Wigner function negativity and contextuality in quantum computation on rebits

Nicolas Delfosse[1], Philippe Allard Guerin[2], Jacob Bian[2], Robert Raussendorf[2]
1: Département de Physique, Université de Sherbrooke, Sherbrooke, Québec, J1K 2R1, Canada,
2: Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada
(Dated: September 19, 2014)

We describe a universal scheme of quantum computation by state injection on rebits (states with real density matrices). For this scheme, we establish contextuality and Wigner function negativity as computational resources, extending results of [M. Howard *et al.*, Nature 510, 351–355 (2014)] to two-level systems. For this purpose, we define a Wigner function suited to systems of n rebits, and prove a corresponding discrete Hudson's theorem. We introduce contextuality witnesses for rebit states, and discuss the compatibility of our result with state-independent contextuality.

PACS numbers: 03.67.Mn, 03.65.Ud, 03.67.Ac

I. INTRODUCTION

In quantum computation by state injection (QCSI) [1], the set of quantum gates is by construction not universal. This restriction is made up for by the injection of states that could not be created within the scheme itself, the so-called magic states.

Besides its promise for the realization of fault-tolerant quantum computation, QCSI is of fundamental theoretical interest. Since the magic states enable universality, one is led to ask: *Precisely which quantum properties of these states are responsible for the gain in computational power?*

Contextuality [2], [3] and negativity of Wigner functions have recently been proposed as the quintessential quantum properties of magic states; See [4],[5],[6]. Contextuality is an obstruction to modelling the inherent randomness of quantum measurement in a statistical mechanics fashion, namely by a probability distribution over configurations with predetermined measurement outcomes for all measurable observables. Wigner functions [7], [8] are the closest quantum analogue of probability distributions over phase space. The key difference is that Wigner functions can assume negative values, and this negativity is taken as an indication of quantumness. Despite their separate origins in the fields of quantum optics and foundations of quantum mechanics, Wigner function negativity and contextuality are closely related indicators of non-classical behaviour [9], [4].

The reason for the appearance of Wigner functions in the discussion of QCSI is their relation [5], [10] to the stabilizer formalism [11]. The stabilizer formalism is also relevant for QCSI, since the restricted gate set therein is typically chosen to be the Clifford gates. These gates are indeed not universal, and—if supplemented only with Pauli measurements and stabilizer states—can be efficiently classically simulated by stabilizer techniques.

An epitome for the link between Wigner functions and QCSI via the stabilizer formalism is the discrete Hudson's theorem [10], which says that in Hilbert spaces of odd prime-power (hence finite) dimension, the pure states with positive Wigner function are exactly the stabilizer states. Thus, stabilizer states are "classical" from both the perspectives of Wigner functions and QCSI. In the wake of this result, contextuality and Wigner function negativity have been established as quantum resources for QCSI with qudits of odd prime dimension [6], [4].

Extending these properties to 2-level systems is pertinent, since quantum algorithms are typically formulated in terms of qubits. But attempts to do so hit barriers: As for the Wigner functions, many constructions cannot be adapted to qubits [10], [12]; and for the remaining ones, [5], [13], the discrete Hudson's theorem breaks down. There are qubit stabilizer states with negative Wigner function. As for contextuality, it now arises in its state-independent form [14]. In result, every quantum state of more than one qubit can be considered contextual [4], which is at odds with viewing contextuality as a resource possessed only by special states.

Here, we establish Wigner function negativity and contextuality as necessary resources for QCSI on two-level systems. We achieve this at the price of restricting from qubits to rebits, i.e., real density matrices of n two-level systems. This restriction does not affect universality [15]. The role that was previously played by the stabilizer states is now played by the CSS-states [16], and the group of Clifford gates is replaced by the subgroup of CSS-ness preserving Clifford gates. Within this new setting, we resurrect a discrete Hudson's theorem, as well as a number of related properties of the Wigner function. The final picture for rebits resembles that for qudits, but with one difference: contextuality and Wigner function negativity are no longer synonymous. There exist non-contextual states with negative Wigner function.

This paper is organized as follows. Section II summarizes the known results on the roles of contextuality and negativity in qudit QCSI, and defines our setting for rebits. In Section III we present a universal scheme of quantum computation by state injection on rebits. In Section IV, we construct a matching Wigner function, equipped with a discrete Hudson's theorem and extended Gottesman-Knill theorem. In Section V we provide necessary and sufficient conditions for contextuality in terms of the Wigner function. Section VI contains our results on contextuality and negativity as resources in rebit-QCSI. We conclude in Section VII.

II. QUANTUM COMPUTATION BY STATE INJECTION

QCSI has four operational quantum components: the restricted unitary gates, the restricted measurements, the cheap states and the magic states. The cheap states are those that can be produced from sequences of measurements from the restricted set and restricted unitary gates, possibly classically conditioned on measurement outcomes. The classical side-processing is unrestricted.

A typical choice for the restricted operations is that they live within the stabilizer world. That is, the restricted set of unitary gates is in the group of Clifford gates, the restricted set of observables is the Pauli observables or a subset thereof, and the cheap states are stabilizer states.

A. Summary of the qudit case

For the case of odd prime local dimension d, QCSI has been investigated for the restricted gate set being the Clifford gates [6],[4]. For this scenario, two essential quantum properties of the magic states have been identified, namely the negativity of their Wigner function, and their contextuality with respect to stabilizer measurements. Specifically, it has been established that

(i) Negativity in the Wigner function of raw magic states is necessary for successful magic state distillation (Theorem 3 in [6]) and for the hardness of classical simulation of QCSI (Theorem 1 in [6]).

(ii) Contextuality of magic states w.r.t stabilizer measurements is necessary for universality of QCSI [4].

The Wigner function plays a dual role for QCSI. It is relevant for the phenomenology observed (see above), but it is also deeply involved in the mathematical description of the computational scheme. This is revealed in the following five properties, which hold for odd prime d when the restricted operations belong to the stabilizer world,

(iii) The set of stabilizer states is singled out by a Hudson's theorem as the set of pure states with non-negative Wigner function [10].

(iv) The set of Clifford gates is singled out as the set of unitaries that transform the Wigner function covariantly [10].

(v) Clifford gates and stabilizer measurements preserve positivity of the Wigner function [6].

(vi) Necessary and sufficient conditions for contextuality w.r.t to the restricted set of measurements can be expressed in terms of the Wigner function [4].

(vii) For one-qudit states, negativity of the Wigner function and contextuality w.r.t. measurements from the restricted set are the same [4].

The above physical properties (i) and (ii) are consequences of the structural properties (iii) - (vii). For example, an efficient classical simulation method for the evolution of states with non-negative Wigner function under the restricted gates can be built on properties (iv) and (v) [6]. Its existence directly implies (i). Furthermore, Hudson's theorem (iii) connects this simulation method with the Gottesman-Knill theorem.

B. The trouble with qubits

For systems of qubits, both the employed contextuality witnesses [17] and Wigner functions run into difficulty. As for contextuality, if the goal is to establish it as a quantum resource, one has to overcome a problem posed by the phenomenon of state-independent contextuality which is revealed, for example, by the Mermin square and star [14]. Mermin's square can be translated into a contextuality witness for which *all* quantum states of $n \geq 2$ qubits come out contextual [4]. If contextuality is generic then it cannot be a resource.

As for the Wigner functions, many of the Wigner functions proposed for Hilbert spaces of finite dimension d^n require for their definition the existence of 2^{-1} in $\mathbb{F}_d$, and thus do not apply to the qubit case $d = 2$; for examples see e.g. [10], [12].

Yet some Wigner functions do survive the transition to $d = 2$; see e.g. [5], [13]. However, in these cases, the general connection with the stabilizer world breaks down. Not all stabilizer states have non-negative Wigner function anymore, and the Wigner function no longer transforms covariantly under all Clifford operations.

For the construction [5], [13], Wigner function negativity of the magic states is necessary for universality. Therein, not a single Wigner function is considered but instead the whole class introduced in [8]. A "classical" state must be positively represented for each of these Wigner functions. The number of pure n-qubit states for which this holds is super-exponentially small compared to the number of n-qubit stabilizer states [10].

C. Rebits

In this paper, we discuss the case of local dimension $d = 2$. We present a universal scheme of QCSI for which contextuality and Wigner function negativity are established as necessary quantum resources. The price we pay is that we have to restrict from qubits to rebits. Specifically, we require that the density matrix of the processed quantum state ρ is real; i.e. at each point in the quantum computation it holds that

$$\langle x|\rho|y\rangle \in \mathbb{R},$$

for all $|x\rangle$, $|y\rangle$ in the computational basis.

For the discussed scheme of rebit QCSI, the set of cheap states is the CSS states, the set of restricted gates

is the CSS-ness preserving Clifford gates, and the allowed measurements are of observables from the set

$$\mathcal{O} = \{X(\mathbf{a}_X), Z(\mathbf{a}_Z) | \mathbf{a}_X, \mathbf{a}_Z \in \mathbb{Z}_2^n\}. \quad (1)$$

That is, in our construction the restricted operations belong to the CSS-stabilizer world, rather than the more general stabilizer world.

D. CSS states and CSS-ness preserving Clifford operations

Calderbank-Shor-Steane (CSS) states are a subset of the stabilizer states. They are defined by the property that for any CSS state $|\psi\rangle$, the corresponding Pauli stabilizer group $\mathcal{S}(|\psi\rangle)$ decomposes into an X- and a Z-part; i.e., $\mathcal{S}(|\psi\rangle) = \mathcal{S}_X(|\psi\rangle) \times \mathcal{S}_Z(|\psi\rangle)$, where all elements of $\mathcal{S}_X(|\psi\rangle)$ and $\mathcal{S}_Z(|\psi\rangle)$ are of the form $X(\mathbf{a}_X)$ and $Z(\mathbf{a}_Z)$, respectively. All CSS states are real, but not all real stabilizer states are of CSS type.

We now characterize the CSS-ness preserving transformations. Denote by Ω the set of pure CSS-states and by G_{CSS} the subgroup of the n-qubit orthogonal group $O_{2^n}(\mathbb{R})$ which preserves the set Ω of CSS states,

$$G_{CSS} = \{g \in C_n |\, g|\Psi\rangle \in \Omega, \forall \Psi \in \Omega\}. \quad (2)$$

By definition, G_{CSS} is a subgroup of the real n-qubit Clifford group C_n. The following can be said about the structure of G_{CSS}.

Lemma 1 *The n-rebit CSS-ness preserving group G_{CSS} is*

$$G_{CSS} = \left\langle \bigotimes_{i=1}^{n} H_i, CNOT(i,j), X_i, Z_i \right\rangle, \quad (3)$$

where $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$. We have the group isomorphism

$$G_{CSS}/\{\pm I\} = \mathbb{Z}_2^{2n} \rtimes (\mathrm{GL}_n(\mathbb{Z}_2) \rtimes \mathbb{Z}_2). \quad (4)$$

In Eq.(4), the component $\mathbb{Z}_2^{2n}$ corresponds to the Pauli operators $T_{\mathbf{u}}$, the component $\mathrm{GL}_n(\mathbb{Z}_2)$ corresponds to the group generated by the $CNOT$, and the subgroup $\mathbb{Z}_2$ is generated by the simultaneous Hadamard gate $\otimes_i H_i$.

Since the set $\mathcal{O} = \{Z(\mathbf{u}) \mid \mathbf{u} \in \mathbb{Z}_2^n\} \cup \{X(\mathbf{v}) \mid \mathbf{v} \in \mathbb{Z}_2^n\}$ is mapped onto itself by conjugation under gates from the group on the r.h.s. of Eq. (3), it is clear that this group is a subgroup of G_{CSS} as defined in Eq. (2). That it is indeed all of G_{CSS} is proved in Appendix C.

We note that the set Ω of "cheap" CSS states, the CSS-ness preserving unitary gates G_{CSS} and the projective measurements of observables in $\mathcal{O}$ form a compatible classical reference structure for QCSI, in the sense that none of these operations can map states inside Ω to states outside Ω.

III. UNIVERSAL QUANTUM COMPUTATION BY STATE INJECTION ON REBITS

It has been shown in [15] that rebits are sufficient for universal quantum computation. In that scheme, first, a quantum state of n qubits,

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

is encoded into a state of $n+1$ rebits,

$$\overline{|\psi\rangle} = \sum_{\mathbf{v} \in \mathbb{Z}_2^n} \left(r_{\mathbf{v}} \cos\theta_{\mathbf{v}} |\mathbf{v}\rangle \otimes |R\rangle + r_{\mathbf{v}} \sin\theta_{\mathbf{v}} |\mathbf{v}\rangle \otimes |I\rangle \right). \quad (5)$$

The additional rebit, with basis states $|R\rangle = |0\rangle$ and $|I\rangle = |1\rangle$, allows to keep track of the real and imaginary parts of the unencoded n-qubit state. Second, an encoded set of gates is constructed which (i) is universal, and (ii) preserves real-ness of the states in Eq. (5).

Using the encoding Eq. (5), we construct a universal scheme of QCSI on rebits. The restricted gate set used therein consists of CNOT-gates, the simultaneous Hadamard-gate H_{all} on all rebits, and Pauli-flips, i.e.,

$$\langle \mathcal{G}_{\mathrm{restricted}} \rangle = G_{CSS}.$$

These unitary gates are supplemented by measurements of observables in the set $\mathcal{O}$, or, w.l.o.g., of observables $\{Z_i |\, i = 1, .., n\}$.

The (unitary) Pauli operators and the simultaneous Hadamard-gate can be dispensed with, because they can be propagated past the readout measurements. This is a consequence of the well-known propagation relations for Pauli operators under conjugation by Clifford gates, and $CNOT(i,j)H_{\mathrm{all}} = H_{\mathrm{all}}CNOT(j,i)$. If those gates are eliminated, we remain with the CNOT-gates and measurements of X_i and Z_i. We note that this is precisely the set of gates which can be performed fault-tolerantly on the surface code [18] using defect braiding [19]. However, for the present purpose, we keep the redundant H_{all} and Pauli flips in the restricted gate set.

For the universal gate set, we pick

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

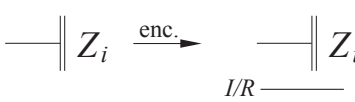
supplemented with measurements of the Pauli observables Z_i, for $i = 1, .., n$.

We now demonstrate that the encoded versions of these gates can be realized only using the gates from the restricted set and the injection of two types of ancilla states, $|A\rangle$ and $|B\rangle$, defined as

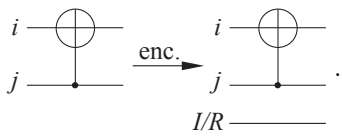
$$\begin{aligned} |A\rangle &= \frac{|0\rangle|R\rangle + \cos\frac{\pi}{4}|1\rangle|R\rangle + \sin\frac{\pi}{4}|1\rangle|I\rangle}{\sqrt{2}}, \\ |B\rangle &= \frac{|0\rangle|+\rangle + |1\rangle|-\rangle}{\sqrt{2}}. \end{aligned} \quad (6)$$

The ancilla $|A\rangle$ is the encoded $(|0\rangle + e^{i\pi/4}|1\rangle)/\sqrt{2}$, with respect to the encoding of Eq. (5).

