Journal of Physics A: Mathematical and Theoretical

PAPER

The non-disjoint ontic states of the Grassmann ontological model, transformation contextuality, and the single qubit stabilizer subtheory

To cite this article: Lucas Kocia and Peter Love 2019 *J. Phys. A: Math. Theor.* **52** 095303

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **52** (2019) 095303 (22pp) https://doi.org/10.1088/1751-8121/aafca7

The non-disjoint ontic states of the Grassmann ontological model, transformation contextuality, and the single qubit stabilizer subtheory

Lucas Kocia[1] **and Peter Love**[2]

[1] National Institute of Standards and Technology, Gaithersburg, MD 20899, United States of America
[2] Department of Physics and Astronomy, Tufts University, Medford, MA 02155, United States of America

E-mail: lucas.kocia@nist.gov

Received 31 May 2018, revised 5 November 2018
Accepted for publication 20 December 2018
Published 4 February 2019

Abstract

We show that it is possible to construct a preparation non-contextual ontological model that does not exhibit 'transformation contextuality' for single qubits in the stabilizer subtheory. In particular, we consider the 'blowtorch' map and show that it does not exhibit transformation contextuality under the Grassmann Wigner–Weyl–Moyal (WWM) qubit formalism. Furthermore, the transformation in this formalism can be fully expressed at order $\hbar^0$, where it satisfies all of Kolmogorov's axioms of classical probability theory, and so does not qualify as a candidate quantum phenomenon. In particular, we find that the Grassmann WWM formalism at order $\hbar^0$ corresponds to an ontological model governed by an additional set of constraints arising from the relations defining the Grassmann algebra. Due to this additional set of constraints, the allowed probability distributions in this model do not form a single convex set when expressed in terms of disjoint ontic states and so cannot be mapped to models whose states form a single convex set over disjoint ontic states. However, expressing the Grassmann WWM ontological model in terms of non-disjoint ontic states corresponding to the monomials of the Grassmann algebra results in a single convex set. We further show that a recent result by Lillystone *et al* that proves a broad class of preparation and measurement non-contextual ontological models must exhibit transformation contextuality lacks the generality to include the ontological model considered here; Lillystone *et al*'s result is appropriately limited to ontological models whose states produce a single convex set when expressed in terms of disjoint ontic states. Therefore, we prove that for the qubit stabilizer subtheory to be

captured by a preparation, transformation and measurement non-contextual ontological theory, it must be expressed in terms of non-disjoint ontic states, unlike the case for the odd-dimensional single-qudit stabilizer subtheory.

Keywords: Wigner–Weyl–Moyal, ontological, non-disjoint, Grassmann, contextuality

(Some figures may appear in colour only in the online journal)

1. Introduction

There has been much interest recently in the study of contextuality by those pursuing the classical simulation of near-term quantum computation. This is because of its central role in the extension of many efficiently simulatable systems to quantum universality. Contextuality has been shown to be the salient ingredient introduced in the magic state injection of Clifford circuits [1, 2], measurement-based quantum computation [3, 4], and the T gate extension of the Clifford gateset [5–9].

Contextuality can be present in the operational forms of preparation contextuality, transformation contextuality and measurement contextuality [10]. Measurement contextuality is perhaps the oldest and best-known form of contextuality, and is the inability to pre-assign outcomes to a set of observables without prior knowledge of the 'context' that they will be taken in [10–13]. In general, contextuality is believed to be a non-classical property of quantum mechanics and has been shown to require higher than order $\hbar^0$ terms in the Wigner–Weyl–Moyal (WWM) representation of the observables [8, 9, 14]. It is most frequently described in the ontological models formalism, wherein measurement contextuality is responsible for multiple possible outcomes in an ontological model (defined in the next section) where a single outcome is expected [10].

One of the simplest quantum subtheories is the single-qubit stabilizer subtheory, which has long been thought to be completely non-contextual [9, 15, 16]. However, recently, Lillystone *et al* proved that for a single qubit, a broad class of ontological models that are preparation and measurement non-contextual still exhibit transformation contextuality under the 'blowtorch' map [17]. Such a result contradicts the association of the presence of contextuality with the presence of non-classical properties.

In this paper we relate the Grassmann WWM formalism at order $\hbar^0$ to an ontological model—a preparation and measurement non-contextual ψ-epistemic ontological model for a single qubit—and perform Lillystone *et al*'s calculations. We find that the Grassmann WWM ontological model does not exhibit transformation contextuality under the 'blowtorch' map or any other map consisting of convex combination of one-qubit stabilizer states.

This suggests that there must be some aspect of the Grassmann WWM ontological model that is neither captured by Lillystone *et al*'s proof nor by many prior ontological models studied in the literature. We find that at order $\hbar^0$ the ontological model described by the Grassmann WWM formalism differs from these others in that it must be defined over non-disjoint ontic states to form a single convex set of probability distributions associated with the single-qubit stabilizer subtheory. We find that such an ontological model possesses unique properties that are not captured by restricting study to ontological models defined only over disjoint ontic states, as in past studies [10, 15, 18]. The ontological model corresponding to the Grassmann WWM formalism appears to be an example of a novel subclass of ontological models that seem to have been overlooked in the literature.

We show that despite its novelty, the Grassmann WWM ontological model still corresponds to a fully valid classical probability theory as it satisfies all of the (Kolmogorov's) axioms that define a classical probability theory. Therefore, the introduction of such an ontological model with ontic states based on Grassmann variables does not 'hide' contextuality or quantumness within its framework, as such an effort would have to, by definition, violate at least one of Kolmogorov's axioms.

We begin with a review of the results of Lillystone *et al* [17] in section 2 where we also introduce transformation contextuality in ontological models with disjoint ontic states. In section 3 we introduce the Grassmann WWM formalism and demonstrate that it does not exhibit transformation contextuality at order $\hbar^0$. In section 4 we introduce a simple ontological model over non-disjoint states and show how re-expressing it over disjoint ontic states produces more than one convex subset. This motivates why such ontological models cannot be represented by models with disjoint states. We then demonstrate in section 5 that the Grassmann WWM formalism is such an ontological model with non-disjoint states and establish more of its properties in section 6. We prove that it is inequivalent to Lillystone *et al*'s representative disjoint eight-state model in section 7. We conclude in section 8.

2. Review

We define ontological models according to [19]: an ontological model is defined by a measurable space Λ of possible physical states, with an associated σ-algebra Σ, and sets of measures or measurable functions $P_A : \Sigma \to [0, 1]$ are used to represent preparations, transformations and measurements in the ontological model. Λ is called the ontic space and elements $\lambda \in \Lambda$ are called ontic states.

An ontological model is a classical probability theory and so must satisfy Kolmogorov's three axioms:

1. non-negativity: $P(\lambda) \in \mathbb{R}$ and $P(\lambda) \geqslant 0 \, \forall \, \lambda \in \Lambda$,
2. $P(\Lambda) = 1$,
3. σ-additivitiy: $P(\cup_i \lambda_i) = \sum_i P(\lambda_i)$ if $\{\lambda_i\}$ are disjoint (i.e. correspond to mutually exclusive events).

From these axioms follow [20]: for any two subsets A, $B \in \Lambda$,

4. probability of an empty set: $P(\emptyset) = 0$,
5. the sum rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, and,
6. monotonicity: if $B \subset A$, then $P(A) \leqslant P(B)$ and $P(A \setminus B) = P(A) - P(B)$.

A and B are disjoint if $P(A \cap B) = \emptyset$ and non-disjoint otherwise. It should be noted that ontological models can be defined over both disjoint and non-disjoint ontic states and some past work has been careful to include both cases [19]. Non-disjoint ontological models are often treated as a 'coarse-graining' of a disjoint ontological model. We will show that in some cases, they must be treated in terms of non-disjoint states in order that their states form a single convex set.

Furthermore, we can distinguish between two different types of ontological models. From Harrigan *et al* [21]:

Definition 1. An ontological model is ψ-ontic if for any pair of preparation procedures, P_ψ and P_ϕ, associated with distinct quantum states Ψ and ϕ, we have $p(\lambda|P_\psi)\, p(\lambda|P_\phi) = 0$ for all λ.

Definition 2. If an ontological model fails to be ψ-ontic, then it is said to be ψ-epistemic.

ψ-ontic and ψ-epistemic ontological models are both also called 'hidden variable theories'. Colloquially, ψ-ontic ontological models can be thought of as hidden variable theories where the 'hidden' variables are not really hidden (because distinct wavefunctions correspond to distinct subsets of Λ) while ψ-epistemic models are models with truly hidden variables.

