Wigner Function Negativity and Contextuality in Quantum Computation with Rebits

by

Philippe Allard Guérin

B. Sc. Physics, Université de Montréal, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Physics)

The University of British Columbia
(Vancouver)

August 2015

Abstract

We study the resources necessary for quantum computation with rebits (qubit states with real amplitudes in the standard basis). We introduce a scheme for universal quantum computation by state injection, and define a Wigner function appropriate for this scheme. We show that the Wigner function obeys a Hudson's theorem and transforms covariantly under CSS-ness preserving unitary gates; these results allows us to establish that Wigner function negativity is necessary for quantum computation. Furthermore, we establish contextuality as another necessary computational resource. We show that in contrast with the case of qudits [M. Howard et al., Nature 510, 351 (2014)], negativity does not imply contextuality. We discuss state independent contextuality and why it does not arise in our computational scheme.

Preface

The material presented in Chapters 2, 3 and 4 has been published as [Nicolas Delfosse, Philippe Allard Guerin, Jacob Bian, and Robert Raussendorf as *Wigner Function Negativity and Contextuality in Quantum Computation on Rebits*. Phys. Rev. X 5, 021003 – 2015]. My contributions to this work include Chapter 2, and discussions with Robert Raussendorf on the contents of Chapter 4, which he wrote. Chapter 3 was principally the work of Nicolas Delfosse, with contributions from Jacob Bian. In both previously mentioned chapters, I have slightly modified or expanded some of the proofs from their original shape in the paper. Chapter 1 is a review of the preexisting literature, and it is original work that I realized for this thesis. Figures 1.3, 4.1, 4.2, 4.3, as well as the circuit diagrams of Chapter 2 are taken from the published paper, and were produced by Robert Raussendorf. The remaining figures were created by me for this thesis. I was closely involved in rereading and giving feedback during the writing of the paper, as well as during the review process.

Table of Contents

List of Figures

Acknowledgments

I wish to thank my Master's research supervisor Dr. Robert Raussendorf for his guidance and for allowing me to explore other interesting areas of quantum information than those presented in this thesis. In particular, I am grateful for having the opportunity to go to Paris for three months to do research on topological quantum computation, for which there was unfortunately no room in this thesis. I thank Dr. Rémy Mosseri of the Université Pierre et Marie Curie for his enthusiastic supervision when I was working with him in Paris.

I acknowledge my collaborators Jacob Bian and Nicolas Delfosse, without whom the work presented in this thesis could not have come to fruition.

I am eternally grateful to my parents Luc Guérin and Sylvie Allard for supporting me throughout my life and for encouraging my passion for learning. Finally, I wish to express gratitude to all my friends at UBC and elsewhere, for the pleasant discussions and debates that have enriched my experience as a graduate student.

Chapter 1

Introduction

Quantum computation is a field of research whose broad aim is to use the quantum properties of matter in order to acheive greater computational power than what is available classically. Many quantum algorithms, Shor's [23] and Grover's [15] being famous examples, have been devised, and these algorithms offer considerable speedup over their classical competitors. We shall give the designation universal quantum computer to any machine that can apply any unitary operation (with arbitrary precision) to the state of a quantum system [20]. Such theoretical machines are able to run Shor's and Grover's algorithms, and are thus believed to be faster than classical computers. However, the reason for this faster speed is not very well understood.

Superposition and entanglement are well known non-classical features of quantum mechanics. However, this work will study the role of a lesser known quantum property called contextuality, and will establish it as a necessary resource for quantum computation. Contextuality arises in the study of hidden variable models (HVM's) of quantum mechanics, the attempt to assign preestablished values to the observables of a quantum system, which are merely revealed by measurement. Bell [4] famously proved that there can be no HVM that reproduces quantum mechanics while obeying the locality postulate of special relativity. But there is a broader class of HVM's that cannot reproduce quantum mechanics: non-contextual HVM's [18]. These

are HVM's for which the values assigned to an observable do not depend on which other observables are simultaneously measured with it.

In the study of quantum computation, there are many equivalent schemes that allow universality. The paradigmatic example is the circuit model [20], but there is also measurement based quantum computation [21], and quantum computation by state injection (QCSI) [7]; we shall focus on the latter in this thesis. For systems with local odd prime dimension (qudits), it has been discovered [16] that for the computational scheme of QCSI, contextuality is necessary for universal quantum computation. One of the goals of this thesis is to try to establish similar results for two-level systems. We will be lead to study systems of rebits, two-level systems with real density matrices in the standard basis.

The result according to which contextuality is a resource for quantum computation draws heavily for it's proof on a construction known as the Wigner function [29]. These functions are meant to be quantum analogs of the phase space distributions that arise in classical mechanics. However, quantum mechanics is such that these distributions can sometimes take negative values, which gives them the name quasi-probability distributions. For qudits, it has been shown [27], that negativity of the Wigner function is necessary for universal quantum computation. Furthermore, contextuality and negativity are equivalent for qudits [16]. In this thesis, we will define a Wigner function for rebits, and study the link between contextuality, negativity, and computation.

In the following sections of this chapter, we will set the stage by mathematically defining our main subjects of study: Wigner functions and contextuality. We will give an introduction to the phase space formulation of quantum computation for continuous variables, qudits and qubits. We also discuss the scheme of quantum computation by state injection, which will be the scheme of computation studied in this thesis and especially in Chapter 2. We will then give an introduction to contextuality and hidden variable models of quantum mechanics. This introduction also presents what is known about the link between negativity of the Wigner function, contextuality, and quantum computation in the case of qudits. Chapter 2

introduces rebits and presents the scheme that allows us universal quantum computation with rebits. Chapter 3 defines and characterizes the rebit Wigner function, and establishes negativity of the Wigner function as a resource for quantum computation. Chapter 4 studies contextuality of rebit quantum computation, and uses the Wigner function to find criteria for contextuality. We finish this thesis with a conclusion in Chapter 5. There is also an Appendix that compiles some mathematical results and definitions about characters and symplectic vector spaces.

1.1 Quantum computation in phase space

In this section we study the phase space formulation for systems of qubits. We start by reviewing some definitions from the stabilizer formalism. We then define a Wigner function, and try to justify why we will focus on rebits for the rest of this thesis.

1.1.1 Stabilizer formalism

The stabilizer formalism [13] is an indispensable tool in the field of quantum error correction [20], and we give a brief introduction to the formalism herein. We start by defining the n-qubit Pauli group . Let $\mathbf{u}_X, \mathbf{u}_Z \in \mathbb{Z}_2^n$ and define $Z(\mathbf{u}) = Z_1^{u_1} \otimes Z_2^{u_2} \otimes \cdots \otimes Z_n^{u_n}$, where Z_i is the spin Z operator acting on qubit i, and define $X(\mathbf{u})$ in an analogous way. Then the n-qubit Pauli group is the group generated by the local gates $\langle X, Y, Z\rangle$ for each qubit of the system.

$$\mathcal{P}_n = \{(i)^k Z(\mathbf{u}_Z)X(\mathbf{u}_X) | (\mathbf{u}_X, \mathbf{u}_Z) \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n, k \in \{0,1,2,3\}\} \tag{1.1}$$

It will be useful to make a distinction between the n-qubit Pauli group $\mathcal{P}_n$ and the n-qubit Pauli operators, which we define as

$$\mathcal{T}_n = \{T_{(\mathbf{a}_Z, \mathbf{a}_X)} = Z(\mathbf{a}_Z)X(\mathbf{a}_X) | (\mathbf{u}_X, \mathbf{u}_Z) \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n\} \tag{1.2}$$

Notice the absence of the factor i in this definition. In the case of a single qubit it means that $T_{(1,1)} = ZX \neq Y$, and that $T_{(1,1)}$ is not hermitian. The set $\mathcal{T}_n$ is not a group because it is not closed under multiplication. To obtain

a group from it, we must at least include phases $\{\pm 1\}$. The properties of the Pauli operators can be obtained from the well known properties of the single qubit Pauli matrices. For instance, all $T \in \mathcal{T}_n$ are traceless, except $T_{\mathbf{0}} = I$, which has trace 2^n. Also, a pair of Pauli operators either commutes or anticommutes, and

$$T_{\mathbf{u}} T_{\mathbf{v}} = (-1)^{[\mathbf{u},\mathbf{v}]} T_{\mathbf{v}} T_{\mathbf{u}}, \tag{1.3}$$

where $[\cdot,\cdot] : \mathbb{Z}_2^n \times \mathbb{Z}_2^n \to \mathbb{Z}_2$ is the symplectic inner product, defined as

$$[\mathbf{u},\mathbf{v}] = (\mathbf{u}_X, \mathbf{v}_Z) + (\mathbf{u}_Z, \mathbf{v}_X), \tag{1.4}$$

with $(\cdot,\cdot)$ being the usual inner product over $\mathbb{Z}_2^n$.

The Pauli operators $\mathcal{T}_n$ form an orthonormal basis for the 2^n dimensional complex vector space of n by n square matrices, in which the inner product is defined as

$$(A, B) = \frac{1}{2^n} \mathrm{Tr}(A^\dagger B) \tag{1.5}$$

Indeed, it can be easily checked that with this definition,

$$(T_{\mathbf{a}}, T_{\mathbf{b}}) = \delta_{\mathbf{a},\mathbf{b}} \quad \forall T_{\mathbf{a}}, T_{\mathbf{b}} \in \mathcal{T}_n \tag{1.6}$$

The n-qubit Clifford group $\mathcal{C}_n$ is defined as the subgroup of the unitary group that maps the Pauli group to itself under conjugation; in technical terms, it is the normalizer of the Pauli group.

$$\mathcal{C}_n = \{g \in U(n) | g \mathcal{P}_n g^\dagger = \mathcal{P}_n\} \tag{1.7}$$

Furthermore, it is known [13] that the Clifford group is generated by the single qubit Pauli operators, single qubit Hadarmard gates, and the two-qubit CNOT gates.

$$\mathcal{C}_n = \langle X_i, Z_i, H_i, CNOT(i,j) \rangle \tag{1.8}$$

Of course, since a global phase does not change the physical state of the system, all the gates in $\mathcal{C}_n$ may be multiplied by a complex number of modulus one.

We say that a state $|\psi\rangle$ is stabilized by a Pauli operator $g \in \mathcal{P}_n$ if $|\psi\rangle$ is an eigenvector of g with eigenvalue one:

$$g\,|\psi\rangle = +\,|\psi\rangle \tag{1.9}$$

A n-qubit state that is stabilized by n independent Pauli operators (in the sense that none of these operators can be obtained by multiplying some of the other $n-1$ operators) is called a stabilizer state. Knowing the stabilizer of a state specifies it completely, up to a phase. The group generated by these n independent Pauli operators is called the stabilizer group of $|\psi\rangle$ and the notation we use is

$$\mathcal{S}(|\psi\rangle) = \{g \in \mathcal{P}_n |\, g\,|\psi\rangle = |\psi\rangle\} \tag{1.10}$$

The Gottesman-Knill Theorem [13] states that the action of Clifford unitaries on stabilizer states is efficiently classically simulable. Briefly put, the reason is because we can efficiently track the evolution of the stabilizer of the state after every Clifford unitary that appears in the computation.

An important subset of stabilizer states are the Calderbank-Shor-Steane (CSS) states [8]. These are stabilizer states with the property that their stabilizer group can be decomposed in a purely Z part and a purely X part: $\mathcal{S}(|\psi\rangle) = \mathcal{S}_X(|\psi\rangle) \times \mathcal{S}_Z(|\psi\rangle)$. We also define the group of CSS-ness preserving transformations as the subgroup of the Clifford group that preserves the set Σ of CSS states.

$$G_{CSS} = \{g \in \mathcal{C}_n |\, g\,|\psi\rangle \in \Sigma\,, \forall\,|\psi\rangle \in \Sigma\} \tag{1.11}$$

We conclude this section with a lemma characterizing G_{CSS} that will be useful in future chapters.

Lemma 1.1. *The n-rebit CSS-ness preserving group G_{CSS} is*

$$G_{CSS} = \langle \bigotimes_{i=1}^{n} H_i, CNOT(i,j), X_i, Z_i \rangle \tag{1.12}$$

Proof. We sketch a simple proof of this fact and refer the reader to [10] for the completely rigorous proof. Starting with the fact that G_{CSS} is a subgroup of the Clifford group $\mathcal{C}_n$, we will look at the generators of $\mathcal{C}_n$ and eliminate the ones that do not preserve CSS-ness, while keeping the others.

Clearly, X_i and Z_i preserve CSS-ness, because if $\mathcal{S}(\psi)$ is initially a CSS stabilzer, then $X_i\mathcal{S}(\psi)X_i$, will contain the same Pauli operators, with possibly different signs depending on their commutation relation with X_i. The same thing is true for Z_i.

The $CNOT(i,j)$ gate, with qubit i as control qubit, is also in G_{CSS}, as we now show. Without loss of generality, we consider the effect of $CNOT(1,2)$ on single-qubit Pauli operators.

$$\begin{aligned}
CNOT(1,2)X_1CNOT(1,2) &= X_1X_2 \\
CNOT(1,2)X_2CNOT(1,2) &= X_2 \\
CNOT(1,2)Z_1CNOT(1,2) &= Z_1 \\
CNOT(1,2)Z_2CNOT(1,2) &= Z_1Z_2 \\
CNOT(1,2)X_iCNOT(1,2) &= X_i \quad \text{if } i > 2 \\
CNOT(1,2)Z_iCNOT(1,2) &= Z_i \quad \text{if } i > 2
\end{aligned}$$

This means that conjugating a Pauli operator with CNOT does not mix X-type and Z-type operators, so we see that if we start from a CSS state, $\mathcal{S} = \mathcal{S}_X \times \mathcal{S}_Z$, we end up with another CSS state after applying the CNOT. Therefore, all the CNOT gates are in G_{CSS}

The single-qubit Hadamard is not CSS-ness preserving. Indeed, if $|\psi\rangle$ is a n-qubit CSS state with stabilizer $\mathcal{S}(|\psi\rangle) = \langle X_1X_2, Z_1Z_2, X_3, X_4, \ldots, X_n\rangle$, then acting on this state with H_1 yields a new stabilizer $\mathcal{S}' = \langle Z_1X_2, X_1Z_2, X_3, X_4, \ldots, X_n\rangle$, which is not CSS. It is still possible that a subgroup of $\langle H_i\rangle \leq \mathcal{C}_n$ is CSS-ness preserving. All such subgroups are of order 2 and consist of multiple Hadamards acting on different qubits. It is not too hard to see that we can always construct a counterexample in a similar way as before, except for the subgroup that is generated by $\otimes_{i=1}^n H_i$. Indeed, $\otimes_{i=1}^n H_i$ transforms all X-operators into Z-operators, and vice-versa, so it preserves

CSS-ness. □

1.1.2 Phase space

The vectorial notation we used in Section 1.1.1 suggests that there is a useful relation between the Pauli operators and a binary vector space with symplectic structure. We define the phase space V as

$$V = \{(\mathbf{a}_Z, \mathbf{a}_X) | \, \mathbf{a}_Z, \mathbf{a}_X \in \mathbb{Z}_2^n\} \cong \mathbb{Z}_2^{2n} \tag{1.13}$$

We can now define a mapping π from the phase space to the Pauli operators:

$$\pi : V \to \mathcal{T} \tag{1.14}$$

$$(\mathbf{a}_Z, \mathbf{a}_X) \mapsto T_{(\mathbf{a}_Z, \mathbf{a}_X)} \tag{1.15}$$

For $\mathbf{u}, \mathbf{v} \in V$, we have two useful related relations

$$T_{\mathbf{u}+\mathbf{v}} = (-1)^{(\mathbf{u}_X, \mathbf{v}_Z)} T_{\mathbf{u}} T_{\mathbf{v}} \tag{1.16}$$

$$T_{\mathbf{u}} T_{\mathbf{v}} = (-1)^{[\mathbf{u}, \mathbf{v}]} T_{\mathbf{v}} T_{\mathbf{u}}, \tag{1.17}$$

where $[\mathbf{u}, \mathbf{v}] = (\mathbf{u}_X, \mathbf{v}_Z) + (\mathbf{u}_Z, \mathbf{v}_X)$ is the symplectic product that gives a symplectic structure to V.

Equation 1.16 tells us that the Pauli group is a projective representation of V. This is because that the mapping $\pi : V \to \mathcal{T}$ defined in Equation 1.14 is a homomorphism of groups, up to a sign. If we consider instead equivalence classes of $\mathcal{T}$ under multiplication by phases, the map induced by π is an isomorphism. This map is a first step towards using V as a phase space over which we will define the discrete Wigner function.

1.2 Wigner functions

In this section, we discuss Wigner functions, which allow the phase space formulation of quantum mechanics. We begin by surveying the field of

continuous variables Wigner functions, and then use this insight as a guide for understanding the discrete Wigner function of qudits and qubits.

1.2.1 The Wigner function for continuous variables

Wigner functions are a major tool of quantum optics and were introduced in 1932 by Wigner [29] as quantum analogues of the phase space distributions that occur in classical mechanics. We will put things in place by first discussing the case of the continuous variable Wigner function for a spinless particle. The phase space is the set of points

$$V = \{(x, p) \in \mathbb{R} \times \mathbb{R}\}. \tag{1.18}$$

In classical mechanics, the state of a point particle is fully specified by a point (or Dirac delta distribution) in phase space. However, in the quantum case, the uncertainty principle manifests itself through the impossibility of having precisely defined position and momentum simultaneously. Therefore, the density matrix of a system is represented as a distribution in the phase space. Furthermore, it turns out that for some states this distribution may take negative values at some point; this feature is called negativity and is associated with non-classical behavior, as we will soon see.

We start by defining the Wigner-Weyl transform [28], which puts a correspondence between operators on a Hilbert space and functions on phase space. The Wigner transformation takes an operator A acting on the Hilbert space to a function on the phase space V. It is defined as

$$\tilde{A}(x, p) = \int_{-\infty}^{\infty} e^{ipy/\hbar} \langle x - \frac{y}{2} | A | x + \frac{y}{2} \rangle \, dy. \tag{1.19}$$

This transformation is invertible:

$$\begin{aligned}\langle x|A|x'\rangle &= \int dy\, \delta(x-x'+y)\, \langle \frac{x+x'-y}{2}|A|\frac{x+x'+y}{2}\rangle \\ &= \frac{1}{2\pi\hbar}\iint dy\, dp\, e^{ip(x-x'+y)/\hbar}\, \langle \frac{x+x'-y}{2}|A|\frac{x+x'+y}{2}\rangle \\ &= \frac{1}{h}\int_{-\infty}^{\infty} e^{ip(x-x')/\hbar}\tilde{A}\left(\frac{x+x'}{2}, p\right) dp\,, \end{aligned} \tag{1.20}$$

and the inverse transformation is known as the Weyl transform.

The Wigner function for a state ρ is defined as the normalized Wigner transform of ρ

$$W_\rho(x,p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} \langle x+y|\rho|x-y\rangle\, e^{-2ipy/\hbar} dy \tag{1.21}$$

For a pure state $\rho = |\psi\rangle\langle\psi|$, the previous formula takes the form

$$W_\psi(x,p) = \frac{1}{\pi\hbar}\int \psi^*(x+y)\psi(x-y)e^{-2ipy/\hbar} dy, \tag{1.22}$$

where $\psi(x) = \langle x|\psi\rangle$.

For the purpose of future comparison with the discrete case, we introduce displacement operators $T(u,v)$ with $(u,v) \in \mathbb{R}^2$

$$T(u,v) = e^{i(v\hat{X}-u\hat{P})/\hbar} = e^{-i\frac{uv}{2\hbar}} e^{iv\hat{X}/\hbar} e^{-iu\hat{P}/\hbar} = e^{i\frac{uv}{2\hbar}} e^{-iu\hat{P}/\hbar} e^{iv\hat{X}/\hbar}, \tag{1.23}$$

where $\hat{X}, \hat{P}$ are the position and momentum operators, and the above inequalities were obtained by use of the Baker-Campbell-Hausdorff formula, which states that for two operators A, B such that $[[A,B],A] = [[A,B],B] = 0$, then

$$e^{A+B} = e^{-[A,B]/2} e^A e^B = e^{[A,B]/2} e^B e^A. \tag{1.24}$$

The operators $T(u,v)$ earn their name because of their action on position

eigenstates:

$$T(u,v)\,|x\rangle = e^{iuv/(2\hbar)} e^{-iu\hat{P}/\hbar} e^{iv\hat{X}/\hbar}\,|x\rangle \tag{1.25}$$

$$= e^{iuv/(2\hbar)} e^{ivx/\hbar} e^{-iu\hat{P}/\hbar}\,|x\rangle \tag{1.26}$$

$$= e^{iv(u/2+x)/\hbar}\,|x+u\rangle \tag{1.27}$$

Furthermore, we can check that the displacement obey the following composition law

$$\begin{aligned} T(u,v)T(y,z) &= \exp\left(\frac{-i\hbar}{2}(uv+yz)\right) e^{iv\hat{X}} e^{-iu\hat{P}} e^{iz\hat{X}} e^{-iy\hat{P}} \\ &= \exp\left(\frac{-i}{2\hbar}(uv+yz+2uz)\right) e^{iv\hat{X}/\hbar} e^{iz\hat{X}/\hbar} e^{-iu\hat{P}/\hbar} e^{-iy\hat{P}/\hbar} \\ &= \exp\left(\frac{i}{2\hbar}(vy-uz)\right) T(u+y, v+z) \\ &= \exp\left(-\frac{i}{2\hbar}[(u,v),(y,z)]\right) T(u+y, v+z), \end{aligned} \tag{1.28}$$

where $[\cdot,\cdot]$ is the symplectic product over $\mathbb{R}^2$. In the discrete case, we will find analogues of the displacement operators, and we will see that their composition reveals a similar symplectic structure.

It is possible to express the Wigner function in a more concise form by defining defining point operators $A(x,p)$ as

$$A(0,0) = \frac{1}{2\pi\hbar} \iint du\, dv\, T(u,v) \tag{1.29}$$

and

$$
\begin{aligned}
A(x,p) &= T(x,p)A(0,0)T(x,p)^\dagger \qquad (1.30)\\
&= \frac{1}{2\pi\hbar}\iint du\,dv\,T(x,p)T(u,v)T(-x,-p)\\
&= \frac{1}{2\pi\hbar}\iint du\,dv\,\exp\left(\frac{i}{2\hbar}(pu-xv)\right)T(u+x,v+p)T(-x,-p)\\
&= \frac{1}{2\pi\hbar}\iint du\,dv\,\exp\Bigg(\frac{i}{2\hbar}[(pu-xv)+(v+p)(-x)\\
&\qquad\qquad -(u+x)(-p)]\Bigg)T(u,v)\\
&= \frac{1}{2\pi\hbar}\iint du\,dv\,\exp\left(\frac{i}{\hbar}(pu-xv)\right)T(u,v). \qquad (1.31)
\end{aligned}
$$

Then,

$$
\begin{aligned}
\frac{1}{2\pi\hbar}\mathrm{Tr}(A(x,p)\rho) &= \frac{1}{2\pi\hbar}\int dy\,\langle y|\,\rho A(x,p)\,|y\rangle\\
&= \frac{1}{(2\pi\hbar)^2}\iiint du\,dv\,dy\,e^{i(pu-vx)/\hbar}\,\langle y|\,\rho T(u,v)\,|y\rangle\\
&= \frac{1}{(2\pi\hbar)^2}\iiint du\,dv\,dy\,e^{ipu/\hbar}e^{iv(y-x+u/2)/\hbar}\,\langle y|\,\rho\,|y+u\rangle\\
&= \frac{1}{2\pi\hbar}\iint du\,dy\,e^{iup/\hbar}\delta(y-x+u/2)\,\langle y|\,\rho\,|y+u\rangle\\
&= \frac{1}{2\pi\hbar}\int du\,e^{iup/\hbar}\,\langle x-u/2|\,\rho\,|x+u/2\rangle\\
&= \frac{1}{\pi\hbar}\int dy\,e^{-2iyp/\hbar}\,\langle x+y|\,\rho\,|x-y\rangle\,.
\end{aligned}
$$

Thus we see that

$$W_\rho(x,p) = \frac{1}{2\pi\hbar}\mathrm{Tr}\left(A(x,p)\rho\right). \qquad (1.32)$$

This is the form of the Wigner function that we will try to reproduce in the discrete finite dimensional setting.

We now present some elementary properties of the continous Wigner function, and refer the interested reader to the review [9] for the proofs.

1. W_ρ is real

2. The Wigner function sums to one:

$$\iint dx\, dp\, W_\rho(x,p) = \mathrm{Tr}(\rho) = 1 \tag{1.33}$$

3. The probability distributions for x and p are given by the marginals of the Wigner function:

$$\langle x|\rho|x\rangle = \int dp\, W_\rho(x,p) \tag{1.34}$$

$$\langle p|\rho|p\rangle = \int dx\, W_\rho(x,p) \tag{1.35}$$

4. If an observable $B = B(\hat{X})$ is purely a function of the position operator $\hat{X}$, then

$$\langle A\rangle = \iint dx\, dp\, W(x,p)A(x), \tag{1.36}$$

and similarly for observables that are pure functions of $\hat{P}$.

To better understand the properties of the Wigner function, we look at some important examples of quantum optics: coherent states and Fock states.