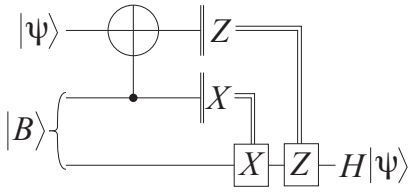
(a) *The measurement of Z_i.* Since the Pauli-operator Z is real, its measurement does not differentiate between the real and imaginary parts of the measured state, and $\overline{Z}_i \equiv Z_i$. Graphically,



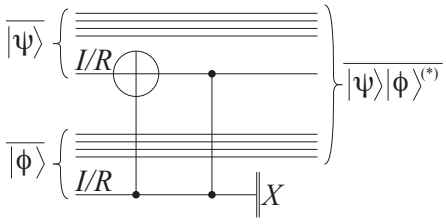
(b) *The CNOT-gate between qubits i and j.* The CNOT-gate is real and hence does not mix the real and imaginary parts of the state it is applied to. Hence, $\overline{CNOT(i,j)} = CNOT(i,j)$. Graphically,



(c) *The Hadamard gate H_i.* The encoded Hadamard gate is realized by injection of an ancilla $|B\rangle$ into the circuit



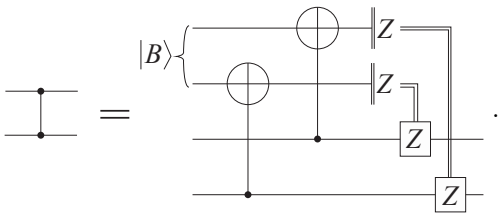
(d) *The gate* $\exp(i\pi/8\, Z_i)$. The encoded version of this gate uses an ancilla states $|A\rangle$ and $|B\rangle$, and proceeds in two steps. The first step is a pre-processing jointly for all the $\exp(i\pi/8\, Z_i)$ gates in the circuit. Namely, at the beginning of the computation, each ancilla state $|A\rangle$ is in its own separate code block. In the pre-processing step, all data and ancila rebits are merged into the same code block. The merging can be done two blocks at a time, and the corresponding circuit is



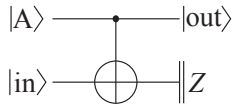
For a pair of encoded input states $\overline{|\psi\rangle}$, $\overline{|\phi\rangle}$, the result of the code merging circuit is $\overline{|\psi\rangle \otimes |\phi\rangle}$ or $\overline{|\psi\rangle \otimes |\phi\rangle^*}$, depending on the outcome of the X-measurement ($|\phi\rangle^*$ denotes the state obtained from $|\psi\rangle$ by complex conjugation w.r.t. the computational basis).

We will only ever use the code merging circuit for encoding the ancilla $|A\rangle = \overline{|\pi/8\rangle}$ into a single code block. Since $|\pi/8\rangle$ and $|-\pi/8\rangle$ allow to perform the $\pi/8$-phase gate with the same efficiency, the probabilistic nature of the code merging circuit does not affect the computation.

The code merging circuit contains a conditional phase gate which is not part of the restricted gate set. It is realized via the following state-injection circuit,



The second step then is the encoded version of the standard state injection circuit for the $\pi/8$-gate [20],



(7)

This circuit consists solely of operations whose encoded versions we have already demonstrated.

IV. A WIGNER FUNCTION FOR REBITS

In the last section we described a universal scheme of quantum computation by state injection on rebits, and here we construct the matching Wigner function. We first propose the rebit Wigner function and examine its basic properties. Second, we prove a discrete Hudson's theorem for rebits. Third, we prove covariance of the rebit Wigner function under CSS-ness preserving Clifford unitaries; and finally show that the evolution of states with positive Wigner function under CSS-ness preserving Clifford unitaries and measurements can be efficiently classically simulated.

A. Definition of a Wigner function for rebits

We now proceed to construct the Wigner W function for n-rebit states, which is suited to describe the computational scheme introduced in the previous section. It is a modification of the Wigner function $\tilde{W}$ [6], [4] for qubits.

In the qubit case, there are 4^n Pauli operators T_a,

$$T_{(\mathbf{a}_Z,\mathbf{a}_X)} = Z(\mathbf{a}_Z)X(\mathbf{a}_X), \quad \text{where } \mathbf{a}_Z, \mathbf{a}_X \in \mathbb{Z}_2^n. \tag{8}$$

Therein, $Z(\mathbf{a}) = Z_1^{a_1} \otimes Z_2^{a_2} \otimes .. \otimes Z_n^{a_n}$, for all $(\mathbf{a}_X, \mathbf{a}_Z) \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n$. We denote

$$\begin{aligned} V &:= \{(\mathbf{a}_X, \mathbf{a}_Z) |\, \mathbf{a}_X, \mathbf{a}_Z \in \mathbb{Z}_2^n\} \cong \mathbb{Z}_2^{2n}, \\ \mathcal{T} &:= \{T_\mathbf{a} |\, \mathbf{a} \in V\}. \end{aligned}$$

The Pauli operators $\mathcal{T}$ form an orthonormal basis of the vector space of square matrices of size 2^n with complex coefficients endowed with the inner product defined by $(A, B) = \frac{1}{2^n} \mathrm{Tr}(A^\dagger B)$,

$$\mathrm{Tr}(T_\mathbf{a}^\dagger T_\mathbf{b}) = 2^n \delta_{\mathbf{a},\mathbf{b}}, \quad \forall\, T_\mathbf{a}, T_\mathbf{b} \in \mathcal{T}. \tag{9}$$

In the present work, we are interested in rebits, which are defined by symmetric real density operators. We consider the set

$$\mathcal{A} := \{T_{\mathbf{a}} \mid (\mathbf{a}_Z, \mathbf{a}_X) = 0 \mod 2\}, \tag{10}$$

which is an orthonormal basis of the space of symmetric matrices (see Lemma 19 in Appendix B), and define

$$W_\rho(\mathbf{u}) := \frac{1}{2^n}\mathrm{Tr}(A_{\mathbf{u}}\rho), \tag{11}$$

with

$$A_{\mathbf{0}} = \frac{1}{2^n}\sum_{T_{\mathbf{a}}\in\mathcal{A}} T_{\mathbf{a}}, \text{ and } A_{\mathbf{u}} = T_{\mathbf{u}} A_{\mathbf{0}} T_{\mathbf{u}}^\dagger. \tag{12}$$

For later use, denote $V_{\mathcal{A}} = \{\mathbf{a} \in V \,|\, (\mathbf{a}_z, \mathbf{a}_x) \bmod 2 = 0\}$. Note that the operator $A_{\mathbf{u}}$ can also be written

$$A_{\mathbf{u}} = \frac{1}{2^n}\sum_{T_{\mathbf{a}}\in\mathcal{A}} (-1)^{[\mathbf{u},\mathbf{a}]} T_{\mathbf{a}} \tag{13}$$

where $[\mathbf{u}, \mathbf{v}] = (\mathbf{u}_Z, \mathbf{v}_X) + (\mathbf{v}_Z, \mathbf{u}_X)$ is the symplectic inner product in $\mathbb{Z}_2^{2n}$.

When considering real states, the family $(A_{\mathbf{u}})_{\mathbf{u}\in V}$ is not a basis of the space of symmetric matrices since it contains too many matrices. Nevertheless, in close analogy with the qudit Wigner function [10], [6], the rebit Wigner function of Eq. (11) has the following properties (compare with [21]):

1. Any real density matrix ρ satisfies

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

   W is thus informationally complete.

2. W transforms covariantly under the group of CSS-ness preserving Clifford transformations.

3. The CSS-states are the only pure states with non-negative W (discrete Hudson's theorem).

4. For all real density matrices ρ, σ,

$$W_{\rho\otimes\sigma} = W_\rho \cdot W_\sigma.$$

5. The trace inner product is given as

$$\mathrm{Tr}(\rho\sigma) = 2^n \sum_{\mathbf{u}\in V} W_\rho(\mathbf{u}) W_\sigma(\mathbf{u}). \tag{14}$$

6. The phase point operators satisfy $\mathrm{Tr} A_{\mathbf{u}} = 1$. Thus, $\mathrm{Tr} B = \sum_{\mathbf{u}} W_B(\mathbf{u})$ for any symmetric operator B.

Property 1 is proven in Lemma 20 in Appendix B, Property 2 in Section IV B, and Property 3 in Section IV C. Property 4 is shown in Appendix B. Properties 5 and 6 are immediate consequences of Property 1.

B. A discrete Hudson's theorem for rebits

The original Hudson's theorem in infinite-dimensional Hilbert space [22] singles out the Gaussian states as the pure states with positive Wigner function. This result has a counterpart in finite, odd prime-power dimension. Namely, the pure states with positive Wigner function are the stabilizer states [10]. In this way, a connection between Wigner functions and the discrete world of the stabilizer formalism is established. For no known Wigner function defined on qubits, this result carries over.

Here, for the Wigner function defined in the previous section, we find that for multiple rebits a discrete Hudson's theorem holds with the stabilizer states replaced by the more special CSS states.

Theorem 1 *A pure real state $|\psi\rangle$ has non-negative Wigner function W_ψ if and only if it is a CSS state.*

Recall that a Wigner function W_ρ for some density operator ρ is said to be non-negative if $W_\rho(\mathbf{u}) \geq 0$ for all $\mathbf{u} \in V$, and is said to be negative otherwise.

In order to prove this result we follow the strategy proposed by Gross for the qudit case [10]. First, we determine the Wigner function of CSS states in Section IV B 1, proving that these Wigner functions are non-negative. Then, in Section IV B 2, we consider a pure state with non-negative Wigner function and we prove that this function is precisely the Wigner function of a CSS state. Finally, the fact that the Wigner function is informationally complete allows us to conclude the proof of Theorem 1 in Section IV B 3.

1. *Wigner function of CSS states*

We start by computing the Wigner function of pure CSS states.

Lemma 2 *The Wigner function of a pure CSS state $|\psi\rangle$ is of the form*

$$W_\psi = \frac{1}{2^n}\delta_{\mathbf{t}+V_S},$$

where $\mathbf{t}$ is a vector of $\mathbb{Z}_2^{2n}$ and $V_S = N^\perp \times N$ for some subspace N of $\mathbb{Z}_2^n$. Moreover, every such function $\frac{1}{2^n}\delta_{\mathbf{t}+V_S}$ is the Wigner function of a CSS state.

In particular, the Wigner function of a pure CSS state is non-negative.

Proof of Lemma 2. Let $|\psi\rangle$ be a CSS state. Its stabilizer group $\mathcal{S}$ is generated by r independent operators $(-1)^{\alpha_i} Z(\mathbf{a}_i)$, for $1 \leq i \leq r$, and $n-r$ independent operators $(-1)^{\alpha_i} X(\mathbf{b}_i)$, for $r+1 \leq i \leq n$. Denote by N the subspace of $\mathbb{Z}_2^n$ generated by the vectors $\mathbf{b}_i$, for $r+1 \leq i \leq n$, so that its orthogonal complement is $N^\perp = \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \rangle$.

The elements of $\mathcal{S}$ are thus of the form $(-1)^{\alpha(\mathbf{v})}T_{\mathbf{v}}$, where $\mathbf{v} \in N^\perp \times N$. Moreover, we can easily check that the phase $(-1)^{\alpha(\mathbf{v})}$ defines a character of $N^\perp \times N$. Since every such character can be written as $\mathbf{v} \mapsto (-1)^{[\mathbf{t},\mathbf{v}]}$, for some vector $t \in \mathbb{Z}_2^{2n}$, we have

$$\mathcal{S} = \{(-1)^{[\mathbf{t},\mathbf{v}]}T_{\mathbf{v}} \mid \mathbf{v} \in N^\perp \times N\}.$$

Denote by $V_{\mathcal{S}}$ the subspace $N^\perp \times N$ of $\mathbb{Z}_2^{2n}$, then

$$|\psi\rangle\langle\psi| = \frac{1}{2^n}\sum_{\mathbf{v}\in V_S}(-1)^{[\mathbf{t},\mathbf{v}]}T_{\mathbf{v}}.$$

This, together with the definition Eq.(13) of $A_{\mathbf{u}}$ leads to

$$\begin{aligned}
W_\psi(\mathbf{u}) &= \frac{1}{2^n}\operatorname{Tr}(A_{\mathbf{u}}|\psi\rangle\langle\psi|) \\
&= \frac{1}{2^{3n}}\sum_{\mathbf{v}\in V_{\mathcal{S}}}\sum_{\mathbf{a}\in V_{\mathcal{A}}}(-1)^{[\mathbf{t},\mathbf{v}]}(-1)^{[\mathbf{u},\mathbf{a}]}\operatorname{Tr}(T_{\mathbf{a}}T_{\mathbf{v}}) \\
&= \frac{1}{2^{2n}}\sum_{\mathbf{v}\in V_{\mathcal{S}}}\sum_{\mathbf{a}\in V_{\mathcal{A}}}(-1)^{[\mathbf{t},\mathbf{v}]+[\mathbf{u},\mathbf{a}]}\delta_{\mathbf{a},\mathbf{v}} \\
&= \frac{1}{2^{2n}}\sum_{\mathbf{v}\in V_{\mathcal{S}}}(-1)^{[\mathbf{v},\mathbf{t}+\mathbf{u}]} \\
&= \frac{1}{2^n}\delta_{V_{\mathcal{S}}}(\mathbf{t}+\mathbf{u}) \\
&= \frac{1}{2^n}\delta_{\mathbf{t}+V_{\mathcal{S}}}(\mathbf{u}).
\end{aligned}$$

To transition from the third to the fourth line above, we have use the property that $V_{\mathcal{S}} \subset V_{\mathcal{A}}$. □

2. Non-negative Wigner functions

To complete the proof of Theorem 1, we consider a pure state which has non-negative Wigner function and we determine its Wigner function. We will show that this function coincides with the Wigner function of a CSS state. By refining the qudit proof of Gross [10] we will show that

Lemma 3 *If a pure real state $|\psi\rangle$ has non-negative Wigner function W_ψ, then its Wigner function is of the form*

$$W_\psi(\mathbf{u}) = \frac{1}{2^n}\delta_T(\mathbf{u}), \tag{15}$$

where $T = (\mathbf{p_0} + N^\perp) \times (\mathbf{q_0} + N)$, $\mathbf{p_0}, \mathbf{q_0}$ are two vectors of $\mathbb{Z}_2^n$ and N is a linear subspace of $\mathbb{Z}_2^n$

The proof of this result comprises the next 5 lemmas. First, we find, by explicit computation that

Lemma 4 *The Wigner function W_ψ of a pure real state $|\psi\rangle$, at some point $(\mathbf{p},\mathbf{q}) \in \mathbb{Z}_2^{2n}$ is*

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

where $\psi(\mathbf{x})$ denotes the inner product $\langle\psi|\mathbf{x}\rangle$.

This result is proved in Appendix B.

This encourages us to study the function $\psi : \mathbb{Z}_2^n \to \mathbb{R}$ defined by $\psi(\mathbf{x}) = \langle\psi|\mathbf{x}\rangle$. The support of ψ, denoted $\operatorname{supp}(\psi)$, is the set of vectors $\mathbf{x} \in \mathbb{Z}_2^n$ such that $\psi(\mathbf{x}) \neq 0$.

For fixed $\mathbf{q}$, we consider the function $K(\mathbf{q},\cdot)$ defined by

$$K_\psi(\mathbf{q},\mathbf{x}) = \psi(\mathbf{q})\psi(\mathbf{q}+\mathbf{x}). \tag{16}$$

It is related to the Wigner function of the state $|\psi\rangle$ via a Fourier transformation

$$\mathcal{F}K_\psi(\mathbf{q},\cdot) = 2^{n/2}W_\psi(\cdot,\mathbf{q}). \tag{17}$$

The definition of the Fourier transform for the present binary setting is recalled in Appendix A.

Eq. (17) allows us to relate properties of $|\psi\rangle$ and W_ψ.

Lemma 5 *Let $|\psi\rangle = \sum_{\mathbf{x}}\psi(\mathbf{x})|\mathbf{x}\rangle$ be a pure real state. If W_ψ is non-negative then the function ψ has constant absolute value over its support* $\operatorname{supp}(\psi)$.

Proof of Lemma 5. By Lemma 4, $W(\cdot,\mathbf{q})$ is the Fourier transform of the function $K(\mathbf{q},\cdot)$ defined in Eq.(16), up to multiplication by $2^{n/2}$. That means that $K(\mathbf{q},\cdot)$ has non-negative Fourier transform. Therefore, we can apply Bochner's theorem, exactly as stated in Theorem 44 of [10] (This result and its proof are unchanged in the binary setting). This proves that the matrix $A_{\mathbf{y}}^{\mathbf{x}} = K(\mathbf{q},\mathbf{x}-\mathbf{y})$ is postive semi-definite, where the tuples $\mathbf{x}$ and $\mathbf{y}$ are viewed as the binary writing of the matrix indices. From a well known characterization of postive semi-definite matrices, every principal minor of the matrix $A_{\mathbf{y}}^{\mathbf{x}}$ is non-negative. In particular, the determinant

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

is non-negative. This implies the following inequality.

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

If $\mathbf{q} \in \operatorname{supp}(\psi)$ and $\mathbf{x} \in \mathbb{Z}_2^n$, then we obtain

$$|\psi(\mathbf{q})| \geq |\psi(\mathbf{q}+\mathbf{x})| \tag{18}$$

since $\psi(\mathbf{q}) \neq 0$.

