langle u_1|u_{n-1}\rangle = \text{tr}(\rho_0 F_{m_1|M_1}F_{m_1|M_{n-1}}) = \text{tr}(\rho_0 F_{m_1|M_1}) + \mathcal{O}\left(d^{1/2}\left(1-\text{tr}(\rho_{m_1|M_1}F_{m_1|M_{n-1}})\right)^{1/2}\right). \tag{A.27}$$

We can also substitute ${\sum_k}' |\Psi_k^{m_1|M_{n-1}}\rangle\langle\Psi_k^{m_1|M_{n-1}}|$ for $\Pi^{(n-1)}_{\text{relev}} F_{m_1|M_1}$, giving us

$$\langle u_1|u_{n-1}\rangle = \text{tr}(\rho_0 F_{m_1|M_i}) + \mathcal{O}\left(d^{1/2}\left(1-\text{tr}(\rho_{m_1|M_i}F_{m_1|M_j})\right)^{1/2}\right),$$

where $i \neq j \in \{1, n-1\}$ and $\text{tr}(\rho_0 F_{m_1|M_i}) \geq \text{tr}(\rho_0 F_{m_1|M_j})$.

Noting that $\langle u_1|u_{n-1}\rangle \leq \left(\text{tr}(\rho_0 F_{m_1|M_1})\,\text{tr}(\rho_0 F_{m_1|M_{n-1}})\right)^{1/2} \leq \max_{i\in\{1,n-1\}}\text{tr}(\rho_0 F_{m_1|M_i})$, we conclude that

$$\begin{aligned}\||\bar{u}_1\rangle\|_2^2 &\geq \left[\text{tr}(\rho_0 F_{m_1|M_1}) - \frac{\text{tr}(\rho_0 F_{m_1|M_i})^2}{\text{tr}(\rho_0 F_{m_1|M_{n-1}})}\right] + 2\,d^{1/2}\frac{\text{tr}(\rho_0 F_{m_1|M_i})}{\text{tr}(\rho_0 F_{m_1|M_{n-1}})}\left(1-\text{tr}(\rho_{m_1|M_i}F_{m_1|M_j})\right)^{1/2} \\ &= \left[\text{tr}(\rho_0 F_{m_1|M_1}) - \frac{\text{tr}(\rho_0 F_{m_1|M_i})^2}{\text{tr}(\rho_0 F_{m_1|M_{n-1}})}\right] + \kappa\,d^{1/2}\left(1-\text{tr}(\rho_{m_1|M_i}F_{m_1|M_j})\right)^{1/2},\end{aligned}$$

where we defined

$$\kappa := 2\,\frac{\mathrm{tr}(\rho_0 F_{m_1|M_i})}{\mathrm{tr}(\rho_0 F_{m_1|M_{n-1}})}.$$

□

Lemma A.3. *Define the positive semi-definite operators $\rho_{m_1|M_1}$, $F_{m_1|M_{n-1}}$, and the cutoff $\mathcal{X}$ like at the beginning of Section 7.2. If we define $\zeta := 1 - \mathrm{tr}(\rho_{m_1|M_1} F_{m_1|M_{n-1}})$, then*

$$\left\| \Pi^{(1)}_{relev} F_{m_1|M_{n-1}} - {\sum_k}' |\Psi_k^{m_1|M_1}\rangle\langle\Psi_k^{m_1|M_1}| \right\| \leq \mathcal{O}\left(\frac{\zeta}{\mathcal{X}}\right),$$

assuming that ϵ is much smaller than ζ.

