

Wigner Function Negativity and Contextuality in Quantum Computation on Rebits

Nicolas Delfosse,¹ Philippe Allard Guerin,² Jacob Bian,² and Robert Raussendorf²

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*

(Received 1 October 2014; published 2 April 2015)

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 (London)* **510**, 351 (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.

DOI: [10.1103/PhysRevX.5.021003](https://doi.org/10.1103/PhysRevX.5.021003)

Subject Areas: Quantum Physics, Quantum Information

I. INTRODUCTION

In quantum computation by state injection (QCSI) [1], the set of quantum gates is, by construction, not universal. This restriction is compensated 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 the following question: Precisely which quantum properties of these states are responsible for the gain in computational power?

Contextuality [2–5] and negativity of Wigner functions have recently been proposed as the quintessential quantum properties of magic states; see Refs. [6–8]. Contextuality is an obstruction to modeling 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 [9–11] 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 nonclassical behavior [6,12].

The reason for the appearance of Wigner functions in the discussion of QCSI is their relation [7,11] to the stabilizer formalism [13]. 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.

The epitome for the link between Wigner functions and QCSI via the stabilizer formalism is the discrete Hudson's theorem [11], which says that in Hilbert spaces of odd prime-power (hence finite) dimension, the pure states with positive Wigner functions are exactly the stabilizer states. Thus, stabilizer states are “classical” from the perspectives of both 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 [6,8].

Extending these properties to two-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 [11,14], and for the remaining ones [7,15], the discrete Hudson's theorem breaks down. There are qubit stabilizer states with negative Wigner functions. As for contextuality, it now arises in its state-independent form [16]. As a result, every quantum state of more than one qubit can be considered contextual [6], 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 [17]. The role that was previously played by the stabilizer states is now played by the Calderbank-Shor-Steane (CSS) states [18], 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

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

function. Furthermore, the restriction to CSS-ness preserving operations permits us to carve out a computational scheme of rebit QCSI that is free of state-independent contextuality, even if this phenomenon exists in rebits.

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 Sec. III, we present a universal scheme of quantum computation by state injection on rebits. In Sec. IV, we construct a matching Wigner function, equipped with a discrete Hudson’s theorem and extended Gottesman-Knill theorem. In Sec. 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 Sec. 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. In other words, the restricted set of unitary gates is in the group of Clifford gates, the restricted set of observables are 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 when the restricted gate set is the Clifford gates [6,8]. 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 Ref. [8]) and for the hardness of classical simulation of QCSI (Theorem 1 in Ref. [8]).
- (ii) Contextuality of magic states with respect to stabilizer measurements is necessary for universality of QCSI [6].

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 second role 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 theorem as the set of pure states with non-negative Wigner functions [11].
- (iv) The set of Clifford gates is singled out as the set of unitaries that transform the Wigner function covariantly [11].
- (v) Clifford gates and stabilizer measurements preserve positivity of the Wigner function [8].
- (vi) Necessary and sufficient conditions for contextuality with respect to the restricted set of measurements can be expressed in terms of the Wigner function [6].
- (vii) For one-qudit states, negativity of the Wigner function and contextuality with respect to measurements from the restricted set are the same [6].

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 functions under the restricted gates can be built on properties (iv) and (v) [8]. The existence of this simulation method directly implies (i). Furthermore, Hudson’s theorem (iii) connects this simulation method with the Gottesman-Knill theorem.

B. Trouble with qubits

For systems of qubits, both the employed contextuality witnesses [19] 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 [16]. Mermin’s square can be translated into a contextuality witness for which *all* quantum states of $n \geq 2$ qubits come out contextual [6]. 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, they do not apply to the qubit case $d = 2$; for examples, see, e.g., Refs. [11,14].

Yet, some Wigner functions do survive the transition to $d = 2$; see, e.g., Refs. [7,15]. However, in these cases, the general connection with the stabilizer world breaks down. Not all stabilizer states have non-negative Wigner functions, and the Wigner function no longer transforms covariantly under all Clifford operations.

For the construction [7,15], 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 Ref. [10]. 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 superexponentially small compared to the number of n -qubit stabilizer states [11].

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

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

In other words, 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. We denote by Ω the set of pure CSS states and by G_{CSS} the subgroup of the n -qubit Clifford group C_n that preserves the set Ω of CSS states,

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

The following can be said about the structure of G_{CSS} .

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

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

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

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

In Eq. (4), the component \mathbb{Z}_2^{2n} corresponds to the Pauli operators $T_{\mathbf{u}}$, the component $\text{GL}_n(\mathbb{Z}_2)$ 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 $\mathcal{O} = \{Z(\mathbf{u}) | \mathbf{u} \in \mathbb{Z}_2^n\} \cup \{X(\mathbf{v}) | \mathbf{v} \in \mathbb{Z}_2^n\}$ is mapped onto itself by conjugation under gates from the group on the right-hand side 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.

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 Ref. [17] 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,

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

The additional rebit, with basis states $|R\rangle = |0\rangle$ and $|I\rangle = |1\rangle$, allows us 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 realness of the states in Eq. (5).

Using the encoding Eq. (5), we construct a universal scheme of QCSI on rebits. The restricted gate set therein consists of CNOT gates, the simultaneous Hadamard gate

$$H_{\text{all}} := \bigotimes_{i=1}^n H_i, \text{ and Pauli flips } X_i, Z_j; \text{ i.e.,}$$

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

These unitary gates are supplemented by measurements of observables in the set \mathcal{O} or, without loss of generality, of observables $\{Z_i | i = 1, \dots, n\}$.

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 $\text{CNOT}(i, j)H_{\text{all}} = H_{\text{all}}\text{CNOT}(j, i)$. If those gates are eliminated, we are left with the CNOT gates and measurements of X_i and Z_i . We note that this is precisely the set of gates that can be performed fault tolerantly on the surface code [20] using defect braiding [21]. 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

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

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

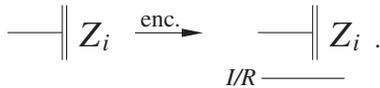
We now demonstrate that the encoded versions of these gates can be realized only by using the gates from the restricted set and the injection of two types of ancilla states, $|A\rangle$ and $|B\rangle$, defined as

$$|A\rangle = \frac{|0\rangle|R\rangle + \cos\frac{\pi}{4}|1\rangle|R\rangle + \sin\frac{\pi}{4}|1\rangle|I\rangle}{\sqrt{2}},$$

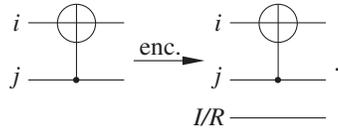
$$|B\rangle = \frac{|0\rangle|+\rangle + |1\rangle|-\rangle}{\sqrt{2}}. \quad (6)$$

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).

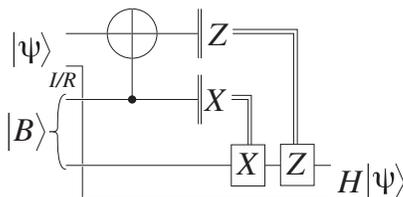
- (a) *The measurement of Z_i .* Since the Pauli operator Z is real, its measurement does not differentiate between the real and imaginary parts of the measured state, and $\bar{Z}_i \equiv Z_i$. Graphically,



- (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, $\overline{\text{CNOT}(i, j)} = \text{CNOT}(i, j)$. Graphically,

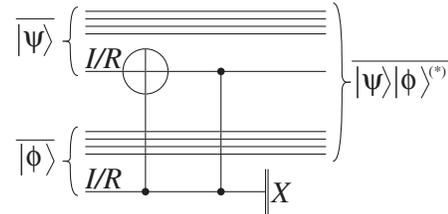


- (c) *The Hadamard gate H_i .* The encoded Hadamard gate is realized by injection of an ancilla $|B\rangle$ into the circuit