Now, consider two vectors $\mathbf{q}$ and $\mathbf{q}'$ of $\operatorname{supp}(\psi)$. Applying Eq.(18) to $\mathbf{q}$ and $\mathbf{x} = \mathbf{q}+\mathbf{q}'$ we find $|\psi(\mathbf{q})| \geq |\psi(\mathbf{q}+\mathbf{q}+\mathbf{q}')| = |\psi(\mathbf{q}')|$ and exchanging the roles of $\mathbf{q}$ and $\mathbf{q}'$, we obtain the reverse inequality of (18), and thus

$$|\psi(\mathbf{q})| = |\psi(\mathbf{q}')|. \tag{19}$$

This proves that ψ has constant absolute value over its support $\operatorname{supp}(\psi)$. □

Lemma 6 *Let $|\psi\rangle = \sum_{\mathbf{x}}\psi(\mathbf{x})|\mathbf{x}\rangle$ be a pure real state. If W_ψ is non-negative then the support of ψ is an affine subspace of $\mathbb{Z}_2^n$,* $\operatorname{supp}(\psi) = \mathbf{q_0} + N$.

Proof of Lemma 6. Let $\mathbf{q}, \mathbf{q}+\mathbf{x}$ and $\mathbf{q}+\mathbf{y}$ be three vectors in $\mathrm{supp}(\psi)$. We have to show that $\mathbf{q}+\mathbf{x}+\mathbf{y}$ is also in $\mathrm{supp}(\psi)$[30]. In order to obtain an equation relating more vectors of $\mathbb{Z}_2^n$, we consider the following 3×3 principal minor of the matrix $A_{\mathbf{y}}^{\mathbf{x}}$, which is also non-negative by Bochner's theorem.

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

The expansion of this determinant leads to the inequality

$$\begin{aligned}&\psi(\mathbf{q})^3\left(\psi(\mathbf{q})^3+2\psi(\mathbf{q}+\mathbf{x})\psi(\mathbf{q}+\mathbf{y})\psi(\mathbf{q}+\mathbf{x}+\mathbf{y})\right)\\&-\psi(\mathbf{q})^4\left(\psi(\mathbf{q}+\mathbf{x})^2-\psi(\mathbf{q}+\mathbf{y})^2-\psi(\mathbf{q}+\mathbf{x}+\mathbf{y})^2\right)\geq 0\end{aligned}$$

By contradiction, assume that $\psi(\mathbf{q}+\mathbf{x}+\mathbf{y})=0$, then we have

$$\psi(\mathbf{q})^6-\psi(\mathbf{q})^4\psi(\mathbf{q}+\mathbf{x})^2-\psi(\mathbf{q})^4\psi(\mathbf{q}+\mathbf{y})^2\geq 0. \quad (20)$$

From Lemma 5, the three real numbers $\psi(\mathbf{q})$, $\psi(\mathbf{q}+\mathbf{x})$ and $\psi(\mathbf{q}+\mathbf{y})$ have the same absolute value. Therefore Eq.(20) cannot be satisfied since the three terms of the left hand side are equal and positive. This contradiction implies that $\psi(\mathbf{q}+\mathbf{x}+\mathbf{y})\in\mathrm{supp}(\psi)$. Hence $\mathrm{supp}(\psi)$ is an affine space: $\mathrm{supp}(\psi)=\mathbf{q_0}+N$ where $\mathbf{q_0}\in\mathbb{Z}_2^n$ and N is a linear subspace of $\mathbb{Z}_2^n$. □

Lemma 7 *Let $|\psi\rangle=\sum_{\mathbf{x}}\psi(\mathbf{x})|\mathbf{x}\rangle$ be a pure real state. If W_ψ is non-negative then for every $\mathbf{q}\in\mathbf{q_0}+N$, the function $W_\psi(\cdot,\mathbf{q})$ is*

$$W_\psi(\cdot,\mathbf{q})=c\delta_{\mathbf{p_0}+\mathbf{N}^\perp},$$

where $c=c(\mathbf{q})\in\mathbb{R}$ and $\mathbf{p_0}=\mathbf{p_0}(\mathbf{q})\in\mathbb{Z}_2^n$ may both depend on $\mathbf{q}$. Moreover, if $\mathbf{q}\notin\mathbf{q_0}+N$, then $W_\psi(\cdot,\mathbf{q})=0$.

Proof of Lemma 7. First, we fix a vector $\mathbf{q}\in\mathbb{Z}_2^n$ and we focus on the support the function $W_\psi(\cdot,\mathbf{q})$. From Lemma 4, this function satisfies

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

Therefore, $W_\psi(\cdot,\mathbf{q})$ is the zero function when $\mathbf{q}$ does not belong to the support of ψ, which is $\mathbf{q_0}+N$ from Lemma 6.

In what follows, the vector $\mathbf{q}$ is chosen in $\mathbf{q_0}+N$. In the above expression of W_ψ, the term $\psi(\mathbf{q})\psi(\mathbf{q}+\mathbf{x})$ can be replaced by $K(\mathbf{q},\mathbf{x})$, defined in Eq.(16). The support of the function K is $\mathrm{supp}(K)=(\mathbf{q_0}+\mathbf{N})\times\mathbf{N}$ where $\mathbf{q_0}+\mathbf{N}$ is the support of ψ. Then, K can be restricted to its support. This gives

$$W_\psi(\mathbf{p},\mathbf{q})=\frac{1}{2^n}\sum_{\mathbf{x}\in\mathbf{N}}(-1)^{(\mathbf{p},\mathbf{x})}K'(\mathbf{q},\mathbf{x})$$

where K' is the restriction of K to its support.

Now note that, for every vector $\mathbf{q}\in\mathbb{Z}_2^{\mathbf{n}}$, the function $W_\psi(\cdot,\mathbf{q})$ is constant over the cosets of $N^\perp$. Therefore, this function induces a function $W_\psi([\cdot],\mathbf{q})$ over $\mathbb{Z}_2^n/(N^\perp)$:

$$\begin{aligned}W_\psi([\cdot],\mathbf{q}):\mathbb{Z}_2^n/(N^\perp)&\longrightarrow\mathbb{R}\\[\mathbf{p}]=\mathbf{p}+N^\perp&\longmapsto W_\psi(\mathbf{p},\mathbf{q}).\end{aligned}$$

The space $\mathbb{Z}_2^n/(N^\perp)$ is isomorphic to the linear space N. Indeed, the application from $\mathbb{Z}_2^n$ to the dual N^* of N, defined by $\mathbf{x}\mapsto(\mathbf{x},\cdot)$ induces an isomorphism between $\mathbb{Z}_2^n/(N^\perp)$ and N^*. Thus, N^* is canonically isomorphic to N.

Up to this isomorphism $\mathbb{Z}_2^n/(N^\perp)\simeq N$, the functions $K'(\mathbf{q},\cdot)$ and $W_\psi([\cdot],\mathbf{q})$ are both defined over the same space and $W_\psi([\cdot],\mathbf{q})$ is the Fourier transform of $K'(\mathbf{q},\cdot)$ up to multiplication by $2^{n/2}$, that is $\mathcal{F}K'(\mathbf{q},\cdot)=2^{n/2}W_\psi([\cdot],\mathbf{q})$. Applying $\mathcal{F}$ to this equality, we obtain

$$2^{-n/2}K'(\mathbf{q},\cdot)=\mathcal{F}W_\psi([\cdot],\mathbf{q}),$$

because $\mathcal{F}$ is involutive, from Lemma 18 in Appendix A.

The function $K'(\mathbf{q},\cdot)$ has constant absolute value over N by Lemma 5, thus we apply the second item of Bochner Theorem (Theorem 44 in [10]) to $W_\psi([\cdot],\mathbf{q})$. This tells us that $W_\psi([\cdot],\mathbf{q})$ is orthogonal to its translations, *i.e.*

$$\sum_{[\mathbf{p}]}W_\psi([\mathbf{p}],\mathbf{q})W_\psi([\mathbf{p}]+[\mathbf{t}],\mathbf{q})=\mathbf{0},$$

for every $[\mathbf{t}]\in\mathbb{Z}_2^n/(N^\perp)$. A positive function which satisfies this orthogonality condition can be either zero or proportional to an indicator function $\delta_{[\mathbf{p_0}]}$. But $W_\psi([\cdot],\mathbf{q})$ cannot be zero. Otherwise $K(\mathbf{q},\cdot)$ is also the zero function by injectivity of the Fourier transform, and this cannot happen when $\mathbf{q}$ is chosen in $\mathbf{q_0}+N$. □

The next lemma concludes the proof of Lemma 3.

Lemma 8 *Let $|\psi\rangle=\sum_{\mathbf{x}}\psi(\mathbf{x})|\mathbf{x}\rangle$ be a pure real state. If W_ψ is non-negative then W_ψ is of the form*

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

where $\mathbf{p_0},\mathbf{q_0}\in\mathbb{Z}_2^n$ and N is a linear subspace of $\mathbb{Z}_2^n$.

Proof of Lemma 8. From Lemma 7, the global support, $\mathrm{supp}(W_\psi)$ is the disjoint union

$$\mathrm{supp}(W_\psi)=\bigsqcup_{\mathbf{q}\in\mathbf{q_0}+\mathbf{N}}\left(\mathbf{p_0}(\mathbf{q})+\mathbf{N}^\perp\right)\times\{\mathbf{q}\}. \quad (21)$$

Our first goal is to prove that $\mathbf{p_0}(\mathbf{q})$ does not depend on $\mathbf{q}$. To this end, it is natural to separate the variables $\mathbf{p}$ and $\mathbf{q}$ in the writing of W_ψ obtained in Lemma 4. This leads to

$$W_\psi(\mathbf{p},\mathbf{q})=(-1)^{(\mathbf{p},\mathbf{q})}\hat{\psi}(\mathbf{p})\psi(\mathbf{q}), \quad (22)$$

where $\hat{\psi}$ is the Fourier transform of ψ. Thus the support of W_ψ is also $\mathrm{supp}(W_\psi)=\mathrm{supp}(\hat{\psi})\times\mathrm{supp}(\psi)$. This

can be satisfied if and only if $\mathbf{p_0}$ is independent of $\mathbf{q}$ in Eq.(21). This proves that the support of W_ψ is the cartesian product

$$\mathrm{supp}(W_\psi) = (\mathbf{p_0} + N^\perp) \times (\mathbf{q_0} + N).$$

Now, let us prove that W_ψ has constant absolute value over its support. Let $(\mathbf{p}, \mathbf{q}) \in \mathrm{supp}(W_\psi)$. Combining Lemma 7 and Eq.(22), we find that the modulus of $W_\psi(\mathbf{p}, \mathbf{q})$ is

$$|c(\mathbf{q})| = |\hat{\psi}(\mathbf{p})| \cdot |\psi(\mathbf{q})|,$$

where $c(\mathbf{q})$ is the constant introduced in Lemma 7. Recall that $c(\mathbf{q})$ is independent of $\mathbf{p}$. We proved in Lemma 5 that $|\psi(\mathbf{q})|$ is constant, therefore $|c(\mathbf{q})|$ is also independent of $\mathbf{q}$. This proves that $|c|$ is constant over $\mathrm{supp}(\psi)$. By positivity of W_ψ, we have $c = |c|$ and

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

for some constant $c \in \mathbb{R}$.

To conclude the proof it remains to evaluate the value of c. By the normalisation of Property 6. of the Wigner function, it suffices to compute the cardinality of the support of W_ψ. We find $|\mathrm{supp}(W_\psi)| = |N^\perp| \cdot |N| = 2^{n-\dim N} \cdot 2^{\dim N} = 2^n$, which gives $c = 1/2^n$. This concludes the proof. □

3. Proof of Hudson's theorem for rebits

Lemma 2 together with Lemma 3 enable us to prove a rebit version of Hudson's Theorem.

Proof of Theorem 1. Lemma 2 implies that every CSS states has non-negative Wigner function.

Now, consider a pure real state $|\psi\rangle$ which admits a non-negative Wigner function. In order to prove that this is a CSS state, it is enough to prove that its Wigner function coincides with the Wigner function of a pure CSS state $|\varphi\rangle$. Indeed, since the Wigner function is informationally complete (Property 1.), this implies $|\psi\rangle = |\varphi\rangle$. We proved in Lemma 3 that W_ψ can be written

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

Since $(\mathbf{p_0}+N^\perp)\times(\mathbf{q_0}+N) = \mathbf{t}+V_S$, where $\mathbf{t} = (\mathbf{p_0}, \mathbf{q_0})$ and $V_S = N^\perp \times N$, this is indeed the Wigner function of a CSS state by Lemma 2. □

C. Covariance of the rebit Wigner function

Our next goal is to demonstrate that the action of CSS-ness preserving Clifford gates on Wigner functions W_ρ can be understood simply from the action of such gates on the underlying phase space, c.f. Lemma 11 below. To prepare for this result, we make two observations.

Lemma 9 *Let $g \in G_{CSS}$. Then, there exists a unique pair $(F, \mathbf{x})$ composed of a vector $\mathbf{x} \in \mathbb{Z}_2^{2n}$ and a symplectic matrix $F \in \mathrm{Sp}_{2n}(\mathbb{Z}_2)$ such that*

$$gT_\mathbf{a}g^\dagger = (-1)^{[\mathbf{x},\mathbf{a}]}T_{F\mathbf{a}}, \quad \forall \mathbf{a} \in V_\mathcal{A}. \tag{23}$$

The proof of Lemma 9 is given in Appendix C.

Furthermore, the action of a $g \in G_{CSS}$ on a translation operator $T_\mathbf{a} \in \mathcal{T}$ by conjugation induces a morphism from the CSS Clifford group to the affine group $\mathrm{AGL}_{2n}(\mathbb{Z}_2)$. Recall that an affine transformation of $\mathbb{Z}_2^{2n}$ is an application of the form $A(F, \mathbf{t}) : \mathbf{a} \mapsto F\mathbf{a} + \mathbf{t}$, where $F \in \mathrm{GL}_{2n}(\mathbb{Z}_2)$ is a linear application and $\mathbf{t}$ is a vector of $\mathbb{Z}_2^{2n}$. In the present work F is often symplectic and this affine map is then called an affine symplectic map. The set of affine symplectic transformations of $\mathbb{Z}_2^{2n}$ is a subgroup of the affine group denoted $\mathrm{ASp}_{2n}(\mathbb{Z}_2^{2n})$.

Lemma 10 *Let $\mathcal{F}$ be the application*

$$\begin{aligned} \mathcal{F} : G_{CSS} &\longrightarrow \mathrm{ASp}_{2n}(\mathbb{Z}_2) \\ g &\longmapsto A(F, \mathbf{t}) \end{aligned}$$

such that $gT_\mathbf{a}g^\dagger = (-1)^{[\mathbf{t},F\mathbf{a}]}T_{F\mathbf{a}}$, for all $\mathbf{a}$. Then $\mathcal{F}$ is a group morphism.

The proof of Lemma 10 is given in Appendix C. The application $\mathcal{F}$ is well defined by unicity in Lemma 9. The translation vector $\mathbf{t}$ and the vector $\mathbf{x}$ of Lemma 9 are related by the equation $\mathbf{t} = F\mathbf{x}$.

We are now ready to state the covariance result.

Lemma 11 *The n-rebit Wigner function W is covariant under G_{CSS}, in the sense that for all ρ, for all $\boldsymbol{u} \in \mathbb{Z}_2^{2n}$, and for all $g \in G_{CSS}$ it holds that*

$$W_{g^\dagger\rho g}(\boldsymbol{u}) = W_\rho\left(\mathcal{F}(g)(\mathbf{u})\right). \tag{24}$$

Applying this result to $g\rho g^\dagger = (g^{-1})^\dagger \rho g^{-1}$, we find

$$W_{g\rho g^\dagger}(\mathbf{u}) = W_\rho\left(\mathcal{F}(g)^{-1}(\mathbf{u})\right) = W_\rho\left(F^{-1}(\mathbf{u}+\mathbf{t})\right),$$

where $\mathcal{F}(g) = A(F, \mathbf{t})$.