Coherent states [30] are eigenstates of the annihilation operator

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i}{m\omega}\hat{p}\right). \tag{1.37}$$

These states minimize the uncertainty principle, and are often thought as the states whose dynamic is the most similar to that of a classical harmonic oscillator. They are labeled by the complex number α such that

$$a\,|\alpha\rangle = \alpha\,|\alpha\rangle\,. \tag{1.38}$$

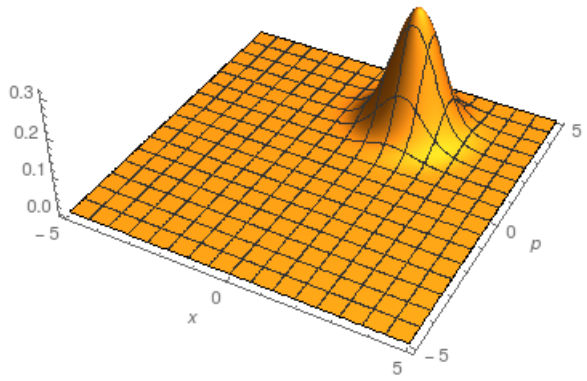


Figure 1.1: Wigner function of a coherent state $|1+2i\rangle$ in arbitrary units.

The wave function of a coherent state $|\alpha\rangle$ is

$$\psi^{(\alpha)}(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}\left(x-\sqrt{\frac{2\hbar}{m\omega}}\Re[\alpha]\right)^2+i\sqrt{\frac{2m\omega}{\hbar}}\Im[\alpha]x}. \tag{1.39}$$

With this, we can directly use Equation 1.22 to calculate the Wigner function of a coherent state. Figure 1.1 shows the Wigner function of coherent state, using natural units such that $m\omega = 1$. The Wigner function of a coherent state is a Gaussian centered in $(x, p) = (\Re[\alpha], \Im[\alpha])$. Note that the Wigner function of a coherent state is non-negative over all the phase space. This property allows us to identify these states as classical states, because the Wigner function can be interpreted as a classical probability distribution.

Another important class of states in quantum optics is the Fock states. These states are eigenstates of the number operator $N = a^\dagger a$. Figure 1.2 shows the Wigner function of the state $|1\rangle$. Notice that the Wigner function for this state takes negative values, with maximal negativity occurring at the origin.

A positive Wigner function can be interpreted as a probability distribution, an thus as a hidden variable model (more about this in section 1.4). It

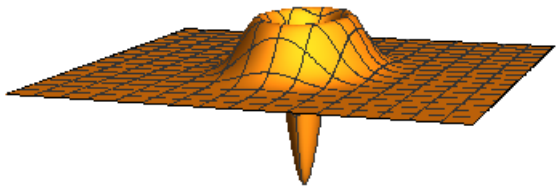


Figure 1.2: Wigner function of the Fock state $|1\rangle$. Units have been chosen such that $\hbar = 1$, and $m\omega = 1$.

is therefore common to interpret negativity of the Wigner function as a non-classical feature. This agrees with the previous observation that Fock states have negativity in their Wigner function, and with the typical interpretation of these states as particle-like, a distinctively quantum feature.

However, we must be careful when identifying non-negativity of the Wigner function with classicality. It has been shown [2], that the EPR state [11] has non-negative Wigner function, but it can nonetheless be used to violate a Bell inequality.

The previous discussion shows that some pure states have non-negative Wigner functions, while others can take negative values at some points in the phase space. The complete set of pure states with non-negative Wigner function is identified by Hudson's theorem [17] and it's multi-particle generalization by Soto and Claverie [24].

Theorem 1.2. *(Hudson's Theorem) The Wigner function of a pure state $\psi \in L^2(\mathbb{R}^n)$ is non-negative if and only if ψ is Gaussian, that is, if its wave function has the form*

$$\psi(x) \propto e^{2\pi i(x\theta x + \omega x)}, \tag{1.40}$$

where $x, \omega \in \mathbb{R}^n$ and θ is a complex symmetric matrix.

A positive Wigner function can also be interpreted as classical in the sense of computation. Indeed, there exists [3] an equivalent of the Gottesman-Knill theorem for continuous variables. Any computing machine acting on

an initial unentangled Gaussian state (whose Wigner function is positive by Hudson's theorem) with gates generated by Hamiltonians that are quadratic in the canonical operators $\hat{X}$ and $\hat{P}$ (these gates send Gaussian states to Gaussian states), and where the allowable measurements are canonical operators can be efficiently classically simulated.

The question of whether there exists equivalent theorems in the case of discrete phase spaces is the starting point for what follows.

1.2.2 The qudit Wigner function

In this chapter, we survey the previously known results for the case of qudits: d-level systems, where d is an odd prime. We give the statement of the discrete Hudson's theorem, and detail the link between universality, contextuality and Wigner function negativity. These results are valid for virtually all discrete quantum systems, except the most paradigmatic one in quantum computation, systems of qubits. The remaining chapters of this thesis are largely motivated by what is know for qudits, and we shall make frequent reference to the results presented here.

For the remaining of this section, we will assume that d is an odd prime number. The definition of the Wigner function we present here is that of Gross [14], with a somewhat adapted notation. For systems of dimension d, the standard basis vectors are $\{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$. Here $\{0, \ldots, d-1\}$ are elements of the field $\mathbb{Z}_d$, in which addition and multiplication are defined modulo d. The analogs of the Pauli operators in dimension d are defined by their action on the standard basis vectors as

$$X|x\rangle = |x+1\rangle \qquad\qquad Z|x\rangle = \omega^x |x\rangle, \tag{1.41}$$

where $\omega = e^{2\pi i/d}$ is the d-th root of unity. Using the X and Z operators as building blocks, we define the Weyl (or Heisenberg-Weyl operators) as

$$T_{(a_z,a_x)} = \omega^{2^{-1}a_z a_x} Z^{a_z} X^{a_x}. \tag{1.42}$$

The set $\{T_{(a_z,a_x)}\}$ is only closed under multiplication up to phases, it is

therefore not a group.

Note the appearance of 2^{-1}, the inverse of 2 in the above expression. This quantity does not exist in dimension 2, and it is the source of most if not all of the mathematical complications that occur for qubits. Although this "detail" might seem surprising to the eyes of a physicist, fields of characteristic two are well known by mathematicians to be pathological in various ways. The characteristic of a field F is the smallest integer n such that $1+1+\cdots+1=0$, with the sum on the left containing n terms. Two simple properties of fields of characteristic two that do not apply for other fields are that $x^{-1}=-x$, $\forall x \in F$ and the fact that 1 only has one square root.

The construction of the Heisenberg-Weyl operators is readily generalized to systems of multiple qudits. In this case, the phase space $V=\mathbb{Z}_d^n \times \mathbb{Z}_d^n$ consists of the vectors $\mathbf{a}=(\mathbf{a_z},\mathbf{a_x})=(a_z^1,a_z^2,\ldots a_z^n,a_x^1,\ldots,a_x^n)$, and the Heisenberg-Weyl operators are defined as

$$T_{\mathbf{a}} = T_{(a_{z^1},a_{x^1})} \otimes T_{(a_{z^2},a_{x^2})} \otimes \cdots \otimes T_{(a_{z^n},a_{x^n})}. \tag{1.43}$$

The phase space $\mathbb{Z}_d^n \times \mathbb{Z}_d^n$ naturally inherits a symplectic structure from the composition of the Weyl operators, because

$$T_{\mathbf{u}} T_{\mathbf{v}} = \omega^{2^{-1}[\mathbf{u},\mathbf{v}]} T_{\mathbf{u}+\mathbf{v}}, \tag{1.44}$$

where

$$[\mathbf{u},\mathbf{v}] = (\mathbf{u}_z,\mathbf{v}_x) - (\mathbf{u}_z,\mathbf{u}_x) \tag{1.45}$$

is the symplectic product with addition and multiplication modulo d.

From Equation 1.44 and the antisymmetry of the symplectic product, we have

$$T_{\mathbf{u}} T_{\mathbf{v}} = \omega^{[\mathbf{u},\mathbf{v}]} T_{\mathbf{v}} T_{\mathbf{u}}. \tag{1.46}$$

The Wigner function for a density operator ρ is the symplectic Fourier transform of the coefficients of ρ in the basis of the Heisenberg operators. More concretely, it's the coefficient of the expansion of ρ in terms of point-

operators $A_{\mathbf{u}}$, defined as

$$A_{\mathbf{0}} = \frac{1}{d^n} \sum_{\mathbf{a} \in V} T_{\mathbf{a}} \tag{1.47}$$

$$A_{\mathbf{u}} = T_{\mathbf{u}} A_{\mathbf{0}} T_{\mathbf{u}}^{\dagger}. \tag{1.48}$$

Therefore, the Wigner function W_ρ of ρ is

$$W_\rho(\mathbf{u}) = \frac{1}{d^n} \mathrm{Tr}(\rho A_{\mathbf{u}}). \tag{1.49}$$

Properties of the qudit Wigner function

It can be shown [14] that the Wigner function has the following properties

1. W_ρ is real and sums to one: $\sum_{\mathbf{u}} W_\rho(\mathbf{u}) = 1$.
2. For all density matrices ρ and σ, $W_{\rho \otimes \sigma} = W_\rho \cdot W_\sigma$.
3. W is informationally complete: since $\rho = \sum_{\mathbf{u}} W_\rho(\mathbf{u}) A_{\mathbf{u}}$, knowing the Wigner function for a state is equivalent to knowing the state
4. $A_{\mathbf{0}} |\mathbf{x}\rangle = |-\mathbf{x}\rangle$, $\forall \mathbf{x} \in \mathbb{Z}_d^{2n}$.
5. Discrete Hudson's Theorem: the only pure states with non-negative Wigner function are stabilizer states

Remark. In dimension $d = 2$, all vectors of the phase space obey $|\mathbf{x}\rangle = |-\mathbf{x}\rangle$. Therefore Property 4 does not hold for qubits since it would imply $A_{\mathbf{0}} = I$.

Clifford covariance

Some unitary gates have a relatively simple effect on the quantum state of a system. This is captured in the definition of Wigner function covariance.

Definition 1.3. A Wigner function W_ρ is covariant under a group G if for any $g \in G$ and for all states ρ, there exists a permutation map $\theta_g : V \to V$

such that

$$W_{g\rho g^\dagger}(\mathbf{u}) = W_\rho(\theta_g(\mathbf{u})). \tag{1.50}$$

We will use this definition to study stabilizer quantum computation for qudits. The stabilizer formalism of Section 1.1.1 generalizes readily to qudits, by defining the Clifford group as the group that preserves generalized Pauli operators (Weyl operators). If U belongs to the Clifford group,

$$UT_{\mathbf{u}}U^\dagger = c(\mathbf{u})T_{S\mathbf{u}}, \tag{1.51}$$

where $c : V \to \mathbb{C}$ is a map whose image has values ω^x for $x \in \mathbb{Z}_d$, and $S : V \to V$ is a permutation of the phase space. It is a fact [14], that for any Clifford unitary U, there exists a symplectic matrix F and a vector $a \in V$ such that

$$UT_{\mathbf{u}}U^\dagger = T_{F\mathbf{u}+\mathbf{a}}. \tag{1.52}$$

For qubits, this property only holds up to a sign ± 1. It possible to use Equation 1.52 to show that the Wigner function is covariant under the Clifford group. Let U be a Clifford operation, then there exists [14] a symplectic matrix F and a vector $\mathbf{a}$ such that

$$W_{U\rho U^\dagger}(\mathbf{u}) = W_\rho(F^{-1}(\mathbf{v} - \mathbf{a})). \tag{1.53}$$

The Clifford covariance of the Wigner function shows that it has a high level of symmetry. It is possible, as was pointed out by Gibbons et al. [12], to define many different Wigner functions. However there is an important theorem, proved by Gross [14] that shows that the definition we have chosen is the only Wigner function possessing Clifford covariance. We give it's full content here for clarity

Theorem 1.4. *Consider another Wigner function W' that satisfies the following properties*

1. *W' is a linear mapping on the space of density matrices: $W'_{\rho+\sigma} = W'_\rho + W'_\sigma$.*

2. *W' is covariant under the action of the Clifford group.*

3. *W' gives the correct marginal probabilities.*

Then $W' = W$.

In this sense, it can be said that the definition of the Wigner that we have presented here is the most natural one.

Negativity as a resource

It is known through the Gottesman-Knill theorem that quantum circuits consisting of only Clifford gates acting on stabilizer states can be simulated classically. Veitch et al. [27] have expanded this result to all (pure and mixed) states that have positive Wigner functions.

Theorem 1.5. *Quantum circuits consisting of Clifford gates acting on an initial state of qudits with positive Wigner function can be efficiently simulated by a classical computer*

The proof of the previous is very similar to the efficient simulation proof that we will give in Chapter 3, so we do not review it here. The proof relies crucially on the property that the Wigner function is covariant under Clifford gates. This covariance allows us to classically track the evolution of the quantum state by calculating the associated phase space permutation after each gate in the computation.

This theorem identifies negativity as a resource in qudit stabilizer quantum computation, since it cannot be created using the allowed gate set. In order to acheive computational, negatively represented "magic states" need to be introduced; we shall have more to say about this process in Section 1.3

1.2.3 The qubit Wigner function

We will now go through a heuristic "derivation" of the qubit Wigner function, and attempt to show why some desirable properties, such as Clifford covariance, fall short in this case. For qubits, the phase space is discrete,

and it is the V that was described in equation 1.13. Our goal is to describe the density matrix of our system as a real-valued distribution on V. Using the fact that Pauli operators $\mathcal{T}$ are a basis for the complex space of matrices, we can write the density matrix of the system as

$$\rho = \sum_{\mathbf{u}} \rho_{\mathbf{u}} T_{\mathbf{u}}, \tag{1.54}$$

where the $\rho_{\mathbf{u}}$ are generally complex coefficients if we consider states of qubits. The idea behind the Wigner function is to find another basis $(A_{\mathbf{u}})$ for ρ, whose coefficients are real and whose total sum over the phase space is one. The expansion coefficients of ρ in this basis are taken to be the definition of the Wigner function W_ρ

$$\rho := \sum_{\mathbf{u}} W_\rho(\mathbf{u}) A_{\mathbf{u}}. \tag{1.55}$$

The requirement that W_ρ is real and ρ hermitian forces the operators $A_{\mathbf{u}}$ to be hermitian. The $A_{\mathbf{u}}$'s are called point operators. If we further require that the $A_{\mathbf{u}}$'s are orthonormal, we can invert equation 1.55 to obtain

$$W_\rho(\mathbf{u}) = \mathrm{Tr}(A_{\mathbf{u}} \rho). \tag{1.56}$$

One possible way to define the point operators is

$$A_{\mathbf{0}} = \frac{1}{2^n} \sum_{\mathbf{u}} (i)^{a_{\mathbf{u}}} T_{\mathbf{u}} \qquad\qquad A_{\mathbf{u}} = T_{\mathbf{u}} A_{\mathbf{0}} T_{\mathbf{u}}^\dagger. \tag{1.57}$$

The multiplicative factor $(i)^{a_{\mathbf{u}}}$ is there to enforce hermiticity of the point operators. It is defined as

$$a_{\mathbf{u}} = \begin{cases} 0 & \text{if } (\mathbf{u}_X, \mathbf{u}_Z) = 0 \\ 1 & \text{if } (\mathbf{u}_X, \mathbf{u}_Z) = 1 \end{cases} \tag{1.58}$$

We stress that there are many different ways to define a Wigner function, corresponding to different choices of $A_{\mathbf{0}}$, and they have been characterized by

Gibbons et al. [12]. For certain applications, however, we will be interested in using Wigner functions that are particularly "symmetrical" in the sense of covariance, a property that we have already seen in the context of the qudit Wigner function.

The qubit Wigner function that we have defined enjoys the nice property of covariance under Pauli operations. In fact, elements of $\mathcal{T}$ have the effect of translating the phase space:

$$\begin{aligned} W_{T_{\mathbf{v}}\rho T_{\mathbf{v}}^\dagger}(\mathbf{u}) &= \frac{1}{2^n}\mathrm{Tr}(T_{\mathbf{u}}A_{\mathbf{0}}T_{\mathbf{u}}^\dagger T_{\mathbf{v}}\rho T_{\mathbf{v}}^\dagger) \\ &= \frac{1}{2^n}\mathrm{Tr}(T_{\mathbf{v}}^\dagger T_{\mathbf{u}}A_{\mathbf{0}}(T_{\mathbf{v}}^\dagger T_{\mathbf{u}})^\dagger\rho) \\ &= \frac{1}{2^n}\mathrm{Tr}(T_{\mathbf{v}}T_{\mathbf{u}}A_{\mathbf{0}}(T_{\mathbf{v}}T_{\mathbf{u}})^\dagger\rho) \\ &= \frac{1}{2^n}\mathrm{Tr}(T_{\mathbf{u}+\mathbf{v}}A_{\mathbf{0}}(T_{\mathbf{u}+\mathbf{v}})^\dagger\rho) \\ &= W_\rho(\mathbf{u}+\mathbf{v}), \end{aligned}$$

where in the second to third line of the previous calculation we used that $T_{\mathbf{v}}^\dagger = \pm T_{\mathbf{v}}$.

There is a more powerful and desirable kind of covariance, called Clifford covariance which is possessed by the qudit Wigner function, as we have seen in Section 1.2.2. Clifford gates send Pauli operators to other Pauli operators, and furthermore they preserve the commutation relations. In light of Equation 1.17, preserving the commutation relations of the Pauli operators is equivalent to preserving the symplectic product over the phase space. Therefore, for any $g \in \mathcal{C}_n$, there exists a symplectic matrix F such that

$$gT_{\mathbf{u}}g^\dagger = \pm T_{F_{\mathbf{u}}}. \tag{1.59}$$

Symplectic matrices are matrices that preserve the symplectic product:

$$[F\mathbf{u}, F\mathbf{v}] = [\mathbf{u}, \mathbf{v}]. \tag{1.60}$$

We are led to wonder if there exists a permutation of the phase space that describes the action of Clifford operations, as it was the case for qudits

in Section 1.2.2. In that section we established that Clifford covariance was a desirable property for the Wigner function, because it leads to an efficient classical simulation method. However, it is a mathematical fact that there exists no definition of the Wigner function that is covariant under the Clifford group for systems of qubits [31]. We can show this easily for the function W_ρ that we have defined.

Lemma 1.6. *The qubit Wigner function W_ρ of Equation 1.56 is not Clifford-covariant*

Proof. The proof proceeds through counterexample, by considering a one-qubit state ρ undergoing transformation by the Clifford unitary H. The most general form of a one qubit state is

$$\rho = \frac{1}{4}(1 + r_x X + i r_y ZX + r_z Z), \tag{1.61}$$

with $r_x^2 + r_y^2 + r_z^2 < 1$. After the Hadamard gate, the state is

$$\rho \to H\rho H = \frac{1}{4}(1 + r_z X - i r_y ZX + r_x Z). \tag{1.62}$$

We can compute the Wigner function of the state before and after the Hadamard gate

$$\begin{aligned}
W_\rho(0,0) &= \frac{1}{4}(1 + r_x + r_y + r_z) \quad & W_{H\rho H}(0,0) &= \frac{1}{4}(1 + r_x - r_y + r_z) \\
W_\rho(1,0) &= \frac{1}{4}(1 + r_x - r_y - r_z) \quad & W_{H\rho H}(1,0) &= \frac{1}{4}(1 - r_x + r_y + r_z) \\
W_\rho(0,1) &= \frac{1}{4}(1 - r_x + r_y + r_z) \quad & W_{H\rho H}(0,1) &= \frac{1}{4}(1 + r_x + r_y - r_z) \\
W_\rho(1,1) &= \frac{1}{4}(1 - r_x + r_y - r_z) \quad & W_{H\rho H}(1,1) &= \frac{1}{4}(1 - r_x - r_y - r_z)
\end{aligned}$$

.

Looking at the above equations for W_ρ and $W_{H\rho H}$, we see that for general one qubit states, there is no permutation of the phase space that sends the first function to the other. Indeed, the function for $W_{H\rho H}(1,1)$ contains three minus signs, while there is no phase space point such that has this

property for W_ρ. □

Is there a way to obtain some limited form of covariance? We can achieve this by setting one of the $r_i = 0$ in Equation 1.61. In this thesis, we will choose to set $r_y = 0$, in which case ρ is a real state, or rebit. If we consider a general states of n rebits, it's expansion in terms of the Pauli operators is

$$\rho = \sum_{\mathbf{a}\in\mathcal{A}} \rho_{\mathbf{a}} T_{\mathbf{a}}, \tag{1.63}$$

with $\rho_{\mathbf{a}} \in \mathbb{R} \forall \mathbf{a}$ and $\mathcal{A} = \{\mathbf{u} \in V | \, (\mathbf{u}_X, \mathbf{u}_Z) = 0\}$. Therefore, we can drop some terms from the definition of $A_{\mathbf{0}}$ in Equation 1.64 without consequence, and obtain

$$A_{\mathbf{0}} = \frac{1}{2^n} \sum_{\mathbf{a}\in\mathcal{A}} T_{\mathbf{a}} \qquad\qquad A_{\mathbf{u}} = T_{\mathbf{u}} A_{\mathbf{0}} T_{\mathbf{u}}^\dagger. \tag{1.64}$$

This definition for the point operators will be the starting point for our discussion of the rebit Wigner function in Chapter 3. In particular, we will establish that this Wigner function is covariant under the subgroup G_{CSS} of the Clifford group.

1.2.4 Some properties of general Wigner functions

We list here some properties held generally by Wigner functions $W_\rho : V \to \mathbb{R}$

1. The Wigner function is a quasi-probability distribution, so it sums to one
$$\sum_{\mathbf{u}\in V} W_\rho(\mathbf{u}) = 1. \tag{1.65}$$
2. W_ρ is informationally complete, in the sense that $\rho = \sum_{\mathbf{u}} W_\rho(\mathbf{u}) A_{\mathbf{u}}$é
3. For two state ρ and σ, their overlap (trace inner product) is
$$\mathrm{Tr}(\rho\sigma) = 2^n \sum_{\mathbf{u}} W_\rho(\mathbf{u}) W_\sigma(\mathbf{u}). \tag{1.66}$$

If $\sigma = |\psi\rangle\langle\psi|$ is a stabilizer state associated to a line L, this equation reduces to the probability that ρ is in the state $|\psi\rangle$

$$\mathrm{Tr}(\rho\,|\psi\rangle\langle\psi|) = \sum_{\mathbf{u}\in L} W_\rho(\mathbf{u}). \tag{1.67}$$

We will prove these properties, and some others, for the rebit Wigner function in Chapter 3.

1.3 Magic states and computation by state injection

As we have mentioned in Section 1.1.1, the set of Clifford gates acting on stabilizer states is classically simulable. If we hold the typical assumption that a general purpose quantum computer is not simulable classically, we are naturally lead to wonder how we can supplement this operational restriction and achieve universal quantum computation. One possible way is to introduce a single qubit gate that does not belong to the Clifford group, such as the T gate [20]. Here we will focus on another road to universality known as quantum computation by state injection (QCSI), and the closely related idea of magic state distillation.

This approach is twofold. First we show that given a supply of perfect copies of a certain quantum state as ancillas, and by only using Clifford gates, we can realize a non-Clifford single qubit gate. These states have the property of "unlocking" universal quantum computation, which is why they are called magic states in the literature. Then, we show that some magic states have the other desirable property that a large number of imperfect copies of a target magic state can be purified to obtain a smaller number of better copies of that target.

We first review a familiar construction in the quantum computation community [20] that allows to realize the family of gates

$$\Lambda(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix} \tag{1.68}$$

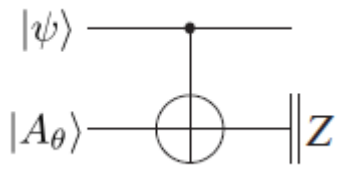


Figure 1.3: Circuit for a state-injection that yields $\Lambda(\theta)\,|\psi\rangle$ or $\Lambda(-\theta)\,|\psi\rangle$ with equal probabilities

by state injection using the magic state $|A_\theta\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\theta}\,|1\rangle\right)$. This family of gates contains $|A_{\pi/4}\rangle := |H\rangle$ and $|A_{\pi/8}\rangle := |T\rangle$, which allow to realize the H and T gates, and obtain universal quantum computation. The circuit for state injection is shown in figure 1.3, and can easily be shown to produce $\Lambda(\theta)\,|\psi\rangle$ or $\Lambda(-\theta)\,|\psi\rangle$ with equal probabilities.

A question that naturally occurs in the context of fault-tolerant quantum computation is: can this scheme work if our ancilla states are not perfect? Bravyi and Kitaev answered this question positively by inventing the scheme of magic state distillation. Broadly, the scheme proceeds as follows.

1. Start with a large number of imperfect states $\rho^{\otimes n}$
2. Perform Clifford gates and measurements
3. Obtain a smaller number of states $\rho'^{\otimes m}$ with improved fidelity.

It has been shown that $|T\rangle$ and $|H\rangle$ states are distillable efficiently, but as the constructions rely heavily on the theory of quantum error correction, we will refer the interested reader to their original paper [7].

Does magic state distillation also work for rebits? We address this question in Section 2.3

1.4 Hidden variable models and contextuality

Many of the insights on the peculiar nature of quantum mechanics have been obtained by the study of hidden variable models. These ideas go back

to the Einstein-Podolsky-Rosen paradox [11] in which the authors famously argued that quantum mechanics was incomplete; their assumption of a local hidden variable model describing quantum theory led to "spooky action at a distance". It was latter proved by Bell in a seminal paper [4], that no local hidden variable model can reproduce the predictions of quantum mechanics. Subsequent experimental work has largely (but as of now not completely) concluded in favor of the formalism of quantum mechanics over local hidden variable models. However, there is another feature that we would expect reasonable hidden variable models to exhibit, but that is also violated by quantum mechanics: contextuality. This section aims to be a quick but self-contained introduction to the aspects of hidden variable models and contextuality that are relevant for the content of this thesis.