- (d) *The gate $\exp(i\pi/8Z_i)$.* The encoded version of this gate uses ancilla states $|A\rangle$ and $|B\rangle$ and proceeds in two steps. The first step is a preprocessing jointly for all the $\exp(i\pi/8Z_i)$ 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 preprocessing

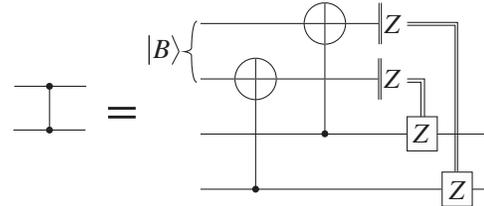
step, all data and ancilla rebits are merged into the same code block. The merging can be done two blocks at a time, and the corresponding circuit is



For a pair of encoded input states $|\overline{\psi}\rangle, |\overline{\phi}\rangle$, the result of the code-merging circuit is $|\overline{\psi}\rangle \otimes |\overline{\phi}\rangle$ or $|\overline{\psi}\rangle \otimes |\overline{\phi}\rangle^*$, depending on the outcome of the X measurement ($|\overline{\phi}\rangle^*$ denotes the state obtained from $|\overline{\psi}\rangle$ by complex conjugation with respect to 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 $|\overline{\pi/8}\rangle$ and $|\overline{\pi/8}\rangle^* = |-\overline{\pi/8}\rangle$ allow us 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 that 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 [22],

$$\begin{array}{c} |A\rangle \text{---} \bullet \text{---} |\text{out}\rangle \\ | \text{in}\rangle \text{---} \oplus \text{---} \| Z \end{array} \quad (7)$$

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

IV. 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; finally, we show that the evolution of states with a 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 function W 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} [6,8] for qubits. Of the properties (i)–(vii) listed in Sec. II A for the Wigner function on qudits, our rebit Wigner function has counterparts for properties (i)–(vi) but not for (vii).

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

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

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

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

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 $(A, B) = (1/2^n)\text{Tr}(A^\dagger B)$,

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

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

$$\mathcal{A} := \{T_a | (\mathbf{a}_Z, \mathbf{a}_X) = 0 \pmod{2}\}, \quad (10)$$

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

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

with

$$A_0 = \frac{1}{2^n} \sum_{T_a \in \mathcal{A}} T_a, \quad \text{and} \quad A_{\mathbf{u}} = T_{\mathbf{u}} A_0 T_{\mathbf{u}}^\dagger. \quad (12)$$

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

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

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 [11,18], the rebit Wigner function of Eq. (11) has the following properties (compare with Ref. [23]):

- (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

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

- (6) The phase point operators satisfy $\text{Tr} A_{\mathbf{u}} = 1$. Thus, $\text{Tr} B = \sum_{\mathbf{u}} W_B(\mathbf{u})$ for any symmetric operator B . Property 1 is proven in Lemma 20 in Appendix B, Property 2 in Sec. IV B, and Property 3 in Sec. IV C. Property 4 is shown in Appendix B. Properties 5 and 6 are immediate consequences of Property 1.

B. Discrete Hudson's theorem for rebits

The original Hudson theorem in infinite-dimensional Hilbert space [24] singles out the Gaussian states as the pure states with a positive Wigner function. This result has a counterpart in finite, odd prime-power dimensions. Namely, the pure states with a positive Wigner function are the stabilizer states [11]. 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 multiple qubits does this result carry over (see Ref. [25], however, for a single qubit).

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 a 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 it is said to be negative otherwise.

In order to prove this result, we follow the strategy pursued by Gross for the qudit case [11]. First, we determine the Wigner function of CSS states in Sec. IV B 1, proving that these Wigner functions are non-negative. Then, in Sec. IV B 2, we consider a pure state with a 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 Sec. 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 $(-1)^{a_i} Z(\mathbf{a}_i)$, for $1 \leq i \leq r$, and $n-r$ independent operators $(-1)^{a_i} X(\mathbf{b}_i)$, 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 $N^\perp = \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle$.

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 $\mathbf{t} \in \mathbb{Z}_2^{2n}$, we have

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

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

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

To transition from the third to the fourth line above, we use the property that $V_S \subset V_A$. ■

2. Non-negative Wigner functions

To complete the proof of Theorem 1, we consider a pure state that has a 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 [11], we will show the following.

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

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

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, the following.

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 $\psi: \mathbb{Z}_2^n \rightarrow \mathbb{R}$ defined by $\psi(\mathbf{x}) = \langle\psi|\mathbf{x}\rangle$. The support of ψ , denoted $\text{supp}(\psi)$, 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

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

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

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

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

Equation (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 a constant absolute value over its support $\text{supp}(\psi)$.

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 $2^{n/2}$, which 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 Ref. [11]. (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 positive 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 positive semi-definite matrices, every principal minor of the matrix $A_{\mathbf{x}}^{\mathbf{x}}$ 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 $\mathbf{q} \in \text{supp}(\psi)$ and $\mathbf{x} \in \mathbb{Z}_2^n$, then we obtain

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

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

Now, consider two vectors \mathbf{q} and \mathbf{q}' of $\text{supp}(\psi)$. 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 Eq. (18). Thus,

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

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

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 , $\text{supp}(\psi) = \mathbf{q}_0 + N$.

Proof of Lemma 6.—Let \mathbf{q} , $\mathbf{q} + \mathbf{x}$ and $\mathbf{q} + \mathbf{y}$ be three vectors in $\text{supp}(\psi)$. We have to show that $\mathbf{q} + \mathbf{x} + \mathbf{y}$ is also in $\text{supp}(\psi)$. [In the qudit case [11], 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 $\mathbf{q} + \mathbf{x}$.] 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 $A_{\mathbf{x}}^{\mathbf{x}}$, which is also non-negative by Bochner's theorem.

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

The expansion of this determinant leads to the inequality

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

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 on the left-hand side are equal and positive. This contradiction implies that $\psi(\mathbf{q} + \mathbf{x} + \mathbf{y}) \in \text{supp}(\psi)$. Hence, $\text{supp}(\psi)$ is an affine space: $\text{supp}(\psi) = \mathbf{q}_0 + N$, 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

$$W_{\psi}(\cdot, \mathbf{q}) = c\delta_{\mathbf{p}_0 + N^{\perp}},$$

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 focus on the support of the function $W_{\psi}(\cdot, \mathbf{q})$. From Lemma 4, this function satisfies

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

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 $\text{supp}(K) = (\mathbf{q}_0 + N) \times N$, where $\mathbf{q}_0 + N$ is the support of ψ . Then, K can be restricted to its support. This restriction gives

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

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

Now note that, for every vector $\mathbf{q} \in \mathbb{Z}_2^n$, 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^* . Moreover, 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 a constant absolute value over N by Lemma 5; thus, we apply the second item of the Bochner Theorem (Theorem 44 in Ref. [111]) to $W_\psi([\cdot], \mathbf{q})$. This tells us that $W_\psi([\cdot], \mathbf{q})$ is orthogonal to its translations, i.e.,

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

for every nonzero $[\mathbf{t}] \in \mathbb{Z}_2^n / (N^\perp)$. A positive function that 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 $\text{supp}(W_\psi)$ is the disjoint union

$$\text{supp}(W_\psi) = \bigsqcup_{\mathbf{q} \in \mathbf{q}_0 + N} (\mathbf{p}_0(\mathbf{q}) + N^\perp) \times \{\mathbf{q}\}. \quad (21)$$

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 separation 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 $\text{supp}(W_\psi) = \text{supp}(\hat{\psi}) \times \text{supp}(\psi)$. 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

$$\text{supp}(W_\psi) = (\mathbf{p}_0 + N^\perp) \times (\mathbf{q}_0 + N).$$

Now, let us prove that W_ψ has a constant absolute value over its support. Let $(\mathbf{p}, \mathbf{q}) \in \text{supp}(W_\psi)$. 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 $\text{supp}(\psi)$. 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 normalization of Property 6, of the Wigner function, it suffices to compute the cardinality of the support of W_ψ . We find $|\text{supp}(W_\psi)| = |N^\perp| \cdot |N| = 2^{n-\dim N} \cdot 2^{\dim N} = 2^n$, which gives $c = 1/2^n$. This concludes the proof. ■

3. Proof of Hudson's theorem for rebits

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

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

Now, consider a pure real state $|\psi\rangle$ that 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, cf. Lemma 11 below. To prepare for this result, we make two observations.