Proof of Lemma 11. Let $g \in G_{CSS}$ and let $\mathcal{F}(g) = A(F, \mathbf{t})$ be its induced affine symplectic map. First, consider the image of $A_\mathbf{u}$ by conjugation by g. Using Eq.(13), we obtain

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

where we have used $[\mathbf{u},\mathbf{a}] = [F\mathbf{u}, F\mathbf{a}]$ and the fact that F induces a bijection of the set $V_{\mathcal{A}}$. This leads to

$$\begin{aligned} 2^n W_{g^\dagger \rho g}(\mathbf{u}) &= \mathrm{Tr}(A_{\mathbf{u}} g^\dagger \rho g) \\ &= \mathrm{Tr}(g A_{\mathbf{u}} g^\dagger \rho) \\ &= \mathrm{Tr}(A_{\mathcal{F}(g)(\mathbf{u})} \rho) \\ &= 2^n W_\rho(\mathcal{F}(g)(\mathbf{u})), \end{aligned}$$

which proves the covariance.

For $n \geq 2$, W is not covariant under all real Clifford operations. As an example, consider $n = 2$ and $g = H_1$, which is real Clifford but not CSS-ness preserving. H_1 converts a Bell state into a 2-qubit graph state. The former has positive and the latter negative Wigner function. Hence, H_1 does not transform W covariantly.

D. Efficient simulation of Clifford circuits

An operational justification for emphasizing positivity of Wigner functions is the following result [6] for qudits: Circuits of Clifford gates and stabilizer measurements acting on an initial state with non-negative Wigner function can be efficiently simulated classically. The discrete Hudson's theorem [10] ensures that for pure states, the simulation method based on Wigner functions has the same scope as the Gottesman-Knill theorem. For mixed states it is an extension of that theorem, since not all states with non-negative Wigner function are mixtures of stabilizer states [10].

Here we prove an analogue of the result [6] for the rebit Wigner function W defined in Eqs. (11), (12).

Theorem 2 *Every circuit consisting of CSS-ness preserving Clifford unitaries and measurements, acting on a product state $\rho = \bigotimes_{i=1}^n \rho_i$ with non-negative Wigner function W_ρ, can be efficiently classically simulated.*

Proof of Theorem 2. We describe a simulation method based on sampling. For a quantum state ρ represented by a Wigner function W_ρ, the probability of an outcome s corresponding to the POVM element $E(s)$ is

$$P(s) = \sum_{\mathbf{u}} W_\rho(\mathbf{u}) W_{E(s)}(\mathbf{u}).$$

For the allowed observables $O \in \mathcal{O}$, the POVM elements $E(s) = (I + s\,O)/2$ all have positive Wigner function $W_{E(s)}$. Therefore, $P(s)$ can be efficiently estimated if W_ρ is positive (i.e., is a probability distribution), and can be efficiently sampled from. We show by induction that this is indeed the case for all Wigner functions generated by the above circuits.

First, the initial Wigner function for the state $\rho(0) = \rho_1 \otimes \rho_2 \otimes .. \otimes \rho_n$, $W_{\rho(0)} = W_{\rho_1} W_{\rho_2} \cdot .. \cdot W_{\rho_n}$, can be efficiently sampled from. It is positive, and the W_{ρ_i} may be sampled from independently, which is efficient.

Now we show that if the Wigner function $W_{\rho(t)}$ after time step t can be efficiently sampled from, then so can the Wigner function $W_{\rho(t+1)}$ after step $t+1$. We distinguish two cases: (a) $\rho(t+1) = g\rho(t)g^\dagger$, with $g \in G_{CSS}$, and (b) $\rho(t+1) \sim \frac{I+sO}{2}\rho(t)\frac{I+sO}{2}$, with $O \in \mathcal{O}$, $s = \pm 1$.

(a) Unitary evolution. The Wigner function transforms covariantly under gates $g \in G_{CSS}$,

$$W_{\rho(t+1)}(F_g \mathbf{u} + \mathbf{t}_g) = W_{\rho(t)}(\mathbf{u}).$$

Thus, sampling from $W_{\rho(t+1)}$ can be efficiently reduced to sampling from $W_{\rho(t)}$. In particular, gates in G_{CSS} preserve the positivity of the Wigner function.

(b) Projective measurement. We note

Lemma 12 *The Wigner function of the state ρ' of the system after measuring $T_{\mathbf{a}} \in \mathcal{O}$ with the outcome $s \in \{\pm 1\}$ is*

$$W_{\rho'}(\mathbf{u}) = \begin{cases} \frac{1}{2}\left(W_\rho(\mathbf{u}) + W_\rho(\mathbf{u}+\mathbf{a})\right) & \text{if } s \cdot (-1)^{[\mathbf{u},\mathbf{a}]} = 1 \\ 0 & \text{else} \end{cases}$$

where ρ is the state before measurement. In particular, measurements of observables in $\mathcal{O}$ preserves the positivity of the Wigner function of the system.

$W_{\rho(t+1)}$ is sampled from as follows. Repeat: (1) The sampling routine for $W_{\rho(t)}$ is called, which returns a $\mathbf{u} \in V$. (2) The measurement outcome $s = (-1)^{[\mathbf{u},\mathbf{a}]}$ is reported. (3) A fair coin is flipped, and, depending on the outcome, $\mathbf{u}$ or $\mathbf{u}+\mathbf{a}$ is reported as sample from $W_{\rho(t+1)}$.

This concludes the proof of Theorem 2, subject to the proof of Lemma 12. □

Remark 1: The locality of the initial state, $\rho(0) = \rho_1 \otimes \rho_2 \otimes .. \otimes \rho_n$ is of no physical significance. It is just one possible way to ensure that the positive $W_{\rho(0)}$ can be efficiently sampled from by a classical algorithm.

Remark 2: In most respects, the present simulation method is the same as its qudit counterpart [21], but a difference occurs in measurement. Here, mere positivity of the effect $W_{E(s)}$ and positivity of $W_{\rho_{\mathrm{in}}}$ for the input state ρ_{in} do not imply positivity of the Wigner function $W_{\rho_{\mathrm{out}}}$ for the output state ρ_{out}. Example: The two-rebit state $\rho = (I + X_1 Z_2)/4$ has positive Wigner function, and the POVM-element $(I + Z_1 X_2)/2$ is also positively represented. However, the state after measurement, a pure stabilizer state with stabilizer group $\mathcal{S} = \langle X_1 Z_2, Z_1 X_2 \rangle$, has *negative* Wigner function. Note that $Z_1 X_2 \notin \mathcal{O}$.

Proof of Lemma 12. For all $T_{\mathbf{a}} \in \mathcal{A}$, $T_{\mathbf{y}} \in \mathcal{O}$, it holds that

$$\text{if } [T_{\mathbf{a}}, T_{\mathbf{y}}] = 0, \text{ then } T_{\mathbf{a}} T_{\mathbf{y}} = T_{\mathbf{a}+\mathbf{y}}. \tag{25}$$

This is a consequence of all $T_{\mathbf{y}} \in \mathcal{O}$ being entirely of X-type or Z-type (by definition of $\mathcal{O}$).

We define the set $\mathcal{A}_{\mathbf{y}}$ as $\mathcal{A}_{\mathbf{y}} = \{T_{\mathbf{a}} \in \mathcal{A} |\, [\mathbf{a},\mathbf{y}] = 0\}$. It has the property that

$$T_{\mathbf{y}} \mathcal{A}_{\mathbf{y}} = \mathcal{A}_{\mathbf{y}}. \tag{26}$$

Eq. (26) holds because $[T_{\mathbf{a}}, T_{\mathbf{y}}] = 0 \Leftrightarrow [T_{\mathbf{y}} T_{\mathbf{a}}, T_{\mathbf{y}}] = 0$, and Eq. (25) ($T_{\mathbf{y}} T_{\mathbf{a}} \in \mathcal{A}$, i.e., has the right sign).

Now, the update $W_\rho \mapsto W_{\rho'}$ under measurement of the observable $T_{\mathbf{y}} \in \mathcal{O}$, with outcome $s = \pm 1$, is

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

When transitioning from the third to the fourth line above, we used the property Eq. (26). □

V. CONTEXTUALITY

A. Scope of hidden variable models for rebit QCSI

A quantum-mechanical setting comprising quantum states and measurements is said to be contextual if it cannot be described by any non-contextual hidden variable model. For the rebit scheme of quantum computation by state injection considered here, we first need to determine the scope of the phenomenology that any purported non-contextual HVM needs to reproduce.

The set of quantum states is unrestricted. The candidate HVM must yield the correct measurement statistics for any real quantum state. However, the observables which can be measured in rebit QCSI, and the sets of observables which can be measured jointly, are restricted. To analyze the situation, we first discuss a few examples, and then impose a general criterion.

First, the set of observables which can be physically measured in rebit QCSI is $\mathcal{O} = \{X(\mathbf{a}_X), Z(\mathbf{a}_Z)\}$. The candidate HVM therefore needs to correctly reproduce the probabilities of measurement outcomes for all observables $O \in \mathcal{O}$, and furthermore the correct joint outcome probability distributions for any number of commuting observables in $\mathcal{O}$.

But there is more. For example, consider the two-rebit observable $X_1 Z_2$, which is in the set $\mathcal{A}$ but not in $\mathcal{O}$. The measurement outcome of $X_1 Z_2$ can be obtained by measuring the commuting observables $X_1, Z_2 \in \mathcal{O}$, and then post-processing the outcomes. Therefore, a measurement of $X_1 Z_2$ can be reduced to measurements of observables in $\mathcal{O}$. The same holds for all observables in $\mathcal{A}$. We therefore require that any candidate HVM must reproduce the correct measurement statistics for all observables in $\mathcal{A}$.

We now turn to the simultaneous measurement of compatible observables. Continuing with the above example, it is possible to simultaneously measure the pair of observables $\{X_1, X_1 Z_2\}$, namely by the same operations that measured $X_1 Z_2$ alone.

Now, is it possible to simultaneously measure the commuting observables $X_1 Z_2$ and $Z_1 X_2$? In the setting of rebit QCSI, this is not the case. The measurement of $X_1 Z_2$ necessitates the measurement of X_1 and Z_2 separately. Since these observables do not commute with $Z_1 X_2$, a subsequent measurement of $Z_1 X_2$ is no longer guaranteed to reveal the original value. Thus, compatible observables in $\mathcal{A}$ need not be simultaneously measurable in the same way as compatible observables in $\mathcal{O}$.

Based on the phenomenology discussed above, we adopt the following operational criterion to define the scope of hidden variable models:

Criterion 1 *Be M a set of commuting observables. Any hidden variable model describing M must correctly predict the joint probability distribution p_M of measurement outcomes, if for all observables $O \in M$ the outcomes can be simultaneously obtained from measurements on a single copy of the given quantum state.*

We denote by $\mathcal{M}$ the set of measurement settings $M \subset \mathcal{A}$ admitted by Criterion 1. Given a quantum state ρ and a set M of compatible observables, we denote by $p_{M,\rho}$ the probability distribution for measurement outcomes corresponding M.

Definition 1 *A hidden variable model describing the physical setting $(\rho, \mathcal{M})$ consists of (a) a non-empty set $\mathcal{S}$ of internal states, (b) a probability distribution q over $\mathcal{S}$, and (c) conditional probabilities $p(\boldsymbol{s}_M | \, \boldsymbol{u})$, $\boldsymbol{u} \in \mathcal{S}$, for outcomes $\boldsymbol{s}_M = (s_1, s_2, .., s_{|M|})$ of measurements in M, $M \in \mathcal{M}$, such that*

(i) For every $\boldsymbol{u} \in \mathcal{S}$, all observables $O \in \mathcal{A}$ have definite values, $\lambda_{\boldsymbol{u}}(O) = \pm 1$, and for all $M \in \mathcal{M}$

$$
p(\boldsymbol{s}_M | \, \boldsymbol{u}) = \prod_{i \,|\, O_i \in M} \delta_{s_i, \lambda(O_i)}. \tag{27}
$$

(ii) For all $M \in \mathcal{M}$, all triples of commuting observables $A, B, AB \in \langle M \rangle$, and all $\boldsymbol{u} \in \mathcal{S}$, the value assignments are consistent,

$$
\lambda_{\boldsymbol{u}}(AB) = \lambda_{\boldsymbol{u}}(A) \lambda_{\boldsymbol{u}}(B). \tag{28}
$$

(iii) Given the quantum state ρ, the probability distribution q_ρ reproduces all probability distributions of measurement outcomes; i.e.

$$
p_{M,\rho}(\boldsymbol{s}_M) = \sum_{\boldsymbol{u} \in \mathcal{S}} p(\boldsymbol{s}_M | \, \boldsymbol{u}) \, q_\rho(\boldsymbol{u}), \tag{29}
$$

for all $M \subset \mathcal{M}$, and all values of $\boldsymbol{s}_M$.

In Sections V B and V C below, we derive necessary and sufficient conditions for the existence of a hidden variable model over $\mathcal{M}$, or, the other way around, for contextuality. These conditions are expressed in terms of the rebit Wigner function.

We conclude this section with a characterization of the sets $M \in \mathcal{M}$ of simultaneously measurable observables in QCSI that are admitted by Criterion 1.

Lemma 13 *Be $M \subset \mathcal{A}$ a set of commuting observables. Then, $M \in \mathcal{M}$ if and only if $T_a T_b = T_{a+b}, \forall\, T_a, T_b \in M$.*

Remark 3: What is excluded here is the possibility of $T_{\mathbf{a}} T_{\mathbf{b}} = -T_{\mathbf{a}+\mathbf{b}}$.

Proof of Lemma 13. "If": Assume that a set $M \subset \mathcal{A}$ has the property that $T_{\mathbf{a}} T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}}$ for all $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$. Since $T_{\mathbf{a}+\mathbf{b}} = (-1)^{\mathbf{a}_X \cdot \mathbf{b}_Z} T_{\mathbf{a}} T_{\mathbf{b}}$, it follows that $\mathbf{a}_X \cdot \mathbf{b}_Z = \mathbf{a}_Z \cdot \mathbf{b}_X = 0 \pmod 2$, for all $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$.

Therefore, for all $T_{\mathbf{a}} \in M$, the operators $X(\mathbf{a}_X)$ and $Z(\mathbf{a}_Z)$ commute with all of M and among themselves. They thus generate a CSS stabilizer

$$S = \big\langle X(\mathbf{a}_X), Z(\mathbf{a}_Z) |\, T_{\mathbf{a}} \in M \big\rangle.$$

By construction, $M \subset S$. Therefore, the measurement outcomes for all observables $O \in M$ can be obtained by measuring the set of observables $\{X(\mathbf{a}_X), Z(\mathbf{a}_Z) |\, T_{\mathbf{a}} \in M\} \subset \mathcal{O}$, and subsequent classical processing. The set M thus satisfies Criterion 1.

"Only if": Since physical measurements are restricted to observables on $\mathcal{O}$, the only way of measuring an observable $T_{\mathbf{a}} \in \mathcal{A}$ is to separately measure its X-part $X(\mathbf{a}_X)$ and Z-part $Z(\mathbf{a}_Z)$, and then post-process the measurement outcomes. We assume that for a given set $M = \{T_{\mathbf{a}}\} \subset \mathcal{A}$ Criterion 1 holds. Then, $[X(\mathbf{a}_X), Z(\mathbf{b}_Z)] = 0$, for all $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$, or, equivalently, $\mathbf{a}_X \cdot \mathbf{b}_Z = 0$, for all $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$. Since $T_{\mathbf{a}+\mathbf{b}} = (-1)^{\mathbf{a}_X \cdot \mathbf{a}_Z} T_{\mathbf{a}} T_{\mathbf{b}}$, it follows that $T_{\mathbf{a}+\mathbf{b}} = T_{\mathbf{a}} T_{\mathbf{b}}$ for all $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$. □

For an illustration of Lemma 13, we previously argued that $X_1 Z_1$ and $Z_1 X_2$ cannot be simultaneously measured in rebit QCSI; $\{X_1 Z_1, Z_1 X_2\} \notin \mathcal{M}$. Lemma 13 detects this as follows: If $T_{\mathbf{a}} = X_1 Z_2$ and $T_{\mathbf{b}} = Z_1 X_2$ then $T_{\mathbf{a}+\mathbf{b}} = -Y_1 Y_2$, and therefore $T_{\mathbf{a}+\mathbf{b}} = -T_{\mathbf{a}} T_{\mathbf{b}}$.

B. A necessary condition for contextuality

Theorem 3 *The setting $(\rho, \mathcal{M})$ is contextual only if $W_\rho < 0$.*

Proof of Theorem 3. If $W_\rho > 0$ then W_ρ *is* a valid non-contextual HVM for the setting $(\rho, \mathcal{M})$. To verify this claim, we need to check that if $W_\rho > 0$ then W_ρ provides the constructs (a) - (c) required in Definition 1, and that the conditions (i) - (iii) therein are satisfied.

