Generation and verification of 27-qubit Greenberger-Horne-Zeilinger states in a superconducting quantum computer

Gary J Mooney, Gregory A L White, Charles D Hill and Lloyd C L Hollenberg

School of Physics, University of Melbourne, VIC, Parkville, 3010, Australia
School of Mathematics and Statistics, University of Melbourne, VIC, Parkville, 3010, Australia
E-mail: mooneyg@unimelb.edu.au, white.g@unimelb.edu.au, cdhill@unimelb.edu.au and lloydch@unimelb.edu.au

Keywords: quantum computing, quantum information, entanglement

Abstract
Generating and detecting genuine multipartite entanglement (GME) of sizeable quantum states prepared on physical devices is an important benchmark for highlighting the progress of near-term quantum computers. A common approach to certify GME is to prepare a Greenberger-Horne-Zeilinger (GHZ) state and measure a GHZ fidelity of at least 0.5. We measure the fidelities using multiple quantum coherences of GHZ states on 11 to 27 qubits prepared on the IBM Quantum ibmq_montreal device. Combinations of quantum readout error mitigation (QREM) and parity verification error detection are applied to the states. A fidelity of 0.546 ± 0.017 was recorded for a 27-qubit GHZ state when QREM was used, demonstrating GME across the full device with a confidence level of 98.6%. We benchmarked the effect of parity verification on GHZ state preparation embeddings on the heavy-hexagon architecture. The results show that the effect of parity verification, while relatively modest, led to a detectable improvement of GHZ fidelity.

Introduction
In the race to scale-up quantum computers and demonstrate quantum advantage, an important technical milestone is the generation of entanglement across a device. Entanglement—or the inability to factorise a multi-qubit system into separable states—is typically seen as the essence of what differentiates quantum behaviour from classical. Indeed, it has been shown that quantum systems with low or no amounts of entanglement can be simulated efficiently on a classical computer [1–3]. For this reason, the ability to generate and maintain genuine multipartite entanglement (GME) is fundamental for quantum information processors (QIPs) to outperform classical computers.

There are various benchmarks that indicate the capabilities of a given QIP. For example, quantum volume [4] is a holistic number that takes into account qubit number and error rates of a device. Although bipartite entanglement can be rigorously quantified, it remains an open problem to do the same for GME. To demonstrate and validate GME, we select a state that is known to be entangled across multiple qubits and assess how well a device can construct that state. In particular, Greenberger-Horne-Zeilinger (GHZ) states are well suited to this purpose—in that they are GME states whose fidelity on a QIP can be efficiently estimated. The fidelity estimate comes from a combination of measuring the populations of the qubit states as well as their coherences. An approach used to detect GME is to use GME witness operators [5–7]. A negative expectation value with respect to a target state is a sufficient but not necessary condition for the state containing GME. It has been shown that measuring a GHZ fidelity of at least 0.5 is equivalent to measuring a negative GME witness expectation value, hence implying that the state exhibits GME [8].

In this work we create large GHZ states up to 27 qubits on a physical device and measure their fidelities. The experiments are performed on the IBM Quantum ibmq_montreal device, which consists of 27 superconducting transmon qubits [9]. The device is from the series of IBM Quantum Falcon processors and was recently benchmarked at having a quantum volume of 64 [10]. When constructing states on QIPs, the actual entanglement of the states within the devices may be acceptably high, however the observed entanglement could
be betrayed by erroneous measurement results. Using a quantum detector tomography (QDT) technique, measurement errors can be sufficiently understood and classically inverted to estimate the pre-measurement states [11]. We employ a well-known quantum readout error mitigation (QREM) technique to implement measurement correction within our experiments [12] which has previously been used in the certification of an 18-qubit GHZ state [13]. We investigate and justify the assumptions of this method when applied to the prepared noisy GHZ states to ensure that the corrections made do not inflate the actual fidelity of the states within the device. GME has been demonstrated copiously in the literature across many different QIP architectures. A plot summarising this history of results for state sizes \( N \geq 3 \) qubits within gate-based quantum systems is shown in figure 1. With QREM applied, we record a fidelity of 0.546 \( \pm 0.017 \) with 98.6\% confidence for being above the 0.5 threshold for a 27-qubit GHZ state, which appears to be the largest demonstration of GME to-date [8, 13–41]. Error bars represent the standard error (of the mean). Beyond implementation of QREM, we investigate the use of parity verification on the fidelity of the GHZ states created. Parity verification is a fundamental error detection protocol used within quantum error correction schemes towards the realisation of large-scale fault-tolerant quantum computing. Ancilla qubits are used to measure the parity of state qubits to detect errors, enabling erroneous computations to be corrected or discarded. We benchmark the effects of parity verification on entanglement generation for various sized GHZ states prepared on the ibmq_montreal device. This work highlights the technical achievement in quantum hardware and the positive progress towards the realisation of practical quantum computers.

