Quantum physics can be discriminated from classical physics using Bell-type inequalities [1]. In particular, the violation of Bell inequalities for two qubits has been extensively verified since they were first checked in atomic physics experiments [2,3]. Later on, the improvement of quantum optics techniques as well as other technologies such as nitrogen-vacancy (NV) centers has made it possible to eliminate many of the loopholes in the experimental verification of two-qubit Bell inequalities [3].

An extension of Bell inequalities to a larger number of particles corresponds to the set of Mermin inequalities [4]. Such inequalities should be maximally violated by Greenberger-Horne-Zeilinger (GHZ)-type states [5]. The experimental verification of multipartite Mermin inequalities faces the problem of a good control of three or more qubits, including the generation of entangled states, and the possibility of performing different measurements on each one. Violation of Mermin inequalities has been reported for three qubits [6] and four qubits [7], where all qubits are made out of photons, and for up to 14 qubits with a quantum computer based on ion traps [8].

In the case of superconducting qubits, violation of the Clauser-Horne-Shimony-Holt (CHSH) inequality was achieved in Ref. [9], whereas the GHZ construction and the three-qubit Mermin inequality violation was demonstrated in Ref. [10]. For a general review of theoretical and experimental progress in Bell inequalities, see Ref. [11].

The construction of the first prototypes of quantum computers allows for the possibility of experimenting with quantum states containing more than two qubits. In particular, IBM has opened the use of its five-qubit quantum computer to the community [12]. We here report results on the use of this quantum computer to test the violation of Mermin inequalities for three, four, and five superconducting qubits.

I. MERMIN POLYNOMIALS

Local realism can be tested using Mermin polynomials. The technique to generate them is explained for example in Ref. [13]. The Mermin polynomial for three qubits is

\[ M_3 = (a_1 a_2 a_3 + a_1 a_2 a_3 + a_1 a_2 a_3) - (a'_1 a'_2 a'_3), \]  

(1)

where \( a_i \) and \( a'_i \) correspond to two different settings for the measurement of each qubit \( i \). Each measurement can take the values \({\pm 1} \). Classical theories obey local realism (LR) which translates into a bound for the expectation value of the Mermin polynomial, \( \langle M_3 \rangle_{LR} \leq 2 \). Instead, for quantum mechanics (QM) the observables \( a_i \) and \( a'_i \) are built out of linear combinations of Pauli matrices. Each measurement is expressed as a Kronecker product of the three local measurements and the expectation value for \( \langle M_3 \rangle \) is the maximum eigenvalue of the resulting \( 8 \times 8 \) matrix. In this case, the maximum possible eigenvalue, and therefore the quantum bound, is \( \langle M_3 \rangle_{QM} \leq 4 \). We briefly construct circuits to check the violation of the classical bound on this inequality.

The Mermin polynomial that will be experimentally checked for four-qubits is

\[ M_4 = -(a_1 a_2 a_3 a_4) + (a_1 a_2 a_3 a'_4 + a_1 a_2 a'_3 a_4 + a_1 a'_2 a_3 a_4 + a_1 a'_2 a_3 a'_4 + a_1 a'_2 a_3 a'_4) \\
+ (a'_1 a_2 a_3 a_4) + (a'_1 a_2 a'_3 a_4 + a'_1 a'_2 a_3 a'_4 + a_1 a'_2 a_3 a_4 + a_1 a'_2 a'_3 a_4) \\
+ (a'_1 a_2 a_3 a_4) + (a'_1 a_2 a'_3 a_4 + a'_1 a'_2 a_3 a'_4 - a_1 a'_2 a_3 a'_4) \\
+ (a'_1 a_2 a_3 a_4) + (a'_1 a_2 a'_3 a_4 - a_1 a'_2 a_3 a'_4), \]

(2)

with a classical bound of \( \langle M_4 \rangle_{LR} \leq 4 \) and a quantum bound of \( \langle M_4 \rangle_{QM} \leq 8/\sqrt{2} \).

