Commutation Groups and Algebraic Contextuality

Șerban Cercelescu

Submitted in Partial Fulfillment for the Degree of

Master of Mathematics and Computer Science

Trinity Term

13$^{\text{th}}$ of May 2024

To Jessica, for making the good days great and
the bad days meh for these four wonderful years.

Page intentionally left blank.

Acknowledgments

First and foremost, I would like to thank Samson Abramsky for giving me the honour of agreeing to supervise my master's thesis, for being a constant academic and professional support for the past year, for keeping the spirits up when I felt that my results went nowhere and for introducing me to the field of contextuality and to the many nice people working in it. I would like to thank Carmen Constantin for her great work and support throughout this research project, for the excellent host she was throughout our meetings and for inspiring me and many others through her past work and presence in the Computer Science Department. I would also like to thank Sam Staton for agreeing to formally supervise this thesis and review it. Finally, I would like to thank "Agenția Națională de Burse și Credite" for the financial support they've offered me for part of my last year in Oxford through the scholarship "H.G. nr. 118/2023".

Disclaimer

Soon after I began writing this document, I realised that there exists an unavoidable conflict between my obligation to best preserve my anonymity for the sake of the fairness of the examination and my obligation to cite the origin of the results I am using. This conflict stems from the fact that I have published a conference paper together with S. Abramsky and C. Constantin which bears my name on it and covers many of the results presented in this work. Citing this paper and claiming originality over some of its content would reveal my name. The compromise I have resorted to is attaching the paper at the end of the document with any reference to my name removed. The attached paper is not part of this thesis and is not considered in the word-count, being present in the pdf solely to resolve the citation issue. One should also not treat this thesis as a more detailed version of the paper, as it contains some different results. *I also adopt the convention that results present in the paper that have*

been discovered by me will be not cited, but all others will be.

Moreover, there is a clear discrepancy between the submitted project proposal and the content of this thesis. While our original goal was to study and generalize the arguments in [2] using contextuality to separate the complexity classes $\mathbf{NC}_0$ and $\mathbf{QNC}_0$, after a couple of months of attempting to familiarise myself with [2], it has proved itself to be too risky of a challenge for me to approach as a topic for my master's thesis.

Contents

Abstract

We introduce *commutation groups* as an algebraic structure which enables one to obtain state-independent contextuality arguments from frequently occurring sets of observables. We provide a classification theorem of commutation groups and study their general algebraic properties, such as their representations, relationship to the Pauli and Heisenberg groups. Finally, we introduce multiple methods of exhibiting contextuality in empirical models stemming from observables that are part of a commutation group.

1 Introduction

1.1 How we got here and why we're still here

The origins of the study of contextuality lie unambiguously in the phenomenology of quantum mechanics. It the author's view that this study originates in Einstein's opposition to the idea that quantum mechanics is a complete theory of mechanics at a small scale; we trace the chronology of this as follows: in June 1926, Max Born published "Zur Quantenmechanik der Stoßvorgänge" [9] (On the Quantum Mechanics of Collision Processes), first describing the probabilistic interpretation of Schrödinger's wave mechanics, description which now bears his name as "the Born rule." Max Born was a friend of Albert Einstein, with whom he frequently corresponded on academic and personal matters, and on the 4^{th} of December that year, Einstein addressed him a letter containing the lines:

"Aber, eine innere Stimme sagt mir, daß das doch nicht der wahre Jakob ist. Die Theorie liefert viel, aber dem Geheimnis des Alten bringt sie uns kau näher. Jedenfalls bin ich überzeugt, daß Der nicht würfelt."

Our English translation of this is:

"However an inner voice tells me that this is not the real Jacob. The theory delivers a lot, however it hardly brings us closer to the secret of the Old One. In any case, I am convinced that He does not throw dice."

More of this correspondence can be found in [1], a book containing many of the letters exchanged between Einstein and Born, commented by Born himself, containing an introduction written by Bertrand Russel and a foreword by Werner Heisenberg. Later in 1926, Niels Bohr allegedly (but famously) told Einstein to *"stop telling God what to do"*. It our opinion that the letter marks the beginning of Einstein's opposition to the notion that quantum mechanics is a complete theory, and it is Bohr's reaction to this letter that marks the beginning of the Bohr-Einstein debates. While this debate continued until Einstein's death and had a remarkable academic showdown during the 5^{th} Solovay conference in 1927, in order to stay on topic, we fast forward to May 1935, when Einstein, Podolsky and Rosen published the paper "Can Quantum Mechanical Description of Reality be Considered Complete?" [11]. The paper tried to present the phenomenon of quantum steering as a paradox. It is indeed highly physically counter-intuitive that for example, given a pair of physically separated qubits in a Bell state, performing a measurement of one of the qubits in one site in a chosen way determines the state of the other qubit in a manner *dependent on the choice of measurement*, while not transferring any information to the other site. Einstein deduced that in order to preserve what is now known as *local realism*, the quantum state of the second qubit cannot fully describe its "real state" and that there must be some sort of hidden information not described by quantum mechanics behind this phenomenon. While Einstein's opposition to the completeness of quantum theory mainly concerned its probabilistic nature, this time, the disputed phenomenon was precisely that of non-locality, which mathematically can be seen as a special case of contextuality [4]. Bohr published later that year an article with the exact same title and in the same journal [8] attacking the premises of Einstein, Podolsky and Rosen's argument and casting doubt on whether the phenomenon was indeed a paradox in

the first place. However, Bohr did not concretely refute the possibility of the existence of a more general physical theory whose description of the state of an object would not violate local realism. The one who would solve this issue definitively would be John Stewart Bell in his paper (submitted in November 1964) titled "On the Einstein Podolsky Rosen Paradox" [7], where he showed that no such more general theory (as described by Einstein) can be consistent with the theory of quantum mechanics. Kochen and Specker later independently proved a similar result in their 1966 article "The Problem of Hidden Variables in Quantum Mechanics" [14] (they used different techniques and cited no work published later than 1963). While these papers essentially settled the question of whether a more general theory than quantum mechanics could solve these apparent paradoxes, interest on the topic persisted, notably in the work of Mermin [15] in the '80s and '90s. More recently, work by Abramsky and Brandenburger [4] described contextuality for the first time in a language independent of quantum mechanics, work which led to new techniques in the study of databases [3] and the complexity of constraint satisfaction [5]. The contextuality of a quantum system of measurements, being a distinguishing feature of quantum systems in comparison to classical ones, has also found its way into the study of quantum complexity theory, most notably in [12] and [2].

1.2 Content

The topic of this thesis is the study of contextuality arising from measurement scenarios in which the observables are part of the newly introduced algebraic structure of commutation groups. The key idea behind commutation groups is the following: in many examples (e.g. the Peres-Mermin square) of contextuality arising from quantum scenarios, one of the key components of the proof that no non-contextual assignment exists is the fact that (group theoretical, i.e. $[g, h] = ghg^{-1}h^{-1}$) commutators of certain observables are of the form $[A, B] = \omega_n^k I_n$. Commutation groups were introduced to abstract away the concrete linear algebra behind such arguments and work within a group-theoretical framework, the definition of contextuality for a commutation group being dependent solely on the group structure and not of any

choice of representation (see Theorem 4.7). In a spirit similar to that of [4], we define the notion of algebraic contextuality in a language that is independent of quantum theory.

While in the incipient stages of our research, we considered commutation groups as generalizations of the Pauli group, the discovery of Theorem 3.23 not that long before the finalization of [17] forced a perspective change, namely that commutation groups are a handy group theoretical gadget that describes certain subgroups of the generalized Pauli group in a manner that enables a family of contextuality proofs.

1.3 Notation and Conventions

- Given integers $n, m \in \mathbb{Z}$, $[n]$ and $[n..m]$ denote the discrete intervals $\{1, 2, 3, \ldots, n\}$ and $\{n, n+1, \ldots, m-1, m\}$ respectively. If $m > n$, $[n..m] = \varnothing$.
- We write $\omega_d := \exp(2\pi i/d)$.
- We write $\mathcal{Q}_d := \mathbb{C}^d$ for the state-space of a d-dimensional qudit and denote the computational basis on $\mathcal{Q}_d$ by $\{|k\rangle \mid k \in [0..d-1]\}$;
- We define the **generalized Pauli group** $\mathcal{P}_d^n$ as follows: $\mathcal{P}_d$ is the group generated by the $d \times d$ (complex) matrices σ_x, σ_z and $\omega_d I_d$, where

$$\sigma_x = \sum_{k=0}^{d-1} |k+1\rangle\langle k| \qquad\qquad \sigma_z = \sum_{k=0}^{d-1} \omega_d^k |k\rangle,$$

  where $\{|0\rangle, \ldots, |d-1\rangle\}$ denotes the canonical basis on $\mathbb{C}^d$ and we identify $|d\rangle := |0\rangle$. The group $\mathcal{P}_d^n$ is defined as $\mathcal{P}_d^{\otimes n}$.
- Given a ring R, $\mathfrak{so}(n, R)$ denotes the set of $n \times n$ anti-symmetric matrices with entries in R and are zero on the main diagonal.
- Given a natural number $n \in \mathbb{N}_{\geqslant 1}$ and a ring R, we will routinely refer to the following elements the canonical basis of R^n:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad \dots \quad \vec{e}_{n-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix} \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

- We routinely identify $n \times n$ matrices with entries in an ring R with their corresponding bilinear forms on the free R-module R^n.
- Given a commutation group $G(\mu)$ for some commutator matrix $\mu : X_\mu \times X_\mu \to \mathbb{Z}_d$, writing $X_\mu = \{x_1, \dots, x_n\}$, we will often write μ_{ij} for $\mu(x_i, x_j)$ and consider μ as a matrix in $\mathfrak{so}(n, \mathbb{Z}_d)$. Analogously, we'll identify matrices $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ with functions $\mu : X_\mu \times X_\mu \to \mathbb{Z}_d$ as in Definition 2.1.
- X^* denotes the Kleene star operator over the alphabet X.
- Given an $n \times n$ matrix μ over a ring R, $\tilde{\mu}$ denotes its upper-triangular part, i.e.

$$\tilde{\mu}_{ij} = \begin{cases} \text{if } i < j : & \mu_{ij} \\ \text{otherwise:} & 0 \end{cases}$$

- Throughout this paper, we assume all Hilbert spaces to be finite dimensional.
- We use **bold** for newly introduced words and *italics* for emphasis.

2 Contextuality

In this section we present a formalization of the notion of contextuality as introduced in [AB2011].

2.1 Measurement Scenarios and Empirical Models

Contextuality is a phenomenon which arises when considering a system where one may perform measurements, thus obtaining certain outcomes, where some of the measurements are mutually exclusive. We formalize this notion as follows:

Definition 2.1 A **measurement scenario** is a tuple $(X, \mathcal{M}, \mathcal{O})$, where X is a set whose elements are called **measurements**, $\mathcal{O}$ is a set whose elements are called **outcomes** and $\mathcal{M} \subseteq \mathcal{P}(X)$ is a downwards closed (i.e. for any $A, B \subseteq X$ such that $A \subseteq B$ and $B \in \mathcal{M}$, $A \in \mathcal{M}$) set whose elements are called **contexts**, such that $\bigcup_{C \in \mathcal{M}} C = X$. For the entirety of this thesis, we will assume that X and $\mathcal{O}$ are finite.

Note that the pair $(X, \mathcal{M})$ forms a simplicial complex. This has a couple of useful consequences which are discussed in [10].

Example 2.2 Suppose that we have access to a pair of qubits in the Bell state $|\Psi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, on which we may perform any of the measurements corresponding to the observables in the set $X = \{\sigma_x \otimes I, \sigma_z \otimes I, I \otimes \sigma_x, I \otimes \sigma_z\}$. Due to commutativity restrictions we may not jointly measure all of the observables in X, but we may only simultaneously measure a subset C of observables in X if for any $A, B \in C$, $AB = BA$; we write $\mathcal{M}$ for the set of such subsets. Upon measuring any one of the observables, we obtain outcomes from the set $\mathcal{O} = \{+1, -1\}$, corresponding to the eigenvalues of our observables. We thus obtain a measurement scenario $(X, \mathcal{M}, \mathcal{O})$ which models the compatibility relations between the observables present in the described system of measurements.

Given a measurement scenario $(X, \mathcal{M}, \mathcal{O})$, one may describe an event stemming from performing the measurements as a pair (C, e_C) where $C \in \mathcal{M}$ is a context consisting of the measurements which were performed and $e_C : C \to \mathcal{O}$ is a function mapping each of the measurements in C to its corresponding outcome. A mathematical structure which can track all the possible events arising from a measurement scenario is a sheaf.

Definition 2.3 Given a measurement scenario $(X, \mathcal{M}, \mathcal{O})$, its corresponding **event sheaf** $\mathcal{E}$ consists of the following data:

- To each subset $U \subseteq X$, we associate a set $\mathcal{E}(U) := \mathcal{O}^U$, whose elements are called the **sections** of $\mathcal{E}$ at U. We may thus think of $\mathcal{E}$ as a (**Set**)-sheaf over the discrete topology on X.
- To each pair of nested subsets $U \subseteq V \subseteq X$, we associate a **restriction mapping** $\rho_U^V : \mathcal{E}(V) \to \mathcal{E}(U)$, which is given by function restriction, i.e. for any $e \in \mathcal{E}(V)$, $\rho_U^V(e) = e|_U$.