**Results**

**Detecting genuine multipartite entanglement in GHZ states**

GHZ states are highly entangled states. They can be prepared in gate-based quantum devices by initialising a single primary qubit to the \( |+\rangle \) state and the other qubits to the \( |0\rangle \) state, then CNOT gates are iteratively applied from the primary qubit (or any other qubit that has already had a CNOT applied in this manner) to each other qubit involved in the state, as shown in figure 2. A GHZ state can be expressed as

\[
|\text{GHZ}_N\rangle = \frac{|0\rangle \otimes N + |1\rangle \otimes N}{\sqrt{2}},
\]

where \( N \) is the number of qubits in the state. This state has a convenient structure for detecting genuine multipartite entanglement since its density matrix only consists of non-zero entries in each of the four corners. The GHZ fidelity, \( F \), can be calculated by measuring two observables population \( P \) and coherence \( C \) as
where $P = \rho_{00...0,00...1} + \rho_{11...1,11...1}$ can be directly measured as the GHZ populations and
$C = |\rho_{11...1,00...1} + |\rho_{00...0,11...1}|$ can be measured through parity oscillations [8, 22], or multiple quantum coherences (MQC) [13, 42–44]. A GHZ fidelity greater than 0.5 is sufficient to demonstrate that the state exhibits GME [8]. The following description of the process for measuring the fidelity is adapted from the details presented in appendix B. To measure the coherence, we use the method of MQC due to its promising robustness to noise by using a refocusing $\pi$-pulse similar to a Hahn echo [45] to refocus low frequency noise and reduce dephasing [13]. Since X gates stabilise the GHZ state, this does not affect fidelity computations, and the usual second pulse used to cancel the first in refocusing pulse sequences can be omitted. The coherence can be calculated by measuring the overlap signals $S_\phi = \text{Tr}(\rho_\phi \rho)$, where $\rho_\phi = e^{-i\phi \sum \sigma_z} \rho e^{i\phi \sum \sigma_z}$ is produced by rotating each qubit of the state $\rho$ by $\phi$ about the $Z$-axis. Essentially, $S_\phi$ is the probability of measuring the zero state after encoding the GHZ state, applying the $\phi$ phase rotations, then decoding the GHZ state. For an ideal GHZ state, the phase produced by rotating each individual qubit accumulates and is equivalent to adding a phase of $N\phi$ to the state, i.e. $\frac{1}{N^2}(|00...0\rangle + e^{-iN\phi}|11...1\rangle)$, which reduces the overlap signal to

$$F \equiv \frac{P + C}{2} \quad (2)$$

Figure 2. Example preparation circuits for 7-qubit GHZ states. Whenever a CNOT is applied within these circuits, the GHZ state grows in size by 1 qubit. (a) An inefficient embedding where all CNOTs are applied from the primary qubit, resulting in a CNOT depth of 6. (b) An optimal embedding where CNOTs are applied in parallel from qubits that are already included in the growing GHZ state, resulting in a CNOT depth of 3. When constructing GHZ states, we aim to prepare them using a similar embedding to this optimal one modulo the qubit layout of the quantum device.
$$S_{\phi}^{\text{ideal}} = \frac{1}{2} [1 + \cos(N\phi)], \quad (3)$$

where the constant term corresponds to the diagonal elements of the density matrix and the cosine term corresponds to the off-diagonal corner elements. By Fourier transforming $S_{\phi}$ we can obtain the MQC amplitudes

$$I_q = N^{-1} \sum_{\phi} e^{i\phi S_{\phi}}, \quad (4)$$

where $\phi = \frac{2\pi j}{N+1}$ for $j = 0, 1, \ldots, 2N + 1$ to detect up to frequency $N + 1$ and $N = 2N + 2$ is the number of angles $\phi$ in the summation. The coherence is then calculated as $C = 2\sqrt{I_0}$. There is a slow free rotation that occurs in idle qubits throughout the computation. To help alleviate these drift effects, a refocusing $\pi$-pulse applied between the GHZ encoding and decoding steps is used. The procedure to compute each of the overlap signals $S_{\phi}$ consists of the following steps.

1. Encode the GHZ state by applying the preparation circuit over the desired qubits.
2. Apply a refocusing $\pi$-pulse as an X gate on each qubit.
3. Apply the phase gate rotation of $\phi$ about the Z-basis.
4. Decode the GHZ state by applying the inverse of its preparation circuit.
5. The overlap signal $S_{\phi}$ is then the probability of measuring the zero state $|00 \ldots 0\rangle$.

Further details involved in expressing the fidelity concretely are presented in appendix B.

Quantum readout-error mitigation (QREM)

On noisy intermediate-scale quantum (NISQ) hardware, measurement represents one of the largest single-component error sources. This is especially significant in low-depth circuits where the number of qubit readouts is comparable to the total number of gates applied. Typical readout error rates of a few percent per qubit can quickly scramble the results of any quantum algorithm. In particular, they may obfuscate the results of an otherwise relatively well-performing device. Fortunately, however, readout error is far more addressable than other quantum errors on QIPs. First, excluding the case of mid-circuit measurement, readout errors do not propagate. Secondly, and more importantly, readout error can be treated as classical error to a good approximation for typical superconducting devices. We employ the quantum readout error mitigation (QREM) procedure developed in [12] in our experiments.

Readout error can be difficult to characterise. Formally, laboratory measurements output estimates of the object $\text{Tr}[E(UU^*)]$, for some noisy state $\rho$, intermediate unitary circuit $U$, and positive operator-valued measure (POVM) elements $E$. Typically, this POVM is the set $\{|0\rangle \langle 0|, |1\rangle \langle 1|\}$, i.e. measurement in the Z-basis. Any deviation from the ideal case in this object cannot be narrowed down to any one of the three components without further investigation. Even without intermediate operations, it is impossible to tell where an error falls on the continuum between a faulty state preparation with a perfect measurement, and perfect state preparation with a faulty measurement.