1.4.1 Hidden variables

Contrary to the usual Copenhagen interpretation of quantum mechanics, in which measurements outcome are brought into being by the act of measurement, hidden variable models (HVMs) work under the reasonable assumption that physical systems have well-defined values of their properties before a measurement is performed on the systems. Thus, in such models we assume quantum mechanics to be incomplete, and augment it with hidden-variables: a set of values $\{\lambda_i\}$ that describe the definite outcome of any measurement that can be performed on the system. In what follows the values $\{\lambda_i\}$ describing a particular system will sometimes be referred to as the ontic state of the system. To account for the fact that the standard formalism of quantum mechanics has always been vindicated by experiment, we require that on a set of identically prepared particles, the values λ_i are sampled from a probability distribution such that all statistical predictions (expectation values of observables, correlations between observables, etc.) agree with quantum mechanics.

In quantum mechanics, as measurable quantity is represented as a hermitian operator, and the possible outcomes of the measurement are the eigenvalues of the operator. The state of the system after a measurement

is given by the projection postulate; it is the (normalized) projection of the original state onto the eigenspace corresponding to the measurement result. Mutually commuting operators are said to be simultaneously measurable, because they have simultaneous eigenvectors. For our purposes, we define a context to be a set $\{M_i\}$ of mutually commuting observables. A context can therefore physically be identified with a possible measurement apparatus, which yields a set of compatible measurement outcomes. We note that there exists more general definitions of contexts [25], which include ours as a special case. This discussion has set the stage for some definitions

Definition 1.7. A physical setting is a pair $(\rho, \mathcal{M})$, where ρ is the density operator describing system and $\mathcal{M}$ is a set of measurement contexts such that all observables O belonging to a context $M \in \mathcal{M}$ are mutually commuting and can be simultaneously measured.

Definition 1.8. A hidden variable model for a physical setting $(\rho, \mathcal{M})$ is a set of internal states $\mathcal{S}$, together with a set value assignment functions $\{\lambda_{\mathbf{u}} : \mathcal{O} \to \mathbb{R} | \, \mathbf{u} \in \mathcal{S}\}$, where $\mathcal{O}$ is the set of observables that appear in $\mathcal{M}$. The internal states should follow a probability distribution such that all statistical prediction of quantum mechanics are reproduced by the HVM.

Definition 1.9. A non-contextual hidden variable model (NCHVM) is a HVM such that for two contexts C_1 and C_2, the values associated to an observable A such that $A \in C_1$ and $A \in C_2$ is independent on which of the two contexts is measured.

Definition 1.10. A physical setting is contextual if it cannot be described by a non-contextual hidden variable model.

As a more concrete example, suppose we have three observables A, B, C such that $[A, B] = [A, C] = 0$, but $[B, C] \neq 0$. Then $C_1 = \{A, B\}$, $C_2 = \{A, C\}$ are two contexts, and a hidden variable model for this system will be non-contextual if the value assigned to A is independent of whether A is measured jointly with B or with C. We note that our definition makes explicit reference to quantum theory, and that it has been substantially generalized by Spekkens [25]. Using the language introduced by that paper, we

make the assumption of outcome determinism, meaning that the a physical system is in a well-defined ontic state, in opposition with it being given by a probability distribution over ontic states.

The Definition 1.7 of a physical setting has the virtue of further restricting the measurement contexts that need to be replicated by the HVM than to the requirement that contexts are sets of mutually commuting observables. Indeed, if for some operational reason we are unable to simultaneously measure a pair of commuting observables A and B, then we will not require our NCHVM to account for measurements containing A and B.

1.4.2 Algebraic constraints on the value assignments

A HVM comes with a set of value assignment functions $\lambda_{\mathbf{u}} : \mathcal{O} \to \mathbb{R}$, where $\mathbf{u} \in \mathcal{S}$ is an internal state of the system and $\mathcal{O}$ is the set of observables on the Hilbert space of the system. Since the only possible outcomes for a measurement of an observable $O \in \mathcal{O}$ are eigenvalues of that observable, this limits the value assignments to $\lambda_{\mathbf{u}}(O) \in \mathrm{Eig}(O)$. Furthermore, if two observables A and B commute, and if $|\psi\rangle$ is a simultaneous eigenvector of both operators with eigenvalues a and b, then $AB\,|\psi\rangle = ab\,|\psi\rangle$. This imposes a consistency constraint on the value assignments:

$$\lambda_{\mathbf{u}}(AB) = \lambda_{\mathbf{u}}(A)\lambda_{\mathbf{u}}(B). \tag{1.69}$$

By the same reasoning, for any set of commuting observables, functional identities of operators $f(A, B, C, ...) = 0$ carry on to the value assignments so that $f(\lambda_{\mathbf{u}}(A), \lambda_{\mathbf{u}}(B), \lambda_{\mathbf{u}}(C), ...) = 0$.

Commuting observables are said to be simultaneously measurable because they have simultaneous eigenvectors, and because the formalism of quantum mechanics gives the same predictions no matter what the order of the measurements in the set is. However, it is not immediately obvious what this requirement of simultaneous measurability means when considering a HVM. It is conceivable that the ontic state of the system might change after the measurement of one of the observables in the setting, such that relation 1.69 would not necessarily hold at the beginning of the measurement proce-

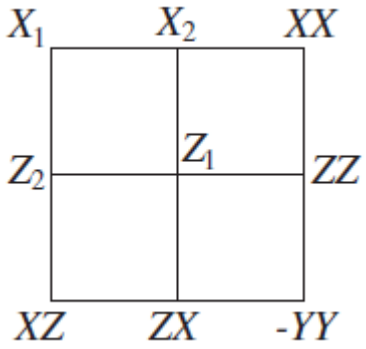


Figure 1.4: The Mermin square.It is impossible to have a consistent value assignment for the observables of the square

dure. However, the usefulness of this constraint is saved by noticing that for finite contexts containing observables with a finite number of eigenvalues, it is always possible to construct a single observable whose measurement would unambiguously determine the eigenvalues of all the other observables in the context. Namely, if the context consists of observables $\{A_i\}_{i=1..n}$ with eigenvalues ranging for 1 to N, then we can construct the observable

$$O = \sum_{i=1}^{n} N^i A_i, \tag{1.70}$$

whose eigenvalue is an N-ary number that represents the set of eigenvalues for the context.

1.4.3 Contextuality proofs

If we consider a system consisting of a single qubit, then it is possible to explicitly construct a NCHVM that reproduces all predicitons of quantum mechanics [4]. Contextuality first arises in systems of dimension 3, as was first demonstrated by Kochen and Specker [18]. The proof is relatively complicated and we do not reproduce it here.

There exists a very simple proof of contextuality due to Mermin [19] which considers a system formed by two qubits. We present this proof

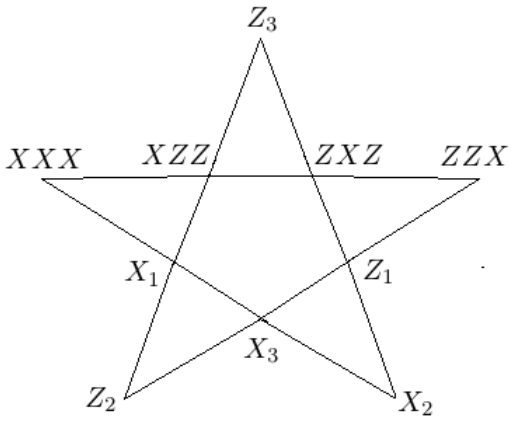


Figure 1.5: Observables for which it is impossible to have a consistent value assignment

in a slightly modified version, to make it more readily applicable to the arguments of Chapter 4. The proof is purely algebraic and reaches a contradiction by showing that it is impossible to satisfy the value constraints of Equation 1.69 for the set of nine observables shown in Figure 1.4, known herein as the Mermin square. It can be easily checked that every row and column is formed of mutually commuting observables, and that the product of the observables in every row and column is $+I$, except for the right column which multiplies to $-I$. Consistency of the value assignments gives us 6 equations of the form $\lambda(X_1)\lambda(X_2)\lambda(X_1X_2) = 1$, etc. constraining the values in the HVM. Multiplying these six equations together and noticing that each observable appears in exactly two equations, we get

$$\prod_{A\in\mathcal{O}} \lambda(A)^2 = -1. \tag{1.71}$$

This is a contradiction, because the fact that Pauli operators have eigenvalues ± 1 implies that $\lambda(A)^2 = 1$ for every observable. We remark that this proof made no mention of a particular state, it is thus an example of something called state-independent contextuality.

The previous argument uses some non-local observables, for instance

$X_1 Z_2$, but there exists another proof of the Kochen-Specker theorem which relies only on local observables. In such a proof the assumption of non-contextuality might be more convincing, because it is guaranteed by locality. Indeed, if the individual particles are space-like separated, then it is impossible for a measurement at one site to affect the values at another site. Consider the observables of the Mermin star in Figure 1.5, and assume that the state of the system is in an eigenstate of the (commuting) observables belonging to the horizontal row. Since the state is an eigenstate, the four value assignments for the operators on the horizontal row are fixed to the associated eigenvalue. All the other observables, whose properties under measurement need to be replicated by the HVM, are local Pauli operators. The rest of the proof proceeds in an analogous way as Mermin's square. We can check that the five lines of the star are formed of commuting observables, and that the product of the observables on every line is $+1$, except for the horizontal line whose product is -1. Multiplying the five equations together yields a similar contradiction as in the Mermin square. The price we had to pay for the limitation to local observables is that this proof is constructed to work only for one state, hence it is called state-dependent contextuality. Note that if we allow ourselves to measure non-local observables, we can directly get a state-independent proof of contextuality from the Mermin star.

We conclude this brief introduction to HVM's by pointing out that hidden variable models that reproduce quantum mechanics are indeed possible, if one drops the assumption of non-contextuality. A well-known early example of such a model is Bohm's pilot wave theory [6].

The discussion of the two Mermin proofs shows that there is a fine line between state-independent and state-dependent contextuality, and that this demarcation has to do with the operations we allow ourselves to carry on the physical system. We discuss these issues further in following section.

1.4.4 Contextuality of subtheories of quantum mechanics

In the previous paragraphs we saw that quantum theory exhibits state independent contextuality. The arguments given above are quite general, as for any system with total dimension higher than four, we can pick a 4-dimensional subspace on which the Mermin square proof applies directly. However, if for a reason or another we have an operational restriction (only a certain subset of gates, states or measurements are allowed), we might be able to construct a non-contextual hidden variable model describing every possible measurement outcome. As a trivial example of such a situtation, consider a system of qubits, initially in state $|000\cdots0\rangle$ on which we are only able to perform X and Z gates, and Z measurements. Since we do not have the capability to perform entangling gates, every qubit can be independently described by the non-contextual HVM of Bell [4] to which we have already referred to.

In the field of fault-tolerant quantum computation [20], certain quantum gates are considered "cheap" because they can be implemented fault-tolerantly in a relatively simple way. One prominent example of such a set of gates occurs in stabilizer quantum computation, which we have discussed in Section 1.1.1, and where the cheap gates are $\langle X_i, Z_i, H_i, CNOT(i,j)\rangle$. As now know, stabilizer quantum computation is not universal, and in order to achieve universality, we need to resort the scheme of QCSI.

Restricting ourselves to a such a subtheory of quantum mechanics, it is conceivable that the subtheory could be described by a NCHVM. Suppose for the following that we have an operational restriction that yields a non-contextual subtheory of quantum mechanics. Then it is immediate that our restricted set of gates will not allow universal quantum computation. Indeed, if it did, we could achieve exactly the measurements of the Mermin square proof and contradict our assumption of non-contextuality.

Howard et al. [16] have expanded on the previous results of Veitch et al. [27] by considering the role of contextuality as a resource in qudit QCSI. In their paper, drawing from an earlier result of Spekkens [26], they show that negativity and contextuality are equivalent notions. Therefore they establish

contextuality of the magic state as a resource for quantum computation. An important point to note is that for qudits, there exists no state-independent contextuality proof involving only measurements of generalized Pauli operators. This is clearly not the case for qubits, as evidenced by the Mermin square and star.

The advantage of the QCSI scheme for fundamental study of the resources of quantum computation is that we put the computational power on the magic states. Whether a state is contextual or not completely depends on the operational restriction we impose on ourselves. One of the main questions we shall ask in Chapter 4 is: "does contextuality imply universality in rebit QCSI?", and the answer will be positive.

Chapter 2

Universal quantum computation by state injection with rebits

2.1 Rebits

Just as real numbers can ben seen as a subset of the complex numbers, it is possible to define rebits as the subset of quantum states of qubits that have density matrices with real entries in the computational basis. If ρ is a n-rebit state, then for every computational basis states $|x\rangle, |y\rangle$ (± 1 eigenstates of $Z_1 \otimes \cdots \otimes Z_n$), it will obey

$$\langle x|\rho|y\rangle \in \mathbb{R}. \tag{2.1}$$

The full Clifford group does not preserve rebit states, as can be easily seen from $Y|0\rangle = i|1\rangle$. The subgroup of the Clifford group that preserves realness is G_{CSS}, which we defined in Equation 1.11.

2.2 Universal quantum computation

Universal quantum computation, in the more traditional case of qubits, refers to the ability to approximate any unitary gate using number of elementary gates that scales polynomially with the desired accuracy [20]. It has been shown [22], that through the use of a simple encoding, that rebits are universal for quantum computation. The encoding consists of mapping a n-qubit state to a state of $n+1$ rebits, with the additional rebit coding for the real and imaginary parts of the initial state. More explicitly, starting from a general n-qubit state

$$|\psi\rangle = \sum_{\mathbf{v}\in\mathbb{Z}_2^n} r_{\mathbf{v}} e^{i\theta_{\mathbf{v}}}, |\mathbf{v}\rangle \tag{2.2}$$

the corresponding encoded state is

$$\overline{|\psi\rangle} = \sum_{\mathbf{v}\in\mathbb{Z}_2^n} (r_{\mathbf{v}} \cos\theta_{\mathbf{v}} |\mathbf{v}\rangle \otimes |R\rangle + r_{\mathbf{v}} \sin\theta_{\mathbf{v}} |\mathbf{v}\rangle \otimes |I\rangle). \tag{2.3}$$

Here $|R\rangle = |0\rangle$ and $|I\rangle = |1\rangle$ is just notation that makes explicit the real and imaginary parts in the encoding. The fact that changing a qubit state by a global phase is reflected by the existence of infinitely many rebit states encoding the same qubit state. As a simple example, $|0\rangle|R\rangle$ and $|0\rangle|1\rangle$ are both encodings of the qubit state $|0\rangle$.

We will now exhibit a set of gates acting on encoded rebit states that is universal and preserves realness of the states. One common choice for a universal gate set [20] is

$$\mathcal{G}_{universal} = \langle \mathrm{CNOT}(i,j), H_i, \exp(i\pi/8 Z_i)\rangle. \tag{2.4}$$

We will restrict our allowed gates to the CSS-ness preserving group G_{CSS}, so that

$$\mathcal{G}_{restricted} = \langle \mathrm{CNOT}(i,j), \otimes_{i=1}^n H_i, Z_i, X_i\rangle, \tag{2.5}$$

where $\otimes_{i=1}^n H_i := H_{all}$ is the Hadamard on all rebits. Our goal is now to

show that, assuming that we have access to a reservoir of "magic states", we can performed encoded versions of the gates H_i and $\exp(i\pi/8Z_i)$. Our method closely follows the scheme of universal quantum computation by state injection of Bravyi and Kitaev [7]. The two magic states we will need are

$$|A\rangle = \overline{\frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\pi/4}|1\rangle\right)} \tag{2.6}$$

$$= \frac{1}{\sqrt{2}}\left(|0\rangle|R\rangle + \cos\frac{\pi}{4}|1\rangle|R\rangle + \sin\frac{\pi}{4}|1\rangle|I\rangle\right) \tag{2.7}$$

and

$$|B\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|+\rangle + |1\rangle|-\rangle\right), \tag{2.8}$$

where $|\pm\rangle$ is the ± 1 eigenstate of the Pauli operator X.

In the following we demonstrate how to obtain $\mathcal{G}_{universal}$, using only gates from $\mathcal{G}_{resricted}$, the states $|A\rangle$ and $|B\rangle$, and measurements in the computational basis. Said differently, for all $g \in \mathcal{G}_{universal}$, we show how to realize the transformation $\overline{|\psi\rangle} \to \overline{g|\psi\rangle}$ using only gates from G_{CSS} and the magic states $|A\rangle$, $|B\rangle$.

1. Pauli Y on qubit i

   The $Y = iXZ$ gate has imaginary components, so it's encoding for rebits has to affect the ancilla A. We will show that the encoded gate $\bar{Y}_i$ is

   $$\bar{Y}_i = Y_i \otimes Y_A = -X_i Z_i X_A Z_A. \tag{2.9}$$

   Indeed, starting from the initial encoded rebit state

   $$|\bar{\psi}\rangle = \sum_{\mathbf{v}} r_{\mathbf{v}}|\mathbf{v}\rangle\left(\cos\theta_{\mathbf{v}}|0\rangle_A + \sin\theta_{\mathbf{v}}|1\rangle_A\right) \tag{2.10}$$

   and applying $Y_i \otimes Y_A$ gives (up to an overall minus sign)

   $$\sum_{\mathbf{v}}(-1)^{v_i} r_{\mathbf{v}}|\mathbf{v}+\mathbf{e_i}\rangle\left(\cos\theta_{\mathbf{v}}|1\rangle_A - \sin\theta_{\mathbf{v}}|0\rangle_A\right), \tag{2.11}$$

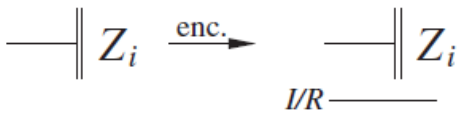
where $\mathbf{e_i} \in \mathbb{Z}_2^n$ is the basis vector that has value 1 in position i and 0 elsewhere. This state the same as the encoding of the state

$$Y_i \left|\psi\right\rangle = \sum_{\mathbf{v}} (-1)^{v_i} r_{\mathbf{v}} e^{i(\theta_{\mathbf{v}}+\pi/2)} \left|\mathbf{v} + \mathbf{e_i}\right\rangle, \tag{2.12}$$

proving our initial claim.

2. Measurement of Z_i

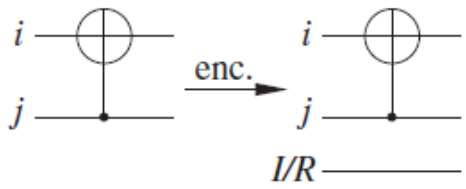
   The Pauli operator Z_i is real in the computational basis, so it does not mix the real and imaginary parts of the encoded state. Therefore the encoded version $\overline{Z_i}$ is the same as Z_i



(2.13)

3. CNOT gate between qubits i and j

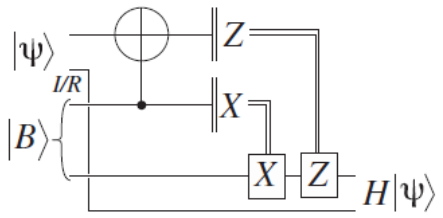
   The CNOT gate between qubits i and j gate belongs to G_{CSS}, and it is real in the computational basis. Therefore it does not mix the real and imaginary parts of the state it is applied to, so $\overline{CNOT(i,j)} = CNOT(i,j)$



(2.14)

4. Hadamard gate

   We will show that the following circuit, using one copy of the magic state $|B\rangle$ realizes an encoded Hadamard on the state $|\psi\rangle$



(2.15)

The initial state is

$$(r_0 \cos\theta_0 \,|0\rangle \,|R\rangle + r_1 \sin\theta_1 \,|1\rangle \,|I\rangle)\,|B\rangle\,. \tag{2.16}$$

After the CNOT gate, the state of the system is

$$\frac{1}{\sqrt{2}}(r_0 \cos\theta_0 \,|0\rangle \,|R\rangle + r_1 \sin\theta_1 \,|1\rangle \,|I\rangle)\,|0\rangle \,|+\rangle \tag{2.17}$$

$$+ \frac{1}{\sqrt{2}}(r_0 \cos\theta_0 \,|1\rangle \,|R\rangle + r_1 \sin\theta_1 \,|0\rangle \,|I\rangle)\,|1\rangle \,|-\rangle\,. \tag{2.18}$$

After the X measurement and the conditional X gate associated with it, we discard the third qubit and the state is

$$\frac{1}{\sqrt{2}}(r_0 \cos\theta_0 \,|0\rangle \,|R\rangle + r_1 \sin\theta_1 \,|1\rangle \,|I\rangle)\,|+\rangle \tag{2.19}$$

$$+ \frac{1}{\sqrt{2}}(r_0 \cos\theta_0 \,|1\rangle \,|R\rangle + r_1 \sin\theta_1 \,|0\rangle \,|I\rangle)\,|-\rangle\,. \tag{2.20}$$

After the Z measurement and the conditional Z gate associated with it, we discard the first qubit. We also swap the two remaining qubits so that the I/R ancilla ends up in the right position. The final state is

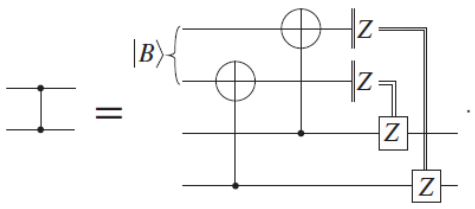
$$r_0 \cos\theta_0 \,|+\rangle \,|R\rangle + r_1 \sin\theta_1 \,|-\rangle \,|I\rangle = \overline{H\,|\psi\rangle}, \tag{2.21}$$

as claimed.

5. Conditional phase gate

   In order to make the next point of state-merging easier to follow, we

demonstrate a state injection circuit for the conditional phase gate. This two-qubit gate leaves all standard basis states invariant, except for $|11\rangle \to -|11\rangle$. We show that



(2.22)

The initial state is

$$\frac{1}{\sqrt{2}}\big(\,|0\rangle\,|+\rangle + |1\rangle\,|-\rangle\,\big)\big(a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle\,\big). \tag{2.23}$$

The two CNOT gates send the state to

$$\begin{aligned}&\frac{1}{\sqrt{2}}\big[a\big(\,|0\rangle\,|+\rangle + |1\rangle\,|-\rangle\,\big)\,|00\rangle + b\big(\,|0\rangle\,|+\rangle - |1\rangle\,|-\rangle\,\big)\,|01\rangle\\ &+ c\big(\,|1\rangle\,|+\rangle + |0\rangle\,|-\rangle\,\big)\,|10\rangle + d\big(\,|1\rangle\,|+\rangle - |0\rangle\,|-\rangle\,\big)\,|11\rangle\,\big].\end{aligned} \tag{2.24}$$

The measurement of Z on the first qubit, followed by a conditional Z gate on the third qubit yields (after discarding the measured qubit from the notation) the state $|\psi^{\pm}\rangle$. We have + if the measured spin was in the state $|0\rangle$, and − if the spin was in state $|0\rangle$. This state is

$$|\psi^{\pm}\rangle = a\,|\pm\rangle\,|00\rangle + b\,|\pm\rangle\,|01\rangle + c\,|\mp\rangle\,|10\rangle - d\,|\mp\rangle\,|11\rangle\,. \tag{2.25}$$

And finally, measuring Z and the first remaining qubit and applying a conditional Z on the last qubit produces the transformation

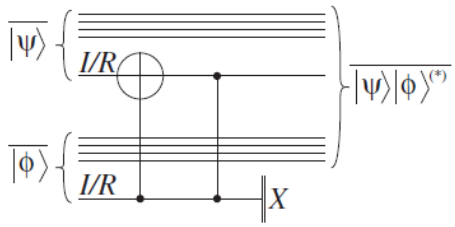
$$|\psi^{\pm}\rangle \to \pm\big(a\,|00\rangle + b\,|01\rangle + c\,|10\rangle - d\,|11\rangle\,\big), \tag{2.26}$$

where we have suppressed the measured qubit from the notation. This is exactly the outcome of a conditional phase gate, up to an irrelevant

global phase.

6. State-merging

   Suppose that we have two multiqubit states $|\psi\rangle$ and $|\phi\rangle$ in their rebit-encoded form $\overline{|\psi\rangle}$ and $\overline{|\phi\rangle}$. We would like to get the encoded state for their direct product: $\overline{|\psi\rangle \otimes |\phi\rangle}$. It is obvious that we need to get rid of one of the ancillas, and we will show that the following circuit realizes the adequate transformation.