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

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

The proof of Lemma 9 is given in Appendix C.

Furthermore, the action of a $g \in G_{\text{CSS}}$ on a translation operator $T_{\mathbf{a}} \in \mathcal{T}$ by conjugation induces a morphism from the CSS Clifford group to the affine group $\text{AGL}_{2n}(\mathbb{Z}_2)$. Recall that an affine transformation of $\text{AGL}_{2n}(\mathbb{Z}_2)$ is an application of the form $A(F, \mathbf{t}): \mathbf{a} \mapsto F\mathbf{a} + \mathbf{t}$, where $F \in GL_{2n}(\mathbb{Z}_2)$ 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 $\text{ASp}_{2n}(\mathbb{Z}_2^{2n})$.

Lemma 10: Let \mathcal{F} be the application

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

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 uniqueness 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 $\mathbf{u} \in \mathbb{Z}_2^{2n}$, and for all $g \in G_{\text{CSS}}$, it holds that

$$W_{g^\dagger \rho g}(\mathbf{u}) = W_\rho(\mathcal{F}(g)(\mathbf{u})). \quad (24)$$

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

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

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

Proof of Lemma 11.—Let $g \in G_{\text{CSS}}$ 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

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

which proves the covariance. \blacksquare

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 two-qubit graph state. The former has a positive and the latter a 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 [8] for qudits: Circuits of Clifford gates and stabilizer measurements acting on an initial state with a non-negative Wigner function can be efficiently simulated classically. The discrete Hudson's theorem [11] 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 functions are mixtures of stabilizer states [11].

Here, we prove an analogue of the result [8] for the rebit Wigner function W defined in Eqs. (11) and (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 a 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 $E(s) = (I + sO)/2$ 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 $\rho(0) = \rho_1 \otimes \rho_2 \otimes \dots \otimes \rho_n$, $W_{\rho(0)} = W_{\rho_1} W_{\rho_2} \dots W_{\rho_n}$ 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 $g \in G_{\text{CSS}}$, and (b) $\rho(t+1) \sim (I + sO)/2 \rho(t) (I + sO)/2$, with $O \in \mathcal{O}$, $s = \pm 1$.

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

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

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 the following.

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

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

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

$W_{\rho(t+1)}$ is sampled from as follows. Repeat: (1) Call the sampling routine for $W_{\rho(t)}$, which returns a $\mathbf{u} \in V$. (2) Report the measurement outcome $s = (-1)^{[\mathbf{u}, \mathbf{a}]}$. (3) Flip a fair coin, and, depending on the outcome, report \mathbf{u} or $\mathbf{u} + \mathbf{a}$ as a 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, $\rho(0) = \rho_1 \otimes \rho_2 \otimes \dots \otimes \rho_n$, 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:—The present simulation method is similar to its qudit counterpart [23], but a difference occurs in measurement. Here, mere positivity of the effect $W_{E(s)}$ and positivity of $W_{\rho_{\text{in}}}$ for the input state ρ_{in} do not imply positivity of the Wigner function $W_{\rho_{\text{out}}}$ for the output state ρ_{out} . For example, the two-rebit state $\rho = (I + X_1 Z_2)/4$ has a 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 a *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

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

This result 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 $\mathcal{A}_{\mathbf{y}} = \{T_{\mathbf{a}} \in \mathcal{A} \mid [\mathbf{a}, \mathbf{y}] = 0\}$. It has the property that

$$T_{\mathbf{y}} \mathcal{A}_{\mathbf{y}} = \mathcal{A}_{\mathbf{y}}. \quad (26)$$

Equation (26) holds because $[T_{\mathbf{a}}, T_{\mathbf{x}}] = 0$, so Eq. (25) guarantees that $T_{\mathbf{y}} T_{\mathbf{a}}$ has the right sign and thus $T_{\mathbf{y}} T_{\mathbf{a}} \in \mathcal{A}$.

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

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

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 noncontextual 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 noncontextual HVM must 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 that can be measured in rebit QCSI, and the sets of observables that 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 that can be physically measured in rebit QCSI is $\mathcal{O} = \{X(\mathbf{a}_{\mathbf{x}}), Z(\mathbf{a}_{\mathbf{z}})\}$. 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 postprocessing the outcomes. Therefore, a measurement of $X_1 Z_2$ can be reduced to measurements of commuting 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 $\{X_1, X_1Z_2\}$, namely, by the same operations that measured X_1Z_2 alone.

Now, is it possible to simultaneously measure the commuting observables X_1Z_2 and Z_1X_2 ? In the setting of rebit QCSI, this is not the case. The measurement of X_1Z_2 necessitates the measurement of X_1 and Z_2 separately. Since these observables do not commute with Z_1X_2 , a subsequent measurement of Z_1X_2 is no longer guaranteed to reveal the original value. Thus, commuting observables in \mathcal{A} need not be simultaneously measurable in the same way as commuting 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 Let M be 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 to M .

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

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

$$p(\mathbf{s}_M|\mathbf{u}) = \prod_{i|O_i \in M} \delta_{s_i, \lambda_{\mathbf{u}}(O_i)}. \quad (27)$$

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

$$\lambda_{\mathbf{u}}(AB) = \lambda_{\mathbf{u}}(A)\lambda_{\mathbf{u}}(B). \quad (28)$$

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

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

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

In Secs. VB and VC below, we derive necessary and sufficient conditions for the existence of a hidden variable

model over \mathcal{M} or, vice versa, 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: Let $M \subset \mathcal{A}$ be a set of commuting observables. Then, $M \in \mathcal{M}$ if and only if $T_{\mathbf{a}}T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}}$, $\forall T_{\mathbf{a}}, T_{\mathbf{b}} \in M$.

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 M and among themselves. They thus generate a CSS stabilizer

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

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 postprocess 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 $T_{\mathbf{a}+\mathbf{b}} = (-1)^{\mathbf{a}_X \cdot \mathbf{b}_Z} T_{\mathbf{a}}T_{\mathbf{b}}$, 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_1Z_1 and Z_1X_2 cannot be simultaneously measured in rebit QCSI; $\{X_1Z_1, Z_1X_2\} \notin \mathcal{M}$. Lemma 13 detects this as follows: If $T_{\mathbf{a}} = X_1Z_2$ and $T_{\mathbf{b}} = Z_1X_2$, then $T_{\mathbf{a}+\mathbf{b}} = -Y_1Y_2$, and therefore $T_{\mathbf{a}+\mathbf{b}} = -T_{\mathbf{a}}T_{\mathbf{b}}$.

B. A necessary condition for contextuality

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

Proof of Theorem 3.—If $W_\rho > 0$, then W_ρ is a valid noncontextual HVM for the setting (ρ, \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 $E(\mathbf{s}_M)$,

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

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

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

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 $E(s) = (I + sT_{\mathbf{a}}/2)$ and $T_{\mathbf{a}} \in \mathcal{A}$. Using the orthogonality relation $\text{Tr}(T_{\mathbf{a}}T_{\mathbf{b}}) = 2^n \delta_{\mathbf{a},\mathbf{b}}$, 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

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

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

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

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 (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

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

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 noncontextual HVM for the setting (ρ, \mathcal{M}) . The claim follows by negation of this statement. ■

Finally, as an application of Theorem 3, we briefly discuss the state-dependent version of Mermin's star [16]. Employing a Greenberger-Horne-Zeilinger (GHZ) state $(|000\rangle + |111\rangle)/\sqrt{2}$ in the rebit setting, there is neither negativity nor contextuality. The GHZ state, being of CSS type, has a non-negative Wigner function and hence, by Theorem 3, is noncontextual. Correspondingly, Mermin's parity proof does not apply to rebits because the local Pauli observables Y_i are imaginary.