Lillystone *et al* introduce an eight-state ontological model for one qubit [17], originally developed in [15], which consists of an ontic space $\Lambda = \{\pm 1\}^3$ that can be indexed by $\lambda = (x, y, z) \in \Lambda_1$ for $x, y, z \in \pm 1$—the eigenvalues of the Pauli matrices $\hat{X}$, $\hat{Y}$ and $\hat{Z}$, respectively. This model is preparation and measurement non-contextual [15]. Ontic states evolve under the maps corresponding to $\hat{X}$, $\hat{Y}$ and $\hat{Z}$ as

$$\Gamma_X : (x, y, z) \to (x, -y, -z), \tag{1}$$

$$\Gamma_Y : (x, y, z) \to (-x, y, -z), \tag{2}$$

and

$$\Gamma_Z : (x, y, z) \to (-x, -y, z), \tag{3}$$

respectively. They evolve under the Hadamard gate H as

$$\Gamma_H : (x, y, z) \to (z, -y, x). \tag{4}$$

Since x, y, and z are each in $\{\pm 1\}$, these maps are not continuous; they are permutations on $\{\pm 1\}^3$ defined by equations (1)–(4).

Lillystone *et al* then consider evolution of an input state ρ under the two operationally equivalent implementations of the following map:

$$T_1(\rho) = \frac{1}{4}(\rho + X\rho X + Y\rho Y + Z\rho Z), \tag{5}$$

and

$$T_2(\rho) = HT_1(\rho)H. \tag{6}$$

$T_1(\rho) = T_2(\rho) = I/2$ and so this is often called the 'blowtorch' map since it is akin to 'taking a blowtorch' to the state ρ and heating it up to become the maximally mixed state (a Gibbs distribution at infinite temperature) [22]. Though $T_1(\rho) = T_2(\rho) = I/2$, the authors point out that under their eight-state model these two transformations are non-equivalent as they produce different outcomes and thus illustrate 'transformation contextuality'. Specifically,

$$\begin{aligned}(x, y, z) \overrightarrow{T_1} \frac{1}{4}\,&[(x, y, z) + (x, -y, -z) \\ &+(-x, y, -z) + (-x, -y, z)]\,.\end{aligned} \tag{7}$$

Thus T_1 maps ontic states with even (odd) sign parity to ontic states with even (odd) sign parity. On the other hand

$$\begin{aligned}(x, y, z) \overrightarrow{T_2} \frac{1}{4}\,&[(-x, -y, -z) + (-x, y, z) \\ &+(x, -y, z) + (x, y, -z)]\,.\end{aligned} \tag{8}$$

T_2 maps ontic states with even (odd) sign parity to ontic states with odd (even) sign parity. These two sets of four points are different and therefore the two maps can produce different probability distributions, as shown in figure 1.

As a result, this model produces different probability distributions over the ontic states depending on whether T_1 or T_2 is taken. However, since both maps result in the same result—the

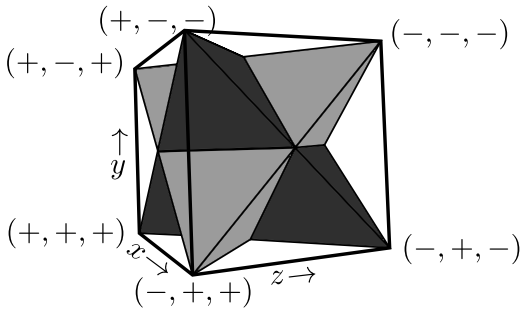

Figure 1. The eight-state ontic space from [17] showing the simplices of the even- (dark grey) and odd- (light grey) parity ontic states. T_1 maps ontic states in a tetrahedron to another ontic state in the same color tetrahedron. T_2 maps ontic states between the two tetrahedral regions even though it is operationally equivalent.

maximally mixed state—the resultant probability distribution should be the same. Thus, the eight-state model exhibits transformation contextuality. Lillystone *et al* then prove that every one-qubit non-preparation contextual ontological model can be mapped to the eight-state model and so all such models exhibit transformation contextuality. This includes both ψ-epistemic and ψ-ontic ontological models. We examine their proof carefully in section 7.

3. The blowtorch map in the Grassmann WWM formalism

The qubit Wigner–Weyl–Moyal (WWM) formalism was originally introduced by Berezin [23] and fully developed in [9]. The Grassmann model at $\mathcal{O}(\hbar^0)$ provides a classical Hamiltonian system that yields spin-$\frac{1}{2}$ under canonical quantization. It makes use of ξ_p, ξ_q and ξ_r, three real generators of a Grassmann algebra $\mathcal{G}_3$ which obey the anticommutation relation:

$$\xi_j\xi_k + \xi_k\xi_j \equiv \{\xi_j, \xi_k\} = 0, \quad \text{for } j, k \in \{p, q, r\}. \tag{9}$$

Any element $g \in \mathcal{G}_3$ may be represented as a finite sum of homogeneous monomials of the Grassmann elements and g is called a Weyl symbol.

In an effort to examine the T_1 and T_2 maps in this qubit WWM hidden variable theory, we consider the Weyl symbol of a single qubit pure state $\hat{\rho}$:

$$\rho = \frac{1}{2}\left(1 + \alpha \mathrm{i}\xi_r\xi_q + \beta \mathrm{i}\xi_p\xi_q + \gamma \mathrm{i}\xi_p\xi_r\right), \tag{10}$$

where $\alpha^2 + \beta^2 + \gamma^2 = 1$, for α, β, $\gamma \in \mathbb{R}$. The i's make the Weyl symbol ρ real, under a generalized conjugation operation [9].

Transformations $\hat{I}\hat{\rho}\hat{I}$, $\hat{X}\hat{\rho}\hat{X}$, $\hat{Y}\hat{\rho}\hat{Y}$, $\hat{Z}\hat{\rho}\hat{Z}$, and $\hat{H}\hat{\rho}\hat{H}$ are all Clifford transformations and so can be captured in the Wigner–Weyl–Moyal formalism at order $\hbar^0$ by solving the following classical equations of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi_k = \{H, \xi_k\}_{\mathrm{P.B}} = \mathrm{i}H\frac{\overleftarrow{\partial}}{\partial\xi_k} \tag{11}$$

where the right derivative $\frac{\overleftarrow{\partial}}{\partial\xi_k}$ is as defined in [9], $H_I = 1$, $H_X = -\mathrm{i}\xi_r\xi_q$, $H_Y = -\mathrm{i}\xi_p\xi_q$, $H_Z = -\mathrm{i}\xi_p\xi_r$, for $t = \pi/2$ and $H_{\hat{H}} = -\frac{\mathrm{i}}{\sqrt{2}}(\xi_r\xi_q + \xi_p\xi_r)$ for $t = \pi$. For $\hbar > 0$, these equations of motion are deformed to the Weyl algebra [9] for non-Clifford unitaries. They are then described by a Weyl bracket instead of a Poisson bracket and the Grassmann elements become the usual

Pauli matrices in quantum mechanics. However, since this is unnecessary for Clifford transformations, we will not need to explore this regime.

Clifford transformations take stabilizer states to stabilizer states. Solving the equations of motion for transformations I, X, Y, Z, and H, can be written in the same way as the (x, y, z) transformations in [17] by using three-tuples $(x, y, z) \in \Lambda_2$ for $x, y, z \in \{\pm\xi_p, \pm\xi_q \pm \xi_r\}$:

$$(\xi_p, \xi_q, \xi_r) \underset{I}{\to} (\xi_p, \xi_q, \xi_r), \tag{12}$$

$$(\xi_p, \xi_q, \xi_r) \underset{X}{\to} (\xi_p, -\xi_q, -\xi_r), \tag{13}$$

$$(\xi_p, \xi_q, \xi_r) \underset{Y}{\to} (-\xi_p, \xi_q, -\xi_r), \tag{14}$$

$$(\xi_p, \xi_q, \xi_r) \underset{Z}{\to} (-\xi_p, -\xi_q, \xi_r), \tag{15}$$

and

$$(\xi_p, \xi_q, \xi_r) \underset{H}{\to} (\xi_q, \xi_p, -\xi_r). \tag{16}$$

Substituting in the maps given by equations (12)–(16), we find that $\hat{I}\hat{\rho}\hat{I}$, $\hat{X}\hat{\rho}\hat{X}$, $\hat{Y}\hat{\rho}\hat{Y}$, $\hat{Z}\hat{\rho}\hat{Z}$, and $\hat{H}\hat{\rho}\hat{H}$ are

$$\rho_I = \frac{1}{2}\left(1 + \alpha \mathrm{i}\xi_r\xi_q + \beta \mathrm{i}\xi_p\xi_q + \gamma \mathrm{i}\xi_p\xi_r\right). \tag{17}$$

$$\rho_X = \frac{1}{2}\left(1 + \alpha \mathrm{i}\xi_r\xi_q - \beta \mathrm{i}\xi_p\xi_q - \gamma \mathrm{i}\xi_p\xi_r\right). \tag{18}$$

$$\rho_Y = \frac{1}{2}\left(1 - \alpha \mathrm{i}\xi_r\xi_q + \beta \mathrm{i}\xi_p\xi_q - \gamma \mathrm{i}\xi_p\xi_r\right), \tag{19}$$

$$\rho_Z = \frac{1}{2}\left(1 - \alpha \mathrm{i}\xi_r\xi_q - \beta \mathrm{i}\xi_p\xi_q + \gamma \mathrm{i}\xi_p\xi_r\right), \tag{20}$$

and

$$\rho_H = \frac{1}{2}\left(1 + \gamma \mathrm{i}\xi_r\xi_q - \beta \mathrm{i}\xi_p\xi_q + \alpha \mathrm{i}\xi_p\xi_r\right), \tag{21}$$

respectively.

