ARTICLE OPEN

Check for updates

Experimental test of non-macrorealistic cat states in the cloud

Huan-Yu Ku [1,2], Neill Lambert [2✉], Feng-Jui Chan[1], Clive Emary[3], Yueh-Nan Chen [1,2✉] and Franco Nori [2,4]

The Leggett–Garg inequality attempts to classify experimental outcomes as arising from one of two possible classes of physical theories: those described by macrorealism (which obey our intuition about how the macroscopic classical world behaves) and those that are not (e.g., quantum theory). The development of cloud-based quantum computing devices enables us to explore the limits of macrorealism. In particular, here we take advantage of the properties of the programmable nature of the IBM quantum experience to observe the violation of the Leggett–Garg inequality (in the form of a 'quantum witness') as a function of the number of constituent systems (qubits), while simultaneously maximizing the 'disconnectivity', a potential measure of macroscopicity, between constituents. Our results show that two- and four-qubit 'cat states' (which have large disconnectivity) are seen to violate the inequality, and hence can be classified as non-macrorealistic. In contrast, a six-qubit cat state does not violate the 'quantum witness' beyond a so-called clumsy invasive-measurement bound, and thus is compatible with 'clumsy macrorealism'. As a comparison, we also consider un-entangled product states with $n = 2, 3, 4$ and 6 qubits, in which the disconnectivity is low.

INTRODUCTION

The availability of public quantum computers, like the 'IBM quantum experience' (IBM QE)[1], promises both applications[2–11] and tests of fundamental physics[12–14]. In particular, as the number of available qubits increases, it potentially allows for a rigorous study of the crossover between classical and quantum worlds[15,16], including tests like the Leggett–Garg inequality (LGI)[17,18]. The LGI was derived as a means to classify experimental outcomes as arising from one of two possible classes of physical theories: those described by macrorealism, and those that are not (e.g., quantum theory).

A macrorealistic theory is one where the system properties are always well-defined (i.e., obey realism), and in which said properties can be observed in a measurement-independent manner (i.e., measurements just reveal pre-existing properties of the system, and do so in a way that does not change those properties). Quantum theory obeys neither of these stipulations, but our intuition about the classical world does. Thus, 'macrorealists' propose that macrorealistic theories apply when the dimension, mass, particle number, or some other indicator of the size of a system is increased, such that the behaviour of suitably macroscopic systems will tend to obey realism and can be observed without disturbance.

Over the last 10 years a large variety of experimental tests of the LGI, and its generalizations[19–34], have been performed and violations observed. In typical tests, such as with photon[35–38] and nuclei–electron spin pairs[39], the macroscopicity of the system has been small. And while larger for superconducting qubits[40–42] and some of the atom-based examples[43,44], testing truly macroscopic systems yet remains a distant goal[45,46].

In this work, we take advantage of the programmable nature of the IBM QE to enable tests of increasing macroscopicity by directly increasing the number of constituent parts of the system in a non-trivial way. To do so we design a circuit that generates n-qubit 'cat state' superpositions of fully polarized configurations, i.e., states

which have genuine multipartite entanglement[47] and a large 'disconnectivity', an indicator of macroscopicity[17,18,48–52]. This allows us to see how the violation of the LGI (here in the form of a 'quantum witness'[23,53,54]) changes as we increase the macroscopicity in terms of the number of constituent qubits.

In addition, we augment the basic quantum witness test with a measurement invasiveness test[41,55], which accounts for 'macroscopically invasive' measurements by modifying the witness bound. We term systems which cannot violate the bound 'clumsy-macrorealistic'. For the experiments we perform on the IBM QE, our tests show that two- and four-qubit 'cat states' clearly violate the quantum witness and are thus non-macrorealistic. On the other hand, as we increase the number of qubits involved in the state to six, the witness value is suppressed, suggesting that this case is compatible with 'clumsy macrorealism'.

Finally, instead of preparing entangled states, we also consider product states with zero entanglement, and hence low disconnectivity (compared with our test using entangled states), which implies these states are less macroscopic. In comparison with the cat states, we observe that the violation of the witness for these states is more robust to decoherence as the number of qubits is increased. We also show that the quantum witness can serve an additional role as a dimensionality (as in the number of states in the Hilbert space discriminated by the intermediate measurement) witness.

RESULTS

Quantum witness

According to Leggett and Garg, macrorealistic systems obey two assumptions: macrorealism per se (MRPS) and non-invasive measurability (NIM)[17,18]. MRPS assumes that the system always exists in a definite macroscopic state, and NIM assumes that measurements reveal what that state is, but do not change it.

---
[1]Department of Physics and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan. [2]Theoretical Quantum Physics Laboratory, RIKEN Cluster for Pioneering Research, Wako-shi, Saitama 351-0198, Japan. [3]Joint Quantum Centre Durham-Newcastle, School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK. [4]Department of Physics, The University of Michigan, Ann Arbor, MI 48109-1040, USA. ✉email: nwlambert@gmail.com; yuehnan@mail.ncku.edu.tw

Under the assumptions of MRPS and NIM, the LGI in the form of a 'quantum witness'[23,53,54] tells us that if we consider measurements on a system at two times, t_1 and t_2, the probability of observing outcome j at time t_2 should be independent of whether the measurement at the earlier time, t_1, was performed or not. This probability is then related to the sum of all joint probabilities in the standard way[53]:

$$p_{t_2}(j) = p_{t_2}^M(j) = \sum_i p_{t_1}(i) p_{t_2,t_1}(j|i). \tag{1}$$

Here, $p_{t_2}(j)$ is the probability of observing the outcome j at time t_2, and $p_{t_1}(i) p_{t_2,t_1}(j|i)$ is the joint probability for observing the measurement outcome i at time t_1 followed by the outcome j at time t_2. The superscript M denotes that a measurement was performed at the earlier time t_1, and conversely the absence of M implicitly denotes the probabilities are collated from experiments where such an earlier measurement was not performed.

Given these definitions, the quantum witness[53] can be defined as the breakdown of the equality $W = 0$ where the witness is defined as

$$W = |p_{t_2}(j) - p_{t_2}^M(j)|. \tag{2}$$

If we find $W \neq 0$, the state at time t_1 is said to be non-macrorealistic, in the sense that the assumptions of either MRPS or NIM (or both) are shown to be invalid for it.

The assumption of NIM is hard to justify, even if we assume MRPS holds. We can modify Eq. (2) to take into account certain types of invasive measurements by allowing the measurement process at time t_1 to change the macroscopic state of the system. In this case, the relationship between marginal and joint probabilities can be extended to[53,56]

$$p_{t_2}^M(j) = \sum_k p_{t_2,t_1}(j|k) \sum_i \epsilon^M(k|i) p_{t_1}(i), \tag{3}$$

which incorporates the probability $\epsilon^M(k|i)$ that observing the system in state i at t_1 can cause the system to change to state k. We dub this the assumption of 'macroscopically invasive measurements'.

As shown in the Methods section, combining this assumption with Eq. (2) gives us an inequality for 'clumsy macrorealism'

$$W \leq \max_i[\mathcal{I}(i)]; \qquad \mathcal{I}(i) = |1 - \epsilon^M(i|i)|. \tag{4}$$

We call $\mathcal{I}(i)$ an invasiveness test, and it can be evaluated in an additional experimental run by preparing the system in state i, performing a measurement, and checking whether it is still in that state immediately after said measurement. If we observe that the inequality in Eq. (4) is violated we can say that the both macrorealism and non-clumsy macrorealism do not hold.

Under the most clumsy of measurements, the clumsy macrorealism bound can be unity if $\epsilon^M(i|i) = 0$, which occurs when the measurement so strongly disturbs the system the given state i is completely changed into some other states $j \neq i$. Therefore, we note that the quantum witness and the measurement invasiveness test should be implemented under the same conditions. One of the goals of the LGI is to identify whether the macroscopic nature of a given system influences whether it behaves in a 'quantum way' or in a macrorealistic fashion. While definitions of macroscopicity are myriad[52], Leggett[48,49] suggested that a minimal starting point are the *extensive difference* and the *disconnectivity*. The former compares the difference in magnitude of the observable outcomes to some fundamental physical scale. Recent experiments have attempted to maximize the extensive difference with a macroscopically large superconducting flux qubit[41].