C. Sufficient condition for contextuality

Below we provide a sufficient criterion for contextuality in terms of the Wigner function. It involves the notion of an isotropic subspace. A subspace $U \subset V = \mathbb{Z}_2^{2n}$ is isotropic if, for all $\mathbf{v}, \mathbf{w} \in U$, $[\mathbf{v}, \mathbf{w}] := (\mathbf{v}_X, \mathbf{w}_Z) + (\mathbf{v}_Z, \mathbf{w}_X) \bmod 2 = 0$. Such a space U is said to be maximally isotropic if it is a maximal isotropic subspace of \mathbb{Z}_2^{2n} with respect to inclusion. This happens if and only if the dimension of the isotropic subspace U is n .

Theorem 4: The n -rebit setting (ρ, \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

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

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 in Sec. V D.

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 $\mathcal{B}(U) = \{\mathbf{a}(1), \mathbf{a}(2), \dots, \mathbf{a}(m)\}$ and can be evaluated on points $\mathbf{x} \in \mathbb{Z}_2^{2n}$, for any density operator ρ . Namely, we define

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

The contextuality witnesses $\mathcal{W}^{\mathcal{B}(U)}$ resemble the witnesses discussed by Cabello, Severini, and Winter in [19] in the following respect. They are linear operators for which the range of expectation values allowed by quantum mechanics is strictly greater than that allowed for noncontextual HVMs. We make the following observation.

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

Before turning to the proof of Lemma 14, we illustrate the contextuality witnesses, Eq. (33), in a specific case.

Example. Consider two rebits, and a maximal isotropic subspace $U = \mathbb{Z}_2^2 = \text{span}(\{\mathbf{a}, \mathbf{b}\})$ such that $T_{\mathbf{a}} = X_1 Z_2$ and $T_{\mathbf{b}} = Z_1 X_2$. With these specifications,

$$\mathcal{W}_\rho^{\{\mathbf{a}, \mathbf{b}\}}(\mathbf{0}) = \langle I_{12} + X_1 Z_2 + Z_1 X_2 - Y_1 Y_2 \rangle_\rho, \quad \lambda_{\mathbf{u}}(T_{\mathbf{a}}) = (-1)^{[\mathbf{u}, \mathbf{a}]}, \quad \forall \mathbf{u} \in \mathcal{S}, \quad (34)$$

Note that $T_{\mathbf{a}+\mathbf{b}} = -Y_1 Y_2 = -T_{\mathbf{a}} T_{\mathbf{b}}$. If we choose $\rho = |K_2\rangle\langle K_2|$ for a graph state $|K_2\rangle$ with stabilizer relations $X_1 Z_2 |K_2\rangle = Z_1 X_2 |K_2\rangle = -|K_2\rangle$, then $\mathcal{W}_{|K_2\rangle}^{\{\mathbf{a}, \mathbf{b}\}}(\mathbf{0}) = -2$. The witness \mathcal{W} can thus indeed take negative values, but what does that say about contextuality?

To answer this question, assume that there exists a noncontextual HVM in which all observables in \mathcal{A} have values $\lambda(\cdot) = \pm 1$ and that these values satisfy the compatibility condition Eq. (28). Then, $\lambda(X_1 Z_2) = \lambda(X_1)\lambda(Z_2)$ and $\lambda(Z_1 X_2) = \lambda(Z_1)\lambda(X_2)$. Similarly, $\lambda(-Y_1 Y_2) = \lambda(X_1 X_2)\lambda(Z_1 Z_2) = \lambda(X_1)\lambda(X_2)\lambda(Z_1)\lambda(Z_2)$. Therefore, the HVM version of the witness $\mathcal{W}\{\mathbf{a}, \mathbf{b}\}(\mathbf{0})$ evaluates to

$$\mathcal{W}_\lambda^{\{\mathbf{a}, \mathbf{b}\}}(\mathbf{0}) = (1 + \lambda(X_1)\lambda(Z_2))(1 + \lambda(Z_1)\lambda(X_2))$$

and is thus non-negative for every value assignment λ to the observables X_1, X_2, Z_1, Z_2 . Hence, it is also non-negative for all probabilistic mixtures over such assignments. A negative value of $\mathcal{W}_\rho^{\{\mathbf{a}, \mathbf{b}\}}(\mathbf{0})$ is therefore an indicator of contextuality.

In addition, we observe that the witness $\mathcal{W}_\rho^{\{\mathbf{a}, \mathbf{b}\}}(\mathbf{0})$ is closely related to state-independent contextuality. Combining the aforementioned relations for $\lambda(T_{\mathbf{a}} = X_1 Z_2)$, $\lambda(T_{\mathbf{b}} = Z_1 X_2)$ and $\lambda(T_{\mathbf{a}+\mathbf{b}} = -Y_1 Y_2)$, we find that $\lambda(T_{\mathbf{a}+\mathbf{b}}) = \lambda(T_{\mathbf{a}})\lambda(T_{\mathbf{b}})$. From Eq. (28), this contradicts the above operator relation $T_{\mathbf{a}+\mathbf{b}} = -T_{\mathbf{a}} T_{\mathbf{b}}$, giving rise to a state-independent parity proof of contextuality. In fact, the proof in question is a locally rotated version of Mermin's square [16] [also see Eq. (44)].

Proof of Lemma 14.—We prove the converse statement, namely, that if (ρ, \mathcal{M}) is noncontextual, 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 noncontextual HVM describing the setting (ρ, \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 $M = \{Z(\mathbf{a}_Z) | \mathbf{a}_Z \in \mathbb{Z}_2^n\}$ satisfies Criterion 1. Therefore, by property (ii) of Definition 1, $\lambda_{\mathbf{u}}(T_{(\mathbf{a}_Z, 0)}) = \lambda_{\mathbf{u}}(Z(\mathbf{a}_Z)) = \prod_{i|[\mathbf{a}_Z]_i=1} \lambda_{\mathbf{u}}(Z_i)$. Likewise, $\lambda_{\mathbf{u}}(T_{(0, \mathbf{a}_X)}) = \lambda_{\mathbf{u}}(X(\mathbf{a}_X)) = \prod_{i|[\mathbf{a}_X]_i=1} \lambda_{\mathbf{u}}(X_i)$. 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), $\lambda_{\mathbf{u}}(T_{(\mathbf{a}_Z, \mathbf{a}_X)}) = \lambda_{\mathbf{u}}(T_{(\mathbf{a}_Z, 0)})\lambda_{\mathbf{u}}(T_{(0, \mathbf{a}_X)})$.

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 $i = 1, \dots, n$. We may write this as

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 noncontextual HVMs.

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

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

We rewrite this condition as

$$\begin{aligned} \lambda_{\mathbf{u}}(T_{\mathbf{a}+\mathbf{b}}) &= \lambda_{\mathbf{u}}(T_{\mathbf{a}})\lambda_{\mathbf{u}}(T_{\mathbf{b}}), \\ \forall T_{\mathbf{a}}, T_{\mathbf{b}} \in \mathcal{A} \text{ such that } [T_{\mathbf{a}}, T_{\mathbf{b}}] &= 0. \end{aligned} \quad (35)$$

We now evaluate the witness $\mathcal{W}_\rho^{\mathcal{B}(U)}(\mathbf{x})$ under the assumption of a noncontextual 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

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

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 noncontextual 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. ■

Remark 4.—The connection between the witnesses $\mathcal{W}^{\mathcal{B}(U)}$ and state-independent contextuality observed in the earlier two-rebit example persists in the general case. While the witnesses measure—as it is their purpose—contextuality possessed by quantum states, they are linked to state-independent parity proofs of contextuality as given in Ref. [16]. Namely, a witness $\mathcal{W}^{\mathcal{B}(U)}$ can assume a negative value only if the associated isotropic space U contains two vectors \mathbf{a}, \mathbf{b} such that $T_{\mathbf{a}+\mathbf{b}} = -T_{\mathbf{a}} T_{\mathbf{b}}$. Whenever that happens, a parity proof can be built from $T_{\mathbf{a}+\mathbf{b}}, T_{\mathbf{a}}, T_{\mathbf{b}}$ and Pauli operators X_i, Z_i ; cf. Eq. (35).