Thus, we see that under the T_1 transformation,

$$\begin{aligned}
\rho \underset{T_1}{\to} \frac{1}{4}\Bigg[&\frac{1}{2}\left(1 + \alpha \mathrm{i}\xi_r\xi_q + \beta \mathrm{i}\xi_p\xi_q + \gamma \mathrm{i}\xi_p\xi_r\right) \\
&+ \frac{1}{2}\left(1 + \alpha \mathrm{i}\xi_r\xi_q - \beta \mathrm{i}\xi_p\xi_q - \gamma \mathrm{i}\xi_p\xi_r\right) \\
&+ \frac{1}{2}\left(1 - \alpha \mathrm{i}\xi_r\xi_q + \beta \mathrm{i}\xi_p\xi_q - \gamma \mathrm{i}\xi_p\xi_r\right) \\
&+ \frac{1}{2}\left(1 - \alpha \mathrm{i}\xi_r\xi_q - \beta \mathrm{i}\xi_p\xi_q + \gamma \mathrm{i}\xi_p\xi_r\right)\Bigg] \\
= \frac{1}{4}&(\rho_I + \rho_X + \rho_Y + \rho_Z) \\
= \frac{1}{2}&.
\end{aligned} \tag{22}$$

This is the Weyl symbol for $\hat{I}/2$. The simplification of the convex combination above is accomplished by the Weyl algebra of $\mathcal{G}_3$. Such a simplification is not possible under $\{\pm 1\}^3$, which lacks such algebraic operations.

On the other hand, acting on this evolution with the Hadamard gate to effect transformation T_2 produces:

$$\begin{aligned}\rho \underset{T_2}{\to} \frac{1}{4}\Bigg[&\frac{1}{2}\left(1+\gamma \mathrm{i}\xi_r\xi_q-\beta \mathrm{i}\xi_p\xi_q+\alpha \mathrm{i}\xi_p\xi_r\right)\\ &+\frac{1}{2}\left(1+\gamma \mathrm{i}\xi_r\xi_q+\beta \mathrm{i}\xi_p\xi_q-\alpha \mathrm{i}\xi_p\xi_r\right)\\ &+\frac{1}{2}\left(1-\gamma \mathrm{i}\xi_r\xi_q-\beta \mathrm{i}\xi_p\xi_q-\alpha \mathrm{i}\xi_p\xi_r\right)\\ &+\frac{1}{2}\left(1-\gamma \mathrm{i}\xi_r\xi_q+\beta \mathrm{i}\xi_p\xi_q+\alpha \mathrm{i}\xi_p\xi_r\right)\Bigg]\\ =&\frac{1}{4}\left(\rho_H+\rho_{HXH}+\rho_{HYH}+\rho_{HZH}\right)\\ =&\frac{1}{2}.\end{aligned} \tag{23}$$

Again, this is the Weyl symbol for $\hat{I}/2$. Both of these results are obtained without quantizing the Weyl symbols and so this result is possible all while working at order $\hbar^0$.

This result raises an interesting question when compared to the result obtained using Lillystone *et al*'s eight-state ontological model: since the WWM formalism is able to obtain the maximally mixed state at order $\hbar^0$ regardless of whether map T_1 or T_2 is taken, does this suggest that there exists an analogous classical probability theory (a preparation non-contextual ontological model) that similarly does not depend on whether transformation T_1 or T_2 is taken? If so, how can this be reconciled with Lillystone *et al*'s proof that every such ontological model can be mapped to their eight-state ontological model, which does exhibit dependence on whether T_1 or T_2 is taken?

We investigate these questions in the following sections by first defining a simple three-state ontological model example in section 4, which introduces the key element that the eight-state ontological model does not possess: non-disjoint ontic states. This then leads us to develop a larger ontological model equivalent to the Grassmann WWM formalism in sections 5 and 6.

4. Example of a simple ontological model with non-disjoint ontic states

The eight-state model is an example of an ontological model with disjoint ontic states. This means that for any two ontic states A and B, $P(A \cup B) = P(A) + P(B)$. By contradistinction, non-disjoint ontic states have non-zero overlaps ($A \cap B \neq \emptyset$) and so satisfy the classical relation: $P(A \cap B) = P(A) + P(B) - P(A \cup B)$. This can be derived directly from Kolmogorov's three axioms as we noted in section 2.

Here we introduce a simple example of a classical probability theory that is defined over only three ontic states, two of which are non-disjoint due to an additional set of relations that satisfies Kolmogorov's axioms. Within this simple model, we show how re-expressing the ontic states in terms of only disjoint states does not produce a convex set of probability distributions due to this additional set of constraints.

Consider a probability space with three elements, $\Lambda = \{A, B, C\}$. We wish to deal with a proper classical probability space, and so must satisfy all of Kolmogorov's axioms given in section 2.

We specify additional constraints on our probability space that we will show are compatible with these axioms:

$$P(C) = 1, \tag{24}$$

$$P(A) + P(B) = 1, \tag{25}$$

$$P(A \cup B) = \max\{P(A), P(B)\}, \tag{26}$$

and

$$P(A \cap B) = \min\{P(A), P(B)\}. \tag{27}$$

These additional constraints impose that our ontic states A and B are disjoint.

Axiom 1 is satisfied since $\Lambda = C$ and $P(C) = 1$ and axiom 2 can be imposed.

Axiom 3 is satisfied since A and B are not disjoint and so satisfy the sum rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{28}$$

Since $P(C) = 1$, all probability distributions only cover a part of our ontic space. We show three example probability distributions in figure 2 that satisfy the additional constraints given by equations (24)–(27).

It is of course perfectly acceptable to split up our ontic space into four 'finer' disjoint ontic states [15, 17, 18], which we label W, X, Y and Z as in figure 3.

However, while with the 'coarse-grained' non-disjoint states (A, B and C) the probability distributions form a single convex set, with the 'atomic' or 'finer' disjoint states (W, X, Y and Z), the additional set of constraints splits this convex set into more than one subset.

To see this, note that for the probability space labelled by the disjoint ontic states W, X, Y, and Z, incorporating the additional system of equations given by equations (24)–(27) produces:

$$P(W) + P(X) + P(Y) + P(Z) = 1, \tag{29}$$

$$P(W) + P(X) + 2P(Y) = 1, \tag{30}$$

$$P(W) + P(X) + P(Y) = \max\{P(W) + P(Y), P(X) + P(Y)\}, \tag{31}$$

$$P(Y) = \min\{P(W) + P(Y), P(X) + P(Y)\}, \tag{32}$$

respectively.

Allowed probability distributions are points in the three-simplex defined by equation (31) that also satisfy equations (29), (30) and (32). There are only two cases of solutions:

1. $P(W) = 0$,
2. $P(X) = 0$.

Let the tuple $(w, x, y, z) \in \Lambda_3$ refer to the probability on W, X, Y, and Z, respectively. In cases 1 and 2 we can choose $\alpha \equiv P(W)$ or $\alpha \equiv P(X)$ respectively, and define a one-parameter family of probability distributions (w, x, y, z):

$$1. \quad (\alpha, 0, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\alpha)), \tag{33}$$

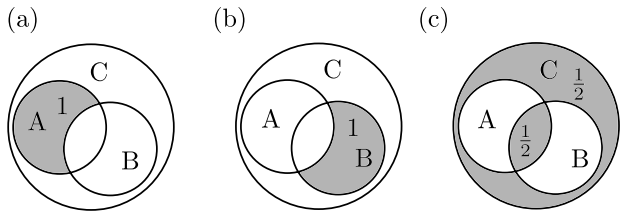

Figure 2. Probability distributions where $P(C) = 1$ and (a) $P(A) = 1$, $P(B) = 0$, (b) $P(A) = 0$, $P(B) = 1$ and (c) $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$ that are supported on the the non-disjoint states A and B and so satisfy equations (24)–(27).

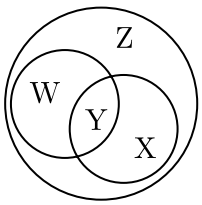

Figure 3. Same ontic space as in figure 5 but now labelled by a 'finer' set of disjoint ontic states W, X, Y and Z.

$$2. \quad (0, \alpha, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\alpha)), \tag{34}$$

respectively, for $0 \leqslant \alpha \leqslant 1$. $\alpha = 0$ corresponds to the only probability distribution that lies in both cases and is the one indicated by figure 2(c). These two cases correspond to two convex subsets of the original single convex set.