Of course, the event sheaf on its own has no real use and might seem as an arbitrary use of sophisticated mathematics, however, we will soon see that its structure allows one to elegantly track probability distributions arising from systems modelled by measurement scenarios. To see this, however, we must first define the following two functors:

Definition 2.4

- The **probability distribution functor** $\mathcal{D} : \mathbf{Set} \to \mathbf{Set}$, is the functor mapping each set X to the set of probability distributions over X (where X is considered with its discrete measure), i.e.

$$\mathcal{D}X = \{p : X \to [0,1] | \sum_{x \in X} p(x) = 1\},$$

and sending each map $f : X \to Y$ to a map $\mathcal{D}f : \mathcal{D}X \to \mathcal{D}Y$, such that for any $p \in \mathcal{D}X$,

$$(\mathcal{D}f)(p)(y) = \sum_{x \in f^{-1}(y)} p(x).$$

- The **possibility distribution functor** $\mathcal{P} : \mathbf{Set} \to \mathbf{Set}$ is the functor mapping each set X to its set of subsets $\mathcal{P}(X) := \{A \subseteq X\}$ and sending each map $f : X \to Y$ to a map $\mathcal{P}(f) : \mathcal{P}(X) \to \mathcal{P}(Y)$ such that for any $A \in \mathcal{P}(X)$, $(\mathcal{P}f)(A) = \{f(a) : a \in A\}$.

Note that the possibility distribution functor is in fact the usual set of subsets endofunctor, however, we rename it for a perspective change; namely we can think of an element $A \in \mathcal{P}(X)$ (where X is some set of events) as the subset of events which are possible. Similarly, we may think of a function $f : X \to Y$ as of a map which relabels events labeled by elements of X to events labeled by elements of Y.

Composing an event sheaf as above with the probability distribution functor results in a *presheaf* $\mathcal{DE}$, which maps each subset $C \subseteq X$ to the set $\mathcal{DE}(C) = \mathcal{D}(\mathcal{O}^C)$ of probability distributions over the joint outcomes of the measurements in C. Moreover, for each nested pair of subsets $U \subseteq V \subseteq X$, we get a restriction map $\mathcal{D}(\rho^V_U) : \mathcal{DE}(V) \to \mathcal{DE}(U)$ corresponding to marginalizing these distributions:

$$\mathcal{DE}(C') = \mathcal{D}(\textstyle\prod_{c \in C'} \mathcal{O}) \xrightarrow{\mathcal{D}(\rho^{C'}_C)} \mathcal{DE}(C) = \mathcal{D}(\textstyle\prod_{c \in C} \mathcal{O})$$

Similarly, by composing an event sheaf with the possibility distribution functor, we obtain a *presheaf* $\mathcal{PE}$ which maps each $C \subseteq X$ to the set $\mathcal{PE}(C)$ of possibility distributions over the outcomes of measurements in C and for each nested pair of subsets $U \subseteq V \subseteq X$ contains a restriction map $\mathcal{P}(\rho^V_U) : \mathcal{PE}(V) \to \mathcal{PE}(U)$.

Definition 2.5 Given a measurement scenario $\mathcal{S} = (X, \mathcal{M}, \mathcal{O})$, a **probabilistic empirical**

model over $\mathcal{S}$ is defined as a family of probability distributions $e = \{e_C \in \mathcal{DE}(C) \,|\, C \in \mathcal{M}\}$, that satisfy the **no-signalling condition**, that is, for any two contexts $U, V \in \mathcal{M}$, $\rho^U_{U\cap V}(e_U) = \rho^V_{U\cap V}(e_V)$. In other words, the probability of a joint outcome of a set of measurements occurring is independent of which other (compatible) measurements are performed. Similarly, a **possibilistic empirical model** $\mathcal{S}$ is defined as a family of possibility distributions $e = \{e_C \in \mathcal{PE}(C) \mid C \in \mathcal{M}\}$, obeying the same condition that for any two contexts $U, V \in \mathcal{M}$, $\rho^U_{U\cap V}(e_U) = \rho^V_{U\cap V}(e_V)$.

Note that given any probabilistic empirical model e over a measurement scenario $\mathcal{S}$ as above, we obtain a possibilistic empirical model e' given by the supports of the probability distributions, i.e. for any $C \in \mathcal{M}$, $e'_C := \operatorname{supp} e_C$.

Example 2.6 Consider the measurement scenario from Example 2.2. By performing the measurements corresponding to the observables in X, we obtain the empirical model given by the following table, from which we can read, for example, that the probability of measuring $+1$ when measuring both observables $\sigma_X \otimes I$ and $I \otimes \sigma_X$ is $1/2$:

	$(\sigma_X \otimes I, I \otimes \sigma_X)$	$(\sigma_X \otimes I, I \otimes \sigma Z)$	$(\sigma_Z \otimes I, I \otimes \sigma_X)$	$(\sigma_Z \otimes I, I \otimes \sigma_X)$
$(+1,+1)$	$1/2$	$1/4$	$1/4$	$1/2$
$(+1,-1)$	0	$1/4$	$1/4$	0
$(-1,+1)$	0	$1/4$	$1/4$	0
$(-1,-1)$	$1/2$	$1/4$	$1/4$	$1/2$

Note that when describing an empirical model e over a measurement scenario $(X, \mathcal{M}, \mathcal{O})$, due to the no-signalling condition, it suffices to specify the values e_C solely for maximal contexts $C \in \mathcal{M}$.

2.2 Types of Contextuality

We are now ready to finally define what contextuality is:

Definition 2.7 Given a probabilistic (or possibilistic) empirical model e over a measurement scenario $\mathcal{S} = (X, \mathcal{M}, \mathcal{O})$, we say that e is **contextual** if there exists no global section $g \in \mathcal{DE}(X)$ (or $g \in \mathcal{PE}(X)$) such that $e_C = \rho^X_C(g)$ for all $C \in \mathcal{M}$.

One can think of a contextual empirical model as a witness to the failure of $\mathcal{DE}$ (or $\mathcal{PE}$) to be a sheaf. This alternative perspective led to the cohomological techniques used to exhibit contextuality in certain measurement scenario in [6].

Note that given a contextual probabilistic empirical model e, its corresponding possibilistic empirical model need not be necessarily contextual. On the other hand, it is straightforward to prove that the converse implication does hold, namely, if e's corresponding possibilistic model is contextual, then e need be contextual as well. This enables us to refine the definition of contextuality as follows:

Definition 2.8 Let e be a possibilistic empirical model over a measurement scenario $\mathcal{S} = (X, \mathcal{M}, \mathcal{O})$.

- Let $C \in \mathcal{M}$ be some context and $s \in e_C$ be some *possible* event in the event sheaf. We say that e is **logically contextual** at s and write $\mathrm{LC}(e, s)$ if there exists no global section $g \in \mathcal{E}(X)$ of the event sheaf such that $g|_C = s$ and $g|_{C'} \in e_{C'}$ for every other context $C' \in \mathcal{M}$.
- We say that e is **strongly contextual** and write $\mathrm{SC}(e)$ if for every context $C \in \mathcal{M}$ and every $s \in e_C$, we have that $\mathrm{LC}(e, s)$. Equivalently, e is strongly contextual if there exists no $g \in \mathcal{E}(X)$ such that $g|_C \in e_C$ for every context $C \in \mathcal{M}$.

We can extend these definitions to probabilistic empirical models as well as follows: let e

be a probabilistic empirical model over $\mathcal{S}$;

- Let $C \in \mathcal{M}$ be some context and $s \in \operatorname{supp} e_C$ be some possible event in the event sheaf. We say that e is **logically contextual** at s and write $\mathrm{LC}(e, s)$ if there exists no global section $g \in \mathcal{E}(X)$ of the event sheaf such that $g|_C = s$ and $g|_{C'} \in e_{C'}$ for every other context $C' \in \mathcal{M}$.

- We say that e is **strongly contextual** and write $\mathrm{SC}(e)$ if for every context $C \in \mathcal{M}$ and every $s \in e_C$, we have that $\mathrm{LC}(e, s)$. Equivalently, e is strongly contextual if there exists no $g \in \mathcal{E}(X)$ such that $g|_C \in \operatorname{supp} e_C$ for every context $C \in \mathcal{M}$. In other words, e is strongly contextual if its corresponding possibilistic empirical model is strongly contextual.

Of course, give an empirical model e, if it is strongly contextual or logically contextual at a section, then e is also contextual.

Example 2.9 Consider a measurement scenario with three measurements $X = \{M_1, M_2, M_3\}$, outcome set $\mathcal{O} = \{0, 1\}$ and measurements $\mathcal{M} = \{\{M_1, M_2\}, \{M_2, M_3\}, \{M_3, M_1\}\}^{\downarrow}$, so that any two distinct measurements may be performed, but never all three of them. Let e be the empirical model over $\mathcal{S} = (X, \mathcal{M}, \mathcal{O})$, where any two joint measurements can only result in different outcomes with equal probability, as in the table below:

	(M_1, M_2)	(M_2, M_3)	(M_3, M_1)
$(0,0)$	0	0	0
$(0,1)$	1/2	1/2	1/2
$(1,0)$	1/2	1/2	1/2
$(1,1)$	0	0	0

It is easy to see that e is strongly contextual, as by the pigeonhole principle there exists no assignment $g \in \mathcal{E}(X)$ that assigns different outcomes in $\{0, 1\}$ to any pair of two mea-

surements. Perhaps intriguingly, with or without quantum resources, it is impossible to construct such a simple measurement scenario in reality without communication between the measurement sites.

It is worth noting that the contextuality is a computable property of an empirical model. In the probabilistic case, [4] describes a linear programming method to decide the contextuality of an probabilistic empirical model and a method to express deciding the contextuality of a possibilistic empirical model as a constraint satisfaction problem.

2.3 Quantum Scenarios

The most general way in which we can model a quantum measurement system is the following: we consider a quantum state described by some density operator ρ over a Hilbert space $\mathcal{H}$ and a set X of observables over $\mathcal{H}$ corresponding to our measurements. We then define the set of outcomes $\mathcal{O}$ as:

$$\mathcal{O} := \{\lambda \in \mathbb{R} \mid \lambda \text{ is an eigenvalue of some } A \in X\}.$$

This data defines a measurement scenario $(X, \mathcal{M}, \mathcal{O})$, where

$$\mathcal{M} := \{C \subseteq X \mid \text{for any } A, B \in C,\ AB = BA\}$$

and a probabilistic empirical model e where for any $C \in \mathcal{M}$ and outcome $s : C \to \mathcal{O}$, the probability $e_C(s)$ is given by:

$$e_C(s) := \mathrm{Tr}\left[\left(\prod_{A \in C} E_A^{\lambda}\right)^{\dagger} \rho \left(\prod_{A \in C} E_A^{\lambda}\right)\right]$$

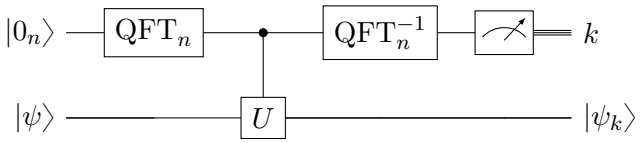
where E_A^λ is the projector onto the λ-eigenspace of the observable A if λ is an eigenvalue of A and the zero map if not. Note that the products in the definition of e_C is well defined because the projectors involved all commute with each other. Given a set of observables X over $\mathcal{H}$ as above, we write $\mathcal{S}_X$ for its corresponding measurement scenario and given a density matrix ρ over $\mathcal{H}$, we write $e^{\rho,X}$ for its corresponding empirical model, or just e^ρ when the set of measurements is clear from context. Note that the empirical model from Example 2.6 is precisely the one obtained by the described procedure above (starting instead from a pure state rather than a density matrix).

Remark 2.10 In a slight deviation from orthodoxy, throughout this paper we will consider observables to be unitary rather than self-adjoint operators over a Hilbert space. We do so as follows: let $\mathcal{H}$ be some Hilbert space and U be a unitary operator over $\mathcal{H}$; the eigenspaces of U form an orthogonal decomposition of $\mathcal{H}$ corresponding to the measurement outcomes. Formally, writing $\mathcal{H}_1, \ldots, \mathcal{H}_n$ for the eigenspaces of U (so that for any $k \in [n]$, there is some $\lambda_k \in \mathbb{C}$ such that $|\psi\rangle \in \mathcal{H}_k$ if and only if $U|\psi\rangle = \lambda_k |\psi\rangle$), we have that $\mathcal{H} = \bigoplus_{k=1}^n \mathcal{H}_k$; and so, given an arbitrary state $|\psi\rangle \in \mathcal{H}$, we may write $|\psi\rangle = \sum_{k=1}^n |\psi_k\rangle$ for some uniquely determined vectors $\{|\psi_k\rangle \in \mathcal{H}_k\}_{k\in[n]}$. By performing the measurement corresponding to U, we obtain the outcome λ_k with probability $\langle\psi_k|\psi_k\rangle$, after which the state of the system becomes:

$$\frac{|\psi_k\rangle}{\| |\psi_k\rangle \|}.$$

It is important to note that unitary observables obey similar rules to self-adjoint ones, in particular, the observables corresponding to the unitaries U and V over $\mathcal{H}$ are compatible if and only if U and V commute.

More concretely, in the context of quantum computing in the circuit model, performing a measurement corresponding to a unitary U on a system in state $|\psi\rangle$, where the eigenvalues of U are all of the form ω_n^k for some values of k, can be performed via phase-kickback, as depicted below:



where the first wire contains an n-qdit initialized in the $|0_n\rangle$ state, the controlled gate corresponds to the unitary

$$\text{c-}U = \sum_{k=0}^{n-1} |k\rangle\langle k| \otimes U^k$$

and the $|\psi_k\rangle$ state is the normalized projection of $|\psi\rangle$ onto the ω_n^k-eigenspace of U.

Remark 2.11 A very useful property of quantum scenario is that the set of contexts is given by the clique complex of the commutation graph; that is, given a set of observables $\mathcal{O}$, we obtain a graph $G = (\mathcal{O}, E)$ where for any distinct $A, B \in \mathcal{O}$, (A, B) forms an edge if and only if $AB = BA$, this is a symmetric relation, so G is an undirected graph. Therefore, a subset of observables $C \subseteq \mathcal{O}$ forms a measurable context in the measurement scenario associated to $\mathcal{O}$ if the induced graph of C in G is a clique. This explains why the possibilistic model in Example 2.9 is not quantum-realisable, as if the induced graphs of $\{M_1, M_2\}, \{M_2, M_3\}$ and $\{M_3, M_1\}$ are all cliques, then $\{M_1, M_2, M_3\}$ must induce a clique as well and thus correspond to a measurement context, thus implying that if the empirical model was quantum realisable, then it could not be contextual.