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 $\mathcal{B}(U) = \{\mathbf{a}(1), \mathbf{a}(2), \dots, \mathbf{a}(m)\}$ and a set $\tilde{\mathcal{B}} = \{\mathbf{b}(1), \mathbf{b}(2), \dots, \mathbf{b}(m)\}$ such that $[\mathbf{a}(i), \mathbf{b}(j)] = \delta_{ij}$ for all $i, j = 1, \dots, m$. For every $\eta(\mathbf{x}) = \sum_i x_i \mathbf{a}(i) \in U$, denote by $\bar{\eta}(\mathbf{x})$ the vector $\bar{\eta}(\mathbf{x}) = \sum_i x_i \mathbf{b}(i) \in \mathbb{Z}_2^{2n}$. Then,

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

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

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

We may further rewrite this expression as

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

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, $\text{Tr} T_{\mathbf{a}} \rho = 0$, for all \mathbf{a} with $(\mathbf{a}_X, \mathbf{a}_Z) \bmod 2 = 1$. ■

Proof of Theorem 4.—The combined conclusion of Lemmas 14 and 15 is that the n -rebit setting (ρ, \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

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

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 $\bar{U} \subset \mathbb{Z}_2^{2n}$ such that $\tilde{U}^\perp = U^\perp \oplus \bar{U}$. Hence,

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

If every term in brackets on the right-hand side is ≥ 0 , so is the left-hand side. Since every isotropic \tilde{U} can be embedded in a maximally isotropic U , the above claim follows. In other words, we may restrict the condition

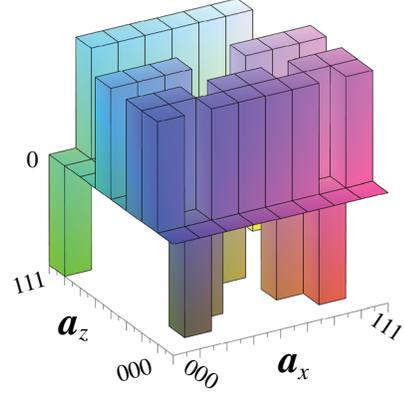


FIG. 1. Wigner function of the three-rebit graph state $|K_3\rangle$ corresponding to the complete graph K_3 . The Wigner function takes negative values, and furthermore this negativity is strong enough to witness contextuality of $|K_3\rangle$ with respect to CSS-ness preserving Pauli measurements.

Eq. (37) to maximally isotropic subspaces U . In those cases, $U^\perp = U$, which yields the condition stated in Theorem 4. ■

Finally, as an application of Theorem 4, we briefly discuss the state-dependent version of Mermin's star [16], in a locally rotated form. It comprises the nonlocal observables XXX, XZZ, ZXZ, ZZX and local observables X_i, Z_i , for $i = 1, \dots, 3$. Furthermore, the rotated GHZ state is a three-rebit graph state $|K_3\rangle$, with K_3 being the fully connected graph of three vertices; hence, $|K_3\rangle$ is a joint eigenstate of the above four nonlocal observables. $W_{|K_3\rangle}$ takes negative values; see Fig. 1. $W_{|K_3\rangle}$ is, in fact, so negative that it implies contextuality of $|K_3\rangle$ by Theorem 4. To see this, for the maximal isotropic subspace U appearing in the condition of Theorem 4, use $U = \text{span}(\{\mathbf{a}, \mathbf{b}, \mathbf{c}\})$ with $T_{\mathbf{a}} = XZZ$, $T_{\mathbf{b}} = ZXZ$, and $T_{\mathbf{c}} = ZZX$. Correspondingly, in contrast to the original version discussed in Sec. VB, the rotated version of Mermin's star fully embeds into real quantum mechanics such that Mermin's parity proof of contextuality applies there. Furthermore, the state-dependent version of this proof applies to rebit QCSI.

D. Are negativity and contextuality the same?

We observe that the sufficient condition for contextuality in Theorem 4 does not, in general, match the necessary condition of Theorem 3. This means that either the sufficient condition is not optimal or, for the present setting, contextuality and negativity are inequivalent.

To address this question, we consider the general one-rebit state

$$\tilde{\rho}(x, z) = \frac{I + xX + zZ}{2}. \quad (38)$$

The corresponding phase diagram is depicted in Fig. 2. The set of physical states is constrained by $x^2 + z^2 \leq 1$.

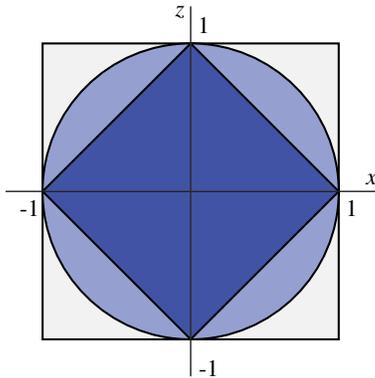


FIG. 2. Phase diagram for the families of states $\tilde{\rho}(x, y)$ of Eq. (38), with $x, y \in \mathbb{R}$. The medium-shaded area shows the physical states; the dark-shaded area shows the states classified as noncontextual by Theorem 3. The states classified as contextual by Theorem 4 lie outside the square of $|x|, |z| \leq 1$ and are thus not physical.

By Theorem 3, $\tilde{\rho}(x, z)$ is noncontextual 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, noncontextual HVMs can be constructed [2,16], and they imply noncontextual 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. Thus, for the present rebit setting, Wigner function negativity and contextuality are not the same.

We have to explain how our finding relates to the result by Spekkens [12] that negativity and contextuality, when suitably defined, *are* equivalent notions of nonclassicality. In Ref. [12], the following observations are made: (i) Non-negativity in the quasiprobability distributions representing quantum states is not sufficient for classicality; the conditional probabilities representing measurements must also be non-negative. (ii) A classical explanation cannot be ruled out by considering a single quasiprobability representation; negativity must be demonstrated for all such representations. (iii) The requirement of outcome determinism for sharp measurements should be dropped from the definition of noncontextuality. In other words, given an internal state $\mathbf{u} \in \mathcal{S}$ of the HVM, the conditional probabilities $p(\mathbf{s}_M|\mathbf{u})$ for the measurement outcomes are not required to be δ distributions.

Our setting satisfies the above criterion (i). All projectors onto eigenspaces of the measurable observables $O \in \mathcal{O}$ are non-negatively represented. This is important for the efficient classical simulation method for states with non-negative Wigner functions evolving under CSS-ness preserving operations (cf. Sec. IV D).

Regarding (iii), here we keep the requirement of outcome determinism. Hence, all conditional probability distributions $p(\mathbf{s}_M|\mathbf{u})$ for measurement outcomes given a fixed internal state are δ distributions, cf. Eq. (27) in Definition 1.

While not as general as Ref. [12], it is in accordance with Ref. [19] (the contextuality measures employed in the qudit counterpart [6] of the present work) and Ref. [16].

In addition, we point out that not just any δ distribution will do for $p(\mathbf{s}_M|\mathbf{u})$. Rather, the δ distributions, Eq. (27), are constrained by outcome compatibility, Eq. (28).

Regarding (ii), in contrast to Ref. [12], here we consider only a single quasiprobability distribution—the Wigner function defined in Eq. (11). This choice is motivated by the present computational setting to which the notions of negativity and contextuality are applied: QCSI. As described in Secs. II C and II D, CSS states, the observables in \mathcal{O} , and the CSS-ness preserving unitaries form a classical reference structure for QCSI on rebits. This result implies, in particular, that, for the present setting, certain bases of Hilbert space are preferred over others for state preparation and measurement. This inequivalence carries over to quasiprobability distributions.

