In order to better understand a complex theory like quantum mechanics, it is sometimes useful to take a step back and create alternative theories, with more intuitive foundations, and examine which features of quantum mechanics can be reproduced by such a 
foil theory. A prominent example is Spekkens’ toy theory, which is based off a simple premise: “What if we took a common classical theory and added the uncertainty principle as a postulate?” In other words, the theory imposes an 
epistemic restriction on our knowledge about a physical system: only half of the variables can ever be known to an observer. Like good science fiction, from this simple principle a rich behaviour emerges, most notoriously when we compose several systems. The toy theory emulates some aspects of quantum non-locality, although crucially it is still a non-contextual model. In this pedagogical review we consolidate different approaches to Spekkens’ toy theory, including the stabilizer formalism and the generalization to arbitrary dimensions, completing them with new results where necessary. In particular, we introduce a general characterization of measurements, superpositions and entanglement in the toy theory.

“Fiction writers, at least in their braver moments, do desire the truth: to know it, speak it, serve it. But they go about it in a peculiar and devious way, which consists in inventing persons, places, and events which never did and never will exist or occur, and telling about these fictions in detail and at length and with a great deal of emotion, and then when they are done writing down this pack of lies, they say, There! That’s the truth.”

Ursula K. Le Guin, The Left Hand of Darkness

Ladina Hausmann: l@dinahausmann.ch
Nuriya Nurgalieva: nuriya@phys.ethz.ch
Lídia del Rio: lidia@phys.ethz.ch
What to expect from this review

Goals and audience. This article gives a pedagogical overview of different approaches to Spekkens’ toy theory. We hope it is useful for readers new to the theory, who wish to get a smooth introduction, and also to those familiar with a few results, who are looking for a consolidated picture or a quick resource for consultation.

Approaches covered. We present three formulations of Spekkens’ toy theory: the original formulation by Spekkens [1], its reformulation in the stabilizer notation by Pusey [2], and a generalization to systems of an arbitrary dimension by Spekkens [3]. Even though the restrictions on knowledge appear different for the different versions, the first two are equivalent and are a special case of the generalized toy theory [2–4]. In short, the original formulation is where we get the intuition, the stabilizer approach is better suited to prove results and find quantum analogs, and the generalization takes the best of both. For each formulation, we define epistemic states, measurements, mixtures, superpositions, entanglement and allowed transformations.

New results. In addition to reviewing current approaches, this article adds a number of contributions and simplifications to Spekkens’ toy theory, mostly concerning the generalization to arbitrary dimensions. These are listed in the conclusions (section 6.2). Our new results are blended into the natural section referring to each topic, among established results (and not in an independent section “results”). Existing results in the literature have explicit references; new lemmas, theorems and definitions are those without a reference. All proofs can be found in the appendix.

Contents

1 Epistemic states 3
2 Measurements 16
3 Superpositions and mixtures 24
4 Entanglement 31
5 Transformations 35
6 Discussion 39
Appendix 43
A Linear algebra lemmas 43
Spekkens’ toy theory is an epistemically-restricted theory [5]. Such theories distinguish two types of states: ontic states, which encode the physical state of a system, and epistemic states – the states of knowledge that an observer has about the system. In the toy theory, epistemic states are constrained by the knowledge balance principle (Figure 1.1), inspired by the Heisenberg uncertainty principle:

“If one has maximal knowledge, then for every system, at every time, the amount of knowledge one possesses about the ontic state of the system at that time must equal the amount of knowledge one lacks.” [1]

We will see how this principle is implemented in different formalizations of the theory, and how it restricts the set of allowed states and the behaviour of the toy theory.

1.1 Original toy theory: epistemically-restricted picture

The knowledge balance principle is a restriction on the amount of knowledge one can have, and to understand this principle we first need to quantify knowledge. As a quantifier, Spekkens takes the
amount of questions in a canonical set that an observer can answer. A canonical set is defined as the minimal set of questions that can fully specify each ontic state with different combinations of answers. For example, assume that we only allow for binary questions, and consider a system with four different ontic states, labelled 1, 2, 3 and 4. As this is a two-bit system, a canonical set has two questions, for example

\{“Is the ontic state in \{1, 2\}?”, “is the ontic state in \{1, 3\}?”\}.

For instance, an answer of (no, no) would single out the ontic state as 4. The knowledge-balance principle restricts an observer to only know the answer to one of these questions. Suppose that we know the answer to the first question is “yes”; then the epistemic state would be “the ontic state is in \{1, 2\}”. We denote this state as 1 \lor 2.

**Graphical representation.** Epistemic states can also be graphically represented; in this representation we assign each ontic state as a cell in a grid,

\[
\begin{array}{cccc}
 1 & 2 & 3 & 4 \\
\end{array}
\] .

(1.1)

The epistemic state 1 \lor 2 is then represented by marking the cells where the ontic state could be,

\[
\begin{array}{cccc}
 1 & 2 & 3 & 4 \\
\end{array}
\] .

(1.2)

**Further restrictions.** To fully describe the toy theory, we need three more assumptions [1]. Firstly, we only consider physical systems that allow for a knowledge-balance principle. In particular, each ontic state must be uniquely described by a set of answers to a canonical set of questions, and the number of questions needs to be divisible by two. If there are only two questions in the canonical set, we say that such a system is elementary, like in the above example. Secondly, we assume that every system is composed of elementary systems, and that the knowledge balance principle applies to each possible set of subsystems. We will see that this assumption additionally constrains the restricts epistemic states. Thirdly, we assume that the epistemic state does not constrain the evolution of a system, as it seems implausible that the state of mind of an observer could influence the system.\(^2\)

**Pure and mixed states.** We call an epistemic state a maximal information state or pure if exactly half of the answers to the questions in a canonical set are known; otherwise we call such a state non-maximal informational or mixed.

\(^1\)In principle we can build theories where we would use canonical questions with three or more answers. However, these theories cannot easily be connected to each other [1].

\(^2\)This assumption will be addressed in future work.
Composed systems. If a system is composed of \( N \) elementary subsystems, it has \( 2N \) binary questions in the canonical set (two for each subsystem), and therefore has \( 2^{2N} \) ontic states. Note that any set of questions that uniquely characterizes any ontic state with \( 2^{2N} \) elements is a canonical set of questions, because differentiating \( 2^{2N} \) ontic states requires \( 2N \) bits of information. We denote an ontic state \( o \) on \( N \) subsystems as the composition of the ontic states of each single systems \( o = (o_1 \times o_2 \times \ldots \times o_N) \), where \( o_i \) is the ontic state of the \( i \)th subsystem. The symbol \( \times \) represents the composition of the toy systems.

We can extend the graphical notation for systems composed of two elementary subsystems. Each ontic state is represented by a field in a \( 4 \times 4 \) grid, such that the ontic state \( (o_A \times o_B) \) is the in the \( o_A \)th row and the \( o_B \)th column as in the grid below:

\[
\begin{array}{cccc}
4 & 3 & 2 & 1 \\
3 & 2 & 1 & 4 \\
2 & 1 & 4 & 3 \\
1 & 4 & 3 & 2 \\
\end{array}
\]

An epistemic state is then represented by the fields where the ontic state could be. For example, the state \( (1 \times 1) \lor (1 \times 2) \lor (2 \times 1) \lor (2 \times 2) \) would be represented by

\[
\begin{array}{cccc}
4 & 3 & 2 & 1 \\
3 & 2 & 1 & 4 \\
2 & 1 & 4 & 3 \\
1 & 4 & 3 & 2 \\
\end{array}
\]

For systems composed of more than two subsystems, states can be represented in a \( N \)-dimensional grid where each side has length 4 [1].

Let us investigate which epistemic states are allowed under these conditions. To do so, we need to characterize the canonical sets of questions, and find out what the knowledge balance principle entails for epistemic states. We show that it breaks permutation symmetry and imposes an inductive structure.

To formalize the notion of question in the toy theory, we introduce a map \( \mathcal{M} \) that maps a question

\[ Q = \text{“Is the ontic state in } M_0 \text{?”} \] (1.5)

to a partition of all ontic states \( O \):

\[ \mathcal{M} : Q \rightarrow (M_0, O \setminus M_0) = (M_0, M_1). \] (1.6)

We can write the answer to a question \( Q_i \) as a bit \( a_i \): 0 for “yes, the ontic state is in \( M_{i,0} \)” and 1 for “no, the ontic state is not in \( M_{i,0} \); it is in \( M_{i,1} \) instead.” For a given set of \( k \) questions \( Q \), we
can represent a hypothetical answer as a \( k \)-bit string \( a \) where each individual bit \( a_i \in \{0,1\} \) is the answer to question \( Q_i \). Defining these partitions allows us to characterize canonical sets (proofs in the appendix).

**Lemma 1.1 (Canonical set of questions).** Let \( Q \) be a set of \( 2N \) questions for a system with \( N \) elementary subsystems, \( Q_i \) the \( i \)-th question and \( \mathcal{M}(Q_i) = (M_{i,0}, M_{i,1}) \).

Then \( Q \) is a canonical set of questions if and only if for every \( 2N \)-bit string \( a \) the intersection \( \cap_i M_{i,a_i} \) contains exactly one element.

We saw that the epistemic state corresponds to a set of answers to some questions in a canonical set, which rule out some ontic states. We call the set of ontic states that are allowed under a given epistemic state the **ontic basis** of the epistemic state \([1]\). We will often write an epistemic state as \( o_1 \lor o_2 \lor \ldots \lor o_k \) for \( o_1, \ldots, o_k \) ontic states in its ontic basis. This allows us to calculate the ontic basis from the set of answers in the following way.

**Lemma 1.2 (Ontic basis of epistemic states).** Let \( Q \) be a canonical set of \( 2N \) questions on \( N \) elementary systems, corresponding to the partitions \( \mathcal{M}(Q_i) = (M_{i,0}, M_{i,1}) \). Let \( E \subseteq \{1, 2, \ldots, 2N\} \) index a subset of these questions, and \( A_E = \{a_i\}_{i \in E} \) represent answers to this subset of questions (that is, an epistemic state, not necessarily a valid one). Then the ontic basis of this epistemic state is given by the set \( E = \cap_{i \in E} M_{i,a_i} \).

Lemmas 1.1 and 1.2 allow us to find a necessary condition for the validity of epistemic states.

**Theorem 1.3 (Dimension of valid epistemic states).** Let \( E \) be an ontic basis of a valid epistemic state on \( N \) elementary systems, which covers the answers to a subset of \( k \) out of \( 2N \) canonical questions. Then \( |E| = 2^{2N-k} \).

**Simplest example: the toy qubit** \([1]\). From the previous discussion, a state on a single elementary system needs to have an ontic basis of size either \( 2^{2-1} = 2 \) or \( 2^{2-0} = 4 \) and any permutation of 2 or 4 ontic states in the ontic basis is valid. The valid states are

\[
\begin{align*}
\text{(1.7)} & \quad \begin{array}{c|c}
\begin{array}{c}
\text{ontic states 1}
\end{array} & \begin{array}{c}
\text{ontic states 2}
\end{array}
\end{array} \\
\text{(1.8)} & \quad \begin{array}{c|c}
\begin{array}{c}
\text{ontic states 3}
\end{array} & \begin{array}{c}
\text{ontic states 4}
\end{array}
\end{array} \\
\text{(1.9)} & \quad \begin{array}{c|c}
\begin{array}{c}
\text{ontic states 5}
\end{array} & \begin{array}{c}
\text{ontic states 6}
\end{array}
\end{array} \\
\end{align*}
\]

and the non-maximal information state

\[
\begin{array}{c|c}
\begin{array}{c}
\text{ontic states 1}
\end{array} & \begin{array}{c}
\text{ontic states 2}
\end{array}
\end{array}
\]

\[
\text{(1.10)}
\]

This single system is analogous to a qubit system quantum mechanics. In particular, we can identify the six maximum information states with the following pure states of a qubit \([1]\):
where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$, and the non-maximal information state with the fully mixed qubit state $|1\rangle$:

$$(1 \lor 2 \lor 3 \lor 4) \iff \frac{1}{2}.$$  

**Valid states at the level of subsystems.** Theorem 1.3 only says something about the knowledge balance principle applied to the global state. However, for a state to be valid, the knowledge balance principle also has to be fulfilled on each subsystem. As an example consider the valid bipartite state $(1 \lor 1) \lor (1 \lor 2) \lor (2 \lor 1) \lor (2 \lor 2)$ graphically represented as

The knowledge balance principle applied to the entire state is fulfilled for any permutation of the ontic states. We can apply the permutation to the partition corresponding to the questions in the canonical set and get a new permuted canonical set of questions, because the condition in lemma 1.1 is still fulfilled. Thus, the permuted epistemic state fulfills the knowledge balance principle at the global. However, some of these permutations, like the state $(1 \cdot 1) \lor (2 \cdot 1) \lor (3 \cdot 1) \lor (4 \cdot 1)$,

are not valid: in this example, the knowledge balance principle is not fulfilled on the second system, where the local ontic state is known to be 1. This example shows that the knowledge balance principle on the entire state is permutation-symmetric, but when we require it to hold for every subsystem we break this permutation symmetry.
**Local permutations.** Nevertheless, some permutations still map valid states to valid states, even for the composed systems. These are local permutations: those that act on individual elementary subsystems, leaving the others untouched. This fact leads to an equivalence relation between states [1].

**Definition 1.4 (Equivalent epistemic states).** A permutation of a composed system is called local if it acts only on an individual elementary subsystem. Two epistemic states are equivalent if and only if there exists a local permutation that maps one to other.

Spekkens shows that this is indeed an equivalence relation [1].

**Example: two elementary systems** [1]. For a system composed of two elementary subsystems (the analogous of two qubits), the equivalence classes of states are those of the pure states

\[
|0\rangle \otimes |0\rangle, \quad \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle),
\]

and the mixed states

\[
|0\rangle \langle 0| \otimes \frac{1}{2}, \quad \frac{1}{2} \otimes |0\rangle \langle 0|,
\]

\[
\frac{1}{2}(|0\rangle \langle 0| \otimes |0\rangle \langle 0| + |1\rangle \langle 1| \otimes |1\rangle \langle 1|),
\]

and

\[
\frac{1}{4}.
\]
Inductive validity checks. The requirement that the knowledge balance principle has to apply to each subsystem leads to an inductive structure: let

\[ E_{1,2,...,N} = (a_{1,1} \cdot ... \cdot a_{1,N}) \lor (a_{2,1} \cdot ... \cdot a_{2,N}) \lor ... \lor (a_{2^k,1} \cdot ... \cdot a_{2^k,N}) \]  

be a state whose validity needs to be checked. First we need to check that the state is globally valid. Then the knowledge balance requires that each reduced state of \( N - 1 \) subsystems be a valid epistemic state too. For example the marginal over system 1 (equivalent to tracing out system 1)

\[ E_{2,...,N} = (a_{1,2} \cdot ... \cdot a_{1,N}) \lor (a_{2,2} \cdot ... \cdot a_{2,N}) \lor ... \lor (a_{2^k,2} \cdot ... \cdot a_{2^k,N}) \]

needs to be a valid state. So we check whether this reduced state is valid, first at a global level... But for this to be a valid state, each of its \( (N - 2) \)-subsystem marginals must be valid, and so on. Thus, the knowledge balance principle leads to an inductive structure, and a lengthy process to verify validity of composed states.

Note that it would not be sufficient to simply check the reduced state of all the elementary subsystems, as “illegal” correlations between those could make a locally valid state be globally invalid at any of the levels described above. We go over such an example later on, when we describe measurements; see (2.9).

1.2 Stabilizer formalism

In general it is hard to directly verify if a given epistemic state over multiple systems is valid, because of the inductive structure of epistemic states. However, Pusey [2] noticed that the epistemic restriction can be expressed in a way analogous to the quantum stabilizer formalism. It is important to note that while the quantum stabilizer formalism only allows the description of a small corner of the Hilbert space [6, 7], the toy stabilizer formalism can describe every state in the toy theory. The stabilizer approach allows us to answer questions that are left open in the previous formalism, and to verify the validity of states and operations more efficiently. The proof of the equivalence of the stabilizer and epirestricted pictures can be found in [2].