This noise falls under a category termed state preparation and measurement (SPAM) error. There are techniques equipped to self-consistently characterise SPAM [46], but are typically computationally costly. Instead, it often suffices as a good approximation for modern hardware to assume the former end of the spectrum: that is, a device capable of perfectly delivering a $|0\rangle$ state but with some error in measuring it. The magnitude of the different error sources are often two or more orders of magnitude apart, justifying the assumption. The benefit of this approximation is that the conditional probability distribution $p(\text{measure } x|\text{prepared } y)$ can be characterised simply by preparing all states $y$, and counting all measurements $x$. Whilst the number of preparation states to consider grows exponentially in the number of qubits, it can simplify matters greatly to assume that each faulty POVM element has a tensor product structure. That is, that the readout errors are either local-only or have correlations with limited spatial extent. In the case of solely local errors, the characterisation can be performed in a constant number of circuits—by preparing and measuring both $|0\rangle^{\otimes n}$ and $|1\rangle^{\otimes n}$. For limited local connections, this is constant with the number of qubits and exponential in the extent of the correlations.

For this work, we operate under the first-order regime of local errors. Each state $|\psi\rangle$ is transformed according to $|\psi\rangle = A_{\text{local}} |\psi\rangle$, where $A_{\text{local}} := \bigotimes_{i=1}^{n} A_i$. Each stochastic matrix $A_i$ is called the calibration matrix for qubit $i$ and is defined as
where the notation $p_i(x|y)$ indicates the probability of measuring the state $|x\rangle$ given the prepared state $|y\rangle$ for qubit $i$. The calibration matrices $A_i$ are constructed using the QDT [11] technique. In principle, the QREM procedure consists of collecting a vector of measurement outcomes and multiplying this by $A_i^{-1}$ in order to obtain the pre-readout state. In practice, however, imperfections in the calibration and the mitigation—caused at the very least by finite sampling error—can result in a final probability vector with negative elements. To combat this, a projection method is usually used to find the closest physical probability vector (on the unit simplex, with positive elements summing to 1) [47]. To summarise: a stochastic calibration matrix is constructed for each individual qubit; the tensor product of the inverse of each calibration matrix is multiplied by the measured results vector; finally, the closest probability vector is found, representing the best estimate of the quantum state before readout error. In appendix A, we estimate the size of the approximations made in performing QREM using this simplified procedure. The maximum difference in findings is small when compared to the error bars of our results.

The initial measured probability vector can only contain probability values that are at least as large as the equivalent of a single shot, $1/($shot count$)$, due to the finite number of shots, hence the maximum number of states with non-zero probability is equal to the shot count. However, applying the $A_i^{-1}$ to the probability vector often results in a high number of states having very small non-zero probabilities (sometimes much less than $1/100$ of a shot). The amount of memory required to precisely store a single probability vector after applying $A_i^{-1}$ no longer scales with the number of shots. Instead, the memory scales as $2^N$ for $N$ qubits, since sparse vectors are no longer useful. To help alleviate this computational resource requirement, we apply the inverse calibration matrices for each qubit $A_i^{-1}$ sequentially on the relevant elements of the probability vector. After every application of $A_i^{-1}$, probability values with magnitude below some small threshold are set to zero. A threshold value of $1/10$ of a shot was used, which equates to a probability of just over $1.2 \times 10^{-5}$ for 8192 shot experiments. This value was chosen because it was high enough to considerably reduce the overall computation time, while the approximation remains close to the full calculation. When testing this approximation on a 19-qubit state, the average computation time for each application of QREM dropped from 66 sec to 3.8 sec (performed on a laptop with 2.70 GHz Intel i7-7500U CPU and 16 GB RAM), and the fidelity was found to match the full calculation up to five decimal places. The reduced computation time enabled each sample from the 27-qubit GHZ state to be computed within two days.

Parity Verification

We examine the extent to which the fidelity of GHZ states can be improved through error-detecting state preparation techniques. In particular, we employ the parity-checking technique used in [48] for the $[4, 2, 2]$ code. The goal here is to extract information about errors occurring within the GHZ state without disturbing the state itself. This can be achieved with a parity check between qubits of the state, where two CNOTs (control $c$ → target $t$, denoted CNOT) are performed which are controlled on GHZ state qubits and target a single ancilla qubit initialised in the ground state. In the case of an even number of bit-flip errors, parity-checking will leave the state of the ancilla qubit unchanged. Alternatively, for an odd number of bit-flips, the state of the ancilla will flip, signalling a detectable error within the GHZ state. More explicitly, a parity check on qubits $q_1$ and $q_2$ using ancilla qubit $a$ will result in the transformations

$$\begin{align*}
|00\rangle_{q_1,q_2} |0\rangle_a &\rightarrow |00\rangle_{q_1,q_2} |0\rangle_a, \\
|01\rangle_{q_1,q_2} |0\rangle_a &\rightarrow |01\rangle_{q_1,q_2} |1\rangle_a, \\
|10\rangle_{q_1,q_2} |0\rangle_a &\rightarrow |10\rangle_{q_1,q_2} |1\rangle_a, \\
|11\rangle_{q_1,q_2} |0\rangle_a &\rightarrow |11\rangle_{q_1,q_2} |0\rangle_a.
\end{align*}$$

Measuring the ancilla in the $|1\rangle_a$ state implies a bit-flip error has occurred within the GHZ state, allowing the erroneous shot measurement to be discarded from the results. In the case of measuring the $|11\rangle_{q_1,q_2} |0\rangle_a$ state, the two bit-flip errors are not detected. However, if $p$ is the single-qubit probability of a bit-flip occurring for qubits $q_1$ and $q_2$, then the probability of bit-flips occurring on both $q_1$ and $q_2$ is small $p^2$. For example, assuming bit-flip probabilities are moderate at $p = 0.25$, if $|0\rangle_a$ is already measured (assuming no readout errors), then there is a probability of $(1 - p)^2/(1 - p^2) = 0.9$ that the state has no bit-flip errors $|00\rangle_{q_1,q_2}$ and a probability of $p^2/(1 - p^2) = 0.1$ that the state has two bit-flip errors $|11\rangle_{q_1,q_2}$.

