

Spekkens' toy model in all dimensions and its relationship with stabiliser quantum mechanics

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 New J. Phys. 19 073035

(<http://iopscience.iop.org/1367-2630/19/7/073035>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 128.41.35.98

This content was downloaded on 25/08/2017 at 12:21

Please note that [terms and conditions apply](#).

You may also be interested in:

[An Introduction to the Formalism of Quantum Information with Continuous Variables: Quantum information with continuous variables](#)

C Navarrete-Benlloch

[Quantum circuit dynamics via path integrals: Is there a classical action for discrete-time paths?](#)

Mark D Penney, Dax Enshan Koh and Robert W Spekkens

[Quasi-probability representations of quantum theory with applications to quantum information science](#)

Christopher Ferrie

[The resource theory of stabilizer quantum computation](#)

Victor Veitch, S A Hamed Mousavian, Daniel Gottesman et al.

[Measures and applications of quantum correlations](#)

Gerardo Adesso, Thomas R Bromley and Marco Cianciaruso

[Entanglement in algebraic quantum mechanics: Majorana fermion systems](#)

F Benatti and R Floreanini

[Optimal bounds for parity-oblivious random access codes](#)

André Chailloux, Iordanis Kerenidis, Srijita Kundu et al.

[Negative quasi-probability as a resource for quantum computation](#)

Victor Veitch, Christopher Ferrie, David Gross et al.

[A framework for phase and interference in generalized probabilistic theories](#)

Andrew J P Garner, Oscar C O Dahlsten, Yoshifumi Nakata et al.

**PAPER**

Spekkens' toy model in all dimensions and its relationship with stabiliser quantum mechanics

OPEN ACCESS**RECEIVED**

24 February 2017

REVISED

18 May 2017

ACCEPTED FOR PUBLICATION

8 June 2017

PUBLISHED

27 July 2017

Lorenzo Catani and Dan E Browne

University College London, Physics and Astronomy department, Gower St, London WC1E 6BT, United Kingdom

E-mail: lorenzo.catani.14@ucl.ac.uk**Keywords:** Spekkens' toy model, stabiliser quantum mechanics, discrete Wigner functions, non-prime dimensional systems, measurements update rules, epistemic view, contextuality

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

**Abstract**

Spekkens' toy model is a non-contextual hidden variable model with an epistemic restriction, a constraint on what an observer can know about reality. The aim of the model, developed for continuous and discrete prime degrees of freedom, is to advocate the epistemic view of quantum theory, where quantum states are states of incomplete knowledge about a deeper underlying reality. Many aspects of quantum mechanics and protocols from quantum information can be reproduced in the model. In spite of its significance, a number of aspects of Spekkens' model remained incomplete. Formal rules for the update of states after measurement had not been written down, and the theory had only been constructed for prime-dimensional and infinite dimensional systems. In this work, we remedy this, by deriving measurement update rules and extending the framework to derive models in all dimensions, both prime and non-prime. Stabiliser quantum mechanics (SQM) is a sub-theory of quantum mechanics with restricted states, transformations and measurements. First derived for the purpose of constructing error correcting codes, it now plays a role in many areas of quantum information theory. Previously, it had been shown that Spekkens' model was operationally equivalent to SQM in the case of odd prime dimensions. Here, exploiting known results on Wigner functions, we extend this to show that Spekkens' model is equivalent to SQM in all odd dimensions, prime and non-prime. This equivalence provides new technical tools for the study of technically difficult compound-dimensional SQM.

1. Introduction

A long tradition of research, starting from the famous 'EPR paper' [1], has consisted of analysing quantum theory in terms of hidden variable models, with the aim of obtaining a more intuitive understanding of it. This has led to some crucial results in foundation of quantum mechanics, namely Bell's and Kochen–Specker's no-go theorems [2, 3]. Nowadays a big question is whether to interpret the quantum state according to the ontic view, i.e. where it completely describes reality, or to the epistemic view, where it is a state of incomplete knowledge of a deeper underlying reality which can be described by the hidden variables. In 2005, Robert Spekkens [4] constructed a non-contextual hidden variable model to support the epistemic view of quantum mechanics. The aim of the model was to replace quantum mechanics by a hidden variable theory with the addition of an epistemic restriction (i.e. a restriction on what an observer can know about reality). The first version of the model [4] was developed in analogy with quantum bits (qubits), with 2-outcome observables. Despite the simplicity of the model, it was able to support many phenomena and protocols that were believed to be intrinsically quantum mechanical (such as dense coding and teleportation). Spekkens' toy model has influenced much research over the years: e.g. people provided a new notation for it [25], studied it from the categorical point of view [26], used it for quantum protocols [27], exploited similar ideas to find a classical model of one qubit [28], and tried to extend it in a contextual framework [29]. Also Spekkens' toy model addresses many key issues

in quantum foundations: whether the quantum state describes reality or not, finding a derivation of quantum theory from intuitive physical principles and classifying the inherent non-classical features.

A later version of the model [5], which we will call Spekkens' theory (ST), introduced a more general and mathematically rigorous formulation, extending the theory to systems of discrete prime dimension, where *dimension* refers to the maximum number of distinguishable measurement outcomes of observables in the theory, and continuous variable systems. Spekkens called these classical statistical theories with epistemic restrictions as *epistricted statistical theories*. By considering a particular epistemic restriction that refers to the symplectic structure of the underlying classical theory, the *classical complementarity principle*, theories with a rich structure can be derived. Many features of quantum mechanics are reproduced there, such as Heisenberg uncertainty principle, and many protocols introduced in the context of quantum information, such as teleportation. However, as an intrinsically non-contextual theory, it cannot reproduce quantum contextuality (and the related Bell non-locality), which, therefore arises as the signature of quantumness. Indeed, for odd prime dimensions and for continuous variables, ST was shown to be operationally equivalent to sub-theories of quantum mechanics, which Spekkens called quadrature quantum mechanics.

In the finite dimensional case quadrature quantum mechanics is better known as *stabiliser quantum mechanics* (SQM). The latter is a sub-theory of quantum mechanics developed for the description and study of quantum error correcting codes [6], but subsequently playing a prominent role in many important quantum protocols. In particular, many studies of quantum contextuality can be expressed in the framework of SQM, including the GHZ paradox [8] and the Peres–Mermin square [9, 10]. This exposes a striking difference between odd and even dimensional SQM. Even-dimensional SQM contains classical examples of quantum contextuality while odd-dimensional SQM exhibits no contextuality at all, necessary for its equivalence with ST. While developed for qubits, SQM was rapidly generalised to systems of arbitrary dimension [6]. However, for non-prime dimensions SQM remains poorly characterised and little studied (recent progress in this was recently reported in [12]).

Quasiprobability representations, such as the Wigner function, have been an important tool for the description of quantum systems for many years [24]. Recently, negative quasi probability representations and contextuality have been shown to have an important resource character in quantum computation [13–16, 17–22]. In particular in certain fault tolerant quantum computation schemes, SQM plays a central role as both the set of operations that can be directly fault tolerantly realised, and the part of the computation which is efficiently simulable by a classical computer [7]. Such computation can be then boosted to quantum universality by 'injecting' a resource state, known as a magic state [11]. In the case of odd prime dimensions, Howard *et al* [13] showed that the contextuality of the injected state is necessary for reaching universal quantum computation. Other similar results have been found in the case of qubits, at the cost of considering smaller subtheories than SQM for the classically simulatable non-contextual part of the computation [14–16]. The operational equivalence between SQM and ST in odd dimensions suggests to study the possible role of ST in this research field, thus also accomplishing the task of characterising its computational power.

In spite of the importance of Spekkens' theory, there remain some important aspects of it which have not yet been characterised and studied. First, all prior work on ST have only considered systems where the dimension is prime. Furthermore, while Spekkens' recent work strengthens the mathematical foundations of the model [5], one key part of the theory has not yet been described in a general and rigorous way. These are the measurement update rules, the rules which tell us how to update a state after a measurement has been made. In prior work, these rules, and the principles behind them have been described but not formalised.

In this paper, we complete this step, deriving a formal description of the measurement rules for prime-dimensional ST. Having done so, we now have a fully formal description of the model, which can be used as a basis to generalise it. We do so, generalising the framework from prime-dimensions to arbitrary dimensions and finding that it is the measurement update rule, where the richer properties of the non-prime dimension can be seen, which provides the key to this generalisation.

Having developed ST for all finite dimensions, we then focus on the general odd-dimensional case, and prove that in all odd-dimensional cases Spekkens' theory is equivalent to SQM. The bridge between SQM and ST is given by Gross' theory (GT) of discrete Wigner functions [23]. Unlike most other studies, Gross' treatment considered both prime and non-prime cases in its original formulation.

To summarise the contributions of this paper, we provide a complete formulation of ST in *all* discrete dimensions, even and odd, endowed with the updating rules for sharp measurements both for prime and non-prime dimensional systems. We extend the equivalence between ST and SQM via Gross' Wigner functions to *all* odd dimensions, and find the measurement updating rules also for the Wigner functions. The above equivalence allows us to shed light onto a complete characterisation of SQM in non-prime dimensions. Finally the incredibly elegant analogy between the three theories in odd dimensions: ST, SQM and GT, is depicted in terms of their updating rules.

The remainder of the paper is structured as follows. In section 1 we precisely and concisely describe the original framework of ST, in particular we define ontic and epistemic states, observables and the rule to obtain the outcome of the measurement of an observable given a state. In sections 2 and 3 we state and prove the updating rules in ST respectively for prime and non-prime dimensional systems. We prove these in two steps: first considering the case in which the state and measurement commute, and then the more general (non-commuting) case. The mathematical difference between the set of integers modulo d , for d prime and non-prime, results in having two levels of observables: the fundamental ones—the fine graining observables—and the ones that encode some degeneracy—the coarse-graining observables. The latter are problematic and are only present in the non-prime case. This is the reason why we need a different formulation in the two cases. The updating rules for the coarse graining observables will need a step in which the coarse-graining observables are written in terms of fine graining ones. In section 4 we state the equivalence of ST and SQM via Gross' Wigner functions in all odd dimensions. We also express the already found updating rules in terms of Wigner functions and we use them to depict the elegant analogies between these three theories. The paper ends with a discussion of the possible applications of our achievements and with a summary of the main results.

2. Spekkens' theory

We start by reviewing and introducing ST for prime-dimensional systems. We take a slightly different approach to [4, 5]. ST is a hidden variable theory, where the hidden variables are points in a phase space. The state of the hidden variables is called the *ontic state*. In Spekkens' model the ontic state is hidden and can never be known by an experimenter. The experimenter's best description of the system is the *epistemic state*, representing a probability distribution over the points in phase space.

For a single d -dimensional system, a phase space can be defined via the values of two conjugate fiducial variables, which we label X and P , in analogy to position and momentum. X and P can each take any value between 0 and $d - 1$, and a single ontic state of the system is specified by a pair (x, p) , where x is the value of X and p is the value of P . This phase space is equivalent to the space \mathbb{Z}_d^2 . In figure 1 three examples of epistemic states of one trit ($d = 3$) are depicted, where X and P are represented by the columns and rows in the phase space \mathbb{Z}_3 .

A collection of n systems is described by n pairs of independent conjugate variables X_j and P_j , with $j \in 0, \dots, n - 1$ a label indexing the systems. The phase space, denoted by Ω , is simply the cartesian product of single system phases spaces and thus $\Omega \equiv (\mathbb{Z}_d)^{2n}$.

The ontic state of the n -party system represents a set of values for each fiducial observables X_j and P_j . In other words, an ontic state is denoted by a point in the phase space $\lambda \in \Omega$. We call X_j and P_j observables because they correspond to measurable quantities, and assume that these observables are sufficient to uniquely define the ontic state. We can refer to Ω as a vector space where the ontic states are vectors (bold characters) whose components (small letters) are the values of the fiducial variables:

$$\lambda = (x_0, p_0, x_1, p_1, \dots, x_{n-1}, p_{n-1}). \quad (1)$$

Not only are the fiducial variables important for defining the state space, they also generate the set of all general observables in the theory. A generic observable, denoted by Σ , is defined by any linear combination of fiducial variables:

$$\Sigma = \sum_m (a_m X_m + b_m P_m), \quad (2)$$

where $a_m, b_m \in \mathbb{Z}_d$ and $m \in 0, \dots, n - 1$. The observables inhabit the dual space Ω^* , which is isomorphic to Ω itself. Therefore we can define them as vectors, in analogy with ontic states,