Proof. The expression

$$\left\| \Pi^{(1)}_{\mathrm{relev}} - \sum_j {\sum_k}' |\Psi_k^{m_1|M_1}\rangle\langle\Psi_k^{m_1|M_1}| \right\| = \left(\sum_{j\neq j'} {\sum_{k,k'}}' |\langle\Psi_k^{m_1|M_1}|\Psi_k^{m_1|M_1}\rangle|^2 \right)^{1/2}$$

already appeared in A.20, and we found an upper bound $\mathcal{O}\left(d^{3/2}\left(\frac{\epsilon}{\mathcal{X}}\right)^{1/2}\right)$. As such,

$$\begin{aligned}\left\| \Pi^{(1)}_{\mathrm{relev}} F_{m_1|M_{n-1}} - {\sum_k}' |\Psi_k^{m_1|M_1}\rangle\langle\Psi_k^{m_1|M_1}| \right\| &\leq \mathcal{O}\left(d^{3/2}\left(\frac{\epsilon}{\mathcal{X}}\right)^{1/2}\right) + \left\| \left(F_{m_1|M_i} - \mathbb{1}\right) {\sum_k}' |\Psi_k^{m_1|M_i}\rangle\langle\Psi_k^{m_1|M_i}| \right\| \\ &\quad + \left\| \sum_{j\neq 1} {\sum_k}' |\Psi_k^{m_j|M_1}\rangle\langle\Psi_k^{m_j|M_1}| F_{m_1|M_i} \right\| \\ &\leq 2\frac{\zeta}{\mathcal{X}},\end{aligned}$$

where we used A.6 and A.7, substituting $\epsilon \leftrightarrow \zeta$.

□

Define $X'_{ij} = \langle u'_i|u'_j\rangle$, where $\{|u'_i\rangle\}_{i=0}^n$ is the set of vectors introduced in Lemma A.2. We left $|u_0\rangle$ untouched, therefore $|u'_0\rangle := |u_0\rangle$. Lastly, we enforce the constraints $X'_{0i} = X'_{ii}$, by multiplying each $|u'_i\rangle$ with $\frac{X'_{0i}}{X'_{ii}}$, giving us $|\tilde{u}_i\rangle := \frac{X'_{0i}}{X'_{ii}}|u'_i\rangle$. This preserves positive semi-definiteness, since any matrix whose entries are given by the inner product between vectors from a given set, in this case $\{|\tilde{u}_i\rangle\}_{i=1}^n$, is by construction positive semi-definite. To leading order

$$\frac{X'_{0i}}{X'_{ii}} = 1 + \frac{X'_{0i} - X'_{ii}}{X'_{0i}},$$

and

$$\||\tilde{u}_i\rangle - |u_i\rangle\|_2 \leq \||u_i\rangle - |u'_i\rangle\|_2 + \||u'_i\rangle\|_2 \left|\frac{X'_{0i} - X'_{ii}}{X'_{0i}}\right|,$$

giving us

$$\||\tilde{u}_i\rangle - |u_i\rangle\|_2 \leq \||u_i\rangle - |u_i'\rangle\|_2 \left[3 + \mathrm{tr}(\rho_0 F_{m_1|M_i})^{-1/2}\right].$$

Finally, Lemma A.2 implies that

$$\||\tilde{u}_i\rangle - |u_i\rangle\|_2 \leq \mathcal{O}\left(d^{13/12}\,\epsilon^{1/3}\,\eta^{-1/4}\right).$$

Bibliography

[1] B. Amaral. *On graph approaches to contextuality and their role in quantum theory*. Springer, 2018. DOI: 10.1007/978-3-319-93827-1.

[2] N. Barnett and J. Crutchfield. Computational mechanics of input-output processes: Structured transformations and the ϵ-transducer. *J. Stat. Phys.*, 161, 2015. DOI: 10.1007/s10955-015-1327-5.

[3] J. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics Physique Fizika*, 1, 1964. DOI: 10.1103/physicsphysiquefizika.1.195.

[4] J. Bell. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.*, 38, 1966. DOI: 10.1103/revmodphys.38.447.

[5] K. Bharti, M. Ray, A. Varvitsiotis, N. Warsi, A. Cabello, and L.-C. Kwek. Robust self-testing of quantum systems via noncontextuality inequalities. *Phys. Rev. Lett.*, 122, 2019. DOI: 10.1103/physrevlett.122.250403.

[6] G. Brassard, R. Cleve, and A. Tapp. Cost of exactly simulating quantum entanglement with classical communication. *Phys. Rev. Lett.*, 83, 1999. DOI: 10.1103/physrevlett.83.1874.