Under the typical framework of quantum error correction, the syndrome measurement of an ancilla qubit will be fed forward to correct the original state. Since error correction is not achievable within the current restrictions of quantum hardware, we instead focus on detection so that runs can be post-selected. This is usually seen as the condition under which fault-tolerance can be demonstrated on near-term quantum devices [48].
Fidelity of GHZ states with parity verification in the ibmq_montreal device

In this section, we measure the fidelity on a range of GHZ state sizes prepared on the ibmq_montreal device and observe the effects of using parity-checking qubits to verify them. The following experiments prepare GHZ states and parity verification on ibmq_montreal’s qubit layout via the diagrams shown in figure 3. The parity verification CNOTs are performed directly after preparing the GHZ states. In the case of measuring the coherence, these CNOTs are in the middle of the circuit, since there is a GHZ decoding sequence of gates applied afterwards. Ideally the parity qubits would be measured immediately after their corresponding CNOTs have been applied to help reduce parity qubit error and abandon the computation as soon as possible in the case of an error, however superconducting devices do not yet robustly support this functionality and hence the parity qubits are measured at the end of the experiment. The population is calculated by summing the measured 0 … 0 and 1 … 1 occupancies from the prepared GHZ states and the coherence is calculated by using the MQC method. Additionally, QREM was used on GHZ state qubits to reduce the effects of noise due to measurement errors. The parity-checking qubit cannot have QREM applied because it is a single qubit within the device calibration and are shown in figure 4, the overlap signals for state sizes 11, 13, 15, and 17 for the corresponding CNOT circuit depths and counts for each experiment are shown in table 1.

GHZ states of incrementally increasing sizes from 11 to 19 qubits were prepared using the embedding shown in figure 3(a). The smallest state of 11 qubits is the minimum size required to include a parity-checking qubit within the ibmq_montreal device’s layout. The initial Hadamard is performed on qubit 2 and CNOTs grow the state in the up-right and down-right directions. The parity qubit is at location 13 and was added to the circuits by applying two CNOTs with control qubits 12 and 14 and target qubit 13 directly after preparing the GHZ states. The resulting fidelity, population and coherence measurements were obtained within the same device calibration and are shown in figure 4, the overlap signals for state sizes 11, 13, 15, and 17 for the calculation of the coherences are shown in figure 6(a). Performing the parity-checking verification requires the CNOT circuit depth to increase by one for states of size 11 to 14, while the depth is unchanged for states of size 15 to 19 since the parity-checker CNOTs are performed in parallel to the preparation circuit. When using the parity check, there is a noticeable improvement to fidelity for state sizes that do not require the CNOT circuit depth to increase, namely for state sizes 15 to 19, where the average fidelity increases by 0.051 ± 0.007 with QREM and 0.031 ± 0.004 without QREM. The population values gain a consistent advantage through parity verification, while the coherence values appear mostly unchanged for sizes 15 to 19 and slightly decrease for sizes 11 to 14. The decrease in coherence is likely due to the inclusion of the additional CNOTs increasing the CNOT circuit depth and introducing noise. Measurement error on the parity-checking qubit could also be a contributing factor by resulting in a small number of the wrong circuits being discarded. For all of the experiments, the standard errors were calculated from 8 independent runs of the experiment where all circuits were performed using 8192 shots. All computations were performed on the Spartan high performance computing system [49]. The corresponding CNOT circuit depths and counts for each experiment are shown in table 1.

Figure 3. Diagrams showing how the GHZ states are constructed on the ibmq_montreal device. Blue vertices represent qubits of the GHZ state, green vertices represent parity checking qubits, and lighter beige vertices represent qubits that are not directly involved in the experiment. Arrows represent the direction of CNOTs (control (c) → target (t), denoted CNOTc→t). Dark blue arrows indicate CNOTs used for constructing the GHZ state for all state sizes, light blue arrows indicate CNOTs growing the GHZ state, orange arrows indicate CNOTs used for parity verification, dotted beige lines indicate that no two-qubit gates are applied. (a) Shows the construction of the state sizes 11 to 19 beginning at qubit 2 and including the single parity checking qubit 13. The CNOT2 gate is performed before the CNOT17 gate. The states are incrementally grown from the initial 11 qubits by including qubits in the order: 15, 16, 18, 19, 21, 22, 23, and 25. (b) Shows the construction of the state sizes 19 to 25 beginning at qubit 13 and including the two parity checking qubits 2 and 14. The initial CNOTs are performed in the order CNOT1, CNOT2, CNOT3, CNOT12, CNOT15, CNOT16, and CNOT25. The states are incrementally grown from the initial 19 qubits by including qubits in the order: 9, 6, 20, 17, 0 and 26. The 26 and 27-qubit states do not use qubits 2 and 24 as parity checking qubits, instead the 26-qubit GHZ state includes qubit 2 with a CNOT1 in its construction while the 27-qubit state additionally includes qubit 24 with a CNOT22 gate.
way around. This leads to population measurements being far more sensitive to measurement error than the coherence curve (since the occupancies of both $|00\ldots0\rangle$ and $|11\ldots1\rangle$ are required to measure the population while only the occupancy of $|00\ldots0\rangle$ is required for the coherence).