The disconnectivity, in contrast, arises from considering that a violation of the witness by a quantum system arises because, at time t_2, quantum dynamics can generate superpositions of 'macroscopic states'. Simultaneously, these superpositions are collapsed by measurement, and hence LGI and quantum witness tests are violated. Thus, if an object is composed of many 'particles', we want the macroscopic nature of the system to contribute to the 'superposition of macroscopic states' in a non-trivial way, i.e., a large number of the particles should have different states in the 'branches' of the superposition. For instance, a Bell state $1/\sqrt{2}(|11\rangle + |00\rangle)$ satisfies the above statement with both qubits having clearly different states in the two branches of the superposition, while the product state $1/\sqrt{2}(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)$ does not. This recalls the idea that in a Schrödinger's cat thought experiment, the whole cat is in superposition, not just one whisker.

To put such a definition in a quantitative format, Leggett[48,49] argued that the disconnectivity can be defined as the 'number' of correlations (between constituents) one needs to measure to distinguish a linear superposition between two branches from a mixture (which are indistiguishable with single-particle measurements alone). A potential quantitative measure proposed by Leggett in ref. [49] is as follows: considering n spins, for any integer $n' \leq n$, the reduced von-Neumann entropy (also known as the entanglement entropy, a measure of entanglement for bi-partite pure states) of the state $\rho_{n'}$ (having traced out the other spins) is $S_{n'} = -\text{Tr}\rho_{n'}\ln\rho_{n'}$. Leggett then defined the disconnectivity Γ as the maximum value of n' such that

$$\delta_{n'} = \frac{S_{n'}}{\min_m(S_m + S_{n'-m})} < \eta, \tag{5}$$

where η is a small value that sets the bound between classical mixtures and entangled states (see below). Here one assigns $\delta_{n'} = 1$ when $S_{n'} = \min_m(S_m + S_{n'-m}) = 0$ and defines $\delta_1 = 0$. With this definition one can see that states which are 'globally' pure but locally mixed give large values of disconnectivity, implying the mixed-ness arises from global entanglement.

Considering an n-body pure entangled state like the GHZ state we use in our experiment, we will have a vanishing numerator and non-vanishing denominator, leading to $\delta_n = 0$ and thus $\Gamma = n$. On the other hand, for a product state, or a mixture of product states, one finds $\delta_2 = 1$ and 0.5, respectively, and hence $\Gamma = 1$ for these cases. (As an aside, this suggests a possible choice of $\eta = 0.5$ as a bound to delineate between mixtures of product states and mixed entangled states in Eq. (5).) It is clear that the disconnectivity is strongly related to definitions of genuinely multipartite pure-state entanglement, a connection which is discussed in-depth in refs [48,49]. In the tests of macrorealism performed to date, most are arguably in the regime of $\Gamma = 1$, particularly those employing single photons, electrons, or nuclear spins, and so on[35–37,39]. On the other hand, the question of the disconnectivity of a single superconducting qubit[40–42] has been open to debate (see the Supplementary Information of ref. [41] for an in-depth discussion). Our approach here, irrespective of the disconnectivity of the constituent qubit, provides a way to increase the overall disconnectivity by constructing large cat states of many entangled qubits.

To translate this onto the IBM QE, we identify the macroscopic states i with the n-qubit computational basis states of the quantum register, as revealed by standard read-out measurements. We denote, where appropriate, these macroscopic states with classical bit-strings, such that what would be $|0\rangle^{\otimes n}$ in the bra-ket notation we write as $\{0\}^n$.

To generate superposition states with high disconnectivity, we design circuits in the IBM QE to produce an evolution which starts with all qubits in the product state $|0\rangle^{\otimes n}$ at time t_0, and then ideally implements a unitary $U(n, \theta)$ that creates an entangled n-qubit 'cat state' at 'time t_1', namely

$$\rho_{t_1} = |\phi(n,\theta)\rangle\langle\phi(n,\theta)|, \tag{6}$$

where $|\phi(n,\theta)\rangle = \cos\frac{\theta}{2}|0\rangle^{\otimes n} + \sin\frac{\theta}{2}|1\rangle^{\otimes n}$, with real coefficient θ (which for $\theta = \pi/2$ and $n > 2$ are GHZ states). According to the
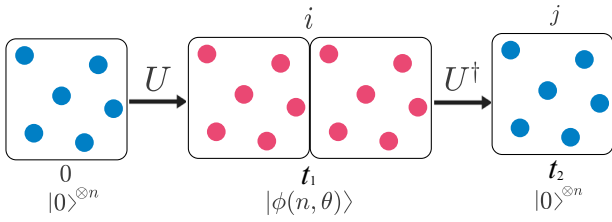
Fig. 1 Schematic setup. We prepare n qubits on the state $|0\rangle^{\otimes n}$ (blue) at time t_0. A unitary U transfers the system into the entangled cat state $|\phi(n,\theta)\rangle = \cos\frac{\theta}{2}|0\rangle^{\otimes n} + \sin\frac{\theta}{2}|1\rangle^{\otimes n}$ (red) at time t_1. Then, an inverse unitary $U^\dagger$ is performed to the entangled system, such that the system returns back to the state $|0\rangle^{\otimes n}$ at time t_2. The outcomes i and j are obtained at t_1 and t_2, respectively.

witness prescription, a measurement is then either performed or not performed on this state. We then choose the evolution for t_1 to t_2 to be given by the inverse unitary transformation $U^\dagger$ so that, in the situation where no intermediate measurement has been performed, the entangled state is, ideally, 'evolved back' to the starting state and $\rho_{t_2} = |0\rangle\langle 0|^{\otimes n}$ (see Fig. 1 for a schematic description). In the witness itself we choose to only look at the probability of being in that particular macroscopic state, i.e., $j \equiv \{0\}^n$ at t_2.

In an ideal quantum system, undergoing evolution described by Eq. (6), one can trivially calculate that $p_{t_2}(\{0\}^n) = 1$, $p_{t_1}(\{0\}^n) = \cos^2(\frac{\theta}{2})$, $p_{t_1}(\{1\}^n) = \sin^2(\frac{\theta}{2})$, $p_{t_2,t_1}(\{0\}^n|\{0\}^n) = \cos^2(\frac{\theta}{2})$ and $p_{t_2,t_1}(\{0\}^n|\{1\}^n) = \sin^2(\frac{\theta}{2})$. The corresponding quantum witness evaluates as

$$W = 1 - \cos^4\left(\frac{\theta}{2}\right) - \sin^4\left(\frac{\theta}{2}\right). \qquad (7)$$

Equation (7) shows that the quantum witness in Eq. (2) is violated for any $\theta \neq k\pi$ with $k = 0, 1, 2, \ldots$.

It is interesting to compare the situation described in Eq. (6) to an example which still uses many qubits but has low disconnectivity. In this case, the maximally entangled states we used previously are now replaced by a product of single-qubit superposition states at time t_1,

$$|\psi(n)\rangle = \frac{1}{\sqrt{2^n}}(|0\rangle + |1\rangle)^{\otimes n}, \qquad (8)$$

which has the lowest disconnectivity[48,49] of $\Gamma = 1$. Surprisingly, this product state saturates the maximum quantum bound of the witness, which is given by[57]

$$W_{\text{max}} = 1 - \frac{1}{D_{\text{Ideal}}}, \qquad (9)$$

where D_{Ideal} is the number of states spanning the Hilbert space, obtained when the intermediate measurement process discriminates all D_{Ideal} states. We note that this bound is derived under the assumptions of quantum mechanics (not clumsy macrorealism). As an aside, we mention that combining Eq. (9) with the clumsy-macrorealistic bound in Eq. (4), we can obtain that we need $\epsilon^M(i|i) \geq [D_{\text{Ideal}}]^{-1}$ for the possibility of our system to violate clumsy macrorealism at all.

In our experiments, since we individually measure every qubit in the computational basis, $D_{\text{Ideal}} = 2^n$ with n number of qubits. Because of this relation between the maximum violation and the dimensionality of the states contributing to the quantum witness, a secondary application as a dimensionality witness arises. In our previous example of cat states, even though we had many qubits, and high disconnectivity, the effective dimension of the states involved in the test of the witness was low, because it was dominated by just two states, $|0\rangle^{\otimes n}$ and $|1\rangle^{\otimes n}$.