In our setting, a classical explanation *can* be ruled out by considering a single quasiprobability representation. While mere negativity of the Wigner function is no guarantee for contextuality, a setting (ρ, \mathcal{M}) is contextual, hence non-classical, if the Wigner function W_ρ is sufficiently negative to satisfy the condition of Theorem 4.

Having established that, for the present situation, negativity and contextuality are not equivalent, we turn to the question of whether there are at least large families of states for which the two notions agree. An example is the family of two-rebit states

$$\rho(a, b) = \frac{(I + aX_1Z_2)(I + bZ_1X_2)}{4}. \quad (39)$$

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. 3. 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 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. This example generalizes as follows.

Lemma 16: Let ρ be a state diagonal in a real 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 $\mathbf{v} \in U$.

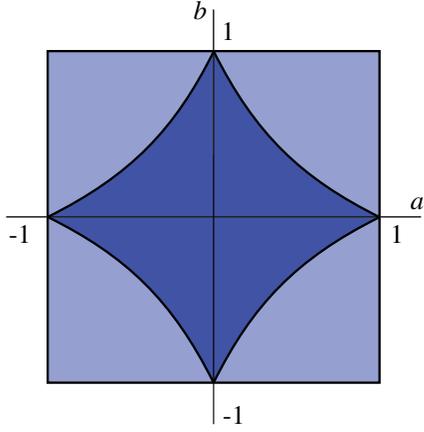


FIG. 3. Phase diagram for the states $\rho(a, b)$ of Eq. (39), with $a, b \in \mathbb{R}$. The medium-shaded area shows the physical states; the dark-shaded area shows the states classified as noncontextual by Theorem 3.

Then, by covariance of the Wigner function under translations,

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

In this case, the expression on the left-hand side 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 statement, combined with Theorem 3, proves the claim. ■

To summarize, unlike for qudits in odd prime dimensions [6], for rebits, contextuality and Wigner function negativity are not the same. Yet, they coincide on all states that are diagonal in a real stabilizer basis. However, note that the definition of contextuality in Ref. [6] is different from ours. Specifically, in Ref. [6], one-qudit states can be classified as contextual based on two-qudit measurements of the given state and a completely depolarized ancilla.

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, we give the following corollary.

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 $g = T_a g_F$, where $g_F T_b g_F^\dagger = T_{Fb}$ for all $T_b \in \mathcal{A}$, and $[Fb, Fc] = [b, c]$ for all $b, c \in V$. Then, using the form Eq. (36) of the contextuality witnesses,

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

On the right-hand side, $F^{-1}\mathcal{B}(U)$ is again the basis of an isotropic subspace since F and F^{-1} preserve the commutation relations. As a 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 creating an encoded graph state $|\bar{G}_2\rangle$, with stabilizer $\langle \bar{XZ}, \bar{ZX} \rangle$. Therein, the encoding is that of 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 $i = 1, \dots, n$, we have

$$\bar{X}_i = X_i, \quad \bar{Z}_i = Z_i, \quad \bar{Y}_i = Y_i \otimes Y_A, \quad (41)$$

where $Y := iXZ$. With $\bar{iI} = iY_A$, this is compatible with the Pauli multiplication table $\bar{Y} = \bar{iZ}\bar{X} = \bar{iI}\bar{X}\bar{Z}$.

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

$$\bar{T} = T, \quad \forall T \in \mathcal{A}. \quad (42)$$

For the state $|\bar{G}_2\rangle$, the contextuality witness based on the operators $T_a = XZ$ and $T_b = ZX$ is negative, namely,

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

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 $|\bar{G}_2\rangle$ thus maximally reveals contextuality.

We now prove that the initial state fed into the computation must also maximally reveal contextuality. The proof is by induction. We consider the circuit that created the state $|\bar{G}_2\rangle$ and assume the gates are performed sequentially, one in each step m . We show that if the state $\rho(m)$ after step m

maximally reveals contextuality, then so does the state $\rho(m-1)$ after step $m-1$. In other words, 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. Thus, 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 $T_c \in \mathcal{O}$. Let the witness for the state $\rho(m)$ be constructed from the isotropic subspace spanned by $\{\mathbf{a}(m), \mathbf{b}(m)\}$, such that $\mathcal{W}_{\rho(m)}^{\{\mathbf{a}(m), \mathbf{b}(m)\}}(\eta) = -2$, for some η . There are two subcases to consider.

Case ii(a): T_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, $\rho(m-1)$ maximally reveals contextuality.

Case ii(b): T_c does not commute with both $T_{\mathbf{a}(m)}$ and $T_{\mathbf{b}(m)}$. Then, T_c anticommutes 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 assumes that T_c anticommutes 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 result 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 [16] provides a beautifully simple proof of the Kochen-Specker theorem [3] in dimension four and higher, but for 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 [19] for which *all* two-qubit states come out contextual [6]. 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 dimensions [26] and therefore was not an issue in Ref. [6]. 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.

$$\begin{array}{ccc}
 X_1 & X_2 & XX \\
 \hline
 Z_2 & Z_1 & ZZ \\
 \hline
 XZ & ZX & -YY
 \end{array} \quad . \quad (44)$$

Since, as in qudit QCSI, 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 not, in fact, void the contextuality-as-resource viewpoint.

To do so, we revisit the results established in Sec. V. First, by Theorem 3, states with a non-negative Wigner function are noncontextual. 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 included 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, $X_1 \cdot X_2 \cdot X_1 X_2 = +I$, etc.

For the bottom row, the included 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 Sec. V, these observables cannot be simultaneously measured in rebit QCSI. Therefore, the predetermined measurement outcomes $\lambda(XZ)$, $\lambda(ZX)$, and $\lambda(-YY)$ 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 with respect to rebit QCSI. The algebraic contradiction vanishes because we have effectively removed the bottom row from the diagram (44).

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

Generalizing the above observation, the phenomenon of state-independent contextuality does not come into play for the present setting of rebit QCSI, even if it does exist for systems of rebits. The reason for this is the restriction of the physical 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, $\lambda_{\mathbf{u}}: \mathcal{A} \rightarrow \{\pm 1\}$, $\forall \mathbf{u} \in \mathcal{S}$, is nonempty. Thus, there is no state-independent contextuality in rebit QCSI. Contextuality may persist at the level of probability.

Proof of Lemma 17.—The value assignments $\lambda_{\mathbf{u}}: \mathcal{A} \rightarrow \{\pm 1\}$, $\mathbf{u} \in \mathbb{Z}_2^n$, 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 $\mathbf{u} \in \mathbb{Z}_2^n$. The value assignments $\lambda_{\mathbf{u}}(\cdot)$ are thus consistent with the operator constraints. Hence, $\mathbb{Z}_2^n \subseteq \mathcal{S}$, and $\mathcal{S} \neq \emptyset$. ■

Finally, in Sec. VC, we argued that the contextuality witnesses \mathcal{W} are closely related to state-independent contextuality, cf. Remark 4. This relation is not in contradiction to the above statement that there is no state-independent contextuality in rebit QCSI. Namely, Remark 4 refers to rebit quantum mechanics (without the CSS restriction), not to rebit QCSI (with CSS restriction).

Let us revisit three facts from the preceding discussion. Assume two vectors $\mathbf{a}, \mathbf{b} \in V_{\mathcal{A}}$, $[\mathbf{a}, \mathbf{b}] = 0$, such that $T_{\mathbf{a}+\mathbf{b}} = -T_{\mathbf{a}}T_{\mathbf{b}}$. Then, (i) (and only then) the witness $\mathcal{W}^{\{\mathbf{a}, \mathbf{b}\}}$ can detect state-dependent contextuality. (ii) A state-independent contextuality proof can be constructed from $T_{\mathbf{a}+\mathbf{b}}$, $T_{\mathbf{a}}$, and $T_{\mathbf{b}}$ and local Pauli operators X_i, Z_i . (iii) The observables $T_{\mathbf{a}}$ and $T_{\mathbf{b}}$ cannot be simultaneously measured in rebit QCSI (cf. Lemma 13).