(2.27)

   Let our initial encoded states be

$$\overline{|\psi\rangle} = \sum_{\mathbf{v}\in\mathbb{Z}_2^n} \left(r_{\mathbf{v}} \cos\theta_{\mathbf{v}} \,|\mathbf{v}\rangle\,|R\rangle + r_{\mathbf{v}} \sin\theta_{\mathbf{v}} \,|\mathbf{v}\rangle\,|I\rangle\right) \tag{2.28}$$

   and

$$\overline{|\phi\rangle} = \sum_{\mathbf{u}\in\mathbb{Z}_2^m} \left(s_{\mathbf{u}} \cos\kappa_{\mathbf{u}} \,|\mathbf{u}\rangle\,|R\rangle + s_{\mathbf{u}} \sin\kappa_{\mathbf{u}} \,|\mathbf{u}\rangle\,|I\rangle\right) \tag{2.29}$$

   Taking the product of these two states gives

$$\begin{aligned}\overline{|\psi\rangle} \otimes \overline{|\phi\rangle} = \sum_{\mathbf{v},\mathbf{u}} r_{\mathbf{v}} s_{\mathbf{u}} \big(&\cos\theta_{\mathbf{v}} \cos\kappa_{\mathbf{u}} \,|\mathbf{v},R\rangle\,|\mathbf{u},R\rangle \\ &+ \cos\theta_{\mathbf{v}} \sin\kappa_{\mathbf{u}} \,|\mathbf{v},R\rangle\,|\mathbf{u},I\rangle \\ &+ \sin\theta_{\mathbf{v}} \cos\kappa_j \,|\mathbf{v},I\rangle\,|\mathbf{u},R\rangle \\ &+ \sin\theta_{\mathbf{v}} \sin\kappa_{\mathbf{u}} \,|\mathbf{v},I\rangle\,|\mathbf{u},I\rangle\big).\end{aligned}$$

   It should be noted that this state is not equal to $\overline{(|\psi\rangle \otimes |\phi\rangle)}$. We now detail the effect of the circuit on this state. Applying the CNOT gate and the conditional phase gate (preserves basis stats except for

$|11\rangle \to -|11\rangle$) on the ancillas yields

$$\sum_{\mathbf{v},\mathbf{u}} r_{\mathbf{v}} s_{\mathbf{u}} (\cos\theta_{\mathbf{v}} \cos\kappa_{\mathbf{u}} |\mathbf{v}, R\rangle |\mathbf{u}, R\rangle - \cos\theta_{\mathbf{v}} \sin\kappa_{\mathbf{u}} |\mathbf{v}, I\rangle |\mathbf{u}, I\rangle + \sin\theta_{\mathbf{v}} \cos\kappa_{\mathbf{u}} |\mathbf{v}, I\rangle |\mathbf{u}, R\rangle + \sin\theta_{\mathbf{v}} \sin\kappa_{\mathbf{u}} |\mathbf{v}, R\rangle |\mathbf{u}, I\rangle). \tag{2.30}$$

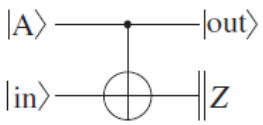
We now perform the X measurement on the last qubit and discard the measured qubit. We also swap the remaining ancilla qubit so it is at the last position in our notation. Let $|\Psi^a\rangle$ be the two possible output states after measurement, with $a = 0$ denoting the positive outcome, and $a = 1$ the negative outcome. This state is

$$\begin{aligned} |\Psi^a\rangle &= \sum_{\mathbf{v},\mathbf{u}} \big(r_{\mathbf{v}} s_{\mathbf{u}} (\cos\theta_{\mathbf{v}} \cos\kappa_{\mathbf{u}} + (-1)^a \sin\theta_{\mathbf{v}} \sin\kappa_{\mathbf{u}}) |\mathbf{v}, \mathbf{u}, R\rangle \\ &\qquad + (\sin\theta_{\mathbf{v}} \cos\kappa_{\mathbf{u}} + (-1)^{a+1} \cos\theta_{\mathbf{v}} \sin\kappa_{\mathbf{u}}) |\mathbf{v}, \mathbf{u}, I\rangle \big) \\ &= \sum_{\mathbf{v},\mathbf{u}} r_{\mathbf{v}} s_{\mathbf{u}} \big[\cos\left(\theta_{\mathbf{v}} + (-1)^{a+1} \kappa_{\mathbf{u}}\right) |\mathbf{v}, \mathbf{u}, R\rangle \\ &\qquad + \sin\left(\theta_{\mathbf{v}} + (-1)^{a+1} \kappa_{\mathbf{u}}\right) |\mathbf{v}, \mathbf{u}, I\rangle \big]. \end{aligned} \tag{2.31}$$

We see that $|\Psi^0\rangle = \overline{|\psi\rangle \otimes |\phi\rangle^*}$, where $|\phi\rangle^*$ is the state obtained from $|\phi\rangle$ by complex conjugation of the coefficients in the standard basis. Also, we have that $|\Psi^1\rangle = \overline{|\psi\rangle \otimes |\phi\rangle}$. We have thus shown that it is possible to merge encoded states, modulo a probabilistic complex conjugation. We will return to this issue of complex conjugation later, where we shall see that it affects neither the computational power nor the efficiency of state state injection.

7. The gate $\exp i\pi/8 Z_i$

   As we showed in Section 1.3, the following qubit circuit is a valid state-injection circuit for the $\exp i\pi/8 Z_i$ gate.



(2.32)

To realize an encoded version of this circuit on rebits, it suffices to perform a state merging routine between the magic state $|A\rangle$ and the input state, followed by the necessary CNOT and measurement.

While the magic state $|A\rangle$ certainly allows us to perform this task, any two-rebit state that encodes for the qubit state $\frac{1}{\sqrt{2}}(|0\rangle + e^{i\pi/4}|1\rangle)$ will also work.

2.3 Magic state distillation

2.3.1 Bell states

We first show, following Bennett et al. [5], that it is possible to take multiple copies of an impure Bell state and perform manipulations on them to get a state with higher purity. Consider an initial mixed state that is mostly composed of $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$,

$$\rho_{in} = (1-\epsilon)\rho_{00} + \frac{\epsilon}{3}(\rho_{01} + \rho_{10} + \rho_{11}), \tag{2.33}$$

where for notational simplicity, we describe the Bell basis in the stabilizer formalism so that ρ_{ij} is the pure state stabilized by $\langle (-1)^i Z_1 Z_2, (-1)^j X_1 X_2 \rangle$. This state has fidelity $F = \langle \Phi^+ | \rho_{in} | \Phi^+ \rangle = 1 - \epsilon$.

The first step of the protocol is

Step 1 on qubits 1 and 2

1. Start in state $\rho_{in} \otimes \rho_{in}$
2. Apply a CNOT gate with qubit 1 as control and qubit 3 as target
3. Apply a CNOT gate with qubit 2 as control and qubit 4 as target

4. Measure the two-qubit Pauli observable Z_3Z_4. If the result is $+1$, continue; otherwise reject the state and start over. A mixed state consisting of mostly the $|\Phi^+\rangle$ Bell state is now located on the first two qubits

We show, using the stabilizer formalism, how this protocol improves the fidelity. The initial state can be written as

$$\rho_{in} \otimes \rho_{in} = \sum_{\mathbf{e} \in \mathbb{Z}_2^4} f(\mathbf{e}) \rho_{e_1 e_2} \otimes \rho_{e_3 e_4}, \tag{2.34}$$

with

$$f(\mathbf{e}) = \begin{cases} (1-\epsilon)^2 & \text{if } (\mathbf{e} = 0 \\ \frac{\epsilon}{3}(1-\epsilon) & \text{if } \mathbf{e} \neq 0 \text{ and } e_1 = e_2 = 0 \text{ or } e_3 = e_4 = 0 \\ \frac{\epsilon^2}{9} & \text{otherwise} \end{cases} \tag{2.35}$$

Now let us see how the stabilizer of a state $\rho_{e_1 e_2} \otimes \rho_{e_3 e_4}$ evolves at each step in the circuit.

1. Initial state: $\langle (-1)^{e_1} Z_1 Z_2, (-1)^{e_2} X_1 X_2, (-1)^{e_3} Z_3 Z_4, (-1)^{e_4} X_3 X_4 \rangle$
2. First CNOT:

$$\langle (-1)^{e_1} Z_1 Z_2, (-1)^{e_2} X_1 X_2 X_3, (-1)^{e_3} Z_1 Z_3 Z_4, (-1)^{e_4} X_3 X_4 \rangle$$

3. Second CNOT:

$$\begin{aligned} &\langle (-1)^{e_1} Z_1 Z_2, (-1)^{e_2} X_1 X_2 X_3 X_4, (-1)^{e_3} Z_1 Z_2 Z_3 Z_4, (-1)^{e_4} X_3 X_4 \rangle \\ &\quad = \langle (-1)^{e_1} Z_1 Z_2, (-1)^{e_2 + e_4} X_1 X_2, (-1)^{e_1 + e_3} Z_3 Z_4, (-1)^{e_4} X_3 X_4 \rangle \end{aligned}$$

4. The Z_3Z_4 measurements discards all $\mathbf{e}$ such that $e_1 + e_3 = 1$, and keeps the other half. If the measurement was a success, we remove qubits 3 and 4, and the result is a two qubit state with stabilizer

$$\langle (-1)^{e_1} Z_1 Z_2, (-1)^{e_2 + e_4} X_1 X_2 \rangle .$$

With the previous information it is possible to calculate that the output state after the first step of the protocol is

$$\rho_{1step} \propto \left((1-\epsilon)^2 + \frac{\epsilon^2}{9}\right)\rho_{00} + \frac{2\epsilon^2}{9}\rho_{10} + \frac{2\epsilon}{3}(1-\epsilon)\rho_{01} + \frac{2\epsilon^2}{9}\rho_{11}. \quad (2.36)$$

This form of the density matrix after one step of purification reveals why it is interesting to have a two step protocol. Indeed, we see that the coefficient of ρ_{01} is of order ϵ, while the coefficients for ρ_{10} and ρ_{11} are of order ϵ^2. There has thus been no significant improvement in the ρ_{01} portion of the state, and we would like to fix this.

The goal of the second step of the protocol is to take to copies of ρ_{1step} and to correct the ρ_{01} errors. This is achieved by a slight modification of the first step.

Step 2

1. Start with input state $\rho_{1step} \otimes \rho_{1step}$
2. Apply a CNOT gate with qubit 1 as control and qubit 3 as target
3. Apply a CNOT gate with qubit 2 as control and qubit 4 as target
4. Measure the two-qubit Pauli observable X_1X2_2. If the result is $+1$, continue; otherwise reject the state and start over. A mixed state consisting of mostly the $|\Phi^+\rangle$ Bell state is now located on the last two qubits

The outcome of this procedure can be calculated as previously, and the result is

$$\begin{aligned}\rho_{out} \propto &\left[\left((1-\epsilon)^2 + \frac{\epsilon^2}{9}\right)^2 + \left(\frac{2\epsilon^2}{9}\right)^2\right]\rho_{00} + \frac{4\epsilon^2}{9}\left((1-\epsilon)^2 + \frac{\epsilon^2}{9}\right)\rho_{10} \\ &+ \left[\frac{2\epsilon}{3}(1-\epsilon) + \left(\frac{2\epsilon^2}{9}\right)^2\right]\rho_{01} + \frac{8\epsilon^3}{27}(1-\epsilon)\rho_{11}. \quad (2.37)\end{aligned}$$

This expression shows that the two step protocol allows for all of the errors to be reduced to order ϵ^2 or more. Assuming $\epsilon \ll 1$, we can expand the output state as

$$\rho_{out} = (1 - \epsilon_{out})\rho_{00} + \ldots, \tag{2.38}$$

with

$$\epsilon_{out} = \frac{8\epsilon^2}{9} + \mathcal{O}(\epsilon^3). \tag{2.39}$$

If we apply the protocol recursively, and let $\epsilon_{out}(k, \epsilon)$ be the error after k recursion levels starting with initial error ϵ, then the efficiency of the distillation is characterized by

$$\epsilon_{out}(k, \epsilon) = \sqrt{\frac{9}{8}} \left(\sqrt{\frac{8}{9}} \epsilon \right)^{2^k}, \tag{2.40}$$

in the limit where $\epsilon \ll 1$.

2.3.2 The state $|B\rangle$

The Bell state distillation method that we just presented can be straightforwardly modified to allow distillation of the magic state

$$|B\rangle = \frac{1}{\sqrt{2}}(|0\rangle |+\rangle + |1\rangle |-\rangle) = H_2 |\Phi^+\rangle . \tag{2.41}$$

As the previous equation shows, $|B\rangle$ is related to $|\Phi^+\rangle$ through the application of a Hadamard gate. For what follows, we need the following circuit identity

$$(H_1 H_2) CNOT(1, 2)(H_1 H_2) = CNOT(2, 1), \tag{2.42}$$

whose validity can be easily verified. Therefore, we only need to invert the direction of one CNOT, and change one measurement, to adapt the Bell state distillation scheme to our needs. Explicitly, we have

Step 1

1. Start in state $\rho_{in} \otimes \rho_{in}$

2. Apply a CNOT gate with qubit 1 as control and qubit 3 as target

3. Apply a CNOT gate with qubit 4 as control and qubit 2 as target

4. Measure the two-qubit Pauli observable Z_3X_4. If the result is $+1$, continue; otherwise reject the state and start over. A mixed state consisting of mostly the $|B\rangle$ state state is now located on the first two qubits

Step 2

1. Start with input state $\rho_{1step} \otimes \rho_{1step}$

2. Apply a CNOT gate with qubit 1 as control and qubit 3 as target

3. Apply a CNOT gate with qubit 4 as control and qubit 2 as target

4. Measure the two-qubit Pauli observable Z_1X2_2. If the result is $+1$, continue; otherwise reject the state and start over. A mixed state consisting of mostly the $|B\rangle$ state is now located on the last two qubits

The above described scheme allows us to distill $|B\rangle$ with the same efficiency of distillation as we calculated in the case of $|\Phi^+\rangle$ Bell states.

We are now in possession of a reliable supply of $|B\rangle$ magic states. Referring back to the rebit encoding scheme of Section 2.2, this allows us to perform encoded controlled-phase gates, encoded Hadamard gates, and state merging

The state $|A\rangle$

The magic state that we need for universal quantum computation is

$$|A\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|R\rangle + \frac{1}{\sqrt{2}}|1\rangle|R\rangle + \frac{1}{\sqrt{2}}|1\rangle|I\rangle\right). \tag{2.43}$$

It is the rebit encoding of the qubit state

$$|A_0\rangle = \frac{1}{2}(|0\rangle + e^{i\pi/4}|1\rangle), \tag{2.44}$$

for which a distillation scheme relying on an intricate CSS-code construction was first presented by Bravyi and Kitaev [7]. Their qubit scheme takes 15 copies of $|A_0\rangle$ and outputs a copy with higher fidelity. The only operations needed for the scheme are Clifford gates (which we can now do using the magic state $|B\rangle$), X and Z Pauli measurements, and the single qubit gate

$$A = \frac{1}{\sqrt{2}}(X + Y) = e^{-i\pi/4}SX, \tag{2.45}$$

where

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{2.46}$$

We refer the reader to the original paper for the details of the construction. Our strategy will be to implement the same method on encoded rebit states.The action of an S-gate on an encoded rebit state is

$$\overline{S\,|\psi\rangle} = C_Z(\overline{|\psi\rangle}), \tag{2.47}$$

where C_Z is a conditional phase gate, which is possible to realize using a magic states $|B\rangle$ as we have already demonstrated in Section 2.2

We now show how we can perform this scheme on impure copies of the two-rebit state $|A\rangle$. We first need to address the issue of how to deal with errors in the larger Hilbert space of two rebits. Consider the following basis for the real vector space of two rebits

$$\begin{aligned}
|A\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle\,|R\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\,|R\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\,|I\rangle\right) \\
|A_1\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle\,|R\rangle - \frac{1}{\sqrt{2}}\,|1\rangle\,|R\rangle - \frac{1}{\sqrt{2}}\,|1\rangle\,|I\rangle\right) = Z_1\,|A_0\rangle \\
|A'\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle\,|I\rangle - \frac{1}{\sqrt{2}}\,|1\rangle\,|R\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\,|I\rangle\right) \\
|A'_1\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle\,|I\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\,|R\rangle - \frac{1}{\sqrt{2}}\,|1\rangle\,|I\rangle\right) = Z_1\,|A_1\rangle
\end{aligned}$$

.

The important thing to notice here is that $|A\rangle$ and $|A'\rangle$ encode from the same one qubit state, and similarly for $|A_1\rangle$ and $|A_1'\rangle$. Indeed,

$$|A_1\rangle = \overline{i\,|A_0\rangle}$$
$$|A_1'\rangle = \overline{iZ\,|A_0\rangle}$$
$$.$$

Therefore, since our rebit quantum computer realizes all its computation by operating on encoded qubit states, the primed states will yield exactly the same outputs as the unprimed states. This allows us to regard them as equivalent for computational purposes. We take the following state as our initial impure state

$$\rho = (1-\epsilon)|A\rangle\langle A| + \epsilon|A_1\rangle\langle A_1|. \tag{2.48}$$

The Bravyi-Kitaev protocol requires 15 copies of this state:

$$\rho_{in} = \rho^{\otimes 15}. \tag{2.49}$$

We perform the state-merging routine which we described in the previous section. We must attend to the potential complex conjugation of states after an encoded conditional-phase gates, which we saw happens with probability one half. The results of a complex conjugation are either

$$|A_0\rangle^* = \frac{1}{2}(|0\rangle + e^{-i\pi/4}\,|1\rangle) = \frac{1}{2}(|0\rangle - ie^{i\pi/4}\,|1\rangle) = S^\dagger\,|A_0\rangle \tag{2.50}$$

or

$$|A_1\rangle^* = \frac{1}{2}(|0\rangle - e^{-i\pi/4}\,|1\rangle) = \frac{1}{2}(|0\rangle + ie^{i\pi/4}\,|1\rangle) = S^\dagger\,|A_1\rangle\,. \tag{2.51}$$

Therefore, if we know that a conjugation has happened, applying the encoded S gate will correct the state to it's non-conjugated form. The result of these manipulations is the state $\overline{\rho_{qubit}^{\otimes 15}}$, where $\rho_{qubit} = (1-\epsilon)\,|A_0\rangle\langle A_0| + \epsilon Z\,|A_0\rangle\langle A_0|\,Z$, for which Bravyi and Kitaev demonstrated their distillation

method. Following their method step by step in an encoded fashion, we obtain $\overline{\rho_{out} \otimes \rho_{junk}}$, with ρ_{junk} a 14-qubit state which is unentangled from the purified state ρ_0. Tracing out these junk states gives us our purified state.

Chapter 3

The rebit Wigner function

The goal of this Chapter is to define and characterize a rebit Wigner function that has properties that are as similar as possible to the qudit Wigner function described in Section 1.2.2. We will state and prove a number of properties about this Wigner function, the most important of which being a discrete Hudson's theorem for rebits. We will also study covariance of the Wigner function under a subgroup of the Clifford group, the CSS-ness preserving Clifford group. We will also give a theorem analogous pertaining to the classical simulability of circuits consisting of CSS-ness preserving Clifford unitaries acting on states with positive Wigner functions. This chapter is rather technical; the reader is advised to skip over the proofs on a first reading, and to consult the appendices for some mathematical background.

3.1 Definition and elementary properties

In this section we revisit the phase space formalism and Wigner function of Section 1.1, this time for rebits. The Wigner function W of a rebit state ρ is defined over the phase space

$$V = \{\mathbf{a} = (\mathbf{a}_Z, \mathbf{a}_X) | \mathbf{a}_Z, \mathbf{a}_X \in \mathbb{Z}_2^n\} \cong \mathbb{Z}_2^{2n}. \tag{3.1}$$

With each of the phase space points $\mathbf{a} \in V$ is associated a Pauli operator

$T_{\mathbf{a}}$ defined as

$$T_{(\mathbf{a}_Z,\mathbf{a}_X)} = Z(\mathbf{a}_Z)X(\mathbf{a}_X), \tag{3.2}$$

where $Z(\mathbf{a}) = Z_1^{a_1} \otimes Z_2^{a_2} \otimes \cdots \otimes Z_n^{a_n}$, and equivalently for $X(\mathbf{a}_X)$. Note that the Pauli operators defined in this way are not all Hermitian. These operators form a basis for the complex vector space of n by n square matrices, with the trace inner product $(A, B) = \frac{1}{2^n}\mathrm{Tr}(A^\dagger B)$. Indeed, it can be checked that

$$\mathrm{Tr}(T_{\mathbf{a}}^\dagger T_{\mathbf{b}}) = 2^n \delta_{\mathbf{a},\mathbf{b}}. \tag{3.3}$$

In this chapter we will be restricting ourselves to real density matrices. The requirement that density matrices should be Hermitian and real forces that their expansion in terms of $T_{\mathbf{a}}$'s is restricted to the subset

$$\mathcal{A} = \{\mathbf{a} \in V |\, (\mathbf{a}_\mathbf{z}, \mathbf{a}_\mathbf{x}) = 0\}. \tag{3.4}$$

By abuse of notation, $\mathcal{A}$ will sometimes be taken to denote the set $\{T_{\mathbf{a}} |\, \mathbf{a} \in \mathcal{A}\}$, but the meaning will always be clear from the context.

In analogy with the qudit situation, we define a family of operators $(A_{\mathbf{u}})_{\mathbf{u}\in V}$ as

$$A_{\mathbf{0}} = \frac{1}{2^n} \sum_{\mathbf{a}\in\mathcal{A}} T_{\mathbf{a}} \tag{3.5}$$

$$A_{\mathbf{u}} = T_{\mathbf{u}} A_{\mathbf{0}} T_{\mathbf{u}}^\dagger. \tag{3.6}$$

Using the commutation relations for the Pauli operators, we can also write $A_{\mathbf{u}}$ into the following form, of which we will make frequent use in this thesis

$$A_{\mathbf{u}} = \frac{1}{2^n} \sum_{\mathbf{a}\in\mathcal{A}} (-1)^{[\mathbf{u},\mathbf{a}]} T_{\mathbf{a}}, \tag{3.7}$$

where $[\mathbf{u}, \mathbf{v}] = (\mathbf{u}_X, \mathbf{u}_Z) - (\mathbf{u}_Z, \mathbf{u}_X)$ is the symplectic product, for which addition and multiplication is modulo two. The $A_{\mathbf{u}}$ operators do not form a basis for the space of real density matrices, because it is overcomplete. We can nonetheless obtain an informationally complete Wigner function

$W_\rho : V \to \mathbb{R}$, by defining

$$W_\rho(\mathbf{u}) = \frac{1}{2^n}\mathrm{Tr}(A_\mathbf{u}\rho) \tag{3.8}$$

We now turn to proving a list of properties concerning the rebit Wigner function W_ρ. We encourage the reader to compare this list with the list of properties for the qudit Wigner function appearing in Section 1.2.2.

1. W_ρ is real and sums to one: $\sum_\mathbf{u} W_\rho(\mathbf{u}) = 1$.

   Proof. For a rebit state ρ, the expansion of its density matrix in terms of Pauli operators can be written as

   $$\rho = \frac{1}{2^n}\sum_{\mathbf{a}\in\mathcal{A}} \rho_\mathbf{a} T_\mathbf{a}, \tag{3.9}$$

   with $\rho_\mathbf{0} = 1$ for the trace condition, and other conditions on the coefficients $\rho_\mathbf{a} \in \mathbb{R}$ to ensure that the state is physical, but that we will not need to consider. The Wigner function is thus

   $$\begin{aligned} W_\rho(\mathbf{u}) &= \frac{1}{2^{2n}}\sum_{\mathbf{a}\in\mathcal{A}} \rho_\mathbf{a}\mathrm{Tr}\left(A_\mathbf{u}T_\mathbf{a}\right) \\ &= \frac{1}{2^{2n}}\sum_{\mathbf{a}\in\mathcal{A}} \rho_\mathbf{a}(-1)^{[\mathbf{u},\mathbf{a}]}, \end{aligned} \tag{3.10}$$

   where in the first to second line we used Equation 3.7 and the orthogonality relation of the Paulis. Therefore, $W_\rho(\mathbf{u})$ is clearly real for all $\mathbf{u} \in V$ because of the realness of the coefficients $\rho_\mathbf{a}$.

   To prove that W_ρ sums to one, we take Equation 3.10 and sum it up over V:

   $$\begin{aligned} \sum_{\mathbf{u}\in V} W_\rho(\mathbf{u}) &= \frac{1}{2^{2n}}\sum_{\mathbf{a}\in\mathcal{A}} \rho_\mathbf{a}\sum_{\mathbf{u}\in V}(-1)^{[\mathbf{u},\mathbf{a}]} \\ &= \sum_{\mathbf{a}\in\mathcal{A}} \rho_\mathbf{a}\delta_{\mathbf{a},\mathbf{0}} \\ &= \rho_\mathbf{0} = 1, \end{aligned}$$

where in going to the second line we used the character orthogonality relations, $\sum_{\mathbf{u}\in V}(-1)^{[\mathbf{u},\mathbf{a}]} = 2^{2n}\delta_{\mathbf{a},\mathbf{0}}$. The basics of character theory are reviewed in Appendix A.1. □

2. W is informationally complete: $\rho = \sum_{\mathbf{u}} W_\rho(\mathbf{u})A_\mathbf{u}$

   Proof. We start from the expression 3.10 for $W_\rho(\mathbf{u})$

$$\begin{aligned}\sum_{\mathbf{u}\in V} W_\rho(\mathbf{u})A_\mathbf{u} &= \frac{1}{2^{2n}}\sum_{\mathbf{u}\in V}\sum_{\mathbf{a}\in\mathcal{A}}(-1)^{[\mathbf{u},\mathbf{a}]}\rho_\mathbf{a}A_\mathbf{u} \\ &= \frac{1}{2^{3n}}\sum_{\mathbf{u}\in V}\sum_{\mathbf{a},\mathbf{b}\in\mathcal{A}}(-1)^{[\mathbf{u},\mathbf{a}+\mathbf{b}]}\rho_\mathbf{a}T_\mathbf{b} \\ &= \frac{1}{2^n}\sum_{\mathbf{a},\mathbf{b}\in\mathcal{A}}\delta_{\mathbf{a},\mathbf{b}}\rho_\mathbf{a}T_\mathbf{b} \\ &= \frac{1}{2^n}\sum_{\mathbf{a}\in\mathcal{A}}\rho_\mathbf{a}T_\mathbf{a} \\ &= \rho.\end{aligned}$$

   We have again used the ever useful character orthogonality relation in going from the second to the third line. □

3. For all real density matrices ρ and σ, $W_{\rho\otimes\sigma} = W_\rho \cdot W_\sigma$

   Proof. Let $W_\rho : V_A \to \mathbb{R}$ and $W_\sigma : V_B \to \mathbb{R}$ be the Wigner functions for ρ and σ, defined on phase spaces of possibly different size $|V_A| = 2^n$ and $|V_B| = 2^m$. Then the domain of $W_{\rho\otimes\sigma}$ is $V_A \otimes V_B$, and

$$\begin{aligned}W_{\rho\otimes\sigma}(\mathbf{u}) &= \frac{1}{2^{n+m}}\mathrm{Tr}(A_\mathbf{u}\rho) \\ &= \frac{1}{2^{2(n+m)}}\sum_{\mathbf{a}\in\mathcal{A}_{A,B}}(-1)^{[\mathbf{u},\mathbf{a}]}\mathrm{Tr}(T_\mathbf{a}\rho\otimes\sigma).\end{aligned}$$

   By definition of the direct sum, we can decompose any $\mathbf{v} \in V_A \otimes V_B$ as $\mathbf{v} = \mathbf{v}_A + \mathbf{v}_B$, with $\mathbf{v}_A$ having trivial projection over the subspace V_B and vice versa.