$$\Sigma = (a_0, b_0, a_1, b_1, \dots, a_{n-1}, b_{n-1}). \quad (3)$$

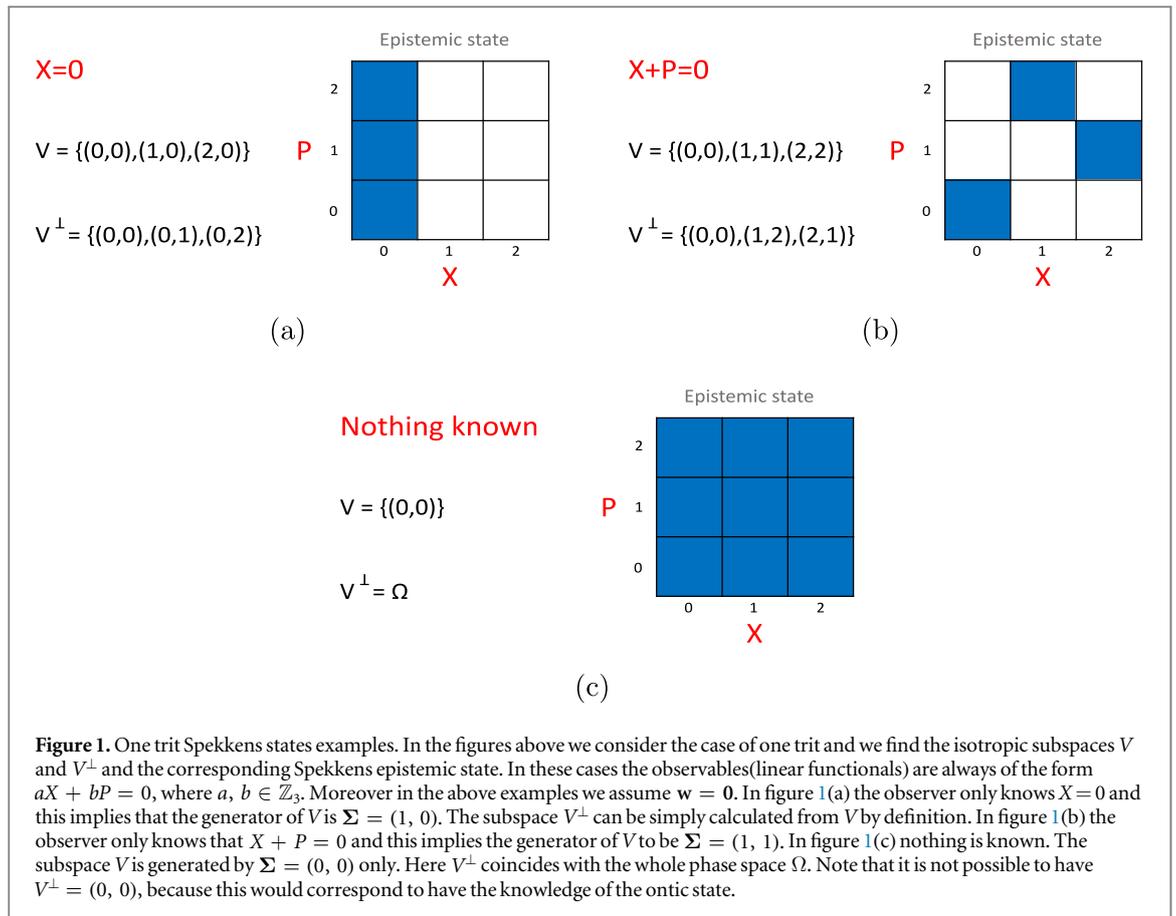
The formalism provides a simple way of *evaluating* the outcome σ of any observable measurement Σ given the ontic state λ , i.e. by computing their *inner product*:

$$\sigma = \Sigma^T \lambda = \sum_j (a_j x_j + b_j p_j), \quad (4)$$

where all the arithmetic is over \mathbb{Z}_d .

ST gains its special properties, and in particular, its close analogy with SQM via the imposition of an *epistemic restriction*, a restriction on what an observer can know about the ontic state of a system. The observer's best description is called the *epistemic state*, which is represented by a probability distribution $p(\lambda)$ over Ω (figure 1).

¹ The dimension d is any positive number, and we will not, in general, restrict it to odd or even, prime or non-prime, unless specified.



The epistemic restriction of ST is called *classical complementarity principle* and it states that two observables can be simultaneously measured only when their Poisson bracket is zero. This is motivated by SQM, since it captures the condition for two observables in SQM to *commute*. We shall adopt the quantum terminology here, and say that if the Poisson bracket between two observables is zero they commute. This can be simply recast in terms of the *symplectic inner product*:

$$\langle \Sigma_1, \Sigma_2 \rangle \equiv \Sigma_1^T J \Sigma_2 = 0, \tag{5}$$

where $J = \bigoplus_{j=1}^n \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}_j$ is the symplectic matrix. Note that each observable Σ_j partitions Ω into d subsets, each of the form $(\text{span}\{\Sigma_j\})^\perp + \mathbf{w}$, where \mathbf{w} is any ontic state such that $\Sigma_j^T \cdot \mathbf{w} = \sigma_j$.

Let us now consider sets of variables that can be jointly known by the observer. Such variables commute, and represent a sub-space of Ω known as an *isotropic subspace*. We denote the subspace of the known variables as $V = \text{span}\{\Sigma_1, \dots, \Sigma_n\} \subseteq \Omega$, where Σ_i denotes one of the generators (commuting observables) of V .

Sets of known commuting variables are important as these define the epistemic states within the theory. In particular, we can define an epistemic state by the set of variables V that are known by the observer and also the values $\sigma_1, \dots, \sigma_n$ that these variables take.

This means that $\Sigma_j^T \cdot \mathbf{w} = \sigma_j$, where $\mathbf{w} \in V$ is an ontic state that evaluates the known observables. We will call \mathbf{w} a *representative ontic state* for the epistemic state. More precisely we can state the following theorem.

Proposition 1. *The set of ontic states consistent with the epistemic state described by (V, \mathbf{w}) is*

$$V^\perp + \mathbf{w}, \tag{6}$$

where the perpendicular complement of V is, by definition, $V^\perp = \{a \in \Omega \mid \mathbf{a}^T \mathbf{b} = 0 \ \forall \ \mathbf{b} \in V\}$.

Proof. Let us start by considering the set of ontic states λ such that $\Sigma_j^T \lambda = 0 \ \forall j$. By definition of perpendicular complements, the ontic states λ belong to V^\perp . If we consider an ontic state \mathbf{w} such that $\Sigma_j^T \mathbf{w} = \sigma_j$, then $\Sigma_j^T (\lambda + \mathbf{w}) = \sigma_j$. Therefore the ontic states consistent with the epistemic state associated to (V, \mathbf{w}) are the ones of the kind $\lambda + \mathbf{w}$, i.e. the ones belonging to $V^\perp + \mathbf{w}$. \square

Note that the presence of $\mathbf{w} \neq \mathbf{0}$ simply implies a translation, that is why we can also call it *shift vector*.

By assumption the probability distribution associated to the epistemic state (V, \mathbf{w}) is uniform (indeed we expect all possible ontic states to be equiprobable), so the probability distribution of one of the possible ontic states in the epistemic state (V, \mathbf{w}) is

$$P_{(V, \mathbf{w})}(\lambda) = \frac{1}{d^n} \delta_{V^\perp + \mathbf{w}}(\lambda), \quad (7)$$

where the delta is equal to one only if $\lambda \in V^\perp + \mathbf{w}$ (note this means that the theory is a *possibilistic* theory). In figure 1 we specify the subspaces V and V^\perp in three different examples of epistemic states of one bit.

We can sum up our approach to Spekkens' model as follows:

- (i) Start from the intuitive formula (4) that relates observables Σ_j , ontic states λ and outcomes σ_j .
- (ii) Epistemic restriction: the compatible observables are the ones whose symplectic inner product is zero.
- (iii) Compute the shift vector \mathbf{w} . This allows us to shift back the set of points λ to obtain a subspace.
- (iv) The set of ontic states compatible with the epistemic state (V, \mathbf{w}) is $V^\perp + \mathbf{w}$, where V is the isotropic subspace spanned by the observables Σ_j (the set of known variables).

We say that this approach is physically intuitive because we start with equation (4), which is physically motivated and states, observables and the corresponding outcomes are defined in terms of it. Equation (4) also allows us to see that the shift comes from the need to recover the subspace structure.

3. Updating rules—*prime dimensional case*

The formulation of ST in [5], made for prime (and infinite) dimensional systems and described in the previous section, does not provide a full treatment of the transformative aspect of measurements, i.e. how the epistemic state has to be updated after a measurement procedure. In the following we will provide a proper formalisation of it, and in the next section we will generalise the formalism to all dimensions, non-prime too.

The set of integers modulo d shows different features depending on d being prime or not. In particular in the non-prime case it is not always possible to uniquely define the inverse of a number. The consequences of this will directly affect the updating rules. In particular the possible observables sometimes will not show full spectrum: some outcomes will not be possible because they would derive from arithmetics involving numbers with not well-defined inverses. This will divide the set of possible observables in two categories depending on whether they have full spectrum or not. We start from the prime case where problematic observables are not present because inverses always exist.

Like in quantum theory, duality in the description of states and measurements characterises ST. This means that we can represent the elements of a measurement Π in an epistemic-state way, (V_Π, \mathbf{r}) , where we can go from one element of the measurement to the other by simply shifting the representative ontic vector \mathbf{r} (see figure 2). In ST the measurement process corresponds to the process of *learning* some information (*aka* asking questions) about the ontic state of the system. According to the classical complementarity principle only the observables that are compatible (i.e. Poisson-commute) with the state of the system can be learned (jointly knowable). This means that the state after measurement will be given by the generators of the measurement and the generators of the state before the measurement, which are compatible with it². It is then fundamental to understand how compatible sets of ontic states (the isotropic subspaces of known variables V and their perpendicular V^\perp) change when independent observables are added and removed from the set of known variables V .

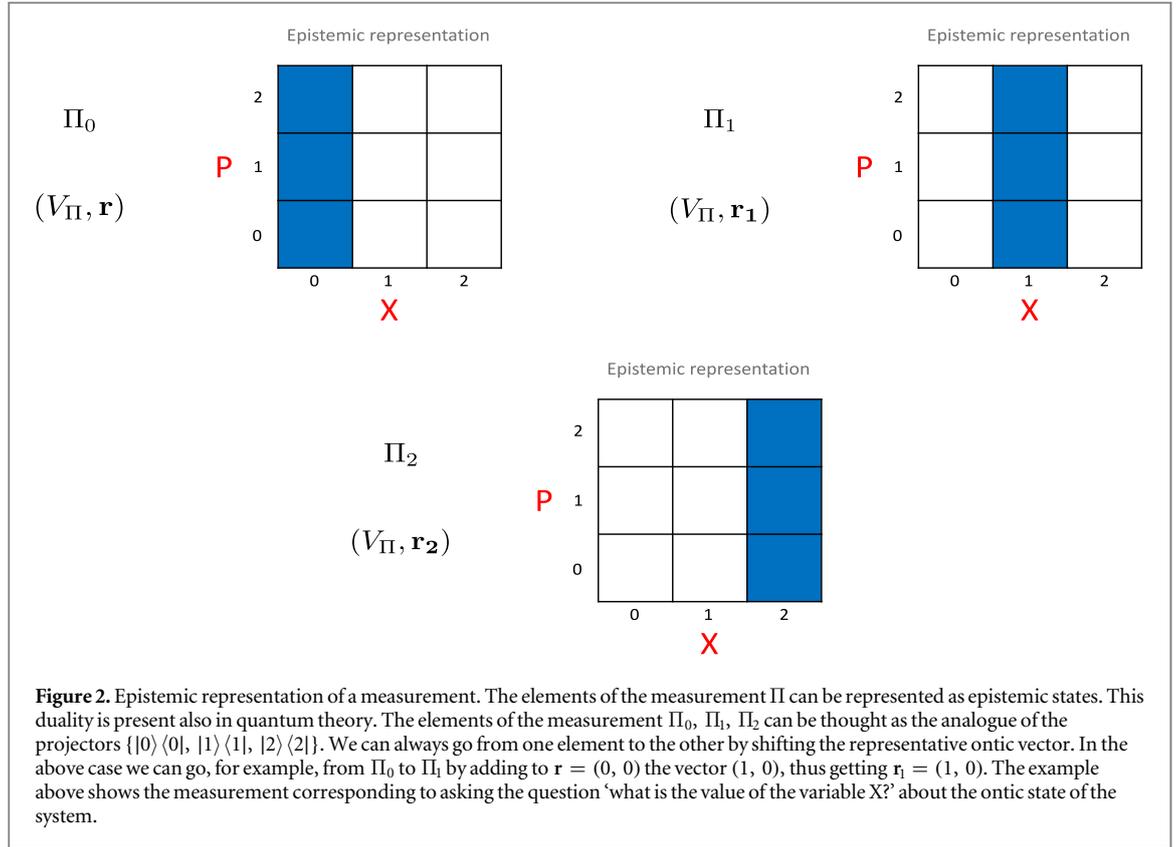
3.1. Adding and removing generators to/from V

- (i) Let us start with the case of *adding* a generator Σ' to the set of generators of $V = \text{span}\{\Sigma_1, \dots, \Sigma_n\}$. We assume that Σ' is linear independent with respect to the set spanned by the Σ_j . Let us see what happens to V^\perp . The subspace V after the addition becomes

$$V' = V \oplus \text{span}\{\Sigma'\}. \quad (8)$$

By definition the direct sum of two subspaces $A \oplus B$ returns a subspace such that for each $a \in A$ and $b \in B$, the sum $a + b$ belongs to $A \oplus B$. The direct sum of two subspaces is a subspace. We are interested in the orthogonal complement of a direct sum. It is well known that $(A \oplus B)^\perp = A^\perp \cap B^\perp$. This means that by

² As an abuse of language we here talk of generators of a state meaning the orthogonal basis set that generates the subspace of known variables associated with the state.



adding a generator to V , its perpendicular V^\perp is given by

$$V'^\perp = V^\perp \cap (\text{span}\{\Sigma'\})^\perp. \tag{9}$$

Note that V'^\perp is smaller than V^\perp .

- (ii) We now analyse what happens if we *remove* a generator, say Σ_n , from the set of generators of V . This means that now $V' = \text{span}\{\Sigma_1, \dots, \Sigma_{n-1}\}$. The set V^\perp is clearly contained in V'^\perp , since any vector orthogonal to all elements of V must also be orthogonal to all elements of V' . By definition, the set V'^\perp is composed by all the ontic states λ such that $\Sigma_j^T \lambda = 0$ for all $j < n$, but $\Sigma_n^T \lambda \neq 0$. This means that we need to remove the constraint $\Sigma_n^T \lambda = 0$ to enlarge V^\perp to V'^\perp , i.e. we simply need to add the ontic states $\lambda' = c\gamma$ to V^\perp , where $c \in \mathbb{Z}_d \neq 0$ and γ is a vector such that $\Sigma_n^T \gamma = 1$. Indeed this implies that

$$\Sigma_n^T(\lambda + \lambda') = \Sigma_n^T(\lambda + c\gamma) = 0 + c \neq 0.$$

In *prime* dimensions γ uniquely exists and it corresponds to $k^{-1}\Sigma_n$, where $k = \Sigma_n^T \Sigma_n$. Indeed the inverse of an integer $k \in \mathbb{Z}_d \neq 0$ always uniquely exists if d is a prime number. The formula for V'^\perp then reads

$$V'^\perp = \bigcup_c (V^\perp + ck^{-1}\Sigma_n) \equiv \bigcup_{w_n \in V_n} (V^\perp + w_n) = V^\perp \oplus V_n, \tag{10}$$

where the addition of $+w_n$ means that the whole set V^\perp is shifted by w_n , and $V_n = \text{span}\{\Sigma_n\}$. The previous trick in general works as follows. Given the ontic state λ , the observable Σ and the outcome σ associated with them, i.e. $\Sigma^T \lambda = \sigma$, then it is possible to shift the value σ by a constant k such that $\Sigma_n^T \Sigma_n = k$, by only adding Σ itself to the ontic state:

$$\Sigma^T(\lambda + \Sigma) = \sigma + \Sigma^T \Sigma = \sigma + k. \tag{11}$$

Note that the above identity allows us to change the value of the outcome associated with an ontic state by a constant factor (that we can also choose) without affecting any commuting observable (in this case Σ).