Quantum stabilizers. Let us first briefly review the quantum stabilizer formalism [7]; the analogy to the toy stabilizers will be useful ahead. Let \( 1_n \) be the \( n \times n \) identity matrix and

\[ X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \]  

the Pauli matrices. Note that they are not linearly independent as \( XZ = -iY \). The Pauli group on \( N \) qubits \( P_N \) is defined as the following set of matrices

\[ P_N = \{ \alpha M_1 \otimes ... \otimes M_N | M_i \in \{ 1_2, X, Y, Z \}, \ \alpha \in \{ 1, -1, i, -i \} \}. \]
This group is generated by \( \{i\mathbb{1}, X_k, Z_k\}_k \) where \( X_k, Z_k \) act on the \( k \)-th qubit with the \( X \) or \( Z \) operator respectively and trivially on the others.

A stabilizer group \( S \) is any subgroup of \( P_N \) satisfying the condition that all elements commute and that \( -\mathbb{1} \) is not included. It can be shown that a stabilizer subgroup can can have at most \( N \) independent elements, or \( 2^N \) elements in total. We then define the density matrix associated with \( S,^3 \)

\[
\rho_S = 2^{-N} \sum_{g \in S} g.
\]

This is a uniform mixture over \( \frac{2^N}{|S|} \) states. In particular, it is a pure state if \( |S| = 2^N \).

Often when the stabilizer formalism is used, one is not directly interested in \( \rho_S \) but rather in the subspace \( V_S \) of states \( |\psi\rangle \) that are stabilized by \( S \), i.e. the vectors \( |\psi\rangle \) such that all \( g \in S \) fulfill \( g|\psi\rangle = |\psi\rangle \). Note that \( \rho_S \) is a uniform mixture of the basis elements of \( V_S \).

**Toy stabilizers.** Now let us relate the formalism above with the language of the original toy theory, following Pusey [2]. We associate the four elementary ontic states \( 1, 2, 3, 4 \) of the previous subsubsection with the vectors \( e_1 = (1, 0, 0, 0)^T, e_2 = (0, 1, 0, 0)^T, e_3 = (0, 0, 1, 0)^T, e_4 = (0, 0, 0, 1)^T \).

We define the toy Pauli matrices as

\[
\mathcal{X} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}, \\
\mathcal{Y} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}, \\
\mathcal{Z} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}.
\]

Similarly to the Pauli matrices, these matrices are not independent as \( \mathcal{Y} = \mathcal{Z}\mathcal{X} \). With these toy Pauli matrices, we define the toy Pauli group

\[
G_N = \{ \alpha \ p_1 \otimes ... \otimes p_N \ | \ p_i \in \{\mathbb{1}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}, \ \alpha \in \{1, -1\} \}.
\]

The set \( \{\mathcal{X}_k, \mathcal{Z}_k, -1\ | k \in 1, ..., N\} \) generates \( G_N \), where \( \mathcal{X}_k \) acts as \( \mathcal{X} \) on the \( k \)-th elementary system and \( \mathcal{Z}_k \) is defined analogously. Each ontic state \( o_N = e_{i_1} \otimes e_{i_2} \otimes ... \otimes e_{i_N} \) is in a different combination of eigenspaces of generators of \( G_N \). Hence, the ontic state is fully characterized by the combination of eigenspaces it lies in.

**“Commutation” relations.** In the quantum stabilizer formalism, commutation relations between elements of the Pauli group play a large role. In the toy theory though, all elements of \( G_N \) commute, as they correspond to diagonal matrices. One can get around this aggravation by imposing “commutation relations” by hand, in analogy to the quantum case. That is, we define the map \( m : G_N \rightarrow P_N \) between the toy and quantum Pauli groups such that \( m(\mathcal{X}_k) = X_k, m(\mathcal{Z}_k) = Z_k, m(-\mathbb{1}_{2N}) = -\mathbb{1}_N \) and \( m(gh) = m(g)m(h) \). We further define a binary relation \( \sim \) between two elements \( g, h \in G_N \) such that \( g \sim f \) if and only if \( m(g) \) and \( m(f) \) commute. We forsake rigour of notation for convenience, and say that in this case \( g \) and \( f \) “commute”.

\(^3\)The normalization factor of \( 2^{-N} \) comes from \( \text{tr}(\rho_s) := 2^{-N} \sum_{g \in S} \text{Tr}(g) = 2^{-N} \text{tr}(\mathbb{1}_{2N}) = 1 \), as all Pauli operators except the identity have trace zero.
Toy stabilizer groups. In analogy to the quantum stabilizer formalism, we can define a stabilizer group $S$ as a commuting subgroup of $G_N$ that does not contain $-\mathbb{1}_{2N}$. The epistemic state associated with a stabilizer group $S$ is given by the union of all ontic states stabilized by operators in the group,

$$E_G = \bigvee \{ o_N = e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_N} : g o_N = o_N, \forall g \in S \}.$$  

For example, in an elementary system, the epistemic state $1 \lor 2$ (in stabilizer notation $e_1 \lor e_2$) corresponds to the stabilizer group span$\{Z\}$. The maximal amount of independent generators of a stabilizer group $S$ is $N$. Epistemic states with such a stabilizer group are precisely the pure states.

Relation to original formalism. We can regard each toy Pauli operator as a question where the two eigenspaces (for eigenvalues $+1$ and $-1$) can be identified with the partition over states corresponding to the question. For example, $Z$ corresponds to the question “Is the ontic state in $\{1, 2\}$?” The fact that there are at most $N$ independent generators in a stabilizer group agrees with the knowledge balance principle, which requires the maximal amount of questions an observer can answer to be $N$, half of the size of the canonical set. The commutation requirement ensures that the knowledge balance principle is also fulfilled on subsystems.

1.3 Generalizing to arbitrary dimensions

Up to now we only considered the toy theory for systems analogous to qubits. However, the toy theory can be generalized to arbitrary $d$-level and continuous systems. The generalization of Spekkens’ toy theory is inspired by canonical quantization [8], where a set of observables is jointly measurable if they all commute, this time relative to the Poisson bracket, instead of the usual matrix commutator [3].

Toy position and momentum. To formalize this notion, we represent a system’s $n$ degrees of freedom in a language of the “positions” $q_1, \ldots, q_n$ and conjugated “momenta” $p_1, \ldots, p_n$. The associated phase space is denoted by $\Omega = \mathbb{R}^{2n}$ in the continuous case and $\Omega = \mathbb{Z}_d^{2n}$ in the discrete case. Each point in the phase state $m = (q_1, p_1, \ldots, q_n, p_n)^T \in \Omega$ is an ontic state of the system. These variables do not have to correspond to actual positions or momenta of particles.

Quadrature variables. Observables are represented by functionals on the state space, $f : \Omega \to \mathbb{R}$ or $\mathbb{Z}_d$. Importantly, we will be looking at quadrature variables: linear combinations of position and momentum variables,

$$f(m) = a_1 q_1 + b_1 p_1 + a_2 q_2 + b_2 p_2 + \cdots + c.$$  

Note that for $d$ not prime $\Omega = \mathbb{Z}_d^{2n}$ is not a vector space, because $\mathbb{Z}_d$ is not a field. However, we can see $\mathbb{Z}_d^{2n}$ as a module over the ring $\mathbb{Z}_d$. In this case special care is needed, as not every linear algebra result also holds for $\mathbb{Z}_d^{2n}$ [3]. In contrary to past research [4], in this work we were able to unify and simplify the treatment of the prime and non-prime case.

4
In the discrete case, this expression has to be understood within mod $d$. Without loss of generality $c$ can be set to zero, because if $f(m) - c$ is known then $f(m)$ is also known. Therefore, each quadrature variable can be associated with a vector $f \in \Omega$ with the coefficients $a_1, b_1, \ldots$ and its evaluation on an ontic state $m$ can be compactly written as

$$f(m) = f^T m.$$ 

For example, the quadrature variable $q_1 + p_1$ would be associated with the vector $f^T = (1, 1, 0, \ldots, 0)$.

**Poisson bracket.** The next step is to define the Poisson bracket of two functionals $f$ and $g$. In the continuous case, it follows the definition of classical mechanics [3],

$$[f, g](m) = \sum_{i=1}^{n} \left( \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right),$$

where $q_i, p_i$ are the vectors in phase space where all entries are zero except the position or momentum of the $i$th degree of freedom, which is $q_i$ or $p_i$ respectively. In the discrete case, the Poisson bracket can be defined with differences in the respective modulo space [3]:

$$[f, g](m) = \sum_{i=1}^{n} \left( f(m + q_i) - f(m) \right) \left( g(m + p_i) - g(m) \right) - \left( f(m + p_i) - f(m) \right) \left( g(m + q_i) - g(m) \right) \text{mod } d.$$ 

**Epistemic restriction.** The complete epistemic restriction in then given by the principle of classical complementarity:

"The valid epistemic states are those wherein an agent knows the values of a set of quadrature variables that commute relative to the Poisson bracket, and is maximally ignorant otherwise." [3]

Here, “maximal ignorance” means that there is a uniform probability over all other values of variables. It can be shown that this complementarity principle requires observables to be linear, i.e. quadrature observables.

**Sympletic inner product.** Representing quadrature observables as a vector $f$ allows for a different expression of the commutation rules. We calculate the Poisson bracket for both the continuous and discrete case, where the sum has to be understood in mod $d$ in the discrete case:

$$[f, g] = \sum_{i=1}^{n} f_{2i-1} g_{2i} - f_{2i} g_{2i-1}$$  \hspace{1cm} (1.22)
where $f_j$ denotes the $j$th entry of $f$. We can rewrite this as a symplectic inner product of the two observables $f$ and $g$,

$$[f, g] = f^T J g = :\langle f, g \rangle,$$

(1.23)

where we defined the matrix

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \quad (1.24)$$

Therefore two variables commute if and only if they are orthogonal with respect to this symplectic inner product [3].

**Isotropic subspaces of compatible observables.** When an observer knows the value of variables $f_1, \ldots, f_k$, they automatically know the values of all variables in the span $(f_1, \ldots, f_k)$. This is because the observer can deduce the value of $af_i + bf_j$ for known $a, b \in \mathbb{Z}_d$ or $\mathbb{R}$ if $f_i$ and $f_j$ are known. Therefore, the known variables form a subvector space (in the continuous case or the discrete one with $d$ prime) or submodule (in the case where $d$ not prime) $V = \langle f_1, \ldots, f_k \rangle$ of $\Omega$. This subvector space or submodule is called isotropic, if it satisfies the condition

$$\forall f, g \in V : \langle f, g \rangle = 0. \quad (1.25)$$

A vector space or module spanned by commuting elements is isotropic by definition. The maximal dimension of an isotropic vector space of $\Omega$ is $n$, the number of degrees of freedom. Therefore, the maximal amount of variables that can be known is $n$, which is exactly half the number of linearly independent variables, echoing the knowledge balance principle [3].

**Valuation vector.** For each possible value assignment to a minimal generating set $\{f_1, \ldots, f_k\}$ of an isotropic subvector space or submodule corresponds to a vector $v \in \Omega$, called the valuation vector. For example, if $V = \langle \left(\begin{array}{c} 1 \\ 1 \end{array}\right) \rangle$ and we know that $p_1 + q_1 = 0$, the vector $v$ is then $\left(\begin{array}{c} 1 \\ -1 \end{array}\right)$. This vector is not unique: $v = \left(\begin{array}{c} 2 \\ -2 \end{array}\right)$ would also be an equivalent choice for the valuation vector, as the variables in $V$ have the same valuation for both vectors.

**Epistemic states.** The classical complementarity principle requires that an observer be maximally ignorant for all the other variables. Therefore, the epistemic state is a probability distribution over phase space where we assign equal probability to all ontic states $m \in \Omega$ which are compatible with the knowledge of the observer. More formally, an ontic state $m \in \Omega$ is compatible with the valuation vector $v$ if for all $f \in V$

$$f^T m = f^T v. \quad (1.26)$$
This condition can be rewritten as

$$f^T(m - v) = 0. \quad (1.27)$$

For all ontic states that fulfill the above condition it must hold that

$$m - v \in V^\perp = \{ m' \in \Omega | \forall f \in V \ f^T m' = 0 \}. \quad (1.28)$$

Thus, we can write the set of compatible ontic states as the affine subvector space or submodule

$$V^\perp + v = \{ m' \in \Omega | m' - v \in V^\perp \}. \quad (1.29)$$

This forms an equivalence class: the set of ontic states that cannot be operationally distinguished through observations in $V$, which all correspond to a single epistemic state. We summarize the above in the following definition for epistemic states.

**Definition 1.5 (Epistemic state (generalized)).** Let $V = \langle f_1, \ldots, f_k \rangle$ be the isotropic space/module generated by commuting observables $f_1, \ldots, f_k$, that is a set of mutually knowable observables. Let $v$ be one possible valuation vector for these observables. Together, $v$ and $V$ form an epistemic state $(V, v)$. The set of ontic states compatible with that valuation is given by $V^\perp + v$. Intuitively, this corresponds to the set of possible ontic states, given that we measured the observables in $V$ and obtained valuation $v$.

The resulting probability distribution of possible ontic states for a given epistemic state, assuming classical complementarity over phase space $\mu_{V,V}$ is then given by

$$\mu_{V,V}(m) = \frac{1}{N_V} \delta_{V^\perp + v}(m) \quad (1.30)$$

with

$$\delta_{V^\perp + v}(m) = \prod_{f \in V} \delta(f^T m - f^T v) \quad (1.31)$$

and $\frac{1}{N_V}$ a normalization constant. In the discrete case $\delta$ is the Kronecker delta and in the continuous case it is the Dirac delta function [3].

**Pure and mixed states.** We call a state maximal information or pure if the corresponding isotropic vector space $V$ has the maximal dimension $n$, otherwise we call the state mixed [3].

**Relation to stabilizer formalism.** The above definition of states is reminiscent of the stabilizer formulation. In the case of $d = 2$, we can consider the functionals $f \in V$ as the stabilizers, and the valuations as a generalization of the sign of the stabilizer. For example, if we know the value of the functional $f(m) = q$ to be 1, then we know that all possible ontic states are of the form $\binom{q}{p} = \binom{1}{b}$.
with \( b \in \{0, 1\} \). We can assign the four possible ontic states \( (0)_0, (0)_1, (1)_0, (1)_1 \) in this formalism to the ontic states 1, 2, 3, 4 in the previous formalism. Then the epistemic state is \( 2 \lor 4 \) which is stabilized by \(-\mathcal{X}\). If the valuation of \( f(m) = q \) were 0, then the state would be stabilized by \( \mathcal{X} \).

**Definition 1.6** (Correspondence between stabilizer and general formalisms). Let \( d = 2 \). We say a stabilizer \( g = \alpha p_1 \otimes \ldots \otimes p_n \) corresponds to an observable \( f \) (or vice-versa) if the entries of the observable \( \{f_{2j-1}, f_{2j}\}_{j \in \{1, \ldots, n\}} \) are related to the stabilizer’s individual qubits’ observables \( \{p_j\}_{j \in \{1, \ldots, n\}} \) as

\[
\begin{align*}
(f_{2j-1}, f_{2j}) = (0, 0) & \iff p_j = 1_2, \quad (1.32) \\
(f_{2j-1}, f_{2j}) = (0, 1) & \iff p_j = Z, \quad (1.33) \\
(f_{2j-1}, f_{2j}) = (1, 0) & \iff p_j = X, \quad (1.34) \\
(f_{2j-1}, f_{2j}) = (1, 1) & \iff p_j = Y, \quad (1.35)
\end{align*}
\]

and \( \alpha \in \{-1, 1\} \).