A remarkable feature of quantum mechanics is that in a quantum scenario, given a set of observables X over some Hilbert space $\mathcal{H}$, contextuality can arise independently of of the state of the system. Formally, that is:

Definition 2.12 Given a set of observables X over some Hilbert space $\mathcal{H}$, we say that X is **state-independently contextual** if for any density matrix ρ over $\mathcal{H}$, we have that e^ρ is contextual.

A ubiquitous source of examples of state-independent contextuality is the following:

Definition 2.13 Given a set of unitary observables X over some Hilbert space $\mathcal{H}$, we say that $o : X \to \mathcal{O}$ is a **non-contextual assignment**, where $\mathcal{O}$ is the set of eigenvalues of the observables in X such that $o(A)o(B) = o(AB)$ whenever A and B commute and $AB \in X$ and for any observable $A \in X$, $o(A)$ is an eigenvalue of A. We say that X is **Kochen-Specker contextual** if there exists no non-contextual assignment $o : X \to \mathcal{O}$.

Example 2.14 An often cited example of Kochen-Specker contextuality is that of the Peres-Mermin square [15]. Consider the set of observables $X = \{\sigma_x \otimes I, I \otimes \sigma_x, \sigma_x \otimes \sigma_x, I \otimes \sigma_z, \sigma_z \otimes I, \sigma_z \otimes \sigma_z, \sigma_x \otimes \sigma_z, \sigma_z \otimes \sigma_x, \sigma_y \otimes \sigma_y\}$ arranged as in the table below:

$\sigma_x \otimes I$	$I \otimes \sigma_x$	$\sigma_x \otimes \sigma_x$
$I \otimes \sigma_z$	$\sigma_z \otimes I$	$\sigma_z \otimes \sigma_z$
$\sigma_x \otimes \sigma_z$	$\sigma_z \otimes \sigma_x$	$\sigma_y \otimes \sigma_y$

Note that the product of the entries of any one of the three rows or of any one of the entries of the first two columns is is $I \otimes I$, but that the product of the entries of the third column is $-I \otimes I$. Noting that for any row or column, all of its entries commute, we have that if there exists no assignment $o : X \to \{+1, -1\}$ satisfying the following set of equations, then X is Kochen-Specker contextual.

$$\begin{cases} o(\sigma_x \otimes I) & o(\sigma_I \otimes \sigma_x) & o(\sigma_x \otimes \sigma_x) & = & o(I \otimes I) & = & +1 \\ o(\sigma_z \otimes I) & o(I \otimes \sigma_z) & o(\sigma_z \otimes \sigma_z) & = & o(I \otimes I) & = & +1 \\ o(\sigma_x \otimes \sigma_z) & o(\sigma_z \otimes \sigma_x) & o(\sigma_y \otimes \sigma_y) & = & o(I \otimes I) & = & +1 \\ o(\sigma_x \otimes I) & o(I \otimes \sigma_z) & o(\sigma_x \otimes \sigma_z) & = & o(I \otimes I) & = & +1 \\ o(I \otimes \sigma_x) & o(\sigma_z \otimes I) & o(\sigma_z \otimes \sigma_x) & = & o(I \otimes I) & = & +1 \\ o(\sigma_x \otimes \sigma_x) & o(\sigma_z \otimes \sigma_z) & o(\sigma_y \otimes \sigma_y) & = & o(-I \otimes I) & = & -1 \end{cases}$$

As the multiplicative structure on $\{-1, +1\}$ is isomorphic to the additive group $\mathbb{Z}_2$, we identify the two structures, $\mathcal{O} = (\{-1, +1\}, \cdot) = (\{0, 1\}, +) \simeq \mathbb{Z}_2$ and obtain the following

system of equations over the field $\mathbb{Z}_2$:

$$\begin{cases} o(\sigma_x \otimes I) + o(\sigma_I \otimes \sigma_x) + o(\sigma_x \otimes \sigma_x) = 0 \\ o(\sigma_z \otimes I) + o(I \otimes \sigma_z) + o(\sigma_z \otimes \sigma_z) = 0 \\ o(\sigma_x \otimes \sigma_z) + o(\sigma_z \otimes \sigma_x) + o(\sigma_y \otimes \sigma_y) = 0 \\ o(\sigma_x \otimes I) + o(I \otimes \sigma_z) + o(\sigma_x \otimes \sigma_z) = 0 \\ o(I \otimes \sigma_x) + o(\sigma_z \otimes I) + o(\sigma_z \otimes \sigma_x) = 0 \\ o(\sigma_x \otimes \sigma_x) + o(\sigma_z \otimes \sigma_z) + o(\sigma_y \otimes \sigma_y) = 1 \end{cases}$$

As every variable appears exactly twice in the LHS of the system of equations, we have that the sum of the left hand sides of the equations above is zero, however the sum of the right hand sides of the equations above is one, so we may derive that $0 = 1$ from the system of equations, thus implying that it is unsatisfiable.

Throughout this paper, similarly to Example 1.5, we adopt the convention of identifying eigenvalues of observables which are roots of unity with elements of a cyclic group. For example, given a set of observables X with eigenvalues in the set $\mathcal{O} \subseteq \{\omega_n^k : k \in [0 \dots n-1]\}$, we will identify $\mathcal{O}$ with a subset of $\mathbb{Z}_n$, thus equating the eigenvalue ω_n^k with $k \in \mathbb{Z}_n$ and considering non-contextual assignments of the form $o : X \to \mathbb{Z}_n$ with respect to this identification.

In the sheaf-theoretical framework of contextuality, Kochen-Specker contextuality finds its relevance through the following result:

Theorem 2.15 Let X be a set of observables over some Hilbert space $\mathcal{H}$, such that X is Kochen-Specker contextual; then X is state-independently contextual. Moreover, for any density operator ρ over $\mathcal{H}$, e^ρ is strongly contextual.

Proof. Let $\mathcal{S} = (X, \mathcal{M}, \mathcal{O})$ be the measurement scenario associated to the set of measurements

X. Suppose that for some density operator ρ over $\mathcal{H}$, e^ρ is not strongly contextual; then there must exist a global section $o \in \mathcal{E}(X)$ such that for any $C \in \mathcal{M}$, $o|_C \in \operatorname{supp} e^\rho$. We will show that o need be a non-contextual assignment. First, note that for any measurement $M \in X$, $o|_M \in \operatorname{supp} e_C$, and so it must hold that $o(M)$ is an eigenvalue of M. Second, note that for any two measurements $A, B \in X$ that commute, if $AB \in X$, then $C = \{A, B, AB\}$ need also pairwise commute and thus be in $\mathcal{M}$; thus, we must have that $o_C(A)o_C(B) = o_C(AB)$ and thus $o(A)o(B) = o(AB)$. □

3 Commutation Groups and their Algebra

In the last section, we've seen how one can obtain state-independent contextuality by choosing our observables from the n-Pauli group in the example of the Peres-Mermin square. The goal of this paper is to generalize the algebraic structure and contextuality arguments associated to the Pauli groups by introducing a new structure, namely that of **commutation group**, as presented in [17].

3.1 Two Equivalent Definitions

Definition 3.1 Given a finite set X, which we will to refer to as **the generator set**, we say that a matrix $\mu : X \times X \to \mathbb{Z}_d$ is a **commutator matrix** if it obeys: $\mu(x, y) = -\mu(y, x)$ for any $x, y \in X$ and $\mu(x, x) = 0$ for any $x \in X$. We define the **commutation group** $G(\mu)$ using the following presentation:

$$G(\mu) := \langle X \sqcup Z_d \,|\, \mathcal{R}_\mu \cup \mathcal{R}_o \cup \mathcal{R}_Z \cup \mathcal{R}_c \rangle$$

where:

- Z_d is the set of symbols $Z_d = \{J_x \mid x \in \mathbb{Z}_d\}$ called the **set of scalars** corresponding to the elements of $\mathbb{Z}_d$.
- $\mathcal{R}_\mu := \{xy = J_{\mu(x,y)}yx \,|\, x, y \in X\}$ is a set of commutation relations between the generators in X defined with respect to μ;
- $\mathcal{R}_o := \{x^d = e \,|\, x \in X\}$ is a set of relations that forces all generators in X to have have period d;
- $\mathcal{R}_Z := \{J_x J_y = J_{x+y} \mid x, y \in \mathbb{Z}_d\} \cup \{J_0 = e\}$ is a set of relations that forces the scalars to behave like the elements of $\mathbb{Z}_d$;

- $\mathcal{R}_c := \{J_x y = y J_x \mid x \in \mathbb{Z}_d, y \in X\}$ is a set of relations that ensures that the scalars lie in the centre of the group.

The prototypical example of a commutation group is the generalized Pauli group:

Example 3.2 Consider the following two unitaries over $\mathbb{C}^n$:

$$\sigma_x = \sum_{k=0}^{n-1} |k+1\rangle\langle k| \qquad\qquad \sigma_z = \sum_{k=0}^{n-1} \omega_n^k |k\rangle\langle k|$$

where we identify $|n\rangle = |0\rangle$. The generalized Pauli group $\mathcal{P}_n$ is defined as the group generated by the matrices $\omega_n I_n, \sigma_x$ and σ_z. Setting $X = \{\sigma_x, \sigma_z\}$ and

$$\mu(\sigma_x, \sigma_z) = -\mu(\sigma_z, \sigma_x) = 1 (\in \mathbb{Z}_n),$$

we obtain an isomorphism $\varphi : \mathcal{P}_n \simeq G(\mu)$ mapping:

$$\begin{aligned} \varphi &: \sigma_x \mapsto \sigma_x \\ \varphi &: \sigma_z \mapsto \sigma_z \\ \varphi &: \omega_n I \mapsto J_1 \end{aligned}$$

Lemma 3.3 Let $X = \{x_1, \ldots, x_n\}$ be a set of generators and $\mu : X \times X \to \mathbb{Z}_d$ a commutator matrix. Then every element $g \in G(\mu)$ can be written as $J_\alpha x_1^{e_1} \ldots x_n^{e_n}$ for some $\alpha \in \mathbb{Z}_d$ and $e_1, \ldots, e_n \in [0..d-1]$. We call this form of an element of $G(\mu)$ its **normal form** and we call J_α the **the scalar part** of g.

Proof. As $G(\mu)$ is defined by a group presentation, we may write any element $g \in G(\mu)$ as a word $g = g_1 \ldots g_m$, where $g_k \in X \sqcup \mathbb{Z}_d$ for every $k \in [m]$. By repeatedly using commutation relations in $\mathcal{R}_\mu$ and $\mathcal{R}_c$, we may move all of the x_1 generators to the left of the word, obtaining thus an equivalent word g' (i.e. $g = g'$ in $G(\mu)$). By iterating this process in a manner similar to insertion-sort, we obtain a word $h = h_1 \ldots h_l$, where whenever $h_i = x_a$ $h_j = x_b$ and $a \leqslant b$, it follows that $i \leqslant j$. By then repeatedly applying the relation which implies that scalars lie in the centre of the commutation group, we obtain the desired result. □

Noting that $J_\alpha x_1^{e_1} \dots x_n^{e_n} = J_\beta x_1^{f_1} \dots x_n^{f_m}$ iff $\alpha = \beta$ and $e_1 = f_1, \dots, e_n = f_n$, it follows that:

Corollary 3.4 Given a finite set X with n elements and a commutator matrix $\mu : X \times X \to \mathbb{Z}_d$, the group $G(\mu)$ has d^{n+1} elements.

Lemma 3.5 Let $X = \{x_1, \dots, x_n\}$ be a finite set and $\mu : X \times X \to \mathbb{Z}_d$ be a commutation matrix. Let $a = J_\alpha x_1^{e_1} \dots x_n^{e_n}$ and $b = J_\beta x_1^{f_1} \dots x_n^{f_n}$ be two elements of $G(\mu)$ in normal form. Then $ab = J_\gamma x_1^{e_1+f_1} \dots x_n^{e_n+f_n}$, where

$$\gamma = \alpha + \beta + \sum_{1 \leqslant i < j \leqslant n} e_i f_j \mu(x_i, x_j).$$

Proof. Similarly to Lemma 3.3, we start from $ab = J_{\alpha+\beta} x_1^{e_1} \dots x_n^{e_n} x_1^{f_1} \dots x_n^{f_n}$. At the first step, we move all f_1 x_1 terms to the left, obtaining a scalar factor of $\sigma_1 = \sum_{i=2}^n e_i f_1 \mu(x_i, x_j)$ from the commutation relations, getting $ab = J_{\alpha+\beta+\sigma_1} x_1^{e_1+f_1} \dots x_n^{e_n} x_2^{f_2} \dots x_n^{f_n}$. Similarly, at the second step, we move all x_2 terms to the left, obtaining a scalar factor of $\sigma_2 = \sum_{i=3}^n e_i f_2 \mu(x_i, x_j)$ from the commutation relations, getting $ab = J_{\alpha+\beta+\sigma_1+\sigma_2} x_1^{e_1+f_1} x_2^{e_2+f_2} \dots x_n^{e_n} x_3^{f_3} \dots x_n^{f_n}$. By iterating this process, we obtain $ab = J_{\alpha+\beta+\sigma_1+\dots+\sigma_n} x_1^{e_1+f_1} \dots x_n^{e_n+f_n}$. As $\sigma_k = \sum_{i=k+1}^n e_i f_k \mu(x_i, x_k)$, it follows that:

$$\alpha + \beta + \sum_{i=1}^n \sigma_i = \alpha + \beta + \sum_{1 \leqslant i < j \leqslant n} e_i f_j \mu(x_i, x_j) = \gamma.$$

□

While the definition of commutation groups involving group presentations elegantly captures the motivation behind them, in practice, it is more convenient to work with an alternative linear-algebraic definition when studying the algebraic properties of these groups:

Definition 3.6 Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, we define its commutation group $H(\mu)$ as the group with carrier set $\mathbb{Z}_d \times \mathbb{Z}_d^n$ and operation given by:

$$(k_a, \vec{a}) \cdot (k_b, \vec{b}) = (k_a + k_b + \tilde{\mu}(\vec{a}, \vec{b}), \vec{a} + \vec{b}).$$

Given an arbitrary element $g = (k, \vec{v})$, we call k the **scalar component** of g and $\vec{v}$ the **vector component** of g. Furthermore, if $\vec{v} = 0$, we call g a **scalar element** of $H(\mu)$.