3.2. Measurement updating rules

We now want to find the updating rules for the state (V, \mathbf{w}) of a prime dimensional system when we perform a measurement (V_Π, \mathbf{r}) on it. We will consider V_Π being spanned by the generators denoted as Σ'_j . The representative ontic vector associated to the measurement, \mathbf{r} , is such that, by definition, $\Sigma'^T_j \mathbf{r} = \sigma'_j$, where the σ'_j are the outcomes associated with the measurement. The subspace of known variables V can be written in

terms of the sets generated by the generators Poisson-commuting with all the Σ'_j , V_{commute} , and non-commuting ones, V_{other} . According to this definition V_{commute} will always be a subspace. We cannot state the same for V_{other} , since the null vector does not belong to it. For this reason we augment V_{other} with the null vector in order to create a subspace. This implies that we can decompose V as

$$V = V_{\text{commute}} \oplus V_{\text{other}}. \quad (12)$$

We can also prove the following lemma.

Lemma 1. *The subspace V_{other} has dimension m , where m is the number of non-commuting generators of the measurement with the state.*

Proof. Let us initially assume the measurement to consist only of one non-commuting generator Σ' , so $m = 1$. Let us prove the lemma by contradiction. Let u, v be two orthogonal non-zero elements of V_{other} . Note that, by definition of a subspace, if $u, v \in V_{\text{other}}$, also a linear combination of u, v has to belong to V_{other} . By definition u, v do not commute with Σ' . Therefore we can write

$$\Sigma'^T J \mathbf{u} = a,$$

$$\Sigma'^T J \mathbf{v} = b,$$

where $a, b \neq 0$. In particular there will exist a constant $c \in \mathbb{Z}_d$ such that $a - bc = 0$. This implies that

$$\Sigma'^T J (\mathbf{u} - c\mathbf{v}) = 0.$$

Hence the linear combination $(u - cv)$ belongs to V_{commute} . This is a contradiction, therefore V_{other} has dimension 1. From the same reasoning, in the case of m non-commuting generators of the measurement, the subspace V_{other} has dimensions at maximum equal to m . Let us assume now that the dimension of V_{other} is $m - 1$. This is not possible because it would mean that, for example, Σ'_{m-1} can be written as a linear combination of $\Sigma'_0, \dots, \Sigma'_{m-2}$. However this is not the case because, by definition of basis set, all the generators are linearly independent. Therefore V_{other} has dimension m . \square

We will now provide the updating rules both for V and \mathbf{w} in two steps: first considering the state and measurement to commute, and then the general (non-commuting) case.

Theorem 1. (Commuting case). *Given the epistemic state (V, \mathbf{w}) and the measurement (V_{Π}, \mathbf{r}) that commutes with it, i.e. their generators all Poisson commute, the epistemic state (V', \mathbf{w}') after the measurement is described by*

$$V'^{\perp} = (V^{\perp} + \mathbf{w} - \mathbf{w}') \cap (V_{\Pi}^{\perp} + \mathbf{r} - \mathbf{w}'), \quad (13)$$

where \mathbf{w}' is given by equation

$$\mathbf{w}' = \mathbf{w} + \sum_i \Sigma_i'^T (\mathbf{r} - \mathbf{w}) \gamma_i, \quad (14)$$

where Σ_i' are the generators of the measurement Π and γ_i is such that $\Sigma_i'^T \gamma_i = 1$.

Proof. When the state and measurement commute we have to add the generators of the measurement to the set of generators of V , as we have seen in the previous section 3.1 (learning stage). Therefore the updating rule for the subspace V is (equation (8))

$$V \rightarrow V' = V \oplus \text{span}\{\Sigma'_0, \Sigma'_1, \dots, \Sigma'_i, \dots\} = V \oplus V_{\Pi}. \quad (15)$$

In terms of perpendicular subspaces this implies that $V'^{\perp} = V^{\perp} \cap V_{\Pi}^{\perp}$.

Let us initially assume the measurement to consist only of one generator Σ' . Let us recall that the outcome associated with Σ' is σ' . We assume \mathbf{w} is not compatible with this outcome, i.e. $\Sigma'^T \mathbf{w} = \sigma' + x$, for some shift $x \in \mathbb{Z}_d$, and we want to find \mathbf{w}' such that

$$\Sigma'^T \mathbf{w}' = \sigma'. \quad (16)$$

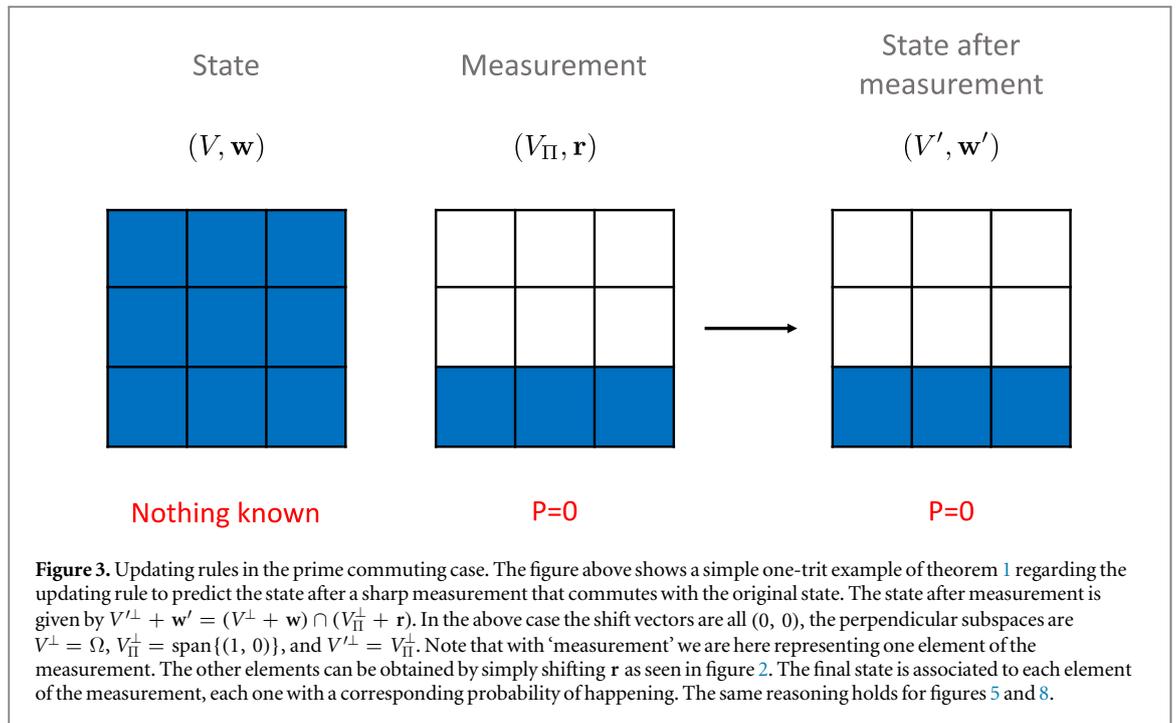
The identity (11) we used in the previous section does the job. More precisely,

$$\mathbf{w}' = \mathbf{w} - x\gamma,$$

where the vector γ is such that $\Sigma'^T \gamma = 1$. The above expression can be also written as

$$\mathbf{w}' = \mathbf{w} - k^{-1}x\Sigma',$$

where $k = \Sigma'^T \Sigma'$. The inverse of k always exists because we are in the prime dimensional case. Without referring to x we can restate the updating rule for the representative ontic vector as



$$\mathbf{w} \rightarrow \mathbf{w} + k^{-1}(\sigma' - \Sigma'^T \mathbf{w}) \Sigma' = \mathbf{w} + k^{-1} \Sigma'^T (\mathbf{r} - \mathbf{w}) \Sigma'. \quad (17)$$

Note that if we consider more than one generator of the measurement, we simply have to sum over all those generators in the second term. This immediately follows from considering the whole measurement Π as a sequence of measurements given by each generator Σ'_i and apply every time the rule (17). We state again that the above formula always holds for prime dimensional systems. We cannot claim the same in non-prime dimensions. The correct updating rule for the subspace V'^{\perp} is found by combining the updating rules for V and \mathbf{w} as in (13). This correction simply sets the subspaces to the same origin in order to correctly compute their intersection, as schematically shown in figure 4. At the end we obtain for the epistemic state (V', \mathbf{w}') that $V'^{\perp} + \mathbf{w}' = (V^{\perp} + \mathbf{w}) \cap (V_{\Pi}^{\perp} + \mathbf{r})$. We recall that the probability associated to each ontic state consistent with the epistemic state is uniform, i.e. given by

$$P(V', \mathbf{w}') = \frac{1}{|V'^{\perp} + \mathbf{w}'|} = \frac{1}{|V'^{\perp}|} = \frac{1}{|(V^{\perp} + \mathbf{w}) \cap (V_{\Pi}^{\perp} + \mathbf{r})|},$$

where $|\cdot|$ indicates the size of the subspace. □

Figure 3 shows a basic example of theorem 1.

Theorem 2. (Non-commuting case). Given the epistemic state (V, \mathbf{w}) and the measurement (V_{Π}, \mathbf{r}) that does not commute with it, i.e. some of the generators do not Poisson commute with the state, the epistemic state (V', \mathbf{w}') after the measurement is described by

$$V'^{\perp} = (V_{\text{commute}}^{\perp} + \mathbf{w} - \mathbf{w}') \cap (V_{\Pi}^{\perp} + \mathbf{r} - \mathbf{w}'), \quad (18)$$

where $V_{\text{commute}}^{\perp}$ is given by

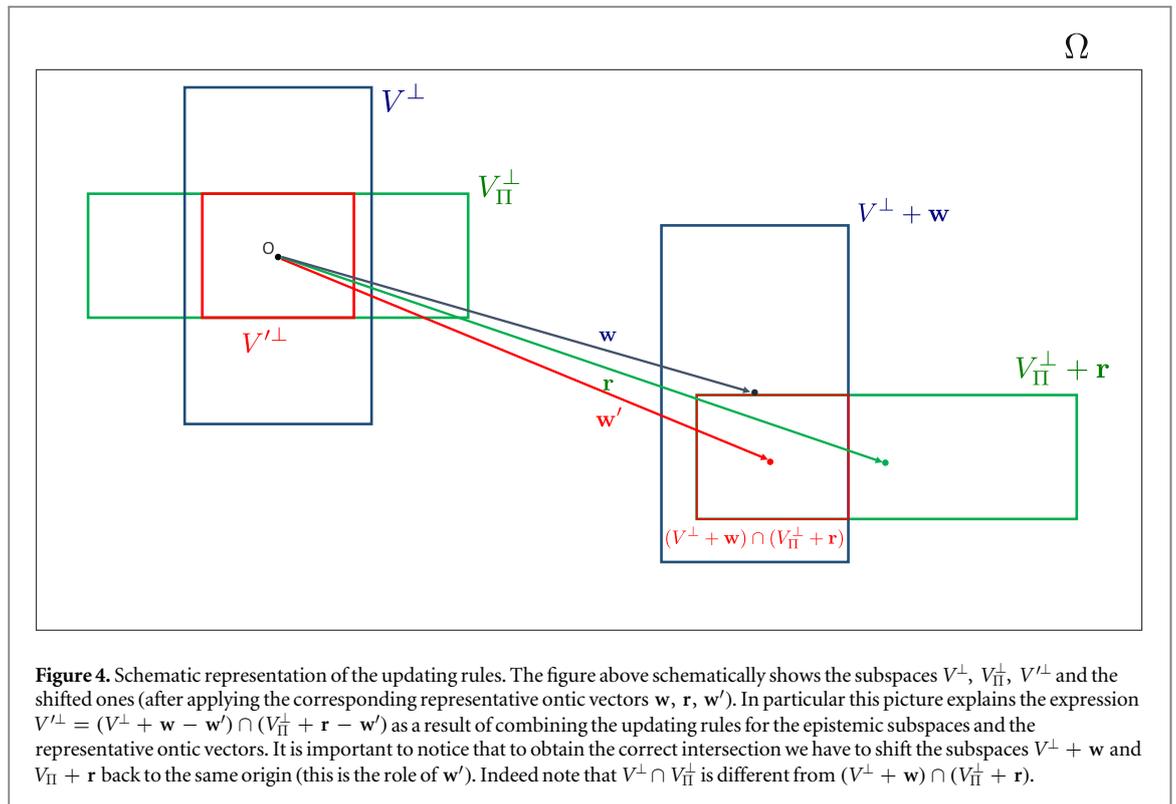
$$V_{\text{commute}}^{\perp} = V^{\perp} \oplus V_{\text{other}}. \quad (19)$$

The representative ontic vector \mathbf{w}' is given by

$$\mathbf{w}' = \mathbf{w} + \sum_i \Sigma_i'^T (\mathbf{r} - \mathbf{w}) \gamma_i, \quad (20)$$

where Σ_i' are the generators (even the non-commuting ones) of the measurement Π and γ_i is such that $\Sigma_i'^T \gamma_i = 1$.