In addition, we say that a stabilizer state \( S = \langle g_1, \ldots, g_N \rangle \) corresponds to a generalized state \((V, v)\) (or vice-versa) if and only if for each observable \( f \in V \) there is a corresponding stabilizer \( g \in S = \alpha p_1 \otimes \ldots \otimes p_n \) satisfying \( \alpha = (-1)^{f^T v} \) (that is, \( f^T v = 0 \iff \alpha = 1 \) and \( f^T v = 1 \iff \alpha = -1 \)).

**Lemma 1.7** (Correspondence between generalized and stabilizer formalisms is sound). If \( d = 2 \), each general state \((V, v)\) has a corresponding stabilizer state and each stabilizer state \( S \) has a corresponding general state. Furthermore, if variables \( f \) and \( g \) correspond to stabilizers \( g', g'' \) respectively, then \( f + g \) corresponds to \( g'g'' \).

**Remark.** It has been stated that for a valid state \((V, v)\), the valuation \( v \) should be chosen in \( V \) \([3, 4]\); however, we found this to be unjustified. If we consider the case \( d = 2 \) and the isotropic space \( V = \langle f \rangle \), with \( f = p + q = (1)_1 \), then the only two vectors in \( V \) are \( v_0 = (0)_0 \) and \( v_1 = (1)_1 \), both of which have valuation 0, as \( f^T v_j \mod 2 = 0 \), \( j = 0, 1 \). There is no reason why there shouldn’t exist a state satisfying \( p + q = 1 \), but the only valuation vectors which produce this outcome (like \( (1)_0 \)) lie outside of \( V \). Thus, if we restrict the valuation vector \( v \) to lie in \( V \), we unnecessarily limit the possible number of outcomes of observables.
2 Measurements

2.1 Original toy theory

A measurement partitions the ontic state space into valid epistemic states, called the measurement basis.\(^5\) The outcome of the measurement is determined by the position of the ontic state and is the epistemic state compatible with the ontic state [1]. Let us illustrate the measurement process with an example. Consider the following epistemic state:

\[
\begin{array}{ccc}
3 & 3 & 4 \\
3 & 3 & 4 \\
1 & 1 & 2 \\
1 & 1 & 2 \\
\end{array}
\]  

and the measurement (where different numbers correspond to the epistemic states of the measurement basis):

\[
\begin{array}{cccc}
3 & 3 & 4 & 4 \\
3 & 3 & 4 & 4 \\
1 & 1 & 2 & 2 \\
1 & 1 & 2 & 2 \\
\end{array}
\]  \hspace{1cm} (2.2)

The only results compatible with the epistemic state are 4 and 1. Suppose the ontic state is \((4 \times 4)\); then the measurement outcome has to be 4.

**Measurement update rule.** Let us consider two cases: a measurement of a single system \((N = 1)\), and a measurement on multiple systems \((N > 1)\).

\(N = 1.\) Let us assume the system is in the epistemic state

\[
(1 \lor 2) = \begin{array}{cc}
3 & 3 \\
\end{array}
\]  

we measure in the basis \((a \text{ and } b \text{ denote two different basis states})

\[
\{(1 \lor 3), (2 \lor 4)\} = \begin{array}{cccc}
a & b & a & b \\
\end{array}
\]  

and the outcome is \(a\). From this measurement result, we can conclude that the ontic state before the measurement was 1. This seems like a violation of the knowledge balance principle. However, the knowledge balance principle does not say anything about the knowledge an observer can have about the past state of a system. In fact, only the updated epistemic state needs to fulfill the knowledge balance principle. As in quantum mechanics, we require that repeated measurement yields the same result.

\(^5\)We will see later that the converse can be problematic: not all partitions result in intuitive measurements.
result. These conditions can be fulfilled through the following principle [1]. A measurement gives an unknown disturbance to the ontic state: either the permutation $1 \leftrightarrow 2$ is applied or no permutation is applied. Under this update rule, the post-measurement epistemic state is

$$(1 \lor 3) = \begin{array}{cc}
1 & 3 \\
1 & 3 \\
1 & 3 \\
1 & 3 \\
\end{array}$$ or $$(2 \lor 4) = \begin{array}{cc}
1 & 4 \\
1 & 4 \\
2 & 2 \\
2 & 2 \\
\end{array}$$

(2.5)

**N > 1.** In the case of multiple systems the update rule is more complicated. If the measurement is a conjunction of measurements on a single subsystem the update rule for one subsystem can be used. As an example consider the measurement in eq. (2.2) applied to the state eq. (2.1) and assume the system is in the ontic state $4 \times 4$. The measurement can be decomposed into the measurement on the first system

$$
\begin{array}{cccc}
1 & 1 & 2 & 2 \\
1 & 1 & 2 & 2 \\
1 & 1 & 2 & 2 \\
1 & 1 & 2 & 2 \\
\end{array},
$$

(2.6)

and a measurement on the second system

$$
\begin{array}{cccc}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
2 & 2 & 2 & 2 \\
2 & 2 & 2 & 2 \\
\end{array}.
$$

(2.7)

The ontic state requires the measurement outcome of the first measurement to be 2. This excludes that $1 \cdot 1$ and $2 \cdot 2$ are the ontic state and the permutation $3 \leftrightarrow 4$ on the first system is randomly applied or not. Therefore, the updated state after the first measurement is

$$
\begin{array}{cc}
1 & 3 \\
1 & 3 \\
2 & 2 \\
2 & 2 \\
\end{array}.
$$

(2.8)

The second measurement will have the outcome 1 and the permutation $3 \leftrightarrow 4$ is randomly applied or not on the horizontal system. However, this permutation does not change the epistemic state and the epistemic state does not change after this measurement.

**Restriction on valid states.** This update rule leads to a restriction on which epistemic states are valid. Consider the state
At first glance one would think that this state was a valid epistemic state, as all marginals are valid and the knowledge balance principle is fulfilled on the bulk. However, if one measures the vertical system in \(((1 \lor 3), (2 \lor 4))\) and the outcome is \((2 \lor 4)\) this update rule says the state should be updated to

\[
\begin{array}{cc}
1 & 0 \\
0 & 1 \\
\end{array}
\]

which violates the knowledge balance principle. One could also argue that then it should not be allowed to measure the vertical system in \(((1 \lor 3), (2 \lor 4))\). However, for each measurement on a single system one can find a state with valid marginals that would violate the knowledge balance principle in a similar manner. Thus, one agent could not measure the system in their possession. For this reason, we do not forbid the measurement but instead say that the state is invalid [1].

**Coarse-grained measurements.** We can distinguish between maximal and non-maximal information measurements. A maximal information measurement is a measurement where the space is partitioned into maximal information (or pure) states. In this case the state is updated to the state in the measurement basis which the outcome corresponds to. A non-maximal information measurement, or coarse-grained measurement, has at least some states in the measurement basis are non-maximal information (or mixed) states. This means there may be multiple valid epistemic states which yield the same result under repeated measurement. In this case, the measurement update rule cannot be uniquely defined by the conditions we imposed above. In quantum mechanics, a similar problem also appears and the most common update rule is to project the original state into the subspace associated with the result. In the toy theory, an analogous update rule can be defined using the fidelity [1].

**Definition 2.1 (Fidelity [1]).** The fidelity of two epistemic states \(E_a, E_b\) is the classical fidelity of the uniform probability distributions \(P_a, P_b\) over the ontic basis of the states \(E_a, E_b\)

\[
F(E_a, E_b) = \sum_{o \in O} \sqrt{P_a(o)} \sqrt{P_b(o)}.
\]

Spekkens defines the updated state as the epistemic state with the highest fidelity with the pre-measurement state that will lead to the same output under repeated measurement [1].
Example. For example, consider the measurement

\[
\begin{array}{cccc}
1 & 1 & 3 & 2 \\
1 & 1 & 2 & 3 \\
3 & 2 & 1 & 1 \\
2 & 3 & 1 & 1 \\
\end{array}
\]

on state .

(2.12)

Suppose that the measurement outcome is 1. The possible post measurement states are those with ontic support in that partition, which are the valid epistemic states

\[
\begin{array}{cccc}
\text{state 1} & \text{state 2} & \text{state 3} & \text{state 4} \\
\text{state 5} & \text{state 6} & \text{state 7} & \text{state 8} \\
\end{array}
\]

where the first is a mixed state, and all others are pure. Indeed the first state can be seen as either a mixture of the second and third or of the fourth and fifth states (see section 3 for a definition of mixtures). According to Spekkens’ update rule we have to choose the state with maximal fidelity to the pre-measurement state, which is the third one.\(^6\)

2.2 Stabilizer formalism

Quantum stabilizer measurement. Suppose we want to measure the Pauli operator \(g\) on a quantum state by the stabilizer group \(S\). If \(\pm g \in S\) the outcome is deterministically \(\pm 1\), otherwise with equal probability the outcome is \(\pm 1\). In the latter case, the stabilizer group needs to be updated. As the measurement determined the value of \(g\), \(g\) itself is added to the stabilizer group. However, this could lead to an invalid stabilizer. Therefore, if there is at least one stabilizer \(h\) in the generating set of \(S\) that anticommutes with \(g\), we multiply all other anticommuting generators with \(h\) and replace \(h\) with \(g\). For example, let \(S = \text{span}\{X_1, X_2\}\) be a stabilizer group corresponding to the state \(|+\rangle|+\rangle\). Let us measure the observable \(Z_1\). Given that \(Z_1 \notin S\) the measurement outcome is random; suppose that it is \(-1\). Since \(X_1\) anticommutes with \(Z_1\), we remove \(X_1\) and replace it with \(-Z_1\). Thus, after the measurement the state has the stabilizer group \(S = \text{span}\{-Z_1, X_2\}\), corresponding to \(|1\rangle|+\rangle\) [7].

Toy theory: binary measurements. In the epirestricted description of the toy theory we have argued that measurements are a partition of all ontic states into valid epistemic states. A toy observable \(g\) then partitions the ontic states into the epistemic states \(\text{span}\{g\}\) and \(\text{span}\{-g\}\). The same procedure as in the quantum stabilizer formalism can be applied to update the epistemic state

\[^6\text{As we will see when we generalize the measurement update for arbitrary dimensions (Theorems 2.2 and 2.3), there is nothing special about fidelity as a particular measure of overlap between two states, at least where this task in concerned. Other measures (with different powers of the distributions instead of a square root) may lead to the same choice of final state.}\]
We list all generators of the original state, remove the first generator $h$ that does not commute with $g$, multiply all other generators that do not commute with $g$ by $h$, and finally add either $g$ or $-g$ (depending on the outcome) to the list of generators.

**Measurements on two elementary systems.** However, not all partitions are of the form $\text{span}\{g\}$ and $\text{span}\{-g\}$, in particular when a measurement has more than two possible outcomes. Consider the partition:

\[
\begin{array}{cccc}
1 & 1 & 3 & 2 \\
1 & 1 & 2 & 3 \\
3 & 2 & 4 & 4 \\
2 & 3 & 4 & 4
\end{array}
\] (2.14)

This partition corresponds to the following stabilizer states:

\[
\begin{align*}
\text{span}\{\mathcal{X}_1\mathcal{X}_2, \mathcal{Z}_1\mathcal{Z}_2\}, & \quad \text{span}\{-\mathcal{X}_1\mathcal{X}_2, \mathcal{Z}_1\mathcal{Z}_2\}, & \quad \text{span}\{\mathcal{Z}_1, -\mathcal{Z}_2\}, & \quad \text{span}\{-\mathcal{Z}_1, \mathcal{Z}_2\}.
\end{align*}
\] (2.15)

We can implement this measurement by first measuring $\mathcal{Z}_1\mathcal{Z}_2$, which corresponds to the partition:

\[
\begin{array}{cccc}
1 & 1 & 2 & 2 \\
1 & 1 & 2 & 2 \\
2 & 2 & 1 & 1 \\
2 & 2 & 1 & 1
\end{array}
\] (2.16)

If the outcome is 1, we measure $\mathcal{X}_1\mathcal{X}_2$, otherwise we measure the stabilizer $\mathcal{Z}_1$. In this way, we can implement more complicated partitions.

**Measurements on more systems.** However, the above procedure does not work in general for more than two elementary systems. For example, the partition

\[
\begin{align*}
\text{span}\{\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{X}_3\}, & \quad \text{span}\{-\mathcal{Z}_1, \mathcal{X}_2, \mathcal{Z}_3\}, & \quad \text{span}\{\mathcal{X}_1, -\mathcal{Z}_2, -\mathcal{Z}_3\}, & \quad \text{span}\{\mathcal{Z}_1, \mathcal{Z}_2, -\mathcal{X}_3\}, \\
\text{span}\{-\mathcal{Z}_1, -\mathcal{X}_2, \mathcal{Z}_3\}, & \quad \text{span}\{-\mathcal{X}_1, -\mathcal{Z}_2, -\mathcal{Z}_3\}, & \quad \text{span}\{\mathcal{Z}_1, -\mathcal{Z}_2, \mathcal{Z}_3\}, & \quad \text{span}\{-\mathcal{Z}_1, \mathcal{Z}_2, -\mathcal{Z}_3\}
\end{align*}
\] (2.17)

is a valid measurement but does not have a non-trivial toy observable that is common to all the eight states.\(^7\) In cases like this, there is no easy way to find the post-measurement state. We will see in the next section that this is because maybe not all partitions should be considered valid measurements.

\(^7\)Interestingly, the ontic states are partitioned into product states, but they cannot be distinguished using only local measurements. In quantum mechanics this property is known as *non-locality without entanglement* [9]; however, the toy theory is a local hidden variable theory, indicating that non-locality is not necessary [2].
2.3 Generalizing to arbitrary dimensions

Valid measurements. The set of valid measurements must respect the principle of classical complementarity. Therefore, we can only jointly measure variables that form an isotropic subspace or submodule $V_\pi$ with outcomes $v_\pi \in \Omega$, in analogy to commuting operators in quantum theory. The probability to get outcome $v_\pi$ if the system is in an ontic state $m$ is given by

$$\xi(v_\pi|m) = \delta_{v_\pi + v_\pi}(m).$$

Intuitively, this means that we can only obtain measurement outcomes that are compatible with the ontic state of the system. We can denote a measurement $V_\pi$ and its outcome $v_\pi$ by the pair $(V_\pi, v_\pi)$. With this conditional probability distribution we can calculate the probability for a measurement outcome given the epistemic state $(V, v)$,

$$P(v_\pi|(V, v)) = \sum_{m \in \Omega} \xi(v_\pi|m) \mu(V, v)(m).$$

Measurement update. The discussion above allows to find the probability for measurement outcome, but does not tell us how to update a state $(V, v)$ after the measurement $V_\pi$ with the outcome $v_\pi$. In the past, there was an attempt at defining a measurement update which differentiated between the prime and non-prime case [4]. Here we unify these two cases and simplify the treatment of the measurement update. Starting from the principle that repeated measurements should result in the same outcome, we obtain the following result (proof in appendix B.2).

Theorem 2.2 (Measurement update rule (generalised)). When the epistemic state $(V, v)$ is subjected to a measurement $V_\pi$, and an outcome $v_\pi$ is obtained, the epistemic state is updated to $(V', v')$, where

$$V' = V_\pi \oplus V_{\text{commute}},$$

$$v' \in (V_\pi^\perp + v_\pi) \cap (V_{\text{commute}}^\perp + v),$$

and

$$V_{\text{commute}} = \{ f \in V : [f, f_\pi] = 0, \forall f_\pi \in V_\pi \} \subseteq V$$

is the subset of $V$ that commutes with $V_\pi$.