Circuit implementation

We use the processor IBM Q5 Tenerife to experimentally test the $n = 2$ cat state with $\theta \in \{0, \pi/8, 2\pi/8, 3\pi/8, 4\pi/8\}$. With the 14-qubit processor IBM Q14 Melbourne, the GHZ states are implemented by considering $\theta = \pi/2$ for $n = 4$ and 6. The IBM QE only allows for a single measurement to be performed on each qubit. This makes collating the two-time correlation functions required by the quantum witness difficult. To overcome this restriction, we use CNOT gates between 'system qubits' and 'ancilla qubits', followed by measurements on the ancilla qubits, to perform the intermediate measurement. As a consequence, we are restricted to a maximal qubit number of 6, with correspondingly 6 ancilla qubits for measurements at time t_1.

From the initial state $|0\rangle^{\otimes n}$, 'cat states' can be obtained by performing the unitary transformation U. The unitary U can be decomposed into several parts. It contains the following single-qubit operation applied to the first qubit:

$$U_3(\lambda, \vartheta, \theta) = \begin{pmatrix} \cos\theta/2 & -e^{i\lambda}\sin\theta/2 \\ e^{i\vartheta}\sin\theta/2 & e^{i(\lambda+\vartheta)}\cos\theta/2 \end{pmatrix}, \qquad (10)$$

with $\lambda = \vartheta = 0$, followed subsequently by a series of CNOT gates between the first qubit and each of the others in turn. The inverse operation $U^\dagger$ is given by applying CNOT gates again before applying the $U_3^\dagger = U_3(0, 0, -\theta)$ gate on the first qubit. In the Methods section, we presents a detailed schematic example for a two-qubit system as well as the details of how these operations function.

To generate the product of superposition states described in Eq. (8), we perform Hadamard gates on each qubit individually at t_1. At time t_2, the Hadamard gates, which are self-inverse, are performed to again obtain a state $|0\rangle^{\otimes n}$.

Noise simulation

Before discussing the experimental results, we introduce a numerical noise model which will assist in understanding two of the main experimental features: suppression of the witness violation due to dephasing and accidental enhancement of the witness violation due to macroscopically invasive measurements.

In the following, to include the influence of decoherence and gate infidelities in a simulation of the quantum circuit, we consider a simple strategy where we assume that each gate is performed perfectly, and instantaneously, after which follows a period of noisy evolution for the prescribed gate time (which can be substantial for two-qubit gates). During such periods, the dynamics of the system can be described by the following Lindblad master equation[58,59]:

$$\begin{aligned} \dot{\rho}_s = & \sum_i^N \frac{\gamma_{T_1}^i}{2}\left[2\sigma_-^i\rho_s(t)\sigma_+^i - \sigma_+^i\sigma_-^i\rho_s(t) - \rho_s(t)\sigma_+^i\sigma_-^i\right] \\ & + \sum_i^N \frac{\gamma_{T_2}^i}{2}\left[2\sigma_z^i\rho_s(t)\sigma_z^i - \sigma_z^{i2}\rho_s(t) - \rho_s(t)\sigma_z^{i2}\right]. \end{aligned} \qquad (11)$$

where σ_+^i, σ_-^i, σ_z^i represent the raising, lowering and Pauli-Z operators of the ith qubit, respectively. Here, we consider the coefficients to be uniformly $\gamma_{T_1}^i = 1/T_1$ and $\gamma_{T_2}^i = [T_2^{-1} - (2T_1)^{-1}]/2 \, \forall i$. The parameters above (like the energy relaxation time T_1, dephasing time T_2, and gate times) are publicly available and can also be reconstructed by the user using well-defined protocols in the IBM QE.

Equation (11) can be derived by assuming that the influence of the environment obeys the standard Born–Markov–Secular approximations. The first and second lines respectively describe energy dissipation and pure dephasing. For the comparison to the experimental data we use values for γ_{T_1} and γ_{T_2} which approximately fit the order of magnitude of the published data

($T_1 = 46\,\mu s$ and $T_2 = 13.5\,\mu s$) such that the numerical simulations approximate the observed experimental results well.

Although in general the noise suppresses the violation of the quantum witness, we will show later that the experimental result with $\theta = 0$, in which the state has no superposition, is not exactly 0 as one might expect. From our numerical simulation we find that this non-zero value is due to imperfect gate operations; in particular, the ancilla CNOT gates for the intermediate measurement. For example, during the intermediate measurement step, for $\theta = 0$ we ideally expect the system to remain in the $|0\rangle^{\otimes n}$ state for the whole duration of the experiment. However, during the intermediate measurement state the system may be accidentally excited from state $|0\rangle^{\otimes n}$ into other states. This is precisely a 'clumsy' macroscopically invasive measurement, such that $W \neq 0$, even though the state of the system obeys MRPS. With the effective identity operations $U = U^\dagger = \mathrm{id}$ in $\theta = 0$ case, this example exactly corresponds to the measurement invasiveness test we described in the section 'Quantum witness'.

In our minimal simulation, we simply model the effects of such errors by the following extra Lindblad terms for all qubits i:

$$\frac{\gamma_{\mathrm{Errors}}^i}{2}\left[2\sigma_+^i\rho_s(t)\sigma_-^i - \sigma_-^i\sigma_+^i\rho_s(t) - \rho_s(t)\sigma_-^i\sigma_+^i\right], \qquad (12)$$

where $\gamma_{\mathrm{Errors}}^i$ is the coefficient to simulate the gate errors for each qubit (which again we take to be uniform $\gamma_{\mathrm{Errors}}^i = \gamma_{\mathrm{Errors}}$). This noise term is phenomenological, and is introduced to capture the noisy imperfect nature of the intermediate measurements which can cause an effective excitation of the qubits, as described above. For our quantum circuits, we determine this value ($\gamma_{\mathrm{Errors}} = 8.5 \times 10^{-2}\,\mu s^{-1}$).

While we primarily use this parameter to fit the witness violation, we point out that when $\theta = 0$, the circuit implementation of the quantum witness is identical to the clumsy-measurement test because the intermediate state is simply $\{0\}^2$ in the computational basis. Thus, the noisy simulation of the parameter γ_{Errors} estimates not only the quantum witness but also the measurement invasiveness in that round of the experiment (see Fig. 1).

Experimental results

To evaluate our proposed modified bound on the witness, we must run additional experiments wherein one prepares and measures all possible quantities $\mathcal{I}(i)$. Although we do not explicitly test the quantum witness for a single-qubit, we do test the measurement invasiveness for this case, to check the effect of the clumsy measurement in the IBM QE. The average and maximum values of the invasiveness of single, two, four and six qubits are shown in Table 1. While we only present the invasiveness of the state $\{0\}^n$, as it usually presents the largest disturbance, we do prepare 'all possible states' in the computational basis for single-, two-, four-qubit systems, in order to test the invasiveness, see 'Methods'.