Now we note, in addition, that (iv) the states ρ stabilized by such $\pm T_{\mathbf{a}}$ and $\pm T_{\mathbf{b}}$ make the witness $\mathcal{W}_{\rho}^{\{\mathbf{a}, \mathbf{b}\}}(\mathbf{x})$ maximally negative, i.e., $\mathcal{W}_{\rho}^{\{\mathbf{a}, \mathbf{b}\}}(\mathbf{x}) = -2$ for suitable values of \mathbf{x} . See Eq. (43) for an example. The sets of commuting real Pauli operators that cannot be jointly measured in rebit QCSI thus become stabilizer generators for states that maximally violate noncontextuality. The power of contextuality is transferred from measurement to (magic) states, exactly as it should be in a scheme of QCSI.

C. CSS vs real Clifford transformations

It is instructive to examine what happens if the restricted gate set of rebit QCSI is extended from the CSS-ness preserving Clifford operations to the larger set of real Clifford operations. This comprises, in particular, increasing the set of physically measurable observables from \mathcal{O} to the larger set \mathcal{A} of all real Pauli observables.

This change opens the door to state-independent contextuality; see Mermin’s square in Eq. (44). As discussed in the previous section, this is not compatible with viewing contextuality as a resource possessed only by special quantum states.

As for the Wigner function in relation to contextuality, for $n \geq 2$ rebits, a positive Wigner function no longer yields a noncontextual hidden variable model. Namely, with the new measurement contexts available, the “predetermined”

measurement outcomes $\lambda_{\mathbf{u}}$ assigned by the Wigner function via Eq. (31) to HVM states $\mathbf{u} \in V$ fail to satisfy Eq. (28) in Definition 1 of a noncontextual HVM.

As for the Wigner function in relation to efficient sampling, (i) W produces the correct quantum mechanical expectation values for all observables in \mathcal{A} via Eq. (14). Hence, it also produces the correct expectation values for all observables defined for real states. (ii) If $W_{\rho} > 0$, then for all $A \in \mathcal{A}$, the expectation values $\langle A \rangle_{\rho}$ can be efficiently estimated by sampling. However, (iii) allowing measurements in the middle of the computation, those in $\mathcal{A} \setminus \mathcal{O}$ can introduce negativity into the Wigner function, preventing efficient sampling from it.

To summarize, as an example for tinkering with rebit QCSI, if the restricted gate set of CSS-ness preserving Clifford operations is replaced by the broader class of real Clifford operations, then contextuality is undone as a resource, the link between positive Wigner functions and noncontextual hidden variable models breaks, and efficient classical simulation, by sampling from the Wigner function, of rebit QCSI without magic states is obstructed (other simulation methods [13,27] remain, though). These observations illustrate the intricate relation among the various constituents of rebit QCSI.

VII. CONCLUSION

We have established that contextuality and Wigner function negativity are necessary resources for computational universality of the discussed scheme of quantum computation by state injection on rebits. To this end, we have constructed the computational scheme itself and supplemented it with a matching Wigner function (complete with Hudson’s theorem and an efficient sampling algorithm for positive Wigner functions) and contextuality witnesses. These parts mutually reinforce each other: Efficient sampling provides operational justification for calling states with positive Wigner functions “classical,” and Hudson’s theorem ensures that the notions of classicality established by the Wigner function and by the gate restrictions in rebit QCSI match. The absence of the Wigner function negativity and contextuality reveals the limitations of the restricted gate set. Furthermore, our computational scheme is constructed in such a way that state-independent contextuality does not come into play, even if it is present in rebits.

We have thus extended all the essential properties that held for QCSI in the case of qudits of odd prime dimension [6,8] to rebits, with the sole exception that for the present rebit scheme, Wigner function negativity does not imply contextuality.

To widen the scope of the discussion, we note that contextuality has also been established as a resource for measurement-based quantum computation (MBQC) [28–30]. Namely, in MBQC, contextuality is necessary for the ability to compute nonlinear Boolean functions.

One may thus want to compare the roles played by contextuality in QCSI and MBQC. But there is an obstacle: The MBQC result has, to date, only been established for the case of two-level systems, where most of the existing results [6] 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.

- (i) 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?
- (ii) Because of 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?
- (iii) 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 [21]. There is more complicated lattice surgery by which, in addition, the Hadamard gate can be realized [31,32]. 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” or nonmagic states. Is there a Wigner function with matching Hudson’s theorem and covariance property?

ACKNOWLEDGMENTS

R. R. thanks Rob Spekkens for discussion. N. D. is supported by the Lockheed Martin Corporation. P. A. G. is supported by Fonds de Recherche du Québec-Nature et Technologies (FRQNT), and R. R. is supported by Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institute for Advanced Research (Cifar), and IARPA.

APPENDIX A: FOURIER TRANSFORM ON THE GROUP \mathbb{Z}_2^n

The Fourier transform of a function $f: M \rightarrow \mathbb{R}$ 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 $f: M \rightarrow \mathbb{R}$ 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. \blacksquare

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 $(A, B) = (1/2^n) \text{Tr}(A^T B)$.

Proof of Lemma 19.—Denote by $E_{i,j}$ the matrix with entry 0 everywhere except 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 $(1/2^n) \text{Tr}(T_u T_v) = 0$ 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. \blacksquare

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

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

Proof of Lemma 20.—We expand the right-hand side of the above equation by inserting the definition Eq. (11) of W_ρ and Eq. (13) for $A_{\mathbf{u}}$, and we obtain

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

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

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

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

Proof of Property 4—(Sec. IVA). With the definition Eq. (12),

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

and thus

$$A_0 = 2^{[(n/2)-1]} (|0_n\rangle \langle +_n| + |+_n\rangle \langle 0_n|). \quad (\text{B1})$$

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 nontrivially only on system A (B). Then, $T_{\mathbf{v}} = \pm T_{\mathbf{v}_A} T_{\mathbf{v}_B}$ and

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

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

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

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

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

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

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

as stated in Lemma 4. \blacksquare

APPENDIX C: CSS-NESS PRESERVING CLIFFORD GATES

Here, we prove Lemma 1 from Sec. [18], and Lemmas 9 and 10 from Sec. 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 $g \in G_{\text{CSS}}$ and let φ_g be the automorphism of the real Pauli group $P_n(\mathbb{R})$ defined by conjugation by g :

$$\varphi_g: P_n(\mathbb{R}) \longrightarrow P_n(\mathbb{R}) Q \mapsto g Q g^\dagger. \quad (\text{C1})$$

This morphism of $P_n(\mathbb{R})$ induces a morphism of its quotient $P_n(\mathbb{R})/\{\pm I\}$, 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, $F \in GL_{2n}(\mathbb{Z}_2)$. Moreover, the conjugation preserves 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 $F \in \text{Sp}_{2n}(\mathbb{Z}_2)$. \blacksquare

Proof of Lemma 10.—Consider a pair $g, g' \in G_{\text{CSS}}$ 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

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

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 $g, g' \in G_{\text{CSS}}$. \blacksquare

Proof of Lemma 1.—Our first 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 $\mathcal{O} = \{Z(\mathbf{u}) | \mathbf{u} \in \mathbb{Z}_2^n\} \cup \{X(\mathbf{v}) | \mathbf{v} \in \mathbb{Z}_2^n\}$

of Pauli observables that 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 show that an operator g that 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, and 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 $M \subset M_Z \oplus M_X$, and S' is a CSS code if and only if we have equality $M = M_Z \oplus M_X$; 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 anticommute. 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

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

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 groups is preserved by conjugation by g . ■

We now return to the subject of Lemma 9 and further characterize the matrices F appearing on the right-hand

side of Eq. (23). These matrices have one of the following two block structures:

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

where $F_Z, F_X \in GL_n(\mathbb{Z}_2)$ 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 $(F, \mathbf{x}) \in F_{\text{CSS}} \times \mathbb{Z}_2^{2n}$.