Convex combination of probability distributions from these two convex sets do *not* satisfy the constraints given by equations (29)–(32) (unless $\alpha = 0$). For instance, consider the convex combination of the probability distributions in figures 2(a) and (b) corresponding to the following tuples in Λ_3:

$$\frac{1}{2}\Big(1,0,0,0\Big) + \frac{1}{2}\Big(0,1,0,0\Big). \tag{35}$$

For the probability space labelled by disjoint ontic states W, X, Y, and Z, since the two terms in the convex combination given by equation (35) correspond to two different cases (1 and 2) with $\alpha \neq 0$, it follows that their result cannot satisfy equations (29)–(32). The only way to obtain a convex combination is to convert the disjoint ontic states W, X, Y, and Z, back into the non-disjoint ontic states A, B and C, perform the convex combination that satisfies the old equations (24)–(27), and then convert back to the disjoint ontic states.

Doing so, we can find that the convex combination given by equation (35), when converted to be in terms of non-disjoint states A and B, produces $P(A = W \cup Y) = P(B = X \cup Y) = \frac{1}{2}$. From equations (26) and (27), this means that $P(A \cap B) = P(A \cup B) = \frac{1}{2}$. Moreover, by equation (24), the resultant probability distribution must have zero support on $A \setminus B$ and $B \setminus A$. Converting back to the disjoint states W, X, Y, Z, this means that $P(W = A \setminus B) = P(X = B \setminus A) = 0$ and

$P(Y = A \cap B) = P(Z = A \cup B) = \frac{1}{2}$. This is represented by the tuple $(0, 0, \frac{1}{2}, \frac{1}{2}) \in \Lambda_3$, which is the probability distribution in figure 2(c).

In other words, given the information that there is a probability $\frac{1}{2}$ of being found in A and a probability $\frac{1}{2}$ of being found in B, this model enforces that A and B are non-disjoint and so produces the physically intuitive result that the probability of being found in A or B is $\frac{1}{2}$ (and so the probability of being found in neither A or B is $\frac{1}{2}$ too). This is a very different outcome from the one obtained if A and B are assumed to be disjoint, which given the information that there is a probability $\frac{1}{2}$ of being found in A and a probability $\frac{1}{2}$ instead implies that the probability of being found in A or B is 1.

Though the constraints given by equations (24)–(27) produce a single convex set with the 'coarse' set of non-disjoint states A, B, and C, the 'finer' disjoint set W, X, Y, and Z, cannot satisfy them with a single convex set.

Indeed, additional relations can only *non-trivially* supplement Kolmogorov's axioms if they produce two or more convex subsets when the ontic states are expressed disjointly (with no overlaps). This is because additional relations that satisfy σ-additivity for all ontic states (i.e. all ontic states are disjoint) add nothing new to the probability theory *unless* they produce more than one convex subset. However, for the theory to still describe the subtheory of interest, i.e. for the additional relations not to be too constraining, there must exist some other set of (non-disjoint) ontic states with respect to which all the probability distributions fall into the same convex set. This example demonstrates that such a middle ground between 'unconstrained' ontological models, which produce one convex set regardless of which set of disjoint or non-disjoint ontic states they are expressed with, and 'overconstrained' ontological models, which produce more than one convex set regardless of which set of ontic states they are expressed with, exists. This middle ground consists of constrained ontological models, which produce one convex set with respect to a particular set of non-disjoint ontic states and more than one for all other sets. This possibility appears to have been overlooked in the literature.

5. Grassmann WWM as an ontological model

In our prior work [9] we showed that it is possible to construct a local hidden variable theory (an ontological model) from the Grassmann WWM formalism to describe qubit stabilizer propagation using a non-negative probability distribution defined over states corresponding to the Grassmann monomials $\xi_j\xi_k$. We now re-present these results with respect to the nomenclature used to examine the simple ontological model in section 4.

A measure on the $\mathcal{G}_3$ algebra can be defined for any state $\rho = |\psi\rangle\langle\psi|$,

$$\mu_\rho(A_i) = \int \rho(\xi)\tilde{A}(\xi)\mathrm{d}^3\boldsymbol{\xi}, \tag{36}$$

where $\tilde{A}(\xi)$ is the dual (odd) Weyl symbol of $A(\xi)$ [9]. When $A(\xi) \equiv A_i(\xi)$ is the Weyl symbol of an element of a positive-operator valued measure (POVM) $\hat{A}_i$, then $\mu_\rho(A_i)$ is a non-negative measure (a probability distribution) over outcomes A_i of state ρ:

$$P_{A_i}(\rho) \equiv \mu_\rho(A_i). \tag{37}$$

However, we cannot rely on the measure $\mu_\rho(A)$ as a probability measure over ontic states $\xi_i\xi_j$ since it can be negative for $A \in \{\xi_j\xi_k\}$ as they are not elements of POVMs. Nevertheless, we can define a one-to-one map between $\mu_\rho(\xi_j\xi_k)$ and a *bone fide* probability measure that also preserves convex combination if we consider the Weyl algebra that the $\xi_j\xi_k$ satisfy.

For the ontic state $\xi_j\xi_k$,

$$-\frac{1}{2} \leqslant \mu_\rho(\xi_j\xi_k) = \int \xi_j\xi_k\tilde{\rho}(\xi)\mathrm{d}^3\boldsymbol{\xi} = \frac{1}{2}\sum_l \epsilon_{jkl}(\alpha_l + \beta_l + \gamma_l) \leqslant \frac{1}{2}, \tag{38}$$

for all one-qubit states ρ.

We note that

$$\mu_\rho(\xi_j\xi_k) = \mu_\rho(-\xi_k\xi_j) = -\mu_\rho(\xi_k\xi_j), \tag{39}$$

and so $\mu_\rho(\xi_k\xi_j) < 0$ is the same statement as $\mu_\rho(\xi_j\xi_k) > 0$. We choose to interpret $\mu_\rho(\xi_j\xi_k) > 0$ as proportional to the non-negative measure of ontic state $\xi_j\xi_i \setminus \xi_i\xi_j$ (and vice versa).

Given a probability $P(\xi_j\xi_k \setminus \xi_k\xi_j) \propto \mu_\rho(\xi_j\xi_k) > 0$, we further choose the probability of the other ontic state, $P(\xi_k\xi_j \setminus \xi_j\xi_k)$, to be zero under the heuristic motivation that μ_ρ does not need to track something if it is zero. Thus, given an ontic state $\lambda \in \{\xi_j\xi_k\}$, we define the non-negative probability of stabilizer state ρ in ontic state $\lambda \setminus -\lambda$ to be

$$P_\rho(\lambda \setminus -\lambda) \equiv \max\{2\mu_\rho(\lambda), 0\}, \tag{40}$$

where the factor of 2 allows the probability to saturate an upper bound of 1. We note that this is perhaps an arbitrary definition, we shall see that it is an acceptable one as it produces a theory consistent with Kolmogorov's axioms once unions and intersections are included, and reproduces the Grassmann WWM formalism for the stabilizer subtheory.

Any single qubit state's Weyl symbol ρ is represented by a linear combination of Grassmann monomials as in equation (10). Thus, our choice of definition for P_ρ equates the 'addition' operator in the Weyl algebra to a 'convex addition' operator since it treats any linear combination involving negative coefficients in front of Grassmann monomials as a unique non-negative convex combination, making use of the Grassmann anticommutation relations.

A stabilizer state has the Weyl symbol

$$\rho_{jk} \equiv \frac{1}{2}(1 + \mathrm{i}\xi_j\xi_k). \tag{41}$$

We now consider a convex combination of the two distinct stabilizer states ρ_{jk} and ρ_{kj} under the Weyl algebra:

$$\begin{aligned}\rho &= \alpha\rho_{jk} + \beta\rho_{kj} \\ &= \frac{1}{2} + (\alpha - \beta)\mathrm{i}\xi_j\xi_k,\end{aligned} \tag{42}$$

for $\alpha, \beta \geqslant 0$ such that $\alpha + \beta = 1$. Note that $(\alpha - \beta)\mathrm{i}\xi_j\xi_k = (\beta - \alpha)\mathrm{i}\xi_k\xi_j$.