Proof. Let us assume that Σ_j' , for $j \in \{0, \dots, m-1\}$, do not commute with the generators of V . In addition to the learning stage of the previous commuting case, we also have a removal stage of the disturbing part of the measurement. We have already seen that we can split the subspace V in $V = V_{\text{commute}} \oplus V_{\text{other}}$, where V_{other} is generated, from lemma 1, by all the Σ_j' , for $j \in \{0, \dots, m-1\}$. Therefore we can reduce to the commuting case if we only consider V_{commute} instead of the whole V . The updating rule for the subspace V then becomes



$$V \rightarrow V' = V_{\text{commute}} \oplus \text{span}\{\Sigma'_0, \Sigma'_1, \dots, \Sigma'_i, \dots\} = V_{\text{commute}} \oplus V_{II}$$

In terms of the perpendicular subspaces note that we can both write

$$V'^\perp = (V^\perp \oplus V_{\text{other}}) \cap V_{II}^\perp,$$

and

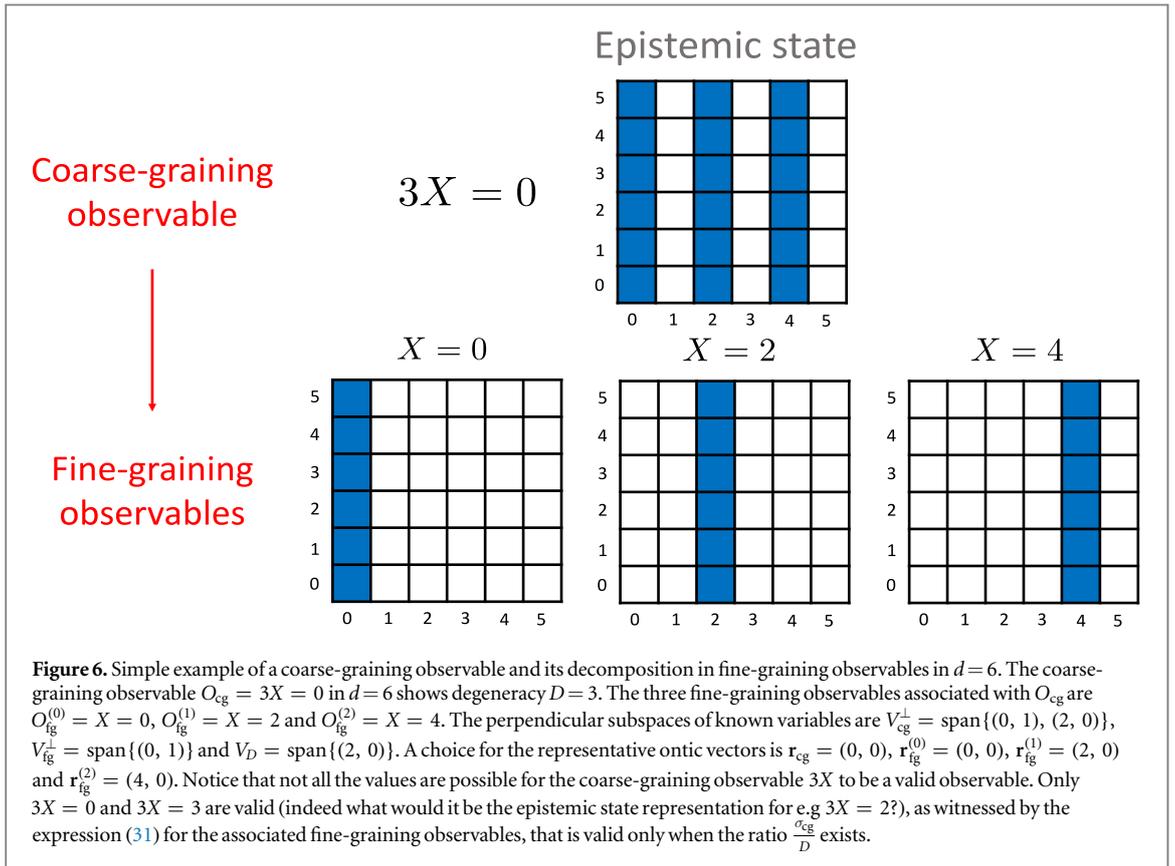
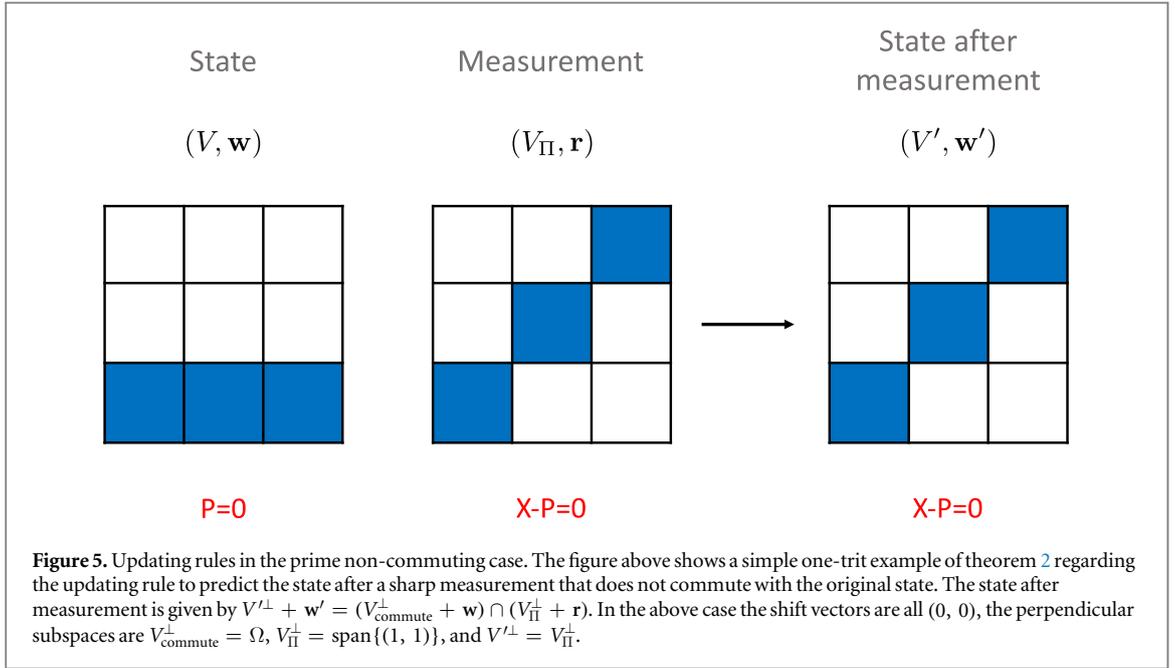
$$V'^\perp = V_{\text{commute}}^\perp \cap V_{II}^\perp,$$

from the usual property that the perpendicular of a direct sum is the intersection of the perpendicular subspaces. The updating rule for the representative ontic vector is the same as in the previous case (equation (14)). The correct updating rule for the subspace V'^\perp is found by combining the updating rules for V and w as in the previous case (13), where V^\perp is replaced by V_{commute}^\perp . At the end we obtain for the epistemic state (V', w') that $V'^\perp + w' = (V_{\text{commute}}^\perp + w) \cap (V_{II}^\perp + r)$. □

Figure 5 shows a basic example of theorem 1.

4. Updating rules—non prime dimensional case

It is quite common in studies of discrete theories, like Spekkens' model and SQM, to only consider the prime dimensional case because of the particular features of the set of integers modulo d , \mathbb{Z}_d , when d is non-prime, like the impossibility of uniquely define inverses of numbers. For example in our present case, figure 6 shows the peculiar properties of the observable $3X$ in $d = 6$, which has not full spectrum of outcomes. The general formulation of Spekkens' model of section 2 does not change; not even the rules for calculating the probabilities of outcome and the updating of the state after a reversible evolutions (which are present in [5]). The new formulation we provide affects the observables and the related measurements updating rules. More precisely our issue, as already noticed, regards the updating-rule formula (14) and (19) for the shift vector w' and the subspace V_{commute}^\perp , which do not always hold when the dimension d is non-prime. In fact the vector γ_i such that $\Sigma_i'^T \gamma_i = 1$ does not always exist in that case. On the other hand, in prime dimensions, it always uniquely exists because $\gamma_i = k_i^{-1} \Sigma_i'$ and the inverse of the integer $k_i = \Sigma_i'^T \Sigma_i'$ always uniquely exists. Unlike the original formulation due to Spekkens, we will now characterise Spekkens' model in non-prime dimensions. In particular we characterise which are the observables that are problematic in the above sense—the *coarse-graining* observables, like $3X$ in $d = 6$ —and we then find the updating rules for a state subjected to the measurement of such observables by rewriting them in terms of non-problematic observables—the *fine-graining* observables.



In the next subsection we assume single-system observables (i.e. of the kind $\Sigma' = aX + bP$, $a, b \in \mathbb{Z}_d$) in order to soften the notation and facilitate the comprehension. This will bring more easily to the updating rules even in the most general case of many systems (section 4.2). In this case we recall, without making any reference to the quantity k^{-1} , but just in terms of the vector γ , the updating rule for the shift vector w' ,

$$w' = w - x\gamma, \tag{21}$$

where, as usual, $x = -\Sigma'^T(\mathbf{r} - \mathbf{w})$, and the expression for V_{commute}^\perp ,

$$V_{\text{commute}}^\perp = \bigcup_c (V^\perp + c\gamma). \quad (22)$$

4.1. Coarse-graining and fine-graining observables

We define a fine-graining observable as an observable that has *full spectrum*, i.e. it can assume all the values in \mathbb{Z}_d . On the contrary a coarse-graining observable has not full spectrum.

Lemma 2. *An observable O_{fg} has full spectrum, i.e. it is a fine-graining observable, if and only if it has the following form,*

$$O_{\text{fg}} = a'X + b'P, \quad (23)$$

where $a', b' \in \mathbb{Z}_d$ are such that they do not share any integer factor or power factor of d .

On the contrary a coarse-graining observable is written as

$$O_{\text{cg}} = aX + bP = D(a'X + b'P), \quad (24)$$

where $a', b' \in \mathbb{Z}_d$ are again such that they do not share any integer factor or power factor of d and D is a factor shared by $a, b \in \mathbb{Z}_d$. More precisely the factor D is called *degeneracy* and it is defined as

$$D = D_1^{n_1} \cdot D_2^{n_2} \cdot \dots, \quad (25)$$

where D_1, D_2, \dots are different integer factors of d shared by a and b , and n_1, n_2, \dots are the maximum powers of these factor such that they can still be grouped out from a and b . We take the maximum powers because we want the remaining part, $a'X + b'P$, to not share any common integer factor or power factor of d between a' and b' . In this way we can associate a fine-graining observable to a coarse graining one by simply dropping the degeneracy D from the latter.

Proof. Let us first prove that an observable of the kind (23), $O_{\text{fg}} = a'X + b'P$, is a full spectrum one. This can be proven by using Bezout's identity [30]: let a' and b' be nonzero integers and let D be their greatest common divisor. Then there exist integers X and P such that $aX + bP = D$. In our case the greatest common divisor D is equal to one, since a', b' are coprime³. Therefore we have proven that there exist values of the canonical variables $X, P \in \mathbb{Z}_d$ such that $O_{\text{fg}} = a'X + b'P = 1$. In order to reach all the other values of the spectrum we simply need to multiply both X and P in the previous equation by $j \in \mathbb{Z}_d$.

We now prove the converse, i.e. that a full spectrum observable implies it to be written as (23). We prove this by seeing that an observable written as (24) has not full spectrum, i.e. we negate both terms of the reverse original implication. Proving the latter is straightforward, since the multiplication modulo d between an arbitrary quantity and a factor D , which is given by powers of integer factors of d , gives as a result a multiple of D . Since the multiples of D do not cover the whole \mathbb{Z}_d , then any observable of the form (24) has not full spectrum⁴. Since an observable of the form (23) is an observable that cannot be written as (24) by definition, we obtain that a full spectrum observable implies the observable to be written as (23). \square

Given lemma 2 we have got the expressions (23) and (24) for coarse-graining and fine-graining observables. We want now to prove the following lemma to ensure that fine-graining observables are characterised by precisely defined updating rules.

Lemma 3. *The vector γ in the updating rule (21) for the shift vector \mathbf{w}' and in the equation (22) for the subspace V_{commute}^\perp exists if and only if the observable is a fine-graining one.*

Proof. Let us prove that if we have a fine graining observable the vector γ exists. In our case $\Sigma' = (a', b')$ and, by definition of a', b' (as usual defined for fine-graining observables) and full spectrum, we can always find a vector $\gamma = (\gamma_a, \gamma_b)$ such that $\Sigma'^T \gamma = a'\gamma_a + b'\gamma_b$ equals 1.

Let us prove the converse. We now have the vector γ such that $\Sigma'^T \gamma = a\gamma_a + b\gamma_b = 1$, where the coefficients $a, b \in \mathbb{Z}_d$ define our observable $aX + bP = \sigma$. We want to prove that σ can achieve all the values of \mathbb{Z}_d . Since $\Sigma'^T \gamma = 1$ we can set the values of (X, P) as equal to (γ_a, γ_b) in order to reach the value $\sigma = 1$. We can

³ It could be that a', b' share a factor which is not a factor of d . In this case the argument follows identically as if they were coprime.

⁴ Multiples of D do not cover the whole spectrum of \mathbb{Z}_d because D has not an inverse D^{-1} (it is not coprime with d) and so we cannot obtain the whole values σ of \mathbb{Z}_d by simply finding X, P such that $a'X + b'P = D^{-1}\sigma$.

now achieve all the other values of the spectrum by simply redefining γ as $\tilde{\gamma} = c\gamma$, where c assumes all the values in \mathbb{Z}_d . \square

The above lemma 3 should convince us that in order to find the updating rules in the presence of a coarse graining observable, it is appropriate to decompose it in terms of fine-graining observables. Let us assume that our coarse-graining observable is $O_{cg} = aX + bP = D(a'X + b'P) = \sigma$, and the associated isotropic subspace and representative ontic vector are $(V_{cg}, \mathbf{r}_{cg})$. To this observable we can associate \bar{D} different fine-graining observables $O_{fg} = a'X + b'P = \sigma_j$, where $j \in \{0, \dots, \bar{D} - 1\}$. The quantity \bar{D} is the degeneracy D without the powers n_1, n_2, \dots , i.e. $\bar{D} = D_1 \cdot D_2 \cdot \dots$. Indeed the powers n_1, n_2, \dots simply represent multiplicities associated to each corresponding fine-graining observable. The associated isotropic subspaces and representative ontic vectors are $(V_{fg}^{(j)}, \mathbf{r}_{fg}^{(j)})$, where $V_{fg}^{(j)} = \text{span}\{(a', b')\}$ (see figure 6).