A projective measurement of a set $M \in \mathcal{M}$ of commuting observables is represented by POVM elements $E(\mathbf{s}_M)$,

$$E(\mathbf{s}_M) = \prod_{i | T_{\mathbf{a}(i)} \in M} \frac{I + s_i T_{\mathbf{a}(i)}}{2}, \tag{30}$$

and $s_i = \pm 1$, for all i. With Eq. (14), the probability of obtaining the outcomes $\mathbf{s}_M$ in the measurement of the set of observables M is

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

We thus identify (a) $V = \mathcal{S}$, (b) $W_\rho = q$, and (c) $2^n W_{E(\mathbf{s}_M)}(\mathbf{u}) = p(\mathbf{s}_M | \mathbf{u})$, for all $\mathbf{u}$. $V = \mathbb{Z}_2^{2n}$ is a valid state space and W_ρ a valid probability distribution, since by assumption $W_\rho > 0$.

It remains to show that $W_{E(\mathbf{s}_M)} > 0$ for all $M \in \mathcal{M}$. First, we compute $W_{E(s)}$ for $E(s) = \frac{I + s T_{\mathbf{a}}}{2}$ and $T_{\mathbf{a}} \in \mathcal{A}$. Using the orthogonality relation $\mathrm{Tr}(T_{\mathbf{a}} T_{\mathbf{b}}) = 2^n \delta_{\mathbf{a},\mathbf{b}}$, we find that $2^n W_{E(s)}(\mathbf{u}) = \delta_{s,(-1)^{[\mathbf{u},\mathbf{a}]}}$. Thus, for all observables $T_{\mathbf{a}} \in \mathcal{A}$ and all states $\mathbf{u} \in V$, we obtain the value assignment

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u},\mathbf{a}]}. \tag{31}$$

We now generalize the above computation of the Wigner function of effects from the observables in $\mathcal{A}$ to all sets $M \in \mathcal{M}$ of measurements. To this end, we note that by Lemma 13 the POVM elements $E(\mathbf{s}_M)$ of Eq. (30) can be rewritten as

$$E(\mathbf{s}_M) = \frac{1}{2^{|M|}} \left(\sum_{N \subset M} \left[\prod_{T_{\mathbf{a}(i)} \in N} s_i \right] T_{\sum_{T_{\mathbf{a}(i)} \in N} \mathbf{a}(i)} \right).$$

Hence we obtain

$$2^n W_{E(\mathbf{s}_M)}(\mathbf{u}) = \prod_{i | T_{\mathbf{a}(i)} \in M} \delta_{s_i, (-1)^{[\mathbf{u}, \mathbf{a}(i)]}}. \tag{32}$$

Thus, $2^n W_{E(\mathbf{s}_M)}$ does indeed represent conditional probabilities, as required for $2^n W_{E(\mathbf{s}_M)}(\mathbf{u}) = p(\mathbf{s}_M | \mathbf{u})$.

Regarding (i), the assignment of Eq. (31) demonstrates that for all states $\mathbf{u} \in \mathcal{S}$, all observables in $\mathcal{A}$ have definite values, as required. Furthermore, for this value assignment, the expression Eqs. (32) for the conditional probability $p(\mathbf{s}_M | \mathbf{u})$ matches the required expression Eq. (27).

Regarding (ii), the value assignment Eq. (31) leads to the constraints

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}+\mathbf{b}}) = \lambda_{\mathbf{u}}(T_{\mathbf{a}}) \lambda_{\mathbf{u}}(T_{\mathbf{b}}),\ \forall \mathbf{u} \in \mathcal{S}, \forall\, T_{\mathbf{a}}, T_{\mathbf{b}}, T_{\mathbf{a}+\mathbf{b}} \in \mathcal{A}.$$

Since, by Lemma 13, $T_{\mathbf{a}+\mathbf{b}} = T_{\mathbf{a}} T_{\mathbf{b}}$ for all $T_{\mathbf{a}}, T_{\mathbf{b}}, T_{\mathbf{a}+\mathbf{b}} \in \langle M \rangle$, the value assignments of Eq. (31) are consistent for all $M \in \mathcal{M}$.

Finally, condition (iii) is satisfied by construction of the Wigner function.

We have thus shown that if $W_\rho > 0$ then W_ρ provides a non-contextual HVM for the setting $(\rho, \mathcal{M})$. The claim follows by negation of this statement. □

C. A sufficient condition for contextuality

Theorem 4 *The n-rebit setting $(\rho, \mathcal{M})$ is contextual if there exists a maximal isotropic subspace $U \subset \mathbb{Z}_2^{2n}$ and a vector $\nu \in \mathbb{Z}_2^{2n}$ such that*

$$\sum_{\boldsymbol{v} \in U} W_\rho(\boldsymbol{v} + \nu) < 0.$$

Comparing Theorems 3 and 4, we find that our necessary and sufficient conditions for contextuality do not match. This indicates the possibility of a Wigner-negative non-contextual phase. Such a phase does indeed exist, as we show later.

To prove Theorem 4, we construct a family of witness functions $\mathcal{W}$ which can detect contextuality. Each such function is based on an isotropic subspace $U \subset \mathbb{Z}_2^{2n}$ with a basis $\mathcal{B}(U) = \{\mathbf{a}(1), \mathbf{a}(2), .., \mathbf{a}(m)\}$, and can be evaluated on points $\mathbf{x} \in \mathbb{Z}_2^m$, for any density operator ρ. Namely, we define

$$\mathcal{W}_\rho^{\mathcal{B}(U)}(\mathbf{x}) = \left\langle \sum_{\mathbf{z} \in \mathbb{Z}_2^m} \left[\prod_{i=1}^m (-1)^{z_i x_i} \right] T_{\sum_i z_i \mathbf{a}(i)} \right\rangle_\rho . \quad (33)$$

Lemma 14 *The setting $(\rho, \mathcal{M})$ is contextual if there exists an isotropic subspace $U \subset \mathbb{Z}_2^{2n}$ such that $\mathcal{W}_\rho^{\mathcal{B}(U)} < 0$.*

Proof of Lemma 14. We prove the converse statement, namely that if $(\rho, \mathcal{M})$ is non-contextual then $\mathcal{W}_\rho^{\mathcal{B}(U)} > 0$ for all isotropic subspaces $U \in \mathbb{Z}_2^{2n}$ and all bases thereof.

Assume there exists a non-contextual HVM describing the setting $(\rho, \mathcal{M})$. Then, by property (i) of Definition 1, the states of this HVM must have definite values ± 1 for all observables in $\mathcal{A}$. Furthermore, for any state $\mathbf{u}$ of the HVM, these values must satisfy the consistency condition (ii) of Definition 1.

Specifically, the set $M = \{Z(\mathbf{a}_Z) | \, a_{\mathbf{Z}} \in \mathbb{Z}_2^n\}$ satisfies Criterion 1. Therefore by Property (ii) of Def. 1, $\lambda_{\mathbf{u}}\left(T_{(\mathbf{a}_Z, 0)}\right) = \lambda_{\mathbf{u}}(Z(\mathbf{a}_Z)) = \prod_{i|[\mathbf{a}_Z]_i = 1} \lambda_{\mathbf{u}}(Z_i)$. Likewise, $\lambda_{\mathbf{u}}\left(T_{(0, \mathbf{a}_X)}\right) = \lambda_{\mathbf{u}}(X(\mathbf{a}_X)) = \prod_{i|[\mathbf{a}_X]_i = 1} \lambda_{\mathbf{u}}(X_i)$. Analogously, for any $T_{(\mathbf{a}_Z, \mathbf{a}_X)} \in \mathcal{A}$, the set $M = \{T_{(\mathbf{a}_Z, 0)}, T_{(0, \mathbf{a}_X)}, T_{(\mathbf{a}_Z, \mathbf{a}_X)}\}$ satisfies Criterion 1, since by definition of $\mathcal{A}$ the Pauli operators $T_{(\mathbf{a}_Z, 0)}$, $T_{(0, \mathbf{a}_X)}$ commute, and $T_{(\mathbf{a}_Z, 0)} T_{(0, \mathbf{a}_X)} = T_{(\mathbf{a}_Z, \mathbf{a}_X)}$. Therefore, by Eq. (28), $\lambda_{\mathbf{u}}\left(T_{(\mathbf{a}_z, \mathbf{a}_x)}\right) = \lambda_{\mathbf{u}}\left(T_{(\mathbf{a}_z, 0)}\right) \lambda_{\mathbf{u}}\left(T_{(0, \mathbf{a}_x)}\right)$.

Combining the above three relations, we find that for all $T_{\mathbf{a}} \in \mathcal{A}$, the value $\lambda(T_{\mathbf{a}})$ follows from the values $\lambda(X_i)$, $\lambda(Z_i)$ assigned to the local observables X_i and Z_i, for $i = 1, .., n$. We may write this as

$$\lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u}, \mathbf{a}]}, \; \forall \mathbf{u} \in \mathcal{S}, \quad (34)$$

and $\mathcal{S} = \mathbb{Z}_2^{2n}$. We find that the same relation Eq. (31) which held for HVMs derived from the Wigner function holds for all non-contextual HVMs.

As a consequence, for all $\mathbf{u} \in \mathcal{S}$, it holds that

$$\lambda_{\mathbf{u}}\left(T_{\mathbf{a}+\mathbf{b}}\right) = \lambda_{\mathbf{u}}\left(T_{\mathbf{a}}\right) \lambda_{\mathbf{u}}\left(T_{\mathbf{b}}\right), \; \forall T_{\mathbf{a}}, T_{\mathbf{b}}, T_{\mathbf{a}+\mathbf{b}} \in \mathcal{A}.$$

We rewrite this condition as

$$\lambda_{\mathbf{u}}\left(T_{\mathbf{a}+\mathbf{b}}\right) = \lambda_{\mathbf{u}}\left(T_{\mathbf{a}}\right) \lambda_{\mathbf{u}}\left(T_{\mathbf{b}}\right), \; \forall T_{\mathbf{a}}, T_{\mathbf{b}} \in \mathcal{A}, \, [T_{\mathbf{a}}, T_{\mathbf{b}}] = 0. \quad (35)$$

We now evaluate the witness $\mathcal{W}_\rho^{\mathcal{B}(U)}(\mathbf{x})$ under the assumption of a non-contextual HVM. Assuming the system is in the state $\mathbf{u} \in \mathcal{S}$ of the HVM, and using the property Eq. (35), the witness of Eq. (33) becomes

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

In transitioning from the first to the second line above, we have used the property that $U = \text{span}(\{\mathbf{a}(i)\})$ is isotropic, such that Eq. (35) can be applied.

As a result of the above inequality, for any probability distribution q_ρ over $\mathcal{S}$, the prediction of any non-contextual HVM is

$$\mathcal{W}^{\mathcal{B}(U)} \geq 0,$$

for all isotropic subspaces $U \subset \mathbb{Z}_2^{2n}$. The negation of this statement proves the claim. □

We now relate the witnesses $\mathcal{W}$ to the rebit Wigner function.

Lemma 15 *Consider an isotropic subspace $U \subset \mathbb{Z}_2^{2n}$ with basis $\mathcal{B}(U) = \{\boldsymbol{a}(1), \boldsymbol{a}(2), .., \boldsymbol{a}(m)\}$, and a set $\tilde{\mathcal{B}} = \{\boldsymbol{b}(1), \boldsymbol{b}(2), .., \boldsymbol{b}(m)\}$ such that $[\boldsymbol{a}(i), \boldsymbol{b}(j)] = \delta_{ij}$ for all $i, j = 1, .., m$. For every $\eta(\boldsymbol{x}) = \sum_i x_i \boldsymbol{a}(i) \in U$, denote by $\overline{\eta}(\boldsymbol{x})$ the vector $\overline{\eta}(\boldsymbol{x}) = \sum_i x_i \boldsymbol{b}(i) \in \mathbb{Z}_2^{2n}$. Then,*

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

Proof of Lemma 15. We may rewrite the witness function $\mathcal{W}$ defined in Eq. (33) in terms of η, $\overline{\eta}$ as

$$\mathcal{W}_\rho^{\mathcal{B}(U)}(\eta) = \left\langle T_{\overline{\eta}} \left(\sum_{\mathbf{u} \in U} T_{\mathbf{u}} \right) T_{\overline{\eta}}^\dagger \right\rangle_\rho . \quad (36)$$

We may further rewrite this expression as

$$\begin{aligned}
\mathcal{W}_\rho^{\mathcal{B}(U)}(\eta) &= \frac{2^m}{2^{2n}} \left\langle T_{\overline{\eta}} \left[\sum_{\mathbf{v} \in U^\perp} T_{\mathbf{v}} \left(\sum_{\mathbf{u} \in \mathbb{Z}_2^{2n}} T_{\mathbf{u}} \right) T_{\mathbf{v}}^\dagger \right] T_{\overline{\eta}}^\dagger \right\rangle_\rho \\
&= \frac{2^m}{2^n} \left\langle T_{\overline{\eta}} \left[\sum_{\mathbf{v} \in U^\perp} T_{\mathbf{v}} A_0 T_{\mathbf{v}}^\dagger \right] T_{\overline{\eta}}^\dagger \right\rangle_\rho \\
&= 2^m \sum_{\mathbf{v} \in U^\perp} W_\rho(\mathbf{v} + \overline{\eta}),
\end{aligned}$$

which demonstrates the claimed relation. In transitioning from the second to the third line above, we have used the fact that ρ is real, and thus $\mathrm{Tr}\, T_{\mathbf{a}}\rho = 0$, for all $\mathbf{a}$ with $(\mathbf{a}_X, \mathbf{a}_Z) \mod 2 = 1$. □

Proof of Theorem 4. The combined conclusion of Lemmas 14 and 15 is that the n-rebit setting $(\rho, \mathcal{M})$ is contextual if there exists an isotropic subspace $U \subset \mathbb{Z}_2^{2n}$ with orthogonal complement $U^\perp$ and a vector $\nu \in \mathbb{Z}_2^{2n}$ such that

$$\sum_{\mathbf{v}\in U^\perp} W_\rho(\mathbf{v}+\nu) < 0. \tag{37}$$

We can further simplify this condition. Suppose that $\sum_{\mathbf{v}\in U^\perp} W_\rho(\mathbf{v}+\nu) \geq 0$ holds for all $\nu \in \mathbb{Z}_2^{2n}$ when U is maximally isotropic in $\mathbb{Z}_2^{2n}$. Then the same holds for all isotropic subspaces of $\mathbb{Z}_2^{2n}$. To verify this claim, consider a maximally isotropic space U and an isotropic subspace $\tilde{U}$ of U. Then, there exists a space $\overline{U} \subset \mathbb{Z}_2^{2n}$ such that $\tilde{U}^\perp = U^\perp \oplus \overline{U}$. Hence,

$$\sum_{\mathbf{v}\in\tilde{U}^\perp} W_\rho(\mathbf{v}+\nu) = \sum_{\mathbf{v}'\in\overline{U}} \left(\sum_{\mathbf{v}\in U^\perp} W_\rho(\mathbf{v}'+\mathbf{v}+\nu) \right).$$

If every term in brackets on the rhs is ≥ 0, so is the lhs. Since every isotropic $\tilde{U}$ can be embedded in a maximally isotropic U, the above claim follows. That is, we may restrict the condition Eq. (37) to maximally isotropic subspaces U. In those cases, $U^\perp = U$, which yields the condition stated in Theorem 4. □

D. When is the sufficient condition tight?

Below we demonstrate that Wigner function negativity is necessary and sufficient for contextuality on all states that are diagonal in a real stabilizer basis.