To further explore the effects of parity verification on larger states, GHZ states were prepared for incrementally increasing sizes from 19 to 25, with and without two parity-checking qubits. The ibmq_montreal device has a convenient layout for preparing larger GHZ states, since it allows a relatively high number of branches when constructing them, allowing more CNOTs to be computed in parallel as shown in figure 3(b). The initial Hadamard is performed on qubit 13 and CNOTs grow the state in the four directions up-left, up-right, down-right, down-left. The parity-checking qubits are located at 2 and 24 and were added to the circuits by applying, directly after preparing the GHZ states, four CNOTs, two with control qubits 1 and 3 and target qubit 2, and two with control qubits 23 and 25 and target qubit 24. Due to the parallel nature of this state embedding, each of the preparation circuits for sizes 19 to 25 have equal CNOT circuit depths, with depth increasing by one when using parity verification for sizes 19 to 23 and increasing by two for sizes 24 and 25. The resulting fidelity, population and coherence measurements were obtained within the same device calibration and are shown in figure 5, the overlap signals for state sizes 19, 22, and 25 are shown in figure 6(b). When using parity verification, the fidelity over state sizes 19 to 25 increases on average by $0.072 \pm 0.009$ with QREM and $0.045 \pm 0.007$ without QREM. In particular, for the 25-qubit GHZ states, fidelities of $0.664 \pm 0.016$ and $0.601 \pm 0.015$ were measured using QREM with and without parity verification respectively. The measured population values are consistently higher for all state sizes, while the measured coherence values show no increase in any of the state sizes and in particular, the coherence values decrease slightly for state sizes 19 to 21 qubits.

To test the extent of possible GME within the device, GHZ states of sizes 26 and 27 qubits were additionally prepared on the device (with embedding shown in figure 3(b)). The experiments were performed on the same calibration as the previous experiments relating to the GHZ state sizes of 19 to 25 and the results are shown in figure 5, with corresponding overlap signals for the calculation of coherence shown in figure 6(b). The measured fidelities using QREM were found to be $0.580 \pm 0.022$ and $0.546 \pm 0.017$ for the 26 and 27-qubit states respectively. These fidelities are above the 0.5 threshold with confidence levels of 99.6% and 98.6% respectively, thus GME was detected in both states.

We suspect it is possible to extend parity verification to consist of three or more qubits with majority rules, instead of a single qubit, to reduce measurement error. This could be beneficial because current measurement error mitigation techniques cannot be applied to parity qubits due to the parity being a single shot measurement. However, it may not be worth it due to introduced CNOT errors potentially being larger than the original measurement error.

| State size (qubits) | Population Depth | Population Count | Population (parity) Depth | Population (parity) Count | Coherence Depth | Coherence Count | Coherence (parity) Depth | Coherence (parity) Count |
|---------------------|------------------|-------------------|----------------------------|----------------------------|----------------|----------------|--------------------------|--------------------------|
| Embedding 1—GHZ state sizes 11 to 19 (one parity-checker qubit) | | | | | | | | |
| 11 | 6 | 10 | 7 | 12 | 12 | 20 | 13 | 22 |
| 12 | 6 | 11 | 7 | 13 | 12 | 22 | 13 | 24 |
| 13 | 7 | 12 | 8 | 14 | 14 | 24 | 15 | 26 |
| 14 | 7 | 13 | 8 | 15 | 14 | 26 | 15 | 28 |
| 15 | 8 | 14 | 8 | 16 | 16 | 28 | 16 | 30 |
| 16 | 8 | 15 | 8 | 17 | 16 | 30 | 16 | 32 |
| 17 | 9 | 16 | 9 | 18 | 18 | 32 | 18 | 34 |
| 18 | 9 | 17 | 9 | 19 | 18 | 34 | 18 | 36 |
| 19 | 10 | 18 | 10 | 20 | 20 | 36 | 20 | 38 |
| Embedding 2—GHZ state sizes 19 to 27 (two parity-checker qubits) | | | | | | | | |
| 19 | 7 | 18 | 8 | 22 | 14 | 36 | 15 | 40 |
| 20 | 7 | 19 | 8 | 23 | 14 | 38 | 15 | 42 |
| 21 | 7 | 20 | 8 | 24 | 14 | 40 | 15 | 44 |
| 22 | 7 | 21 | 8 | 25 | 14 | 42 | 15 | 46 |
| 23 | 7 | 22 | 8 | 26 | 14 | 44 | 15 | 48 |
| 24 | 7 | 23 | 9 | 27 | 14 | 46 | 16 | 50 |
| 25 | 7 | 24 | 9 | 28 | 14 | 48 | 16 | 52 |
| 26 | 7 | 25 | … | … | 14 | 50 | … | … |
| 27 | 7 | 26 | … | … | 14 | 52 | … | … |
Discussion