In general, as we increase the number n of qubits involved in the experiment, testing the invasiveness can be challenging because there are a total of 2^n circuits to be generated to prepare all possible states i. However, if one finds a $\mathcal{I}(i)$ which is already greater than the observed witness violation, it is of course unnecessary to continue. For example, for the six-qubit case, instead of preparing all possible macroscopic states i, we only consider the state $\{0\}^6$ because the experimental value of the quantum witness we observe for the six-qubit case later is already lower than the invasiveness quantity:

$$\mathcal{I}(\{0\}^6) = |\,1 - \epsilon^M(\{0\}^6|\{0\}^6)\,|. \qquad (13)$$

In all cases ($n = 1, 2, 4, 6$), we test the measurement invasiveness in 25 different experiments across different days, each consisting of 8192 runs. This was done because the tests

Table 1. Table of the measurement invasiveness parameters of states $\{0\}^n$ for single-, two-, four- and six-qubit systems, respectively.

n	$\mathcal{I}_{\mathrm{Max}}(\{0\}^n)$	$\mathcal{I}_{\mathrm{Ave}}(\{0\}^n)$
1	0.023 ± 0.004	0.016 ± 0.003
2	0.077 ± 0.008	0.068 ± 0.006
4	0.146 ± 0.006	0.106 ± 0.019
6	0.686 ± 0.007	0.310 ± 0.225

The results of all possible states i in the computational basis for single-, two- and four-qubit cases are given in the 'Methods' section. Here we perform 25 experiments, across multiple days, to take into account the variability in the IBM quantum system parameters. Each experiment consists of 8192 runs. The maximal and average values, over experiments, of the measurement invasiveness are obtained from the IBM Q5 Tenerife for the single- and two-qubit systems, while the results for four- and six-qubit systems are from the IBM Q14 Melbourne. The simulation results, using the noise model described in the main text, are respectively $\mathcal{I}_{\mathrm{Sim}}(\{0\}^1) = 0.031$ and $\mathcal{I}_{\mathrm{Sim}}(\{0\}^2) = 0.061$ for the single- and two-qubit cases.

were not performed at the same time as the data collection for testing the actual witness itself. On these long timescales, the IBM QE exhibits fluctuations in parameters, including coherence times, gate fidelities, etc., and thus we introduce a variance in this test that represents these fluctuations. We note that the results of single- and two-qubit systems are obtained from the IBM Q5 Tenerife, while the results for four- and six-qubit systems are obtained from the IBM Q14 Melbourne.

Figure 2 shows experimental data for the $n = 2$ cat state with $\theta \in \{0, \pi/8, 2\pi/8, 3\pi/8, 4\pi/8\}$. We also show the theoretical predictions both with and without noise simulation, as well as the modified witness bound based on the measurement invasiveness tests. From Fig. 2 we observe that the maximum value of the quantum witness occurs when the entanglement parameter $\theta = \frac{\pi}{2}$, which is the maximally entangled state.

At $\theta = 0$, we find that the value of the quantum witness is lower than the average experimental measurement invasiveness tests, implying it is consistent with macrorealism. Interestingly, there is a residual small violation of the witness even though this is not predicted by the simple 'pure states' $\{0\}^2$ expression in Eq. (7). This 'invasiveness' represents a classically invasive measurement. For example, in our simulation plotted in Fig. 2, we observe that the $\theta = 0$ non-zero witness value arises directly from γ_{Errors} in Eq. (12) (i.e., if we set $\gamma_{\mathrm{Errors}} = 0$ the witness value in the simulation falls to zero). Thus, as discussed in the section Quantum witness, the γ_{Errors} is related to the clumsy measurability $\epsilon^M(\{0\}^2|\{0\}^2)$ since at $\theta = 0$ the intermediate state corresponds to $\{0\}^2$.

When we consider the results of the $n = 4$ and $n = 6$ cat states in Table 2, we see that the magnitude of the quantum witness violation drastically decreases, as compared to the $n = 2$ case. The value of the quantum witness for the four-qubit cat state is still larger than the invasiveness test, implying a non-macrorealistic behaviour. In contrast, while the invasiveness test for six qubits is only performed with the $\{0\}^6$ state, we see that the six-qubit cat state does not exceed these tests. Thus, we can conclude the six-qubit system is compatible with a clumsy-macrorealistic description. This result shows that the IBM QE tends to a clumsy-macroscopic realistic behaviour as the number of qubits increases, due to the increased influence of decoherence and dephasing processes as the circuit complexity, or 'depth'[15], increases.

The values we observe for the quantum witness with the product states are shown in Table 3. We find that, compared to cat states, the product states witness values are more robust as we increase n, reflecting the increased sensitivity of cat states used in the previous section to dephasing and decoherence, and the lower circuit depth (and hence less time being spent exposed to
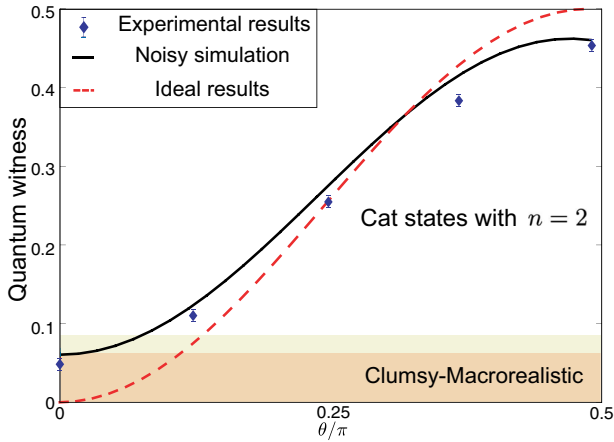
Fig. 2 The value of the quantum witness of $n = 2$ cat states. The circuit is designed to produce the state $|\phi(2, \theta)\rangle = \cos\frac{\theta}{2}|0\rangle^{\otimes 2} + \sin\frac{\theta}{2}|1\rangle^{\otimes 2}$ at an intermediate time. The experimental results obtained by the IBM QE are shown by blue diamonds. The theoretical results, with and without noise simulation, are shown by red dashed and black solid curves, respectively. Obviously, the quantum witness increases with the parameter θ, but also shows a residual violation due to the macroscopically invasive measurements backaction and gate error at $\theta = 0$. We simulate the influence of decoherence and gate infidelities by Lindblad-form master equations (11) and (12). The coefficients of relaxation time $T_1 = 46\,\mu s$, dephasing time $T_2 = 13.5\,\mu s$ and gate-error coefficient $\gamma_{\text{Errors}} = 8.5 \times 10^{-2}\,\mu s^{-1}$ are determined by approximately fitting the experimental results. The grey and orange shaded areas at the bottom are the clumsy-macrorealistic regimes determined by the maximal and average invasiveness tests in Table 1. Note that the invasiveness test including the standard deviation does not depend on θ. The experimental uncertainties are derived from the multinomial distribution and error propagation except for the average disturbance case which is the variance across 25 repeated experiments.

Table 2. Table of the quantum witness W, maximal measurement invasiveness $\mathcal{I}_{\text{Max}}(\{0\}^n)$, and their relative difference for the larger cat states $n = 4$ and $n = 6$.

	$n = 4$	$n = 6$
W	0.245 ± 0.008	0.181 ± 0.007
$\mathcal{I}_{\text{max}}(\{0\}^n)$	0.146 ± 0.006	0.686 ± 0.007
Relative difference	0.099 ± 0.010	-0.505 ± 0.010

noise) of this product state example. Due to inevitable noise described above, the observed values of the quantum witness for each n do not reach precisely the corresponding theoretically predicted maximum possible values of $W_{\text{max}} = 1 - D_{\text{Ideal}}^{-1}$ with $D_{\text{Ideal}} = 2^n$. Nevertheless the value of the witness increases with n (and hence the number of states in the Hilbert space) as expected. More specifically, as we increase the number n of qubits, the value of the corresponding quantum witness not only increases but is always larger than the maximum value with qubit number $(n - 1)$. This confirms that in practice the quantum witness can function as a dimensionality witness. We note that several other approaches to witnessing dimensionality, using different types of temporal correlations, were recently implemented[60,61].

DISCUSSION

By taking advantage of the programmable nature of the IBM QE, our results have shown how the violation of an LGI, in the form of a 'quantum witness', changes as we increase the number of qubits

Table 3. Table of the quantum witness for a product of superposition states.

n	2	3	4	6
D_{Ideal}	4	8	16	64
W_{Max}	0.75	0.875	0.937	0.984
W_{Exp}	0.746 ± 0.005	0.857 ± 0.004	0.902 ± 0.003	0.940 ± 0.003
D_{Exp}	≥ 3	≥ 6	≥ 10	≥ 16

Here, $D_{\text{Ideal}} = 2^n$ is the ideal dimension of the system with qubit number n. The corresponding ideal value of the quantum witness is $W_{\text{Max}} = 1 - 1/D_{\text{Ideal}}$. Here, W_{Exp} is the value of the quantum witness obtaining from the IBM Q14 Melbourne with the estimating dimension $D_{\text{Exp}} = \lfloor(1 - W_{\text{Exp}})^{-1}\rfloor$, where $\lfloor Y \rfloor$ is the integer of the number Y.

contributing to a highly entangled state. This allows us to see directly how the system becomes more macrorealistic as we increase the macroscopicity.