Fig. 1 shows two extremal situations in the application of Theorems 3, 4. An example for where the necessary and sufficient conditions for contextuality match, i.e. where Wigner function negativity is both necessary and sufficient for contextuality, is the family of two-rebit states

$$\rho(a,b) = \frac{(I + a\,XZ)(I + b\,ZX)}{4}. \tag{38}$$

In this case, the conditions of Theorems 3 and 4 for contextuality both read

$$1 + \alpha\, a + \beta\, b - \alpha\beta\, ab < 0,$$

for all combinations of $\alpha, \beta = \pm 1$. The corresponding phase diagram is depicted in Fig. 1a. The physical states fill the square with $|a|, |b| \leq 1$. The corners of that square represent the joint eigenstates of the Pauli operators XZ and ZX, and they sit deep in the contextual phase. This

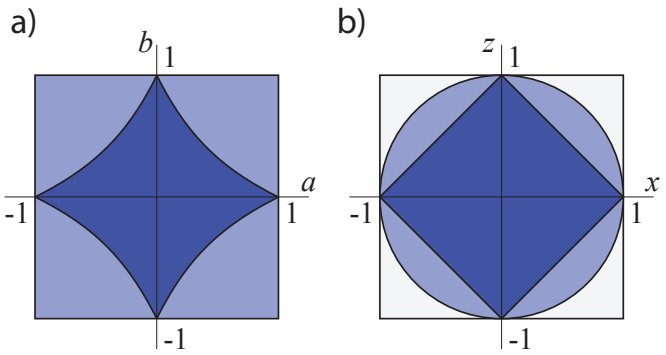


FIG. 1: Phase diagram for the families of states Eq. (38) and (39). Medium shade: the physical states, dark shade: the states classified as non-contextual by Theorem 3. (a) $\rho(a,b)$ with $a, b \in \mathbb{R}$. (b) $\tilde{\rho}(x,y)$ with $x, y \in \mathbb{R}$. The states classified as contextual by Theorem 4 lie outside the square of $|x|, |z| \leq 1$, and are thus not physical.

fits with our earlier observation that the commuting observables XZ and ZX cannot be simultaneously measured in rebit QCSI. Hence their joint eigenstates cannot be prepared by the restricted gates.

The opposite extreme is represented by the general one-rebit state

$$\tilde{\rho}(x,z) = \frac{I + x\,X + z\,Z}{2}. \tag{39}$$

The corresponding phase diagram is depicted in Fig. 1b. The set of physical states is constrained by $x^2 + z^2 \leq 1$. By Theorem 3, $\tilde{\rho}(x,z)$ is non-contextual if $|x| + |z| \leq 1$, and, by Theorem 4, contextual if $|x| > 1 \vee |z| > 1$. We thus find that not a single physical one-rebit state can be classified as guaranteed contextual by Theorem 4.

But this is not a failure of Theorem 4 to get traction. For single qubits, non-contextual HVMs can be constructed [2], [14], and they imply non-contextual HVMs for single rebits as a special case. The states $\tilde{\rho}(x,y)$ with $x^2 + z^2 \leq 1$ and $|x| + |z| > 1$ are thus negatively represented but non-contextual. This example settles to the negative the question of whether Wigner function negativity and contextuality are the same for rebits.

The former example generalizes as follows.

Lemma 16 *Be ρ an n-rebit state diagonal in a stabilizer eigenbasis. Then, ρ is contextual if and only if $W_\rho < 0$.*

Proof of Lemma 16. Denote by $S_U = \{T_{\mathbf{v}} | \, \mathbf{v} \in U\} \subset \mathcal{A}$ the stabilizer in whose joint eigenbasis the state ρ is diagonal; i.e., the corresponding maximal isotropic subspace U is such that $T_{\mathbf{v}} \rho T_{\mathbf{v}}^\dagger = \rho$ for all $\mathbf{v} \in U$.

Then, by covariance of the Wigner function under translations,

$$W_\rho(\nu) = W_{T_{\mathbf{v}}\rho T_{\mathbf{v}}^\dagger}(\nu) = W_\rho(\mathbf{v}+\nu), \quad \forall \mathbf{v} \in U.$$

In this case, the expression on the lhs of the condition in Theorem 4 simplifies to

$$\sum_{\mathbf{v}\in U} W_\rho(\mathbf{v}+\nu) = 2^n W_\rho(\nu).$$

And thus, Theorem 4 itself simplifies to the statement that if $W_\rho(\nu) < 0$ for some $\nu \in \mathbb{Z}_2^{2n}$ then ρ is contextual. This combined with Theorem 3 proves the claim. □

To summarize, unlike for qudits in odd prime dimension, for rebits contextuality and Wigner function negativity are in general not the same. However, they coincide on all states that are diagonal in a real stabilizer basis.

VI. CONTEXTUALITY AND NEGATIVITY IN QUANTUM COMPUTATION

A. Resources

We are now prepared to establish contextuality and Wigner function negativity as necessary resources for universality of QCSI on rebits.

Theorem 5 *In quantum computing via state injection on rebits, contextuality of the initial state is necessary for computational universality.*

Furthermore,

Corollary 1 *In quantum computing via state injection on rebits, Wigner function negativity of the initial state is necessary for computational universality.*

Corollary 1 is the combination of Theorems 3 and 5.

In preparation for the proof of Theorem 5, we note that the witness functions $\mathcal{W}^{\mathcal{B}(U)}$ transform covariantly under CSS-ness preserving unitaries, similar to the Wigner function. Namely, every CSS-ness preserving unitary g can be written as $g = T_{\mathbf{a}} g_F$, where $g_F T_{\mathbf{b}} g_F^\dagger = T_{F\mathbf{b}}$ for all $T_{\mathbf{b}} \in \mathcal{A}$, and $[F\mathbf{b}, F\mathbf{c}] = [\mathbf{b}, \mathbf{c}]$ for all $\mathbf{b}, \mathbf{c} \in V$. Then, using the form Eq. (36) of the contextuality witnesses,

$$\mathcal{W}_\rho^{\mathcal{B}(U)}(\eta) = \mathcal{W}_{g^{-1}\rho g}^{F^{-1}\mathcal{B}(U)}(\eta + \overline{\mathbf{a}}). \tag{40}$$

On the r.h.s., $F^{-1}\mathcal{B}(U)$ is again the basis of an isotropic subspace, since F, F^{-1} preserve the commutation relations. In result, for two density matrices ρ and ρ' related by a CSS-ness preserving Clifford unitary, if there is a witness $\mathcal{W}$ that evaluates to x on ρ then there is a witness $\mathcal{W}'$ that evaluates to the same value x on ρ'.

Proof of Theorem 5. If the discussed computational scheme is universal, it must in particular be capable of executing an encoded Hadamard gate. In turn, the encoded Hadamard gate can be used to convert an encoded Bell state (a CSS state) into an encoded graph state $|\overline{G}_2\rangle$ with stabilizer $\langle \overline{XZ}, \overline{ZX}\rangle$. We are using the encoding of n qubits into $n+1$ rebits by Rudolph and Grover stated in Eq. (5),

$$\sum_k r_k e^{i\phi_k}|k\rangle \longrightarrow \sum_k r_k |k\rangle \otimes (\cos\phi_k |0\rangle_A + \sin\phi_k |1\rangle_A)\,.$$

For this encoding, for all qubits $i = 1..n$ we have

$$\overline{X}_i = X_i,\ \overline{Z}_i = Z_i,\ \overline{Y}_i = Y_i \otimes Y_A. \tag{41}$$

With $\overline{i\,I} = i\,Y_A$, this is compatible with the Pauli multiplication table $\overline{Y} = \overline{iZX} = \overline{i\,I}\,\overline{X}\,\overline{Z}$.

All observables in $\mathcal{A}$ have an even number of Y's, and therefore

$$\overline{T} = T,\ \ \forall T \in \mathcal{A}. \tag{42}$$

For the state $|\overline{G}_2\rangle$, the contextuality witness based on the operators $T_{\mathbf{a}} = XZ$ and $T_{\mathbf{b}} = ZX$ is negative, namely

$$\mathcal{W}^{\{\mathbf{a},\mathbf{b}\}}_{|\overline{G}_2\rangle}((1,1)) = \langle \overline{G}_2 | I - X_1 Z_2 - Z_1 X_2 - Y_1 Y_2 | \overline{G}_2\rangle = -2. \tag{43}$$

For two-dimensional isotropic subspaces U, -2 is the most negative value that a witness $\mathcal{W}^{\mathcal{B}(U)}$ can yield. The final state $|\overline{G}_2\rangle$ thus reveals contextuality maximally.

We now prove that also the initial state fed into the computation must reveal contextuality maximally. The proof is by induction. We consider the circuit which created the state $|\overline{G}_2\rangle$, and assume the gates are performed sequentially, one in each step m. We show that if the state $\rho(m)$ after step m reveals contextuality maximally then so does the state $\rho(m-1)$ after step $m-1$. That is, if there exists a witness $\mathcal{W}$ such that $\mathcal{W}_{\rho(m)}(\eta) = -2$ then there exists another witness $\mathcal{W}'$ such that $\mathcal{W}'_{\rho(m-1)}(\eta') = -2$.

For the gates in the circuit, we distinguish between unitaries and projective measurements. *Case i:* the gate in step m is a unitary. Then, by construction of the computational scheme, the gate is a CSS-ness preserving Clifford unitary. Then, the claim of the induction step follows from the covariance of the witness functions, Eq. (40).

Case ii: The gate in step m is a projective measurement. Then, by construction of the computational scheme, it is the measurement of an observable $T_{\mathbf{c}} \in \mathcal{O}$. Let the witness for the state $\rho(m)$ be constructed from the isotropic subspace spanned by $\{\mathbf{a}(m), \mathbf{b}(m)\}$, such that $\mathcal{W}^{\{\mathbf{a}(m),\mathbf{b}(m)\}}_{\rho(m)}(\eta) = -2$, for some η. There are two sub-cases to consider.

Case ii/a: $T_{\mathbf{c}}$ commutes with both $T_{\mathbf{a}(m)}$ and $T_{\mathbf{b}(m)}$. Then the value of the witness $\mathcal{W}^{\{\mathbf{a}(m),\mathbf{b}(m)\}}(\eta)$ is the same for $\rho(m)$ and $\rho(m-1)$, hence $\rho(m-1)$ reveals contextuality maximally.

Case ii/b: $T_{\mathbf{c}}$ does not commute with both $T_{\mathbf{a}(m)}$ and $T_{\mathbf{b}(m)}$. Then, $T_{\mathbf{c}}$ anti-commutes with two of the three operators $T_{\mathbf{a}(m)}$, $T_{\mathbf{b}(m)}$, $T_{\mathbf{a}(m)+\mathbf{b}(m)}$, and commutes with the third. Wlog assume $T_{\mathbf{c}}$ anti-commutes with $T_{\mathbf{a}(m)}$ and $T_{\mathbf{b}(m)}$, and commutes with $T_{\mathbf{a}(m)+\mathbf{b}(m)}$.

Then, $\langle T_{\mathbf{a}(m)}\rangle_{\rho(m)} = \langle T_{\mathbf{b}(m)}\rangle_{\rho(m)} = 0$. The witness for the state $\rho(m)$ therefore reduces to $\mathcal{W}_{\rho(m)}^{\{\mathbf{a}(m),\mathbf{b}(m)\}}(\eta) = \langle I \pm T_{\mathbf{a}(m)+\mathbf{b}(m)}\rangle_{\rho(m)} \geq 0$. This contradicts the induction assumption. Hence, case ii/b cannot occur.

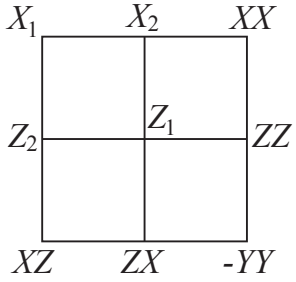
Thus, irrespective of whether a given step in the circuit is a unitary transformation or a projective measurement, if the state after completing the step witnesses contextuality with the maximum negative value, so does the state before the step. By induction, the state before the first gate, i.e. the injected state, witnesses contextuality. □

B. Coping with Mermin's square

Mermin's square [14] provides a beautifully simple proof of the Kochen-Specker theorem [3] in dimension four and higher, but for the programme of establishing contextuality of magic states as a quantum computational resource it poses a problem. Namely, the square can be converted into a contextuality witness of CSW type [17] for which *all* two-qubit states come out contextual [4]. But if contextuality is generic, then it is not a resource.

In more general terms, Mermin's square exhibits the phenomenon of state-independent contextuality. It represents an obstacle to viewing contextuality as a resource possessed by some quantum states but not others.

When restricting to Pauli observables, state-independent contextuality only occurs in Hilbert spaces of even dimension [23], and therefore was not an issue in [4]. However, in the present situation, the Hilbert space dimension is even, and furthermore, by a simple local rotation, Mermin's square can be embedded into real quantum mechanics,



(44)

Since, as in qudit QCSI, also in rebit QCSI contextuality is attributed to quantum states, state-independent contextuality seems likely to cause difficulty. Yet, in Theorem 5 we established contextuality of magic states as a necessary resource. We thus have to explain why Mermin's square, and more generally the phenomenon of state-independent contextuality, did in fact not void the contextuality-as-resource viewpoint.

To do so, we revisit the results established in Section V. First, by Theorem 3, states with non-negative Wigner function are non-contextual. Hence contextuality is not generic, as required for a resource.

Next, we consider the rotated Mermin square, Eq. (44). For all columns and all rows except the bottom one, the belonging observables pairwise commute and generate a stabilizer group of CSS type. They can therefore be simultaneously measured in rebit QCSI. The measurement outcomes ± 1 must multiply to $+1$ in each of these contexts, which is implied by the identities among the observables, $X_1 \cdot X_2 \cdot X_1X_2 = +I$ etc.

For the bottom row, the belonging observables XZ, ZX and $-YY$ still commute and thus generate a stabilizer group, but this group is not of CSS type. As discussed at the beginning of Section V, these observables cannot be simultaneously measured in rebit QCSI. Therefore, the pre-determined measurement outcomes $\lambda(XZ)$, $\lambda(ZX)$, $\lambda(-YY)$ *need not* satisfy the constraint $\lambda(XZ)\lambda(ZX)\lambda(-YY) = -1$ implied by the operator relation $XZ \cdot ZX \cdot (-YY) = -I$. Therefore, $\lambda(\cdot) = +1$ for all observables in the rotated Mermin square is a consistent value assignment w.r.t. rebit QCSI. The algebraic contradiction vanishes because we have effectively removed the bottom row from the diagram.

This situation is handled by our definitions as follows: By Criterion 1, $\{XZ, ZX, -YY\} \notin \mathcal{M}$, c.f. Lemma 13. Therefore, $\lambda_{\mathbf{u}}(XZ)\lambda_{\mathbf{u}}(ZX)\lambda_{\mathbf{u}}(-YY) = -1$ is not required by Definition 1 of a non-contextual HVM (c.f. condition (ii)).

Generalizing the above observation, the phenomenon of state-independent contextuality does not come into play for the present setting of QCSI, even if it does exist for systems of rebits. The reason is the restriction of the physically measurements to observables in $\mathcal{O}$, the set of pure-X and pure-Z Pauli operators.

Lemma 17 *Consider a system of n rebits where the measurable observables are restricted to the set $\mathcal{O}$. Then, the set $\mathcal{S}$ of consistent value assignments of a non-contextual HVM, $\lambda_{\boldsymbol{u}} : \mathcal{A} \to \{\pm 1\}$, $\forall \boldsymbol{u} \in \mathcal{S}$, is non-empty.*

Thus, there is no state-independent contextuality in rebit QCSI. Contextuality may persist at the levels of possibility and probability [24].

Proof of Lemma 17. The value assignments $\lambda_{\mathbf{u}} : \mathcal{A} \longrightarrow \{\pm 1\}$, $\mathbf{u} \in \mathbb{Z}_2^{2n}$, of Eq. (34) all satisfy the consistency condition Eq. (35), $\lambda_{\mathbf{u}}(T_{\mathbf{a}+\mathbf{b}}) = \lambda_{\mathbf{u}}(T_{\mathbf{a}})\,\lambda_{\mathbf{u}}(T_{\mathbf{b}})$, for all $T_{\mathbf{a}}, T_{\mathbf{b}} \in \mathcal{A}$ such that $[T_{\mathbf{a}}, T_{\mathbf{b}}] = 0$. By Lemma 13, for all $M \in \mathcal{M}$ and all $T_{\mathbf{a}}, T_{\mathbf{b}} \in M$ ($[T_{\mathbf{a}}, T_{\mathbf{b}}] = 0$), it holds that $T_{\mathbf{a}+\mathbf{b}} = T_{\mathbf{a}}T_{\mathbf{b}}$, for all $\mathbf{u} \in \mathbb{Z}_2^{2n}$. The value assignments $\lambda_{\mathbf{u}}(\cdot)$ are thus consistent with the operator constraints. Hence, $\mathbb{Z}_2^{2n} \subseteq \mathcal{S}$, and $\mathcal{S} \neq \emptyset$. □

VII. CONCLUSION

We have established that contextuality and Wigner function negativity are necessary for the universality of the discussed scheme of quantum computation by state injection on rebits. To this end, we have introduced an n-rebit Wigner function which provides necessary and sufficient conditions for contextuality with respect to CSS-

ness preserving stabilizer measurements. This Wigner function comes with a Hudson's theorem that singles out the "cheap" CSS-states and a covariance property that singles out the restricted gate set on which the discussed computational scheme is based.