A variety of GHZ states were prepared on the 27-qubit \textit{ibmq\_montreal} device and the GHZ fidelities, populations and coherences were measured. We report the preparation of a 27-qubit state with a GHZ fidelity of $0.546 \pm 0.017$, demonstrating GME across the full device with a confidence level of 98.6%. This is currently the largest state prepared on a quantum system that has been reported to exhibit GME. We further benchmarked the effects of parity verification on the fidelity, population and coherence. Two GHZ state preparation embeddings on the physical qubit layout were used. The first embedding generated GHZ states of incrementally increasing sizes from 11 to 19 qubits while using a single parity-checking qubit. When the parity-checking qubits did not require the CNOT depth of the circuit to increase, which was the case for state sizes 15 to 19, the fidelity increased by an average of $0.051 \pm 0.007$ with QREM and $0.031 \pm 0.004$ without QREM. The fidelity was mostly unchanged for sizes 11 to 13 and was higher for size 14. The lack of fidelity improvement for state sizes 11 to 13 was most likely due to the noise introduced by the CNOTs of the parity-checking qubit and their implementation resulting in the CNOT circuit depth increasing by one for state sizes 11 to 14. The population values increased for all state sizes when the parity-checking qubit was used, while the coherence values appear mostly unchanged for state sizes from 15 to 19 and slightly decrease for sizes from 11 to 14. The second embedding generated GHZ states of incrementally increasing sizes of 19 to 25 qubits while using two parity-checking qubits. It utilised the layout of the \textit{ibmq\_montreal} device by performing as many CNOTs as possible in
parallel to minimise the CNOT circuit depth. The parity-checking qubits increased the CNOT circuit depth for state sizes 19 to 23 by one and for sizes 24 and 25 by two. We observed that the parity-checking procedure with two qubits increased the fidelity slightly for all state sizes resulting in an average increase of $0.072 \pm 0.009$ with QREM and $0.045 \pm 0.007$ without QREM. In particular, the 25-qubit state, which was the largest state tested using parity checkers, had its fidelity increase from $0.601 \pm 0.015$ to $0.664 \pm 0.016$ with QREM. The population values consistently increased for all state sizes, while the coherence values decreased slightly for state sizes 19, 20, 21 and 25 qubits and were unchanged for state sizes 22, 23 and 24. The results show that the effect of parity verification led to a detectable improvement of GHZ fidelity, although relatively modest on current IBM Quantum hardware.

**Acknowledgments**

This work was supported by the University of Melbourne through the establishment of an IBM Q Network Hub at the University. CDH is supported by a research grant from the Laby Foundation.

Figure 5. Embedding 2. Measured observables for GHZ states varying in sizes from 19 to 27 qubits on the ibmq_montreal device. These states are prepared using the more efficient embedding shown in figure 4(b) lowering CNOT circuit depths resulting in higher fidelity values for large states than what is shown in figure 4. Each plot compares combinations of applying quantum readout-error mitigation (QREM) and using parity-checker qubits 2 and 24 to verify the equality of states of qubits 1 and 3, and qubits 23 and 25 respectively. The parity $\Delta$ refers to the difference between using and not using parity verification. The preparation of each state without parity verification requires the same CNOT circuit depth since the additional qubits can be included by applying CNOT gates in parallel. When including the two parity-checking qubits, the CNOT circuit depth increases by one for sizes 19 to 23 and increase by two for sizes 24 and 25. The standard errors were calculated from 8 independent runs of the experiment where all circuits were performed using 8192 shots. All corresponding CNOT circuit depths and counts are displayed in table 1. (a) The population is calculated as the probability of obtaining $|00\ 0\rangle$ or $|11\ 1\rangle$ upon measurement of the prepared GHZ state. Parity verification using the two parity-checking qubits is shown to significantly increase the population values for all state sizes from 19 to 25 qubits. (b) The coherence is calculated using the quantum multiple coherences (QMC) method. No increase in coherence is observed using parity verification for state sizes 19 to 25. For state sizes 19 to 21 in particular, the measured coherence decreases slightly. (c) The fidelity is calculated as the average over the population and coherence. When using parity verification, the fidelity increases on average by $0.072 \pm 0.009$ with QREM and $0.045 \pm 0.007$ without QREM over state sizes 19 to 25.
The data that support the findings of this study are available upon reasonable request from the authors.

Author contributions statement

GJM, GALW, CDH and LCLH conceived the experiments. GJM obtained and processed the main results, while GALW obtained and processed the QREM analysis of appendix A. GJM and GALW wrote the manuscript with input from CDH and LCLH.
Additional information

Non-financial competing interests: The authors are supported by the University of Melbourne through the establishment of an IBM Q Network Hub at the University.

Appendix A. Examining QREM assumptions on device hardware

In the main text, we briefly outlined some simplifying assumptions made in order to implement quantum error mitigation more efficiently. Respectively, these were: measurement error is significantly larger than preparation error, measurement error is predominantly classical, and measurement error is uncorrelated between qubits.

Here, we examine these assumptions on chosen subsets of the device. To this end, we employ gate set tomography (GST) for characterisation. GST is a self-consistent extension of regular quantum process tomography (QPT) which includes the SPAM operations in its maximum likelihood estimation. A characterisation of the ith qubit is performed across an entire gate set, \( \{ \rho_i^0, E_i^0, E_i^1, G_i^0, G_i^1, \cdots \} \), whereby the preparation and measurement operators are also estimated. Importantly, this allows us to robustly evaluate the effects of the liberties we have taken. For an overview of GST and its implementation, see [46, 50, 51]. We use the well-known quantum characterisation, verification and validation (QCVV) Python package PyGSTi for this part of the implementation [52].