For $n = 2$, we observed a violation of the quantum witness for $\theta = \pi/8, 2\pi/8, 3\pi/8$ and $\pi/2$, for $n = 4$ for $\theta = \pi/2$. Thus, we can claim that when manipulating and observing two qubits in the IBM Q5 Tenerife device, and four qubits in the 14-qubit processor IBM Q14 Melbourne used for this experiments, the results must be described with a non-macrorealistic theory. On the other hand, we found that six qubits, prepared in a GHZ state, did not violate the witness beyond a measurement invasiveness test, and thus these observations can, in principle, be described with macrorealistic theories. As the capabilities of the IBM QE improve (e.g., when ancilla qubits are not required for the intermediate measurements), error correction and error mitigation techniques are employed, the boundary between quantum theory and potential clumsy-macrorealistic theories could be tested with a much larger number of qubits.

The classical invasiveness, or clumsiness, we observed in the data (e.g., clearly exemplified by the non-zero quantum witness value at $\theta = 0$ for $n = 2$) can be explained by our 'minimal' Lindblad master equation noise model, where the infidelity of the CNOT operations used in the intermediate measurements causes changes in the state of the qubits. Moreover, our minimal model can also explain the suppression of the witness violation due to dephasing and energy relaxation.

To complement our primary results, instead of preparing entangled states, we also tested a product of superposition states, which has a low disconnectivity. We found that, as expected for such a state, the maximal violation increases with the number of qubits, and hence the dimensionality. In addition, the influence of noise on these results is substantially smaller than the GHZ-state based test. This is because single-qubit coherence tends to be less susceptible to noise than GHZ states, and because of the lower total circuit depth.

Finally, it is important to note that recent work has shown nonnegligible non-Markovian effects in the IBM QE[62]. This can introduce a secondary loophole in the LGI due to the non-instantaneous nature of the measurements we perform at time t_1. For example, in the IBM QE, the measurements at time t_1 take about 0.4 and 0.9 μs for the 5- and 14-qubit devices, respectively. This long timescale appears because of the CNOT operation between primary and ancilla qubits needed for our intermediate measurement. Recent works suggest that non-Markovian effects are important on timescales of $\simeq 5\,\mu s$[62]. Thus, differences in environment evolution on the timescale of our intermediate measurements may cause differences in the outcomes in the two contributions to the witness (i.e., differences to the final probability distributions between when the measurement is performed and when it is not). Like with clumsy measurements,

because of this non-instantaneous measurement time, the origin of violations in this test due to breakdown of macrorealism, or due to non-Markovian environmental influences, cannot be delineated. Our measurement invasiveness test may compensate for this to some degree, but further work is needed to take into account this potential loophole with such a test. Alternatively, the non-Markovian effect can be diminished by using faster measurements, should such become available (either via faster CNOT operations, or the availability of direct measurements on the primary qubits at intermediate times).

In the Methods section, we consider an alternative approach to implement the witness which removes the need to use ancilla qubits, and hence reduce the circuit depth. From a simple inspection of the definition of the quantum witness, one can see that we can, instead of directly measuring the two-time correlation functions by using the ancilla qubits, first run an experiment where the probabilities $p_{t_1}^M(i)$ are collected. Then we run another experiment where one deterministically prepares the system in the state i, and measures $p_{t_2,t_1}(j|i)$. This scenario, which we call 'prepare-and-measure', replaces the non-invasive measurement assumption with an ideal-state preparation and a more explicit non-Markovian evolution assumption (see refs [53,63]).

Overall, our results suggest that the current iteration of the IBM QE tends towards clumsy-macrorealistic behaviour for more than four qubits. This is inevitably also a function of the resulting circuit depth[15] (i.e., overall run-time) on which the witness can be tested, which increases as the number of qubits is increased. A significant contribution to the circuit depth arises from the ancilla-based measurements, thus future improvements to the IBM QE which allow multiple measurements on a single qubit may significantly reduce this circuit depth.

Finally, we point out that since a CNOT gate is its own inverse, one can reinterpret the combination of the quantum witness, and our choice of circuit, as a test of a classical circuit identity under the conditions of macrorealism. In other words, we tested whether $CNOT^2 = id$ still holds under the condition of an intermediate projection onto a classical basis between the two CNOT gates. Under quantum mechanics, of course, such relations are violated. Thus, we arrive at a different perspective on quantum witness tests, namely that they can be viewed as tests of reversible classical circuit identities under intermediate measurements.

METHODS

Modifying the quantum witness for clumsy measurements
Here, we explicitly derive Eq. (4) in the main text. Inserting Eq. (3) into the definition of the witness, one finds

$$W = \left| \sum_{i,k} p_{t_2,t_1}(j|k)\epsilon^M(k|i)p_{t_1}(i) - \sum_i p_{t_2,t_1}(j|i)p_{t_1}(i) \right|$$
$$\leq \max_i \left| \sum_k p_{t_2,t_1}(j|k)\epsilon^M(k|i) - p_{t_2,t_1}(j|i) \right|,$$
(14)

where the maximum over states i at time t_1 comes from the upper bound on a convex combination. This we can rewrite as

$$W \leq \max_i \left| p_{t_2,t_1}(j|i)\left[1 - \epsilon^M(i|i)\right] - \sum_{k \neq i} p_{t_2,t_1}(j|k)\epsilon^M(k|i) \right|.$$

Since $p_{t_2,t_1}(j|i) \leq 1$, $\sum_k \epsilon^M(k|i) = 1$, and the remaining terms are positive, we can bound the right-hand side further as

$$W \leq \max_i \left| 1 - \epsilon^M(i|i) \right|.$$
(15)

Thus we obtain an upper bound for the witness under the assumption of a *macroscopically invasive* nature of the intermediate measurements. This bound assumes nothing about the evolution from t_1 to t_2.

The bound in Eq. (15) can be said to be a weaker bound than that in Eq. (14), but is more experimentally efficient because we do not need to consider the effect of potential arbitrary evolution between t_1 and t_2.

We note that Eq. (14) alone is equivalent to the test employed in ref. [63]. One just needs to sum up the outcomes j in Eq. (14) for the multi-outcome scenario considered in that work. Our additional derivation of a weaker bound in Eq. (15) can be similarly generalized to multiple final outcome measurements. This method [40,55] is also related to the 'adroit' measurement test proposed in refs [40,55], when one assumes a particular intermediate measurement and that the states before that measurement are macrorealistic. Our bound is not as strong as the adroit one, but is easier to implement for the many-qubit situation we explore in this work.

Quantum circuits: direct-measure scenario
From the initial state $|0\rangle^{\otimes n}$, 'cat states' at time t_1 can be obtained by performing the unitary transformation $U|0\rangle^{\otimes n} = |\phi(n,\theta)\rangle = \cos(\frac{\theta}{2})|0\rangle^{\otimes n} + \sin(\frac{\theta}{2})|1\rangle^{\otimes n}$. In the IBM QE, we implemented the unitary U by applying the $U_3(0, 0, \theta)$ gate in Eq. (10) on the first qubit, and subsequently performing $n - 1$ CNOT gates between the first qubits and all others. Therefore, the unitary $U = C_{Q_{n-1}}^{Q_{n-2}}......C_{Q_2}^{Q_1}C_{Q_1}^{Q_0}U_3^{Q_0}$ with the superscript Q_0 of a single quantum gate U_3 representing the operation acting on the qubit Q_0. Here, the super- and subscripts of a CNOT operation represent the control and target qubits, respectively. The inverse operation $U^\dagger$ is applied after time t_1 and it is given by reverse the gate implementation above. We note that if one were to directly implement the circuit without 'barriers' on the IBM QE it would be automatically 'optimized' to be an identity operation. In Fig. 3a, we present an explicit example of a two-qubit cat state.

Now, we can introduce how to perform the intermediate measurement at time t_1 and obtain the two-time correlation function. Since the IBM QE only allows at most one measurement operation on any given qubit, we have to perform a CNOT gate on each measured qubit and an ancilla qubit. Here, the ancilla and measured qubits are respectively the target and control qubits [see Fig. 3b and ref. [64]]. The measurement results on the ancilla qubit refer to the outcomes i and leave behind the corresponding post-measurement states $|\gamma\rangle_i$. After the measurement at time t_1, we apply the $U^\dagger$ on the post-measurement state. We denote this approach as a direct-measure scenario.