WLOG, let us assume that $\alpha \geqslant \beta$. Equation (40) for $P_\rho(\lambda \setminus -\lambda)$ means that the probability of being in ontic state $\xi_j\xi_k \setminus \xi_k\xi_j$ after this convex combination is two times the coefficient in front of the resultant Weyl symbol's $\xi_j\xi_k$ term, $\alpha - \beta$, and the probability of being in ontic state $\xi_k\xi_j \setminus \xi_j\xi_k$ is 0. Before we simplified the convex combination, $P(\xi_j\xi_k) = \alpha$ and so

$$P_\rho(\xi_j\xi_k \cap \xi_k\xi_j) = P_\rho(\xi_j\xi_k) - P_\rho(\xi_j\xi_k \setminus \xi_k\xi_j) = \beta = \min\{\alpha, \beta\}. \tag{43}$$

This further agrees with

$$P_\rho(\xi_j\xi_k \cap \xi_k\xi_j) = P_\rho(\xi_k\xi_j) - P_\rho(\xi_k\xi_j \setminus \xi_j\xi_k) = \beta = \min\{\alpha, \beta\}, \tag{44}$$

since $P_\rho(\xi_k\xi_j) = \beta$. In other words, the convex combination takes a probability density of $\min\{\alpha, \beta\}$ from ontic state $\xi_j\xi_k \setminus \xi_k\xi_j$ to the intersection between the two ontic states. This means that

$$P(\xi_j\xi_k \cup \xi_k\xi_j) = P_\rho(\xi_j\xi_k \setminus \xi_k\xi_j) + P_\rho(\xi_k\xi_j \setminus \xi_j\xi_k) = \alpha = \max\{\alpha, \beta\}. \tag{45}$$

Therefore, for a map between μ and the probabilities to preserve μ's convex combinations under its Weyl algebra, it follows that

$$P(\xi_j\xi_k \cup \xi_k\xi_j) = \max\{P(\xi_j\xi_k), P(\xi_k\xi_j)\} \tag{46}$$

and

$$P(\xi_j\xi_k \cap \xi_k\xi_j) = \min\{P(\xi_j\xi_k), P(\xi_k\xi_j)\}. \tag{47}$$

As a result, we have the same probability space as that considered in the simple example of section 4, except that instead of one independent pair A and B, we have three independent pairs. Moreover, we accomplished this via a one-to-one mapping between our probabilities and our measure μ in $\mathcal{G}_3$ such that the set of probability distributions, when considered over the non-disjoint $\xi_j\xi_k$ ontic states, is a convex set. Most importantly, as we showed in the previous section, these additional constraints satisfy Kolmogorov's axioms and so form a valid classical probability theory or ontological model.

Using non-disjoint ontic states, we can set $A_1 = \xi_{pq}$, $B_1 = \xi_{qp}$, and then add two additional pairs: $\{A_2 = \xi_{pr}, B_2 = \xi_{rp}\}$ and $\{A_3 = \xi_{qr}, B_3 = \xi_{rq}\}$ so that:

$$P(A_i) + P(B_i) = 1, \tag{48}$$

$$P(A_i \cup B_i) = \max\{P(A_i), P(B_i)\}, \tag{49}$$

and

$$P(A_i \cap B_i) = \min\{P(A_i), P(B_i)\}. \tag{50}$$

$C = \Lambda$ now, and $P(\Lambda) = 1$ is enforced by Kolmogorov's first axiom.

The allowed probability distributions all belong in the same family and for $(a_1, a_2, a_3, b_1, b_2, b_3) \in \Lambda_4$, where a_1 is the probability to be in A_1 and so on, they take the form:

$$(\alpha, \beta, \gamma, 1-\alpha, 1-\beta, 1-\gamma), \tag{51}$$

where $0 \leqslant \alpha, \beta, \gamma \leqslant 1$.

Since there are no relations that govern the probabilities between the different pairs, these three sets of ontic states (1, 2, and 3) are independent of each other. Convex combinations of any probability distribution defined on these disjoint ontic states produce another probability distribution on the disjoint ontic states that satisfies the constraints given by equations (48)–(50); there is only one convex set of probability distributions.

On the other hand, using disjoint ontic states, we can set $W_1 = \xi_{rq} \setminus \xi_{qr}$, $X_1 = \xi_{qr} \setminus \xi_{rq}$, $Y_1 = \xi_{rq} \cap \xi_{qr}$, and $Z_1 = (\xi_{rq} \cup \xi_{qr})^c$, and then add two additional pairs: $\{W_2 = \xi_{pq} \setminus \xi_{qp}, X_2 = \xi_{qp} \setminus \xi_{pq}\}$ and $\{W_3 = \xi_{pr} \setminus \xi_{rp}, X_3 = \xi_{rp} \setminus \xi_{pr}\}$, where we define Y_3, Z_3, Y_4, and Z_4 in a similar manner.

W_i, X_i, Y_i, and Z_i satisfy all the constraints that W, X, Y, and Z did:

$$P(W_i) + P(X_i) + P(Y_i) = \max\{P(W_i) + P(Y_i), P(X_i) + P(Y_i)\}, \tag{52}$$

$$P(Y_i) = \min\{P(W_i) + P(Y_i), P(X_i) + P(Y_i)\}, \tag{53}$$

$$P(W_i) + P(X_i) + P(Y_i) + P(Z_i) = 1, \tag{54}$$

$$P(W_i) + P(X_i) + 2P(Y_i) = 1. \tag{55}$$

Now there are $2^3 = 8$ families of solutions that satisfy equations (52)–(55). As before in equations (33) and (34), we can find that $P(Z_i) = P(Y_i)$ and so we discard $P(Z_i)$ when listing these eight cases $(w_1, w_2, w_3, x_1, x_2, x_3, y_1, y_2, y_3) \in \Lambda_5$, where w_1 is the probability of being in W_1 and so on. The set of solutions corresponds to all possible permutations of the two solutions given in equations (33) and (34) extended to three independent pairs:

1. $(\alpha, \beta, \gamma, 0, 0, 0, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (56)

2. $(\alpha, \beta, 0, 0, 0, \gamma, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (57)

3. $(\alpha, 0, \gamma, 0, \beta, 0, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (58)

4. $(\alpha, 0, 0, 0, \beta, \gamma, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (59)

5. $(0, \beta, \gamma, \alpha, 0, 0, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (60)

6. $(0, \beta, 0, \alpha, 0, \gamma, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (61)

7. $(0, 0, \gamma, \alpha, \beta, 0, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (62)

8. $(0, 0, 0, \alpha, \beta, \gamma, \frac{1}{2}(1-\alpha), \frac{1}{2}(1-\beta), \frac{1}{2}(1-\gamma)),$ (63)

where $0 \leqslant \alpha, \beta, \gamma \leqslant 1$. These cases only contain one common probability distribution: the distribution $(0, 0, \frac{1}{2}, 0, 0, \frac{1}{2}, 0, 0, \frac{1}{2}) \in \Lambda_5$ when $\alpha = \beta = \gamma = 0$.

Again, these are eight convex subsets of the 8-simplex of all distributions over eight ontic states; convex combinations of the probability distributions above do not satisfy equations (52)–(55) (unless $\alpha = \beta = \gamma = 0$).

We have thus established that the Grassmann WWM formalism is equivalent to an ontological model defined by three pairs of non-disjoint ontic states for the stabilizer subtheory and produces eight convex subsets when expressed in terms of disjoint ontic states. In the subsequent section 4, we develop more of its properties.

6. Properties of the Grassmann WWM ontological model

In the eight-state model, the ontic space is partitioned into eight disjoint states that are indexed by the eight three-tuples in Λ_1:

$$\begin{aligned}\Lambda = \{&(+,+,+), (+,+,-), (+,-,+), (-,+,+),\\ &(+,-,-), (-,+,-), (-,-,+), (-,-,-)\}.\end{aligned} \tag{64}$$

Convex combinations of these eight tuples defines any valid probability distribution in the eight-state model.

These three-tuples can be converted into equivalent six-tuples by defining the six-tuples to be $(x_+, y_+, z_+, x_-, y_-, z_-) \in \Lambda_6$, where $x_+ = 1$ and $x_- = 0$ if the first entry of the corresponding three-tuple is '+' and $x_+ = 0$ and $x_- = 1$ if it the first entry is '−' and so on. This produces a partition of the ontic space into eight six-tuples

$$\begin{aligned}\Lambda' = \{&(1,1,1,0,0,0),(1,1,0,0,0,1),\\ &(1,0,1,0,1,0),(1,0,0,0,1,1),(0,1,1,1,0,0),\\ &(0,1,0,1,0,1),(0,0,1,1,1,0),(0,0,0,1,1,1)\}.\end{aligned} \tag{65}$$

Using six-tuples (Λ_6) instead of three-tuples (Λ_1) simplifies the resultant probability distribution of convex combinations because they can now be represented by a single six-tuple. For instance, the probability distribution $\frac{1}{2}(+,+,+) + \frac{1}{2}(+,-,-) \in \Lambda_1$ cannot be simplified any further but $\frac{1}{2}(1,1,1,0,0,0) + \frac{1}{2}(1,0,0,0,1,1) = (1, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}) \in \Lambda_6$. For general probability distributions, the equation for component-wise convex addition is

$$\begin{aligned}&\alpha(x_+, y_+, z_+, x_-, y_-, z_-) + \beta(x'_+, y'_+, z'_+, x'_-, y'_-, z'_-) = (\alpha + \beta)\\ &(x_+ + x'_+, y_+ + y'_+, z_+ + z'_+, x_- + x'_-, y_- + y'_-, z_- + z'_-).\end{aligned} \tag{66}$$

The six-tuple notation is still useful in simplifying convex combinations of ontic states into a single tuple when applied to the Grassmann WWM ontological model's probability distributions, defined to be $(w_1, w_2, w_3, x_1, x_2, x_3) \in \Lambda_6$. However, now convex combinations of probability distributions must additionally satisfy equations (48)–(50) and so the same simple component-wise addition rule of equation (66) does not hold.