Testing these assumptions takes place in three parts. First, we perform GST across a range of qubits to obtain an estimate for \( \rho_i^0, E_i^0, \) and \( E_i^1 \), where \( \rho_i^0 \) is the initial density matrix, and \( E_i^0 \) is the POVM element corresponding to when the detector clicks with the result \( m \). Next, using these results, we compare the magnitudes of errors in \( \rho_i^0 \) with \( E_i^0 \). We also determine the off–diagonals of \( E_i^0 \); non-zero values correspond to coherence in the POVM effect, which is a strictly non–classical error. These values roughly represent the information discarded when the decision is made to use the stochastic calibration matrix (equation (5)) rather than the complete POVM. From the POVM, we construct the calibration matrix \( A_i^{\text{POVM}} \). This has the same definition as \( A_i \), in equation (5), except its values are obtained with a more comprehensive procedure. We compare these matrices with the end result obtained from the simpler method in the main text. A similar investigation into the rigours of QREM has been conducted in [53]. Ideally, the GST estimates for the noisy probes would be used as the basis for QREM, however for our purposes this would have greatly increased the experimental overhead. The number of experiments required for the chosen GST characterisation of a single qubit is 350, whereas the prepare-and-measure path takes only two circuits with no post-processing.

Finally, having validated the use of the simpler, cheaper stochastic matrix, we examine the assumption of locality of the measurement error. To do this, we select sets of four qubits. We construct the full calibration matrix across the four qubits—i.e. by preparing \( \{ |i\rangle \}_{i=0}^{15} \) and measuring. We then compare this to the tensor product of the four local calibration matrices. We note in hindsight that there have been recent developments for characterising correlated measurement error in an efficient manner [54]. Although our results do not suggest that the tensor product structure significantly impacts the outcome, this could be a more desirable method in future.

A.1. SPAM magnitudes and the calibration matrix

Across the device, we performed GST on six different qubits. From this, we obtain estimates of the initial density matrix \( \rho_i^0 \), and the two-outcome POVM \( \{ \{ 0 \} \{ 0 \}, \{ 1 \} \{ 1 \} \} \). We also obtain estimates for the gates that comprise the process, but these are not relevant to the discussion. Let \( E_i^m \) denote one of the noisy two-outcome effect operations for the ith qubit. Our aim is firstly to determine how classical the error on \( E_i^m \) is, as well as estimating the size of the error made in assuming perfect state preparation. The latter is not necessarily as simple as finding the size of the preparation error. Rather, we use our estimates \( E_i^m \) to determine the calibration matrix that we would obtain if state preparation were perfect. Using the Born rule, the elements of this matrix \( A_i^{\text{POVM}} \) are:

\[
A_i^{\text{POVM}} = \begin{pmatrix}
A_{0,0} & A_{0,1} \\
A_{1,0} & A_{1,1}
\end{pmatrix}
\]

\[
A_{0,0}^i = p_i(0|0) = \text{Tr} [E_i^0|0\rangle \langle 0|] = E_{0,0}^i,
\]
\[
A_{0,1}^i = p_i(0|1) = \text{Tr} [E_i^0|1\rangle \langle 1|] = E_{0,1}^i,
\]
\[
A_{1,0}^i = p_i(1|0) = \text{Tr} [E_i^1|0\rangle \langle 0|] = E_{1,0}^i,
\]
\[
A_{1,1}^i = p_i(1|1) = \text{Tr} [E_i^1|1\rangle \langle 1|] = E_{1,1}^i.
\]

(A1)

A subtlety in this practice is that \( \rho^0 \) and \( E_i^m \) are not uniquely fixed by any possible experiment; a type of gauge freedom exists in the estimate. Let \( \Lambda(\cdot) \) be some error channel acting on either SPAM operation, and let \( G \) be some gate sequence. If \( \Lambda \) commutes with \( G \) then we have
\[ p_i^m = \text{Tr}[G\rho_i^0 G^\dagger \cdot A(E_i^m)] = \text{Tr}[G\Lambda(\rho_0) G^\dagger \cdot E_i^m]. \] (A2)

That is, there is a class of estimates for \( \rho_i^0 \) and \( E_i^m \) which are completely physically equivalent under commuting transformations. An error map which fits this description is the depolarising channel \( \mathcal{E}_p \):

\[ \mathcal{E}_p(\rho) = (1 - p)\rho + \frac{p}{2}\mathbf{1}. \] (A3)

Therefore, in comparing the GST-produced calibration matrix \( A_i^{\text{POVM}} \) with \( A_i \), we search all variations of

\[ E_i^0 \mapsto (1 - p_i)E_i^0 + \frac{p_i}{2}\mathbf{1}, \]
\[ E_i^1 \mapsto (1 - p_i)E_i^1 + \frac{p_i}{2}\mathbf{1}, \text{ and} \]
\[ \rho_i^0 \mapsto \rho_i^0 + p_i|0\><0| - \frac{p_i}{2}\mathbf{1}. \] (A4)

which preserve the positivity of \( \rho_i^0 \). Since these representations of the SPAM operations are physically indistinguishable, we can determine the closest \( A_i^{\text{POVM}} \) to \( A_i \) for each qubit \( i \). The remaining difference between the two will estimate the error produced in constructing the calibration matrix using the simpler methods of the main text. For qubits 1, 3, 12, 14, 23, and 25, we summarise the results of this process in table 2.