Simple example. Let us consider the state

$$\left( V = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle, v = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) \iff \begin{array}{c} \text{a} \\ \text{a} \\ \text{b} \\ \text{b} \end{array}$$

and the measurement

$$V_\pi = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle \iff \begin{array}{c} \text{a} \\ \text{a} \\ \text{b} \\ \text{b} \end{array}.\)
Then, $V_{\text{commute}}$ is given by $\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rangle$, as no non-zero vector in $V$ commutes with a vector in $V_\pi$. Suppose we get the measurement outcome $v_\pi = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Therefore, the updated state is

$$(V_\pi \oplus V_{\text{commute}} = V_\pi, v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}) \iff \quad (2.25)$$

**Complex example.** Even though the last measurement we considered in the stabilizer formalism cannot be represented by an isotropic subspace $V_\pi$ with outcomes $v_\pi \in \Omega$, we can still describe the update of state after the measurement, as it is a partition of state space. Let the outcome of the measurement be the state corresponding to $\text{span}\{Z_1, Z_2, X_3\}$:

$$W = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$  \quad (2.26)

Then, we can use the before-defined measurement update, as the measurement update is just dependent on the outcome state. Let us assume that the state before the measurement is

$$V = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$  \quad (2.27)

The vector space $V_{\text{commute}}$ is given by

$$V_{\text{commute}} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$  \quad (2.28)

and the vector $v'$ can be chosen to be $v$. To sum up, the update state is

$$V = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$  \quad (2.29)
Relation to original formalism. The measurement in the general formalism suffers from the same problem as the measurement in the stabilizer formalism; that is, not all partitions of the ontic space can be formalized as a measurement. For the partitions that can be seen as a measurement $V_\pi$, the update rule is the same as in the original formulation. The generalized approach of this section is to be understood as more fundamental, and the original epirestricted formalism as a special case for small dimensions $^8$. Therefore, not all partitions should be considered valid measurements. For example, it is not clear whether allowing for all partitions as measurements would result in a theory equivalent to a quantum sub-theory, like stabilizer quantum mechanics. At the very least, we would no longer be able to interpret it as measuring a set of commuting observables $f_1, \ldots, f_n$.

**Theorem 2.3** (Generalised measurement update rule reduces to the original formalism). *For the case $d = 2$, the measurement update rule of 2.2 reduces to the measurement update rule in the original formalism.*

Coarse-grained observables. If $d$ is not prime and the greatest common divisor of the components $f_i$ of $f$ is not 1, the variable cannot attain all values in $\mathbb{Z}_d$. We call such variables *coarse-grained*. For example, consider the case $d = 4$ and $V = \langle (\frac{2}{2}) \rangle$. Then there exist no two numbers $a, b \in \mathbb{Z}_4$ such that $\left(\begin{array}{c} 2 \\ 2 \end{array}\right)^T \left(\begin{array}{c} a \\ b \end{array}\right) = 1$ [4]. Note that if the greatest common divisor of the components $f_i$ of an observable $f$ is 1 then such a variable can attain all values by Bézout’s lemma [4, 10]. We call these variables *fine-grained*. In the previous formulations the coarse-grained variables caused a problem when defining the measurement update rules [4]. However, in our more general formulation given above, coarse-grained variables are not problematic and can be observed.

$^8$This view is also supported by Spekkens (private correspondence).
3 Superpositions and mixtures

A key feature of quantum theory is the existence of coherent superpositions and incoherent mixtures of pure states. These can also be found in Spekkens’ toy theory, but one needs to be careful: if we mix or superpose arbitrary states, we risk violating the knowledge balance principle. Examples and intuitions for these restrictions are given in section 3.3 (generalized theory).

Before we dig in, we would like to remark on two aspects of superpositions in quantum theory. The first one is that any state can be seen as either a diagonal state or a superposition, depending on the choice of basis; this aspect carries over to the toy theory, where every state can be seen as a superposition of elements of a different basis. The second aspect is that any number of quantum states can be superposed with any coefficients (provided we normalize the resulting state); for example $3\ket{0} + 2i\ket{+}$ is a valid, if unnormalized, quantum state. This is the aspect of superposition that is not always true in the toy theory, precisely because arbitrary superpositions could violate the knowledge-balance principle. The same issue emerges for probabilistic mixtures of states.

3.1 Original toy theory

Superpositions. In the original formulation, there is no definition for superposition of epistemic states on more than one system. For a single system, four different coherent superpositions can be defined, under some constraints on the pure states involved [1].

Definition 3.1 (Superpositions [1]). Let $(a \lor b)$ and $(c \lor d)$ be two epistemic states with disjoint ontic bases and written such that $a < b$ and $c < d$. Then, we can define four coherent superpositions

\[
(a \lor b) +_1 (c \lor d) := (a \lor c) \\
(a \lor b) +_2 (c \lor d) := (b \lor d) \\
(a \lor b) +_3 (c \lor d) := (b \lor c) \\
(a \lor b) +_4 (c \lor d) := (a \lor d).
\]

Quantum analogy. Comparing the resulting state of superposition with the analogous qubit state it becomes clear that these four different superpositions correspond to qubit superpositions $|a\rangle + e^{i\phi}|b\rangle$ with different relative phases $\phi$ [1]

\[
+_1 \iff \phi = 0 \\
+_2 \iff \phi = \pi \\
+_3 \iff \phi = \pi/2 \\
+_4 \iff \phi = 3\pi/2
\]

As we will see later, in the stabilizer formalism coherent superpositions can be defined for certain pairs of states composed of an arbitrary number of elementary systems [2].

9As we will see later, in the stabilizer formalism coherent superpositions can be defined for certain pairs of states composed of an arbitrary number of elementary systems [2].
Curiously, the analogy between the qubit and the single system epistemic state breaks down under superpositions [1]. Consider the superposition

\[(1 \lor 3) +_3 (2 \lor 4) = (2 \lor 3) \iff |+i\rangle.\]

Therefore, the analogous qubit superposition gives a different result[1]:

\[|+\rangle + e^{i\pi/2}|-\rangle = e^{i\pi/4}|-i\rangle.\]

**Mixtures.** Incoherent mixtures can be defined for any number of subsystems, but there is still a requirement on which states can be mixed [1]:

**Definition 3.2 (Mixtures [1]).** Let \(E_a\) and \(E_b\) be two epistemic states with disjoint ontic bases \(O_a\) and \(O_b\), such that the union of the ontic basis is a valid epistemic state, then the mixture \(E_a + E_b\) between these two states is the state with the ontic basis \(O_a \cup O_b\).

From this definition two questions arise. Firstly, it is clear from this definition that any state that can be written as the mixture of pure states is a mixed state. However, the converse is not obviously true: are all toy mixed states mixtures of pure states? We will show in the stabilizer formalism that this is indeed true. Similarly to quantum mechanics, this decomposition is not unique as \((1 \lor 2 \lor 3 \lor 4)\) has multiple decompositions:

\[(1 \lor 2 \lor 3 \lor 4) = (1 \lor 2) + (3 \lor 4)\]
\[(1 \lor 2 \lor 3 \lor 4) = (1 \lor 3) + (2 \lor 4)\]
\[(1 \lor 2 \lor 3 \lor 4) = (1 \lor 4) + (2 \lor 3).\]

Secondly, the definition requires that the union of two ontic bases forms again an ontic basis, but it is in general hard to check if a state is valid due to the inductive nature of the epistemic states. The stabilizer formalism can give us a simpler condition [1, 2].

### 3.2 Stabilizer formalism

**Mixtures.** Mixtures can only be defined for states where the corresponding stabilizer groups are rephasings of each other. The stabilizer group \(S'\) is a rephasing of \(S\) if and only if for any \(g \in S, \pm g \in S'\). For two stabilizer groups \(S, S'\) which are rephasings of each other the mixture \(\frac{1}{2}(\rho_S + \rho_{S'})\) is given by the state corresponding to the stabilizer group \(T = S \cap S'\). This definition allows us to state the following theorem (proof in the appendix).

**Theorem 3.3 (Mixed states as convex combinations).** Any mixed state with stabilizer group \(S = \text{span}\{g_1, \ldots, g_{N-k}\}, k \leq N\) can be written as the mixture of \(2^k\) pure states.
Example. Consider the stabilizer groups \( S = \text{span}\{-Z_1\} \) and \( S' = \text{span}\{Z_1\} \), corresponding to the states \(|1\rangle\) and \(|0\rangle\) respectively. The state generated by the rule above has the stabilizer group \( T = \{1_2\} \), which corresponds to the state \( \frac{1}{\sqrt{2}}(|0\rangle \langle 0| + |1\rangle \langle 1|) \) \[2\].

Superpositions. For superpositions, we allow orthogonal pure states which are different and a rephasing of each other. Let \( S \) and \( S' \) be two such pure states and let \( S \) be spanned by \( g_1, ..., g_N \) which are independent generators. Then in general \( S' \) is spanned by the same generators, but with potentially different signs. These two states are either identical or there is at least one generator \( g_i \in S \) where \(-g_i \in S'\). We can multiply all generators with \( g_j \neq g_i \) of \( S \) such that \(-g_j \in S'\) with \( g_i \) for \( S \) and \(-g_i \) for \( S \). The resulting generators are still independent and span the same state. Therefore, in general we can consider two states which are a rephasing of each other to be of the following form \( S = \text{span}\{g_1, ..., g_N\} \) and \( S' = \text{span}\{g_1, ..., -g_N\} \) where only the sign of the last generator changes and \( g_1, ..., g_N \) are independent generators. Then the coherent superposition of the two states represented by these groups is given by the stabilizer group \( T = \text{span}\{g_1, ..., g_{N-1}, h\} \) where \( h \) is such that \( h \notin S \cup S' \). Such an \( h \) exists by a similar argument as in the stabilizer formalism. In every case there are four different \( h \) that can be chosen. In the case of a single system, they correspond directly to the four different superpositions defined in section 3.1 \[2\].

Example. Consider \( S = \text{span}\{Z_1Z_2, Z_2\} \) and \( S' = \text{span}\{Z_1Z_2, -Z_2\} \); these correspond to the states \(|00\rangle\) and \(|11\rangle\) respectively. By our definition the superposition would be \( \text{span}\{Z_1Z_2, h\} \) with \( h = \pm X_1X_2 \) or \( h = \pm Y_1Y_2 \). These superpositions correspond to the states \( \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \) and \( \frac{1}{\sqrt{2}}(|00\rangle \pm i|11\rangle) \) respectively \[2\].

3.3 Generalizing to arbitrary dimensions

Mixtures. In the generalized version of the toy theory, it is easy to see why we cannot simply mix any two states. Recall that the ontic basis of a valid epistemic state \((V, v)\) is the affine subvector space (or submodule) \( V^\perp + v \). If we were to mix two arbitrary states \((V_1, v_1)\) and \((V_2, v_2)\) by taking the union of their ontic basis, the resulting ontic basis of the mixture would be \((V_1^\perp + v_1) \cup (V_2^\perp + v_2)\). However, this is not necessarily a subvector space or submodule, and therefore may not represent a valid epistemic state.

We will see that the conditions for valid mixtures result in the following restriction: we can only mix epistemic states that correspond to making the same observations but obtaining different results; we also have to impose that each of these observables be either constant across the the family of states to be mixed (that is, it has the same valuation for all the states) or totally unknown (all possible outcomes are present in the family of states).

When we mix these states we will obtain a new state \((V, w)\), for which we should still know the valuation of the constant variables (as they have the same outcome for all states in the family). As for the other variables, all we know is that all outcomes are possible; but this is the same as not
knowing their valuation at all (hence the name “totally unknown”). Therefore, the mixed epistemic state \((V, w)\) would not change if we removed all the totally unknown observables from \(V\). This implies that an alternative way to obtain the mixture in the first place is by taking any one of the original states \((V, v_j)\) and simply remove the totally unknown variables from \(V\), leaving only the valuation of the variables that are constant across the family. This procedure is akin to taking a partial trace over the non-constant observables. We can formalize this discussion in the following definitions.

**Definition 3.4** (Constant and totally unknown observables). Let \(\{(V, v_j)\}_j\) be a family of epistemic states. We say that an observable \(f_i \in V\) is

- constant or known if it has the same valuation for all states in the family, \(f_i^T v_j = k_i, \quad \forall v_j;\)
- totally unknown if for all possible outcomes \(k_{i,\ell}\) of \(f_i\), there exists a member of the family satisfying \(f_i^T v_j = k_{i,\ell};\)
- partially known otherwise.

**Definition 3.5** (Mixture of states (generalized)). Let \(\{(V, v_j)\}_j\) be a family of epistemic states such that each observable \(f_i \in V\) is either constant or totally unknown. Then the epistemic mixture of the states in the family is the state \((W, v_1)\), where

\[
W = \langle \{ f_i \in V : f_i \text{ constant} \} \rangle.
\]

The following lemma ensures that this definition results in the intuitive mixture of states.

**Lemma 3.6** (Soundness of generalized mixtures). Definition 3.5 is sound: it results in a valid epistemic state \((W, w)\). The ontic basis of the mixture is the union of the ontic basis of each element of the family,

\[
W^\perp + w = \bigcup_j (V_j^\perp + v_j).
\]

The choice of \(w = v_1\) is arbitrary, as by definition all the \(v_j\) are all equivalent valuation vectors for the constant variables.

If we would not require that all different possible valuations could be attained for the non-constant variables, then we would know that some valuations of these variables were impossible. This would be more knowledge than not knowing the variables, and less than knowing them completely — and such intermediate, partial knowledge cannot be described in the toy theory.

**Example of an illegal mixture.** Suppose that, for \(d = 2\), we wanted to mix the states

\[
\left( \langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right), \left( \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \rangle, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right),
\]

(3.9)
and
\[
\left( \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle \right).
\] (3.10)

We can represent these states in phase space in the following grid, the same way it’s done in the original toy theory \(^{10}\):

\[
(p_1, q_1)_{\begin{array}{c} 11 \\ 10 \\ 01 \\ 00 \end{array}}
\]

\[
(p_2, q_2)_{\begin{array}{c} 11 \\ 10 \\ 01 \\ 00 \end{array}}.
\] (3.11)

In this representation, the above mixing would correspond to mixing the states

\[
\text{, and .}
\] (3.12)

However, the union of these states is

\[
\text{, which is not a valid epistemic state. Therefore, their mixture cannot be valid.}
\] (3.13)

**Superpositions.** In quantum mechanics, position and momentum eigenstates are connected via a Fourier transform, i.e
\[
|q\rangle = \frac{1}{2\pi} \int_{\mathbb{R}} dp \, e^{ipq} |p\rangle.
\] (3.14)

Thus, certain superpositions of momentum states are position eigenstates. We can define superposition in the toy model in an analogous way, starting from the toy analogs for $|p\rangle$ and $|q\rangle$.

In the continuous case we could have a vector space $V_p = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$, which is the toy analog of $|p\rangle$. Then we can superpose the states $(V_p, v_p)_{p \in \mathbb{R}}$, where $v_p = \left( \begin{array}{c} p \\ p \end{array} \right)$. The resulting state of the superposition is the state $(V_q = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle, v_q = \left( \begin{array}{c} q \\ q \end{array} \right))$. The valuation vector of the superimposed

\(^{10}\)This representation is helpful to connect the original formulation and the generalization for $d = 2$ [3].
state can be chosen arbitrarily and corresponds to the different choices of \( q \) in \( e^{ipq} \) in the quantum superposition.

**Definition 3.7** (Superpositions over one observable). Let \( \{(V, v_j)\}_j \) be a family of epistemic states with \( V = (f_1, \ldots, f_k) \) such that one of the observables \( f_k \) is totally unknown, and all others are constant. Let \( f \) be another observable such that \( V' := (f_1, \ldots, f_{k-1}) \oplus f \) is another isotropic subspace (not equal to \( V \)).

We call the new state \((V', v_j)\), where \( v_j \) is one of the valuations in the original family, the superposition over \( f_k \) with a choice of valuation \( v_j \) and an additional observable \( f \). These different choices correspond to different phases in the superposition.

**Lemma 3.8** (Soundness of generalized superpositions). For the continuous and discrete prime cases, definition 3.7 is sound. That is, there is always an observable \( f \) such that \( V' := (f_1, \ldots, f_{k-1}) \oplus f \) is an isotropic subspace not equal to \( V \). Furthermore, \((V', v_j)\) is a valid epistemic state for every valuation \( v_j \) in the original family.