Nevertheless, the six-tuple is useful in another way for the Grassmann WWM ontological model because for probability distributions that correspond to quantum states $\hat{\rho}$, its entries correspond to the coefficients in front of the ontic states in the Weyl symbol of the state when it is written with the minimal number of terms such that all coefficients are non-negative (a unique form) [9]:

$$\bar{g} = (g_p,\, g_r,\, g_q,\, g_{\text{-}p},\, g_{\text{-}r},\, g_{\text{-}q}) \in \Lambda_6, \tag{67}$$

where $g_p = P_\rho(\xi_r\xi_q \setminus \xi_q\xi_r) = \max\{0, 2\mu_g(\xi_r\xi_q)\}$, $g_{-p} = P_\rho(\xi_q\xi_r \setminus \xi_r\xi_q) = \max\{0, 2\mu_g(\xi_q\xi_r)\}$, etc. Since a stabilizer state ρ_{stab} is given by equation (41), and the entries in a six-tuple in Λ_6 correspond to $P_\rho(\xi_j\xi_k \setminus \xi_k\xi_j)$, the six stabilizer states correspond to the probability distributions,

$$\begin{aligned}\rho_{\text{stab}} \in \{&(1,0,0,0,0,0),(0,1,0,0,0,0),(0,0,1,0,0,0)\\ &(0,0,0,1,0,0),(0,0,0,0,1,0),(0,0,0,0,0,1)\}.\end{aligned} \tag{68}$$

Therefore, for stabilizer states the entries in the six-tuple are five 0s and a single 1. This leads to a generalized discrete notion of conserved area or symplecticity for Clifford gates on stabilizer states [9]. For all these reasons, we will proceed to use this six-tuple notation from this point onwards.

Note that $(w_1, w_2, w_3, x_1, x_2, x_3) \in \Lambda_6$ uniquely identifies any probability distribution in the Grassmann WWM ontological model since y_i and z_i can be determined from w_i and x_i ($y_i = z_i = \frac{1}{2}(1 - \max\{w_i, x_i\})$) as we showed in the last section and equations (56)–(63).

The eight ontic states of the eight-state model given by equation (65) in the six-tuple notation, also serve as a valid basis for the convex combination (vector space) operation in the Grassmann WWM ontological model with the six-tuple appropriately redefined to be $(w_1, w_2, w_3, x_1, x_2, x_3)$. This can be shown by noting that the three sets of ontic states W_i, X_i, Y_i and Z_i (or A_i and B_i) are independent and that, for (w_i, x_i), convex combinations of $(1,0)$ and $(0,1)$ determine all the possible probability distributions given by equations (33) and (34)

(after they are converting back to their non-disjoint counterparts, convex added according to the constraints given by equations (24)–(27), and the converted back to the disjoint w_i and x_i). Thus, convex combinations of the Cartesian product $\{(1,0),(0,1)\}^3$ must determine all the possible probability distributions given by the larger set of three pairs of independent ontic states. This Cartesian product corresponds to the eight states given by the six-tuples in equation (65).

To find the overlap between a probability distribution $\rho = (w_1^\rho, w_2^\rho, w_3^\rho, x_1^\rho, x_2^\rho, x_3^\rho)$ and one of the eight ontic states $\lambda \equiv (w_1^\lambda, w_2^\lambda, w_3^\lambda, x_1^\lambda, x_2^\lambda, x_3^\lambda) \in \Lambda'$ of equation (65), one must be careful to include their probability densities in the intersections y_i and complements z_i, which as we pointed out, are uniquely determined by w_i and x_i. In particular, the eight ontic states λ in equation (65) have support of 1 on three w_i and/or x_is, and 0 on all the others. Hence they must have support of 0 on all y_is and z_is (since $y_i = z_i = \frac{1}{2}(1 - \max\{w_i, x_i\})$).

We have shown that every stabilizer state probability distribution has support of 1 on one w_i or x_i and 0 on all the other w_is and x_is. This means that each stabilizer state distribution has support of $\frac{1}{2}$ on two pairs of y_i and z_i. Therefore, stabilizer state probability distributions have non-zero overlap with the four ontic states in equation (65) that also have a '1' in the same entry of their six-tuple. For instance,

$$\begin{aligned}(1,0,0,0,0,0) = \frac{1}{4}\,[&(1,1,1,0,0,0) + (1,1,0,0,0,1)\\ &+(1,0,1,0,1,0) + (1,0,0,0,1,1)]\,.\end{aligned} \tag{69}$$

These actually (superficially) correspond to the same convex combinations as in the eight-state model—as we have seen, the reasoning in the Grassmann WWM ontological involves tallying up additional y_i and z_i regions that do not exist in the eight-state model.

According to definition 2, the fact that stabilizer states have support on more than one ontic state means that the Grassmann model is a ψ-epistemic ontological model for the stabilizer state subtheory; the probability distributions of different (but non-orthogonal) stabilizer states overlap.

In summary, even with disjoint ontic states that must be converted to non-disjoint states when taking convex combinations to account for the constraints given by equations (48)–(50), component-wise addition of support on the eight ontic states $\lambda \in \Lambda'$ given by equation (65), which are equivalent to the eight-state model's, still holds as a way to determine probability distribution overlap.

The Grassmann WWM ontological model is preparation non-contextual for the stabilizer subtheory. Any convex combination of stabilizer state probability distributions produces a unique probability distribution. If two probability distributions are not the same, they do not correspond to the same state since they directly one-to-one map to the Weyl symbol of the state.

The model is also measurement non-contextual for the stabilizer subtheory since the stabilizer state probability functions given by equation (68) correspond to the conditional probability functions $\xi_k^M : \Lambda' \to [0,1]$ of Pauli measurement M, where the probability of outcome k given measurement M,

$$Pr(k|M) = \sum_{\lambda\in\Lambda'} \xi_k^M(\lambda)\rho(\lambda). \tag{70}$$

This is only the same as the probability of outcome k under another measurement M' for all stabilizer states ρ if the two are equivalent measurements ($\xi_k^M = \xi_k^{M'} \Leftrightarrow (k,M) \cong (k,M')$) because no two stabilizer states probability distributions produce the same overlaps with all other stabilizer state probability distributions.

We also demonstrated in our prior work that the Grassmann WWM formalism, and therefore the Grassmann ontological model, exhibits measurement and preparation non-contextuality for one qubit [9].

There are a few different ways of illustrating the unions and intersections of the eight ontic states λ given by equation (65) in a Venn diagram. We choose to use the Edwards–Venn diagram approach, which takes a hemispheric approach to illustrating overlapping regions between ontic states [24]; every additional ontic state added to the Venn diagram has more 'leafs' or hemispheres that overlap with all previous ontic states thereby capturing all possible combinations of intersections with them. We show our ontic space in the Edwards–Venn diagram of figures 4 and 5.

We can use the properties introduced in this section, along with the eight-state Edwards–Venn diagram, to show that the Grassmann WWM ontological model does not exhibit transformation contextuality under the 'blowtorch' map. By using the tuples in Λ_6 (the same as the set $\bar{g}$ used in [9] and defined by equation (67)) to organize the probability distributions of states ρ, we can make use of the fact that their entries correspond to the coefficients in front of the monomials of a state's corresponding Weyl symbol, and thereby rely on equations (17)–(21) to see how the states evolve under the Clifford gates X, Y, Z and the Hadamard H. In this way, we see that under the T_1 transformation,

$$
\begin{aligned}
&(w_1, w_2, w_3, x_1, x_2, x_3) \\
&\underset{T_1}{\rightarrow} \frac{1}{4}[(w_1, w_2, w_3, x_1, x_2, x_3) && (71)\\
&+ (x_1, x_2, w_3, w_1, w_2, x_3) \\
&+ (w_1, x_2, x_3, x_1, w_2, w_3) \\
&+ (x_1, w_2, x_3, w_1, x_2, w_3)] && (72)\\
&= (0, 0, 0, 0, 0, 0). && (73)
\end{aligned}
$$

This is the probability distribution for $\hat{I}/2$ and is illustrated in figure 6(a). The final simplification exhibited in equation (73) can be calculated in at least two ways:

1. Convert from disjoint (W_i and X_i) to non-disjoint (A_i and B_i) ontic states, employ equations (48)–(50) to simplify, and then convert back to disjoint states, or
2. Find the Weyl symbol $\rho' = T_1\rho$ and then use equation (40) to obtain the probabilities $P_{\rho'}(\lambda)$ that make up the entries of the resultant six-tuple $\bar{g}$.

These two methods are equivalent because, as discussed, by construction, equations (48)–(50) are a probability theory that captures the Weyl algebra.

Notice that it is not possible to obtain this solution without appealing to the 'coarse' ontic states A_i and B_i either through method 1 or 2. Otherwise, a convex combination of four probability distributions from four different classes of solutions cannot be evaluated while satisfying the constraints given by equations (52)–(55).