If $\mathbf{a} \in V_A \otimes V_B$, then $T_{\mathbf{a}} = T_{\mathbf{a}_A} T_{\mathbf{a}_B}$, where $T_{\mathbf{a}_A}$ and $T_{\mathbf{a}_B}$ commute because they act on different systems. If we further wish to ensure that $\mathbf{a} \in \mathcal{A}_{A,B}$, it is necessary that either $(\mathbf{a}_A, \mathbf{a}_B) \in \mathcal{A}_A \otimes \mathcal{A}_B$, or $(\mathbf{a}_A, \mathbf{a}_B) \in \bar{\mathcal{A}}_A \otimes \bar{\mathcal{A}}_B$, where $\bar{\mathcal{A}}_A = V_A - \mathcal{A}_A$, and similarly for $\bar{\mathcal{A}}_B$. With the observations just made,

$$W_{\rho\otimes\sigma}(\mathbf{u}) = \frac{1}{2^{2(n+m)}} \sum_{\mathbf{a_A}\in\mathcal{A}_A} \sum_{\mathbf{a_B}\in\mathcal{A}_B} (-1)^{[\mathbf{u_A},\mathbf{a_A}]} (-1)^{[\mathbf{u_B},\mathbf{a_B}]} \mathrm{Tr}_A(T_{\mathbf{a_A}}\rho)\mathrm{Tr}_B(T_{\mathbf{a_B}}\sigma)$$
$$+ \frac{1}{2^{2(n+m)}} \sum_{\mathbf{a_A}\in\bar{\mathcal{A}}_A} \sum_{\mathbf{a_B}\in\bar{\mathcal{A}}_B} (-1)^{[\mathbf{u_A},\mathbf{a_A}]} (-1)^{[\mathbf{u_B},\mathbf{a_B}]} \mathrm{Tr}_A(T_{\mathbf{a_A}}\rho)\mathrm{Tr}_B(T_{\mathbf{a_B}}\sigma).$$

The second sum is zero because ρ and σ are real. We can now factorize the Wigner function, as claimed

$$\begin{aligned} W_{\rho\otimes\sigma}(\mathbf{u}) &= \left(\frac{1}{2^{2n}} \sum_{\mathbf{a_A}\in\mathcal{A}_A} (-1)^{[\mathbf{u_A},\mathbf{a_A}]} \mathrm{Tr}_A(T_{\mathbf{a_A}}\rho) \right) \\ &\qquad \times \left(\frac{1}{2^m} \sum_{\mathbf{a_B}\in\mathcal{A}_B} (-1)^{[\mathbf{u_B},\mathbf{a_B}]} \mathrm{Tr}_B(T_{\mathbf{a_B}}\sigma) \right) \\ &= W_\rho(\mathbf{u}_A) W_\sigma(\mathbf{u}_B). \end{aligned}$$

□

4. The trace inner product is

$$\mathrm{Tr}(\rho\sigma) = 2^n \sum_{\mathbf{u}\in V} W_\rho(\mathbf{u}) W_\sigma(\mathbf{u}) \tag{3.11}$$

Proof. Using Property 2, we obtain

$$\begin{aligned}\mathrm{Tr}(\rho\sigma) &= \sum_{\mathbf{u},\mathbf{v}\in V} W_\rho(\mathbf{u})W_\sigma(\mathbf{v})\mathrm{Tr}(A_\mathbf{u}A_\mathbf{v}) \\ &= \frac{1}{2^{2n}}\sum_{\mathbf{u},\mathbf{v}\in V} W_\rho(\mathbf{u})W_\sigma(\mathbf{v})\sum_{\mathbf{a},\mathbf{b}\in\mathcal{A}}(-1)^{[\mathbf{u},\mathbf{a}]+[\mathbf{v},\mathbf{b}]}\mathrm{Tr}(T_\mathbf{a}T_\mathbf{b}) \\ &= \frac{1}{2^{n}}\sum_{\mathbf{u},\mathbf{v}\in V} W_\rho(\mathbf{u})W_\sigma(\mathbf{v})\sum_{\mathbf{a}\in\mathcal{A}}(-1)^{[\mathbf{u}+\mathbf{v},\mathbf{a}]} \\ &= 2^n\sum_{\mathbf{u}\in V} W_\rho(\mathbf{u})W_\sigma(\mathbf{u})\end{aligned}$$

□

,where in the previous calculation we have used the orthogonality of the Pauli operators and the character orthogonality relation.

5. Hudson's theorem: The only pure states with positive Wigner function are CSS-stabilizer states. The next section is dedicated to proving this theorem.

3.2 Hudson's theorem

The goal of this section is to prove Hudson's theorem:

Theorem 3.1. *A pure rebit state has a non-negative Wigner function if and only if it is a CSS state.*

To prove this, we will need to collect a significant chain of lemmas. We start by proving

Lemma 3.2. *A state ρ has Wigner function*

$$W_\rho = \frac{1}{2^n}\delta_{\mathbf{t}+N^\perp\times N}(\mathbf{u}), \tag{3.12}$$

where $\mathbf{t}\in V$ and N is a subspace of $\mathbb{Z}_2^n$, if and only if ρ is a pure CSS state.

Proof. A CSS state $|\psi\rangle$ has a stabilizer that is the direct product of an X part and a Z part: $\mathcal{S}(|\psi\rangle) = \mathcal{S}_X \times \mathcal{S}_Z$. Taking $N \subset V$ to be the subspace associated with $\mathcal{S}_X$, then its orthogonal complement $N^\perp = \{\mathbf{u} \in V | \, [\mathbf{u}, \mathbf{v}] = 0, \forall \mathbf{v} \in N\}$ will be the subspace associated with $\mathcal{S}_Z$. Therefore, any Pauli operator in $\mathcal{S}$ has the form $\pm T_{\mathbf{a}_X + \mathbf{a}_X}$, with $\mathbf{a}_X \in N$ and $\mathbf{a}_Z \in N^\perp$. By the completeness of characters, there exists a $\mathbf{t} \in V$ such that

$$\rho = |\psi\rangle \langle\psi| = \frac{1}{2^n} \sum_{\mathbf{u} \in N^\perp \times N} (-1)^{[\mathbf{t},\mathbf{u}]} T_{\mathbf{u}}. \tag{3.13}$$

Then the Wigner function W_ρ is

$$\begin{aligned}
W_\rho(\mathbf{u}) &= \frac{1}{2^{2n}} \sum_{\mathbf{v} \in N^\perp \times N} (-1)^{[\mathbf{t},\mathbf{v}]} \mathrm{Tr}(A_{\mathbf{u}} T_{\mathbf{a}}) \\
&= \frac{1}{2^{3n}} \sum_{\mathbf{u} \in N^\perp \times N} \sum_{\mathbf{a} \in \mathcal{A}} (-1)^{[\mathbf{t},\mathbf{v}]} (-1)^{[\mathbf{a},\mathbf{u}]} \mathrm{Tr}(T_{\mathbf{a}} T_{\mathbf{v}}) \\
&= \frac{1}{2^{2n}} \sum_{\mathbf{u} \in N^\perp \times N} \sum_{\mathbf{a} \in \mathcal{A}} (-1)^{[\mathbf{t},\mathbf{v}]} (-1)^{[\mathbf{u},\mathbf{a}]} \delta_{\mathbf{a},\mathbf{v}} \\
&= \frac{1}{2^n} \sum_{\mathbf{u} \in N^\perp \times N} (-1)^{[\mathbf{v},\mathbf{t}+\mathbf{u}]} \\
&= \frac{1}{2^n} \delta_{\mathbf{t} + (N^\perp \times N)}(\mathbf{u})
\end{aligned}$$

This proves one direction of the claim. For reverse implication, we use the fact the the Wigner function is informationally complete. Starting from $W_\rho(\mathbf{u}) = \frac{1}{2^n} \delta_{\mathbf{t}+(N^\perp \times N)}(\mathbf{u})$, we do the inverse Fourier transform to get $\rho = \sum_{\mathbf{u} \in (\mathbf{N}^\perp \times \mathbf{N})} (-1)^{[\mathbf{t},\mathbf{u}]} T_{\mathbf{u}}$, and such a state is a CSS state. □

We have thus proved that all CSS states have positive Wigner function. We now set out for the nontrivial task of proving that all pure states with positive Wigner function are stabilizer states

3.2.1 Bochner's theorem

Here we prove a very limited form of a theorem in harmonic analysis known as Bochner's theorem. It will allow us to better understand the Wigner

function of a pure state by looking at it's Fourier transform.

The Wigner function of a pure rebit state $|\psi\rangle = \sum_{\mathbf{x}\in\mathbb{Z}_2^n} \psi(\mathbf{x})\,|x\rangle$, evaluated at $\mathbf{u} = (\mathbf{p}, \mathbf{q}) \in V$ is

$$\begin{aligned} W_\psi(\mathbf{u}) &= \frac{1}{2^n} \sum_{\mathbf{x},\mathbf{y}\in Z_2^n} \psi(\mathbf{x})\psi(\mathbf{y}) \mathrm{Tr}\left(A_\mathbf{u}\,|\mathbf{x}\rangle\,\langle\mathbf{y}|\right) \\ &= \frac{1}{2^{2n}} \sum_{\mathbf{x},\mathbf{y}\in Z_2^n} \sum_{\mathbf{a}\in\mathcal{A}} \psi(\mathbf{x})\psi(\mathbf{y})(-1)^{[\mathbf{u},\mathbf{a}]} \mathrm{Tr}\left(T_\mathbf{a}\,|\mathbf{x}\rangle\,\langle\mathbf{y}|\right). \end{aligned}$$

Because the state is real, the sum over $\mathcal{A}$ can be extended to a sum over the full phase space without changing the result. This gives

$$\begin{aligned} W_\psi(\mathbf{u}) &= \frac{1}{2^{2n}} \sum_{\mathbf{x},\mathbf{y}\in Z_2^n} \sum_{\mathbf{a}\in V} \psi(\mathbf{x})\psi(\mathbf{y})(-1)^{[\mathbf{u},\mathbf{a}]} \mathrm{Tr}\left(T_\mathbf{a}\,|\mathbf{x}\rangle\,\langle\mathbf{y}|\right) \\ &= \frac{1}{2^{2n}} \sum_{\mathbf{x},\mathbf{y}\in Z_2^n} \sum_{\mathbf{a}\in V} \psi(\mathbf{x})\psi(\mathbf{y})(-1)^{[\mathbf{u},\mathbf{a}]}(-1)^{(\mathbf{a}_Z,\mathbf{y})} \mathrm{Tr}\left(|\mathbf{x}+\mathbf{a}_X\rangle\,\langle\mathbf{y}|\right) \\ &= \frac{1}{2^{2n}} \sum_{\mathbf{x},\mathbf{y}\in Z_2^n} \sum_{\mathbf{a}\in V} \psi(\mathbf{x})\psi(\mathbf{y})(-1)^{[\mathbf{u},\mathbf{a}]}(-1)^{(\mathbf{a}_Z,\mathbf{y})} \delta_{\mathbf{x}+\mathbf{a}_X,\mathbf{y}} \\ &= \frac{1}{2^{2n}} \sum_{\mathbf{y}\in Z_2^n} \sum_{\mathbf{a}\in V} \psi(\mathbf{y}+\mathbf{a}_X)\psi(\mathbf{y})(-1)^{[\mathbf{u},\mathbf{a}]}(-1)^{(\mathbf{a}_Z,\mathbf{y})} \\ &= \frac{1}{2^{2n}} \sum_{\mathbf{y}\in Z_2^n} \sum_{\mathbf{a}_X\in Z_2^n} (-1)^{(\mathbf{u}_Z,\mathbf{a}_X)} \psi(\mathbf{y}+\mathbf{a}_X)\psi(\mathbf{y}) \left(\sum_{\mathbf{a}_Z\in Z_2^n} (-1)^{(\mathbf{a}_Z,\mathbf{y}+\mathbf{u}_X)} \right) \\ &= \frac{1}{2^n} \sum_{\mathbf{y}\in Z_2^n} \sum_{\mathbf{a}_X\in Z_2^n} (-1)^{(\mathbf{u}_Z,\mathbf{a}_X)} \psi(\mathbf{y}+\mathbf{a}_X)\psi(\mathbf{y}) \delta_{\mathbf{y},\mathbf{u}_X} \\ &= \frac{1}{2^n} \sum_{\mathbf{a}_X\in Z_2^n} (-1)^{(\mathbf{u}_Z,\mathbf{a}_X)} \psi(\mathbf{u}_X+\mathbf{a}_X)\psi(\mathbf{u}_X) \\ &= \frac{1}{2^n} \sum_{\mathbf{x}\in Z_2^n} (-1)^{(\mathbf{u}_Z,\mathbf{x})} \psi(\mathbf{u}_X+\mathbf{x})\psi(\mathbf{u}_X). \end{aligned}$$

We are inspired to define the function

$$K_\psi(\mathbf{q}, \mathbf{x}) = \psi(\mathbf{q})\psi(\mathbf{q}+\mathbf{x}) \tag{3.14}$$

So that the Wigner function for fixed $\mathbf{q}$, $W_\psi(\cdot, \mathbf{q})$, is the Fourier transform of $K_\psi(\mathbf{q}, \cdot)$

$$W_\psi(\mathbf{p}, \mathbf{q}) = \frac{1}{2^n} \sum_{\mathbf{x} \in \mathbb{Z}_2^{\mathbf{n}}} (-1)^{(\mathbf{p}, \mathbf{x})} K_\psi(\mathbf{q}, \mathbf{x}). \tag{3.15}$$

We can also do the inverse Fourier transform (in dimension two it is the same as the Fourier transform up to a multiplicative factor), to get

$$\begin{aligned} \sum_{\mathbf{p} \in \mathbb{Z}_2^n} (-1)^{(\mathbf{p}, \mathbf{x})} W_\psi(\mathbf{p}, \mathbf{q}) &= \frac{1}{2^n} \sum_{\mathbf{p}, \mathbf{y} \in \mathbb{Z}_2^n} (-1)^{(\mathbf{p}, \mathbf{x}+\mathbf{y})} K_\psi(\mathbf{q}, \mathbf{y}) \\ &= \sum_{\mathbf{y} \in \mathbb{Z}_2^n} \delta_{\mathbf{x}, \mathbf{y}} K_\psi(\mathbf{q}, \mathbf{y}) \\ &= K_\psi(\mathbf{q}, \mathbf{x}), \end{aligned}$$

where in the first to second line we have used the character orthogonality relation. We restate the previous equation for future reference:

$$K_\psi(\mathbf{q}, \mathbf{x}) = \sum_{\mathbf{x} \in \mathbb{Z}_2^n} (-1)^{(\mathbf{p}, \mathbf{x})} W_\psi(\mathbf{p}, \mathbf{q}). \tag{3.16}$$

Theorem 3.3. *(Bochner's theorem) The matrix $A_y^x = K(\mathbf{q}, \mathbf{x} - \mathbf{y})$ is positive semi-definite if and only if $W_\psi(\mathbf{y}, \mathbf{q}) \geq 0 \quad \forall \mathbf{y} \in \mathbb{Z}_2^n$. A real matrix A with entries A_y^x is positive semi-definite if $\sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} b_{\mathbf{x}} b_{\mathbf{y}} A_y^x \geq 0$ for all real vectors $\mathbf{b}$ with coefficients indexed by $\mathbb{Z}_2^n$.*

Proof. Let $\mathbf{b}$ be a real vector with coefficients indexed by $\mathbb{Z}_2^n$. Then

$$\begin{aligned} \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n} b_{\mathbf{x}} b_{\mathbf{y}} K(\mathbf{q}, \mathbf{x} - \mathbf{y}) &= \frac{1}{2^n} \sum_{\mathbf{x}, \mathbf{y}, \mathbf{p}} b_{\mathbf{x}} b_{\mathbf{y}} (-1)^{(\mathbf{p}, \mathbf{x}-\mathbf{y})} W_\psi(\mathbf{p}, \mathbf{q}) \\ &= \frac{1}{2^n} \sum_{\mathbf{p}} W_\psi(\mathbf{p}, \mathbf{q}) \left(\sum_{\mathbf{x}} (-1)^{(\mathbf{p}, \mathbf{x})} b_{\mathbf{x}} \right)^2 \end{aligned}$$

Therefore, if $W_\psi(\mathbf{p}, \mathbf{q}) \geq 0$ for all $\mathbf{p} \in \mathbb{Z}_2^n$, the right hand side is greater than zero, showing that $K(\mathbf{q}, \mathbf{x} - \mathbf{y})$ is positive semi-definite.

Conversely, suppose there exists an $\mathbf{r}$ such that $W_\psi(\mathbf{r}, \mathbf{q}) < 0$. Then,

recalling that the characters are a basis for the space of functions over $\mathbb{Z}_2^n$, we can choose a vector $\mathbf{a} \in \mathbb{Z}_2^n$ such that

$$\left(\sum_{\mathbf{x}} (-1)^{(\mathbf{p},\mathbf{x})} a_{\mathbf{x}}\right)^2 = \delta_{\mathbf{p},\mathbf{r}}, \tag{3.17}$$

showing that $K(\mathbf{q}, \mathbf{x} - \mathbf{y})$ is not positive semi-definite in this case. □

3.2.2 Modulus and support

Lemma 3.4. *For a pure rebit state $|\psi\rangle = \sum_{\mathbf{x}} \psi(\mathbf{x}) |\mathbf{x}\rangle$, non-negativity of W_ψ implies that the function ψ has constant absolute value over its support.*

Proof. It is a known fact about positive semi-definite matrices that all of their principal minors are non-negative. Applied to the matrix $A_{\mathbf{y}}^{\mathbf{x}} = K(\mathbf{q}, \mathbf{x} - \mathbf{y})$, this means that the determinant

$$\begin{vmatrix} A_{\mathbf{0}}^{\mathbf{0}} & A_{\mathbf{x}}^{\mathbf{0}} \\ A_{\mathbf{0}}^{\mathbf{x}} & A_{\mathbf{x}}^{\mathbf{x}} \end{vmatrix} = \begin{vmatrix} \psi(\mathbf{q})^2 & \psi(\mathbf{q})\psi(\mathbf{q}+\mathbf{x}) \\ \psi(\mathbf{q})\psi(\mathbf{q}+\mathbf{x}) & \psi(\mathbf{q})^2 \end{vmatrix} \geq 0. \tag{3.18}$$

Expanding this equation gives

$$\psi(\mathbf{q})^4 \geq \psi(\mathbf{q})^2 \psi(\mathbf{q}+\mathbf{x})^2. \tag{3.19}$$

Let $\mathbf{q}, \mathbf{q}' \in \mathrm{supp}(\psi)$ and use $\mathbf{x} = \mathbf{q} + \mathbf{q}'$ in Equation 3.19. Then,

$$\begin{aligned} \psi(\mathbf{q})^4 &\geq \psi(\mathbf{q})^2 \psi(\mathbf{q}')^2 \\ |\psi(\mathbf{q})| &\geq |\psi(\mathbf{q}')|. \end{aligned} \tag{3.20}$$

Going through the same reasoning, but exchanging $\mathbf{q}$ and $\mathbf{q}'$ we get $|\psi(\mathbf{q}')| \geq |\psi(\mathbf{q})|$, and thus

$$|\psi(\mathbf{q})| = |\psi(\mathbf{q}')| \quad \forall \mathbf{q}, \mathbf{q}' \in \mathrm{supp}(\psi) \tag{3.21}$$

□

Lemma 3.5. *For a pure rebit state* $|\psi\rangle = \sum_{\mathbf{x}} \psi(\mathbf{x}) |\mathbf{x}\rangle$*, non-negativity of* W_ψ *implies that the support of* ψ *is an affine subspace of* $\mathbb{Z}_2^n$*. This means that* $\mathrm{supp}(\psi) = \mathbf{q_0} + N$ *where* $\mathbf{q_0} \in \mathbb{Z}_2^n$ *and* N *is a subspace of* $\mathbb{Z}_2^n$*.*

Proof. Let $\mathbf{q}_0 \in \mathrm{supp}(\psi)$. To show that $\mathrm{supp}(\psi)$ is an affine subspace, we must show that if $(\mathbf{q}_0 + \mathbf{x}) \in \mathrm{supp}(\psi)$ and $(\mathbf{q}_0 + \mathbf{y}) \in \mathrm{supp}(\psi)$, then $\mathbf{q}_0 + \mathbf{x} + \mathbf{y} \in \mathrm{supp}(\psi)$. From Bochner's theorem, $A_{\mathbf{y}}^{\mathbf{x}}$ is positive semi-definite, so every principal minor is non-negative. In particular,

$$\begin{vmatrix} A_{\mathbf{0}}^{\mathbf{0}} & A_{\mathbf{x}}^{\mathbf{0}} & A_{\mathbf{y}}^{\mathbf{0}} \\ A_{\mathbf{0}}^{\mathbf{x}} & A_{\mathbf{x}}^{\mathbf{x}} & A_{\mathbf{y}}^{\mathbf{x}} \\ A_{\mathbf{0}}^{\mathbf{y}} & A_{\mathbf{y}}^{\mathbf{0}} & A_{\mathbf{y}}^{\mathbf{y}} \end{vmatrix} \geq 0. \tag{3.22}$$

Expanding this determinant yields

$$\begin{aligned} &\psi(\mathbf{q}_0)^3[\psi(\mathbf{q}_0)^3 + 2\psi(\mathbf{q}_0 + \mathbf{x})\psi(\mathbf{q}_0 + \mathbf{y})\psi(\mathbf{q}_0 + \mathbf{x} + \mathbf{y})] \\ &\quad - \psi(\mathbf{q}_0)^4[\psi(\mathbf{q}_0 + \mathbf{x})^2 - \psi(\mathbf{q}_0 + \mathbf{y})^2 - \psi(\mathbf{q}_0 + \mathbf{x} + \mathbf{y})^2] \geq 0. \end{aligned} \tag{3.23}$$

We will now suppose that $(\mathbf{q}_0 + \mathbf{x} + \mathbf{y}) \notin \mathrm{supp}(\psi)$, and reach a contradiction. If this is the case, then

$$\psi(\mathbf{q}_0)^6 - \psi(\mathbf{q}_0)^4[\psi(\mathbf{q}_0 + \mathbf{x})^2 - \psi(\mathbf{q}_0 + \mathbf{y})^2] \geq 0. \tag{3.24}$$

But that's impossible, since by the constant modulus property, $\psi(\mathbf{q}_0 + \mathbf{x})^2 - \psi(\mathbf{q}_0 + \mathbf{y})^2 = 0$. This proves the claim that $\mathbf{q}_0 + \mathbf{x} + \mathbf{y} \in \mathrm{supp}(\psi)$, and therefore that $\mathrm{supp}(\psi) = \mathbf{q}_0 + N$ for some subspace $N \subset \mathbb{Z}_2^n$. □

Lemma 3.6. *If it is non-negative, the Wigner function of a pure real state* $|\psi\rangle = \sum_{\mathbf{x}} \psi(\mathbf{x}) |\mathbf{x}\rangle$ *is*

$$W_\psi(\mathbf{p}, \mathbf{q}) = \frac{1}{2^n} \delta_{(\mathbf{p_0} + \mathbf{N}^\perp) \times (\mathbf{q_0} + \mathbf{N})}, \tag{3.25}$$

with $\mathbf{q}_0 + N = \mathrm{supp}(\psi)$ *as given by Lemma 3.5, and* $\mathbf{p}_0 \in \mathbb{Z}_2^n$

Proof. Suppose that $\mathbf{p}, \mathbf{q} \in \text{supp}(W_\psi)$. Then

$$W_\psi(\mathbf{p}, \mathbf{q}) = \frac{1}{2^n} \sum_{\mathbf{x} \in \mathbb{Z}_2^n} (-1)^{(\mathbf{p}, \mathbf{x})} \psi(\mathbf{q}) \psi(\mathbf{q} + \mathbf{x}) \tag{3.26}$$

is non zero by assumption, implying that $\psi(\mathbf{q}) \neq 0$. Therefore, by Lemma 3.5, $\mathbf{q} \in \mathbf{q}_0 + N$. Furthermore, $\psi(\mathbf{q} + \mathbf{x})$ is zero unless $\mathbf{x} \in N$. This gives

$$W_\psi(\mathbf{p}, \mathbf{q}) = \frac{1}{2^n} \sum_{\mathbf{x} \in N} (-1)^{(\mathbf{p}, \mathbf{x})} \psi(\mathbf{q}) \psi(\mathbf{q} + \mathbf{x}). \tag{3.27}$$

By the constant modulus property of Lemma 3.4,

$$\psi(\mathbf{q}) \psi(\mathbf{q} + \mathbf{x}) = \psi(\mathbf{q})^2 \chi(\mathbf{x}), \tag{3.28}$$

where χ is a character of N. Because it is a character, $\chi(\mathbf{x}) = (-1)^{(\mathbf{p}_0, \mathbf{x})}$ for some $\mathbf{p}_0 \in N$. Plugging this in Equation 3.27, we get

$$W_\psi(\mathbf{p}, \mathbf{q}) = \frac{1}{2^n} \psi(\mathbf{q})^2 \sum_{\mathbf{x} \in N} (-1)^{(\mathbf{p} + \mathbf{p}_0, \mathbf{x})}. \tag{3.29}$$

We would like to use the character orthogonality relation. Using the fact that $\mathbb{Z}_2^n = N \oplus N^\perp$ we see that the sum will be non-zero if and only if $\mathbf{p} \in \mathbf{p}_0 + N^\perp$. Combining this with the earlier established fact that $\mathbf{q} \in \mathbf{q}_0 + N$, we get

$$W_\psi(\mathbf{p}, \mathbf{q}) = c \delta_{(\mathbf{p_0} + \mathbf{N}^\perp) \times (\mathbf{q_0} + \mathbf{N})}, \tag{3.30}$$

where c is an undetermined positive constant, which we now fix through the requirement that W_ψ sums to one. The dimension of the support of W_ψ is

$$|\text{supp}(W_\psi)| = |N| \cdot |N^\perp| = |\mathbb{Z}_2^n| = 2^n. \tag{3.31}$$

We thus obtain the desired result

$$W_\psi = \frac{1}{2^n} \delta_{(\mathbf{p_0} + \mathbf{N}^\perp) \times (\mathbf{q_0} + \mathbf{N})}. \tag{3.32}$$

□

We have now collected all the lemmas needed to prove Hudson's Theorem

Proof of Theorem 3.1. The content of Lemma 3.2 is that

$$W_\psi = \frac{1}{2^n}\delta_{\mathbf{t}+(\mathbf{N}^\perp\times+\mathbf{N})} \iff |\psi\rangle \text{ is a pure CSS state.} \tag{3.33}$$

Combining this with Lemma 3.6, and identifying $\mathbf{t} = (\mathbf{p}_0, \mathbf{q}_0)$, we get that every non-negative pure state is a pure CSS state, which proves the theorem.