For rebits, as was the case for qudits in odd prime dimension, all states with non-negative Wigner function are non-contextual. Therefore, for both the qudit and the rebit case, the possibility that contextuality is *sufficient* for universality remains open. For the case of qubits as discussed in [4], this is not the case due to the existence of contextual states with non-negative Wigner function.

In summary, we have extended all the essential properties that held for QCSI in the case of qudits of odd prime dimension [4], [6] to rebits, with the sole exception that for the present rebit scheme, Wigner function negativity does not imply contextuality. A schematic phase diagram w.r.t. Wigner negativity and contextuality for the three cases of qudits, qubits and rebits is shown in Fig. 2.

To widen the scope of the discussion, we note that contextuality has also been established as a resource for measurement-based quantum computation (MBQC) [25], [26], [27]. Namely, in MBQC contextuality is necessary for the ability to compute non-linear Boolean functions. One may thus want to compare the roles played by contextuality in QCSI and MBQC. But there was an obstacle: The MBQC result has to date only been established for the case of qubits, where most of the existing results [4] do not apply. The present paper removes this mismatch, and thus prepares the ground for a comparison between the two computational schemes.

We conclude with three open questions.

- In QCSI contextuality is about speedup, as one might expect for a scheme of quantum computation. But in MBQC it is about computability. What is the reason for this dichotomy?
- Due to the formulation in terms of a Wigner function, covariance plays an important role for QCSI. Is covariance also a useful concept in the discussion of MBQC?
- We noted that the restricted gate set in the present scheme of rebit QCSI is precisely the gate set that can be implemented by defect braiding and fusion with surface codes [19]. There is more complicated lattice surgery by which, in addition, the Hadamard gate can be realized [28], [29]. In this way, the full real subgroup of the qubit stabilizer group becomes available as the restricted gate set, and the real stabilizer states are the "cheap" / non-magic states. Is there a Wigner function with matching Hudson's theorem and covariance property?

Acknowledgments: RR thanks Rob Spekkens for discussion. ND is supported by the Lockheed Martin Corporation. PAG is supported by FRQNT (Quebec), and RR is supported by NSERC, Cifar and IARPA.

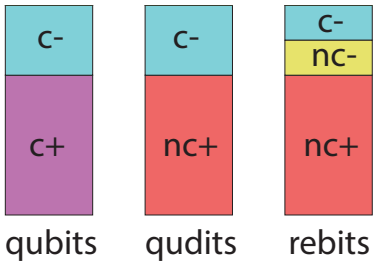


FIG. 2: Schematic phase diagram for qudit, qubit and rebit states. We distinguish 4 phases: nc+ (non-negative Wigner function, non-contextual), nc− (negative Wigner function, non-contextual), c+ (non-negative Wigner function, contextual), c− (negative Wigner function, contextual). Both qudits and rebits have a phase nc+ and no phase c+. To the contrary, the qubit has a phase c+ but no phase nc+.

Appendix A: Fourier transform on the group $\mathbb{Z}_2^n$

The Fourier transform of a function $f : M \to \mathbb{R}$ defined on a linear subspace M of $\mathbb{Z}_2^n$ is the function $\mathcal{F}f$ or $\hat{f}$ defined by

$$\mathcal{F}f(\mathbf{u}) = \frac{1}{\sqrt{|M|}} \sum_{\mathbf{x}\in M} (-1)^{(\mathbf{u},\mathbf{x})} f(\mathbf{x}).$$

Lemma 18 *The Fourier transform $\mathcal{F}$ is involutive,* i.e. *$\mathcal{F} \circ \mathcal{F} = Id$, or equivalently $\mathcal{F}$ is its own inverse.*

Proof of Lemma 18. Let us determine the image of a function $f : M \to \mathbb{R}$ by $\mathcal{F} \circ \mathcal{F}$.

$$\begin{aligned}
((\mathcal{F}\circ\mathcal{F})(f))(\mathbf{u}) &= \mathcal{F}(\mathcal{F}(f))(\mathbf{u}) \\
&= \frac{1}{\sqrt{|M|}} \sum_{\mathbf{x}\in M} (-1)^{(\mathbf{u},\mathbf{x})} \mathcal{F}f(\mathbf{x}) \\
&= \frac{1}{|M|} \sum_{\mathbf{x}\in M} \sum_{\mathbf{y}\in M} (-1)^{(\mathbf{u},\mathbf{x})} (-1)^{(\mathbf{x},\mathbf{y})} f(\mathbf{y}) \\
&= \frac{1}{|M|} \sum_{\mathbf{y}\in M} \left(\sum_{\mathbf{x}\in M} (-1)^{(\mathbf{x},\mathbf{u}+\mathbf{y})} \right) f(\mathbf{y}) \\
&= \sum_{\mathbf{y}\in M} \delta_{\mathbf{u},\mathbf{y}} f(\mathbf{y}) \\
&= f(\mathbf{u}),
\end{aligned}$$

which demonstrates the claim. □

Appendix B: Properties of the rebit Wigner function

Lemma 19 *The set of Pauli operators $\mathcal{A}$ is an orthonormal basis of the space $S_{2^n}(\mathbb{R})$ of symmetric matrices of size 2^n endowed with the inner product $(A, B) = \frac{1}{2^n} \mathrm{Tr}(A^T B)$.*

Proof of Lemma 19. Denote by $E_{i,j}$ the matrix with entry 0 everywhere expect at the intersection of the i-th row and j-th column where it is 1. The space $S_N(\mathbb{R})$ is generated by the matrices $(E_{i,j}+E_{j,i})$ with $1 \leq i < j \leq N$ and $E_{i,i}$ with $1 \leq i \leq N$. Moreover, we can easily check that these $N(N+1)/2$ matrices are independent. Thus the dimension of $S_N(\mathbb{R})$ is $N(N+1)/2$. In our case $N = 2^n$ and $\dim S_{2^n}(\mathbb{R}) = 2^{2n-1} + 2^{n-1}$.

The set $\mathcal{A}$ contains $2^{2n-1} + 2^{n-1}$ symmetric matrices. These matrices are pairwise orthogonal, *i.e.* they satisfy $\frac{1}{2^n}\operatorname{Tr}(T_u T_v) = 0$ when $u \neq v$, thus they are linearly independent. This proves that they form a basis of the space of symmetric matrices. The orthonormality is a consequence of the orthonormality of the Paulis. □

We show that from a given Wigner function we can obtain the corresponding real density operator, proving that the Wigner function is informationally complete.

Lemma 20 *Let ρ be a real density operator and let W_ρ be its Wigner function. Then ρ satisfies*

$$\rho = \sum_{\mathbf{u}\in\mathbb{Z}_2^{2n}} W_\rho(\mathbf{u}) A_\mathbf{u}.$$

Proof of Lemma 20. We expand the r.h.s. of the above equation by inserting the definition Eq. (11) of W_ρ and Eq. (13) for $A_\mathbf{u}$, and obtain

$$\sum_{\mathbf{u}\in\mathbb{Z}_2^{2n}} W_\rho(\mathbf{u}) A_\mathbf{u} = \frac{1}{2^{3n}} \sum_{\substack{\mathbf{u}\in\mathbb{Z}_2^{2n} \\ \mathbf{v},\mathbf{w}\in V_\mathcal{A}}} (-1)^{[\mathbf{u},\mathbf{v}+\mathbf{w}]} \operatorname{Tr}(T_\mathbf{v}\rho) T_\mathbf{w}$$

Now note that the sum $\sum_{\mathbf{u}\in\mathbb{Z}_2^{2n}} (-1)^{[\mathbf{u},\mathbf{v}+\mathbf{w}]}$ is $2^{2n}\delta_{v,w}$, which is a standard property of characters. Hence,

$$\sum_{\mathbf{u}\in\mathbb{Z}_2^{2n}} W_\rho(\mathbf{u}) A_\mathbf{u} = \frac{1}{2^n} \sum_{\mathbf{w}\in V_\mathcal{A}} \operatorname{Tr}(T_\mathbf{w}\rho) T_\mathbf{w}.$$

From Lemma 19, this sum is the decomposition of ρ in the orthonormal basis $\mathcal{A}$. This proves that we recover the state ρ. □

Proof of Property 4 (Section IV A). With the definition Eq. (12),

$$\begin{aligned} A_0 &= \tfrac{1}{2^n} \textstyle\sum_{\mathbf{u}\in V \,|\, (\mathbf{u}_X,\mathbf{u}_Z) \bmod 2 = 0} T_\mathbf{u} \\ &= \tfrac{1}{2^{n+1}} \big(\textstyle\prod_{i=1}^n (I+Z_i) \prod_{j=1}^n (1+X_j) + \\ &\quad + \textstyle\prod_{i=1}^n (I+X_i) \prod_{j=1}^n (1+Z_j) \big) \\ &= 2^{n-1} \left(|0_n\rangle\langle 0_n| |+_n\rangle\langle +_n| + |+_n\rangle\langle +_n| |0_n\rangle\langle 0_n| \right), \end{aligned}$$

and thus

$$A_0 = 2^{\frac{n}{2}-1} \left(|0_n\rangle\langle +_n| + |+_n\rangle\langle 0_n| \right). \tag{B1}$$

Further using the properties that ρ is Hermitian and real,

$$W_\rho(\mathbf{v}) = \frac{1}{\sqrt{2^n}} \langle 0_n | T_\mathbf{v}^\dagger \rho T_\mathbf{v} | +_n \rangle.$$

Now, we consider the case where ρ_{AB} factorizes, $\rho_{AB} = \sigma_A \otimes \tau_B$. We may write any phase space point $\mathbf{v}$ as $\mathbf{v} = \mathbf{v}_A + \mathbf{v}_B$, where $\mathbf{v}_A$ ($\mathbf{v}_B$) acts non-trivially only on system A (B). Then, $T_\mathbf{v} = \pm T_{\mathbf{v}_A} T_{\mathbf{v}_B}$ and

$$\begin{aligned} W_{\sigma\otimes\tau}(\mathbf{v}) &= \frac{\langle 0_{n_A}, 0_{n_B} | T_{\mathbf{v}_A} T_{\mathbf{v}_B} \sigma \otimes \tau \, T_{\mathbf{v}_A} T_{\mathbf{v}_B} | +_{n_A}, +_{n_B} \rangle}{\sqrt{2}^{n_A+n_B}} \\ &= W_\sigma(\mathbf{v}_A) W_\tau(\mathbf{v}_B). \end{aligned}$$

Proof of Lemma 4. By inserting into the Wigner function the expansion $|\psi\rangle\langle\psi| = \sum_{\mathbf{x},\mathbf{y}\in\mathbb{Z}_2^\mathbf{n}} \psi(\mathbf{x})\psi(\mathbf{y}) |\mathbf{x}\rangle\langle\mathbf{y}|$, we obtain

$$W_\psi(\mathbf{u}) = \frac{1}{2^n} \sum_{\mathbf{x},\mathbf{y}\in\mathbb{Z}_2^\mathbf{n}} \psi(\mathbf{x})\psi(\mathbf{y}) \operatorname{Tr}(T_\mathbf{u} A_0 T_\mathbf{u}^\dagger |\mathbf{x}\rangle\langle\mathbf{y}|).$$

We use the expression Eq. (B1) for A_0, and verify by direct calculation that

$$\operatorname{Tr}(T_\mathbf{u} |0_n\rangle\langle +_n| T_\mathbf{u}^\dagger |\mathbf{x}\rangle\langle\mathbf{y}|) = \frac{(-1)^{(\mathbf{x}+\mathbf{u}_X,\mathbf{u}_Z)}}{2^{n/2}} \delta_{\mathbf{u}_X,\mathbf{y}}. \tag{B2}$$

Inserting Eq.(B2) in the above $W_\psi(\mathbf{u})$ gives

$$\begin{aligned} W_\psi(\mathbf{u}) &= \frac{1}{2^{n+1}} \left(\sum_{\mathbf{x}\in\mathbb{Z}_2^\mathbf{n}} \psi(\mathbf{x})\psi(\mathbf{u}_X) (-1)^{(\mathbf{x}+\mathbf{u}_X,\mathbf{u}_Z)} \right) \\ &\quad + \frac{1}{2^{n+1}} \left(\sum_{\mathbf{y}\in\mathbb{Z}_2^\mathbf{n}} \psi(\mathbf{u}_X)\psi(\mathbf{y}) (-1)^{(\mathbf{y}+\mathbf{u}_X,\mathbf{u}_Z)} \right) \\ &= \frac{1}{2^n} \sum_{\mathbf{x}\in\mathbb{Z}_2^\mathbf{n}} (-1)^{(\mathbf{x},\mathbf{u}_Z)} \psi(\mathbf{u}_X)\psi(\mathbf{u}_X+\mathbf{x}). \end{aligned}$$

as stated in Lemma 4. □

Appendix C: CSS-ness preserving Clifford gates

Here, we prove Lemma 1 from Section II D, and Lemmas 9, 10 from Section IV C, about the structure of the CSS-ness preserving subgroup of the Clifford group.

Proof of Lemma 9. First, we omit the phase ± 1 of $T_\mathbf{a}$ and focus on the effect of the conjugation on the vector $\mathbf{a}$. Let $g \in G_{CSS}$ and let φ_g be the automorphism of P_n defined by conjugation by g:

$$\begin{aligned} \varphi_g : P_n &\longrightarrow P_n \\ Q &\longmapsto g Q g^\dagger \end{aligned} \tag{C1}$$

This morphism of group P_n induces a morphism of its quotient $P_n/\{\pm I\}$, which is isomorphic to $\mathbb{Z}_2^{2n}$, that is φ_g induces a matrix $F \in M_{2n}(\mathbb{Z}_2)$ such that

$$g T_\mathbf{a} g^\dagger = \lambda(\mathbf{a}) T_{F\mathbf{a}},$$

where $\lambda(\mathbf{a}) \in \{\pm 1\}$. Since φ_g is an automorphism, $F \in \mathrm{GL}_{2n}(\mathbb{Z}_2)$. Moreover, the conjugation conserves

the commutation relation and we know that $T_{\mathbf{a}}$ and $T_{\mathbf{b}}$ commute if and only if $[\mathbf{a},\mathbf{b}] = 0$. This proves that $F \in \mathrm{Sp}_{2n}(\mathbb{Z}_2)$. □

Proof of Lemma 10. Consider a pair $g, g' \in G_{CSS}$ and denote by $\mathcal{F}(g) = A(F,\mathbf{t})$ and $\mathcal{F}(g') = A(F',\mathbf{t}')$ their images. The value of $\mathcal{F}(gg')$ is defined by the conjugation by gg'. We obtain

$$\begin{aligned} gg'T_{\mathbf{a}}(gg')^\dagger &= g\left(g'T_{\mathbf{a}}g'^\dagger\right)g^\dagger \\ &= g\left((-1)^{[\mathbf{t}',F'\mathbf{a}]}T_{F'\mathbf{a}}\right)g^\dagger \\ &= (-1)^{[\mathbf{t}',F'\mathbf{a}]+[\mathbf{t},FF'\mathbf{a}]}T_{FF'\mathbf{a}} \\ &= (-1)^{[F\mathbf{t}'+\mathbf{t},FF'\mathbf{a}]}T_{FF'\mathbf{a}} \end{aligned}$$

Therein, we have used $[\mathbf{t}', F'\mathbf{a}] = [F\mathbf{t}', FF'\mathbf{a}]$. This gives $\mathcal{F}(gg') = A(FF', F\mathbf{t}' + \mathbf{t})$, which is indeed the composition of $\mathcal{F}(g) = A(F,\mathbf{t})$ and $\mathcal{F}(g') = A(F',\mathbf{t}')$. Hence, $\mathcal{F}(gg') = \mathcal{F}(g)\mathcal{F}(g')$, for all $g, g' \in G_{CSS}$. □

Proof of Lemma 1. Our fist goal is to describe G_{CSS} as the normalizer of the special Pauli operators $\mathcal{O}$.