Theorem 3.7 Given a finite set $X = \{x_1, \ldots, x_n\}$, and a commutator matrix $\mu : X \times X \to \mathbb{Z}_d$. Committing the mild sin of overloading the symbol μ, we define the matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ as the one given by $\mu_{ij} = \mu(x_i, x_j)$; we have that $G(\mu) \simeq H(\mu)$.

Proof. Let $\vec{e_1}, \ldots, \vec{e_n}$ be the canonical basis of $\mathbb{Z}_d^n$, so that $\vec{e_i}^\intercal \mu \vec{e_j} = \mu_{ij}$. Consider the mapping $\varphi : H(\mu) \xrightarrow{\sim} G(\mu)$, given by:

$$\varphi : \left(k, \sum_i \alpha_i \vec{e_i}\right) \longmapsto J_k x_1^{\alpha_1} \ldots x_n^{\alpha_n}.$$

The injectivity of φ is visible and its surjectivity follows from an immediate counting argument using Corollary 3.4. Finally, we show that φ satisfies the homomorphism condition: let $a = (k_a, \sum_i \alpha_i \vec{e_i})$ and $b = (k_b, \sum_i \beta_i \vec{e_i})$ be two arbitrary elements of $H(\mu)$. We have that:

$$\begin{aligned}
\varphi(a)\varphi(b) &= J_{k_a} x_1^{\alpha_1} \ldots x_n^{\alpha_n} J_{k_b} x_1^{\beta_1} \ldots x_n^{\beta_n} \\
&= J_\gamma x_1^{\alpha_1 + \beta_1} \ldots x_n^{\alpha_n + \beta_n} \\
&= \varphi(\gamma, \sum_i (\alpha_i + \beta_i) \vec{e_i}) \\
&= \varphi(ab),
\end{aligned}$$

where, as in Lemma 3.3,

$$\gamma = k_a + k_b + \sum_{1 \leqslant i < j \leqslant n} \alpha_i \beta_j \mu_{ij} = k_a + k_b + \left(\sum_i \alpha_i \vec{e_i}\right)^\intercal \tilde{\mu} \left(\sum_i \beta_i \vec{e_i}\right).$$

□

The linear-algebraic definition of commutation groups has a couple of nice direct consequences:

Lemma 3.8 Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and any element $(k, \vec{v}) \in H(\mu)$, its

inverse is $(-k - \tilde{\mu}(\vec{v}, \vec{v}), -\vec{v})$.

Lemma 3.9 Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and any two elements $(k_a, \vec{a}), (k_b, \vec{b})$, we have that $[(k_a, \vec{a}), (k_b, \vec{b})] = (\mu(\vec{a}, \vec{b}), \vec{0})$, where $[-, -]$ is the group commutator given by $[g, h] = ghg^{-1}h^{-1}$. Moreover, the commutator of two group elements is always a scalar element of the group.

Proof.

$$\begin{aligned}
[(k_a, \vec{a}), (k_b, \vec{b})] &= (k_a, \vec{a})(k_b, \vec{b})(-k_a - \tilde{\mu}(\vec{a}, \vec{a}), -\vec{a})(-k_b - \tilde{\mu}(\vec{b}, \vec{b}), -\vec{b}) \\
&= (k_a + k_b, \vec{a} + \vec{b})(-k_a - k_b - \tilde{\mu}(\vec{a}, \vec{a}) - \tilde{\mu}(\vec{b}, \vec{b}), -\vec{a} - \vec{b}) \\
&= (\vec{a}^\intercal \tilde{\mu} \vec{b} - \vec{b}^\intercal \tilde{\mu} \vec{a}, 0) \\
&= (\mu(\vec{a}, \vec{b}), 0)
\end{aligned}$$

□

Corollary 3.10 Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and an arbitrary element $g = (k, \vec{v}) \in H(\mu)$, g lies in the centre of $H(\mu)$ if and only if $\mu\vec{v} = 0$.

3.2 Classification of Commutation Groups

When considering any family of algebraic structures, a useful step in understanding it is finding an isomorphism criterion. The goal of this subsection is precisely finding such a necessary-sufficient criterion. To do so, we define the following categories and functors:

Definition 3.11 Given an integer $d \in \mathbb{N}_{\geqslant 2}$, we define the **category of commutation groups** $\mathcal{G}_d$ as follows:

- The objects of $\mathcal{G}_d$ are commutation groups $H(\mu)$ for $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ for $n \in \mathbb{N}$.

- Given two objects $H(\mu_0)$, $H(\mu_1)$ in $\mathcal{G}_d$, $\mathrm{Hom}(H(\mu_0), H(\mu_1))$ is the set of group homomorphisms $\phi : H(\mu_0) \to H(\mu_1)$ that preserve scalar elements, i.e. that for any element $(k, 0) \in H(\mu_0)$, $\phi(k, 0) = (k', 0)$ for some $k' \in \mathbb{Z}_d$.

Definition 3.12 Given an integer $d \in \mathbb{N}_{\geqslant 2}$, we define the **category of commutation modules** $\mathcal{V}_d$ as follows:

- The objects of $\mathcal{V}_d$ are pairs $(\mathbb{Z}_d^n, \mu)$ of a finite-dimensional free $\mathbb{Z}_d$-module $\mathbb{Z}_d^n$ and a matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$. We will write $V(\mu)$ for the object $(\mathbb{Z}_d^n, \mu)$.
- Given two commutation modules $V(\mu_0), V(\mu_1)$ for some $\mu_0 \in \mathfrak{so}(n, \mathbb{Z}_d)$ and $\mu \in \mathfrak{so}(m, \mathbb{Z}_d)$, $\mathrm{Hom}(V(\mu_0), V(\mu_1))$ consists of the $\mathbb{Z}_d$-linear maps $\phi : \mathbb{Z}_d^n \to \mathbb{Z}_d^m$ that preserve the values of the bilinear forms - i.e.

$$\forall \vec{v}, \vec{w} \in \mathbb{Z}_d^n \,.\mu_0(\vec{v}, \vec{w}) = \mu_1(\phi(\vec{v}), \phi(\vec{w}))$$

Definition 3.13 Given an integer $d \in \mathbb{N}_{\geqslant 2}$, we define the functor $G_d : \mathcal{V}_d \to \mathcal{G}_d$ as:

- Any object $V(\mu) = (\mathbb{Z}_d^n, \mu)$ is mapped to $G_d V(\mu) := H(\mu)$;
- Given any morphism $\phi : V(\mu_0) \to V(\mu_1)$ for some $\mu_0 \in \mathfrak{so}(n, \mathbb{Z}_d)$ and $\mu_1 \in \mathfrak{so}(m, \mathbb{Z}_d)$, the morphism $G_d\phi : H(\mu_0) \to H(\mu_1)$ is defined as follows:

$$(G_d\phi) : (k, \vec{v}) \longmapsto (k, \phi(\vec{v})).$$

The functoriality of G_d is visible, we do however need to show that $G_d\phi : H(\mu_0) \to H(\mu_1)$ is indeed a morphism in $\mathcal{G}_d$. Consider any two elements $(a, \vec{v}), (b, \vec{w}) \in H(\mu_0)$. We have that:

$$\begin{aligned}((G_d\phi)(a, \vec{v})) \cdot ((G_d\phi)(b, \vec{w})) &= (a, \phi(\vec{v})) \cdot (b, \phi(\vec{w})) \\ &= (a + b + \mu(\phi(\vec{v}), \phi(\vec{w})), \phi(\vec{v} + \vec{w})) \\ &= (a + b + \mu(\vec{v}, \vec{w}), \phi(\vec{v}) + \phi(\vec{w}))\end{aligned}$$

$$= (G_d\phi)((a, \vec{v}) \cdot (b, \vec{w})).$$

Definition 3.14 Given an integer $d \in \mathbb{N}_{\geqslant 2}$, we define the functor $V_d : \mathcal{G}_d \to \mathcal{V}_d$ where:

- Any object $H(\mu)$ in $\mathcal{G}_d$, where $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ is mapped to $V_d G(\mu) := (\mathbb{Z}_n, \mu)$;
- Given any morphism $\phi : H(\mu_0) \to H(\mu_1)$ for some $\mu_0 \in \mathfrak{so}(n, \mathbb{Z}_d)$ and $\mu_1 \in \mathfrak{so}(m, \mathbb{Z}_d)$, writing $\vec{e_1}, \ldots, \vec{e_n}$ for the canonical basis of $\mathbb{Z}_d^n$, we define for all $i \in [n]$, the vector $\vec{f_i} \in \mathbb{Z}_d^m$ as the one given by the identity $\phi(0, \vec{e_i}) = (k, \vec{f_i})$ for some $k \in \mathbb{Z}_d$. Finally, we define $(V_d\phi) : V(\mu_0) \to V(\mu_1)$ to be given by:

$$V_d\phi : \sum_{i=1}^{n} \alpha_i \vec{e_i} \longmapsto \sum_{i=1}^{n} \alpha_i \vec{f_i}$$

We may now state our isomorphism criterion:

Theorem 3.15 Given commutator matrices $\mu_0, \mu_1 \in \mathfrak{so}(n, \mathbb{Z}_d)$, $H(\mu_1)$ and $H(\mu_2)$ are isomorphic if and only if there exists an invertible $n \times n$ matrix U over $\mathbb{Z}_d^n$ such that $U^\intercal \mu_1 U = \mu_2$ (i.e. if μ_1 and μ_2 are cogredient matrices).

Proof. We first prove the ($\implies$) direction: let $\phi : H(\mu_0) \xrightarrow{\sim} H(\mu_1)$ be an isomorphism. We obtain a linear map $V_d\phi : \mathbb{Z}_d^n \to \mathbb{Z}_d^n$, which, by the functoriality of V_d, must also be an isomorphism. Writing U for the matrix corresponding to $V_d\phi$ in the canonical basis, we have that U need be invertible. Moreover, as $V_d\phi$ is a map in $\mathcal{V}_d$, it follows that for any two vectors $\vec{v}, \vec{w} \in \mathbb{Z}_d$,

$$\begin{aligned}
\mu_0(\vec{v}, \vec{w}) &= \mu_1((V_d\phi)(\vec{v}), (V_d\phi)(\vec{w})) \\
&= (U\vec{v})^\intercal \mu_1 (U\vec{w}) \\
&= \vec{v}^\intercal U^\intercal \mu_1 U \vec{w}.
\end{aligned}$$

As this equality holds independently of the choice of $\vec{v}$ and $\vec{w}$, we have that $U^\intercal \mu_0 U = \mu_1$.

The ($\Longleftarrow$) direction is proved in a similar manner: an invertible $n \times n$ matrix U such that $U^{\intercal}\mu_0 U = \mu_1$ corresponds to a linear isomorphism $U : V(\mu_0) \xrightarrow{\sim} V(\mu_1)$, which by the functoriality of G_d, corresponds to an isomorphism $G_d U : H(\mu_0) \xrightarrow{\sim} H(\mu_1)$. $\square$

While the above stated isomorphism criterion is necessary-sufficient, it does not provide any manner of actually deciding whether such a matrix U exists. To that end, the following Theorem in conjunction with the methods of computing the basis change in Appendix A provide a method to canonically find a normal form of a matrix μ in the equivalence class suggested by Theorem 3.15, namely:

$$\{U^{\intercal}\mu U \mid \ U \text{ invertible } \}.$$

Theorem 3.16 Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ be a commutator matrix. There is a basis of $V = \mathbb{Z}_d^n$:

$$\mathcal{F} = \{\vec{a}_1, \ldots, \vec{a}_m, \vec{b}_1, \ldots, \vec{b}_m, \vec{c}_1, \ldots, \vec{c}_l\}$$

such that for any $i, j \in [m]$ and $k, t \in [l]$

$$\begin{aligned} \mu(\vec{a}_i, \vec{b}_j) &= \delta_{ij}\Delta_i \\ \mu(\vec{a}_i, \vec{a}_j) = \mu(\vec{b}_i, \vec{b}_j) &= \mu(\vec{c}_t, \vec{c}_k) = 0 \\ \mu(\vec{a}_i, \vec{c}_k) &= 0 \\ \mu(\vec{b}_i, \vec{c}_k) &= 0 \end{aligned}$$

for some $\Delta_1, \ldots, \Delta_m \in \mathbb{Z}_d$.

Remark 3.17 The matrix obtained from μ by change of basis from the canonical basis to $\mathcal{F}$ in Theorem 3.16 is known as its **Darboux normal form**.

We defer the proof of Theorem 3.16 to Appendix A, the proof is constructive and defines an

efficient algorithm that computes the Darboux normal form of a commutator matrix, whose code can be found in Appendix C.

3.3 Representations of Commutation Groups

We describe three ways of obtaining representations of commutation groups, the image of two of them lying in the Pauli group. The first such representation is the following:

Definition 3.18 Given a commutation matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, the **regular representation** (introduced in [17]) $\rho : H(\mu) \to \mathrm{GL}(\mathcal{Q}_d^{\otimes n})$ is given by:

$$\rho(k, \vec{v}) = \sum_{\vec{u} \in \mathbb{Z}_d^n} \omega_d^{k+\tilde{\mu}(\vec{v},\vec{u})} |\vec{u} + \vec{v}\rangle\langle\vec{u}|$$

where $\{|\vec{v}\rangle \mid \vec{v} \in \mathbb{Z}_d^n\}$ is the computational basis on $\mathcal{Q}_d^{\otimes n}$, so that for example $|(0,2)\rangle \in \mathcal{Q}_3^2$ denotes an unentangled state over 2 (3-dimensional) qudits, the first of which is in state $|0\rangle$ and the second one in state $|2\rangle$.