On the other hand, acting on line (72) subsequently with the Hadamard gate to effect transformation T_2 produces:

$$
\begin{aligned}
&(w_1, w_2, w_3, x_1, x_2, x_3) \\
&\underset{T_2}{\rightarrow} \frac{1}{4}[(w_3, x_2, w_1, x_3, w_2, x_1) && (74)
\end{aligned}
$$

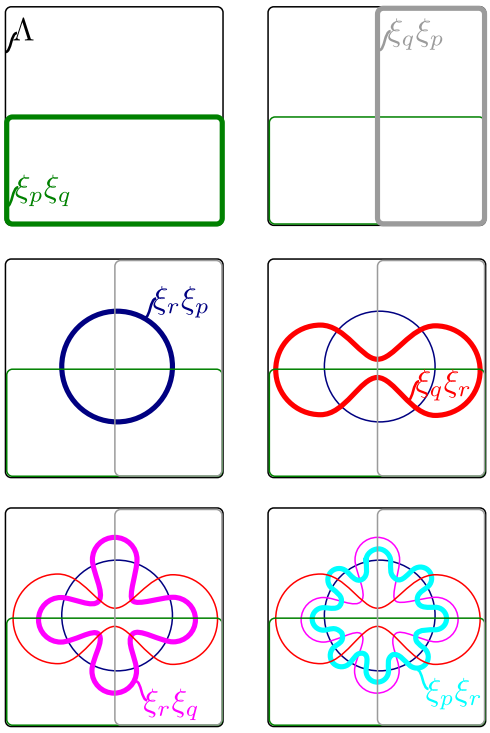

Figure 4. An Edwards–Venn diagram of the ontic space of the Grassmann WWM model that is able to illustrate all the possible overlaps between the non-disjoint ontic states.

$$\begin{aligned}&+(w_3,w_2,x_1,x_3,x_2,w_1)\\&+(x_3,w_2,w_1,w_3,x_2,x_1)\\&+(x_3,x_2,x_1,w_3,w_2,w_1)]\end{aligned} \tag{75}$$

$$=(0,0,0,0,0,0). \tag{76}$$

Again, this is the probability distribution for $\hat{I}/2$ and the final simplification exhibited in equation (76) is the unique $\bar{g}_\rho$ tuple for the final state. This can be seen in figure 6(b). Notably, figure 6 also shows that ignoring the additional constraints given by equations (48)–(50) leads to the inequivalent parity simplices in probability space that are found when T_1 and T_2 are implemented in the eight-state model.

Therefore, the resultant probability distribution $(0,0,0,0,0,0)$ is attained no matter whether transformation T_1 or T_2 is taken in the Grassmann ontological model and so no transformation contextuality is present. The final solution is very similar to the one found in section 4's equation (35) when using the 'coarse' ontic states A, B and C. On the other hand, if the additional constraints are ignored and so a single convex set under the disjoint ontic states is assumed to exist, then different probability distributions are obtained under T_1 and T_2.

We further note that *any* map consisting of convex combinations of states within the stabilizer subtheory will necessarily produce a unique probability distribution for every unique

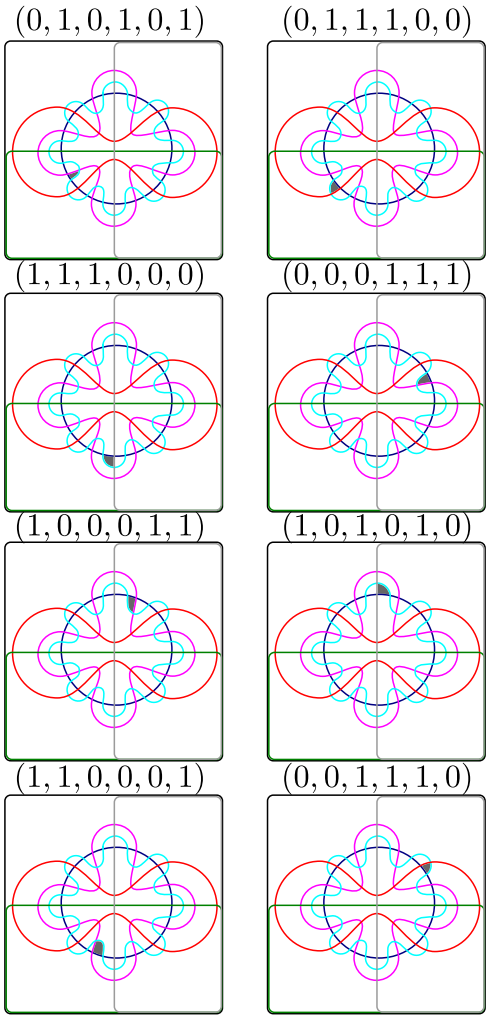

Figure 5. The eight ontic states $\lambda \in \Lambda'$ of the Grassmann WWM ontological model given by equation (65). Though these disjoint ontic states do not appear to cover all of ontic space, since they satisfy the additional equations given by equations (48)–(50) when expressed as non-disjoint states, knowledge of the support of a probability distribution on these eight states is sufficient to determine it everywhere else in ontic space.

quantum state expected, since every Weyl symbol and one-qubit operator is bijectively represented by a probability distribution by the definition of $\bar{g}$ [9] and we have shown that one-qubit convex combination is fully treated at order $\hbar^0$ in section 5. So this result of no transformation contextuality generalizes to all one-qubit maps within the one-qubit stabilizer subtheory.

We proceed to now show why the proof used by Lillystone *et al* explicitly excludes classical probability distributions defined over non-disjoint elements, such as those satisfying additional constraints like those given by equations (48)–(50). We then argue that the Grassmann WWM ontological model is proof that single qubit non-contextuality can be handled by ontological models with non-disjoint ontic states.

7. Inequivalence between ontological models with disjoint ontic states and non-disjoint ontic states

Lillystone *et al* consider an arbitrary preparation non-contextual ontological model of a single qubit stabilizer subtheory. The WWM formalism for a single qubit is such a theory at $\hbar = 0$. They then consider Δ_p to be the support of the quantum state ρ in the ontological model,

$$\Delta_p = \{\lambda | \mu_\rho(\lambda) > 0, \lambda \in \Lambda\}. \tag{77}$$

The proof then proceeds to delete any state $\lambda \in \Lambda$ such that $P_{I/2}(\lambda) = 0$ and partition the remaining set into eight disjoint spanning sets. Since $P_1(\lambda \setminus -\lambda) = 0$ for all $\lambda \in \Lambda$, it follows by equations (48)–(50) that $P_1(\lambda \cap -\lambda) = \frac{1}{2}$ for all $\lambda \in \Lambda$. Therefore, none of the eight Grassmann WWM ontic states given in equation (65) are disqualified.

Lillystone *et al* then proceed to produce a disjoint partition into eight sets. In particular, they rely on repeated application of the following feature of both ψ-ontic and ψ-epistemic ontologicla models: given

$$\mu_\rho(\lambda)\mu_{\rho'}(\lambda) = 0 \ \ \forall \lambda \in \Lambda, \tag{78}$$

this implies that

$$\text{supp}(\mu_\rho) \cap \text{supp}(\mu_{\rho'}) = \emptyset, \tag{79}$$

if $\rho \neq \rho'$ [10].

Since this is true for three pairs of basis states [17], argues that, given preparation non-contextuality, the ontic space can therefore be organized into $2^3 = 8$ disjoint states. The argument is more clearly laid out in [15] and follows the reasoning that since six non-negative states have full support on only one unique basis element of one pair and the same partial support on all the other pairs, it must be possible to partition the space into eight disjoint sets.

For instance, in the eight-state model, the ontic states $x = +$ and $x = -$ are disjoint and so are $y = +$ and $y = -$. Hence, the ontic space can be partitioned into the four disjoint sets $\{(x = +, y = +), (x = +, y = -), (x = -, y = +), (x = -, y = -)\}$. The partition into eight disjoint sets follows from then considering the disjoint sets $z = +$ and $z = -$.

In the Grassmann WWM qubit model, the ontic states $\xi_r\xi_q \setminus \xi_q\xi_r$ and $\xi_q\xi_r \setminus \xi_r\xi_q$ correspond to the eight-state model's $x = +$ and $x = -$ respectively, and $\xi_p\xi_q \setminus \xi_q\xi_p$ and $\xi_q\xi_p \setminus \xi_p\xi_q$ correspond to the eight-state model's $y = +$ and $y = -$ respectively. They can certainly be divided into the disjoint subsets by Lillystone *et al*'s argument and, along with the states $xi_p\xi_r \setminus \xi_r\xi_p$ and $\xi_r\xi_p \setminus \xi_p\xi_r$ that are analogous to the states $z = +$ and $z = -$ respectively, produce the eight disjoint ontic states given by equation (65).

But for the Grassmann WWM model, though equation (79) still holds (disjointness) and reexpressing ontic states in terms of disjoint ontic states produces bipartitions of the ontic space, it does not preserve convex combination of single-qubit state probability distributions. This is because the Weyl algebra over anti-commuting elements is equivalent to imposing additional constraints on top of Kolmogorov's axioms, as we have seen, which have more than one family of solutions when reexpressed in terms of disjoint ontic states and convex combinations between families of solutions is not preserved. This is true even though the model is preparation and measurement non-contextual for one qubit.