The definition of superpositions in the generalized formalism can be reduced to its stabilizer formalism definition for the case \( d = 2 \), which is formally stated in the following lemma.

**Lemma 3.9** (Reduction of general superpositions to the stabilizer case). Definition 3.7 reduces to the one in the stabilizer formalism for the case \( d = 2 \). That is, the corresponding stabilizer state to the superposition of two general states is the superposition their corresponding stabilizer states, and the corresponding general state of the superposition of two stabilizer states is the superposition of the corresponding general states.

**Example.** Let us consider the case where \( d = 3, n = 2 \), the isotropic vector space

\[
V = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}. \tag{3.15}
\]

and the three states with valuation vectors

\[
v_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.16}
\]

The first observable in \( V \) is totally unknown for all states, while the second one is constant. Therefore, we can take the superposition over the three states. The vector

\[
f = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{3.17}
\]
commutes with the second observable generating $V$. Therefore,

$$V' = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}. \quad (3.18)$$

can be seen as the isotropic space of the superposition of the rows. All three states $(V', v_0)$, $(V', v_1)$ and $(V', v_2)$ are valid superpositions. Note that there are other choices available for the new variable, for example

$$f' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (3.19)$$

also commutes with the second observable generating $V$ and $V'$. Hence,

$$V'' = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}. \quad (3.20)$$

is also an isotropic space, and as such the superpositions $(V'', v_0)$, $(V'', v_1)$ and $(V'', v_2)$ are all valid states.
4 Entanglement

In quantum mechanics, the phenomenon of entanglement implies the existence of global states that cannot be written as a product of the states of subsystems, and demonstrates non-classical statistical relations between the subsystems [11, 12]. As the toy theory mimics quantum theory to an extent, we are also able to establish the notion of entanglement there. The proofs for new results in this section can be found in Appendix B.

Meaning of entanglement. One may wonder what it means to define “entanglement” in a local hidden variable theory. In Spekkens’ toy theory, to have knowledge about an individual system is to have answers to some of the questions about its ontic state. To have knowledge of two systems being entangled is to be able to answer some questions on how these two systems are correlated. In other words, the information contained in a composite system can be divided into the information carried by individual systems, and the information contained in the correlations between observations made on individual subsystems [13]. Thanks to the knowledge balance principle, it is hence possible to either possess maximal knowledge about the states of subsystems – which results in their epistemic states being uncorrelated; or about the correlations between them, which is described as a correlated epistemic state – and an entangled one, if it is also a state of maximal knowledge. To sum up, to be in an entangled state in the toy theory simply means to have maximal possible information about the correlations between systems. In particular, this is not at odds with the toy theory being an example of a local hidden variable theory.

4.1 Original toy theory

To define entanglement in the epirestricted picture, we first need to define product states, i.e. non-entangled states.

Definition 4.1 (Product states). We say that a state is of a product form between two systems A and B (or is a product state) if it can be written as

\[ E_A \times E_B = \bigvee_{i=1}^{k_A} \bigvee_{j=1}^{k_B} (a_i \times b_j) = (a_1 \times b_1) \lor (a_1 \times b_2) \lor \cdots \lor (a_2 \times b_1) \lor \ldots, \]

where \( E_A = (a_1 \lor a_2 \lor \ldots a_{k_A}) \) and \( E_B = (b_1 \lor b_2 \lor \ldots b_{k_B}) \) are two valid epistemic states on systems A and B respectively.

Theorem 4.2 (Product states are sound). Let \( E_A = (a_1 \lor a_2 \lor \ldots a_{k_A}) \) and \( E_B = (b_1 \lor b_2 \lor \ldots b_{k_B}) \) be two valid epistemic states on two systems A and B. Then the state \( E_A \times E_B = \bigvee_{i=1}^{k_A} \bigvee_{j=1}^{k_B} (a_i \times b_j) \) is a valid epistemic state.

Using this definition, we can define entanglement and correlations.
Definition 4.3 (Entanglement). We call a pure epistemic state defined on systems $A$ and $B$ entangled if it cannot be written as product of epistemic states on $A$ and $B$.

We call a mixed state correlated if it is not of product form. We say a mixed state is entangled if it cannot be written as a mixture of product states.

Examples. An example for a correlated state would be the state

$$\text{(4.1)}$$

This is the toy analog of the classically correlated state $|00\rangle\langle00| + |11\rangle\langle11|$. An example of an entangled state would be the state

$$\text{(4.2)}$$

which can be seen as the toy theory analog of the Bell state $|00\rangle + |11\rangle$. This state can be generalized to states composed of more systems. In general these states are analogous to cat states $|0\rangle^\otimes N + |1\rangle^\otimes N$ and are given by

$$\text{(4.3)}$$

were $\pi'_i$ and $\pi_i$ are injective functions which map from $\{0,1\} \rightarrow \{1,2,3,4\}$ and fulfill $\pi_i(\{0,1\}) \cap \pi'_i(\{0,1\}) = \emptyset$. We show in the stabilizer formalism that the states of this type are valid. For all different $\pi'_i$ and $\pi_i$, these states are equivalent under local transformations. Note that for systems with $N \geq 4$, eq. (4.3) is not the only allowed type of an entangled state.

4.2 Stabilizer formalism

Quantum pure state entanglement. In quantum stabilizer formalism, pure state entanglement has a straightforward interpretation. Let $S$ be a stabilizer group of $n$ qubits, and $A, B$ – two subsystems consisting of $k_A$ and $k_B = n - k_A$ qubits respectively. Then the stabilizer group can be split into: a local subgroup $S_A \cdot S_B = \text{span}\{S_A, S_B\}$, such that $S_A$ and $S_B$ only contain stabilizers of the systems $A$ and $B$ respectively; and the subgroup of the rest $S_{AB}$. It can be shown that the size of the minimal generating set of $S_{AB}$ is double the entanglement entropy of the state stabilized by $S$ [14].
**Toy pure state entanglement.** We can use the same idea as in quantum mechanics to characterize entanglement in the toy theory.

**Theorem 4.4** (Characterization of pure entangled states). Let $S$ be a stabilizer group of a maximal information state and $A$, $B$ two parties that hold each $N_A$ and $N_B$ such that $N_A + N_B = N$ different elementary subsystems respectively, then the state is entangled if and only if $S \neq S_A \cdot S_B$.

**Toy mixed state entanglement.** For non-maximal information states the situation is more complicated because the state may not be factorizable, but can still be written as the sum of two product states and, therefore, not be entangled. In those cases, the stabilizer group may not factorize. Consequently, the above theorem holds for mixed states only in one direction:

**Theorem 4.5** (Characterizing mixed entangled states). Let $S$ be a stabilizer group of a non-maximal information state and $A$, $B$ two parties that hold each $k$ and $n-k$ elementary subsystems respectively. If $S = S_A \cdot S_B$ then the state is not entangled. The converse is not necessarily true.

**Examples.** In the stabilizer formalism we can also describe the epistemic “cat states” defined in eq. (4.3) and, thus, show that they are valid. Let us define the injections $\pi_j, \pi'_j$ in the following way

\[
\pi(0) = \pi_j(0) = 1 \quad \pi(1) = \pi_j(1) = 2 \quad (4.4)
\]

\[
\pi'(0) = \pi'_j(0) = 3 \quad \pi'(1) = \pi'_j(1) = 4. \quad (4.5)
\]

In this case, each state in the ontic basis is stabilized by $\text{span}\{Z_i Z_{i+1} | i \in \{1, \ldots, N-1\}\}$, because for each ontic state in the ontic basis all subsystems have their ontic state either in the $+1$ eigenspace of $Z$ or in the $-1$ eigenspace of $Z$. This state is additionally stabilized by $\prod_{i=1}^{N} \mathcal{X}_i$, because the ontic state of the last system in each state in the ontic basis ensures that there is always an even amount of systems that are in the eigenstate of $\mathcal{X}$. Thus, the epistemic state in eq. (4.3) with our particular choice of injections is stabilized by $\prod_{i=1}^{N} \mathcal{X}_i$ and all elements in $\{Z_i Z_{i+1} | i \in \{1, \ldots, N-1\}\}$, which commute and are independent. There is no additional independent stabilizers as we already identified $N$ independent commuting stabilizers. Note that any other choice of injections corresponds to a local permutation of the initial choice of injectors – this provides us with a procedure of finding the stabilizers for an arbitrary choice of injectors $\pi_j, \pi'_j$.

### 4.3 Generalization for an arbitrary dimension

To define entanglement for a general case, we first need to understand what it means for a state to be completely uncorrelated. Consider two parties, Alice and Bob, where Alice has access to a subsystem $A$, and Bob – to a subsystem $B$.

**Definition 4.6** (Product states (generalized)). Let $(V, v)$ be a state over two subsystems $A$ and $B$. We say that it is a product state if $V = V_A \oplus V_B$, where $V_A$ is spanned by vectors which only have non zero entries on $A$, and analogously for $V_B$.
In this case Alice, who has access to observables in $V_A$, knows nothing about the ontic state of system $B$, and vice-versa. This definition does not depend on $v$ at all, in the same way as entanglement in the stabilizer formalism did not depend on the sign of the stabilizers (that is, their valuation). We now proceed to defining entanglement and correlations.

**Definition 4.7** (Generalized entanglement). *We call an epistemic state $(V,v)$, shared by $A$ and $B$, correlated if it is not product, and entangled if cannot be written as a mixture of product states.*

Let us illustrate these definitions for the generalized toy theory with few simple examples for $d = 2$, which we accompany by their original toy theory counterparts.

- **Entangled state:**
  \[
  \left( \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle, \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \right) \iff \text{graph}
  \]

- **Correlated state:**
  \[
  \left( \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \right) = \left( \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \right) + \left( \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \right)
  \]

- **Product state:**
  \[
  \left( \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle \right) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle \iff \text{graph}
  \]
5 Transformations

To characterize all possible transformations in Spekkens’ toy theory, we first have to consider what kind of transformations are allowed for ontic states.

5.1 Original toy theory

Valid physical transformations. We call transformations $T$ that act on ontic states as physical transformations. We denote their effect on epistemic states in the natural way: any epistemic state $E = \bigvee_i o_i$ is mapped to $T(E) = \bigvee_i T(o_i)$ [1]. We say a physical transformation is valid if and only if it maps all valid epistemic states to valid epistemic states (also at the level of subsystems).

Physical transformations must be bijective. Suppose that many-to-one transformations on ontic states are possible, for example $T : 1 \mapsto 3, 2 \mapsto 3$. Then the epistemic state $(1 \lor 2)$ is mapped to 3. Therefore, the output of this transformation violates the knowledge-balance principle and as such, many-to-one transformations are excluded. Now consider the case of one-to-many transformations: for example, $T' : 1 \mapsto (3 \lor 4), 2 \mapsto 2$. Then the epistemic state $(1 \lor 2)$ would be mapped to $(2 \lor 3 \lor 4)$. This state is invalid, because it has an ontic basis of size 3. Hence, one-to-many transformations are also excluded, and the only allowed ontic state transformations are bijective, or one-to-one. In summary, physical transformations are reversible permutations of ontic states. However, being a permutation is not a sufficient condition for validity.

Remark on irreversible transformations. There is still a way in the toy theory for irreversible transformations in which agents lose knowledge — for example when their systems of interest interact with an environment they cannot track. Similarly to the Stinespring dilation in quantum mechanics, we can formalize a transformation that does not preserve knowledge as a bijection of ontic states applied to a larger system, followed by tracing out a subsystem. In other words, this is a physical, reversible transformation at the level of ontic states, followed by an epistemic “action” of forgetting a subsystem. These transformations do increase the size of the ontic basis of a state, and will be the subject of an upcoming work; we will also get back to them when we generalize to arbitrary dimensions. In the following we will focus only on what we call physical transformations, that is, permutations of ontic states.

Entangling transformations. A transformation is local if it can be decomposed into transformations acting only on individual subsystems. Beyond that, like in quantum theory, we are interested in characterizing the set of entangling transformations.

Definition 5.1 (Entangling transformations). A transformation is called non-entangling if and only if every product state on any division of subsystems $A_1, \ldots, A_k$ is mapped to a product state of subsystems $A_1, \ldots, A_k$; otherwise a transformation is called entangling.
The following theorem characterizes a class of local transformations.

**Theorem 5.2** (Sufficient condition for local transformations). *All non-entangling transformations that preserve the amount of information stored in each subsystem (that is, the size of the ontic support of each marginal doesn’t change) are local.*

**Corollary 5.3** (Non-entangling $\implies$ swaps and local permutations). *All non-entangling transformations are composed of local permutations and system swaps. A system swap $S_{i,j}$ is defined in the following way*

\[
S_{i,j} : o = a_1 \cdot a_2 \cdot \ldots \cdot a_i \cdot \ldots \cdot a_j \cdot \ldots \cdot a_n \rightarrow S_{i,j}o = a_1 \cdot a_2 \cdot \ldots \cdot a_j \cdot \ldots \cdot a_i \cdot \ldots \cdot a_n.
\] (5.1)

This corollary characterizes the non-entangling transformations in the original formulation of the toy theory. Entangling transformations, on the other hand, are easier to characterize in the stabilizer formalism.

### 5.2 Stabilizer formalism

**Conditions for allowed transformations.** In the stabilizer picture, any physical transformation can be written as a permutation matrix $P$ that maps an ontic state $v$ to another, $Pv$. If $v$ is stabilized by $g$ then $Pv$ is stabilized by $PgP^\dagger$, because $P$ is unitary. To ensure that this is still a valid stabilizer we require that $PgP^\dagger \in G_N$. We want to ensure that stabilizer groups get mapped to stabilizer groups, therefore, we require that $PgP^\dagger$ and $PhP^\dagger$ commute if and only if $g$ and $h$ commute. These two conditions fully characterize the subgroup of permutations that are valid \[2\]. In particular, a valid transformation $P$ is fully characterized by its action on a generating set of the toy Pauli group $G_N$.

**Toy CNOT gate.** The two-system permutation

\[
\begin{array}{c|c|c|c}
\hline
\hline
1 & 1 & 1 & 1 \\
\hline
\end{array}
\]

(5.2)

can be seen as an analog to the CNOT-gate. In the stabilizer formalism the CNOT gate can be defined by the action on generators of $G_2$ \[2\]:

\[
\begin{align*}
\mathcal{X}_1 &\rightarrow \mathcal{X}_1 \mathcal{X}_2, \quad \mathcal{X}_2 \rightarrow \mathcal{X}_2 \\
\mathcal{Z}_1 &\rightarrow \mathcal{Z}_1, \quad \mathcal{Z}_2 \rightarrow \mathcal{Z}_1 \mathcal{Z}_2
\end{align*}
\] (5.3)
Clifford group. The group of valid physical transformations in the toy theory is analogous to the quantum Clifford group. In quantum mechanics, the transformations of stabilizer states are elements of the Clifford group $C_N$, which is the group of all unitary operations $U$ fulfilling $U^\dagger g U \in P_N$, $\forall g \in P_N$, where $P_N$ is the Pauli group for $N$ qubits. These unitary operations transform stabilizer groups into stabilizer groups [15]. The same applies in the toy theory:

**Lemma 5.4 ([2]).** A stabilizer group $S = \text{span}\{g_1, \ldots, g_k\}$ is transformed to a new stabilizer group $S' = \text{span}\{U^\dagger g_1 U, \ldots, U^\dagger g_k U\}$ if and only if the transformation $U$ belongs to the set

$$C_N = \{U : U^\dagger g U \in G_N, \quad \forall g \in G_N\},$$

where $G_N$ is the toy Pauli group. Therefore $C_N$ forms the set of allowed unitary transformations in the toy theory.