Theorem 3.19 Given a commutation matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, the regular representation ρ of $G(\mu)$ is a well-defined faithful scalar-preserving representation [17].

Proof. It is clear from the definition that ρ is injective and that it maps $(k, \vec{0})$ to $\omega_d^k I_d$, so we only need to check the homomorphism condition: let $(a, \vec{v}), (b, \vec{w})$ denote two arbitrary elements of $H(\mu)$; we have that:

$$\begin{aligned}
\rho(a, \vec{v})\rho(b, \vec{w}) &= \left(\sum_{\vec{u} \in \mathbb{Z}_d^n} \omega_d^{a+\tilde{\mu}(\vec{v},\vec{u})} |\vec{u} + \vec{v}\rangle\langle\vec{u}| \right) \left(\sum_{\vec{u} \in \mathbb{Z}_d^n} \omega_d^{b+\tilde{\mu}(\vec{w},\vec{u})} |\vec{u} + \vec{w}\rangle\langle\vec{u}| \right) \\
&= \sum_{\vec{u} \in \mathbb{Z}_d^n} \omega_d^{a+b+\tilde{\mu}(\vec{v},\vec{u}+\vec{w})+\tilde{\mu}(\vec{w},\vec{u})} |\vec{v} + \vec{w} + \vec{u}\rangle\langle\vec{u}| \\
&= \sum_{\vec{u} \in \mathbb{Z}_d^n} \omega_d^{a+b+\tilde{\mu}(\vec{v},\vec{w})+\tilde{\mu}(\vec{w}+\vec{v},\vec{u})} |\vec{v} + \vec{w} + \vec{u}\rangle\langle\vec{u}|
\end{aligned}$$

$$= \rho(a + b + \tilde{\mu}(\vec{v}, \vec{w}), \vec{v} + \vec{w})$$

□

We will denote the generalized d-Pauli matrices as X_d and Z_d, so that for the canonical basis $\{|k\rangle \mid k \in [0..n-1]\}$ of the state space $\mathcal{Q}_d$ of a qudit, they are given by:

$$X_d = \sum_{k=0}^{n-1} |k+1\rangle\langle k| \qquad\qquad Z_d = \sum_{k=0}^{n-1} \omega_d^k |k\rangle\langle k|$$

where we identify $|n\rangle = |0\rangle$. We define the **canonical Pauli representation** of a commutation group as follows:

Definition 3.20 Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ be an arbitrary commutation matrix. We have that $\pi : H(\mu) \to \mathrm{GL}(\mathcal{Q}_d^{\frac{n(n-1)}{2}})$, as given below is a faithful representation of $H(\mu)$:

$$\pi(k, \vec{v}) := \omega_d^k \bigotimes_{a=1}^{n} \bigotimes_{b=a+1}^{n} X_d^{v_a} Z_d^{v_b \mu_{ab}}.$$

where $(k, \vec{v})$ is an arbitrary element of $H(\mu)$ and $\vec{v} = (v_1, \ldots, v_n) \in \mathbb{Z}_d^n$.

Lemma 3.21 For all $d \in \mathbb{N}_{\geqslant 2}$ and $k \in \mathbb{N}_{\geqslant 2}$, $Z_d^m X_d^n = \omega_d^{mn} X_d^n Z_d^m$.

Theorem 3.22 The mapping π as described in Definition 3.8 is a faithful representation of the group $H(\mu)$.

Proof. The injectivity of π is visible. We need only show that π does indeed satisfy the homomorphism criterion: let $(a, \vec{v}), (b, \vec{w})$ denote two arbitrary elements of $H(\mu)$, where $\vec{v} = (v_1, \ldots, v_n)$ and $\vec{w} = (w_1, \ldots, w_n)$. We have that:

$$\pi(k_v, \vec{v})\pi(k_w, \vec{w}) = \left(\omega_d^{k_v} \bigotimes_{a=1}^{n} \bigotimes_{b=a+1}^{n} X_d^{v_a} Z_d^{v_b \mu_{ab}}\right) \left(\omega_d^{k_w} \bigotimes_{a=1}^{n} \bigotimes_{b=a+1}^{n} X_d^{w_a} Z_d^{w_b \mu_{ab}}\right)$$

$$= \omega_d^{k_v+k_w} \bigotimes_{a=1}^{n} \bigotimes_{b=a+1}^{n} \omega_d^{v_a\mu_{ab}w_b} X_d^{v_a+w_a} Z_d^{(v_b+w_b)\mu_{ab}}$$
$$= \omega_d^{k_v+k_w+\tilde{\mu}(\vec{v},\vec{w})} \bigotimes_{a=1}^{n} \bigotimes_{b=a+1}^{n} X_d^{v_a+w_a} Z_d^{(v_b+w_b)\mu_{ab}}$$
$$= \pi((k_v, \vec{v}) \cdot (k_w, \vec{w})).$$

□

From the above, it immediately follows that:

Theorem 3.23 Given a commutation group $H(\mu)$ for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_n)$, $H(\mu)$ is isomorphic to a subgroup of the generalized Pauli group $\mathcal{P}_d^n$.

While the canonical Pauli representation of a commutation group suffices in proving Theorem 3.23, it is hard not to note that it acts on a huge $\frac{dn(n-1)}{2}$-dimensional complex vector space. By using Theorem 3.16, we may obtain a significantly more "compact" dn-dimensional representation for a commutation group, whose properties can actually be used to study the algebraic properties of a commutation group (for example partially computing the character table of $G(\mu)$). The rest of this section is dedicated to constructing this representation.

Let $\mu_0 \in \mathfrak{so}(n, \mathbb{Z}_d)$ denote an arbitrary commutator matrix. By Theorem 3.16, there exists an invertible matrix $U \in \mathrm{GL}(\mathbb{Z}_d^n)$ such that $U^\intercal \mu_0 U = \mu$, where μ is in Darboux normal form, and an isomorphism $\varphi : H(\mu_0) \xrightarrow{\sim} H(\mu)$. Let

$$\mathcal{F} = \{\vec{a}_1, \ldots, \vec{a}_m, \vec{b}_1, \ldots, \vec{b}_m, \vec{c}_1, \ldots, \vec{c}_l\}$$

be a basis of $\mathbb{Z}_d^n$ and $\Delta_1, \ldots, \Delta_m \in \mathbb{Z}_d$ exactly as in the statement of Theorem 3.16. The following is a faithful representation $\pi : H(\mu) \to \mathcal{P}_d^m$, which when composed with φ results in a faithful dm-dimensional representation of $H(\mu_0)$ whose image lies in the generalized Pauli

group:

$$\pi : (k, \sum_{j=1}^{m} \alpha_j \vec{a}_j + \sum_{j=1}^{m} \beta \vec{b}_j + \sum_{j=1}^{l} \gamma_j \vec{c}_j) \longmapsto \omega_d^k \bigotimes_{j=1}^{m} X_d^{\beta_j} Z_d^{\Delta_j \alpha_j}.$$

where $(k, \sum_{j=1}^{m} \alpha_j \vec{a}_j + \sum_{j=1}^{m} \beta \vec{b}_j + \sum_{j=1}^{l} \gamma_j \vec{c}_j)$ denotes an arbitrary element of $H(\mu)$.

4 The Contextuality of Commutation Groups

Having laid out the algebraic properties of commutation groups and the foundations of contextuality, we may now explore how contextuality arises from commutation groups. It turns out that there are two radically different but equivalent methods of producing proofs of contextuality, which we will present separately in subsections 4.2 and 4.3.

4.1 Contextuality of Commutation Groups

To finally rigorously define contextuality, we must first define the mathematical gadget of compatible monoids, which were introduced in [2] with different notation and formalized in the form presented here in [17]. We note that while compatible monoids are very similar to commutative partial monoids which also occur in the literature on contextuality [18], the associativity axiom (rule (iv) in Definition 4.1) is different.

Definition 4.1 A **compatible monoid** is a tuple $(M, \odot, \cdot, 1)$ where M is the carrier set, $\odot$ is the **compatibility relation** ($\odot \subseteq M \times M$) and $\cdot : \odot \to M$ is a partial binary operation called the **product operation**, and 1 is an element of M which we will call **the unit** of the compatible monoid. A compatible monoid needs to satisfy the following conditions:

i. $\odot$ is a symmetric and reflexive (but not necessarily transitive) binary relation over M;

ii. $1 \odot x$ for any $x \in M$;

iii. For any $x, y \in M$ such that $x \odot y$, $x \cdot y = y \cdot x$;

iv. For any $x, y, z \in M$, if $x \odot y$, $y \odot z$ and $z \odot x$, then $x \odot (y \cdot z)$, $y \odot (z \cdot x)$, $z \odot (x \cdot y)$ and $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

Essentially, a compatible monoid models the partial operation resulting from forgetting that

one can multiply non-commuting elements of a monoid; thus, for any monoid $(M, \cdot, 1)$, we may obtain an associated compatible monoid where $x \odot y$ for $x, y \in M$ if and only if x and y commute in M.

Definition 4.2 Let $\mathcal{M} = (M, \odot, \cdot, 1)$ and $\mathcal{N} = (N, \odot, \cdot, 1)$ be two compatible monoids. A homomorphism from $\mathcal{M}$ to $\mathcal{N}$ is a function $f : M \to N$ obeying that $f(1) = 1$ and that for any $x, y \in M$ such that $x \odot y$, $f(x) \odot f(y)$ and $f(x \cdot y) = f(x) \cdot f(y)$.

Remark 4.3 The raison de d'Être of compatible monoids is that their homomorphisms model the fundamental notion present in the definition of Kochen-Specker contextuality, namely, that of *maps which preserve commuting products*:

Consider a set of unitary observables X over an n-dimensional Hilbert space and the set $\mathcal{O} = \{\omega_n^k \mid k \in [0..n-1]\}$ and assume that X is closed under commuting products $(\forall A, B \in X \,.\, AB = BA \implies AB \in X)$ and that $\{\omega_n^k I_n \mid k \in [0..n-1]\} \subseteq X$. Due to their unitarity, we have for all observables in X, their eigenvalues must be elements of $\mathcal{O}$. Writing $C(X)$ and $C(\mathcal{O})$ for the compatible monoids corresponding to X and $\mathcal{O}$, and considering the inclusion mapping:

$$\iota : C(\mathcal{O}) \hookrightarrow C(X) \qquad\qquad \iota : \omega_n^k \longmapsto \omega_n^k I_n$$

we have that a non-contextual assignment for X is a map $o : C(X) \to C(\mathcal{O})$ such that $o \circ \iota = 1_{\mathcal{O}}$, i.e. o is a splitting of the inclusion ι.

We may now define what it means for a commutation group to be contextual:

Definition 4.4 Let G be a group and $X \subseteq G$ be a subset of elements of X. We define the **commutative closure of X in G** — $\langle X \rangle_C$ as the smallest subset of G which includes X such that for any $g, h \in \langle X \rangle_C$ such that g and h commute, $gh \in \langle X \rangle_C$ as well.

Definition 4.5 Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ be a commutation matrix. We define the **generator set**

X_μ **of** $H(\mu)$ as:

$$X_\mu := \{(0, \vec{e_1}), \dots (0, \vec{e_n})\}$$

(where $\{\vec{e_1}, \dots, \vec{e_n}\}$ is the canonical basis of $\mathbb{Z}_d^n$) and we define $C(\mu)$ as the compatible monoid with carrier set given by the commutative closure of the generator set and scalars in $H(\mu)$, i.e. $\langle X_\mu \cup \mathbb{Z}_d \rangle_C$.

Definition 4.6 Given a commutation group $H(\mu)$, we define a **non-contextual assignment for** $H(\mu)$ to be a compatible monoid morphism $o : C(\mu) \to \mathbb{Z}_d$ such that $o \circ \iota = 1_{\mathbb{Z}_d}$, where $\iota : \mathbb{Z}_d \to C(\mu)$ is the inclusion of scalars in $C(\mu)$. If no such non-contextual assignment exists, we say that $H(\mu)$ is **algebraically contextual**.

The following theorem serves as motivation for the above definition:

Theorem 4.7 Given a commutation group $H(\mu)$ for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, if $H(\mu)$ is algebraically contextual, then any set of observables resulting from a representation of the commutative closure of the generating set and scalars of $H(\mu)$ need be Kochen-Specker contextual. In other words, for any scalar-preserving representation $\rho : C(\mu) \to \mathrm{GL}(\mathcal{H})$ (where $\mathcal{H}$ is some Hilbert space and $\mathrm{GL}(\mathcal{H})$ denotes the compatible monoid associated to the general linear group over $\mathcal{H}$), the set of observables $\{\rho(A) : A \in X_\mu\}$ is state-independently contextual.

Proof. Let $\mathcal{H} = \mathbb{C}^m$ for some $m \in \mathbb{N}_{\geqslant 1}$. By scalar-preserving compatible monoid representation, we mean that $\rho : C(\mu) \to \mathrm{GL}(\mathcal{H})$ is a map from the commutative closure of the union of the set of scalars in $H(\mu)$ and the set of generators X_μ to invertible matrices over $\mathcal{H}$, such that for any two *commuting* elements $(k_v, \vec{v}), (k_u, \vec{u}) \in C(\mu)$, $\rho(k_v, \vec{v}) \cdot \rho(k_u, \vec{u}) = \rho(k_u + k_v + \vec{v}^\intercal \tilde{\mu} \vec{u})$ and for any scalar element $(k, \vec{0})$, $\rho(k, \vec{0}) = \omega_d^k I_m$.