[7] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86, 2014. DOI: 10.1103/revmodphys.86.419.

[8] M. Buchanan. Leibniz's principle of indistinguishability. *Nat. Phys.*, 15, 2019. DOI: 10.1038/s41567-019-0711-5.

[9] A. Cabello. Specker's fundamental principle of quantum mechanics. *arXiv preprint arXiv:1212.1756*, 2012. URL https://arxiv.org/abs/1212.1756.

[10] A. Cabello. Simple explanation of the quantum violation of a fundamental inequality. *Phys. Rev. Lett.*, 110, 2013. DOI: 10.1103/physrevlett.110.060402.

[11] A. Cabello, S. Severini, and A. Winter. Graph-theoretic approach to quantum correlations. *Phys. Rev. Lett.*, 112, 2014. DOI: 10.1103/physrevlett.112.040401.

[12] A. Cabello, M. Gu, O. Guehne, J.-A. Larsson, and K. Wiesner. Thermodynamical cost of some interpretations of quantum theory. *Phys. Rev. A*, 94, 2016. DOI: 10.1103/physreva.94.052127.

[13] A. Cabello, M. Kleinmann, and J. Portillo. Quantum state-independent contextuality requires 13 rays. *J. Phys. A*, 49, 2016. DOI: 10.1088/1751-8113/49/38/38lt01.

[14] A. Cabello, M. Gu, O. Guehne, and Z.-P. Xu. Optimal classical simulation of state-independent quantum contextuality. *Phys. Rev. Lett.*, 120, 2018. DOI: 10.1103/physrevlett.120.130401.

[15] J. Clauser, M. Horne, A. Shimony, and R. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23, 1969. DOI: 10.1103/physrevlett.23.880.

[16] R. Colbeck. Bell non-locality. Lecture as part of Solstice of Foundations summer school, ETH Zurich, 2019. URL https://video.ethz.ch/conferences/2019/quantum.html.

[17] J. Conway and S. Kochen. *The Strong Free Will Theorem*. Cambridge University Press, 2011. DOI: 10.1017/cbO9780511976971.014.

[18] A. Einstein. On the electrodynamics of moving bodies. *Annalen der Physik*, 17, 1905. DOI: 10.1002/andp.19053221004.

[19] A. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67, 1991. DOI: 10.1103/physrevlett.67.661.

[20] A. Fine. Hidden variables, joint probability, and the bell inequalities. *Phys. Rev. Lett.*, 48, 1982. DOI: 10.1103/physrevlett.48.291.

[21] B. Hensen, H. Bernien, A. Dreau, A. Reiserer, N. Kalb, M. Blok, J. Ruitenberg, R. Vermeulen, R. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. Mitchell, M. Markham, D. Twitchen, D. Elkouss, S. Wehner, T. Taminiau, and R. Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526, 2015. DOI: 10.1038/nature15759.

[22] C. Heunen, T. Fritz, and M. Reyes. Quantum theory realizes all joint measurability graphs. *Phys. Rev. A*, 89, 2014. DOI: 10.1103/physreva.89.032121.

[23] R. Horn. *Matrix analysis*. Cambridge University Press, second edition edition, 2013. DOI: 10.1017/9781139020411.

[24] M. Kleinmann, O. Guehne, J. Portillo, J.-A. Larsson, and A. Cabello. Memory cost of quantum contextuality. *New J. Phys.*, 13, 2011. DOI: 10.1088/1367-2630/13/11/113011.

[25] A. Klyachko, M. Can, S. Binicioglu, and A. Shumovsky. Simple test for hidden variables in spin-1 systems. *Phys. Rev. Lett.*, 101, 2008. DOI: 10.1103/physrevlett.101.020403.

[26] S. Kochen and E. Specker. The problem of hidden variables in quantum mechanics. *Indiana Univ. Math. J.*, 17, 1968. DOI: 10.1512/iumj.1968.17.170044.

[27] R. Kunjwal. Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements. *Quantum*, 3, 2019. DOI: 10.22331/q-2019-09-09-184.