By definition the perpendicular isotropic subspaces are

$$V_{cg}^\perp = \{\mathbf{v} = (v_a, v_b) \in \Omega | v_a a + v_b b = D(v_a a' + v_b b') = 0 \pmod{d}\}, \quad (26)$$

$$V_{fg}^\perp = \{\mathbf{v}' = (v'_a, v'_b) \in \Omega | v'_a a' + v'_b b' = 0 \pmod{d}\}. \quad (27)$$

It is clear that $V_{cg}^\perp \supset V_{fg}^\perp$ and we can therefore construct V_{cg}^\perp as

$$V_{cg}^\perp = \bigcup_{j=0}^{\bar{D}-1} (V_{fg}^\perp + \mathbf{v}_j) = V_{fg}^\perp \oplus V_D, \quad (28)$$

where the subspace V_D provides all the vectors that we need to combine with the vectors of V_{fg}^\perp to reach the whole V_{cg}^\perp . We call the subspace V_D the *degeneracy subspace* because it encodes the degeneracy of V_{cg} with respect to V_{fg} . It has dimension 1 and size \bar{D} . This is consistent with the fact that the dimensions of V_{cg}^\perp and V_{fg}^\perp are respectively 2 and 1. The sizes are respectively $\bar{D} \cdot d$ and d . The size of V_{fg}^\perp is d because it is always a maximally isotropic subspace and its dimension is 1 because from one generator we get all the other vectors of the subspace by multiplication with $j \in \mathbb{Z}_d$. The dimension V_{cg}^\perp is 2 because it cannot be 1 (it would be the same subspace as V_{fg}^\perp) and it cannot be greater than 2 since also the whole phase space $\Omega = \mathbb{Z}_d^2$ has dimension 2. In order to know the size of V_{cg}^\perp we need to count all the $j\mathbf{v}$, where $j \in \{0, 1, \dots, \bar{D} - 1\}$, that means $\bar{D} \cdot d$. Therefore it can be written as $V_D = \text{span}\{\mathbf{v}\}$, and all its \bar{D} vectors are of the kind $\mathbf{v}_j = j\mathbf{v}$. The above reasoning easily extends to the case of n systems, where the dimensions are $\dim(V_{cg}^\perp) = 2n$, $\dim(V_{fg}^\perp) = n$, $\dim(V_D) = n$, and the sizes are $|V_{cg}^\perp| = \bar{D}^n d^n$, $|V_{fg}^\perp| = d^n$, $|V_D| = \bar{D}^n$. We can now prove that V_D is a vector space.

Proof. The definition of V_D is

$$V_D = \{\mathbf{v} \in \Omega | \alpha \mathbf{w} + \beta \mathbf{v} = \mathbf{t}, \text{ where } \mathbf{w} \in V_{fg}^\perp, \alpha, \beta \in \mathbb{Z}_d, \mathbf{t} \in V_{cg}^\perp\}. \quad (29)$$

To see that it is a vector space we just need to see that $(0, 0)$ belongs to V_D and that V_D is closed under addition and multiplication, i.e. under linear combinations. The null vector belongs to V_D because in the definition (29) we would remain with $\alpha \mathbf{w} = \mathbf{t}$, where $\mathbf{w} \in V_{fg}^\perp$ and $V_{fg}^\perp \subset V_{cg}^\perp$. Let us imagine that we have two vectors $\mathbf{v}, \mathbf{z} \in V_D$. Is the vector $\gamma \mathbf{v} + \delta \mathbf{z}$, where $\gamma, \delta \in \mathbb{Z}_d$, still belonging to V_D ? It is easy to see that if we apply the definition (29) we would get

$$\alpha \mathbf{w} + \beta(\gamma \mathbf{v} + \delta \mathbf{z}),$$

which can be rewritten as

$$(\alpha \mathbf{w} + \beta \gamma \mathbf{v}) + (0 \cdot \mathbf{w} + \beta \delta \mathbf{z}),$$

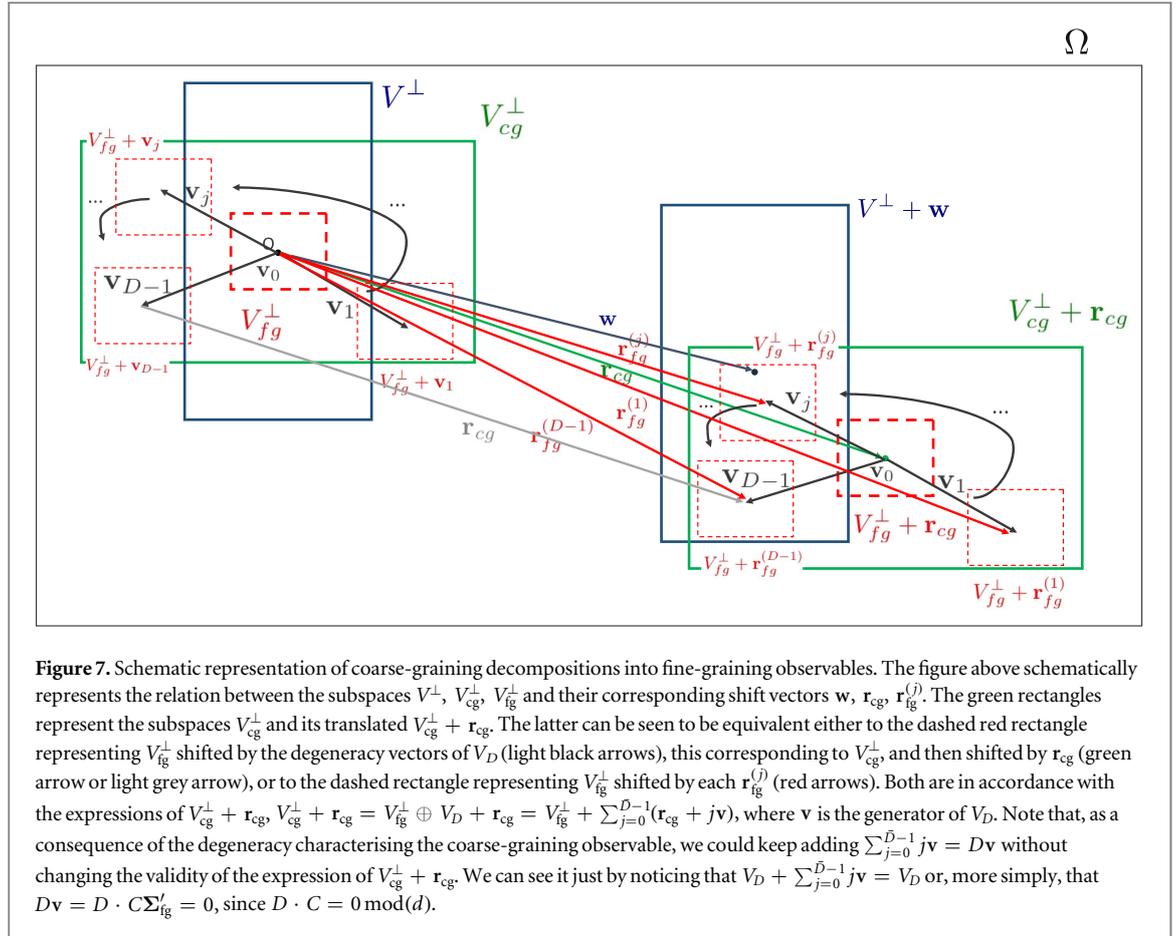
where each of the two terms in parenthesis belong to V_{cg}^\perp , and therefore the whole expression belongs to it too. \square

We now define the shift vectors $\mathbf{r}_{fg}^{(j)}$ in terms of \mathbf{r}_{cg} and see that we can encode the degeneracy expressed by V_D in there. The idea is schematically depicted in figure 7.

Given the shift vector associated to the coarse-graining observable \mathbf{r}_{cg} , the shift vectors $\mathbf{r}_{fg}^{(j)}$ associated to the corresponding fine-graining observables are of the kind

$$\mathbf{r}_{fg}^{(j)} = \mathbf{r}_{cg} + \mathbf{v}_j, \quad (30)$$

where $\mathbf{v}_j \in V_D$ and are therefore of the kind $j\mathbf{v}$, where $j \in \{0, \dots, \bar{D} - 1\}$. This implies that if we assume the outcome associated to the coarse-graining observable to be σ_{cg} , i.e. $\sum_{cg}^T \mathbf{r}_{cg} = \sigma_{cg}$, where $\sum_{cg} = (a, b)$, then the outcomes associated to the fine graining-observables are



$$\Sigma_{fg}^T \mathbf{r}_{fg}^{(j)} = \Sigma_{fg}^T (\mathbf{r}_{cg} + j\mathbf{v}) = \frac{\sigma_{cg}}{D} + jC, \quad (31)$$

where C is the *anti-degeneracy* and it is defined as a non-zero number belonging to \mathbb{Z}_d such that $D \cdot C = 0 \pmod{d}$. The idea is that the vector $\mathbf{v} \in V_D$ is such that $\Sigma_{fg}^T \mathbf{v} = C \neq 0$, so it does not belong to V_{fg}^\perp , but it does belong to V_{cg}^\perp , since $D \cdot C = 0 \pmod{d}$. An easy way to find one of the possible \mathbf{v} is to calculate it as $C \Sigma_{fg}^\perp$, where Σ_{fg}^\perp is the generator of V_{fg}^\perp . In this way we know that $D\mathbf{v} = 0$, but \mathbf{v} does not belong to V_{fg}^\perp , i.e. $v_a a' + v_b b' \neq 0$ because Σ_{fg}^\perp is not in V_{fg}^\perp . It is important to notice that equation (31) implies that not all the outcomes are allowed for the fine-graining observables associated to the coarse-graining one; they are allowed only when the ratio $\frac{\sigma_{cg}}{D}$ exists. Figure 6 also explains this fact.

4.2. Measurement updating rules

Let us assume to have n systems and to measure the coarse-graining observable $O_{cg} = a_1 X_1 + b_1 P_1 + \dots + a_n X_n + b_n P_n = D(a_1' X_1 + b_1' P_1 + \dots + a_n' X_n + b_n' P_n) = \sigma_{cg}$, with corresponding isotropic subspace of known variables V_{cg} and shift vector \mathbf{r}_{cg} , on the state $\rho = \alpha_1 X_1 + \beta_1 P_1 + \dots + \alpha_n X_n + \beta_n P_n = \sigma$, with corresponding isotropic subspace of known variables $V = \text{span}\{\Sigma_1, \dots, \Sigma_n\}$ and shift vector \mathbf{w} . The idea in order to find the updating rules for the state after measurement, the subspace of known variable V' and the representative ontic vector \mathbf{w}' is to compute the updating rule of the initial state ρ with the fine-graining observables that are associated to the coarse graining observable O_{cg} , i.e.

$O_{fg}^{(j)} = a_1' X_1 + b_1' P_1 + \dots + a_n' X_n + b_n' P_n = \sigma_{fg}^{(j)}$ (indeed we know that the updating rules are valid for them from theorem 3), and then combine them together. More precisely, the following theorem holds.

Theorem 3. Given the epistemic state (V, \mathbf{w}) and a coarse-graining measurement $(V_{cg}, \mathbf{r}_{cg})$, the epistemic state (V', \mathbf{w}') after the measurement is described by

$$V'^\perp = \bigcup_{j=0}^{D-1} [(V_{\text{commute}}^\perp + \mathbf{w} - \mathbf{w}') \cap (V_{fg}^\perp + \mathbf{r}_{fg}^{(j)} - \mathbf{w}')], \quad (32)$$

where the shift vector \mathbf{w}' is the shift vector deriving from the updating rule of the state after the measurement of the fine-graining observable $O_{fg}^{(j)}$,

$$\mathbf{w}' = \mathbf{w}'_j = \mathbf{w} + \sum_{i=0}^n \Sigma_i'^T (\mathbf{r}_{\text{fg}}^{(j)} - \mathbf{w}) \gamma_i, \quad (33)$$

where the vectors γ_i are defined such that $\Sigma_i'^T \gamma_i = 1$, and Σ_i' are the n generators of the subspace V_{fg} associated to the fine-graining observable $O_{\text{fg}}^{(j)}$. The subspace V_{commute}^\perp is given by the original V after having removed the non-commuting part, i.e. equation (19),

$$V_{\text{commute}}^\perp = \bigcup_{c=1}^d \left(V^\perp + \sum_{l=N+1}^n c \gamma_l \right) = V^\perp \bigoplus_{l=N+1}^n V_l = V^\perp \oplus V_{\text{other}}, \quad (34)$$

where γ_l is such that $\Sigma_l'^T \gamma_l = 1$ and V_l are the subspaces spanned by the $(n - N)$ non-commuting generators Σ_l . Obviously if the state and measurement commute, then $V_{\text{commute}}^\perp = V^\perp$.

The above theorem tells us that the way we combine the updating subspaces of the state with each individual fine-graining observables is through their union. This result is clear in terms of schematic diagrams (figure 7). The updated shift vector is just one of the updated shift vectors of the state with the fine-graining observables, because the information needed to update the shift vector of the state is encoded in just one of the fine-graining shift vectors. The degeneracy includes a meaningless multiplicity in the coarse-graining shift vector, and therefore every fine-graining observable can do the job of correctly updating the shift vector of the state. Actually every combination of the shift vectors \mathbf{w}'_j can do the job, apart from the ones that sum to 0 mod(d), like $\sum_{j=0}^{\bar{D}-1} \mathbf{w}'_j$. Note also that in the definition of V_{commute}^\perp the vector γ_l is, in general, degenerate. This is not a problem because any degenerate value of γ_l brings to the same subspace V_{commute}^\perp , since by definition its role is to add the vectors $\lambda' = c \gamma_l$ to V^\perp such that $\sum_n^T \lambda' \neq 0$.