Any transformation can be determined by its action on a generating set of the toy Pauli group $G_n$, for example $\{X_k, Z_k\}_{k=1,\ldots,N}$. The quantum Clifford group is generated by the Hadamard gate, phase gate and the CNOT gate [15]; analogously, all valid physical transformations in the toy theory are generated by the toy CNOT gate and single system transformations [2].

### 5.3 Generalizing to arbitrary dimensions

**Reversible physical transformations.** The underlying theory of ontic states is symplectic, therefore allowed transformations must also be symplectic, that is they should conserve the symplectic inner product. Because allowed transformations have to map epistemic states to epistemic states, the transformations must be affine

$$m \rightarrow Sm + a$$

where $a \in \Omega$ and $S$ is a symplectic matrix, that is $S^TJS = J$, where $J$ is given by eq. (1.24). Thus, $S$ conserves symplectic inner products. Importantly, $S$ also has an inverse if $d$ is not prime, because $S^{-1} = J^T S^T J$, and does not require for the numbers $a \in \mathbb{Z}_d$ to have an inverse (which does not necessarily exist if $d$ is not prime). These transformations are the equivalent of unitary transformations in quantum theory [3].

The probability distribution $\mu$ corresponding to the epistemic state $(V, \upsilon)$ is transformed as [3]

$$\mu(m) \rightarrow \mu'(m) = \mu(S^{-1} m - a).$$

From this definition we can derive the transformation rule of $(V, \upsilon)$, and determine the structure of the transformed vector space.

**Lemma 5.5 (Reversible transformations (generalized)).** Valid reversible transformations are symplectic transformations represented by a pair $(S, a)$, where $a \in \Omega$ and $S$ is a symplectic matrix. An epistemic state $(V, \upsilon)$ transforms under such a symplectic transformation as

$$(V, \upsilon) \rightarrow ( (S^T)^{-1} V, S(\upsilon + a) ).$$

37
Lemma 5.6 (Validity of transformed states). The transformed vector space defined in lemma 5.5 is isotropic.

Irreversible transformations. Similarly to CPTP maps in quantum mechanics there are maps where the information about the system is not preserved. These maps are defined in analogy to the Stinespring dilation for CPTP maps. An irreversible map can be seen as a reversible map applied to the system and an ancilla, resulting in the probability distribution \( \mu'(m_s, m_a) \), where \( m_s \) is the ontic state of the system, and \( m_a \) the ontic state of the ancilla. We then marginalize over the ancilla system \( \mu'(m_s) = \sum_{m_a} \mu'(m_s, m_a) \) [3].
6 Discussion

6.1 Relation to other theories

Mimicking quantum behaviour. We saw that the toy theory is equivalent to the stabilizer sub-theory of quantum mechanics. The toy theory reproduces many quantum phenomena [1, 3, 4, 16] including interference, non-commutativity of observables, coherent superposition, collapse of the wave function, complementary bases, no-cloning [17], teleportation [18], key distribution [19], locally indistinguishable product bases [9], unextendible product bases [20], some aspects of entanglement, the Choi-Jamiołkowski isomorphism between states and operations [21, 22], the Stinespring dilation of irreversible operations [23], error correction and $k$-threshold key sharing [16]. However, there are quantum features that cannot be reproduced in the toy theory, such as Bell inequality violations [12] and contextuality [5, 24, 25]. This is because the toy theory is a local hidden variable theory, and its underlying ontological theory is non-contextual.\footnote{It has been possible to extend Spekkens' toy model to produce a Bell inequality violation; however, this extension is no longer epistemically restricted [26].}

Quadrature quantum mechanics. In some cases Spekkens' toy theory is equivalent to quadrature quantum mechanics, a subtheory of quantum theory [3, 4]. The starting point of quadrature quantum mechanics are the projectors onto the position and momentum observables:

$$
\Pi_q(q) = |q\rangle \langle q|, \quad \Pi_p(p) = |p\rangle \langle p|,
$$

where we are considering a single (discrete or continuous) quantum system, $\{|q\rangle\rangle_q$ is an arbitrary basis which we name “position basis”, and the momentum basis $\{|p\rangle\rangle_p$ is a conjugate basis, satisfying $\langle q|p\rangle = \frac{1}{\sqrt{\hbar}} e^{i\frac{\pi}{\hbar}qp}$ in the continuous case and $\langle q|p\rangle = \frac{1}{\sqrt{d}} e^{i2\pi qd}$ in the discrete case, where $d$ is the dimension of the system. To form the quadrature subtheory, we restrict ourselves to the subset of states of our Hilbert space for each we either know the position or momentum, that is, the states that can be represented by one of these projectors. In other words, the valid states in this subtheory are those analogous to states in Spekkens' toy theory where one knows the value of the variables $q \sim q$ or $p \sim p$ respectively. If we have several such quantum systems, the valid states of the subtheory are those given by the tensor product of some of these projectors. It can be shown that for any two sets of commuting variables, there exists a symplectic transformation $S$ which transforms one set into the other. Therefore, with the the unitary representation of the symplectic group, we can apply this to the above projectors and find states analogous to any epistemic state $(V, v)$. The allowed transformations are given by the unitary representation of phase space displacements and the symplectic group; quadrature quantum mechanics is closed under these transformations. All allowed measurements are sets of projectors which are obtained by applying a transformation to the projector valued measurement $\{\Pi_q(q)\rangle_q$ [3]. For even values of the dimension $d$ of individual systems, this connection cannot be made, since then quadrature quantum mechanics becomes contextual, and thus cannot be equivalent to the non-contextual toy theory [4].
**Liouville mechanics and Gaussian quantum mechanics.** A classical example of a theory that distinguishes between ontic and epistemic states is Liouville mechanics [27]. There, the epistemic state is characterized by a probability distribution over the phase space of ontic states. If we take Liouville mechanics and introduce an epistemic restriction similar to the Heisenberg uncertainty principle, we end up with a theory equivalent to Gaussian quantum mechanics [28].

### 6.2 Summary of new results

Besides the review, this paper introduced the following new technical contributions, which we hope bring Spekkens’ toy theory into a more refined state for future work:

Theorem 1.3 We translated the knowledge balance principle of the entire state into a statement about sets, and used it to find a necessary condition for epistemic states.

Corollary 5.3 We proved in the original formulation that non-entangling transformations are decomposed into local transformations and system swaps.

Lemma 5.5 We translated the definition of transformation: from an action on the probability distribution over phase space into an action on the epistemic state.

Section 3.3 We generalized the definition of superpositions and mixtures from the stabilizer formalism to the generalized framework. However, we were only able to prove soundness of the definition of superpositions in the case where the dimension is prime and the continuous case, and we leave the other cases for future work.

Definition 4.7 We defined entanglement in the original formulation and its generalization to arbitrary dimensions. As far as we’re aware, previously there was no definition of entanglement, but some epistemic states were known to be analogous to Bell states in quantum mechanics. We used this as a starting point for our definition of entanglement.

Lemma 4.2 We introduced a criterion for entanglement in the original formulation, based on a similar statement for stabilizer quantum mechanics.

Theorem 2.2 We simplified the measurement update rule and unified the prime and non-prime dimensional cases.

### 6.3 Open questions and next steps

**Observer and memories.** This review ends where our next paper begins: the groundwork has been laid to model observers’ memories as physical systems, and each measurement as the joint evolution of the system measured and the memory. As we will see, this is not without its peculiarities. For example, we will see that it is impossible for a physical agent in the toy theory to model arbitrary probability distributions (and mixtures of states).
**Dynamics.** Is there an equivalent of the Schrödinger equation, and Hamiltonian operators, for the toy theory? What constraints should they satisfy? As far as we know these are open questions which would help characterize the theory better.

**Subsystem decomposition.** In quantum mechanics, we can factor a global Hilbert space as the tensor product of several subsystems in uncountable ways: a state can be entangled in a factorization $\mathcal{H} \sim \mathcal{H}_A \otimes \mathcal{H}_B$, and product in a different decomposition $\mathcal{H} \sim \mathcal{H}_C \otimes \mathcal{H}_D$. In the toy theory, the subsystem structure is imposed rigidly, but it would be interesting to relax this assumption and allow for multiple subsystem decompositions of a global system. For example, consider an epistemic state like that of (1.14), which is valid if taken globally, but invalid at the level of our pre-determined subsystem decomposition (because the knowledge balance principle doesn’t hold for every subsystem). However, there is a mapping to another decomposition of global ontic states into different subsystems for which the epistemic state is valid. This tells us that the validity of an epistemic state is relative to a choice of subsystem decomposition of the global space. It could be fruitful to study the implications of this, both in the toy theory and in quadrature quantum theory.

**Where to learn more.** In this review we went over the formalism(s) of the toy theory. To explore applications like how to implement different information-processing tasks within the theory, we recommend the following resources:

1. The original article by Spekkens covers protocols for teleportation, dense coding, steering and entanglement swapping. It also includes proofs of the no-cloning theorem, lack of Bell inequality violations and non-contextuality of the theory.

2. Wherein Pusey introduces the stabilizer formalism and proves its equivalence to the original formalism for the toy theory. It also introduces the toy Mermin-Peres square.

3, 4 Equivalent of the toy theory to stabilizer quantum theory (for odd prime dimensions in Spekkens’ article [3] and generalized to odd integers by Catani and Browne [4]). Spekkens also proves that quadrature and stabilizer quantum mechanics are equivalent.

**Acknowledgements**

We thank Rob Spekkens and Lorenzo Catani for valuable discussions. NN and LdR acknowledge support from the Swiss National Science Foundation through SNSF project No. 200020_165843 and through the National Centre of Competence in Research Quantum Science and Technology (QSIT). LdR further acknowledges support from the FQXi large grant Consciousness in the Physical World.
Author contributions

This review is the first part of LH’s semester project as a masters student. As such, the bibliographic research, novel technical contributions and first draft of the paper are hers. LdR and NN have supervised her and revised the manuscript.
A Linear algebra lemmas

Here we list the results from linear algebra we used in the refinement of the generalization of Spekkens' toy theory.

**Lemma A.1** ([39]). Let $W \subset \Omega$ be a subvector space or submodule and $w \in \Omega$. Then for any $a \in W + w$

$$W + w = W + a \quad (A.1)$$

*Proof.* Let $b \in W + w$ then

$$b = w_1 + w \quad (A.2)$$

for some $w_1 \in W$. As $a \in W + w$ we know that

$$a = w_2 + w$$

$$\iff w = a - w_2 \quad (A.3)$$

for some $w_2 \in W$. Plugging the expression for $w$ into the expression for $b$ we find:

$$b = w_1 - w_2 + a \in W + a \quad (A.4)$$

Therefore, $W + w \subset W + a$.

Let $c \in W + a$ then we can write

$$c = w_3 + a = w_3 + w_2 + w \in W + w \quad (A.5)$$

where we used the expression for $a$ from above. From this equation we can conclude that $W + a \subset W + w$.

**Lemma A.2** ([30]). Let $W, V \subset \Omega$ be two subvector spaces or submodules and $v, w \in \Omega$. Then if $(W + w) \cap (V + v) \neq \emptyset$, it holds that

$$(W + w) \cap (V + v) = (W \cap V) + u \quad (A.6)$$

with $u \in (W + w) \cap (V + v)$.

*Proof.* If $(W + w) \cap (V + v) \neq \emptyset$ then there exists a $u \in (W + w) \cap (V + v)$. Lemma A.1 allows us to write

$$W + w = W + u$$

$$V + v = V + u. \quad (A.7)$$

This means each element in $V + v$ is of the form $u_1 = v_1 + u$ for $v_1 \in V$, and each element in $W + w$ is of the form $u_2 = w_1 + u$ for $w_1 \in W$. Therefore $u_1$ is in $W + w$ if and only if $v_1 \in W$ and $u_2$ is in $V + v$ if and only if $w_1 \in V$. Therefore we can conclude that $(W + w) \cap (V + v) = (V \cap W) + u$. 

43
Lemma A.3 ([4]). Let $V, W \subset \Omega$ be two subvector spaces or submodules then it holds that

$$(V \oplus W)^{\perp} = V^{\perp} \cap W^{\perp} \tag{A.8}$$

Proof. Let $a \in (V \oplus W)^{\perp}$ and, therefore, for all vectors $u \in (V \oplus W)$ it holds that $a^T u = 0$. In particular, this holds for $a \in V^{\perp}$ and $a \in W^{\perp}$ as $V$ and $W$ are subsets of $(V \oplus W)$. Therefore, we can conclude that $(V \oplus W)^{\perp} \subset V^{\perp} \cap W^{\perp}$.

Let $b \in V^{\perp} \cap W^{\perp}$ and let $u \in (V \oplus W)$ be arbitrary. Then we find that

$$u^T b = u_w^T b + u_v^T b = 0 \tag{A.9}$$

for $u_w \in W$ and $u_v \in V$ such that $u_w + u_v = u$. Therefore, $V^{\perp} \cap W^{\perp} \subset (V \oplus W)^{\perp}$. \hfill \square

Lemma A.4 ([31, 32]). Let $V \subset \Omega$ be a subset or submodule. Then it holds that $(V^{\perp})^{\perp} = V$.

Proof. The proof for this in the case for general $d$ can be found in [31]. In the case for general vector spaces ($d$ prime or the continuous case) the proof can be found in [32]. \hfill \square

B Proofs of new results

B.1 States

Here you can find the proofs of statements listed in section 1.

Lemma 1.1 (Canonical set of questions). Let $Q$ be a set of $2N$ questions for a system with $N$ elementary subsystems, $Q_i$ the $i$th question and $\mathcal{M}(Q_i) = (M_{i,0}, M_{i,1})$.

Then $Q$ is a canonical set of questions if and only if for every $2N$-bit string $a$ the intersection $\bigcap_i M_{i,a_i}$ contains exactly one element.

Proof. “$\Longrightarrow$” Let $Q$ be a canonical set of questions, then it fulfills that each ontic states is uniquely defined through a set of answers. Thus, the intersection $\bigcap_i M_{i,a_i}$ will contain exactly one element.

“$\Longleftarrow$” Let $Q$ be a set of $2N$ questions such that the correspond set of partitions $(M_{i,0}, O \setminus M_{i,0}) = (M_{i,0}, M_{i,1}) \in \mathcal{M}(Q)$ that fulfills $|\bigcap_i M_{i,a_i}| = 1$ for all $2N$-bit strings $a$. No two bit strings $a, a'$ generate the same element in the intersection, otherwise this element would have to be in a set and its complement at the same time. Therefore there are $2^{2N}$ different states in the intersection which by the pigeon hole principle have to correspond to all different ontic states. Thus for each ontic state there is a set of answers which characterizes the ontic state uniquely. \hfill \square

Lemma 1.2 (Ontic basis of epistemic states). Let $Q$ be a canonical set of $2N$ questions on $N$ elementary systems, corresponding to the partitions $\mathcal{M}(Q_i) = (M_{i,0}, M_{i,1})$. Let $E \subseteq \{1, 2, \ldots, 2N\}$ index a subset of these questions, and $A_{E} = \{a_i\}_{i \in E}$ represent answers to this subset of questions (that is, an epistemic state, not necessarily a valid one). Then the ontic basis of this epistemic state is given by the set $E = \bigcap_{i \in E} M_{i,a_i}$. 44
Proof. For each answer in $a_i \in \mathcal{A}_E$ we can conclude that the state is in $M_{i,a_i}$ and not in $O \setminus M_{i,a_i}$. Therefore all compatible states need to be contained in the intersection $\cap_{i \in E} M_{i,a_i}$. Conversely, all states in $\cap_{i \in E} M_{i,a_i}$ are compatible with all the answers in $\mathcal{A}_E$ by definition.

**Theorem 1.3** (Dimension of valid epistemic states). Let $E$ be an ontic basis of a valid epistemic state on $N$ elementary systems, which covers the answers to a subset of $k$ out of $2N$ canonical questions. Then $|E| = 2^{2N-k}$.