In the five-qubit case, the Mermin polynomial reads

\[ M_5 = -(a_1 a_2 a_3 a_4 a_5) + (a_1 a_2 a_3 a'_4 a'_5 + a_1 a_2 a'_3 a_4 a'_5 + a_1 a'_2 a_3 a_4 a'_5 + a_1 a'_2 a_3 a'_4 a'_5) \\
+ (a'_1 a_2 a_3 a_4 a'_5) + (a'_1 a_2 a'_3 a_4 a'_5 + a'_1 a'_2 a_3 a_4 a'_5) \\
+ (a'_1 a_2 a_3 a_4 a_5) + (a'_1 a_2 a'_3 a_4 a_5 + a_1 a'_2 a_3 a_4 a_5 + a'_1 a'_2 a_3 a'_4 a_5 - (a_1 a'_2 a_3 a'_4 a'_5)) \\
+ (a'_1 a_2 a_3 a'_4 a'_5) + (a'_1 a_2 a'_3 a'_4 a'_5 + a'_1 a'_2 a_3 a'_4 a'_5), \]

(3)

with a classical bound of \( \langle M_5 \rangle_{LR} \leq 4 \) and a quantum bound of \( \langle M_5 \rangle_{QM} \leq 16 \).

II. CIRCUIT IMPLEMENTATION

There are a number of technical issues associated with the specific implementation of the IBM five-qubit quantum computer. This quantum computer is based on superconducting flux qubits that live on a fridge with a temperature of about 15 mK, where only one of the qubits can be used to act as...
the target qubit of any controlled-\textsc{not} (CNOT) gate. In the test of Mermin inequalities, only GHZ-like states have to be created. This requires the use of a Hadamard gate on a control qubit followed by \textsc{cnot}s targeted to the rest. In order to implement this kind of action we need to operate \textsc{cnot} gates targeted to other qubits. This can be done using the relation \textsc{cnot}  \rightarrow S_H \textsc{cnot} S_H, where \textsc{cnot} and \textsc{h} are Hadamard gates on qubits 1 and 2, whereas \textsc{cnot} $\rightarrow$ $S_I$ is the controlled-\textsc{not} gate which is controlled by qubit 1.

In our choice of settings, the needed GHZ-like states have relative phases, as in the case of three qubits, where $\ket{\phi} = 1/\sqrt{2} (\ket{000} + i \ket{111})$. These phases are implemented using $S$ and $T$ gates, which are one-qubit gates that multiply the $1$ term with $\pi/2$ and $\pi/4$ phases, respectively. Measurements can only be done on the $s_i$ basis, but they can be simulated in another basis with the help of additional gates, namely an $H$ gate for $s_i$ and an $S_I$ gate followed by an $H$ gate for $s_i$.

Another relevant issue to be considered is that not all of the qubits are equally robust in the present quantum computer, some have relaxation and decoherence times larger than others, although all of them are of the order of $T = O (100 \mu s)$. We adapt our circuits to minimize the number of gates on the qubits that behave more poorly. For example, gates that implement phases that can be put freely in any qubit are allocated to the most robust ones.

Figures 1 and 2 represent the three circuits for the three-, four-, and five-qubit Mermin inequalities. In principle we need to perform as many experiments as the number of terms in the Mermin inequalities (1), (2), and (3). However due to our limited access to the computer and the symmetry of particle exchange of the states and the inequalities, only one experiment for a term representative of each number of primes ($a'_j$) was run. In our choice of settings, the number of primes amounts to the number of $s_i$ measurements, whereas the nonprimes ($a_i$) correspond to $s_i$ measurements. We thus have two experiments for three qubits, five experiments for four qubits, and three experiments for five qubits. Each experiment was run 8192 times, the maximum available, except for the three-qubit experiments, which were run only 1024 times. When computing the expected value of the whole polynomial, each experiment was given the corresponding weight. In the errors discussion we compare results obtained when using the symmetry with results obtained without using it, computing all the terms, for the three-qubit case.

### III. Results

We now give a more detailed discussion of the results for the three-qubit case and an abridged one for the four- and five-qubit cases, as much of it is basically the same.