[28] D. Mayers and A. Yao. Self testing quantum apparatus. *arXiv preprint quant-ph/0307205*, 2003. URL https://arxiv.org/abs/quant-ph/0307205.

[29] M. Mazurek, M. Pusey, R. Kunjwal, K. Resch, and R. Spekkens. An experimental test of noncontextuality without unphysical idealizations. *Nat. Commun.*, 7, 2016. DOI: 10.1038/ncomms11780.

[30] M. McKague and M. Mosca. Generalized self-testing and the security of the 6-State Protocol. In *Theory of quantum computation, communication, and cryptography*. Springer, 2011. DOI: 10.1007/978-3-642-18073-6_10.

[31] D. Mermin. Hidden variables and the two theorems of John Bell. *Rev. Mod. Phys.*, 65, 1993. DOI: 10.1103/revmodphys.65.803.

[32] N. Miklin and M. Oszmaniec. A universal scheme for robust self-testing in the prepare-and-measure scenario. *arXiv preprint arXiv:2003.01032*, 2020. URL https://arxiv.org/abs/2003.01032.

[33] M. Nielsen and I. Chuang. *Quantum computation and quantum information: 10th anniversary edition*. Cambridge University Press, 2010. DOI: 10.1017/cbO9780511976667.

[34] A. Peres. Two simple proofs of the Kochen-Specker theorem. *J. Phys. A*, 24, 1991. DOI: 10.1088/0305-4470/24/4/003.

[35] A. Peres. *Quantum theory: Concepts and methods*. Kluwer, 2002. DOI: 10.1007/0-306-47120-5.

[36] M. Pusey. Robust preparation noncontextuality inequalities in the simplest scenario. *Phys. Rev. A*, 98, 2018. DOI: 10.1103/physreva.98.022112.

[37] M. Pusey. Contextuality. Lecture as part of Solstice of Foundations summer school, ETH Zurich, 2019. URL https://video.ethz.ch/conferences/2019/quantum.html.

[38] M. Pusey, L. del Rio, and Bettina Meyer. Contextuality without access to a tomographically complete set. *arXiv preprint arXiv:1904.08699*, 2019. URL https://arxiv.org/abs/1904.08699.

[39] E. Specker. The logic of propositions which are not simultaneously decidable. In *The logico-algebraic approach to quantum mechanics: Volume I: Historical Evolution*. Springer, 1975. DOI: 10.1007/978-94-010-1795-4_8.

[40] R. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Phys. Rev. A*, 71, 2005. DOI: 10.1103/physreva.71.052108.

[41] R. Spekkens. Foundations of quantum mechanics. Lecture Series, Perimeter Institute, 2012. URL https://www.perimeterinstitute.ca/video-library/collection/11/12-psi-foundations-quantum-mechanics.

[42] R. Spekkens. The status of determinism in proofs of the impossibility of a noncontextual model of quantum theory. *Foundations of Physics*, 44, 2014. DOI: 10.1007/s10701-014-9833-x.

[43] R. Spekkens. The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein. *arXiv preprint arXiv:1909.04628*, 2019. URL https://arxiv.org/abs/1909.04628.

[44] I. Supic and J. Bowles. Self-testing of quantum systems: A review. *Quantum*, 4, 2020. DOI: 10.22331/q-2020-09-30-337.

[45] B. Toner and D. Bacon. Communication cost of simulating Bell correlations. *Phys. Rev. Lett.*, 91, 2003. DOI: 10.1103/physrevlett.91.187904.

[46] J. Watrous. *The theory of quantum information*. Cambridge University Press, 2018. DOI: 10.1017/9781316848142.

[47] K. Wiesner, M. Gu, E. Rieper, and V. Vedral. Information-theoretic lower bound on energy cost of stochastic computation. *Proc. R. Soc. A*, 468, 2012. DOI: 10.1098/rspa.2012.0173.

[48] S. Yu and C. Oh. State-independent proof of Kochen-Specker theorem with 13 rays. *Phys. Rev. Lett.*, 108, 2012. DOI: 10.1103/physrevlett.108.030402.

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.