Proof. We find the expression for the updated subspace V'^\perp by simply reusing the already found formulas (13) and (19) of the prime-dimensional case and substituting V_{Π}^\perp with V_{cg}^\perp and \mathbf{r} with \mathbf{r}_{cg} ,

$$V'^\perp = (V_{\text{commute}}^\perp + \mathbf{w} - \mathbf{w}') \cap (V_{\text{cg}}^\perp + \mathbf{r}_{\text{cg}} - \mathbf{w}').$$

If we now consider the decomposition of $V_{\text{cg}}^\perp + \mathbf{r}_{\text{cg}}$ as in (28) and (30), we obtain

$$V'^\perp = (V_{\text{commute}}^\perp + \mathbf{w} - \mathbf{w}') \cap \left[\bigcup_{j=0}^{\bar{D}-1} (V_{\text{fg}}^\perp + \mathbf{r}_{\text{fg}}^{(j)}) - \mathbf{w}' \right].$$

Since the intersection of a union is the union of the intersections, we have proven the first part of the theorem,

$$V'^\perp = \bigcup_{j=0}^{\bar{D}-1} [(V_{\text{commute}}^\perp + \mathbf{w} - \mathbf{w}') \cap (V_{\text{fg}}^\perp + \mathbf{r}_{\text{fg}}^{(j)} - \mathbf{w}')].$$

The second part of the proof regards \mathbf{w}' being equal to any of the \mathbf{w}'_j . Because of the degeneracy, any \mathbf{w}'_j is equivalent to the others (with different value of j) in order to provide us with \mathbf{w}' , indeed it is possible to find one from another just by adding a vector $\mathbf{v} \in V_D$. The latter can be proven as follows. For simplicity let us assume to be in the case $n = 1$ and that \mathbf{v} is the generator of V_D . We know that, by the definition of state after measurement of a fine-graining observable, the updated shift vector \mathbf{w}'_j is such that $\Sigma_{\text{fg}}^T \mathbf{w}'_j = \frac{\sigma_{\text{cg}}}{D} + jC = \sigma_{\text{fg}}^{(j)}$, where $C = \Sigma_{\text{fg}}^T \mathbf{v}$ is the antidegeneracy (equation (31)). It is straightforward to see that if we add \mathbf{v} to \mathbf{w}'_j , we get $\mathbf{w}'_j + \mathbf{v} = \mathbf{w}'_{j+1}$, indeed $\Sigma_{\text{fg}}^T (\mathbf{w}'_j + \mathbf{v}) = \frac{\sigma_{\text{cg}}}{D} + (j + 1)C = \sigma_{\text{fg}}^{(j+1)}$. \square

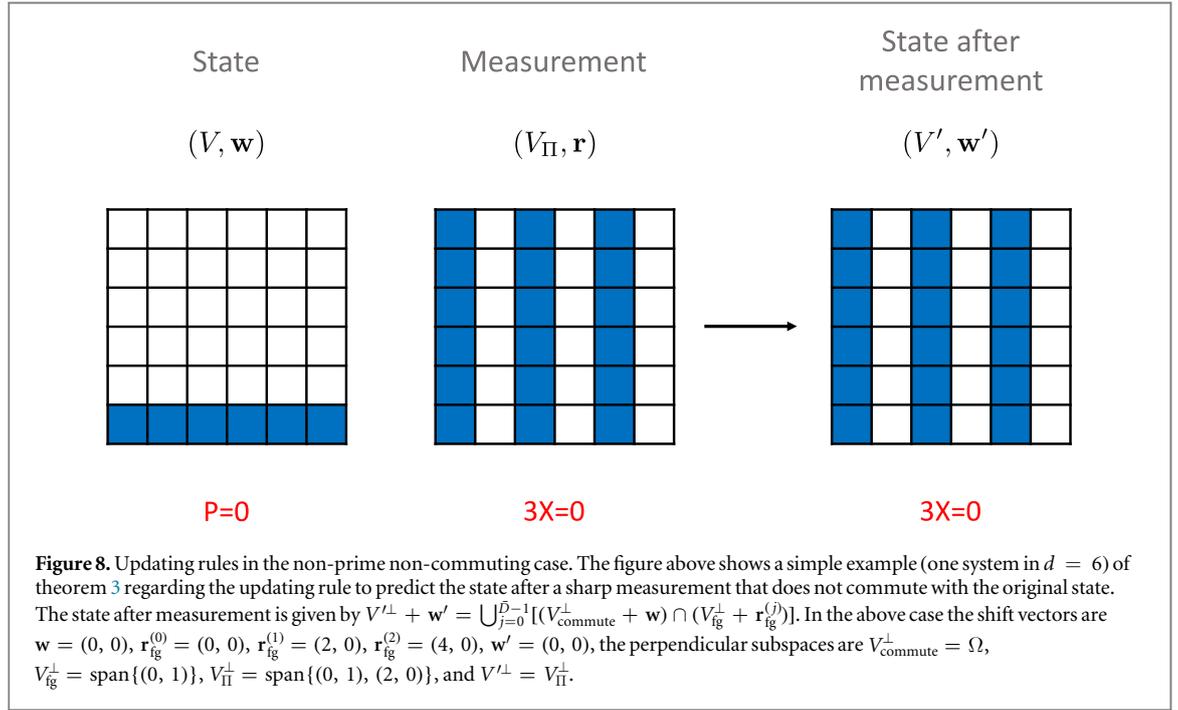
Figure 8 shows a basic example of theorem 3.

5. Equivalence of ST and SQM in all odd dimensions

In [5] it has been shown that SQM and Spekkens' toy model are two operationally equivalent theories in odd prime dimensions via GT of discrete non-negative Wigner functions. We have generalised Spekkens' model to all discrete dimensions. The above equivalence does not hold in even dimensions, but we will now see that it holds in *all* odd dimensions. We will also state the equivalence in terms of the updating rules, where all its elegance arises. We recall that SQM and GT of non-negative Wigner functions are equivalent in all odd dimensions [23].

5.1. SQM—updating rules

SQM is a subtheory of quantum mechanics where we only consider common eigenstates of tensors of Pauli operators, unitaries belonging to the Clifford group, and Pauli measurements [6]. We can always write a



stabiliser state ρ as

$$\rho = \frac{1}{\mathcal{N}} \rho_1 \cdot \rho_2 \cdot \dots \cdot \rho_N, \tag{35}$$

where $\mathcal{N} = \text{Tr}[\rho_1 \cdot \rho_2 \cdot \dots \cdot \rho_N]$, $j \in \{1, \dots, N \leq n\}$, n is the number of qudits and

$$\rho_j = (\mathbb{I}_d + g_j + g_j^2 + \dots + g_j^{d-1}), \tag{36}$$

where g_j is a stabiliser generator, more precisely a Weyl operator:

$$\hat{W}(\lambda) = \chi(pq) \hat{S}(q) \hat{B}(p), \tag{37}$$

where $\chi(pq) = e^{\frac{2\pi i}{d} pq}$, q, p are the coordinates of the phase space point $\lambda = (q, p)$, and \hat{S}, \hat{B} are respectively the shift and boost operators (generalised Pauli operators) and the arithmetics is modulo d ,

$$\hat{S}(q) = \sum_{q' \in \mathbb{Z}_d} |q' - q\rangle \langle q'| \tag{38}$$

$$\hat{B}(p) = \sum_{q \in \mathbb{Z}_d} \chi(pq) |q\rangle \langle q|. \tag{39}$$

When considering more than one qudit, the Weyl operator is given by the tensor product of the single Weyl operators. We can write the stabiliser state ρ in a more compact way as

$$\rho = \frac{1}{\mathcal{N}} \prod_j^n \sum_i^{d-1} g_j^i. \tag{40}$$

However we will mostly use the following notation in terms of stabiliser generators,

$$\rho \rightarrow \langle g_1, \dots, g_N \rangle. \tag{41}$$

We now analyse the updating rules for the state ρ under the stabiliser measurement Π ,

$$\Pi \rightarrow \langle p_1, \dots, p_M \rangle \tag{42}$$

where p_k is a stabiliser generator of Π and $k \in \{1, \dots, M \leq n\}$. We analyse the updating rules first in the commuting case ($[\rho, \Pi] = 0$) and then in the general case.

- (i) For *non-disturbing* (commuting) measurements, the state after measurement ρ' is given by adding the stabiliser generators of the measurement Π and the state ρ , unless some generators coincide. In the latter case we obviously count them only once.

$$\rho' \rightarrow \langle g_1, g_2, \dots, g_N, p_1, p_2, \dots, p_M \rangle, \tag{43}$$

where we have here considered the case in which no generators coincide. This formula means that the state

ρ' is now

$$\rho' = \frac{1}{\mathcal{N}} \prod_j^{N^*} \sum_i^{d-1} r_j^i,$$

where $N^* = N + M$ and r_j is a stabiliser generator of ρ' , i.e. it is either a valid (commuting) generator g_j or p_j . In the case where e.g. F generators coincide, then $N^* = N + M - F$.

- (ii) For *disturbing* (non-commuting) measurements (the most general case) the idea is that if we remove the non-commuting factors ρ_j from the state ρ , i.e. $[\rho_j, \Pi] \neq 0$, this case reduces to the previous commuting one. We assume the state ρ to have only one non-commuting factor, say ρ_N , which corresponds to the stabiliser generator g_N . The state after measurement ρ' is given by removing the non-commuting generator and adding the remaining ones of the state and measurement, unless some generators coincide. In the latter case we obviously count them only once.

$$\rho' \rightarrow \langle g_1, g_2, \dots, g_{N-1}, p_1, p_2, \dots, p_M \rangle, \tag{44}$$

where we have here considered the case in which no generators coincide. This formula means that the state ρ' is now

$$\rho' = \frac{1}{\mathcal{N}} \prod_j^{N^*} \sum_i^{d-1} r_j^i,$$

where $N^* = N + M - 1$ and r_j is a stabiliser generator of ρ' , i.e. it is either a valid (commuting) generator g_j or p_j . In the case where e.g. F generators coincide, then $N^* = N + M - 1 - F$.

To sum up, in the commuting case we add generators of state and measurement to obtain the state after measurement. In the non-commuting case we remove the non-commuting generator of the state and add all the others as in the commuting case. This structure is perfectly analogue to Spekkens' updating rules, which are just motivated by the classical complementarity principle.

5.2. Gross' Wigner functions—updating rules

Gross theory. In GT the Wigner function of a state ρ in a point of the phase space $\lambda \in \Omega$ is given by

$$W_\rho(\lambda) = \text{Tr}[\hat{A}(\lambda)\rho], \tag{45}$$

where $\hat{A}(\lambda)$ is the phase point operator associated to each point λ ,

$$\hat{A}(\lambda) = \frac{1}{d^n} \sum_{\lambda' \in \Omega} \chi(\langle \lambda, \lambda' \rangle) \hat{W}(\lambda'), \tag{46}$$

where $\hat{W}(\lambda)$ are the Weyl operators defined in equation (37). Note that the normalisation is such that $\text{Tr}[\hat{A}(\lambda)] = 1$. We recall that a stabiliser state is a joint eigenstate of a set of commuting Weyl operators. Two Weyl operators commute if and only if the corresponding phase-space points \mathbf{a}, \mathbf{a}' have vanishing symplectic inner product:

$$[\hat{W}(\mathbf{a}), \hat{W}(\mathbf{a}')] = 0 \text{ if and only if } \langle \mathbf{a}, \mathbf{a}' \rangle = \mathbf{a}^T J \mathbf{a}' = 0. \tag{47}$$

This result derives from the product rule of Weyl operators:

$$\hat{W}(\mathbf{a}) \hat{W}(\mathbf{a}') = \chi(\langle \mathbf{a}, \mathbf{a}' \rangle) \hat{W}(\mathbf{a} + \mathbf{a}').$$

From this result, the sets of commuting Weyl operators, and, as a consequence, the stabiliser states, are parametrised by the isotropic subspace M of Ω . More precisely, for each M and each $\mathbf{w} \in \Omega$ we can define a stabiliser state (Gross construction) $\rho_{M,\mathbf{w}}$ as the projector onto the joint eigenspace spanned by $\{\hat{W}(\mathbf{a}) : \mathbf{a} \in M\}$, where $\hat{W}(\mathbf{a})$ has eigenvalue $\chi(\langle \mathbf{w}, \mathbf{a}' \rangle)$. The Wigner function associated to the state $\rho_{M,\mathbf{w}}$ is always positive (necessary and sufficient condition in odd dimensions) and it is of the kind

$$W_{(m,\mathbf{w})}(\lambda) = \frac{1}{d^n} \delta_{M^c + \mathbf{w}}(\lambda), \tag{48}$$

where M^c is the symplectic complement of M . Moreover the transformations that preserve the positivity of the Wigner functions are the Clifford unitaries. GT of non-negative Wigner functions is a faithful way of representing SQM.

Equivalence of ST and GT. The Wigner function (48) has the same form of the probability distribution (7) associated to the epistemic state (V, \mathbf{w}) in ST. More precisely, they are equivalent if we assume $M = JV^5$, indeed

⁵ Note that the action of J is simply to map a variable into its conjugated.

this transformation implies that $V^\perp = M^C$. The equivalence between GT and Spekkens theory, using the symplectic matrix J as the bridge, also extends in terms of transformations and measurement statistics [5]. This equivalence also implies the equivalence between ST and SQM in odd dimensions. Therefore we can see the description based on known variables (Spekkens) and the description based on Wigner functions (Gross) as two equivalent descriptions of SQM in odd dimensions. We will now translate the already found updating rules of ST into Gross' Wigner functions.