Lemma 21 *The group G_{CSS} is the normalizer in $O_{2^n}(\mathbb{R})$ of the set $\mathcal{O} = \{Z(\mathbf{u}) \mid \mathbf{u} \in \mathbb{Z}_2^n\} \cup \{X(\mathbf{v}) \mid \mathbf{v} \in \mathbb{Z}_2^n\}$ of Pauli-observables which have only an X-part or only a Z-part.*

Proof of Lemma 21. If g belongs to the normalizer of $\mathcal{O}$, then it conserves CSS codes and CSS states.

In order to obtain the inverse implication, we will show that an operator g which preserves CSS states, stabilizes the set of all CSS groups by conjugation. Applying this argument to rank one groups $\langle X_i\rangle$ and $\langle Z_i\rangle$, we obtain the lemma. Thus, we want to prove that the image under conjugation by g of a CSS group is also a CSS group. This is true when S has rank n. We work by induction. Assume the result for every CSS group of rank r and let us prove that it is also true for a CSS group S of rank $r-1$. Let S' be the group $gSg^\dagger$ obtained after conjugation et let M be the subspace

$$M = \{\mathbf{a} \in \mathbb{Z}_2^{2n} \mid \pm T_{\mathbf{a}} \in S'\}.$$

We associate with M two subspaces

$$M_Z = \{\mathbf{u}_Z \in \mathbb{Z}_2^n \mid \exists\, (\mathbf{u}_Z, \mathbf{u}_X) \in M\}$$

and

$$M_X = \{\mathbf{u}_X \in \mathbb{Z}_2^n \mid \exists\, (\mathbf{u}_Z, \mathbf{u}_X) \in M\}.$$

Note that $M \subset M_Z \oplus M_Z$ and S' is a CSS codes if and only if we have equality $M = M_Z \oplus M_Z$ and in that case M_Z and M_X are two orthogonal subspaces. Assume that S' is not a CSS code then $M_Z \oplus M_X$ contains strictly M and has dimension

$$\dim M_Z \oplus M_X > \dim M = r - 1.$$

Now, choose two logical operators $\bar{X}$ and $\bar{Z}$ for the code S which anti-commute. The two CSS groups $\langle S, \bar{X}\rangle$ and $\langle S, \bar{Z}\rangle$ are sent onto CSS groups by conjugation. Denote by N and R respectively the corresponding subspaces of $\mathbb{Z}_2^{2n}$, defined as M. These two spaces can be decomposed as

$$N = N_Z \overset{\perp}{\oplus} N_X \text{ and } R = R_Z \overset{\perp}{\oplus} R_X.$$

These spaces both contain $M_Z \oplus M_X$ and have dimension r, hence we have $M_Z \oplus M_X = N = R$. To find a contradiction, consider the operators $g\bar{X}g^\dagger = \lambda(\mathbf{a})T_{\mathbf{a}}$ and $g\bar{Z}g^\dagger = \lambda(\mathbf{b})T_{\mathbf{b}}$. By construction, we have $\mathbf{a} \in N$ and $\mathbf{b} \in R$. Using the equality $N = R$, we can see that the two inner products $(\mathbf{a}_Z, \mathbf{b}_X)$ and $(\mathbf{b}_Z, \mathbf{a}_X)$ are 0, which implies that $T_{\mathbf{a}}$ and $T_{\mathbf{b}}$ commute. This is a contradiction since g preserves the commutation relation. Finally, we proved that S' is a CSS group. The set of all CSS group is preserved by conjugation by g. □

We now return to the subject of Lemma 9, and further characterize the matrices F appearing on the r.h.s. of Eq. (23). These matrices have one of the two following block structures.

$$F = \begin{pmatrix} F_Z & 0 \\ 0 & F_X \end{pmatrix} \quad \text{or} \quad F = \begin{pmatrix} 0 & F_X \\ F_Z & 0 \end{pmatrix} \tag{C2}$$

where $F_Z, F_X \in \mathrm{GL}_n(\mathbb{Z}_2)$ and $F_X = (F_Z^{-1})^t$. In what follows, we denote by F_{CSS} the set of symplectic matrices introduced in Eq. (C2). The result is that every CSS Clifford operator induces a pair $(F,\mathbf{x}) \in F_{CSS} \times \mathbb{Z}_2^{2n}$.

To demonstrate Eq. (C2), note that the conjugation φ_g of Eq. (C1) preserves the set of CSS operators $X(\mathbf{u})$ and $Z(\mathbf{v})$. Suppose an operator $X(\mathbf{u})$ is sent onto $X(\mathbf{u}')$ and that $X(\mathbf{v})$ is sent onto $Z(\mathbf{v}')$. Then, the image of $X(\mathbf{u}+\mathbf{v})$ is $X(\mathbf{u}')Z(\mathbf{v}')$ which is impossible. Therefore, φ_g has two possible structures, either it conserves both sets $\{X(\mathbf{u}) \mid \mathbf{u} \in \mathbb{Z}_2^n\}$ and $\{Z(\mathbf{u}) \mid \mathbf{u} \in \mathbb{Z}_2^n\}$, or it exchanges these two sets. This proves that the matrix F has one of the two following block structures.

$$F = \begin{pmatrix} F_Z & 0 \\ 0 & F_X \end{pmatrix} \quad \text{or} \quad F = \begin{pmatrix} 0 & F_X \\ F_Z & 0 \end{pmatrix}$$

where $F_Z, F_X \in \mathrm{GL}_n(\mathbb{Z}_2)$. Finally, $F_Z = (F_X^t)^{-1}$ is a consequence of the requirement that the F's must preserve the symplectic form.

The knowledge of the structure of the matrix F will now be useful to determine the phase $\lambda(\mathbf{a})$ of the operator $gT_{\mathbf{a}}g^\dagger = \lambda(\mathbf{a})T_{F\mathbf{a}}$. Since every character of $\mathbb{Z}_2^{2n}$ is of the form $\mathbf{a} \mapsto (-1)^{[\mathbf{x},\mathbf{a}]}$ for some vector $\mathbf{x}$ of $\mathbb{Z}_2^{2n}$, it suffices to show that λ is the restriction of such a character to the set $V_{\mathcal{A}}$. Denote by $(\mathbf{e}_i)_{i=1}^{2n}$ the canonical basis of the space $\mathbb{Z}_2^{2n}$ and denote by μ the character of $\mathbb{Z}_2^{2n}$ defined by $\mu(\mathbf{e}_i) = \lambda(\mathbf{e}_i)$. To prove that $\mu = \lambda$ on the set $V_{\mathcal{A}}$, it is enough to show that

- If $\mathbf{a}_X = \mathbf{b}_X = 0$ or if $\mathbf{a}_Z = \mathbf{b}_Z = 0$, then we have $\lambda(\mathbf{a}+\mathbf{b}) = \lambda(\mathbf{a})\lambda(\mathbf{b})$.

- If $\mathbf{a} = (\mathbf{a}_Z, \mathbf{a}_X) \in V_{\mathcal{A}}$, then we have $\lambda(\mathbf{a}) = \lambda((\mathbf{a}_Z, 0))\lambda((0, \mathbf{a}_X))$.

In what follows, we assume that F is block diagonal. The proof is similar in the anti-diagonal case. If $\mathbf{a}_X = \mathbf{b}_X = 0$, then we have $\varphi_g(T_{\mathbf{a}+\mathbf{b}}) = \lambda(\mathbf{a}+\mathbf{b})T_{F(\mathbf{a}+\mathbf{b})}$ which is also $\varphi_g(T_{\mathbf{a}}T_{\mathbf{b}}) = \lambda(\mathbf{a})\lambda(\mathbf{b})T_{(F_Z\mathbf{a}_Z,0)}T_{(F_Z\mathbf{b}_Z,0)} = \lambda(\mathbf{a})\lambda(\mathbf{b})T_{F(\mathbf{a}+\mathbf{b})}$. The equality $\lambda(\mathbf{a}+\mathbf{b}) = \lambda(\mathbf{a})\lambda(\mathbf{b})$ follows. The proof of the second implication is similar.

This implies that λ coincides with the character μ on the set $V_{\mathcal{A}}$, which means that $\lambda(\mathbf{a}) = (-1)^{[\mathbf{x},\mathbf{a}]}$ for some vector $\mathbf{x} \in \mathbb{Z}_2^{2n}$.

To illustrate the above with examples, we list the pairs $(F, \mathbf{x})$ for a few gates $g \in G_{CSS}$ of special interest.

- If $g = \otimes_i H_i$ then

$$F = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} \text{ and } \mathbf{x} = 0. \tag{C3}$$

- If $g = CNOT(i,j)$ then

$$F = \begin{pmatrix} I_n + E_{i,j} & 0 \\ 0 & I_n + E_{j,i} \end{pmatrix} \text{ and } \mathbf{x} = 0. \tag{C4}$$

  The matrix $E_{i,j}$ denote the $n \times n$ binary matrix whose only non-zero coefficient is in position (i,j)

- If $g = T_{\mathbf{u}}$ then

$$F = I_n \quad \text{and} \quad \mathbf{x} = \mathbf{u}. \tag{C5}$$

The fact that the pair $(F, \mathbf{x})$ associated with a Pauli operator $T_{\mathbf{u}}$ is $(0, \mathbf{u})$ is a direct consequence of the commutation relations between Pauli operators.

We now return to the subject of Lemma 10, and describe the image of the map $\mathcal{F}$. It holds that

$$\operatorname{Im}\mathcal{F} = \{A(F, \mathbf{t}) \mid F \in F_{CSS}, \mathbf{t} \in \mathbb{Z}_2^{2n}\}.$$

Recall that this application $\mathcal{F}$ is well defined by unicity in Lemma 9. The translation vector $\mathbf{t}$ and the vector $\mathbf{x}$ of Lemma 9 are related by the equation $\mathbf{t} = F\mathbf{x}$.

In order to determine the image of $\mathcal{F}$, note that we already know some elements of $\operatorname{Im}\mathcal{F}$. Indeed, from Eq. (C4), all the transformations $A(F, 0)$ with

$$F = \begin{pmatrix} I_n + E_{i,j} & 0 \\ 0 & I_n + E_{j,i} \end{pmatrix}$$

belong to this subgroup. The matrices $I_n + E_{i,j}$ are called transvection matrices and they are known to generate the group $\mathrm{SL}_n(\mathbb{Z}_2)$, which coincides with $\mathrm{GL}_n(\mathbb{Z}_2)$. This implies that $\operatorname{Im}\mathcal{F}$ contains all the $A(F, 0)$ associated with

$$F = \begin{pmatrix} M & 0 \\ 0 & (M^{-1})^t \end{pmatrix}, \tag{C6}$$

where $M \in \mathrm{GL}_n(\mathbb{Z}_2)$.

This means that $\operatorname{Im}\mathcal{F}$ contains all the block diagonal matrices of F_{CSS}. The anti-diagonal matrices of F_{CSS} can be obtained by multiplication with the matrix F of Eq. (C3). This shows that $\operatorname{Im}\mathcal{F}$ contains all the affine maps $A(F, 0)$, with $F \in F_{CSS}$. Finally, to reach $A(F, \mathbf{t}) = A(0, \mathbf{t})A(F, 0)$, note that $A(0, \mathbf{t}) \in \operatorname{Im}\mathcal{F}$ by Eq. (C5). □

Thanks to this group morphism, we obtain a complete description of the group G_{CSS}. First, we have the group isomorphism

$$G_{CSS}/\operatorname{Ker}\mathcal{F} \simeq \operatorname{Im}\mathcal{F}. \tag{C7}$$

By construction of $\mathcal{F}$, its kernel is the set of orthogonal matrices commuting with every matrix. This is $\{\pm I_{2^n}\}$. We have seen above that $\operatorname{Im}\mathcal{F}$ is generated by the images of $\otimes_i H_i, CNOT(i,j)$ and $T_{\mathbf{u}}$. Thus, from the previous isomorphism, the group G_{CSS} is generated by these 3 types of operators and by $\operatorname{Ker}(\mathcal{F}) = \{\pm I\}$. This proves the first part of Lemma 1.

The Affine group $\mathrm{AGL}(\mathbb{Z}_2)$ is known to be the semi-direct product of the group of translations by the general linear group. We obtain a similar structure for $\operatorname{Im}\mathcal{F}$. It is the semi-direct product of the group of translations $A(0, \mathbf{t})$ by F_{CSS}, which implies

$$\operatorname{Im}\mathcal{F} \simeq \mathbb{Z}_2^{2n} \rtimes F_{CSS}. \tag{C8}$$

To prove this decomposition, it is suficient to check that these two sets are subgroups of $\operatorname{Im}\mathcal{F}$ which jointly generate $\operatorname{Im}\mathcal{F}$ and that the subgroup of translations is a normal subgroup.

By definition the group F_{CSS} can also be decomposed as a semi-direct product

$$F_{CSS} \simeq \mathrm{GL}_n(\mathbb{Z}_2) \rtimes Z_2. \tag{C9}$$

The set of block-diagonal matrices of F_{CSS} is a subgroup isomorphic to $\mathrm{GL}_n(\mathbb{Z}_2)$ and it is normal since it is a subgroup of index 2 of F_{CSS}. The second component is the subgroup of F_{CSS} generated by the matrix of order 2

$$\begin{pmatrix} O & I_n \\ I_n & O \end{pmatrix},$$

and is isomorphic to $\mathbb{Z}_2$. The second item of Lemma 1 follows from Eq. (C7), Eq. (C8), and Eq. (C9). □

[1] S. Bravyi and A. Kitaev, Phys. Rev. A 71, 022316 (2005).
[2] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[3] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[4] M. Howard *et al.*, Nature **510**, 351355 (2014).
[5] Ernesto F, Galvão, Phys. Rev. A **71**, 043202 (2005).
[6] V. Veitch, C. Ferrie, D. Gross and J. Emerson, New J. Phys. **14**, 113011 (2012).
[7] E. Wigner, Phys. Rev. 40, 749 (1932).
[8] K.S. Gibbons, M.J. Hoffman, and W.K. Wootters, Rev. A **70**, 062101 (2004).
[9] R. W. Spekkens, Phys. Rev. Lett. **101**, 020401 (2008).
[10] D. Gross, PhD Thesis, Imperial College London, 2005.
[11] D. Gottesman, Ph.D. thesis, California Institute of Technology, 1997.
[12] A. Mari, J. Eisert, Phys. Rev. Lett. **109**, 230503 (2012).
[13] Cecilia Cormick, Ernesto F. Galvão, Daniel Gottesman, Juan Pablo Paz, and Arthur O. Pittenger, Phys. Rev A **73**, 012301 (2006).
[14] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[15] T. Rudolph, L. Grover, *A 2 rebit universal for quantum computing*, arXiv:quant-ph/0210187
[16] A. R. Calderbank, E. M Rains, P. W. Shor, N. J. A. Sloane, Phys. Rev. Lett. **78** 405 (1997).
[17] Adan Cabello, Simone Severini, Andreas Winter, Phys. Rev. Lett. **112**, 040401 (2014).
[18] A. Kitaev, Ann. Phys. (N.Y.) 303, 2 (2003).
[19] R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007).
[20] M.A. Nielsen and I.L. Chuang, *Quantum Information and Computation*, Cambridge University Press, 2000.
[21] Victor Veitch, Seyed Ali Hamed Mousavian, Daniel Gottesman, and Joseph Emerson, New J. Phys. **16**, 013009 (2014).
[22] R. L. Hudson, Rep. Math. Phys. 6, 249 (1974).
[23] Mark Howard, Eoin Brennan and Jiri Vala, Entropy **15**, 2340 (2013).
[24] Samson Abramsky, Adam Brandenburger, New J. Phys. 13, 113036 (2011).
[25] J. Anders and D. E. Browne, Phys. Rev. Lett. **102**, 050502 (2009).
[26] M. J. Hoban and D. E. Browne, Phys. Rev. Lett. **107**, 120402 (2011).
[27] R.Raussendorf, Phys. Rev A **88**, 022322 (2013).
[28] H. Bombin, Phys. Rev. Lett. 105, 030403 (2010).
[29] A.G. Fowler, Quant. Inf. Comp. 12, 970 (2012).
[30] In the qudit case [10], this result is deduced from the qudit version of Eq. (18). This strategy cannot be adapted here since Eq. (18) only involves two vectors $\mathbf{q}$ and $\mathbf{q}+\mathbf{x}$.