Lillystone *et al* complete the proof by directly relying on convex linearity to argue that there exists an implementation of T_1 and T_2 that has the same contextual implementation as theirs, when defined over the eight disjoint sets; they assume that the states in their convex sum fall into the same convex set when considered in terms of disjoint ontic states. They

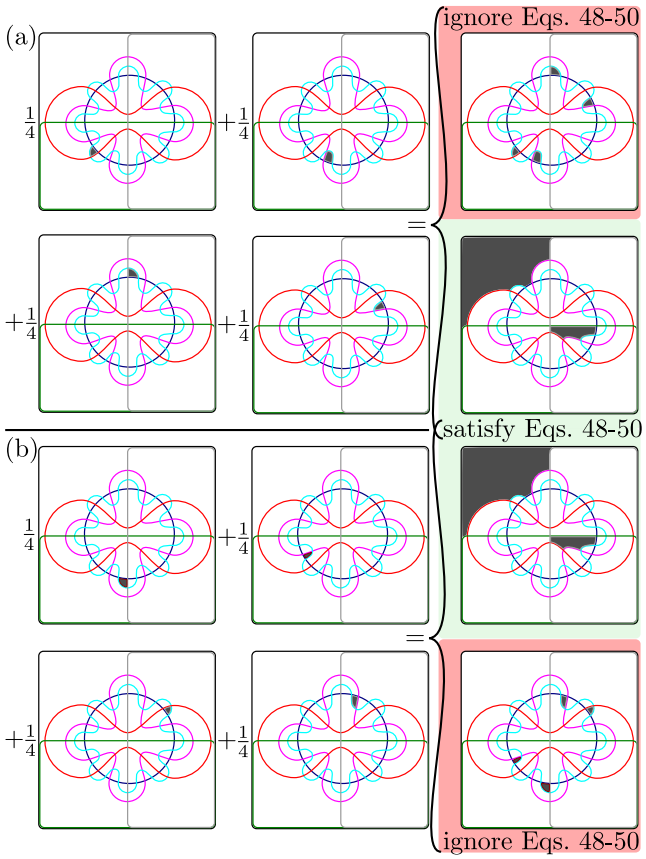

Figure 6. The convex combinations produced by (a) T_1 and (b) T_2 in the ontological model defined over disjoint ontic states. Note that the resultant probability distribution from the convex combination is the same for (a) and (b) when the ontic states are converted to their non-disjoint counterparts and equations (48)–(50) are used. On the other hand, if this additional system of equations is ignored then the resultant probability distributions are different for T_1 and T_2. In fact, the two different resultant probability distributions correspond to the light and dark grey regions of the eight-state model's ontic space indicated in figure 1.

thus implicitly neglect the possibility of an additional set of constraints, commensurate with Kolmogorov's axioms, that does not result in a single convex set that contains all the probability distributions they consider when their ontic states are expressed as disjoint ontic states. Therefore, the Grassmann WWM ontological model lies outside the scope of their argument.

8. Conclusion

This paper answers the question of how to interpret the Grassmann WWM formalism at order $\hbar^0$ in the framework of ontological models established by Leifer [19]. We show that the Grassmann WWM ontological model is ψ-epistemic, that it is most simply expressed in terms of overlapping or non-disjoint ontic states due to an additional set of constraints it must satisfy, that this additional set of constraints only serve to restrict the model which remains a

classical probability model, and that the probability distributions over these ontic states form a single convex set, but that probability distributions over disjoint ontic states do not.

If the additional set of constraints is ignored (i.e. explicitly not satisfied), convex combinations of probability distributions from different families of solutions produce different results when only one is expected. This is the origin of transformation contextuality under the 'blowtorch' map in Lillystone *et al*'s ontological eight-state model over disjoint ontic states. We showed that transformation contextuality is not present in the Grassmann WWM classical probability theory at order $\hbar^0$—an ontological model defined over non-disjoint states with such an additional set of constraints.

The Grassmann model offers a case where Lillystone *et al*'s proof—that preparation noncontextual qubit ontological models exhibit transformation contextuality in the one-qubit stabilizer subtheory—does not hold. Indeed, we have shown that ontological models defined over non-disjoint ontic states appear to be able to treat single-qubit noncontextuality properly, and so not exhibit transformation contextuality in the one-qubit stabilizer subtheory. We therefore contest Lillystone *et al*'s conclusion that 'the single-qubit stabilizer subtheory, a very simple subtheory of the smallest quantum system, exhibits generalized contextuality (and) demonstrates that generalized contextuality is so prevalent that even an essentially trivial quantum subtheory is classified as contextual, and therefore non-classical' [17].

In summary, we have shown that for the qubit stabilizer subtheory to be captured by a preparation, transformation and measurement non-contextual ontological theory, it must be handled in terms of non-disjoint ontic states, unlike the case for the odd-dimensional single-qudit stabilizer subtheory.

As a final point, one can ask more precisely why supplementation of Kolmogorov's axioms by an additional set of constraints does not seem to be present in the literature on ontological models so far. We point out that such an additional set of constraints can always be formulated for any ontological model after it is reexpressed in terms of non-disjoint ontic states. However, prior work has almost always considered ontological models where such an additional set of constraints is trivial because they are too weak; it only produces one family of solutions when the model is reexpressed in term of its original disjoint ontic states. And therefore it is natural that such constraints have not been discussed. Nevertheless, this ability to include an additional set of constraints, available due to the freedom provided by the sum rule (a consequence of σ-additivity), has always been there. In a way, this is an unused 'degree of freedom' of classical probability theories that has been hidden in plain sight all along, or at least since the introduction of the Grassmann algebra in physics.

Acknowledgments

Parts of this manuscript are a contribution of NIST, an agency of the US government, and are not subject to US copyright. PJL acknowledges support from NSF award PHY1720395 and from a Google Faculty Research Award (2017).

ORCID iDs

Lucas Kocia https://orcid.org/0000-0002-2774-0565

References

[1] Howard M, Wallman J, Veitch V and Emerson J 2014 Contextuality supplies the magic for quantum computation *Nature* **510** 351–5
[2] Raussendorf R, Browne D E, Delfosse N, Okay C and Bermejo-Vega J 2017 Contextuality as a resource for qubit quantum computation *Phys. Rev.* A **95** 052334
[3] Raussendorf R and Briegel H J 2001 A one-way quantum computer *Phys. Rev. Lett.* **86** 5188
[4] Okay C, Roberts S, Bartlett S D and Raussendorf R 2017 Topological proofs of contextuality in quantum mechanics (arXiv:1701.01888)
[5] Gross D 2006 Hudson's theorem for finite-dimensional quantum systems *J. Math. Phys.* **47** 122107
[6] Veitch V, Ferrie C, Gross D and Emerson J 2012 Negative quasi-probability as a resource for quantum computation *New J. Phys.* **14** 113011
[7] Mari A and Eisert J 2012 Positive Wigner functions render classical simulation of quantum computation efficient *Phys. Rev. Lett.* **109** 230503
[8] Kocia L, Huang Y and Love P 2017 Semiclassical formulation of the Gottesman–Knill theorem and universal quantum computation *Phys. Rev.* A **96** 032331
[9] Kocia L and Love P 2017 Discrete Wigner formalism for qubits and noncontextuality of Clifford gates on qubit stabilizer states *Phys. Rev.* A **96** 062134
[10] Spekkens R W 2005 Contextuality for preparations, transformations, and unsharp measurements *Phys. Rev.* A **71** 052108
[11] Redhead M 1987 *Incompleteness, Nonlocality, and Realism: a Prolegomenon to the Philosophy of Quantum Mechanics* (Oxford: Clarendon)
[12] Peres A 1990 Incompatible results of quantum measurements *Phys. Lett.* A **151** 107–8
[13] Mermin N D 1990 Simple unified form for the major no-hidden-variables theorems *Phys. Rev. Lett.* **65** 3373
[14] Kocia L and Love P 2018 Measurement contextuality and Planck's constant *New J. Phys.* **20** 073020
[15] Wallman J J and Bartlett S D 2012 Non-negative subtheories and quasiprobability representations of qubits *Phys. Rev.* A **85** 062121
[16] Blasiak P 2013 Quantum cube: a toy model of a qubit *Phys. Lett.* A **377** 847–50
[17] Lillystone P, Wallman J J and Emerson J 2018 Contextuality and the single-qubit stabilizer formalism (arXiv:1802.06121)
[18] Karanjai A, Wallman J J and Bartlett S D 2018 Contextuality bounds the efficiency of classical simulation of quantum processes (arXiv:1802.07744)
[19] Leifer M 2014 Is the quantum state real? an extended review of ψ-ontology theorems *Quanta* **3** 67–155
[20] Ash R B 1970 *Basic Probability Theory* (New York: Wiley)
[21] Harrigan N and Spekkens R W 2010 Einstein, incompleteness, and the epistemic view of quantum states *Found. Phys.* **40** 125–57
[22] van den Brink A M 2005 personal communication
[23] Berezin F A and Marinov M S 1977 Particle spin dynamics as the Grassmann variant of classical mechanics *Ann. Phys.* **104** 336–62
[24] Edwards A W F 2004 *Cogwheels of the Mind: the Story of Venn Diagrams* (Johns Hopkins University Press)