**Proof.** Let $\mathcal{P} = \{\mathcal{M}(Q_i) | a_i \notin \mathcal{A}_E\}$ the set partitions corresponding to questions in $Q$ where the answer an is unknown. Then it must hold for each $2N - k$ bit string $b$ and $\mathcal{M}(Q_i) \in \mathcal{P}, i \in \{0, ..., 2N - k\}$

$$|E \cap (M_{0,b_0} \cap \ldots \cap M_{2N-k,b_{2N-k}})| = 1. \quad (B.1)$$

The single element in the intersection is different for each bit string. Therefore, there are $2^{2N-k}$ different elements in $E$ that correspond to different bit strings $b$. Let there be an ontic state in $E$ that does not correspond to a bit string $b$. Then this ontic state needs to have different answers to the questions known by the observer, as it is characterized by a complete set of answers and the bit strings $b$ correspond to all different sets of complete answers compatible with the knowledge of the observer. Therefore all ontic states in $E$ are characterized by a bit string $b$ and the cardinality of $E$ is $2^{2N-k}$. \hfill \Box

**Lemma 1.7** (Correspondence between generalized and stabilizer formalisms is sound). If $d = 2$, each general state $(V, \bm{v})$ has a corresponding stabilizer state and each stabilizer state $S$ has a corresponding general state. Furthermore, if variables $f$ and $g$ correspond to stabilizers $g', g''$ respectively, then $f + g$ corresponds to $g' g''$.

**Proof.** Let $(V, \bm{v})$ be a general state. Then we can construct the corresponding stabilizer state. For each $\bm{v} \in V$ the corresponding stabilizer $g = \alpha p_1 \otimes \ldots \otimes p_n$ has the $p_j$ defined by

$$f_{2j-1} = 0, f_{2i} = 0 \implies p_j = \mathcal{I}_4 \quad (B.2)$$
$$f_{2j-1} = 1, f_{2i} = 0 \implies p_j = \mathcal{X} \quad (B.3)$$
$$f_{2j-1} = 0, f_{2i} = 1 \implies p_j = \mathcal{Z} \quad (B.4)$$
$$f_{2j-1} = 1, f_{2i} = 1 \implies p_j = \mathcal{Y} \quad (B.5)$$

and $\alpha = -2f^T v + 1$. Now we need to show that this is a valid stabilizer state. We observe that for $n = 1$, the observables corresponding to $g$ and $g'$ commute if and only if the observables commute. For $n \geq 1$, two observables $f, g$ with corresponding stabilizers $f, g$ respectively commute if and only if an even for an even amount of is $f_{2i-1} g_{2i} - f_{2i} g_{2i-1} = 1$, which happens if and only if $p_{i}^f$ and $p_{i}^g$ the of the corresponding stabilizers $f, g$ respectively do not commute. As $f, g$ commute if and only if an even amount of $p_{i}^f$ and $p_{i}^g$ do not commute, $f, g$ commute if and only if $f, g$ commute.
Therefore, the corresponding stabilizer state to \((V, v)\) is valid. The argument the other way around is analogous, except that we need to show that a valuation vector \(v\) exists. This is the case because \(d = 2\) is prime and each observable attains all valuations.

The second statement can be seen in the following way. It can be easily checked for the correspondence between pairs of components \(f_{2j-1} = 0, f_{2i} = 0\) and the \(p_j\) the component wise addition of pairs is multiplication of the \(p_j\). However, the multiplication of two stabilizers is just the multiplication of \(p_j\) and the statement follows.

\[\Box\]

B.2 Measurements

Here you can find the proofs of statements listed in section 2.

**Theorem 2.2** (Measurement update rule (generalised)). When the epistemic state \((V, v)\) is subjected to a measurement \(V_{\pi}\), and an outcome \(v_{\pi}\) is obtained, the epistemic state is updated to \((V', v')\), where

\[V' = V_{\pi} \oplus V_{\text{commute}},\]

\[v' \in (V_{\pi}^{\perp} + v_{\pi}) \cap (V_{\text{commute}}^{\perp} + v),\]  

and

\[V_{\text{commute}} = \{f \in V : \langle f, f_{\pi} \rangle = 0, \forall f_{\pi} \in V_{\pi}\} \subseteq V\]

is the subset of \(V\) that commutes with \(V_{\pi}\).

**Proof.** Similarly to quantum mechanics, we impose that repeated measurements of the same quantity should give the same result. One probability distribution that fulfills this property is

\[\mu'_{(V', v')}(m) \propto \delta_{V_{\pi}^{\perp} + v_{\pi}}(m) \mu_{(V, v)}(m) = \frac{1}{N_{V}} \delta_{V_{\pi}^{\perp} + v_{\pi}}(m) \delta_{V^{\perp} + v_{\perp}}(m),\]  

where \(N_{V}\) is the normalization constant, as in eq. (1.30). With this updated probability distribution, we can define the update rule for \(V\).

First we consider the case where \(V\) commutes with all variables in \(V_{\pi}\). To find the updated state \((V', v')\), we want to find all \(m \in \Omega\) such that \(\mu'_{(V', v')}(m) \neq 0\). This condition is fulfilled if and only if \(m \in V_{\pi}^{\perp} + v_{\pi}\) and \(m \in V^{\perp} + v\), that is,

\[V_{\pi}^{\perp} + v = (V_{\pi}^{\perp} + v_{\pi}) \cap (V^{\perp} + v).\]

Notice that \(V_{\pi}^{\perp} + v\) is non empty, as otherwise the outcome \(v_{\pi}\) would appear with zero probability.

With lemmas A.1 to A.3 we find that

\[V_{\pi}^{\perp} + v = (V_{\pi}^{\perp} \cap V^{\perp}) + r\]

\[= (V_{\pi} \oplus V)^{\perp} + r\]  

(B.8)
with \( r \) such that \( r \in (V_{\pi}^{\perp} + \nu_{\pi}) \cap (V^{\perp} + \nu) \). In fact, we can choose any such \( r \), as all vectors in \((V_{\pi}^{\perp} + \nu_{\pi}) \cap (V^{\perp} + \nu)\) lead to an equivalent measurement update. Using lemma A.4, we find the following measurement update rules:

\[
V' = V_{\pi} \oplus V , \quad \nu' \in (V_{\pi}^{\perp} + \nu_{\pi}) \cap (V^{\perp} + \nu) . \tag{B.9}
\]

As all vectors in \( V \) commute with those in \( V_{\pi} \) the \( V' = V_{\pi} \oplus V \) is isotropic.

In the case where some the vectors in \( V_{\pi} \) do not commute with vectors in \( V \) we have to modify the measurement update, because the rule of eq. (B.11) produces a non-isotropic subspace. Like in the commuting case, we want that a repeated measurement produces the same outcome. Therefore, to satisfy the principle of classical complementarity, information about the pre-measurement state needs to get lost during a measurement. To get a valid post-measurement state we must remove all the variables \( f \) from \( V \) that do not commute. Let \( V_{\text{commute}} \) be the subspace of \( V \) where all variables commute with all variables \( V_{\pi} \). Then we can follow the same reasoning as in the commuting case with \( V_{\text{commute}} \) instead of \( V \) and find the final result. As \( V_{\text{commute}} \) and \( V_{\pi} \) commute by definition of \( V_{\text{commute}} \) the updated state

\[
V' = V_{\pi} \oplus V_{\text{commute}}, \quad \nu' \in (V_{\pi}^{\perp} + \nu_{\pi}) \cap (V_{\text{commute}}^{\perp} + \nu) . \tag{B.10}
\]

is a valid state, by the same reasoning as above.

**Theorem 2.3** (Generalised measurement update rule reduces to the original formalism). For the case \( d = 2 \), the generalised measurement update rule of 2.2 reduces to the measurement update rule in the original formalism.

**Proof.** The measurement \( V_{\pi} \) and a set of outcome vectors \( \nu_{\pi}^{1}, ..., \nu_{\pi}^{k} \) where each \( \nu_{\pi}^{j} \) has a different valuation of the vectors in \( V_{\pi} \). The updated state is a valid state, if we define it via the correspondence to the stabilizer formalism. By construction, the updated state is such that its ontic support is in the right part of the partition. This means we only need to show that it has maximal fidelity. Let us first calculate the fidelity of two arbitrary states \((V, \nu), (V', \nu')\). Either, \( |(V^{\perp} + \nu) \cap (V'^{\perp} + \nu')| = 0 \) and the fidelity is zero, or the fidelity is non zero and \( |(V^{\perp} + \nu) \cap (V'^{\perp} + \nu')| = |(V \oplus V')^{\perp}| = 2^{2n - \dim(V \oplus V')} \). In the latter case, the fidelity is given by \( 2^{2n - \dim(V \oplus V')} \sqrt{2 - 2n + \dim V \cdot 2 - 2n + \dim V'} \). Let us now apply this to the case where \((V, \nu)\) is the state we do a measurement on, and \((V', \nu')\) is the updated state. Let us define \( d = \dim W \), where \( W = V \cap V' \). Therefore, \( \dim V' = d + d_{V' \cap W} \) and \( \dim V = d + d_{V \cap W} \).

We can now calculate the fidelity of these two states. If the fidelity between the two states is zero, then \((V', \nu')\) it is not a valid measurement update. Therefore the fidelity between the initial and updated state is given by

\[
F((V, \nu), (V', \nu')) = 2^{-d + d_{V \cap W} + d_{V' \cap W}} / 2^{(d + d_{V' \cap W})/2} = 2^{-d_{V \cap W} + d_{V' \cap W}} / 2 . \tag{B.13}
\]
Therefore, the fidelity is the largest if $d_{V/W} \frac{d_{V'/W}}{d_{V/W}}$ are the smallest. This means that the minimal amount of vectors are removed from or added to $V$. For $V'$ to be a valid measurement update it must hold that $V_\pi \subseteq V'$, therefore we need to add all vectors from $V_\pi$ to $V$ which are not already in $V$ and we need to remove all vectors from $V$ which do not commute with $V_\pi$. However, this is just the vector space $V_\pi \oplus V_{\text{commute}}$, which is $V'$ in our definition of the measurement update. \qed

B.3 Mixtures and superpositions

Here you can find the proofs of statements listed in section 3.

**Theorem 3.3** (Mixed states as convex combinations). Any mixed state with stabilizer group $S = \text{span}\{g_1, ..., g_{N-k}\}$, $k \leq N$ can be written as the mixture of $2^k$ pure states.

**Proof.** Let $h_1, ..., h_k$ be independent stabilizer such that together with $g_1, ..., g_{N-k}$ they span a stabilizer group $T = \text{span}\{g_1, ..., g_{N-k}, h_1, ..., h_k\}$ of a maximal information state. With the same argument as in the stabilizer formalism one can show that it is in fact possible to find such $h_1, ..., h_k$.

We define $2^k$ rephasings of $T$

$$\mathcal{T}^{(0)}_{i(0)} = \text{span}\{g_1, ..., g_{N-k}, (-1)^{i(0)}_1 h_1, ..., (-1)^{i(0)}_k h_k\},$$

(B.14)

where $i^{(0)}_j$ is the $j$-th digit of the binary representation of $i^{(0)}$. Then we can pair up all $\mathcal{T}^{(0)}_{i(0)}$ such that the pairs have the same sign in front of all generators except the last generator. Then we take the mixture of each of the pairs and define these as $\mathcal{T}^{(1)}_{i(1)}$ where $i^{(1)}$ is obtained from $i^{(0)}$ by setting $i^{(0)}_k$ to zero in the binary representation. There are now only $2^{k-1} \mathcal{T}^{(1)}_{i(1)}$ left. Repeating this procedure reduces the number of $h$ by one and halves the number of $\mathcal{T}^{(1)}_{i(1)}$. Thus, if this procedure is applied $k$ times there is only one $\mathcal{T}^{(k)}_{i(k)}$ left, which is generated by $g_1, ..., g_{N-k}$. \qed

**Lemma 3.8** (Soundness of generalized superpositions). For the continuous and discrete prime cases, definition 3.7 is sound. That is, there is always an observable $f$ such that $V' = \{f_1, \ldots, f_{k-1}\} \oplus f$ is an isotropic subspace not equal to $V$. Furthermore, $(V', v_j)$ is a valid epistemic state for every valuation $v_j$ in the original family.

**Proof.** For this definition to be sound, we need to prove that there always exists a $w \notin V$ which commutes with $f^{(1)}, ..., f^{(k-1)}$. Let $U = \{f^{(1)}, ..., f^{(k-1)}\}$ then $w$ has to be contained in $U^{\perp_{\text{symplectic}}} = \{m \in \Omega | \forall v \in V \ v^T J m = 0\}$. Because $U$ is isotropic it holds that $U \subset U^{\perp_{\text{symplectic}}}$. Let $\ell = \dim(U^{\perp_{\text{symplectic}}}) - \dim(U)$. It holds that $\dim(U) + \dim(U^{\perp_{\text{symplectic}}}) = (k-1) + (k-1 + \ell) = 2n$ \cite{32}. Rearranging this expression, we can calculate $\ell = 2n - 2k + 2$. This implies that $\ell \geq 2$, because $k$ is upper bounded by $n$. Therefore, there always exists a vector $w \in U \setminus U^{\perp_{\text{symplectic}}}$ which is linearly independent of $f^{(k)}$ and, thus, not contained in $V$. \qed

**Lemma 3.9** (Reduction of general superpositions to the stabilizer case). Definition 3.7 reduces to the one in the stabilizer formalism for the case $d = 2$. That is, the corresponding stabilizer state
to the superposition of two general states is the superposition their corresponding stabilizer states, and the corresponding general state of the superposition of two stabilizer states is the superposition of the corresponding general states.

**Proof.** Let \((V, v_1)\) and \((V, v_2)\) with \(V = \langle f_1, \ldots, f_k \rangle\) two states where \(f^k\) is completely unknown. Let \(g_k\) be the stabilizer corresponding to \(f_k\) with valuation vector \(v\). Then \(g_1, \ldots, g_k\) are independent due to the second point in lemma 1.7. The corresponding stabilizer state to \((V, v_1)\) is then \(S = \langle g_1, \ldots, g_k \rangle\). As, \(f_k\) is totally unknown the corresponding stabilizer state of \((V, v_2)\) is \(S' = \langle g_1, \ldots, -g_k \rangle\). A superposition of \((V, v_1)\) and \((V, v_2)\) is \(V' = \langle f_1, \ldots, f_{k-1} \rangle + w, v_1\) wit \(w\) such that \(V'\) isotropic. Then, let \(h\) be the corresponding stabilizer to \(w\) with valuation \(v_1\). Then, \(h\) is independent of and commutes with \(g_1, \ldots, g_{k-1}\). Which means that \(h\) fulfills the definition of \(h\) in the definition of superposition for stabilizers.

Let \(S = \langle g_1, \ldots, g_k \rangle\) and \(S' = \langle g_1, \ldots, -g_k \rangle\) be two stabilizer states. Then the corresponding general states \((V, v_1)\) and \((V, v_2)\) with \(V = \langle f_1, \ldots, f_k \rangle\) are such that \(f^k\) is completely unknown. Let \(h\) be a stabilizer as in the definition of the superposition of stabilizer states. Then the corresponding observable to \(h\) fulfills the definition of \(w\) in the definition of superposition of general states. Let \(v'\) be the valuation vector corresponding to \(\langle g_1, \ldots, g_{k-1}, h \rangle\). Then \(v'\) has the same valuation as \(v_1\) and \(v_2\) on \(f_1, \ldots, f_{k-1}\) and either the valuation of \(v_1\) or \(v_2\) on \(f_k\). Therefore, \(v'\) is an equivalent valuation vector to either \(v_1\) or \(v_2\) and \(\langle f_1, \ldots, f_{k-1}, w, v' \rangle\) is the superposition of \((V, v_1)\) and \((V, v_2)\).

\(\square\)

### B.4 Entanglement

Here you can find the proofs of statements listed in section 4.