Updating rules. Let us consider a stabiliser state $\rho = \rho_1 \cdot \rho_2 \cdot \dots \cdot \rho_n$, where n is the number of qudits (odd prime dimensions), and a measurement Π on the stabiliser state $\Pi = \Pi_1 \cdot \Pi_2 \cdot \dots \cdot \Pi_m$, where, in general, $m \leq n$. Let us assume $m = n$ in order to consider 'total' measurements (not only to a part of the state).

Theorem 4. (Commuting case). *Let us assume the state and measurement to commute, i.e. $[\rho, \Pi] = 0$. The Wigner function of the state after measurement is*

$$W_{\rho'}(\lambda) = \frac{1}{N} W_\rho(\lambda) R_\Pi(\lambda), \tag{49}$$

where $\lambda \in \Omega$ and R_Π denotes the Wigner function (also called response function) associated with the measurement Π . The normalisation factor N is

$$N = \sum_{\lambda \in \Omega} W_\rho(\lambda) R_\Pi(\lambda).$$

Proof. We rewrite the formula (49) by replacing the Wigner functions with their definition in terms of Spekkens' subspaces,

$$\delta_{\lambda, V'^\perp + \mathbf{w}'} = \delta_{\lambda, V^\perp + \mathbf{w}} \cdot \delta_{\lambda, V_\Pi^\perp + \mathbf{r}}. \tag{50}$$

The proof is straightforward. The rhs is one if and only if both the deltas are one; this means that λ has to belong simultaneously to $V^\perp + \mathbf{w}$ and $V_\Pi^\perp + \mathbf{r}$, i.e. $\lambda \in (V^\perp + \mathbf{w}) \cap (V_\Pi^\perp + \mathbf{r})$. If we recall equation (13) (and figure 4), we see that

$$(V^\perp + \mathbf{w}) \cap (V_\Pi^\perp + \mathbf{r}) = (V'^\perp) + \mathbf{w}'$$

and we can conclude that the rhs of equation (50) is one if and only if the lhs is one. At this point we can insert the normalisation factors on the rhs and the lhs. These guarantee that $\sum_{\lambda \in \Omega} W_{\rho'}(\lambda) = 1$ and the uniformity as expected. □

In the commuting case the updating rule in SQM consists of the *addition* of the stabiliser generators of state and measurement (equation (43)). In ST the updating rule consists of the *intersection* of the perpendicular isotropic subspaces (equation (13)). In GT addition and intersection translate into the *product* of the Wigner functions (equation (49)). In particular this stage consists of introducing zeros to the Wigner function in correspondence of the addition of generators to the subspace of known variables V (and so removing generators from the subspace V^\perp). We will call this process—where we *learn* information about the state—the *localisation stage*.

Theorem 5. (Non-commuting case). *Let us assume the measurement, in general, not to commute with the state, i.e. $[\rho, \Pi] \neq 0$. The Wigner function of the state after measurement is*

$$W_{\rho'}(\lambda) = \frac{1}{N} \sum_{\mathbf{t} \in V_{\text{other}}} W_\rho(\lambda - \mathbf{t}) R_\Pi(\lambda), \tag{51}$$

where $\lambda \in \Omega$, V_{other} is the set spanned by the non-commuting generators of Spekkens' subspace V associated to the state ρ . The normalisation factor N is

$$N = \sum_{\lambda \in \Omega} \sum_{\mathbf{t} \in V_{\text{other}}} W_\rho(\lambda - \mathbf{t}) R_\Pi(\lambda).$$

Note that we could have stated the theorem in terms of stabiliser generators instead of Spekkens' generators. The former being related to the latter as follows,

$$g_j = \hat{W}(J^{-1}\Sigma_j), \tag{52}$$

where J is the usual symplectic matrix, Σ_j are Spekkens' generators and g_j the corresponding stabiliser generators. The relation (52) follows from the relation between ST and GT previously described, where the bridge between the two formulations is given by the matrix J .

Proof. In general the state after measurement in quantum mechanics (up to a normalisation) is $\rho' = \Pi\rho\Pi$. If $[\rho, \Pi] = 0$ then $\rho' = \rho\Pi$.

In order to simplify the proof, let us assume the case of only one non-commuting generator, say ρ_n . In the present case we know, from the structure of SQM and Spekkens' updating rules (adding the commuting factors between state and measurement and removing the non-commuting ones), that the state after measurement is $\rho' = \rho^* \Pi$, where $\rho^* = \rho_1 \cdots \rho_{n-1}$. This means that we can write the state after measurement as a product of two commuting terms: ρ^* and Π . Therefore we can write the Wigner function of ρ' according to the product rule for the commuting case (equation (49)):

$$W_{\rho'}(\lambda) = \frac{1}{N} W_{\rho^*}(\lambda) R_{\Pi}(\lambda),$$

where $N = \sum_{\lambda} W_{\rho^*}(\lambda) R_{\Pi}(\lambda)$. We want now to prove that equation (51) is equal to the latter. This means we want to prove the following:

$$W_{\rho'}(\lambda) = \sum_{\mathbf{t} \in V_{\text{other}}} W_{\rho}(\lambda - \mathbf{t}) R_{\Pi}(\lambda) = W_{\rho^*}(\lambda) R_{\Pi}(\lambda).$$

We can simplify the terms $R_{\Pi}(\lambda)$, thus getting

$$\sum_{\mathbf{t} \in V_{\text{other}}} W_{\rho}(\lambda - \mathbf{t}) = W_{\rho^*}(\lambda). \tag{53}$$

At this point, in order to prove the above theorem, we rewrite the formula (51) by replacing the Wigner functions with their definition, i.e. Kronecker deltas,

$$\sum_{\mathbf{t} \in V_{\text{other}}} \delta_{\lambda - \mathbf{t}, V^{\perp} + \mathbf{w}} = \delta_{\lambda, V_{\text{commute}}^{\perp} + \mathbf{w}}, \tag{54}$$

where $V_{\text{commute}}^{\perp} = V^{\perp} \oplus V_{\text{other}}$. Note that we have removed the response function of the measurement. This also implies that we do not have to change \mathbf{w} , because we have only modified V^{\perp} into $V_{\text{commute}}^{\perp}$ and \mathbf{w}' is not affected. We now want to see that the lhs of equation (54) is different from zero exactly when the rhs is. The lhs is different from zero when at least one $\mathbf{t} \in V_{\text{other}}$ is such that $\lambda - \mathbf{t} \in V^{\perp} + \mathbf{w}$. The latter corresponds to $\lambda \in V^{\perp} + \mathbf{w} + \mathbf{t}$. This means that $\lambda \in V^{\perp} \oplus V_{\text{other}} + \mathbf{w}$, i.e. $\lambda \in V_{\text{commute}}^{\perp} + \mathbf{w}$, which is precisely what makes the rhs different from zero. \square

In the most general non-commuting case, in addition to the localisation stage, in SQM we also have to *remove* the non-commuting generators from the state (equation (44)). In ST this consists of the *union and shifts* in the perpendicular subspace (equation (22)). In GT removal and union translate into the *averaging out* of the Wigner function (equation (51)). In particular this stage consists of introducing ones to the Wigner function in correspondence of the removal of generators from the subspace of known variables V (and so adding generators to the subspace V^{\perp}). We can think of this process as the one where, after having learned some information in the localisation stage, we need to forget something, otherwise we would get too much information about the ontic state, which is forbidden by the classical complementarity principle. This also explains why non-commuting measurements are also called *disturbing* measurements. We will call this forgetting-part of the process the *randomisation stage*. Finally note that the general-case formula (51) reduce to the product rule (49) in the commuting case. Figure 9 summarises the updating rules in the three theories in prime dimensions.

In the non-prime dimensional case, we can rephrase all the reasonings already done in ST in terms of Wigner functions.

Lemma 4. *The Wigner function $W_{\text{cg}}(\lambda)$ of the coarse-graining observable*

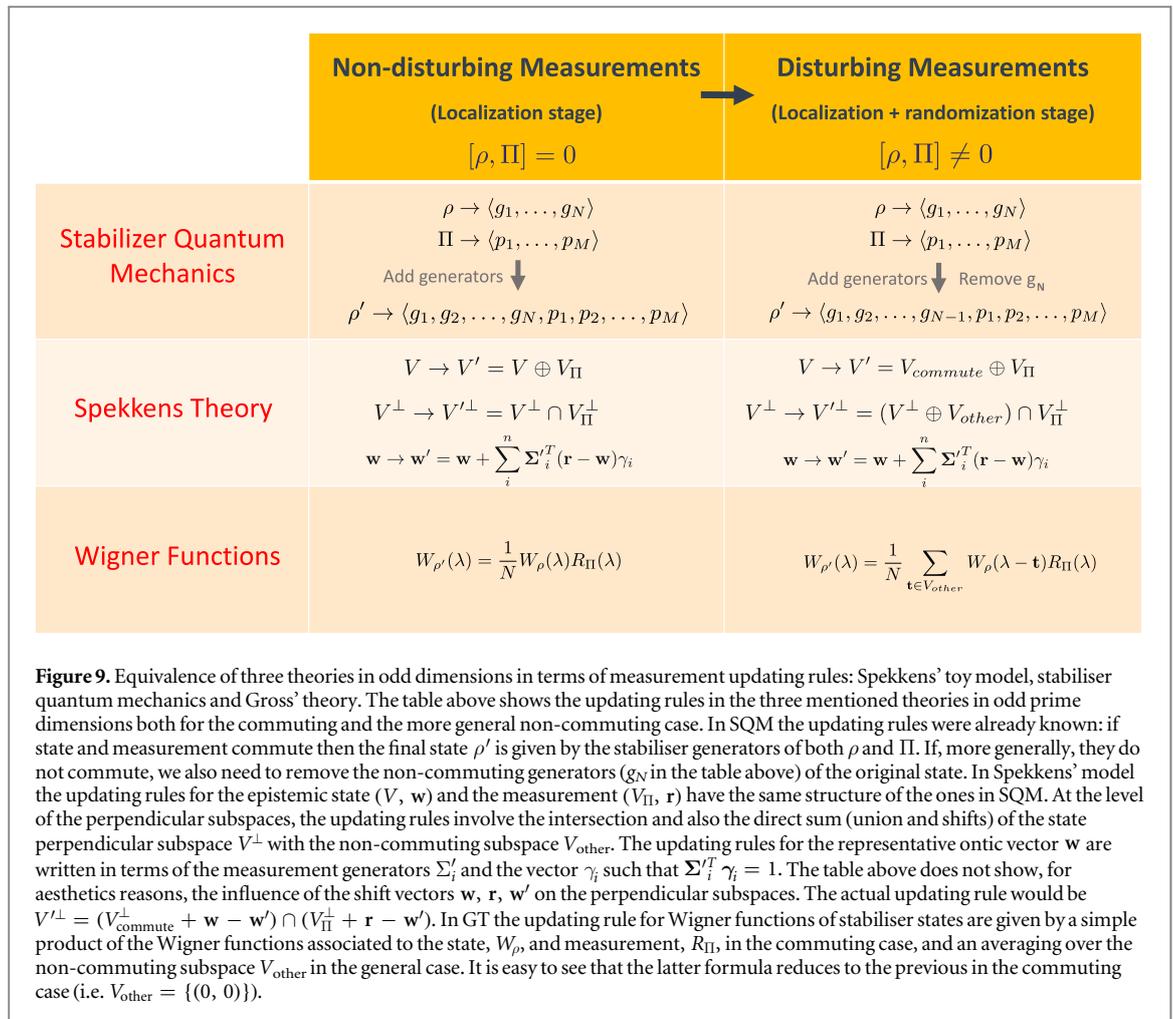
$O_{\text{cg}} = a_1 X_1 + b_1 P_1 + \dots + a_n X_n + b_n P_n = D(a'_1 X_1 + b'_1 P_1 + \dots + a'_n X_n + b'_n P_n) = \sigma_{\text{cg}}$, *can be written in terms of the Wigner functions $W_{\text{fg}}^{(j)}(\lambda)$ of the associated fine graining observables*

$$O_{\text{fg}}^{(j)} = a'_1 X_1 + b'_1 P_1 + \dots + a'_n X_n + b'_n P_n = \sigma_{\text{fg}}^{(j)} \text{ as}$$

$$W_{\text{cg}}(\lambda) = \frac{1}{\bar{D}} \sum_{j=0}^{\bar{D}-1} W_{\text{fg}}^{(j)}(\lambda). \tag{55}$$

Proof. First of all the normalisation factor $\frac{1}{\bar{D}}$ is due to the fact that we are adding \bar{D} Wigner functions, each of them having a normalisation factor of $\frac{1}{d}$, since they are Wigner functions of maximally isotropic subspaces (of dimension d). The proof of the rest of the formula is straightforward. According to the definition of Wigner functions, we need to prove that

$$\delta_{V_{\text{cg}}^{\perp} + \mathbf{r}_{\text{cg}}} \propto \sum_j \delta_{V_{\text{fg}}^{\perp} + \mathbf{r}_{\text{fg}}^{(j)}}. \tag{56}$$



From the decomposition of the isotropic subspaces and shift vectors in Spekkens’ model, equations (28) and (30), we already know that $V_{cg}^\perp + \mathbf{r}_{cg} = V_{fg}^\perp \oplus V_D + \mathbf{r}_{cg} = V_{fg}^\perp + \sum_{j=0}^{\bar{D}-1} (\mathbf{r}_{cg} + j\mathbf{v})$, which exactly proves that the rhs of (56) is one if and only if the lhs is one. □

From the above construction and theorem 3 we can immediately write the Wigner function of a stabiliser state after a coarse-graining measurement, thus generalising theorem 5.