For instance, if the target and control qubits are respectively $|0\rangle$ and $\alpha|0\rangle + \beta|1\rangle$, with $|\alpha|^2 + |\beta|^2 = 1$, the state after the CNOT operation is $|\kappa\rangle = \alpha|00\rangle + \beta|11\rangle$. Now we perform a measurement on the target qubit in the computational basis. Following Born's rule, we have

$$\gamma_i = \text{Tr}_{\text{Target}}(\text{id} \otimes |i\rangle\langle i|\rho),$$
(16)

where $\rho = |\kappa\rangle\langle\kappa|$ is the state at time t_1, $|i\rangle\langle i|$ is a projector onto the computational basis, and $\gamma_i = |\gamma\rangle_i\langle\gamma|$ is the remained state with the corresponding outcome i.

The second measurement with outcome j at time t_2 can be implemented, without the need for ancillas. From this, the IBM QE can return the result $p_{t_2}^M(j)$. Finally, we note that while IBM Q14 Melbourne has 14 qubits, one cannot perform CNOT gates between arbitrary qubits because the direction of a CNOT gate is limited by the physical processor design (see the physical structures in ref. [1]), limiting us to 6 qubit in our cat state, and 6 ancilla qubits. We note that in the current IBM QE, all of the qubits are measured in the end regardless of whether the measurement gates are actually implemented in the quantum circuit. After measuring all of the qubits, post-processing of the resulting data is applied according to the measurement gates one has chosen.

The measurement invasiveness of the other states
Here, we present the values of the measurement invasiveness of the single-, two-, four-qubit states in Tables 4 and 5. We prepare all 'macrorealistic' states i in the computational basis to test the invasiveness of the intermediate measurement at time t_1.

In addition, it is important to note that the uncertainties given for the average values of the measurement invasiveness test represent the variance across 25 different experiments (each individually consisting of 8192 runs) performed on different days, and thus reflect the variance in various properties of the IBM QE across these long timescales[12], and are thus different from the ones in the rest of the paper.

Quantum circuits: prepare-and-measure scenario
An alternative approach (which can in principle allow for a larger number of measured qubits since no ancilla qubits are needed) relies on trading the measurement at time t_1 with ideal-state preparation. In this scenario, the first circuit is performed with a unitary transformation U before the measurements at time t_1. The IBM QE returns the probability distribution
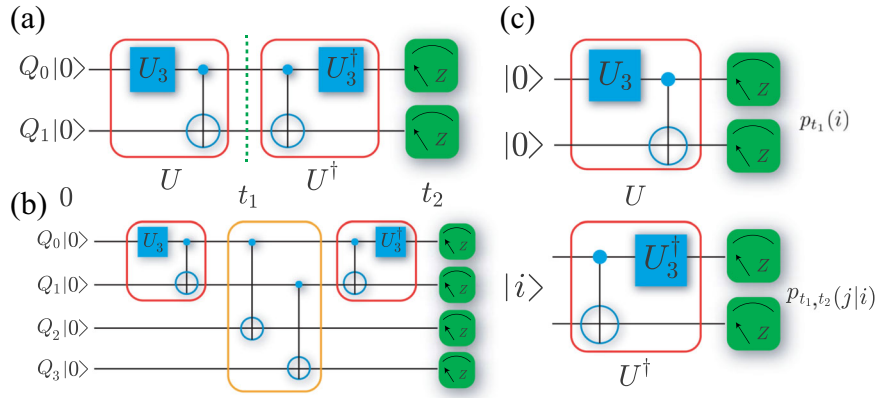
Fig. 3 **Example of schematic quantum circuits for $n = 2$. (a)** is for measuring $p_{t_2}(j)$. In the IBM QE, the qubits denoted by Q_i for $i = 0$ and 1 are initially prepared in $|0\rangle$. The left and right red areas, respectively, represent the unitary transformations U and $U^\dagger$, which can be decomposed by $U_3(0, 0, \theta)$ (U_3 in short) and a series of CNOT operations. In the beginning, U_3 is performed on Q_0, followed by a CNOT gate on the control Q_0 and the target qubits Q_1. The green dots represent the barrier between U and $U^\dagger$ to avoid the automatic optimization. The $U^\dagger$ is performed after the barrier. In the end, the measurements on the computational basis are performed such that the value $p_{t_2}(j)$ is obtained. **(b)** shows the quantum circuit for measuring $\sum_i p_{t_1}(i)p_{t_2,t_1}(j|i)$ in the direct-measure scenario. Since the IBM QE cannot measure the same qubit twice, the intermediate measurement at time t_1 can be implement by the CNOT operation with the ancilla qubit Q_2 and Q_3. We use the yellow box to represent the intermediate measurement. Here, the ancilla qubits are initially in $|0\rangle$. Since we only consider the projective measurement onto the computational basis, one can implement the CNOT operation to transfer the classical information of the state to the ancilla qubit. The measurement operations on the ancilla qubits Q_2 and Q_3 remain in the post-measurement state $|\gamma_i\rangle$ with outcomes i. Finally, with the measurement on the qubits Q_1 and Q_2, the quantum circuit returns the result $\sum_i p_{t_1}(i)p_{t_2,t_1}(j|i)$. **(c)** shows the quantum circuits for, respectively, measuring $p_{t_1}(i)$ and $p_{t_2,t_1}(j|i)$ in the prepare-and-measured scenario. The unitary transformation U is performed on the state $|0\rangle$, followed by measurement operations with outcome i at time t_1. In the second experiments, the eigenstates $|i\rangle$ are prepared according to the probability $p_{t_1}(i)$, followed by the inverse unitary transformation $U^\dagger$. The measurement results are the probability with outcome j conditional on i.

Table 4. Table of the measurement invasiveness parameters for single, two-qubit systems, respectively.			
$\{i\}$	$\mathcal{I}_{\mathrm{Max}}(\{i\})$	$\mathcal{I}_{\mathrm{Ave}}(\{i\})$	$\mathcal{I}_{\mathrm{Sim}}(\{i\})$
$\{0\}$	0.023 ± 0.004	0.016 ± 0.003	0.031
$\{1\}$	0.024 ± 0.004	0.014 ± 0.004	0.011
$\{00\}$	0.077 ± 0.008	0.068 ± 0.006	0.061
$\{11\}$	0.051 ± 0.005	0.033 ± 0.009	0.021
$\{10\}$	0.033 ± 0.004	0.020 ± 0.006	0.019
$\{01\}$	0.040 ± 0.004	0.030 ± 0.005	0.019

Here we perform the 25 experiments, across multiple days, to take into account variability in the IBM quantum system parameters. Each collation consists of 8192 runs. The maximal and average values over experiments of the measurement invasiveness are obtained from the IBM Q5 Tenerife for the single- and two-qubit systems. The simulation results, using the noise model described in the main text, are also presented.

Table 5. Table of the measurement invasiveness parameters for four-qubit systems.		
$\{i\}$	$\mathcal{I}_{\mathrm{Max}}(\{i\})$	$\mathcal{I}_{\mathrm{Ave}}(\{i\})$
$\{0000\}$	0.146 ± 0.005	0.106 ± 0.019
$\{0001\}$	0.115 ± 0.005	0.075 ± 0.029
$\{0010\}$	0.123 ± 0.006	0.104 ± 0.011
$\{0100\}$	0.136 ± 0.006	0.093 ± 0.026
$\{1000\}$	0.100 ± 0.006	0.078 ± 0.015
$\{0011\}$	0.091 ± 0.006	0.067 ± 0.013
$\{0101\}$	0.124 ± 0.006	0.079 ± 0.031
$\{1001\}$	0.097 ± 0.006	0.061 ± 0.023
$\{0110\}$	0.121 ± 0.006	0.103 ± 0.012
$\{1010\}$	0.118 ± 0.006	0.086 ± 0.022
$\{1100\}$	0.122 ± 0.006	0.080 ± 0.020
$\{1110\}$	0.109 ± 0.007	0.078 ± 0.022
$\{1101\}$	0.068 ± 0.007	0.052 ± 0.010
$\{1011\}$	0.077 ± 0.006	0.046 ± 0.018
$\{0111\}$	0.095 ± 0.007	0.057 ± 0.021
$\{1111\}$	0.070 ± 0.007	0.035 ± 0.018

$p_{t_1}(i)$ with outcomes i. According to the probability distribution $p_{t_1}(i)$, we then prepare a new circuit with an initial state in the eigenstates $|i\rangle$. The $U^\dagger$ operation is then performed before the measurements at time t_2 on the system. The results from the IBM QE represent the conditional probability distributions $p_{t_2,t_1}(j|i)$. Here, only the outcome $j = 0$ is used to analyse the quantum witness in Eq. (2).