In order to check the violation of the inequality, one has to choose the settings and the corresponding state that maximally violate it. One possibility is to choose the settings $a_i = s_i$ and $a'_j = s_j$ for all the qubits. The state that maximizes the quantum violation in this case is $\ket{\phi} = 1/\sqrt{2} (\ket{000} + i \ket{111})$.

The three-qubit Mermin inequality has four terms as shown in Eq. (1). In principle, four different circuits are needed, one for each term. The state will be the same for all of them, but the settings change. However, one can use the symmetry of the state and the inequality to reduce the number of measurements needed if there is limited access to the experimental setting as is our case. All the terms that have the same number of primes ($a'_j$) are represented by the same circuit by symmetry. We then considered only two different experiments, with 1024 runs each, the $s_i s_j s_k$ experiment and the $s_i s_j s_k$ experiment. The results are shown in Table I.

![Image 1](https://example.com/image1.png)

**FIG. 1.** The two circuits used for the three-qubit Mermin inequality. The first circuit corresponds to the $s_i s_j s_k$ experiment, and the second circuit corresponds to the $s_i s_j s_k$ experiment. The $S_I$ gates make the difference between a $s_i$ measurement and a $s_j$ measurement.

![Image 2](https://example.com/image2.png)

**FIG. 2.** Two of the circuits used for the four-qubit and five-qubit Mermin inequalities. The first circuit corresponds to the $s_i s_j s_k$ experiment, whereas the second corresponds to the $s_i s_j s_k$ experiment. The $S_I$ gates make the difference between a $s_i$ measurement and a $s_j$ measurement. In order to change from $s_i$ to $s_j$, one has to add an $S_I$ gate, or remove it to do the opposite. With this technique one can obtain all the circuits needed to test the inequalities.

<table>
<thead>
<tr>
<th>Result $XXX$</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.197</td>
<td>0.042</td>
<td>0.024</td>
<td>0.194</td>
<td>0.043</td>
<td>0.203</td>
<td>0.231</td>
<td>0.033</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Result $YYY$</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.050</td>
<td>0.188</td>
<td>0.028</td>
<td>0.258</td>
<td>0.026</td>
<td>0.041</td>
<td>0.221</td>
<td></td>
</tr>
</tbody>
</table>
We may get an estimation of the statistical error as a dispersion around the mean. We may, as well, treat the results as a multinomial distribution, using the expression $\delta p = \sqrt{p(1 - p)/N}$, which for $N = 8192$ gives $\delta p = O(10^{-2})$. The different Mermin inequalities for three, four, and five qubits require a different number of experiments to be done, which are considered as independent. We may then add in quadrature its errors, which is the figure we associate with the explicit results. In this sense, the five-qubit result obtained with the present quantum computer does not have sufficient statistical significance to discard local realism.

Furthermore, some of the issues related to the elimination of loopholes cannot be addressed. Experiments suffer from errors related to stability, loss of coherence, and lack of full fidelity of the quantum gates. This is clearly seen as the violation of Mermin inequalities deteriorate progressively as the numbers of qubits, and gates used in the experiment, increase. We may think of the experimental verification of Mermin inequalities as a test of the overall fidelity of the whole Mermin circuit.

### IV. CONCLUSIONS

Experimental verification of Mermin inequalities for three, four, and five qubits has been tested on a five-qubit IBM quantum computer. Results do show violation of local realism in all cases, with a clear degradation in quality as the number of qubits (and needed gates) increases. Nonetheless, this produces the first experimental violation of four- and five-qubit Mermin inequalities with superconducting qubits, though the statistical significance of the second one is still poor. It should be noted however that, in the case of the four-qubit inequality, the result shows generic nonlocality but does not provide evidence for genuine four-particle nonlocality, because this would only be implied by $M_4 > 8$. [14]. It can be argued that the measurements of Mermin polynomials for many qubits can be used as a figure of merit to assess the fidelity of a quantum computer.

### ACKNOWLEDGMENTS

D.A. acknowledges financial help from the APIF Scholarship of the University of Barcelona. J.I.L. acknowledges financial support from Grant No. FIS2013-41757-P. We acknowledge use of the IBM Quantum Experience for this work.

The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Quantum Experience team.