**Theorem 4.2** (Product states are sound). Let \(E_A = (a_1 \lor a_2 \lor \ldots a_{k_A})\) and \(E_B = (b_1 \lor b_2 \lor \ldots b_{k_B})\) be two valid epistemic states on two systems \(A\) and \(B\). Then the state \(E_A \times E_B = \bigvee_{i=1}^{k_A} \bigvee_{j=1}^{k_B} (a_i \times b_j)\) is a valid epistemic state.

**Proof.** See the \(\Leftarrow\) part of the proof of theorem 4.4.

**Theorem 4.4** (Characterization of pure entangled states). Let \(S\) be a stabilizer group of a maximal information state and \(A, B\) two parties that hold each \(N_A\) and \(N_B\) such that \(N_A + N_B = N\) different elementary subsystems respectively, then the state is entangled if and only if \(S \neq S_A \cdot S_B\).

**Proof.** \(\Rightarrow\) If \(S = S_A \cdot S_B\) then \(S = \text{span}\{s_{A,1}, \ldots, s_{A,k}, s_{B,1}, \ldots, s_{B,N-k}\}\) where \(s_{A,i}, s_{B,i}\) are stabilizer operator only acting non trivially on the \(A\) and \(B\) system respectively. Therefore, there are \(2^{2N_A-k}\) ontic states that are stabilized by \(S_A\) on system \(A\) and \(2^{2N_B-N+k}\) that are stabilized by \(S_B\) on system \(B\). Taking the product of these states results in a state with an ontic basis of \(2^{2N_A-k+2N_B-N+k} = 2^{2(N_A+N_B)-N} = 2^N\). The product state is stabilized by \(S\), as each element of the ontic basis of the product state is on the \(A\) part stabilized by all \(s_{A,i}\) and on the \(B\) part
stabilized by all $s_{B,i}$. As $S$ is a pure state we have found there are no more ontic states that can be stabilized by $S$ and therefore the state stabilized by $S$ is pure.

⇐ Let $S$ be the stabilizer group corresponding to a product state, then we can find the stabilizer group $S_A$ that stabilizes $A$ and the stabilizer group $S_B$ that stabilizes the state on the $B$ system. These stabilizers from a new group $S' = S_A \cdot S_B$. We have found a different stabilizer group that stabilizes the same state as $S$. If $S' \neq S$ then there must be a stabilizer in $S'$ that is not in $S'$. Therefore, there must be an ontic state stabilized by $S$ but not by $S$, but then the states stabilized are not the same. Thus, the two groups have to be the same and $S = S_A \cdot S_B$.  

**Theorem 4.5** (Characterizing mixed entangled states). Let $S$ be a stabilizer group of a non-maximal information state and $A, B$ two parties that hold each $k$ and $n - k$ elementary subsystems respectively. If $S = S_A \cdot S_B$ then the state is not entangled. The converse is not necessarily true.

Proof. The same argument as in the $\Leftarrow$ part of the previous theorem holds. The reason the $\Rightarrow$ direction does not hold is because for mixed states non-entangled states are not necessarily product states. \hfill $\square$

### B.5 Transformations

Here you can find the proofs of statements listed in section 5.

**Theorem 5.2** (Sufficient condition for local transformations). All non-entangling transformations that preserve the amount of information stored in each subsystem (that is, the size of the ontic support of each marginal doesn’t change) are local.

Proof. Local transformations are non-entangling and subsystem information preserving, as it just permutes the ontic states of each elementary system.

Let $T$ be a non-entangling subsystem information preserving transformation on $N$ systems, $O_N$ the set of ontic states on $N$ systems and $C = (a \lor b) \cdot (1 \lor 2 \lor 3 \lor 4)^{N-1}$ a valid epistemic state (i.e $a \neq b$). Then

$$T(c) = \lor_{o \in O_{N-1}} T((a \lor b) \cdot o)$$
$$= (c \lor d) \cdot (1 \lor 2 \lor 3 \lor 4)^{N-1}$$

with $c \neq d$, because $T$ is a valid non-entangling subsystem preserving transformation. Therefore, for any ontic state $o$ it holds that

$$T((a \lor b) \cdot o) = (c \lor d) \cdot o'$$

$T$ is a permutation, therefore the amount of states in the intersection of two ontic bases is invariant under transformation. Furthermore, it does not matter if the states are first transformed and then the intersection is taken or if first the intersection is taken and then the states are
transformed. Let $C' = (a \lor f) \cdot (1 \lor 2 \lor 3 \lor 4)^{N-1}$ be an valid epistemic state such that $a \neq f$ and $b \neq f$. Then the same conclusions that hold for $C$ also hold for $C'$. Therefore, for any ontic state $o$

$$T(a \cdot o) = T((a \lor b) \cdot o) \cap T((a \lor f) \cdot o) = g \cdot o'$$  \hfill (B.17)

with $g$ either $c$ or $d$ and $o'$ some ontic state. The same argument also applies for $b$. Therefore, the transformation can be written as

$$T((a \lor b) \cdot o) = T_1(a \lor b) \cdot T_{N-1}(o).$$  \hfill (B.18)

Repeating the same argument for $T_{N-1}(o)$ the transformation $T$ factorizes into local transformations.

\textbf{Corollary 5.3} (Non-entangling $\Rightarrow$ swaps and local permutations). All non-entangling transformations are composed of local permutations and system swaps. A system swap $S_{i,j}$ is defined in the following way

$$S_{i,j} : o = a_1 \cdot a_2 \cdot ... \cdot a_i \cdot ... \cdot a_j \cdot ... \cdot a_n \rightarrow S_{i,j} o = a_1 \cdot a_2 \cdot ... \cdot a_j \cdot ... \cdot a_i \cdot ... \cdot a_n.$$  \hfill (5.1)

\textbf{Proof.} Let $T$ be a non entangling transformation. Then we can detect changes in the system by applying transformation to $C = (a \lor b) \cdot (1 \lor 2 \lor 3 \lor 4)^{N-1}$ and detecting which system $i$ is in the pure state after the transformation. Only one system can be in the pure state, as the transformation is a permutation of ontic states.

Consider the transformation $S_{1,i}T$, for which the same argument as in the proof above holds. Thus, this transformation can be decomposed in the following way

$$S_{1,i}T = T_1 \cdot T_{N-1} \hfill (B.19)$$

where $T_1$ a transformation on the first and $T_{N-1}$ a transformation on the last $N - 1$ subsystems. Because $S_{1,i}$ is its own inverse, we can rewrite the above decomposition

$$T = S_{1,i}(T_1 \cdot T_{N-1}). \hfill (B.20)$$

Applying the same argument as above inductively to $T_{N-1}$ factorizes the transformation in swaps $S_{i,j}$ and local transformations.

\textbf{Lemma 5.4} ([2]). A stabilizer group $S = \text{span}\{g_1, ..., g_k\}$ is transformed to a new stabilizer group $S' = \text{span}\{U^\dagger g_1 U, ..., U^\dagger g_k U\}$ if and only if the transformation $U$ belongs to the set

$$C_N = \{U : U^\dagger g U \in G_N, \forall g \in G_N\},$$

where $G_N$ is the toy Pauli group. Therefore $C_N$ forms the set of allowed unitary transformations in the toy theory.
Proof. The elements of the set \( \{ U^\dagger g_1 U, \ldots, U^\dagger g_k U \} \) still commute as

\[
[U^\dagger g_i U, U^\dagger g_j U] = U^\dagger [g_i, g_j] U = 0.
\]

(B.21)

Therefore, \( S' \) is a stabilizer group. \( \square \)

**Lemma 5.5** (Reversible transformations (generalized)). *Valid reversible transformations are symplectic transformations represented by a pair \((S, a)\), where \( a \in \Omega \) and \( S \) is a symplectic matrix. An epistemic state \((V, v)\) transforms under such a symplectic transformation as*

\[
(V, v) \to ( (S^T)^{-1} V, S(v + a) ).
\]

(5.6)

**Proof.** We can find the updated epistemic \((V', v')\) state by calculating the set \( V'^\perp + v' \) which is characterized by vectors \( m \) such that \( \mu'(m) \neq 0 \). For such vectors we can do the following chain of reasoning

\[
\mu'(m) \neq 0 \iff \mu(S^{-1} m - a) \neq 0 \\
\iff S^{-1} m - a \in V'^\perp + v
\]

(B.22)

Therefore, we can conclude that if and only if \( \mu'(m) \neq 0 \) it holds that \( m = S(v'^\perp + v + a) \) with \( v'^\perp \in V'^\perp \). This means we can infer \( m \in SV'^\perp + S(v + a) \) if \( \mu'(m) \neq 0 \). On the other hand, if \( m \in SV'^\perp + S(v + a) \) then

\[
\mu'(m) = \mu'(S(v'^\perp + v + a)) \\
= \mu(S^{-1}(S(v'^\perp + v + a)) - a) \\
= \mu(v'^\perp + v) \neq 0.
\]

(B.23)

We can conclude that

\[
V'^\perp + v' = SV'^\perp + S(v + a).
\]

(B.24)

This means that the transformed valuation vector \( v' \) is given by

\[
v' = S(v + a)
\]

(B.25)

and the updated vector space \( V' \)

\[
V' = (SV'^\perp)'^\perp.
\]

(B.26)

This expression can be evaluated using the definition of an orthogonal complement

\[
w \in (SV'^\perp)^\perp \iff \forall v \in SV'^\perp : v^T w = 0 \\
\iff \forall v \in V'^\perp : (Sv)^T w = 0 \\
\iff \forall v \in V'^\perp : v^T(S^T w) = 0 \\
\iff (S^T w) \in V \\
\iff w \in (S^T)^{-1} V.
\]

(B.27)
Therefore we can give a more direct expression for $V'$

$$V' = (S^T)^{-1}V. \quad \text{(B.28)}$$

In summary a state $(V, v)$ transforms under a symplectic transformation $S \cdot + a$ in the following way

$$(V, v) \rightarrow ((S^T)^{-1}V, S(v + a)) \quad \text{(B.29)}$$

Lemma 5.6 (Validity of transformed states). The transformed vector space defined in lemma 5.5 is isotropic.

Proof. Let $V = \langle f_1, ..., f_k \rangle$ be an isotropic space. Then $V'$ is given by $V' = \langle (S^T)^{-1}f_1, ..., (S^T)^{-1}f_k \rangle$. The symplectic inner product between $(S^T)^{-1}f_i$ and $(S^T)^{-1}f_j$ can then be calculated

$$((S^T)^{-1}f_i)^T(J((S^T)^{-1}f_j) = f_i^T(S^{-1}J(S^T)^{-1})f_j$$
$$= f_i^T(J^T S^T J)(S^T S J) f_j$$
$$= f_i^T(J^T J) f_j$$
$$= f_i^T J f_j$$
$$= 0, \quad \text{(B.30)}$$

were we used that $J^T = J^{-1}$ and the expression for $S^{-1}$.

53
References

[1] Robert W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Physical Review A*, 75(3), mar 2007. DOI: 10.1103/physreva.75.032110.

[2] Matthew F. Pusey. Stabilizer notation for spekkens’ toy theory. *Foundations of Physics*, 42(5):688–708, mar 2012. DOI: 10.1007/s10701-012-9639-7.

[3] Giulio Chiribella and Robert W. Spekkens, editors. *Quantum Theory: Informational Foundations and Foils*, chapter Quasi-quantization: classical statistical theories with an epistemic restriction. Springer Netherlands, 2016. DOI: 10.1007/978-94-017-7303-4.

[4] Lorenzo Catani and Dan E Browne. Spekkens’ toy model in all dimensions and its relationship with stabiliser quantum mechanics. *New Journal of Physics*, 19(7):073035, 2017. DOI: 10.1088/1367-2630/aa781c.

[5] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71(5), may 2005. DOI: https://doi.org/10.1103/PhysRevA.71.052108.

[6] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Physical Review A*, 70(5), 2004. DOI: 10.1103/PhysRevA.70.052328.

[7] Daniel Gottesman. Stabilizer codes and quantum error correction. 1997. arXiv:quant-ph/9705052.

[8] Paul Adrien Maurice Dirac. The fundamental equations of quantum mechanics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 109(752):642–653, 1925. DOI: 10.1098/rspa.1925.0150.

[9] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Physical Review A*, 59(2):1070–1091, 1999. DOI: 10.1103/physreva.59.1070.

[10] Étienne Bézout. *Théorie générale des équations algébriques*. 1779.

[11] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 47:777–780, 1935. DOI: 10.1103/PhysRev.47.777.

[12] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1(3):195–200, nov 1964. DOI: 10.1142/9789812386540_0002.

[13] Caslav Brukner, Marek Zukowski, and Anton Zeilinger. The essence of entanglement, 2001. arXiv:quant-ph/0106119.

[14] David Fattal, Toby S. Cubitt, Yoshihisa Yamamoto, Sergey Bravyi, and Isaac L. Chuang. Entanglement in the stabilizer formalism. 2004. arXiv:quant-ph/0406168.

[15] Daniel Gottesman. Theory of fault-tolerant quantum computation. *Physical Review A*, 57(1):127–137, 1998. DOI: 10.1103/PhysRevA.57.127.

[16] Leonardo Disilvestro and Damian Markham. Quantum protocols within spekkens’ toy model. *Physical Review A*, 95(5), may 2017. DOI: 10.1103/PhysRevA.95.052324.
[17] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, oct 1982. DOI: https://doi.org/10.1038/299802a0.

[18] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Physical Review Letters*, 70(13):1895–1899, mar 1993. DOI: https://doi.org/10.1103/PhysRevLett.70.1895.

[19] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, dec 2014. DOI: 10.1016/j.tcs.2014.05.025.

[20] Charles H. Bennett, David P. DiVincenzo, Tal Mor, Peter W. Shor, John A. Smolin, and Barbara M. Terhal. Unextendible product bases and bound entanglement. *Physical Review Letters*, 82(26):5385–5388, jun 1999. DOI: https://doi.org/10.1103/PhysRevLett.82.5385.

[21] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, jun 1975. DOI: https://doi.org/10.1016/0024-3795(75)90075-0.

[22] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefinite-ness of operators. *Reports on Mathematical Physics*, 3(4):275–278, dec 1972. DOI: https://doi.org/10.1016/0034-4877(72)90011-0.

[23] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, jun 1975. DOI: https://doi.org/10.1016/0024-3795(75)90075-0.

[24] Simon Kochen and E. P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967. ISSN 00959057, 19435274.

[25] Yeong-Cherng Liang, Robert W. Spekkens, and Howard M. Wiseman. Specker’s parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. *Physics Reports*, 506(1-2):1–39, sep 2011. DOI: https://doi.org/10.1016/j.physrep.2011.05.001.

[26] S. J. van Enk. A toy model for quantum mechanics. *Foundations of Physics*, 37(10):1447–1460, aug 2007. DOI: 10.1007/s10701-007-9171-3.

[27] Joseph Liouville. Note sur la théorie de la variation des constantes arbitraires. *Journal de mathématiques pures et appliquées*, pages 342–349, 1838.

[28] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Reconstruction of gaussian quantum mechanics from liouville mechanics with an epistemic restriction. *Physical Review A*, 86(1), jul 2012. DOI: 10.1103/PhysRevA.86.012103.

[29] Gerd Fischer. *Lineare Algebra: Eine Einführung für Studienanfänger*. Imprint Springer Spektrum, Wiesbaden, 2014. ISBN 9783658039448.

[30] Jean Gallier. Basics of Affine Geometry. In *Texts in Applied Mathematics*, pages 7–63. Springer New York, 2011. DOI: 10.1007/978-1-4419-9961-0_2.

[31] David Wilding, Marianne Johnson, and Mark Kambites. Exact rings and semirings. *Journal of Algebra*, 388:324–337, 2013. DOI: 10.1016/j.jalgebra.2013.05.005.
[32] T.Y. Lam. *Introduction to Quadratic Forms over Fields*. American Mathematical Society, 2004. DOI: 10.1090/gsm/067.