We prepare all possible eigenstates $|i\rangle$ for $n = 2$ and 4 qubits systems. For the 6-qubit case, we only prepare the eigenstates $|i\rangle$ if $p_{t_1}(i) \geq 10^{-3}$, which is chosen to be much smaller than the ideal outcome of, e.g., $p_{t_1}^M(0) = 0.5$ (note that the error induced in the witness due to omission of these small terms can in principle be of the same order as the uncertainty in the experimental data we show later; but given that the observed violation is already lower than the measurement invasiveness, this error does not cause a false witness). Finally, we note that there are at most ($i +$ 1) quantum circuits in this scenario with i being the total number of the states we need to prepare. However, there are only two experimental circuits needed to collate the corresponding statistical data

$\sum_i p_{t_1}(i_{t_1})p_{t_2,t_1}(j|i)$, and $p_{t_2}(j)$ in the direct-measure scenario. Therefore, the prepare-and-measure scenario is not efficient as the number of qubits increases because the number of quantum circuits we need to collate all possible correlations increases with the number of outcomes i.

As with the direct-measure scenario, which suffers from a 'clumsiness loophole' arising from the non-invasive measurement assumption, the prepare-and-measure scenario can similarly suffer from a clumsiness loophole related to non-ideal-state preparation which lead a non-zero value for $\theta = 0$ in our experiment. Moreover, in principle, non-Markovian effects also lead to a false-positive violation of the quantum witness. For instance, if the history from time t_0 to t_1 influences the evolution from time

Table 6. Table of quantum witness of cat states with $n = 2$.

	W_{PM}	W_{DM}
$\theta = 0$	0.061 ± 0.006	0.058 ± 0.007
$\theta = \pi/8$	0.112 ± 0.006	0.102 ± 0.007
$\theta = 2\pi/8$	0.261 ± 0.006	0.250 ± 0.008
$\theta = 3\pi/8$	0.430 ± 0.006	0.376 ± 0.008
$\theta = 4\pi/8$	0.454 ± 0.006	0.453 ± 0.008

The W_{PM} and W_{DM} are quantum witness obtained by prepare-and-measure and direct-measure scenarios, respectively. We note that the measurement invasiveness is 0.077 ± 0.008.

Table 7. Table of quantum witness of cat states with $n = 4$ and $n = 6$.

	W_{PM}	W_{DM}
$n = 4$	0.251 ± 0.007	0.245 ± 0.008
$n = 6$	0.183 ± 0.067	0.181 ± 0.007

The W_{PM} and W_{DM} are quantum witness obtained by prepare-and-measure and direct-measure scenarios, respectively.

t_1 to t_2, this may also lead to differences in the probability distributions[53,63] $p_{t_2}(j)$ and $p_{t_2}^M(j)$.

Finally, we present a schematic example of the quantum circuit for the two-qubit case [see Fig. 3c]. The initial state of the total system on the IBM QE is $|0\rangle^{\otimes 2}$. The state becomes a 'cat state' in Eq. (6) by applying a unitary transformation U. Instead of evolving back to the state $|0\rangle^{\otimes 2}$, we measure the cat states at time t_1 to obtain the probability $p_{t_1}(i)$ [see the top half of the Fig. 3c]. After the first experiment, we prepare a quantum state $|i\rangle$, which is the eigenstate of the corresponding outcomes i, and perform a $U^\dagger$ operation. Finally, the measurement operation is performed to obtain the probability $p_{t_2,t_1}(j|i)$ [see the bottom of the Fig. 3c]. One can easily expand the two-qubit system to a GHZ one.

In general, the prepare-and-measure scenario can also test for qubit number $n > 6$. However, we do not do this cumbersome procedure because the direct-measure results shows that for the $n = 6$ case the system is already classified as macrorealistic.

Interestingly, the witness values from the prepare-and-measure scenario are almost all slightly higher than the direct-measure ones [see Tables 6 and 7]. From the circuit-implementation point of view, the prepare-and-measure scenario significantly reduces the number of CNOT gates, which take almost four times longer than the U_3 gates. Therefore, the prepare-and-measure scenario effectively reduces the overall effect of noise on the witness and has a much lower circuit depth. However even the prepare-and-measure scenario does not produce a violation for six qubits.

DATA AVAILABILITY

All data supporting the findings of this study including cat states, measurement invasiveness, and product of superposition states have been deposited in creative commons with the https://doi.org/10.25405/data.ncl.9994739.

CODE AVAILABILITY

The code for analysing the simulation and the IBM qiskit code is available at https://github.com/huan-yu-20/IBM-cat-state.