To prove the above theorem, it suffices to show that for any such representation ρ and any element $g \in C(\mu)$, the eigenvalues of $\rho(g)$ are included in $\{\omega_d^0, \omega_d^1, \dots, \omega_d^{d-1}\}$. We do so using

a structural induction argument over $C(\mu)$:

First, note that for any unitary matrix U, if $U^m = I$, then the eigenvalues of U are included in the set $\{\omega_d^0, \ldots, \omega_d^{d-1}\}$. Moreover, it holds for any element g of $X_\mu \cup \{(k, \vec{0}) \mid k \in [d]\}$ that $g^d = (0, \vec{0})$, so $\rho(g)^d = I_m$. As the carrier set of $C(\mu)$ is defined as the commutative closure of $X_\mu \cup \{(k, \vec{0}) \,|\, k \in [d]\}$, it suffices to then prove (as the inductive step) that whenever $g, h \in C(\mu)$ commute and $g^d = h^d = (0, \vec{0})$, then $(gh)^d = (0, \vec{0})$ (and thus $\rho(gh)^d = I_m$ and so the eigenvalues of $\rho(gh)$ lie in $\{\omega_d^0, \ldots, \omega_d^{d-1}\}$). This is obviously true, as $(gh)^d = g^d h^d = (0, \vec{0})$ by the commutativity of g and h. □

4.2 Algebraic Contextuality

This section covers proofs of contextuality for commutation groups which rely on their linear algebraic definition and properties. We give several criteria for contextuality and show that for a commutation group $H(\mu)$ where $\mu \in \mathfrak{so}(n, \mathbb{Z}_\ell)$ with ℓ odd, $H(\mu)$ cannot be algebraically contextual.

Lemma 4.8 Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and $\vec{v} \in \mathbb{Z}_d^n$. If d is odd, then $\vec{v}^\intercal \tilde{\mu} \vec{v} = 0$.

Proof. We have that:

$$\begin{aligned} 2\vec{v}^\intercal \tilde{\mu} \vec{v} &= \vec{v}^\intercal \tilde{\mu} \vec{v} + \vec{v}^\intercal \tilde{\mu} \vec{v} \\ &= \vec{v}^\intercal \tilde{\mu} \vec{v} + (\vec{v}^\intercal \tilde{\mu} \vec{v})^\intercal \\ &= \vec{v}^\intercal \tilde{\mu} \vec{v} + \vec{v}^\intercal \tilde{\mu}^\intercal \vec{v} \\ &= \vec{v}^\intercal \tilde{\mu} \vec{v} = 0 \end{aligned}$$

where the last equality stems from the fact that μ is skew-symmetric. As we are working over

odd characteristic, we may multiply the equality $2\vec{v}^\intercal\tilde{\mu}\vec{v} = 0$ by the multiplicative inverse of 2 on both sides and thus obtain the claim. □

Lemma 4.9 Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and $(k_v, \vec{v}), (k_w, \vec{w}) \in H(\mu)$. If d is odd and $(k_v, \vec{v})$ and $(k_w, \vec{w})$ commute, then $\vec{v}^\intercal\tilde{\mu}\vec{w} = 0$.

Proof. By Lemma 3.21, $(k_v, \vec{v})$ and $(k_w, \vec{w})$ commute iff $\vec{v}^\intercal\mu\vec{w} = 0$. We have that:

$$\begin{aligned}\vec{v}^\intercal\mu\vec{w} &= \vec{v}^\intercal(\tilde{\mu} - \tilde{\mu}^\intercal)\vec{w} \\ &= \vec{v}^\intercal\tilde{\mu}\vec{w} - \vec{v}^\intercal\tilde{\mu}^\intercal\vec{w} \\ &= \vec{v}^\intercal\tilde{\mu}\vec{w} - \vec{w}^\intercal\tilde{\mu}\vec{v}\end{aligned}$$

As by 3.9 the first term of the equality is zero, we deduce that $\vec{v}^\intercal\tilde{\mu}\vec{w} = \vec{w}^\intercal\tilde{\mu}\vec{v}$. It follows that:

$$\begin{aligned}(\vec{v} + \vec{w})^\intercal\tilde{\mu}(\vec{v} + \vec{w}) &= \vec{v}^\intercal\tilde{\mu}\vec{v} + \vec{w}^\intercal\tilde{\mu}\vec{w} + \vec{v}^\intercal\tilde{\mu}\vec{w} + \vec{w}^\intercal\tilde{\mu}\vec{v} \\ &= 2\vec{v}^\intercal\tilde{\mu}\vec{w}\end{aligned}$$

Where the last equality stems from Lemma 4.8 and our prior observation that $\vec{v}^\intercal\tilde{\mu}\vec{w} = \vec{w}^\intercal\tilde{\mu}\vec{v}$. As by Lemma 4.8. the first term of the equality is zero, it follows that $2\vec{v}^\intercal\tilde{\mu}\vec{w} = 0$ and so, as d is odd, $\vec{v}^\intercal\tilde{\mu}\vec{w} = 0$ which is q.e.d. □

We may now show that contextuality cannot arise in odd characteristic:

Theorem 4.10 Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ for some odd d. Then:

$$o : H(\mu) \to \mathbb{Z}_d \qquad\qquad o : (k, \vec{v}) \mapsto k$$

is a valid non-contextual assignment.

Proof. Clearly, $o \circ \iota = 1_{\mathbb{Z}_d}$, so we only need to check that for all commuting pairs $(k_v, \vec{v}), (k_w, \vec{w})$

in $C(\mu)$, $o(k_v, \vec{v}) + o(k_w, \vec{w}) = o((k_v, \vec{v}) \cdot (k_w, \vec{w}))$; this is true, because:

$$\begin{aligned} o((k_v, \vec{v}) \cdot (k_w, \vec{w})) &= o((k_v + k_w + \vec{v}^{\intercal}\tilde{\mu}\vec{w}, \vec{v} + \vec{w}) \\ &= k_v + k_w \\ &= o(k_v, \vec{v}) + o(k_w, \vec{w}) \end{aligned}$$

□

While the above theorem is theoretically significant, it also leads to a very useful consequence in computing whether or not a commutation group is contextual or not:

Lemma 4.11 Let $d = 2^k\ell \in \mathbb{N}_{\geqslant 2}$, where $k \in \mathbb{N}_{\geqslant 0}$ and ℓ is an odd number. By the Chinese remainder theorem (CRT), $\mathbb{Z}_d$ is isomorphic to $\mathbb{Z}_{2^k} \times \mathbb{Z}_\ell$; we write $q_e : \mathbb{Z}_d \to \mathbb{Z}_{2^k}$ and $q_o : \mathbb{Z}_d \to \mathbb{Z}_\ell$ for the projection mappings corresponding to this product. Writing $\mu_e \in \mathfrak{so}(n, \mathbb{Z}_{2^k})$ and $\mu_o \in \mathfrak{so}(n, \mathbb{Z}_\ell)$ for the "μ modulo 2^k and ℓ respectively" (i.e. $(\mu_o)_{ij} := q_o(\mu_{ij})$ and $(\mu_e)_{ij} := q_e(\mu_{ij})$), we have that $H(\mu) \simeq H(\mu_o) \times H(\mu_e)$.

Proof. In a minor abuse of notation, we will also write q_e and q_o for the projection maps from $\mathbb{Z}_d^n$ to $\mathbb{Z}_{2^k}^n$ and to $\mathbb{Z}_\ell^n$, i.e.:

$$q_\ell : (v_1, \ldots v_n) \mapsto (q_\ell(v_1), \ldots, q_\ell(v_n)) \qquad q_{2^k} : (v_1, \ldots v_n) \mapsto (q_{2^k}(v_1), \ldots, q_{2^k}(v_n)).$$

The isomorphism $\varphi : H(\mu) \xrightarrow{\sim} H(\mu_o) \times H(\mu_e)$ is given by:

$$\varphi : (k, \vec{v}) \mapsto ((q_o(k), q_o(\vec{v})), (q_e(k), q_e(\vec{v}))).$$

It is clear from the CRT that this map is bijective, we do however need to show that it is also an isomorphism. Let $(k_v, \vec{v}), (k_u, \vec{u})$ denote two arbitrarily chosen elements of $H(\mu)$. We

have that:

$$\begin{aligned}
&\varphi(k_v, \vec{v}) \cdot \varphi(k_u, \vec{u}) \\
&\quad = ((q_o(k_v), q_o(\vec{v})), (q_e(k_v), q_e(\vec{v}))) \cdot ((q_o(k_u), q_o(\vec{u})), (q_e(k_u), q_e(\vec{u}))) \\
&\quad = ((q_o(k_v + k_u) + q_o(\vec{v})^\intercal \tilde{\mu}_o q_o(\vec{u}), q_o(\vec{v} + \vec{u})), (q_e(k_v + k_u) + q_e(\vec{v})^\intercal \tilde{\mu}_e q_e(\vec{u}), q_e(\vec{v} + \vec{u}))) \\
&\quad = \phi(k_v + k_u \vec{v}^\intercal \tilde{\mu} \vec{u}, \vec{v} + \vec{v})
\end{aligned}$$

Where in the last step we've applied the CRT. □

Theorem 4.12 Let $d = 2^k \ell \in \mathbb{N}_{\geqslant 2}$, where $k \in \mathbb{N}_{\geqslant 0}$ and ℓ is an odd number and let $\mu \in \mathfrak{so}(n, \mathbb{Z}_n)$. Considering the factoring of $H(\mu)$ into the product $H(\mu_e) \times H(\mu_o)$ as in Lemma 4.11, we have that $H(\mu)$ is algebraically contextual if and only if $H(\mu_e)$ is contextual.

Proof. By the CRT, a non-contextual assignment $o : H(\mu) \to \mathbb{Z}_d$ is the same thing as a pair of non-contextual assignments $o_o : H(\mu_o) \to \mathbb{Z}_\ell$ and $o_e : H(\mu_e) \to \mathbb{Z}_{2^k}$.Given that by Theorem 4.10, a non-contextual assignment $o_o : H(\mu_o) \to \mathbb{Z}_\ell$ always exists, the claim follows. □

The theorem above essentially says that in order to determine whether or not a commutation group $H(\mu)$ is algebraically contextual, it suffices to determine whether $H(\mu_e)$ is algebraically contextual, where μ_e is the "even part" of μ. We end this section with the following remark on decision algorithms for the algebraic contextuality of a commutation group:

Remark 4.13 Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_{2^k})$, the existence of a non-contextual assignment $o : C(\mu) \to \mathbb{Z}_d$ is equivalent to the satisfiability of the following system of linear equations over $\mathbb{Z}_{2^k}$ over the variables $\{o(g) \mid g \in \langle X_\mu \cup \mathbb{Z}_d \rangle_C \}$

$$\{o(g) + o(h) - o(gh) = 0 \mid g, h \in \langle X_\mu \cup \mathbb{Z}_d \rangle_C, g \odot h\} \cup \{o((k, \vec{0})) = k : k \in \mathbb{Z}_d\}$$

This can be expressed as a co-chain condition on o and enables one to explore methods

from homological algebra or algebraic topology in checking if such an o exists in general cases, and in concrete cases, one could compute the Smith normal form corresponding to this system of equations to directly verify if it is satisfiable.

4.3 Contextual Words and Combinatorial Contextuality

There is a combinatorial method (introduced in [13] and generalized in [17]) of proving that a commutation group is algebraically contextual, namely, finding a contextual word. Save for Theorem 3.22, all results present in this subsection are present in [17], although some of the proofs are different.

Definition 4.14 Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and a string of generators $w \in X_\mu^*$, we say that w is a **contextual word for** $G(\mu)$ if there exists a bracketing β of w consisting solely of commuting products and if $\theta(w) = J_k$ for some $k \neq 0$, where $\theta : X_\mu^* \to G(\mu)$ is the map given by the group presentation of $G(\mu)$ in Definition 2.1 from strings of generators to group elements.

Formally, bracketings $\mathbf{BE}_\Sigma$ over an alphabet Σ are defined inductively by the rules:

$$\frac{}{x \in \mathbf{BE}_\Sigma}\ (x \in \Sigma) \qquad\qquad \frac{\beta_1 \in \mathbf{BE}_\Sigma \qquad \beta_2 \in \mathbf{BE}_\Sigma}{(\beta_1\beta_2) \in \mathbf{BE}_\Sigma}$$

Writing ∂ for the "unparenthisation" function given by $\partial : \mathbf{BE}_\mu \to X_\mu^*$, $\partial(x_k) = x_k$, $\partial((\beta_1\beta_2)) = \partial(\beta_1) +\!\!+ \partial(\beta_2)$ (where $+\!\!+$ denotes the string concatenation operator), we say that w **is bracketed by** β if $\partial(\beta) = w$. By "consisting solely out of commuting products", we mean that for any product $(\beta_1\beta_2)$ occurring in our contextual word w, β_1 and β_2 must commute.

Theorem 4.15 Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, if a contextual word w over $G(\mu)$ exists, then $G(\mu)$ is algebraically contextual.

Proof. Because of the commutative bracketing condition, we have that for any non-contextual assignment $o : C(\mu) \to \mathbb{Z}_d$, $o(\theta(w)) = \sum_{x \in X_\mu} n_x o(x)$, where n_x is the number of times that the generator x occurs in the word w. Due to the fact that $\theta(w)$ is a scalar and the equivalence between the group presentation definition and the linear-algebraic definition of commutation groups, d must divide n_x for all $x \in X_\mu$, thus $o(\theta(w)) = J_0$, which contradicts the assumption that w is a contextual word. □

Example 4.16 Consider the commutation group $G(\mu)$, corresponding to the Pauli group $\mathcal{P}_2^2$, given by generator set $X = \{x_1, x_2, y_1, y_2\}$ and the commutator matrix $\mu : X \times X \to \mathbb{Z}_2$ as in the table below:

	x_1	y_1	x_2	y_2
x_1	0	1	0	0
y_1	1	0	0	0
x_1	0	0	0	1
y_2	0	0	1	0

We have that the word $w = (((x_1y_2)(y_1x_2))((x_1x_2)(y_1y_2)))$ is a contextual word over $G(\mu)$.

Remarkably, the reverse implication of Theorem 4.15 is also true:

Theorem 4.17 Let $G(\mu)$ be an arbitrary commutation group for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$. If $G(\mu)$ is algebraically contextual, then there exists a contextual word over $G(\mu)$.

The proof of this result is the subject of Appendix B. Combining Theorem 4.15 and Theorem 4.17, we get that *a commutation group is algebraically contextual if and only if it admits a contextual word.*

The rest of this section is dedicated to proving (again) that commutation groups over odd characteristic cannot be algebraically contextual using solely the framework of contextual

words.