□

3.3 Covariance

In this section we show that CSS-ness preserving unitaries have a very simple action on the Wigner function, in that they can be interpreted as permutations on the phase space. This property will be crucial in establishing the simulability results of section 3.4. We start by proving the following lemma

Lemma 3.7. *Let $g \in G_{CSS}$. Then there is a unique pair $(F, \mathbf{x})$, composed of a symplectic matrix $F \in Sp_{2n}(\mathbb{Z}_2)$ and a vector $\mathbf{x} \in V$ such that for all $\mathbf{a} \in \mathcal{A}$,*

$$gT_\mathbf{a}g^\dagger = (-1)^{[\mathbf{x},\mathbf{a}]}T_{F\mathbf{a}}. \tag{3.34}$$

Proof. Since g is a member of G_{CSS}, we have

$$gT_\mathbf{a}g^\dagger = \chi(\mathbf{a})T_{\phi(\mathbf{a})} \tag{3.35}$$

for some functions $\chi : V \to \{\pm 1\}$ and $\phi : V \to V$. Using this observation on $T_{\mathbf{a}+\mathbf{b}}$ yields

$$gT_{\mathbf{a}+\mathbf{b}}g^\dagger = \chi(\mathbf{a}+\mathbf{b})T_{\phi(\mathbf{a}+\mathbf{b})}. \tag{3.36}$$

But the following also holds

$$\begin{aligned} gT_{\mathbf{a}+\mathbf{b}}g^\dagger &= (-1)^{[\mathbf{a},\mathbf{b}]}gT_{\mathbf{a}}g^\dagger gT_{\mathbf{b}}g^\dagger \\ &= (-1)^{[\mathbf{a},\mathbf{b}]}\chi(\mathbf{a})\chi(\mathbf{b})T_{\phi(\mathbf{a})}T_{\phi(\mathbf{b})} \\ &= (-1)^{[\mathbf{a},\mathbf{b}]}(-1)^{[\phi(\mathbf{a}),\phi(\mathbf{b})]}\chi(\mathbf{a})\chi(\mathbf{b})T_{\phi(\mathbf{a})+\phi(\mathbf{b})}. \end{aligned}$$

Comparing with Equation 3.36, this tells us that

$$\phi(\mathbf{a}+\mathbf{b}) = \phi(\mathbf{a}) + \phi(\mathbf{b}) \tag{3.37}$$

and

$$\chi(\mathbf{a}+\mathbf{b}) = (-1)^{[\mathbf{a},\mathbf{b}]}(-1)^{[\phi(\mathbf{a}),\phi(\mathbf{b})]}\chi(\mathbf{a})\chi(\mathbf{b}). \tag{3.38}$$

Furthermore, conjugation by g preserves the commutation relations of Pauli operators, and these commutation relations are given by the symplectic product. Therefore, $[\mathbf{a},\mathbf{b}] = [\phi(\mathbf{a}),\phi(\mathbf{b})]$. This observation, combined with the linearity of Equation 3.37 proves that ϕ is a symplectic transformation $\phi = F \in Sp_{2n}(\mathbb{Z}_2^n)$. It also allows us to deduce that

$$\chi(\mathbf{a}+\mathbf{b}) = \chi(\mathbf{a})\chi(\mathbf{b}), \tag{3.39}$$

so χ is a character. By a property of characters, there exists $\mathbf{x} \in V$ such that $\chi(\mathbf{a}) = (-1)^{[\mathbf{x},\mathbf{a}]}$. This proves the claim. □

We note in passing that the matrix F appearing in Lemma 3.7 is injective (has trivial kernel), because Clifford unitaries have a bijective action over the Paulis under conjugation.

Here is the statement of the theorem that establishes the conditions for covariance of the rebit Wigner function.

Theorem 3.8. *For a n-rebit state ρ, W_ρ is covariant under G_{CSS}, in the sense that for all $g \in G_{CSS}$,*

$$W_{g\rho g^\dagger}(\mathbf{u}) = W_\rho(F\mathbf{u}+\mathbf{t}), \tag{3.40}$$

for a unique symplectic transformation F and unique vector $\mathbf{t} \in V$

Proof. For any $g \in G_{CSS}$, let $(F_g, \mathbf{x}_g)$ be the symplectic transformation and vector appearing in Lemma 3.7. We start by using the latter lemma to determine the action of $g \in G_{CSS}$ on the point operators.

$$\begin{aligned} gA_{\mathbf{u}}g^\dagger &= \frac{1}{2^n}\sum_{\mathbf{a}\in\mathcal{A}}(-1)^{[\mathbf{u},\mathbf{a}]}gT_{\mathbf{a}}g^\dagger \\ &= \frac{1}{2^n}\sum_{\mathbf{a}\in\mathcal{A}}(-1)^{[\mathbf{u}+\mathbf{x}_g,\mathbf{a}]}T_{F_g\mathbf{a}} \\ &= \frac{1}{2^n}\sum_{\mathbf{a}\in\mathcal{A}}(-1)^{[F_g(\mathbf{u}+\mathbf{x}_g),F_g\mathbf{a}]}T_{F_g\mathbf{a}} \\ &= \frac{1}{2^n}\sum_{\mathbf{a}\in\mathcal{A}}(-1)^{[F_g(\mathbf{u}+\mathbf{x}_g),\mathbf{b}]}T_{\mathbf{b}} \\ &= A_{F_g(\mathbf{u}+\mathbf{x}_g)} \\ &= A_{F_g\mathbf{u}+\mathbf{t}_g}, \end{aligned}$$

where in the last line we defined the vector $\mathbf{t}_g = F_g\mathbf{x}_g$. It only remains to plug this in the expression for the Wigner function

$$\begin{aligned} W_{g\rho g^\dagger}(\mathbf{u}) &= \frac{1}{2^n}\mathrm{Tr}(A_{\mathbf{u}}g\rho g^\dagger) \\ &= \frac{1}{2^n}\mathrm{Tr}((g^{-1})A_{\mathbf{u}}(g^{-1})^\dagger\rho) \\ &= \frac{1}{2^n}\mathrm{Tr}(A_{F_{g^{-1}}\mathbf{u}+\mathbf{t}_{g^{-1}}}\rho) \\ &= W_\rho(F_{g^{-1}}\mathbf{u}+\mathbf{t}_{g^{-1}}). \end{aligned}$$

□

As we remarked in Section 1.2.3, it is impossible to have a Wigner function that is covariant under all (real) Clifford operations. As another example, $g = H_1$ acting on a state of two rebits does not transform the Wigner function covariantly, because it does not preserve positivity. Indeed,

$$H_1\frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{|10\rangle + |01\rangle}{\sqrt{2}}. \tag{3.41}$$

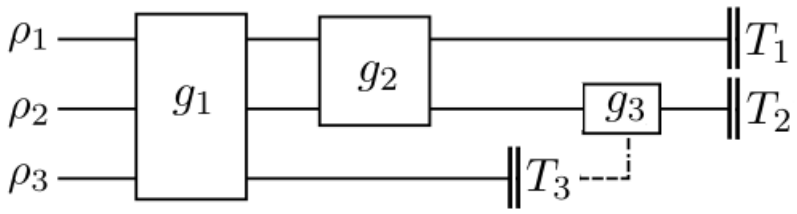


Figure 3.1: Example of a circuit whose distribution of outcomes can be sampled by our efficient classical algorithm. The gates g_i and the measurements T_i are CSS-ness preserving unitaries belonging to $\mathcal{O}$

It can be checked that the initial state has positive Wigner function, while the final state has a negative Wigner function.

3.4 Efficient classical simulation of CSS-Clifford circuits

In this section we prove the analog of Theorem 1.5 in the case of rebits. Here, instead of considering the full set of Clifford gates acting on stabilizer states as in Chapter 1.2.2, we will consider the smaller set of CSS-ness preserving Clifford gates and CSS measurements acting of CSS states, and we will call such circuits CSS-Clifford circuits. This restriction comes in play because the covariance of the Wigner function under our restricted set of gates is the crucial property to establish an efficient classical simulation method. The remainder of this section is dedicated to the proof of the following theorem

Theorem 3.9. *Every CSS-Clifford circuit acting on an initial product state $\rho = \otimes_{i=1}^{n} \rho_i$ with non-negative Wigner function can be efficiently classically simulated.*

We prove this by explicitly giving a classical simulation algorithm. Typical quantum circuits have probabilistic outcomes for the final measurements, a particular run of the circuit giving a string of outputs with some probabil-

ities. A classical simulation algorithm is an algorithm that allows efficient sampling from the probability distribution for the outcomes of the quantum circuit.

The algorithm we present is inspired from the interpretation of a positive Wigner function as being equivalent to a hidden variables model, and is very similar to the qudit algorithm of Veitch et al. [27]. Considering the phase space as the hidden variable, we can update our ontic state at each step of the quantum circuit according to the covariance of the Wigner function.

Classical simulation algorithm

1. Sample a phase point $\mathbf{u} \in V$ according to the initial Wigner function

$$W_{\rho(0)}(\mathbf{u}) = W_{\rho_1}(\mathbf{u_1}) W_{\rho_2}(\mathbf{u_2}) \dots W_{\rho_n}(\mathbf{u_n}). \tag{3.42}$$

2. For every unitary gate $g \in G_{\mathrm{CSS}}$, update the phase point according to

$$\mathbf{u} \to F_g \mathbf{u} + \mathbf{t}_g, \tag{3.43}$$

   where F_g is the symplectic matrix and $\mathbf{t}_g$ is the vector in V that describe the covariant action of g on W_ρ such that

$$W_{g\rho g^\dagger}(F_g \mathbf{u} + \mathbf{t}_g) = W_\rho(\mathbf{u}). \tag{3.44}$$

3. For every measurement of a CSS-ness preserving Clifford unitary $T_{\mathbf{a}} \in \mathcal{O}$, report the measurement outcome $s \in \{\pm 1\}$ with probability $W_{E(s)}(\mathbf{u})$, where $E(s) = \frac{1}{2}(I + sT_{\mathbf{a}})$ is the POVM element corresponding to outcome s and $\mathbf{u}$ is the phase point as it was before the measurement. Then update $\mathbf{u}$ in a probabilistic manner according to

$$\mathbf{u} \to \begin{cases} \mathbf{u} & \text{with probability } \frac{W_\rho(\mathbf{u})}{W_\rho(\mathbf{u}+\mathbf{a})} \\ \mathbf{u}+\mathbf{a} & \text{with probability } 1 - \frac{W_\rho(\mathbf{u})}{W_\rho(\mathbf{u}+\mathbf{a})}. \end{cases} \tag{3.45}$$

   It is possible to condition further steps of the computation according

to the outcome of this measurement.

To show that this algorithm is valid, we need to prove the following lemma about the Wigner function of the post-measurement state

Lemma 3.10. *The Wigner function W_ρ of the state obtained after measuring $T_{\mathbf{a}} \in \mathcal{O}$ with outcome $s \in \{\pm 1\}$ on state ρ is*

$$W_{\rho'}(\mathbf{u}) = k \begin{cases} \frac{W_\rho(\mathbf{u}) + W_\rho(\mathbf{u}+\mathbf{a})}{2} & \text{if } s(-1)^{[\mathbf{u},\mathbf{a}]} = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{3.46}$$

where k is a constant independent of $\mathbf{u}$.

Proof. The proof is by explicit calculation. After the measurement, the state of the system is

$$\rho' = \frac{(I + sT_{\mathbf{a}})\rho(I + sT_{\mathbf{a}})}{\mathrm{Tr}((I + sT_{\mathbf{a}})\rho(I + sT_{\mathbf{a}}))}. \tag{3.47}$$

So the Wigner function of the post-measurement state is

$$
\begin{aligned}
W_{\rho'}(\mathbf{u}) &\propto \frac{1}{2^n}\mathrm{Tr}\left(A_{\mathbf{u}}\frac{I+sT_{\mathbf{a}}}{2}\rho\frac{I+sT_{\mathbf{a}}}{2}\right)\\
&= \frac{1}{2^{2n}}\mathrm{Tr}\left(\frac{I+sT_{\mathbf{a}}}{2}T_{\mathbf{u}}\left[\sum_{\mathbf{v}\in\mathcal{A}}T_{\mathbf{v}}\right]T_{\mathbf{u}}^{\dagger}\frac{I+sT_{\mathbf{a}}}{2}\rho\right)\\
&= \frac{1}{2^{2n}}\mathrm{Tr}\left(T_{\mathbf{u}}\frac{I+s(-1)^{[\mathbf{u},\mathbf{a}]}T_{\mathbf{a}}}{2}\left[\sum_{\mathbf{v}\in\mathcal{A}}T_{\mathbf{v}}\right]\frac{I+s(-1)^{[\mathbf{u},\mathbf{a}]}T_{\mathbf{a}}}{2}T_{\mathbf{u}}^{\dagger}\rho\right)\\
&= \frac{1}{2^{2n}}\sum_{\mathbf{v}\in\mathcal{A}}\mathrm{Tr}\left(T_{\mathbf{u}}T_{\mathbf{v}}\frac{I+s(-1)^{[\mathbf{u}+\mathbf{v},\mathbf{a}]}T_{\mathbf{a}}}{2}\frac{I+s(-1)^{[\mathbf{u},\mathbf{a}]}T_{\mathbf{a}}}{2}T_{\mathbf{u}}^{\dagger}\rho\right)\\
&= \frac{1}{2^{2n}}\sum_{\mathbf{v}\in\mathcal{A}}\mathrm{Tr}\left(T_{\mathbf{u}}\frac{1+(-1)^{[\mathbf{v},\mathbf{a}]}}{4}T_{\mathbf{v}}\left[I+s(-1)^{[\mathbf{u},\mathbf{a}]}T_{\mathbf{a}}\right]T_{\mathbf{u}}^{\dagger}\rho\right)\\
&= \frac{1}{2^{2n+1}}\sum_{\mathbf{v}|[\mathbf{v},\mathbf{a}]=0}\mathrm{Tr}\left(T_{\mathbf{u}}T_{\mathbf{v}}\left[I+s(-1)^{[\mathbf{u},\mathbf{a}]}T_{\mathbf{a}}\right]T_{\mathbf{u}}^{\dagger}\rho\right)\\
&= \frac{1}{2^{2n+1}}\left[\sum_{\mathbf{v}|[\mathbf{v},\mathbf{a}]=0}\mathrm{Tr}\left(T_{\mathbf{u}}T_{\mathbf{v}}T_{\mathbf{u}}^{\dagger}\rho\right)+s(-1)^{[\mathbf{u},\mathbf{a}]}\sum_{\mathbf{v}|[\mathbf{v},\mathbf{a}]=0}\mathrm{Tr}\left(T_{\mathbf{u}}T_{\mathbf{v}+\mathbf{a}}T_{\mathbf{u}}^{\dagger}\rho\right)\right]\\
&= \frac{1}{2^{2n+1}}\left[\sum_{\mathbf{v}|[\mathbf{v},\mathbf{a}]=0}\mathrm{Tr}\left(T_{\mathbf{u}}T_{\mathbf{v}}T_{\mathbf{u}}^{\dagger}\rho\right)+s(-1)^{[\mathbf{u},\mathbf{a}]}\sum_{\mathbf{v}|[\mathbf{v},\mathbf{a}]=0}\mathrm{Tr}\left(T_{\mathbf{u}}T_{\mathbf{v}}T_{\mathbf{u}}^{\dagger}\rho\right)\right]\\
&= \frac{1}{2^{2n+1}}(1+s(-1)^{[\mathbf{u},\mathbf{a}]})\sum_{\mathbf{v}|[\mathbf{v},\mathbf{a}]=0}\mathrm{Tr}\left(T_{\mathbf{u}}T_{\mathbf{v}}T_{\mathbf{u}}^{\dagger}\rho\right)\\
&= \frac{1}{2^{2n+2}}(1+s(-1)^{[\mathbf{u},\mathbf{a}]})\sum_{\mathbf{v}\in\mathcal{A}}\mathrm{Tr}\left(T_{\mathbf{u}}\left[T_{\mathbf{v}}+T_{\mathbf{a}}T_{\mathbf{v}}T_{\mathbf{a}}\right]T_{\mathbf{u}}^{\dagger}\rho\right)\\
&= \frac{1}{4}(1+s(-1)^{[\mathbf{u},\mathbf{a}]})\mathrm{Tr}\left(\left[A_{\mathbf{u}}+A_{\mathbf{u}+\mathbf{a}}\right]\rho\right)\\
&= \frac{1}{2}\delta_{s,(-1)^{[\mathbf{u},\mathbf{a}]}}\mathrm{Tr}\left(\left[A_{\mathbf{u}}+A_{\mathbf{u}+\mathbf{a}}\right]\rho\right)\\
&= \frac{1}{2}\delta_{s,(-1)^{[\mathbf{u},\mathbf{a}]}}\left(W_{\rho}(\mathbf{u})+W_{\rho}(\mathbf{u}+\mathbf{a})\right).
\end{aligned}
$$

The normalization for W'_{ρ} is not obvious to calculate, but it is of no importance for the validity of the classical simulation, as the transition probabilities only depend on the ratio $\frac{W_{\rho}(\mathbf{u})}{W_{\rho}(\mathbf{u}+\mathbf{a})}$. □

The previous lemma guarantees that the non-negativity of the Wigner function is preserved under measurements of observables in $\mathcal{O}$. However, in contrast with the qudit case, non-negativity of the POVM and of the Wigner function is not enough to guarantee that the output state has a non-negative Wigner function. For instance, both $\rho = (I + X_1 Z_2)/4$ and the POVM element $E = (1 + Z_1 X_2)/4$ are positively represented, but the state after measurement is $\rho' = (1+X_1 Z_2)(1+Z_1 X_2)/4$, which has negative Wigner function. Indeed,

$$A_{((1,1),(0,0))} = \frac{1}{4}\left(I - X_1 Z_2 - Z_1 X_2 + Z_1 Z_2 X_1 X_2 + \text{irrelevant terms}\right),$$

so

$$W_{\rho'}((1,1),(0,0)) = \frac{1}{2^n}\left(1 - 1 - 1 - 1\right) < 0.$$

Chapter 4

Contextuality for rebits

In this chapter, we prove that there is a link between contextuality and universal quantum computation. We find that negativity of the Wigner function is necessary for contextuality of the state, but that it is not sufficient. We prove the existence of a condition that guarantees contextuality of the state, provided its Wigner function is sufficiently negative, in a well-defined sense. We study a couple of examples that illustrate the situation, and how it differs from the case of qudits. Finally, we address a reasonable objection concerning the Mermin square and state-independent contextuality in rebit QCSI.

4.1 Hidden-variable models of rebit QCSI

In Section 1.4.4 we discussed that some subtheories of quantum mechanics may or may not be contextual. In particular, some subtheories exhibit state-dependent contextuality. As our scheme of rebit quantum computation is indeed a subtheory of quantum mechanics, we should first endeavor to identify what are the features that should be reproduced by a non-contextual hidden variable model (NCHVM). First, we want our HVM to reproduce the correct measurement statistics for any rebit quantum state, but it is not immediately obvious what set of measurements should be reproduced by the HVM.

Recall that in the scheme of rebit QCSI, the allowed measurements were the CSS Pauli operators

$$\mathcal{O} = \{X(\mathbf{a_x}), Z(\mathbf{a_z}) |\, \mathbf{a_x}, \mathbf{a_z} \in Z_2^n\}. \tag{4.1}$$

We thus want our HVM to reproduce the measurement statistics for all observables in $\mathcal{O}$, and to obey the value consistency condition on commuting observables

$$\lambda(A)\lambda(B) = \lambda(AB) \quad \forall A, B \in \mathcal{O} \text{ such that } [A, B] = 0. \tag{4.2}$$

Can we perform a measurement of observables in $\mathcal{A}$ that do not belong to $\mathcal{O}$? Yes, for instance, if we want to measure $X_1 Z_2$ on a system of two rebits, we can first measure $X_1 \in \mathcal{O}$,then measure $Z_2 \in \mathcal{O}$, and finally multiply the two outcomes together. The key point here is that set of observables for our measurement $M = \{X_1 Z_2\}$ can be decomposed into a set of commuting observables in $\mathcal{O}$ as $M' = \{X_1, Z_2\}$.

Let's see what happens in an example where this decomposition fails to hold. For instance, the observables in $M = \{X_1 Z_2, Z_1 X_2\}$ commute, so regular quantum mechanics states that simultaneous measurement is possible. But if as previously, we attempt to decompose M into observables of $\mathcal{O}$, we get $M' = \{X_1, Z_1, X_2, Z_2\}$, which clearly is not a commuting set of observables. These observables cannot be measured without changing the state of the system, and therefore $M = \{X_1 Z_2, Z_1 X_2\}$ is not a possible measurement in rebit QCSI.

Let us formalize the above discussion with the following definitions.

Definition 4.1. (Allowable measurements) The set of commuting observables $M \subset \mathcal{A}$ is an allowable measurement if there exists a commuting set of observables $M' \subset \mathcal{O}$ such that the outcomes of M can be computed by knowing the outcomes of M'. The set of all allowed measurement settings is denoted by $\mathcal{M}$

Given two operators $T_{\mathbf{a}}$ and $T_{\mathbf{b}}$, we already know that

$$T_{\mathbf{a}}T_{\mathbf{b}} = (-1)^{(\mathbf{u_x},\mathbf{b_z})}T_{\mathbf{a+b}}.$$

The following lemma gives a simplification of this property for observables in $\mathcal{M}$ that will be very useful in the next sections.

Lemma 4.2. *Let $M \subset \mathcal{A}$ be a set of commuting observables. Then $M \in \mathcal{M}$ if and only if $T_{\mathbf{a}}T_{\mathbf{b}} = T_{\mathbf{a+b}}$*

Proof.

$$\begin{aligned} T_{\mathbf{a}}, T_{\mathbf{b}} \in M \subset \mathcal{M} &\iff \{X(\mathbf{a}), X(\mathbf{b}), Z(\mathbf{a}), Z(\mathbf{b})\} \text{ all commute} \\ &\iff (\mathbf{u_x}, \mathbf{b_z}) = (\mathbf{u_z}, \mathbf{b_x}) = 0 \\ &\iff T_{\mathbf{a}}T_{\mathbf{b}} = +T_{\mathbf{a+b}} \quad (\text{and } [T_{\mathbf{a}}, T_{\mathbf{b}}] = 0) \end{aligned}$$

□

Applying this lemma to an example we previously considered, $\{X_1, Z_2, Z_1X_2\} \notin \mathcal{M}$ because if $T_{\mathbf{a}} = X_1Z_2$ and $T_{\mathbf{b}} = Z_1X_2$, then

$$T_{\mathbf{a}}T_{\mathbf{b}} = -Y_1Y_2 = -T_{\mathbf{a+b}}. \tag{4.3}$$

We now give a mathematical description of the features that are required to be present in a NCHVM of rebit QCSI.

Definition 4.3. (Non-contextual hidden variables model) A HVM describing the quantum state ρ and the set of measurement settings $\mathcal{M}$ consists of:

- A set of internal (ontic) states $\mathcal{S}$.
- A probability distribution q_ρ over $\mathcal{S}$.
- Conditional probabilities $p(\mathbf{s}_M|\mathbf{u}), \mathbf{u} \in \mathcal{S}$ that describe the probabilities of the outcome $\mathbf{s}_M = (s_1, s_2, \ldots, s_{|M|})$ of a measurement in $M \subset \mathcal{M}$.