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 following two 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 $F_Z, F_X \in GL_n(\mathbb{Z}_2)$. Finally, $F_Z = (F_X^{-1})^t$ 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 $\mathbf{a} \mapsto (-1)^{\mathbf{x} \cdot \mathbf{a}}$ 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

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

The above implies that λ coincides with the character μ on the set $V_{\mathcal{A}}$, which means that $\lambda(\mathbf{a}) = (-1)^{\mathbf{x} \cdot \mathbf{a}}$ 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 $g \in G_{\text{CSS}}$ of special interest.

(i) If $g = \otimes_i H_i$ then

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

(ii) If $g = \text{CNOT}(i, j)$, then

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

The matrix $E_{i,j}$ denotes the $n \times n$ binary matrix whose only nonzero coefficient is in position (i, j) .

(iii) If $g = T_{\mathbf{u}}$, then

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

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

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

Recall that this application \mathcal{F} is well defined by uniqueness 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 $\text{Im}\mathcal{F}$. 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 $\text{SL}_n(\mathbb{Z}_2)$, which coincides with $\text{GL}_n(\mathbb{Z}_2)$. This implies that $\text{Im}\mathcal{F}$ contains all the $A(F, 0)$ associated with

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

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

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

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

$$G_{\text{CSS}}/\text{Ker}\mathcal{F} \simeq \text{Im}\mathcal{F}. \quad (\text{C7})$$

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

The affine group $\text{AGL}(\mathbb{Z}_2)$ is known to be the semidirect product of the group of translations by the general linear group. We obtain a similar structure for $\text{Im}\mathcal{F}$, which is the semidirect product of the group of translations $A(0, \mathbf{t})$ by F_{CSS} , implying

$$\text{Im}\mathcal{F} \simeq \mathbb{Z}_2^{2n} \rtimes F_{\text{CSS}}. \quad (\text{C8})$$

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

By definition, the group F_{CSS} can also be decomposed as a semidirect product

$$F_{\text{CSS}} \simeq \text{GL}_n(\mathbb{Z}_2) \rtimes \mathbb{Z}_2. \quad (\text{C9})$$

The set of block-diagonal matrices of F_{CSS} is a subgroup isomorphic to $\text{GL}_n(\mathbb{Z}_2)$, 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 F of Eq. (C3) of order 2, and it is isomorphic to \mathbb{Z}_2 . The second item of Lemma 1 follows from Eqs. (C7), (C8), and (C9). ■

-
- [1] S. Bravyi and A. Kitaev, *Universal Quantum Computation with Ideal Clifford Gates and Noisy Ancillas*, *Phys. Rev. A* **71**, 022316 (2005).
 - [2] J. S. Bell, *On the Problem of Hidden Variables in Quantum Mechanics*, *Rev. Mod. Phys.* **38**, 447 (1966).
 - [3] S. Kochen and E. P. Specker, *The Problem of Hidden Variables in Quantum Mechanics*, *J. Math. Mech.* **17**, 59 (1967).
 - [4] S. Abramsky and A. Brandenburger, *The Sheaf-Theoretic Structure of Non-locality and Contextuality*, *New J. Phys.* **13**, 113036 (2011).
 - [5] A. Acín, T. Fritz, A. Leverrier, and A. B. Sainz, *A Combinatorial Approach to Nonlocality and Contextuality*, arXiv:1212.4084.
 - [6] M. Howard, J. Wallman, V. Veitch, and J. Emerson, *Contextuality Supplies the “Magic” for Quantum Computation*, *Nature (London)* **510**, 351 (2014).
 - [7] E. F. Galvão, *Discrete Wigner Functions and Quantum Computational Speedup*, *Phys. Rev. A* **71**, 042302 (2005).
 - [8] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, *Negative Quasi-probability as a Resource for Quantum Computation*, *New J. Phys.* **14**, 113011 (2012).
 - [9] E. Wigner, *On the Quantum Correction for Thermodynamic Equilibrium*, *Phys. Rev.* **40**, 749 (1932).

- [10] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, *Discrete Phase Space Based on Finite Fields*, *Phys. Rev. A* **70**, 062101 (2004).
- [11] D. Gross, *Computational Power of Quantum Many-Body States and Some Results on Discrete Phase Spaces*, Ph.D thesis, Imperial College London, 2005.
- [12] R. W. Spekkens, *Negativity and Contextuality Are Equivalent Notions of Nonclassicality*, *Phys. Rev. Lett.* **101**, 020401 (2008).
- [13] D. Gottesman, *Stabilizer Codes and Quantum Error Correction*, Ph.D thesis, California Institute of Technology, 1997.
- [14] A. Mari and J. Eisert, *Positive Wigner Functions Render Classical Simulation of Quantum Computation Efficient*, *Phys. Rev. Lett.* **109**, 230503 (2012).
- [15] C. Cormick, E. F. Galvão, D. Gottesman, J. P. Paz, and A. O. Pittenger, *Classicality in Discrete Wigner Functions*, *Phys. Rev. A* **73**, 012301 (2006).
- [16] N. D. Mermin, *Hidden Variables and the Two Theorems of John Bell*, *Rev. Mod. Phys.* **65**, 803 (1993).
- [17] T. Rudolph and L. Grover, *A 2 Rebit Gate Universal for Quantum Computing*, [arXiv:quant-ph/0210187](https://arxiv.org/abs/quant-ph/0210187).
- [18] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, *Quantum Error Correction and Orthogonal Geometry*, *Phys. Rev. Lett.* **78**, 405 (1997).
- [19] A. Cabello, S. Severini, and A. Winter, *Graph-Theoretic Approach to Quantum Correlations*, *Phys. Rev. Lett.* **112**, 040401 (2014).
- [20] A. Kitaev, *Fault-Tolerant Quantum Computation by Anyons*, *Ann. Phys. (N.Y.)* **303**, 2 (2003).
- [21] R. Raussendorf and J. Harrington, *Fault-Tolerant Quantum Computation with High Threshold in Two Dimensions*, *Phys. Rev. Lett.* **98**, 190504 (2007).
- [22] M. A. Nielsen and I. L. Chuang, *Quantum Information and Computation* (Cambridge University Press, Cambridge, England, 2000).
- [23] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, *The Resource Theory of Stabilizer Computation*, *New J. Phys.* **16**, 013009 (2014).
- [24] R. L. Hudson, *When Is the Wigner Quasi-Probability Density Non-Negative?*, *Rep. Math. Phys.* **6**, 249 (1974).
- [25] J. J. Wallman and S. D. Bartlett, *Nonnegative Subtheories and Quasiprobability Representations of Qubits*, *Phys. Rev. A* **85**, 062121 (2012).
- [26] M. Howard, E. Brennan, and J. Vala, *Quantum Contextuality with Stabilizer States*, *Entropy* **15**, 2340 (2013).
- [27] J. Bermejo-Vega and M. Van den Nest, *Classical Simulations of Abelian-Group Normalizer Circuits with Intermediate Measurements*, *Int. J Quantum Inf. Comput.* **14**, 181 (2014).
- [28] J. Anders and D. E. Browne, *Computational Power of Correlations*, *Phys. Rev. Lett.* **102**, 050502 (2009).
- [29] M. J. Hoban and D. E. Browne, *Stronger Quantum Correlations with Loophole-Free Postselection*, *Phys. Rev. Lett.* **107**, 120402 (2011).
- [30] R. Raussendorf, *Contextuality in Measurement-Based Quantum Computation*, *Phys. Rev. A* **88**, 022322 (2013).
- [31] H. Bombin, *Topological Order with a Twist: Ising Anyons from an Abelian Model*, *Phys. Rev. Lett.* **105**, 030403 (2010).
- [32] A. G. Fowler, *Low-Overhead Surface Code Logical Hadamard*, *Quantum Inf. Comput.* **12**, 970 (2012).