Definition 4.18 We identify the following notation: given a word $w = w_1 w_2 \dots w_{n-1} w_n$ over an alphabet X, we'll write $w^\dagger$ for the reversed word $w_n w_{n-1} \dots w_2 w_1$. Furthermore, given a commutation group $G(\mu)$ for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and a word $w = w_1 w_2 \dots w_{n-1} w_n \in X_\mu^*$, writing $X_\mu = \{x_1, \dots, x_n\}$ and considering the order $x_1 < \dots < x_n$ on X_μ, we define the set of inversions in w as:

$$\mathcal{I}(w) = \{(i, j) \mid 1 \leqslant i \leqslant j \leqslant n, w_i > w_j\}$$

and the sum over the inversions as:

$$\sum \mathcal{I}(w) = \sum_{(i,j) \in \mathcal{I}(w)} \mu(w_i, w_j)$$

Lemma 4.19 Given a commutation group $G(\mu)$ for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and a word $w \in X_\mu^*$, the scalar part of $\theta(w)$ is equal to $\sum \mathcal{I}(w)$.

Proof. The proof of this is similar to that of Lemma 3.3. We can essentially think of this as bringing the word into normal form by applying bubble sort to order the generators; every swap in the bubble sort corresponding to an inversion in the original word, which contributes $\mu(x, y)$ to the global scalar factor for each swap of generators $x, y \in X_\mu$. □

Lemma 4.20 For any commutation group $G(\mu)$ for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ and word $w \in X_\mu^*$,

$$\sum \mathcal{I}(w) + \sum \mathcal{I}(w^\dagger) = \sum_{x,y \in X_\mu} n_x n_y \mu(x, y),$$

where n_x denotes the number of occurrences of the generator x in the word w.

Proof. This follows from the fact that every pair of distinct indices $1 \leqslant i \leqslant j \leqslant n$ such that $w_i \neq w_j$ corresponds to an inversion either in w or in $w^\dagger$. □

Lemma 4.21 Given a contextual word w over a commutation group $G(\mu)$, where $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, we have that:

$$\sum \mathcal{I}(w) = \sum \mathcal{I}(w^\dagger)$$

Proof. The intuitive way to think about the proof of this fact is based on the fact that if we have a sequence stored in the leaves of a binary tree and we reverse the right and left son of every internal node (i.e. non-leaf), we obtain a binary tree which stores the reversed sequence. More concretely in the case of a commutation group, note that if we swap the factors x_1 and x_2 for every commuting product $(x_1 x_2)$ in the parenthesization of the word w, we obtain the reversed word while not modifying the image of the word through θ into the group $G(\mu)$, i.e. $\theta(w) = \theta(w^\dagger)$. Given that w and $w^\dagger$ have the same image in $G(\mu)$, Lemma 4.19 implies the claim. □

Theorem 4.22 Given a commutation group $G(\mu)$ for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, if d is odd, then no contextual words over $G(\mu)$ exist.

Proof. Suppose for a contradiction that such a contextual word w does exist. By Lemma 4.21, we have that $\sum \mathcal{I}(w) - \sum \mathcal{I}(w^\dagger) = 0$; moreover, as $\theta(w)$ is a scalar element in $G(\mu)$, we must have that the number of occurrences n_x of each generator $x \in X_\mu$ in w is a multiple of d, therefore, we also have that $\sum \mathcal{I}(w) + \sum \mathcal{I}(w^\dagger) = 0$. Adding these two equalities, we get that $2 \sum \mathcal{I}(w) = 0$, which due to the fact that we're in odd characteristic, implies that $\sum \mathcal{I}(w) = 0$, which coupled with Lemma 4.19 contradicts the assumption that w is a contextual word (more precisely, that its scalar part is not zero). □

4.4 Unrestricted Contextuality

When we defined algebraic contextuality for a commutation group $G(\mu)$ in Section 4.1, we have done so in a manner which made algebraic contextuality equivalent to state-independent

contextuality for any observable in X_μ for any representation of $G(\mu)$. In this section, we concern ourselves with finding out whether or not the set of all observables (again, for any representation) in $G(\mu)$ is state-independently contextual. Formally, that is:

Definition 4.23 Given a commutation group $H(\mu)$ for some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, we define a **complete non-contextual assignment** (CNA) as a compatible monoid morphism $o : H(\mu) \to \mathbb{Z}_{2d}$ such that $o(k, \vec{0}) \mapsto \iota(k)$ ($\iota : \mathbb{Z}_d \to \mathbb{Z}_{2d}$ is the homomorphism mapping $1 \in \mathbb{Z}_d$ to $2 \in \mathbb{Z}_{2d}$). We say that $H(\mu)$ displays **complete algebraic contextuality** (CAC) if no such complete non-contextual assignment exists.

Remark 4.24 Note that in the above definition, the compatible monoid we denote by $H(\mu)$ is not the same thing as $C(\mu)$ (hence why Definition 4.6 and Definition 4.23 are different); the difference lies in the fact that $H(\mu)$ is the compatible monoid we obtain by simply forgetting that we may multiply non-commuting elements of the group $H(\mu)$, whereas $C(\mu)$ has a smaller carrier set, namely the commutative closure of X_μ and the scalar elements.

Remark 4.25 Noting that every element g in a commutation group $H(\mu)$ (where $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$) obeys that $g^{2d} = e$, we obtain a result analogous to Theorem 4.7 motivating Definition 4.23, namely that any compatible monoid representation of $H(\mu)$ that preserves scalars need be state-independently contextual.

Another distinguishing aspect of complete algebraic contextuality in contrast to the notion of contextuality discussed so far is that deciding whether a commutation group is CAC is significantly computationally easier to decide than contextuality as in Definition 4.6. The rest of this subsection is dedicated to answering the problem: "Given some $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, is $H(\mu)$ completely algebraically contextual?".

By an analogous argument to that in Theorem 4.17, we may assume that $d = 2^k$ for some $k \in \mathbb{N}_{\geqslant 1}$. Moreover, by Theorem 3.16, we may assume that μ is in Darboux normal form.

Padding with zeros if necessary, suppose that μ has the form:

$$\mu = \begin{bmatrix} 0 & \lambda_1 & 0 & 0 & \dots & 0 & 0 \\ -\lambda_1 & 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & & 0 & 0 \\ 0 & 0 & -\lambda_2 & 0 & \dots & 0 & 0 \\ \vdots & & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_n \\ 0 & 0 & 0 & 0 & \dots & -\lambda_n & 0 \end{bmatrix}$$

for some $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{Z}_{2^k}$. If two or more of these λ terms are non-zero, then we may obtain a contextual word as in Example 4.16, which implies that $H(\mu)$ is CAC, because contextuality as in Definition 4.6 clearly implies CAC. If none of the terms are non-zero, then a CNA $H(\mu) \to \mathbb{Z}_d$ clearly exists. This leaves us with the case of a single λ value being non-zero. In this case, we do not know if (but strongly suspect that) a complete non-contextual assignment $H(\mu) \to \mathbb{Z}_{2^{k+1}}$ exists. Essentially, the decision problem is reduced to the following conjecture, which is equivalent to asking whether or not state-independent contextuality can occur in the generalized Pauli group over a single qudit.

Conjecture 4.26 Given a commutator matrix

$$\mu = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathfrak{so}(2, \mathbb{Z}_{2^k}),$$

there exists a complete non-contextual assignment $o : H(\mu) \to \mathbb{Z}_{2^{k+1}}$.

Using backtracking, we've managed to find such non-contextual assignments for $k \leqslant 4$, but no general method of finding such assignments (see Appendix D).

5 Further Work

In this section, we discuss a couple of open problems and unexplored research directions in the study of commutation groups:

- Given a commutator matrix $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$, is there an algorithm running in time polynomial in the size of μ for deciding whether or not $C(\mu)$ is algebraically contextual?
- Solving Conjecture 4.26 and thus solving the problem of efficiently deciding complete contextuality;
- Exploring cohomological approaches to the study of algebraic contextuality using methods such as those described in [2, 6, 10, 16];
- Generalizing the structure of commutation groups to other discrete structures, for example, describing and exploring notions of algebraic contextuality which arise when considering any group G such that for any $g, h \in G$, the commutator $[g, h]$ lies in the centre of the group;
- Analysing state-dependent contextuality arising in sets of observables which form a commutation group.

Appendices

A Proof of Theorem 3.16

We begin with some general facts about rings:

Definition A.1 Given a commutative ring R, and elements $x, y \in R$, we say that x **divides** y and write $x|y$ if $y = rx$ for some $r \in R$. If $x|y$ and $y|x$, we say that x and y are **congruent** and write $x \sim y$.

Definition A.2 Given some integer $d \geqslant 2$, we define the function $\gcd : \mathbb{Z}_d \times \mathbb{Z}_d \to \mathbb{Z}_d$ as follows: let $a, b \in \mathbb{Z}_d$. By the ideal correspondence theorem and the fact that $\mathbb{Z}_d \simeq \mathbb{Z}/(d)$, we have that (a, b) is a principal ideal in $\mathbb{Z}_d$. Let $\mathcal{C} = \{c_1, \ldots, c_m\}$ be the set of elements in $\mathbb{Z}_d$ that generate (a, b). We set $\gcd(a, b) := c_i$, where c_i is the element of $\mathcal{C}$ such that the smallest non-negative element of $\phi^{-1}(c_i)$ is minimum (where $\phi : \mathbb{Z} \to \mathbb{Z}_d$ is the natural quotient map).

Lemma A.3 Let $\phi : \mathbb{Z} \to \mathbb{Z}_d$ be the natural quotient map onto $\mathbb{Z}_d$. Let $a, b \in \mathbb{Z}_d$ and $a', b' \in \mathbb{Z}$ such that $\phi(a') = a, \phi(b') = b$. Then $\gcd(a, b) \sim \phi(\gcd(a', b'))$.

Definition A.4 Given numbers $a, b \in \mathbb{Z}$, we define $\Gamma_{a,b} \in \mathrm{GL}_2(\mathbb{Z})$ as:

$$\Gamma_{a,b} := \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \text{if } b = 0 \\ \Gamma_{-a,b} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} & \text{if } a < 0 \\ \Gamma_{a,-b} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & \text{if } b < 0 \\ \Gamma_{b,a \bmod b} \begin{bmatrix} 0 & 1 \\ 1 & -\lfloor b/a \rfloor \end{bmatrix} & \text{if } a < 0 \end{cases}$$

Lemma A.5 For any $d \geqslant 2$ and $a, b \in \mathbb{Z}_d$,

$$\Gamma_{a,b} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \gcd(a,b) \\ 0 \end{bmatrix}$$

Lemma A.2 can be proved by induction using the Euclidean algorithm.

Lemma A.6 Let $V = \mathbb{Z}_d^n$, where $n \geqslant 3$, let $\mu : V \times V \to \mathbb{Z}_d$ be a skew-symmetric bilinear form on V (in the context of this paper, we may think of μ as of a commutator matrix) and $\mathcal{E} = \{\vec{e}_1, \ldots, \vec{e}_m\}$ be a basis for V. Let $i, j, k \in [n]$ be any three distinct indices and let $a := \mu(\vec{e}_i, \vec{e}_j)$ and $b := \mu(\vec{e}_i, \vec{e}_k)$. Then there exist $\vec{e'}_j, \vec{e'}_k \in V$ such that $(\mathcal{E} \backslash \{\vec{e}_j, \vec{e}_k\}) \cup \{\vec{e'}_j, \vec{e'}_k\}$ is again a basis of V, $\mu(\vec{e}_i, \vec{e'}_j) \sim \gcd(a, b)$ and $\mu(\vec{e}_i, \vec{e'}_k) = 0$.

Proof. Let $\phi : \mathbb{Z} \to \mathbb{Z}_d$ be the natural quotient map onto $\mathbb{Z}_d$. Let $a' \in \phi^{-1}(a)$ and $b' \in \phi^{-1}(b)$. Writing

$$\Gamma_{a',b'} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix},$$

let

$$\tilde{\Gamma}_{a,b} = \begin{bmatrix} \phi(\gamma_{11}) & \phi(\gamma_{12}) \\ \phi(\gamma_{21}) & \phi(\gamma_{22}) \end{bmatrix};$$

by lemmas A.1 and A.2, we have that $\tilde{\Gamma}_{a,b} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$, where $x \sim \gcd(a, b)$. Noting that $\tilde{\Gamma}_{a,b} \in \mathrm{GL}_2(\mathbb{Z}_d)$ and writing:

$$\gamma : \vec{e}_j \mapsto \gamma_{11}\vec{e}_j + \gamma_{21}\vec{e}_k \qquad\qquad \gamma : \vec{e}_k \mapsto \gamma_{12}\vec{e}_j + \gamma_{22}\vec{e}_k,$$

we have that γ is an *invertible* linear transformation $\langle \vec{e}_j, \vec{e}_k \rangle_{\mathbb{Z}_d} \to \langle \vec{e}_j, \vec{e}_k \rangle_{\mathbb{Z}_d}$, so setting $\vec{e'}_j = \gamma(\vec{e}_j)$ and $\vec{e'}_k = \gamma(\vec{e}_k)$, we get that $(\mathcal{E} \backslash \{\vec{e}_j, \vec{e}_k\}) \cup \{\vec{e'}_j, \vec{e'}_k\}$ is again a basis of V and

$$\begin{aligned} \mu(\vec{e}_i, \vec{e'}_j) &= \mu(\vec{e}_i, \gamma_{11}\vec{e}_j + \gamma_{21}\vec{e}_k) \\ &= \gamma_{11}\mu(\vec{e}_i, \vec{e}_j) + \gamma_{21}\mu(\vec{e}_i, \vec{e}_k) \\ &= \gamma_{11}a + \gamma_{21}b \\ &= x, \end{aligned}$$

$$\begin{aligned} \mu(\vec{e}_i, \vec{e'}_k) &= \mu(\vec{e}_i, \gamma_{12}\vec{e}_j + \gamma_{22}\vec{e}_k) \\ &= \gamma_{12}\mu(\vec{e}_i, \vec{e}_j) + \gamma_{22}\mu(\vec{e}_i, \vec{e}_k) \\ &= \gamma_{12}a + \gamma_{22}b \\ &= 0. \end{aligned}$$

□

Lemma A.7 Let $V = \mathbb{Z}_d^n$ and $\mu : V \times V \to \mathbb{Z}_d$ be a skew-symmetric bilinear form on V. There is a basis $\mathcal{F} = \{\vec{f}_1, \ldots, \vec{f}_n\}$ of V such that for any two indices $i, j \in [n]$, $\mu(\vec{f}_i, \vec{f}_j) = 0$ whenever $|i - j| > 1$.