These objects are required to obey the following conditions

1. There is a value assignment function λ, which assigns for every $\mathbf{u} \in \mathcal{S}$ a definite value to all observables $O \in \mathcal{A}$, $\lambda_{\mathbf{u}}(O) = \pm 1$. For all $M \in \mathcal{M}$, the conditional probabilities take the simple form

$$p(\mathbf{s}_M|\mathbf{u}) = \prod_{i|O_i \in M} \delta_{s_i, \lambda_{\mathbf{u}}(O_i)}. \tag{4.4}$$

2. (Consistency of value assignments) For all $M \in \mathcal{M}$, and all pairs of observables $A, B \in M$ such that $AB \in M$,

$$\lambda_{\mathbf{u}}(A)\lambda_{\mathbf{u}}(B) = \lambda_{\mathbf{u}}(AB) \quad \forall \mathbf{u} \in \mathcal{S}. \tag{4.5}$$

3. (Reproduces quantum mechanics) Let $p_{M,\rho}$ be the probability distribution that quantum mechanics predicts for the measurements outcomes in M, given an initial state ρ. Then the probabilities of the HVM are such that for all $M \subset \mathcal{M}$ and all possible measurement outcomes $\mathbf{s_M}$,

$$p_{M,\rho}(\mathbf{s}_M) = \sum_{\mathbf{u} \in \mathcal{S}} p(\mathbf{s_M}|\mathbf{u}) q_\rho(\mathbf{u}). \tag{4.6}$$

Lemma 4.4. *For any NCHVM of a n-rebit setting $(\rho, \mathcal{M})$, $\mathcal{S} = \mathbb{Z}_2^{2n}$, and for all $\mathbf{u} \in \mathcal{S}$,*

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u},\mathbf{a}]} \quad \forall T_{\mathbf{a}} \in \mathcal{A}, \tag{4.7}$$

so that

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}+\mathbf{b}}) = \lambda_{\mathbf{u}}(T_{\mathbf{a}})\lambda_{\mathbf{u}}(T_{\mathbf{b}}) \quad \forall T_{\mathbf{a}}, T_{\mathbf{b}}, T_{\mathbf{a}+\mathbf{b}} \in \mathcal{A}, \tag{4.8}$$

or equivalently

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}+\mathbf{b}}) = \lambda_{\mathbf{u}}(T_{\mathbf{a}})\lambda_{\mathbf{u}}(T_{\mathbf{b}}) \quad \forall T_{\mathbf{a}}, T_{\mathbf{b}} \in \mathcal{A} \text{ such that } [T_{\mathbf{a}}, T_{\mathbf{b}}] = 0. \tag{4.9}$$

Proof. Both sets of observables $M_X = \{X(\mathbf{a}_X)|\mathbf{a}_X \in \mathbb{Z}_2^n\}$ and $M_Z = \{Z(\mathbf{a}_Z)|\mathbf{a}_Z \in \mathbb{Z}_2^n\}$ are allowable sets of measurements, so the consistency condition (2) of Definition 4.3 requires that the values of these observables

are determined by the single rebit observables:

$$\lambda_{\mathbf{u}}(Z(\mathbf{a}_Z)) = \prod_{i=1}^{n} \lambda_{\mathbf{u}}(Z_i) \tag{4.10}$$

$$\lambda_{\mathbf{u}}(X(\mathbf{a}_X)) = \prod_{i=1}^{n} \lambda_{\mathbf{u}}(X_i). \tag{4.11}$$

Furthermore, for any $T_{(\mathbf{a}_X,\mathbf{a}_Z)} \in \mathcal{A}$, the set of observables

$$M = \{Z(\mathbf{a}_Z), X(\mathbf{a}_X), T_{(\mathbf{a}_X,\mathbf{a}_Z)}\} \tag{4.12}$$

is an allowable measurement, because $[X(\mathbf{a}_X), Z(\mathbf{a}_Z))] = 0$, by the definition of $\mathcal{A}$. Since also $Z(\mathbf{a}_Z)X(\mathbf{a}_X) = T_{(\mathbf{a}_X,\mathbf{a}_Z)}$, consistency of the value assignment gives

$$\lambda_{\mathbf{u}}(T_{(\mathbf{a}_X,\mathbf{a}_Z)}) = \lambda_{\mathbf{u}}(Z(\mathbf{a}_Z))\lambda_{\mathbf{u}}(X(\mathbf{a}_X)). \tag{4.13}$$

Using this equation in conjunction with equations 4.10 and 4.11, we see that for all observables in $\mathcal{A}$ their value is determined by the product of individual local observables X_i and Z_i. Since these local observables take values ± 1, the set of internal states $\mathcal{S}$ has dimension 2^{2n} and is isomorphic (as a set) to Z_2^{2n}. Also, characters span the space of linear functions from $\mathbb{Z}_2^{2n}$ to $\{\pm 1\}$, so we conclude that we can write

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u},\mathbf{a}]} \quad \forall \mathbf{u} \in \mathcal{S}. \tag{4.14}$$

To finish the proof, we can see that Equation 4.8 follows directly from this form for the $\lambda_{\mathbf{u}}$. Equation 4.9 also follows by noticing that if $T_{\mathbf{a}}, T_{\mathbf{b}} \in \mathcal{A}$, then $[T_{\mathbf{a}}, T_{\mathbf{b}}] \iff T_{\mathbf{a}+\mathbf{b}} \in \mathcal{A}$.

□

4.2 Conditions for contextuality

4.2.1 A necessary condition for contextuality

As it was the case for qudits, the Wigner function is a valid NCHVM when it is non-negative. We therefore have the following necessary condition for contextuality

Theorem 4.5. *The physical setting* $(\rho, \mathcal{M})$ *is contextual only if* $W_\rho(\mathbf{u}) < 0$ *for some* $\mathbf{u} \in V$.

Proof. The proof proceeds by showing that if the Wigner function is non-negative, then it satisfies all the points in Definition 4.3 of a NCHVM. We take the set of internal states as equal to the phase space: $\mathcal{S} = V$, and the probability distribution q_ρ over $\mathcal{S}$ is the Wigner function W_ρ, which is a valid probability distribution because it is non-negative by our assumption. The measurement of a set $M = \{T_{\mathbf{a}(i)}\} \in \mathcal{M}$ is represented by POVM elements $E(\mathbf{s}_M)$ associated with the measurement outcomes. Their general form is

$$E(\mathbf{s}_M) = \prod_i \frac{I + s_i T_{\mathbf{a}(i)}}{2}. \tag{4.15}$$

For simplicity, let us first consider the POVM element $E(s)$ corresponding to the measurement of $T_\mathbf{a}$ with outcome $s \in \{\pm 1\}$

$$E(s) = \frac{1}{2}\left(I + sT_\mathbf{a}\right). \tag{4.16}$$

Then

$$\begin{aligned} W_{E(s)}(\mathbf{u}) &= \frac{1}{2^{n+1}} \mathrm{Tr}\left(A_\mathbf{u}(I + sT_\mathbf{a})\right) && (4.17)\\ &= \frac{1}{2^{2n+1}} \sum_{b \in \mathcal{A}} \mathrm{Tr}\left((-1)^{[\mathbf{u},\mathbf{b}]} T_\mathbf{b}(I + sT_\mathbf{a})\right) && (4.18)\\ &= \frac{1}{2^{n+1}} \left(1 + s(-1)^{[\mathbf{u},\mathbf{a}]}\right) && (4.19)\\ &= \frac{1}{2^n} \delta_{s,(-1)^{[\mathbf{u},\mathbf{a}]}}(\mathbf{u}). && (4.20) \end{aligned}$$

We can directly check that by defining $p(\mathbf{s_M}|\mathbf{u}) = 2^n W_{E(s)}(\mathbf{u})$, and $\lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u},\mathbf{a}]}$, we satisfy points 1 and 3 of Definition 4.3.

For a general POVM $E(\mathbf{s}_M) = E_1(s_1)E_2(s_2)\dots E_{|M|}(s_{|M|})$, the Wigner function factorizes as

$$2^n W_{E(\mathbf{s}_M)} = 2^{n|M|} W_{E_1(s_1)} W_{E_2(s_2)} \dots W_{E_{|M|}(s_{|M|})}. \tag{4.21}$$

This can be proved recursively by using Lemma 4.2:

$$\begin{aligned} W_{E_1(s_1)E_2(s_2)}(\mathbf{u}) &= \frac{1}{2^{n+2}} \mathrm{Tr}\left(A_{\mathbf{u}}(I + s_1 T_{\mathbf{a}})(I + s_2 T_{\mathbf{b}})\right) \\ &= \frac{1}{2^{2n+2}} \sum_{\mathbf{c}\in\mathcal{A}} \mathrm{Tr}\left((-1)^{[\mathbf{u},\mathbf{c}]} T_{\mathbf{c}}(I + s_1 T_{\mathbf{a}})(I + s_2 T_{\mathbf{b}})\right) \\ &= \frac{1}{2^{n+2}} \left(1 + s_1 (-1)^{[\mathbf{u},\mathbf{a}]}\right)\left(1 + s_2 (-1)^{[\mathbf{u},\mathbf{b}]}\right) \\ &= \frac{1}{2^n} \delta_{s_1,(-1)^{[\mathbf{u},\mathbf{a}]}}(\mathbf{u}) \delta_{s_2,(-1)^{[\mathbf{u},\mathbf{b}]}}(\mathbf{u}). \end{aligned}$$

Therefore, we obtain that

$$2^n W_{E(\mathbf{s}_M)}(\mathbf{u}) = \prod_{i|T_{\mathbf{a}(i)}} \delta_{s_i,(-1)^{[\mathbf{u},\mathbf{a}(i)]}}. \tag{4.22}$$

Furthermore,

$$p_{M,\rho}(\mathbf{s}_M) = \mathrm{Tr}(E(\mathbf{s}_M)\rho) = 2^n \sum_{\mathbf{u}\in V} W_{E(\mathbf{s}_M)}(\mathbf{u}) W_\rho(\mathbf{u}). \tag{4.23}$$

And we see that $p(\mathbf{s}_M|\mathbf{u}) = 2^n W_{E(\mathbf{s}_M)}(\mathbf{u})$ is indeed a valid conditional probability in the sense of Definition 4.3. It only remain to check point 2 of the definition, concerning the consistency of the assignment. Since we have $\lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u},\mathbf{a}]}$ and $T_{\mathbf{a}}T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}}$ for all allowable measurements,

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}}T_{\mathbf{b}}) = \lambda_{\mathbf{u}}(T_{\mathbf{a}+\mathbf{b}}) = (-1)^{[\mathbf{u},\mathbf{a}+\mathbf{b}]} = \lambda_{\mathbf{u}}(T_{\mathbf{a}})\lambda_{\mathbf{u}}(T_{\mathbf{b}}). \tag{4.24}$$

We thus see that a non-negative W_ρ fulfills all requirements of a NCHVM. □

4.2.2 A sufficient condition for contextuality

We now turn to the proof of a sufficient condition for contextuality. Perhaps surprisingly, the necessary condition of the previous section and the sufficient condition do not coincide. Indeed, as will be shown in Section 4.3.1, there exists states with negative Wigner function that are non-contextual.

Our derivation of the sufficient condition will first require constructing a family of contextuality witnesses. Contextuality witnesses are linear functions of the quantum state ρ that take positive values for all NCHVMs, but that are allowed by quantum mechanics to take negative values. To motivate the construction of the witnesses, let us got back to an example we previously considered of a system of two rebits and the observables $T_{\mathbf{a}} = X_1 Z_2$ and $T_{\mathbf{b}} = Z_1 X_2$. We know already from Section 4.1 that these two operators cannot be simultaneously measured in rebit QCSI, but we are still allowed to calculate their expectation values. Define the witness

$$\mathcal{W}_\rho = \langle I + T_{\mathbf{a}} + T_{\mathbf{b}} + T_{\mathbf{a}+\mathbf{b}} \rangle_\rho = \langle I + X_1 Z_2 + Z_1 X_2 - Y_1 Y_2 \rangle_\rho. \tag{4.25}$$

This function can take negative values. Indeed, if $\rho = |K_2\rangle \langle K_2|$, where $|K_2\rangle$ is the stabilizer state with stabilizer $\langle -X_1 Z_2, -Z_1 X_2 \rangle$, then $\mathcal{W}_\rho = -2$. But now consider the values that $\mathcal{W}$ can take if we restrict ourselves to NCHVMs. The compatibility condition for the value assignments requires that

$$\begin{aligned} \lambda(X_1 Z_2) &= \lambda(X_1)\lambda(Z_2) \\ \lambda(Z_1 X_2) &= \lambda(Z_1)\lambda(X_2) \\ \lambda(Y_1 Y_2) &= \lambda(X_1)\lambda(X_2)\lambda(Z_1)\lambda(Z_2). \end{aligned}$$

Therefore, for any internal state in $\mathcal{S}$, the witness evaluates to $(1+\lambda(X_1)\lambda(Z_2))(1+\lambda(Z_1)\lambda(X_2)) \geq 0$. This is why $\mathcal{W}$ is a contextuality witness; in particular it identifies $|K_2\rangle$ as possessing state dependent contextuality.

We can generalize this construction to create a large family of witness functions. Looking back, we see that the important necessary property is that we find a subspace U such that for all $\mathbf{a}, \mathbf{b} \in U$, $\lambda(T_{\mathbf{a}+\mathbf{b}}) = \lambda(T_{\mathbf{a}})\lambda(T_{\mathbf{b}})$. Lemma 4.4 tells us that this will be the case if all $[T_{\mathbf{a}}, T_{\mathbf{b}}] = 0 \; \forall \mathbf{a}, \mathbf{b} \in U$, i.e.

if U is an isotropic subspace of V. We have thus motivated the statement of the following lemma.

Lemma 4.6. *The n-rebit setting $(\rho, \mathcal{M})$ is contextual if there exists an isotropic subspace $U \subset V$ with a basis $\mathcal{B}(U) = \{\mathbf{a}(1), \mathbf{a}(2), \ldots, \mathbf{a}(m)\}$ such that the associated witness function*

$$\mathcal{W}_{\rho}^{\mathcal{B}(U)}(\mathbf{x}) = \left\langle \sum_{\mathbf{z} \in \mathbb{Z}_2^m} \left[\prod_{i=1}^{m} (-1)^{z_i x_i} \right] T_{\sum_i z_i \mathbf{a}(i)} \right\rangle_{\rho}, \tag{4.26}$$

which is defined for arguments $\mathbf{x} \in \mathbb{Z}_2^m$, takes at least one negative value.

The factors $(-1)^{z_i x_i}$ are there to allow all possible signs in the sum of operators that defines the witness.

Proof. We show that the witness $W_{\rho}^{\mathcal{B}(U)}$ is indeed a contextuality witness, by assuming a NCHVM and showing that it can only take non-negative values. Evaluating the witness function for one particular state $\mathbf{u} \in \mathcal{S}$, we get

$$\mathcal{W}_{\lambda_{\mathbf{u}}}^{\mathcal{B}(U)}(\mathbf{x}) = \sum_{\mathbf{z} \in \mathbb{Z}_2^m} \left[\prod_{i=1}^{m} (-1)^{z_i x_i} \right] \lambda_{\mathbf{u}} \left(T_{\sum_i z_i \mathbf{a}(i)} \right). \tag{4.27}$$

Applying the fact that U is isotropic with Lemma 4.4 allows us to factorize the witness as

$$\begin{aligned} \mathcal{W}_{\lambda_{\mathbf{u}}}^{\mathcal{B}(U)}(\mathbf{x}) &= \sum_{\mathbf{z} \in \mathbb{Z}_2^m} \prod_{i=1}^{m} \left[(-1)^{x_i} \lambda_{\mathbf{u}} \left(T_{\mathbf{a}(i)} \right) \right]^{z_i} && (4.28) \\ &= \prod_{i=1}^{m} \left[(-1)^{x_i} \lambda_{\mathbf{u}} \left(T_{\mathbf{a}(i)} \right) \right] && (4.29) \\ &\geq 0. && (4.30) \end{aligned}$$

Since the witness is positive for all $\mathbf{u} \in \mathcal{S}$, but can take negative values in quantum mechanics, this concludes the proof. □

The following lemma establishes a relationship between the witnesses and the Wigner function

Lemma 4.7. *Let U be an isotropic subspace of $\mathbb{Z}_2^{2n}$ with basis $\mathcal{B}(U) = \{\mathbf{a}(1), \mathbf{a}(2), \ldots, \mathbf{a}(m)\}$. Then there exists a set $\tilde{\mathcal{B}} = \{\mathbf{b}(1), \mathbf{b}(2), \ldots, \mathbf{b}(m)\}$ of vectors of $\mathbb{Z}_2^{2n}$ such that $[\mathbf{a}(i), \mathbf{b}(j)] = \delta_{ij}\ \forall i, j \in \{1, ..., m\}$. For every $\eta(\mathbf{x}) = \sum_i x_i \mathbf{a}(i) \in U$, define the vector $\bar{\eta}(\mathbf{x}) = \sum_i x_i \mathbf{b}(i)$. Then*

$$\mathcal{W}_\rho^{\mathcal{B}(U)}(\eta(\mathbf{x})) = 2^m \sum_{\mathbf{v} \in U^\perp} W_\rho(\mathbf{v} + \bar{\eta}(\mathbf{x})). \tag{4.31}$$

Proof. We can use the operators T_η to reproduce the factor $\prod_i (-1)^{x_i z_i}$ in Equation 4.26, and obtain

$$\mathcal{W}_\rho^{\mathcal{B}(U)}(\eta(\mathbf{x})) = \left\langle T_{\bar{\eta}} \left(\sum_{\mathbf{u} \in U} T_{\mathbf{u}} \right) T_{\bar{\eta}}^\dagger \right\rangle_\rho . \tag{4.32}$$

We can simplify the above expression by noticing that

$$\begin{aligned}
\sum_{\mathbf{u} \in U} T_{\mathbf{u}} &= \sum_{\mathbf{u} \in \mathbb{Z}_2^{2n}} \delta_{\mathbf{u}, U} T_{\mathbf{u}} && (4.33)\\
&= \frac{1}{|U^\perp|} \sum_{\mathbf{u} \in \mathbb{Z}_2^{2n}} \sum_{\mathbf{v} \in U^\perp} (-1)^{[\mathbf{u}, \mathbf{v}]} T_{\mathbf{u}} && (4.34)\\
&= \frac{2^m}{2^{2n}} \sum_{\mathbf{u} \in \mathbb{Z}_2^{2n}} \sum_{\mathbf{v} \in U^\perp} T_{\mathbf{v}} T_{\mathbf{u}} T_{\mathbf{v}}^\dagger && (4.35)\\
&= \frac{2^m}{2^{2n}} \sum_{\mathbf{v} \in U^\perp} T_{\mathbf{v}} A_{\mathbf{0}} T_{\mathbf{v}}^\dagger . && (4.36)
\end{aligned}$$

Substituting into the expression for the witness $\mathcal{W}$, this gives

$$\begin{aligned}
\mathcal{W}_\rho^{\mathcal{B}(U)}(\eta(\mathbf{x})) &= \frac{2^m}{2^{2n}} \left\langle \sum_{\mathbf{v} \in U^\perp} T_{\bar{\eta}} T_{\mathbf{v}} A_{\mathbf{0}} T_{\mathbf{v}}^\dagger T_{\bar{\eta}}^\dagger \right\rangle_\rho && (4.37)\\
&= \frac{2^m}{2^n} \sum_{\mathbf{v} \in U^\perp} \mathrm{Tr}\left(A_{\bar{\eta} + \mathbf{v}} \rho \right) && (4.38)\\
&= 2^m \sum_{\mathbf{v} \in U^\perp} W_\rho(\mathbf{v} + \bar{\eta}). && (4.39)
\end{aligned}$$

□

We can finally prove the following theorem that gives a sufficient condition for contextuality

Theorem 4.8. *The n-rebit setting $(\rho, \mathcal{M})$ is contextual if there exists a vector $\mathbf{t} \in V$ and a maximally isotropic $U \subset V$ such that*

$$\sum_{\mathbf{v} \in U} W_\rho(\mathbf{v} + \mathbf{t}) < 0. \tag{4.40}$$

Proof. Putting Lemmas 4.6 and 4.7 together, we obtain that $(\rho, \mathcal{M})$ is contextual if there exists an isotropic subspace $U \subset V$, and a vector $\mathbf{t} \in V$ such that

$$\sum_{\mathbf{v} \in U^\perp} W_\rho(\mathbf{v} + \mathbf{t}) < 0. \tag{4.41}$$

We are almost done, but we can make the condition stronger by considering maximally isotropic subspaces. Every isotropic space U can be embedded in a maximally isotropic subspace, $U_M = U_M^\perp$, and furthermore there exists a space $\bar{U} \subset V$ such that

$$U^\perp = U_M \oplus \bar{U}. \tag{4.42}$$

Therefore,

$$\sum_{\mathbf{v} \in U^\perp} W_\rho(\mathbf{v} + \mathbf{t}) = \sum_{\mathbf{u} \in \bar{U}} \left(\sum_{\mathbf{w} \in U_M} W_\rho(\mathbf{u} + \mathbf{w} + \mathbf{t}) \right). \tag{4.43}$$

In order for the left hand side to be negative, the bracketed terms in the above equation need take negative values. This means that there exists a $t' \in V$ such that

$$\sum_{\mathbf{w} \in U_M} W_\rho(\mathbf{w} + \mathbf{t}') < 0, \tag{4.44}$$

which proves the claim □

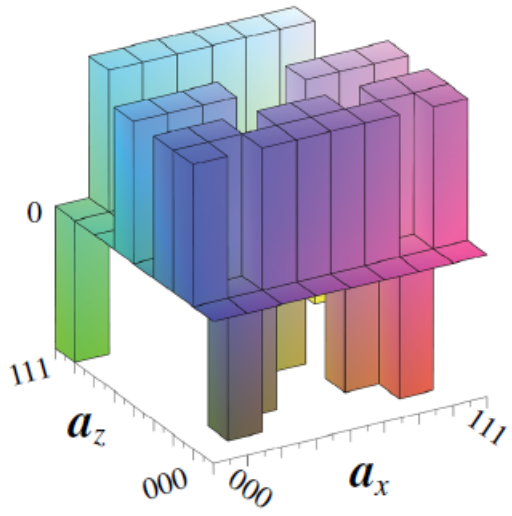


Figure 4.1: Wigner function for the three-rebit GHZ-like state $\rho = \frac{1}{8}(I - X_1Z_2Z_3)(I - Z_1X_2Z_3)(I - Z_1Z_2X_3)$, which appears in the state-dependent version of the Mermin star proof

4.2.3 Example: Mermin star

As an example of Theorem 4.8, we use it to prove the state-dependent contextuality of the Mermin star, which we have discussed already in Section 1.4.3. Consider the three-rebit GHZ-like state

$$\rho = \frac{1}{8}(I - X_1Z_2Z_3)(I - Z_1X_2Z_3)(I - Z_1Z_2X_3) := \frac{1}{8}(I - T_{\mathbf{a}})(I - T_{\mathbf{b}})(I - T_{\mathbf{c}}), \tag{4.45}$$

whose Wigner function is shown in Figure 4.1.

Let us take the maximally isotropic subspace $U = \mathrm{span}(\{\mathbf{a}, \mathbf{b}, \mathbf{c}\})$ and the vector $\mathbf{t} = 0$. The following relations can be easily checked to hold

$$\begin{aligned} T_{\mathbf{a}}T_{\mathbf{b}} &= -T_{\mathbf{a+b}} & T_{\mathbf{a}}T_{\mathbf{c}} &= -T_{\mathbf{a+c}} \\ T_{\mathbf{b}}T_{\mathbf{c}} &= -T_{\mathbf{b+c}} & T_{\mathbf{a}}T_{\mathbf{b}}T_{\mathbf{c}} &= T_{\mathbf{a+b+c}}. \end{aligned}$$

These relations allow us to write $\rho = \frac{1}{8}(I - T_{\mathbf{a}} - T_{\mathbf{b}} - T_{\mathbf{c}} - T_{\mathbf{a+b}} - T_{\mathbf{a+c}} -$

$T_{\mathbf{b}+\mathbf{c}} - T_{\mathbf{a}+\mathbf{b}+\mathbf{c}}$). With this, we obtain that

$$\begin{aligned}\sum_{\mathbf{u}\in U} W_\rho(\mathbf{u}) &= \frac{1}{2^{2n}} \sum_{\mathbf{u}\in U} \sum_{\mathbf{v}\in\mathcal{A}} (-1)^{[\mathbf{u},\mathbf{v}]} \mathrm{Tr}(T_\mathbf{v}\rho) \\ &= \frac{1}{2^{2n}} \sum_{\mathbf{u}\in U} \sum_{\mathbf{v}\in\mathcal{A}} \mathrm{Tr}(T_\mathbf{v}\rho) \\ &= \frac{1}{2^{2n}} \left(\frac{-6}{8}\right) \\ &< 0,\end{aligned}$$

so the setting is contextual by virtue of Theorem 4.8.

4.3 Contextuality and negativity

4.3.1 Negativity and contextuality are not equivalent

The necessary condition for contextuality of Theorem 4.5 and the sufficient condition of Theorem 4.8 do not match, as we have already alluded to.

Theorem 4.5 establishes that all contextual setting have states with negative Wigner function. However, if we know the Wigner W_ρ of a state, merely possessing negativity is not enough to guarantee contextuality, and Theorem 4.8 tells us how much negativity we need. The necessary condition for contextuality of Theorem 4.5 and the sufficient condition of Theorem 4.8 do not match, but is it only because of our failure to find a stronger sufficient condition? It turns out that it is not the case, and we will illustrate this by studying family of all one-rebit states

$$\tilde{\rho}(x,z) = \frac{I + xX + zZ}{2}. \tag{4.46}$$

The requirement that $\mathrm{Tr}(\rho) \leq 1$ for physical states enforces gives a constraint $x^2 + z^2 \leq 1$ on the range of the parameters x and z. Figure 4.2 is a phase diagram for the states $\tilde{\rho}(x,y)$, showing the regions of in parameter space that Theorems 4.5 and 4.8 single out, as well as set of states with negative Wigner function. State with $|x| + |z| \leq 1$ have non-negative

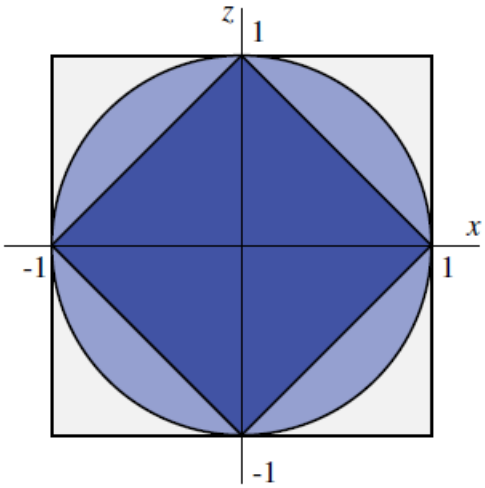


Figure 4.2: Phase diagram for the space of one-rebit states $\tilde{\rho}$. The dark blue area corresponds to states with positive Wigner function, which are non-contextual by virtue of Theorem 4.5. The pale blue region shows the physical states that have negative Wigner function. The states that are identified as contextual by Theorem 4.8 fall outside the pale grey square.