REFERENCES

1. IBM quantum team. *IBM quantum experience*. Available at: https://quantum-computing.ibm.com. Accessed (2020).
2. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).
3. Wecker, D., Bauer, B., Clark, B. K., Hastings, M. B. & Troyer, M. Gate-count estimates for performing quantum chemistry on small quantum computers. *Phys. Rev. A* **90**, 022305 (2014).
4. Balu, R., Castillo, D. & Siopsis, G. Physical realization of topological quantum walks on IBM-Q and beyond. *Quantum Sci Technol.* **3**, 035001 (2018).
5. Hsieh, J.-H., Chen, S.-H. & Li, C.-M. Quantifying quantum-mechanical processes. *Sci. Rep.* **7**, 13588 (2017).
6. Morris, J., Pollock, F. A. & Modi, K. Non-Markovian memory in IBMQX4. Preprint at https://arxiv.org/abs/1902.07980 (2019).
7. Mitarai, K., Negoro, M., Kitagawa, M. & Fujii, K. Quantum circuit learning. *Phys. Rev. A* **98**, 032309 (2018).
8. Devitt, S. J. Performing quantum computing experiments in the cloud. *Phys. Rev. A* **94**, 032329 (2016).
9. Steiger, D. S., Häner, T. & Troyer, M. ProjectQ: an open source software framework for quantum computing. *Quantum* **2**, 49 (2018).
10. Knill, E. Quantum computing with realistically noisy devices. *Nature* **434**, 39–44 (2005).
11. Harper, R. & Flammia, S. T. Fault-tolerant logical gates in the IBM quantum experience. *Phys. Rev. Lett.* **122**, 080504 (2019).
12. Alsina, D. & Latorre, J. Experimental test of Mermin inequalities on a five-qubit quantum computer. *Phys. Rev. A* **94**, 012314 (2016).
13. Wang, Y., Li, Y. & Zeng, B. 16-qubit IBM universal quantum computer can be fully entangled. *npj Quantum Inf.* **4**, 46 (2018).
14. Mooney, G. J., Hill, C. D. & Hollenberg, L. C. L. Entanglement in a 20-qubit superconducting quantum computer. *Sci. Rep.* **9**, 13465 (2019).
15. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
16. Schrödinger, E. Discussion of probability relations between separated systems. *Proc. Cambridge Philos. Soc.* **31**, 555 (1935).
17. Leggett, A. J. & Garg, A. Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Phys. Rev. Lett.* **54**, 857 (1985).
18. Emary, C., Lambert, N. & Nori, F. Leggett-Garg inequalities. *Rep. Prog. Phys.* **77**, 016001 (2014).
19. Emary, C., Lambert, N. & Nori, F. Leggett-Garg inequality in electron interferometers. *Phys. Rev. B* **86**, 235447 (2012).
20. Lambert, N., Emary, C., Chen, Y.-N. & Nori, F. Distinguishing quantum and classical transport through nanostructures. *Phys. Rev. Lett.* **105**, 176801 (2010).
21. Lambert, N., Chen, Y.-N. & Nori, F. Unified single-photon and single-electron counting statistics: from cavity qed to electron transport. *Phys. Rev. A* **82**, 063840 (2010).
22. Uola, R., Vitagliano, G. & Budroni, C. Leggett-Garg macrorealism and the quantum nondisturbance conditions. *Phys. Rev. A* **100**, 042117 (2019).
23. Kofler, J. & Brukner, C. Condition for macroscopic realism beyond the Leggett-Garg inequalities. *Phys. Rev. A* **87**, 052115 (2013).
24. Halliwell, J. J. Leggett-Garg inequalities and no-signaling in time: a quasiprobability approach. *Phys. Rev. A* **93**, 022123 (2016).
25. Budroni, C. & Emary, C. Temporal quantum correlations and Leggett-Garg inequalities in multilevel systems. *Phys. Rev. Lett.* **113**, 050401 (2014).
26. Lambert, N. et al. Leggett-Garg inequality violations with a large ensemble of qubits. *Phys. Rev. A* **94**, 012105 (2016).
27. Hoffmann, J., Spee, C., Gühne, O. & Budroni, C. Structure of temporal correlations of a qubit. *N. J. Phys.* **20**, 102001 (2018).
28. Chen, Y.-N. et al. Temporal steering inequality. *Phys. Rev. A* **89**, 032112 (2014).
29. Chen, S.-L. et al. Quantifying non-Markovianity with temporal steering. *Phys. Rev. Lett.* **116**, 020503 (2016).
30. Bartkiewicz, K., Černoch, A., Lemr, K., Miranowicz, A. & Nori, F. Temporal steering and security of quantum key distribution with mutually unbiased bases against individual attacks. *Phys. Rev. A* **93**, 062345 (2016).
31. Ku, H.-Y. et al. Temporal steering in four dimensions with applications to coupled qubits and magnetoreception. *Phys. Rev. A* **94**, 062126 (2016).
32. Ku, H.-Y., Chen, S.-L., Lambert, N., Chen, Y.-N. & Nori, F. Hierarchy in temporal quantum correlations. *Phys. Rev. A* **98**, 022104 (2018).
33. Li, C.-M., Chen, Y.-N., Lambert, N., Chiu, C.-Y. & Nori, F. Certifying single-system steering for quantum-information processing. *Phys. Rev. A* **92**, 062310 (2015).
34. Uola, R., Lever, F., Gühne, O. & Pellonpää, J.-P. Unified picture for spatial, temporal, and channel steering. *Phys. Rev. A* **97**, 032301 (2018).
35. Goggin, M. E. et al. Violation of the Leggett-Garg inequality with weak measurements of photons. *Proc. Natl Acad. Sci. USA* **108**, 1256–1261 (2011).
36. Dressel, J., Broadbent, C. J., Howell, J. C. & Jordan, A. N. Experimental violation of two-party Leggett-Garg inequalities with semiweak measurements. *Phys. Rev. Lett.* **106**, 040402 (2011).
37. Bartkiewicz, K., Černoch, A., Lemr, K., Miranowicz, A. & Nori, F. Experimental temporal quantum steering. *Sci. Rep.* **6**, 38076 (2016).

38. Ringbauer, M., Costa, F., Goggin, M. E., White, A. G., Fedrizzi, A. Multi-time quantum correlations with no spatial analog. *npj Quantum Inf.* **4**, 37 (2018).

39. Knee, G. C. et al. Violation of a Leggett–Garg inequality with ideal non-invasive measurements. *Nat. Commun.* **3**, 606 (2012).

40. Huffman, E. & Mizel, A. Violation of noninvasive macrorealism by a super-conducting qubit: Implementation of a Leggett-Garg test that addresses the clumsiness loophole. *Phys. Rev. A* **95**, 032131 (2017).

41. Knee, G. C. et al. A strict experimental test of macroscopic realism in a super-conducting flux qubit. *Nat. Commun.* **7**, 13253 (2016).

42. Palacios-Laloy, A. et al. Experimental violation of a Bell's inequality in time with weak measurement. *Nat. Phys.* **6**, 442–447 (2010).

43. Robens, C., Alt, W., Meschede, D., Emary, C. & Alberti, A. Ideal negative mea-surements in quantum walks disprove theories based on classical trajectories. *Phys. Rev. X* **5**, 011003 (2015).

44. Budroni, C. et al. Quantum nondemolition measurement enables macroscopic Leggett-Garg tests. *Phys. Rev. Lett.* **115**, 200403 (2015).

45. Emary, C., Cotter, J. P. & Arndt, M. Testing macroscopic realism through high-mass interferometry. *Phys. Rev. A* **90**, 042114 (2014).

46. Bose, S., Home, D. & Mal, S. Nonclassicality of the harmonic-oscillator coherent state persisting up to the macroscopic domain. *Phys. Rev. Lett.* **120**, 210402 (2018).

47. Gühne, O. & Tóth, G. Entanglement detection. *Phys. Rep.* **474**, 1–75 (2009).

48. Leggett, A. J. Testing the limits of quantum mechanics: motivation, state of play, prospects. *J. Phys. Condens. Matter* **14**, R415–R451 (2002).

49. Leggett, A. J. Macroscopic quantum systems and the quantum theory of mea-surement. *Prog. Theor. Phys. Suppl.* **69**, 80–100 (1980).

50. White, T. C. et al. Preserving entanglement during weak measurement demon-strated with a violation of the Bell–Leggett–Garg inequality. *npj Quantum Inf.* **2**, 15022 (2016).

51. Dressel, J. & Korotkov, A. N. Avoiding loopholes with hybrid Bell-Leggett-Garg inequalities. *Phys. Rev. A* **89**, 012125 (2014).

52. Nimmrichter, S. & Hornberger, K. Macroscopicity of mechanical quantum super-position states. *Phys. Rev. Lett.* **110**, 160403 (2013).

53. Li, C.-M., Lambert, N., Chen, Y.-N., Chen, G.-Y. & Nori, F. Witnessing quantum coherence: from solid-state to biological systems. *Sci. Rep.* **2**, 1092 (2012).

54. Marcus, M., Knee, G. C. & Datta, A. Towards a spectroscopic protocol for unam-biguous detection of quantum coherence in excitonic energy transport. *Faraday Discuss.* **221**, 110–132 (2020).

55. Wilde, M. M. & Mizel, A. Addressing the clumsiness loophole in a Leggett-Garg test of macrorealism. *Found. Phys.* **42**, 256–265 (2011).

56. Oreshkov, O., Costa, F. & Brukner, Č. Quantum correlations with no causal order. *Nat. Commun.* **3**, 885 (2012).

57. Schild, G. & Emary, C. Maximum violations of the quantum-witness equality. *Phys. Rev. A* **92**, 032101 (2015).

58. Johansson, J. R., Nation, P. D. & Nori, F. QuTiP: an open-source python framework for the dynamics of open quantum systems. *Comput. Phys. Commun.* **183**, 1760–1772 (2012).

59. Johansson, J. R., Nation, P. D. & Nori, F. QuTiP 2: a python framework for the dynamics of open quantum systems. *Comput. Phys. Commun* **184**, 1234–1240 (2013).

60. Strikis, A., Datta, A. & Knee, G. C. Quantum leakage detection using a model-independent dimension witness. *Phys. Rev. A* **99**, 032328 (2019).

61. Spee, C. et al. Genuine temporal correlations can certify the quantum dimension. *New J. Phys.* **22**, 023028 (2020).

62. Pokharel, B., Anand, N., Fortman, B. & Lidar, D. A. Demonstration of fidelity improvement using dynamical decoupling with superconducting qubits. *Phys. Rev. Lett.* **121**, 220502 (2018).

63. Knee, G. C., Marcus, M., Smith, L. D. & Datta, A. Subtleties of witnessing quantum coherence in non-isolated systems. *Phys. Rev. A* **98**, 052328 (2018).

64. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* 10th Anniversary Edition (Cambridge University Press, 2010).

AUTHOR CONTRIBUTIONS

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to N.L. or Y.-N.C.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.