Proof. Let $\mathcal{E} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ be the canonical basis of $\mathbb{Z}_d^n$. By an application of Lemma A.3,

we may replace the last two elements of the basis by $\vec{e}_{n-1} \mapsto \vec{t}_{n-1}$ and $\vec{e}_n \mapsto \vec{t}_n$, so that $\{\vec{e}_1, \ldots, \vec{e}_{n-2}, \vec{t}_{n-1}, \vec{t}_n\}$ is again a basis of V such that $\mu(\vec{t}_n, \vec{t}_{n-1}) = \gcd(\mu(\vec{e}_1, \vec{e}_{n-1}), \mu(\vec{e}_1, \vec{e}_n))$ and $\mu(\vec{e}_1, \vec{t}_n) = 0$. By iterating this process, first with the second to last and third to last basis elements, then with the third to last and fourth to last basis elements and so on, we obtain a new basis $\mathcal{E}' = \{\vec{e}_1, \vec{f}_2, \ldots, \vec{f}_n\}$ such that $\mu(\vec{e}_1, \vec{f}_2) = \gcd_{i=2}^{n} \mu(\vec{e}_1, \vec{e}_i)$ and $\mu(\vec{e}_1, \vec{f}_k) = 0$ for all $k \geqslant 3$.

Note that this procedure does not change the first element. We now repeat this procedure with the suffix $\vec{f}_2, \ldots, \vec{f}_n$ of the basis obtaining a new basis of the form $\vec{e}_1, \vec{f}_2, \vec{g}_3, \ldots, \vec{g}_n$ such that $\mu(\vec{e}_1, \vec{g}_k) = 0$ for all $k \geqslant 3$ (as we have used the procedure in Lemma A.3) and $\mu(\vec{f}_2, \vec{g}_k) = 0$ for all $k \geqslant 4$.

By iteratively applying this process of basis change which leaves the first basis element unchanged while ensuring that the second basis element is the only one whose μ-form value is non-zero with the first basis element on each suffix of the basis, we obtain the desired result. □

And now for the final result:

Theorem 3.16 Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ be a commutator matrix. There is a basis of $V = \mathbb{Z}_d^n$:

$$\mathcal{F} = \{\vec{a}_1, \ldots, \vec{a}_m, b_1, \ldots, \vec{b}_m, \vec{c}_1, \ldots, \vec{c}_l\}$$

such that for any $i, j \in [m]$ and $k, t \in [l]$

$$\mu(\vec{a}_i, \vec{b}_j) = \delta_{ij}\Delta_i$$
$$\mu(\vec{a}_i, \vec{a}_j) = \mu(\vec{b}_i, \vec{b}_j) = \mu(\vec{c}_t, \vec{c}_k) = 0$$
$$\mu(\vec{a}_i, \vec{c}_k) = 0$$
$$\mu(\vec{b}_i, \vec{c}_k) = 0$$

for some $\Delta_1, \ldots, \Delta_m \in \mathbb{Z}_d$.

Proof. We will define a chain as a maximal sequence of consecutive indices $i, i+1, \ldots, j$ such that for $u, v \in [i \ldots j]$, $\mu(\vec{e}_u, \vec{e}_v) \neq 0$ iff $|u - v| = 1$ and for any $u \in [i \ldots j]$ and $v \notin [i \ldots j]$, $\mu(\vec{e}_u, \vec{e}_v) = 0$. The length of a chain $i, \ldots, j$ is defined to be $j - i + 1$. We say that a basis $\mathcal{E}$ of V is partitioned into chains if every element of it is part of a chain. By Lemma A.4, we may consider a basis $\mathcal{E} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ of V such that $\mu(\vec{e}_i, \vec{e}_j) = 0$ whenever $|i - j| \neq 1$, so that $\mathcal{E}$ is a partitioned into chains.

We will describe a procedure that given a chain $i \ldots j$ of length $\geqslant 3$, performs a change of basis of the subspace $\langle \vec{e}_i, \ldots, \vec{e}_j \rangle$ by some $\vec{e}_i \mapsto \vec{e'}_i, \ldots, \vec{e}_j \mapsto \vec{e'}_j$, (so that $\langle \vec{e'}_i \ldots, \vec{e'}_j \rangle_{\mathbb{Z}_d} = \langle \vec{e}_i, \ldots, \vec{e}_j \rangle_{\mathbb{Z}_d}$) such that the resulting basis $\mathcal{E}' := \mathcal{E} \backslash \{\vec{e}_i, \ldots, \vec{e}_j\} \cup \{\vec{e'}_i, \ldots, \vec{e'}_j\}$ is still partitioned into chains, but the chain $i \ldots j$ has been broken into smaller chains. Iterating this procedure, we obtained a basis partitioned into chains of length at most 2, from which the claim immediately follows.

Consider an arbitrary chain of length $\geqslant 3$ and without loss of generality (rearranging the basis elements if needed), suppose that this chain is: $1, \ldots, m$ for some $m \geqslant 3$. Let $\lambda_1 := \mu(\vec{e}_1, \vec{e}_2) (\neq 0)$ and $\lambda_2 := \mu(\vec{e}_3, \vec{e}_2) (\neq 0)$. We split the process into two cases. First, if $\lambda_1 | \lambda_2$, we may perform the basis change which just replaces $\vec{e}_2$ with $\vec{f_2} := \vec{e}_2 - (\lambda_2/\lambda_1)\vec{e}_1$. The resulting sequence has $\mu(\vec{e}_1, \vec{f_2}), \mu(\vec{e}_3, \vec{e}_4), \ldots, \mu(\vec{e}_{m-1}, \vec{e}_m)$ the same as before, but $\mu(\vec{f_2}, \vec{e}_3) = 0$, thus resulting in two smaller chains. In the second case (i.e. when $\lambda_1 \nmid \lambda_2$), applying lemma A.3 with $i = 2, j = 1, k = 3$, we may consider integers x_0, y_0, x_1, y_1 such that the substitution

$$\vec{e}_1 \mapsto \vec{f_1} := x_0 \vec{e}_1 + y_0 \vec{e}_3$$
$$\vec{e}_3 \mapsto \vec{f_3} := x_1 \vec{e}_1 + y_1 \vec{e}_3$$

is a valid change of basis and $\mu(\vec{f_1}, \vec{e}_2) = \gcd(\mu(\vec{e}_1, \vec{e}_2), \mu(\vec{e}_3, \vec{e}_2))$ which *strictly* divides λ_1 and $\mu(\vec{f_3}, \vec{e}_2) = 0$. We then apply the same procedure as in the proof of Lemma A.4 over the basis

elements $\vec{f}_1, \vec{e}_2, \vec{f}_3, \vec{e}_4, \vec{e}_5, \ldots, \vec{e}_m$ to obtain a new basis $\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_m$ (with $\vec{g}_1 = \vec{f}_1$) obeying $\mu(\vec{g}_i, \vec{g}_j) = 0$ whenever $|i - j| > 1$ and

$$\mu(\vec{g}_1, \vec{g}_2) | \gcd(\lambda_1, \lambda_2)$$

which need be strictly smaller than λ_1 in the divisibility order. As we repeat this process, $\mu(\vec{e}_1, \vec{e}_2)$ cannot decrease indefinitely, so eventually, the first case must apply, resulting in a smaller chain. □

B Proof of Theorem 4.17

Let $\mu \in \mathfrak{so}(n, \mathbb{Z}_d)$ denote an arbitrary commutator matrix.

The observation lying at the heart of this proof is the fact that a contextual word exists if and only if a non-trivial ($k \neq 0$) scalar element J_k lies in the commutative closure of the generating set X_μ. We will prove the contrapositive of Theorem 3.5, which in light of our observation is: *if the commutative closure of the generating set X_μ of a commutation group $H(\mu)$ contains no non-trivial scalars, then a non-contextual assignment $o : C(\mu) \to \mathbb{Z}_d$ must exist.* We will prove this constructively:

Lemma B.1 If $(k_a, \vec{v}), (k_b, \vec{v}) \in \langle X_\mu \rangle_C$, then $k_a = k_b$.

Proof. Being part of a finite group, $(k_b, \vec{v})$ has an inverse that can be written as $(k_b, \vec{v})^m$ for some m, therefore, $(k_b, \vec{v})^{-1} \in \langle X_\mu \rangle_C$. Therefore,

$$(k_a, \vec{v}) \cdot (k_b, \vec{v})^{-1} = (k_a - k_b, \vec{0}) \in \langle X_\mu \rangle_C,$$

which by the hypothesis that $\langle X_\mu \rangle_C$ contains no non-trivial scalars implies that $k_a = k_b$. □

Let $S = \{\vec{v} | \exists k \in \mathbb{Z}_d \,.(k, \vec{v}) \in \langle X \rangle_C\}$. Lemma B.1 allows us to consider the following functional relation between S and $\mathbb{Z}_d$: given some $\vec{v} \in S$, we define $k_{\vec{v}}$ to be the unique number such that $(k_{\vec{v}}, \vec{v}) \in \langle X_\mu \rangle_C$. This makes the mapping o well-defined:

Noting that the carrier set of $C(\mu)$ is $\langle X \cup \{(k, \vec{0}) \mid k \in \mathbb{Z}_d\} \rangle_C = \{(k, \vec{v}) \mid v \in S, k \in \mathbb{Z}_d\}$, we define the following mapping $o : C(\mu) \to \mathbb{Z}_d$:

$$o : (k_{\vec{v}} + s, \vec{v}) \mapsto s \qquad (\forall \vec{v} \in S, s \in \mathbb{Z}_d)$$

We will now show that this is a non-contextual assignment and thus prove Theorem 3.5: First note that as $\vec{0} \in S$ and $k_{\vec{0}} = 0$, we have that $o(k, \vec{0}) = k$. Second, we need to check the compatible monoid homomorphism condition; consider two arbitrary elements $(k_{\vec{v}}+a, \vec{v})$, $(k_{\vec{u}}+b, \vec{u}) \in C(\mu)$; we have that:

$$\begin{aligned} o((k_{\vec{v}} + a, \vec{v}) \cdot (k_{\vec{u}} + b, \vec{u})) &= o((a + b, \vec{0}) \cdot (k_{\vec{v}}, \vec{v}) \cdot (k_{\vec{u}}, \vec{u})) \\ &= o((a + b, \vec{0}) \cdot (k_{\vec{v}+\vec{u}}, \vec{u} + \vec{v})) \\ &= o((k_{\vec{v}+\vec{u}} + a + b, \vec{u} + \vec{v})) \\ &= a + b \\ &= o(k_{\vec{v}} + a, \vec{v}) + o(k_{\vec{u}} + b, \vec{u}) \end{aligned}$$

and so o is a non-contextual assignment.

References

[1] M.Born A. Einstein, H. Born. *Briefwechsel 1916-1955*. Rohwolt, 1972.

[2] Sivert Aasnæss. Comparing two cohomological obstructions for contextuality, and a generalised construction of quantum advantage with shallow circuits, 2022. arXiv:2212.09382.

[3] Samson Abramsky. Relational databases and bell's theorem, 2013. arXiv:1208.6416.

[4] Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, November 2011. URL: http://dx.doi.org/10.1088/1367-2630/13/11/113036, doi:10.1088/1367-2630/13/11/113036.

[5] Samson Abramsky, Georg Gottlob, and Phokion Kolaitis. Robust constraint satisfaction and local hidden variables in quantum mechanics. pages 440–446, 08 2013.

[6] Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, cohomology and paradox. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CSL.2015.211, doi:10.4230/LIPICS.CSL.2015.211.

[7] John S Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1(3):195, 1964.

[8] N. Bohr. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 48:696–702, Oct 1935. URL: https://link.aps.org/doi/10.1103/PhysRev.48.696, doi:10.1103/PhysRev.48.696.

[9] Max Born. Quantenmechanik der stoßvorgänge. *Zeitschrift für physik*, 38(11):803–827, 1926.

[10] G Caru. *Logical and topological contextuality in quantum mechanics and beyond*. PhD thesis, University of Oxford, 2019.

[11] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935. URL: https://link.aps.org/doi/10.1103/PhysRev.47.777, doi:10.1103/PhysRev.47.777.

[12] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510(7505):351–355, June 2014. URL: http://dx.doi.org/10.1038/nature13460, doi:10.1038/nature13460.

[13] William M. Kirby and Peter J. Love. Contextuality test of the nonclassicality of variational quantum eigensolvers. *Physical Review Letters*, 123(20), November 2019. URL: http://dx.doi.org/10.1103/PhysRevLett.123.200501, doi:10.1103/physrevlett.123.200501.

[14] SIMON KOCHEN and E. P. SPECKER. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967. URL: http://www.jstor.org/stable/24902153.

[15] N. David Mermin. Hidden variables and the two theorems of john bell. *Reviews of Modern Physics*, 65(3):803–815, July 1993. URL: http://dx.doi.org/10.1103/RevModPhys.65.803, doi:10.1103/revmodphys.65.803.

[16] Cihan Okay, Sam Roberts, Stephen D Bartlett, and Robert Raussendorf. Topological proofs of contextuality in quantum mechanics. *arXiv preprint arXiv:1701.01888*, 2017.

[17] Me Samson Abramsky and Carmen Constantin. Commutation groups and state-independent contextuality. 2024.

[18] Sam Staton and Sander Uijlen. Effect algebras, presheaves, non-locality and contextuality. *Information and Computation*, 261:336–354, 2018.