Wigner functions, so Theorem 4.5 identifies these as non-contextual. However, states falling in the pale blue region have negative Wigner function, but are not identified as contextual by Theorem 4.8. Indeed, the latter theorem does not identify any one-rebit state as contextual. This should not be as surprise, since in Section 1.4.3, we pointed out the fact that single qubits can be described by a NCHVM. The take home message from this discussion is that for rebits, Wigner function negativity and contextuality are not equivalent.

4.3.2 States for which negativity and contextuality coincide

Let us consider the family of two-rebit states defined by

$$\rho(a,b) = \frac{(I + aX_1Z_2)(I + bZ_1X_2)}{4}. \tag{4.47}$$

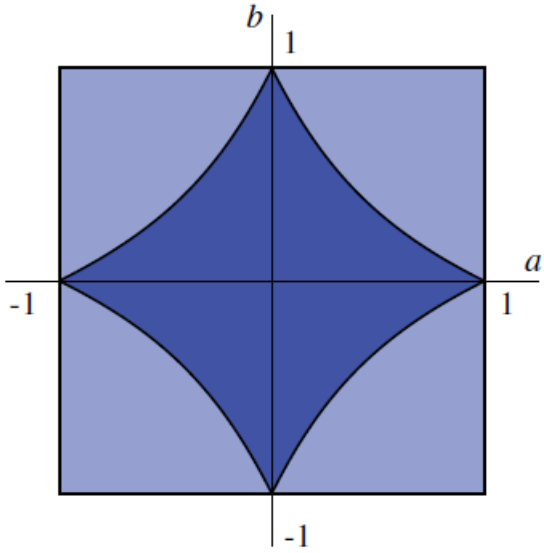


Figure 4.3: Phase diagram for the two-rebit states $\rho(a, b)$ defined by equation 4.47. All states in the square are physical, and states in the dark blue area are non-contextual.

Let $\mathbf{a}, \mathbf{b} \in V$ be such that $T_{\mathbf{a}} = X_1 Z_2$ and $T_{\mathbf{b}} = Z_1 X_2$. Then, the necessary condition for contextuality, that W_ρ be negative, reads as

$$\begin{aligned} W_\rho(\mathbf{u}) &= \frac{1}{2^{2n}} \sum_{\mathbf{c} \in \mathcal{A}} \mathrm{Tr}\left((-1)^{[\mathbf{u},\mathbf{c}]} T_{\mathbf{a}} \rho\right) \\ &= \frac{1}{2^n}\left(1 + (-1)^{[\mathbf{u},\mathbf{a}]} a + (-1)^{[\mathbf{u},\mathbf{b}]} b - (-1)^{[\mathbf{u},\mathbf{a}+\mathbf{b}]} ab\right) \\ &= \frac{1}{2^n}\left(1 + (-1)^{[\mathbf{u},\mathbf{a}]} a + (-1)^{[\mathbf{u},\mathbf{b}]} b - (-1)^{[\mathbf{u},\mathbf{a}]}(-1)^{[\mathbf{u},\mathbf{b}]} ab\right) \qquad (4.48) \\ &= \frac{1}{2^n}\left(1 + \alpha(\mathbf{u}) a + \beta(\mathbf{u}) b - \alpha(\mathbf{u})\beta(\mathbf{u}) ab\right) < 0, \qquad (4.49) \end{aligned}$$

where in the last line we have defined $\alpha(\mathbf{u}) = (-1)^{[\mathbf{u},\mathbf{a}]}$ and $\beta(\mathbf{u}) = (-1)^{[\mathbf{u},\mathbf{b}]}$, functions that can take values ± 1.

Now let $U = \mathrm{span}(\{\mathbf{a}, \mathbf{b}\})$ and $\mathbf{t} \in V$, and let us calculate the sufficient condition of Theorem 4.8. Because of the isotropicity of U, all the characters

in Equation 4.48 will be of the form $(-1)^{[\mathbf{u_1}+\mathbf{t},\mathbf{u_2}]} = (-1)^{[\mathbf{t},\mathbf{u_2}]}$. Therefore,

$$\sum_{\mathbf{v}\in U} W_\rho(\mathbf{v}+\mathbf{t}) = \left(1 + (-1)^{[\mathbf{t},\mathbf{a}]}a + (-1)^{[\mathbf{t},\mathbf{b}]}b - (-1)^{[\mathbf{t},\mathbf{a}]}(-1)^{[\mathbf{t},\mathbf{b}]}ab\right).$$

Depending on the value of $\mathbf{t}$, the characters in the previous equation can take any value in $\{\pm 1\}$. Comparing the two conditions, we see that they are the same, and we conclude that $\rho(a,b)$ is contextual if

$$1 + \alpha a + \beta b - \alpha\beta ab < 0 \tag{4.50}$$

for any possible value of $\alpha, \beta = \pm 1$.

It remains to determine what values of (a,b) describe physical states. Physical density matrices are positive semidefinite, meaning that

$$\langle\psi|\rho|\psi\rangle \geq 0 \quad \text{for all states } |\psi\rangle\,. \tag{4.51}$$

Due to the following easily verified relations,

$$\langle +,0|\rho(a,b)|+,0\rangle = \frac{1+a}{2} \qquad \langle +,1|\rho(a,b)|+,1\rangle = \frac{1-a}{2}$$
$$\langle 0,+|\rho(a,b)|0,+\rangle = \frac{1+b}{2} \qquad \langle 1,+|\rho(a,b)|1,+\rangle = \frac{1-b}{2},$$

we have that physical states obey $|a| \leq 1$, $|b| \leq 1$.

Figure 4.3 shows the phase diagram for the family of states $\rho(a,b)$ The physical states fill the unit square, while the non-contextual states are located in the dark blue area. The corners of the square represent the pure states that are joint eigenstates of X_1Z_2 and Z_1X_2, these states are located in the contextual phase at the points that are the furthest from the boundary with the non-contextual phase.

4.4 Contextuality and negativity as resources for quantum computation

Theorem 4.9. *In order to achieve computational universality in the scheme of QCSI with rebits, the resource states need to be contextual.*

Remark. Contextuality of the resource states implies that they have negative Wigner function, using Theorem 4.5.

Proof. First, we show that if our scheme is capable of universal quantum computation, it is able to create a state that maximally reveals contextuality, i.e has $\mathcal{W}_\rho^{\mathcal{B}(U)}(\mathbf{x}) = -2$ for some $\mathbf{x}$. In section 4.2.2 we have already demonstrated that the state $\tilde{\rho} = \frac{1}{4}(I - X_1Z_2)(I - Z_1X_2)$ maximally reveals contextuality. Therefore, any universal scheme of computation with rebits should be able to create the encoded version of the state $\tilde{\rho}$, for which

$$\mathcal{W}_{\tilde{\rho}}^{\mathcal{B}(U)}(\mathbf{0}) = -2. \tag{4.52}$$

The second part of the proof is to show that if a quantum circuit creates $\tilde{\rho}$, the state at every previous step of the computation also maximally reveals contextuality. Suppose that the state after the m-th step of a circuit, $\rho(m) = \tilde{\rho}$. We show that the state at the step just before $\rho(m-1)$ must also reveal contextuality maximally. There are two cases to consider:

1. Step m of the circuit is a unitary gate.

   Because the witness function obeys Equation 4.31, the covariance under CSS-ness preserving unitaries of the Wigner function carries on to $\mathcal{W}_\rho^{\mathcal{B}(U)}$

   $$\mathcal{W}_\rho^{\mathcal{B}(U)}(\eta) = \mathcal{W}_{g^{-1}\rho g}^{F^{-1}\mathcal{B}(U)}(\eta + \bar{\mathbf{a}}). \tag{4.53}$$

   This covariance guarantees that the state $\rho(m-1)$ also maximally witnesses contextuality.

2. Step m of the circuit is a projective measurement of an observable $T_\mathbf{c} \in \mathcal{O}$. There are two possible situations

   (a) $T_\mathbf{c}$ commutes with all elements of the stabilizer of $\tilde{\rho}$.

   Then the measurement does not change the state. Indeed, if $|\psi\rangle$ is stabilized by U, then

   $$U(1 \pm T_\mathbf{c})\,|\psi\rangle = (I \pm T_\mathbf{c})U\,|\psi\rangle = (I \pm T_\mathbf{c})\,|\psi\rangle\,, \tag{4.54}$$

so $|\psi\rangle$ still has the same stabilizer, and is therefore unchanged.

(b) $T_{\mathbf{c}}$ does not commute with all elements of the stabilizer of $\tilde{\rho}$.
Then $T_{\mathbf{c}}$ commutes with two of the non-trivial stabilizers of $\tilde{\rho}$, and anticommutes with the other. Let $T_{\mathbf{a}}$ and $T_{\mathbf{b}}$ be the two stabilizers that commute with $T_{\mathbf{c}}$, with $T_{\mathbf{a+b}}$ being the anticommuting one. Then

$$\langle T_{\mathbf{a}} \rangle_{\tilde{\rho}} = \langle T_{\mathbf{b}} \rangle_{\tilde{\rho}} = 0, \tag{4.55}$$

which reduces the expression for the witness function to

$$\mathcal{W}_{\tilde{\rho}}^{\mathbf{a},\mathbf{b}}(\eta) = \langle I \pm T_{\mathbf{a+b}} \rangle_{\tilde{\rho}} \geq 0. \tag{4.56}$$

But this contradicts our assumption 4.52. Therefore, case b) impossible.

We have thus established that any circuit that creates a state that maximally witnesses contextuality, must have an initial state that also maximally witnesses contextuality. □

4.5 Mermin square revisited

At first glance, the Mermin square proof of contextuality that was discussed in Section 1.4.3 only uses real observables, and should therefore be also applicable to rebit quantum mechanics. If state-independent contextuality is present, this would pose serious problems for the interpretation of our previous result that contextuality of the intial state is necessary for quantum computation. The square is reproduced again here in Figure 4.4 for the convenience of the reader.

The key feature of rebit QCSI that allows us to avoid the conclusions of the Mermin proof is that not all commuting observables can be simultaneously measured. Indeed, by the definition 4.1 of an allowable measurement,

$$\{X_1 Z_2, Z_1 X_2, -Y_1 Y_2\} \notin \mathcal{M}. \tag{4.57}$$

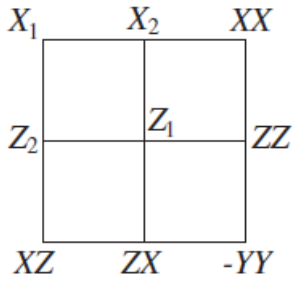


Figure 4.4: Reminder of the observables that appear in the "Mermin square" state-independent contextuality proof.

The lower row of the diagram is thus removed, and we can no longer produce the algebraic contradiction necessary for the original Mermin proof. We generalize these observations in the following lemma

Lemma 4.10. *Consider a system of n rebits with measurements restricted to the set of observables $\mathcal{O}$. Then there exists a consistent non-contextual value assignment $\lambda_{\mathbf{u}} : \mathcal{A} \to \pm 1$, $\forall \mathbf{u} \in \mathcal{S}$. This means that there is no state-independent contextuality in the scheme of rebit QCSI.*

Proof. We prove this by collecting the results of previous lemmas about the possible value assignments and the commutation properties of observables. Lemma 4.4 requires that the value assignments $\lambda_{\mathbf{u}} : \mathcal{A} \to \{\pm 1\}, \mathbf{u} \in V$ take the form

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u},\mathbf{a}]}, \tag{4.58}$$

and satisfy the consistency condition

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}+\mathbf{b}}) = \lambda_{\mathbf{u}}(T_{\mathbf{a}})\lambda_{\mathbf{u}}(T_{\mathbf{b}}) \quad \forall T_{\mathbf{a}}, T_{\mathbf{b}} \in \mathcal{A} \text{ such that } [T_{\mathbf{a}}, T_{\mathbf{b}}] = 0. \tag{4.59}$$

Furthermore, for all allowable measurements $M \in \mathcal{M}$, Lemma 4.2 guarantees that for all $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$,

$$T_{\mathbf{a}+\mathbf{b}} = T_{\mathbf{a}} T_{\mathbf{b}}. \tag{4.60}$$

So we see that all $\lambda_{\mathbf{u}}$ are valid value assignments for a NCHVM that describes

rebit QCSI with measurable observables restricted to $\mathcal{O}$. $\square$

We conclude this section with some general comments about Mermin-type proofs of contextuality, also known as parity proofs. The key ingredient for these proofs is that there be a measurement $M \in \mathcal{M}$ such that for $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$,

$$T_{\mathbf{a}} T_{\mathbf{b}} = -T_{\mathbf{a}+\mathbf{b}}. \tag{4.61}$$

Then, using $T_{\mathbf{a}}, T_{\mathbf{b}}, T_{\mathbf{a}+\mathbf{b}}$, and local Pauli observables X_i, Z_i, it is possible to formulate a state-independent proof of contextuality.

According to the proof of Lemma 4.6, a witness function $\mathcal{W}^{\mathcal{B}(U)}$ will only be able to take negative values if the isotropic subspace $U \subset V$ contains $T_{\mathbf{a}}, T_{\mathbf{b}}$ such that $T_{\mathbf{a}} T_{\mathbf{b}} = -T_{\mathbf{a}+\mathbf{b}}$. In this way, we see that the existence of a contextuality witness is very closely linked to the existence of a state-independent parity proof. What our operational restriction of rebit QCSI does is to prevent any measurement that has $T_{\mathbf{a}} T_{\mathbf{b}} = -T_{\mathbf{a}+\mathbf{b}}$. Indeed, if we allowed all real Clifford operations in our scheme, then the Mermin square proof would apply, and it would be impossible to assert that contextuality is a resource for quantum computation.

Chapter 5

Conclusion

Let us summarize what have been the results of this work. We have seen that two-dimensional systems are different from odd-prime dimensional in ways that manifest themselves in the properties of the Wigner function. In order to have a useful covariant Wigner function for two level systems, we had to restrict ourselves to systems of rebits, and we described a universal scheme of quantum computation for them. We have been able to define a rebit Wigner that is covariant under CSS-ness preserving unitaries, and for which CSS states are positively represented. This allowed us to devise an efficient simulation method for positive states undergoing CSS-Clifford circuits, and therefore establish negativity of the Wigner function as a resource for quantum computation.

We have also studied contextuality, and found necessary and sufficient conditions that were based on the Wigner function. We found that the limited gate set in our QCSI scheme prevented state-independent contextuality to occur. We could thus establish contextuality of the magic state as a resource for universal quantum computation. The previous observations show that we have managed to recover virtually all the properties enjoyed by the qudit case, with the exception that contextuality and negativity are not equivalent: some states are negative but non-contextual.

This research bridges some of the gap of understanding that was present on the literature, as most of the preexisting work on negativity and contex-

tuality had been made on qudits, while quantum computation is typically thought in terms of qubits. A limitation of this work is that we use rebits, which are relatively esoteric, and on which there has been virtually no experimental work. We also have not addressed very deeply the reasons for the differences between qubits, qudits, and rebits, but there are already efforts (not involving the author of this thesis) to use the machinery of group cohomology to better understand the particularities of each case, and to extend the results of this thesis from rebits to qubits. Another possible future research direction would be to use Wigner functions to study measurement based quantum computation (MBQC) [21]. It has been established [1] that contextuality in MBQC is necessary for *classical* universal computation. It would be interesting to use some of the insights from our research to understand the reason of this apparent difference.

Bibliography

[1] J. Anders and D. E. Browne. Computational power of correlations. *Phys. Rev. Lett.*, 102:050502, Feb 2009. doi:10.1103/PhysRevLett.102.050502. URL http://link.aps.org/doi/10.1103/PhysRevLett.102.050502. → pages 91

[2] K. Banaszek and K. Wódkiewicz. Nonlocality of the einstein-podolsky-rosen state in the wigner representation. *Phys. Rev. A*, 58:4345–4347, Dec 1998. doi:10.1103/PhysRevA.58.4345. URL http://link.aps.org/doi/10.1103/PhysRevA.58.4345. → pages 14

[3] S. D. Bartlett, B. C. Sanders, S. L. Braunstein, and K. Nemoto. Efficient classical simulation of continuous variable quantum information processes. *Phys. Rev. Lett.*, 88:097904, Feb 2002. doi:10.1103/PhysRevLett.88.097904. URL http://link.aps.org/doi/10.1103/PhysRevLett.88.097904. → pages 14

[4] J. S. Bell. On the Einstein-Podolsky-Rosen Paradox. *Physics*, 1(3): 195–200, 1964. ISSN 00653276. → pages 1, 26, 29, 32

[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.*, 76:722–725, Jan 1996. doi:10.1103/PhysRevLett.76.722. URL http://link.aps.org/doi/10.1103/PhysRevLett.76.722. → pages 42

[6] D. Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. ii. *Phys. Rev.*, 85:180–193, Jan 1952. doi:10.1103/PhysRev.85.180. URL http://link.aps.org/doi/10.1103/PhysRev.85.180. → pages 31

[7] S. Bravyi and A. Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, 2005. → pages 2, 25, 36, 47

[8] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.*, 78:405–408, Jan 1997. doi:10.1103/PhysRevLett.78.405. URL http://link.aps.org/doi/10.1103/PhysRevLett.78.405. → pages 5

[9] W. B. Case. Wigner functions and weyl transforms for pedestrians. *American Journal of Physics*, 76:937–946, 2008. doi:10.1119/1.2957889. → pages 11

[10] N. Delfosse, P. A. Guerin, J. Bian, and R. Raussendorf. Wigner Function Negativity and Contextuality in Quantum Computation on Rebits. *Physical Review X*, 021003(5):1–23, 2015. doi:10.1103/PhysRevX.5.021003. → pages 6

[11] A. Einstein, B. Podolsky, and N. Rosen. Phys. Rev. 47, 777 (1935): Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 47:777–780, 1935. ISSN 0031-899X. → pages 14, 26

[12] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters. Discrete phase space based on finite fields. *Physical Review A*, 70(6):60, 2004. ISSN 1050-2947. → pages 18, 21

[13] D. Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, 1997. URL http://arxiv.org/abs/quant-ph/9705052. → pages 3, 4, 5

[14] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47:1–17, 2006. ISSN 00222488. → pages 15, 17, 18

[15] L. K. Grover. A fast quantum mechanical algorithm for database search. *28th Annual ACM Symposium on the Theory of Computing*, 1996. → pages 1

[16] M. Howard, J. Wallman, V. Veitch, and J. Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510:351–355, 2014. ISSN 0028-0836. → pages 2, 32

[17] R. Hudson. When is the wigner quasi-probability density non-negative? *Reports on Mathematical Physics*, 6(2):249 – 252, 1974. ISSN 0034-4877. doi:http://dx.doi.org/10.1016/0034-4877(74)90007-X. URL http://www.sciencedirect.com/science/article/pii/003448777490007X. → pages 14

[18] S. Kochen and E. P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17: 59–87, 1967. → pages 1, 29

[19] N. D. Mermin. Hidden variables and the two theorems of John Bell. *Reviews of Modern Physics*, 65:803–815, 1993. ISSN 00346861. → pages 29

[20] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011. ISBN 1107002176, 9781107002173. → pages 1, 2, 3, 24, 32, 35

[21] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001. doi:10.1103/PhysRevLett.86.5188. URL http://link.aps.org/doi/10.1103/PhysRevLett.86.5188. → pages 2, 91

[22] T. Rudolph and L. Grover. A 2 rebit gate universal for quantum computing. *arXiv*, page 2, 2002. URL http://arxiv.org/abs/quant-ph/0210187. → pages 35

[23] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J.Sci.Statist.Comput.*, 26, 1997. → pages 1

[24] F. Soto and P. Claverie. *J. Math. Phys.*, 24, 1983. → pages 14

[25] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A - Atomic, Molecular, and Optical Physics*, 71(5), 2005. ISSN 10502947. → pages 27

[26] R. W. Spekkens. Negativity and contextuality are equivalent notions of nonclassicality. *Phys. Rev. Lett.*, 101:020401, Jul 2008. doi:10.1103/PhysRevLett.101.020401. URL http://link.aps.org/doi/10.1103/PhysRevLett.101.020401. → pages 32

[27] V. Veitch, C. Ferrie, D. Gross, and J. Emerson. Negative quasi-probability as a resource for quantum computation. *New Journal of Physics*, 14:1–15, 2012. ISSN 13672630. → pages 2, 19, 32, 66

[28] H. Weyl. Quantenmechanik und gruppentheorie. *Zeitschrift für Physik*, 46(1-2):1–46, 1927. ISSN 0044-3328. doi:10.1007/BF02055756. URL http://dx.doi.org/10.1007/BF02055756. → pages 8

[29] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–759, Jun 1932. doi:10.1103/PhysRev.40.749. URL http://link.aps.org/doi/10.1103/PhysRev.40.749. → pages 2, 8

[30] W.-M. Zhang, D. H. Feng, and R. Gilmore. Coherent states: Theory and some applications. *Rev. Mod. Phys.*, 62:867–927, Oct 1990. doi:10.1103/RevModPhys.62.867. URL http://link.aps.org/doi/10.1103/RevModPhys.62.867. → pages 12

[31] H. Zhu. Permutation symmetry determines the discrete Wigner function. *arXiv*, page 6, 2015. URL http://arxiv.org/abs/1504.03773. → pages 22

Appendix A

Characters and Fourier transforms

This appendix some of the mathematical concepts necessary to make sense of some sections of this thesis. References have been omitted, but it is all standard material. We have tried at all times to avoid excessive generality so as to be readily applicable to the body of this thesis.

A.1 Characters

Given a finite abelian group G with the group operation written in additive notation, a character of G is a function

$$\chi : G \to \mathbb{C} \tag{A.1}$$

such that $\chi(g + h) = \chi(g)\chi(h)$. In other words, it is a one dimensional representation of G. The set of characters of G is denoted G^*, and it is a fact about finite groups that they are isomorphic

$$G \cong G^*. \tag{A.2}$$

In the context of phase-space distributions, $G = \mathbb{Z}_d^n$, and the characters are of the form $\chi : x \mapsto e^{i2\pi xy/d}$ for some $y \in \mathbb{Z}_d^n$.

Lemma A.1. *Given a group G and any character χ of G,*

$$\sum_g \chi(g) = 0 \tag{A.3}$$

for any character other that $\chi_0 : g \mapsto 1$, in which case $\sum_g \chi_0(g) = |G|$

Proof. If $\chi \neq \chi_0$, then there exists an $h \in G$ such that $\chi(h) \neq 1$. Then

$$\chi(h) \sum_g \chi(g) = \sum_g \chi(h)\chi(g) = \sum_g \chi(h+g) = \sum_g \chi(g). \tag{A.4}$$

Comparing the right-most side of the equation with the left-most side, we conclude that $\chi(h) = 1$, a contradiction. □

Given two character $\chi, \eta \in G^*$, we can use this property to define an inner product on the space of characters of G

$$(\chi, \eta) = \frac{1}{|G|} \sum_{g \in G} \chi(g)\eta(-g). \tag{A.5}$$

This inner product makes the characters an orthonormal basis for the set of functions $G \to C_n$, where $C_n = \{e^{i2\pi x/n} | x \in \{0, 1, \ldots, n-1\}\}$.

A.2 Symplectic spaces, definitions

For the following section, let $V = \mathbb{Z}_d^{2n}$. We will use the notation $Sp_N(F)$ for the space of symplectic matrices of dimension $N \times N$ with entries in the field F. Symplectic matrices preserve the symplectic form over the vector space that they act on:

$$[F\mathbf{u}, F\mathbf{v}] = [\mathbf{u}, \mathbf{v}]. \tag{A.6}$$

Definition A.2. (Nondegenerate bilinear form) A bilinear form $[\cdot, \cdot]$ over V is non degenerate if $[\mathbf{u}, \mathbf{v}] = 0 \quad \forall \mathbf{v}$ implies that $\mathbf{u} = 0$

The usual symplectic product over V defined as $[\mathbf{u}, \mathbf{v}] = \mathbf{u}_z \cdot \mathbf{v}_x - \mathbf{u}_x \cdot \mathbf{v}_z$ is a nondegenerate bilinear form.

Definition A.3. (Isotropic subspace) A subspace U of V is isotropic if $\forall \mathbf{u}, \mathbf{v} \in U, [\mathbf{u}, \mathbf{v}] = 0$

Definition A.4. (Maximally isotropic subspace) An isotropic subspace U of V is maximally isotropic if it has the maximal dimension for which a subspace of V can be isotropic. This happens if and only if $\dim(U) = n$

Lemma A.5. *Any character χ of a non-degenerate symplectic space V can be written as $\chi(\mathbf{u}) = e^{2\pi i[\mathbf{u},\mathbf{w}]/d}$ for some $\mathbf{w} \in V$*

Proof. Using the fact that $|V^*| = |V| = 2n$, we only need to prove that the set of characters $\{\chi_{\mathbf{w}} = e^{2\pi i[\cdot,\mathbf{w}]/d} | \mathbf{w} \in V\}$ contains 2n different characters.

Suppose $\exists \mathbf{w}, \mathbf{w}'$ such that $\chi_{\mathbf{w}} = \chi_{\mathbf{w}'}$. Then we have

$$e^{2\pi i[\mathbf{u},\mathbf{w}]/d} = e^{2\pi i[\mathbf{u},\mathbf{w}']/d} \implies [\mathbf{u}, \mathbf{w} - \mathbf{w}'] = 0. \tag{A.7}$$

But by the nondegeneracy of the symplectic product, this implies that $\mathbf{w} = \mathbf{w}'$. □