Theorem 6. Given the state ρ of n -qudit systems, where the dimension d is a non-prime integer, and the (non-commuting) measurement Π , the Wigner function of the state ρ' after the measurement is given by

$$W_{\rho'}(\lambda) = \frac{1}{N} \frac{1}{\bar{D}} \sum_{\mathbf{t} \in V_{other}} \sum_{j=0}^{\bar{D}-1} W_\rho(\lambda - \mathbf{t}) R_{fg}^{(j)}(\lambda), \tag{57}$$

where $\lambda \in \Omega$, V_{other} is the set spanned by the non-commuting generators of Spekkens’ subspace V associated to the state ρ . The response function of the j th fine-graining measurement is denoted by $R_{fg}^{(j)}$. The normalisation factor N is

$$N = \sum_{\lambda \in \Omega} \sum_{\mathbf{t} \in V_{other}} W_\rho(\lambda - \mathbf{t}) R_\Pi(\lambda),$$

where $R_\Pi(\lambda) = \frac{1}{\bar{D}} \sum_{j=0}^{\bar{D}-1} R_{fg}^{(j)}(\lambda)$.

Proof. We just need to apply lemma 4 to the response function of the coarse graining measurement of theorem 5. □

Figure 10 summarises the updating rules in ST and GT in prime and non-prime dimensions.

	Prime dimensional systems	Non-prime dimensional systems
Spekkens Theory	$V'^{\perp} = (V_{commute}^{\perp} + \mathbf{w} - \mathbf{w}') \cap (V_{\Pi}^{\perp} + \mathbf{r} - \mathbf{w}')$ $\mathbf{w} \rightarrow \mathbf{w}' = \mathbf{w} + \sum_i^n \Sigma_i^T (\mathbf{r} - \mathbf{w}) \gamma_i$	$V'^{\perp} = \bigcup_{j=0}^{\bar{D}-1} [(V_{commute}^{\perp} + \mathbf{w} - \mathbf{w}') \cap (V_{fg}^{\perp} + \mathbf{r}_{fg}^{(j)} - \mathbf{w}')]$ $\mathbf{w}' = \mathbf{w}'_j = \mathbf{w} + \sum_{i=0}^n \Sigma_i^T (\mathbf{r}_{fg}^{(j)} - \mathbf{w}) \gamma_i$
Wigner Functions	$W_{\rho'}(\lambda) = \frac{1}{N} \sum_{\mathbf{t} \in V_{other}} W_{\rho}(\lambda - \mathbf{t}) R_{\Pi}(\lambda)$	$W_{\rho'}(\lambda) = \frac{1}{N} \frac{1}{\bar{D}} \sum_{\mathbf{t} \in V_{other}} \sum_{j=0}^{\bar{D}-1} W_{\rho}(\lambda - \mathbf{t}) R_{fg}^{(j)}(\lambda)$

Figure 10. Measurement updating rules in Spekkens’ toy model and Gross’ theory in prime and non-prime dimensions. The table above shows the updating rules (for the general non-commuting case) of ST and Gross’ theory in prime dimensions, first column, and non-prime dimensions, second column. The former have been already depicted in table 9. The latter regard the case of a coarse-graining measurement observable O_{cg} . In terms of perpendicular subspaces the updating rules consist of the union of the updating subspaces of the original state (V, \mathbf{w}) with each of the \bar{D} individual fine-graining observables $(V_{fg}, \mathbf{r}_{fg}^{(j)})$. The updated shift vector \mathbf{w}' is just one of the updated shift vectors \mathbf{w}'_j of the state with the fine-graining observables. In terms of Wigner functions, the union translates into a sum of \bar{D} terms, and the response functions of the fine-graining observables are denoted as $R_{fg}^{(j)}$.

6. Discussion

The importance of completing ST with updating rules to determine the state after a sharp measurement depends upon their application in future works. In particular we think that it would be interesting to explore how quantum computational schemes can be represented by ST. In order to do this it is appropriate to first characterise ST in terms of its computational power.

We can perform a simple analysis of the computational complexity of simulating ST on a classical computer following the same approach as Aaronson and Gottesman’s analysis of the simulation of stabiliser circuits [31]. In ST, each epistemic state is described by the isotropic subspace of known variables V , which is defined by n generators, and the shift vector \mathbf{w} , which attributes n values to the n variables. Each generator is specified by $2n$ components. Therefore we need $2n^2 + n$ digits to specify an epistemic state (V, \mathbf{w}) . To find the perpendicular subspace V^{\perp} , we need a further n^2 operations to check all the inner products between the generators.

Simulating dynamics requires computing symplectic affine transformations involving about $(2n)^2 + n \simeq n(n + 1)$ digits for each generator of the epistemic state that has $2n \simeq n$ components, since the product between a matrix and a vector involves $O(n)^2$ modular arithmetic operations and the affine translation $2n$ operations. Therefore the total is $n(n + 1) \cdot n \simeq n^3$.

It should be possible to find more efficient algorithm using some of the ideas in [31]. However we are not aiming to optimise this simulation complexity in the current work, just show that it is classically efficient. The updating rules for the measurements in the prime case (2) involve first adding the generators of the subspaces V and V_{Π} and then removing the non-commuting ones (this involves to check their symplectic inner product, which means about n^3 operations). In total we would have $n^3 + 2n = n(n^2 + 2) \simeq n^3$ operations for finding V' , the isotropic subspaces of known variables after the measurement. The updated shift vector involves the sum of two inner products between vectors of $2n$ components, which roughly means n^3 operations. In the non-prime case (3) further operations are needed, namely the ones to recover the degeneracy factor \bar{D} , which consist of dividing the $2n$ components of the generators by each of the d possible integer factors and then do the division again for surely less than d times, which means no more than $2n \cdot d^2$ operations. A final operation of checking whether the results of the divisions of the $2n$ components give the same value must be considered. It implies another factor of $2nd$. This allows us to compute the operations to perform the union of the perpendicular subspaces in (3), i.e. $\bar{D}n \simeq (nd)^3$ operations. This approximate analysis wants just to show that, even with basic simulation schemes, the computational complexity to perform a classical simulation of ST is polynomial in the number of systems. This is in line, as expected, with the computational power of SQM.

A possible application of ST in the above direction is to use it as a non-contextual hidden variable model to represent the classically simulatable part of some state-injection schemes of quantum computation. ST and its

subtheories which are operationally equivalent to subtheories of QM can play the role of witnesses of non-negativity of the Wigner functions and non-contextuality, and can therefore be used as a unifying framework for state injection schemes where negativity and contextuality are resources for universal quantum computation, similarly to [13–16, 18–20]. In particular, this would extend [13] by considering systems of compound dimensions.

The result about the equivalence between ST and SQM and the associated updating rules in prime and *non-prime* odd dimensions can provide a powerful new way to use and analyse SQM in non-prime dimensions, about which almost nothing is known. For example we are now facilitated to state, given a set of commuting Pauli operators, whether the joint eigenstate that they represent is pure. In non-prime dimensions the latter issue is not trivial because for coarse-graining observables the number of independent generators is not equal to the number of observables. However, from our construction to decompose coarse-graining into fine-graining observables, we know that the number of independent generators is equal to the number of fine-graining observables. Therefore if the set of commuting Pauli operators has the number of independent generators that equals the number of fine-graining observables, then the state is pure. Indeed fine-graining observables are associated to pure states. In addition, in the field of quantum error correction it could be interesting to study if the coarse-graining observables have any usefulness. The coarse-grained observables considered here are an example of degenerate observables. Degenerate observables, such as a parity measurement, play a central role in quantum error correction theory. The degeneracy means that errors can be detected without collapsing the logical state. It would be interesting to investigate whether the coarse-grained observables in compound dimension SQM have any utility for novel forms of quantum error correction.

Finally, the enforced equivalence of SQM, ST and GT in odd dimensions can be exploited to address a given problem from different perspectives, where, depending on the cases, one theory can be more appropriate than another. An example is the already mentioned one of addressing protocols based on SQM with Spekkens theory instead of SQM or Wigner functions.

7. Conclusion

Spekkens' toy model is a very powerful model which has led to meaningful insights in the field of quantum foundations and that seems to have interesting applications in the field of quantum computation. We have extended it from prime to arbitrary dimensional systems and we have derived measurement updating rules for systems of prime dimensions when the state and measurement commute, equations (13) and (14), when they do not, equations (18) and (14), and for systems of non-prime dimensions (theorem 3). These results directly derive from the basic axiom of the theory: the classical complementarity principle. The latter characterises a structure for the updating rules which is the same as in SQM: the state after measurement is composed by the generators of the measurement and the compatible (i.e. commuting) generators of the original state.

Spekkens showed the equivalence between SQM and ST in odd prime dimensions via Gross' Wigner functions. We have extended this result to all odd dimensions and we have translated the updating rules of ST in terms of Wigner functions (theorems 4–6). We stress again that Spekkens' model and our measurement updating rules hold in all dimensions, in even dimensions too. However the equivalence between ST and SQM only holds in odd dimensions. The main reason is that SQM in even dimensions shows contextuality, while ST does not. One of the main future challenges is to find an epistemic hidden variable toy model which is also equivalent to qubit SQM.

We treat the problem with systems of non-prime dimensions, which arises from the problem of defining an inverse in \mathbb{Z}_d , by decomposing the problematic (coarse-graining) observables in terms of the non-problematic (fine-graining) ones. This approach naturally suggests the form of the updating rules. By comparing the updating rules in the three mentioned theories we highlight the beauty and the elegance of this equivalence, where addition and removal of generators in SQM correspond to intersection and union in ST and product and randomisation in GT. This correspondence is schematically depicted, for the prime-dimensional case, in table 9. The non-prime case correspondence is represented in table 10. We believe that the fresh perspective gained by moving from one theory to another can give powerful new tools for new insights in the field of quantum computation.

Acknowledgments

We would like to thank Misja Steinmetz for suggesting Bezout's identity. This work was supported by EPSRC Centre for Doctoral Training in Delivering Quantum Technologies [EP/L015242/1].

References

- [1] Einstein A, Podolsky B and Rosen N 1935 Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47** 777–80
- [2] Bell J S 1966 On the problem of hidden variables in quantum mechanics *Rev. Mod. Phys.* **38** 447–52
- [3] Kochen S and Specker E P 1967 The problem of hidden variables in quantum mechanics *J. Math. Mech.* **17** 59–87
- [4] Spekkens R W 2007 Evidence for the epistemic view of quantum states: a toy theory *Phys. Rev. A* **75** 032110
- [5] Spekkens R W 2016 Quasi-quantization: classical statistical theories with an epistemic restriction *Fund. Theor. Phys.* **181** 83–135
- [6] Gottesman D 1997 Stabilizer codes and quantum error correction *PhD Thesis* California Institute of Technology
- [7] Gottesman D 1999 The Heisenberg representation of quantum computers *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics* ed S P Corney, R Delbourgo and P D Jarvis (Cambridge, MA: International Press) 32–43
- [8] Greenberger D, Horne M, Shimony A and Zeilinger A 1990 Bell's theorem without inequalities *Am. J. Phys.* **58** 1131
- [9] Mermin N D 1990 Simple unified form for the major no-hidden-variables theorems *Phys. Rev. Lett.* **65** 3373–6
- [10] Peres A 1990 Incompatible results of quantum measurements *Phys. Lett. A* **151** 107–8
- [11] Bravyi S and Kitaev A 2005 Universal quantum computation with ideal Clifford gates and noisy ancillas *Phys. Rev. A* **71** 022316
- [12] Gottesman D Stabilizer codes with prime power qudits *Invited Talk at Caltech IQIM Seminar (Pasadena, California)*
- [13] Howard M, Wallman J, Veitch V and Emerson J 2014 Contextuality supplies the magic for quantum computation *Nature* **510** 351–5
- [14] Delfosse N, Guerin P A, Bian J and Raussendorf R 2015 Wigner function negativity and contextuality in quantum computation on rebits *Phys. Rev. X* **5** 2160–3308
- [15] Raussendorf R, Browne D, Delfosse N, Okay C and Bermejo-Vega J 2017 Contextuality as a resource for qubit quantum computation *Phys. Rev. A* **95** 052334
- [16] Bermejo-Vega J, Delfosse N, Browne D E, Okay C and Raussendorf R 2016 Contextuality as a resource for qubit quantum computation [arXiv:1610.08529](https://arxiv.org/abs/1610.08529)
- [17] Raussendorf R 2013 Contextuality in measurement-based quantum computation *Phys. Rev. A* **88** 022322
- [18] Veitch V, Mousavian S A H, Gottesman D and Emerson J 2014 The resource theory of stabilizer quantum computation *New J. Phys.* **16** 013009
- [19] Veitch V, Wiebe N, Ferrie C and Emerson J 2013 Efficient simulation scheme for a class of quantum optics experiments with non-negative Wigner representation *New J. Phys.* **15** 013037
- [20] Ferrie C, Morris R and Emerson J 2010 Necessity of negativity in quantum theory *Phys. Rev. A* **82** 044103
- [21] Veitch V, Ferrie C, Gross D and Emerson J 2012 Negative quasi-probability as a resource for quantum computation *New J. Phys.* **14** 113011
- [22] Wallman J J and Bartlett S D 2012 Non-negative subtheories and quasiprobability representations of qubits *Phys. Rev. A* **85** 062121
- [23] Gross D 2006 Hudson's theorem for finite-dimensional quantum systems *J. Math. Phys.* **47** 122107
- [24] Wootters W K 1987 A Wigner-function formulation of finite-state quantum mechanics *Ann. Phys.* **176** 1
- [25] Pusey M F 2012 Stabilizer notation for Spekkens' toy theory *Found. Phys.* **42** 688
- [26] Coecke B, Edwards B and Spekkens R W 2011 Phase groups and the origin of non-locality for qubits, electron *Notes Theor. Comput.* **270** 29
- [27] Disilvestro L and Markham D 2017 Quantum protocols within Spekkens' toy model *Phys. Rev. A* **95** 052324
- [28] Blasiak P 2013 Quantum cube: a toy model of a qubit *Phys. Lett. A* **377** 847–50
- [29] Larsson J 2012 A contextual extension of Spekkens' toy model *AIP Conf. Proc.* **1424** 211
- [30] Bezout E 1779 *Theorie Generale des Equations Algebriques* (Paris, France: Ph.D. Pierres)
- [31] Aaronson S and Gottesman D 2004 Improved simulation of stabilizer circuits *Phys. Rev. A* **70** 052328