The first three data columns respectively give direct estimates of the POVM 0-effect, the GST-derived calibration matrix, and the simplified calibration matrix. The final two columns respectively quantify from these the size of the approximation made from assuming perfect preparation and from assuming classical errors. The experiments were conducted with 8192 shots. Each of these values are small, both compared to the sampling error, and compared to the results in the main text.

We further examine these assumptions by conducting simulated experiments using the experimentally characterised data. Here, we aim to replicate the conditions of our results as much as possible in order to determine if there is any possibility of overestimating entanglement by using QREM. The procedure for this is as follows:

- Generate a set of possible 6-qubit lab states: \( \rho = (1 - p)\cdot \rho_{\text{GHZ}} + p \cdot \rho_{\text{noise}} \) for a range of \( p \) and different noise (in particular: maximal mixtures and random states),
- Simulate out MQC with measurement error by using the GST estimates of the measurement probes \( \{E_i\} \) to produce the complete coherence curve,
- Then, using the measure-and-prepare calibration matrix \( A_{\text{local}} \), we apply QREM and invert the noise of the measurement effects,
- Finally, determine the final error-mitigated fidelity, and compare to the actual state \( \rho \).

We performed this across a variety of different \( \rho_{\text{noise}} \) and values for \( p \), each sampled 500 times. The maximum average fidelity overestimate found with QREM was \( 6 \times 10^{-4} \). We suspect that the MQC method of estimating state fidelity is particularly robust to erroneous QREM. The reason for this is that QREM carries the possibility of overestimating the number of \( |0\rangle^\otimes n \) readouts. However, MQC creates a curve measuring \( |0\rangle^\otimes n \) amplitudes from peak to trough and then estimates the fidelity from the frequency here, which is more robust to translations of the entire curve.

### A.2. Correlation of measurement error

The final assumption made in the name of circuit reduction is that of locality of the errors. In the previous subsection, we justified that single-qubit measurement errors can be accurately characterised for the purpose of QREM on this hardware using only two circuits. An extrapolation to \( n \) qubits would still require \( 2^n \) circuits to create the full calibration matrix. Instead, if we assume that measurement errors on different qubits are uncorrelated, we can take \( A_{\text{local}} = \bigotimes_{i=1}^{n} A_i \) for the \( n \) qubit system.

To check whether this is valid, we collected calibration matrices for four sets of four qubits and compared them to the tensor product of their marginals. Here, \( A_{\text{full}} \) is the calibration matrix for all 16 prepared and measured states. Meanwhile, \( A_{\text{local}} \) is the tensor product of local calibration matrices obtained from the preparation and measurement of \( |0000\rangle \) and \( |1111\rangle \). These results are summarised in figure 7. Here, we present matrix plots of each calibration matrix with the ideal case subtracted off. Each construction presents an almost identical depiction of the noise. One exception is the small occurrence of non-zero elements in \( A_{\text{full}} \) for the set (15, 18, 17, 21) which is not present in \( A_{\text{local}} \)—however we believe this is not large enough to warrant accounting for. The difference between the local and full calibration matrices are quantified by their Frobenius distance.
Table 2. Experimentally determined values in QDT. This includes estimates for the $|0\rangle \langle 0|$ measurement, a comparison of calibration matrices constructed with both GST and prepare-and-measure methods, and an estimate of the quantumness of the measurement effect.

| Qubit # | $E_i^0$ | $A_i^{POVM}$ | $A_i$ | $\|A_i - A_i^{POVM}\|$ | $\frac{1}{2}(|E_i^0|_1 + |E_i^0|_{1,0})$ |
|---------|---------|-------------|------|----------------|------------------|
| 1       | 0.9913  | 0.9913      | 0.9803| 0.00231         | 0.0046           |
|         | $-0.0046 + 0.0255i$ | 0.0087 | 0.0197 | 0.9499         |
| 3       | 0.9964  | 0.9964      | 0.9945| 0.0051          | 0.0002           |
|         | $-0.0002 + 0.0024i$ | 0.0036 | 0.0055 | 0.9877         |
| 12      | 0.9648  | 0.9648      | 0.9683| 0.0074          | 0.0012           |
|         | $-0.0012 + 0.0014i$ | 0.0352 | 0.0317 | 0.9349         |
| 14      | 0.9928  | 0.9928      | 0.9915| 0.0075          | 0.0004           |
|         | $0.0004 + 0.0019i$ | 0.0072 | 0.0085 | 0.9779         |
| 23      | 0.9713  | 0.9713      | 0.9752| 0.0082          | 0.0003           |
|         | $-0.0003 + 0.0049i$ | 0.0287 | 0.0248 | 0.9550         |
| 25      | 0.9960  | 0.9960      | 0.9961| 0.0002          | 0.0038           |
|         | $0.0038 + 0.0024i$ | 0.0040 | 0.0039 | 0.9855         |
all cases, this difference (representing the difference between 256 matrix elements) is small, and we are confident that results are not affected by making this simplifying assumption.

Appendix B. Using multiple quantum coherences to detect GME within noisy GHZ states

GHZ states defined over arrays of qubits are more susceptible to noise than graph states defined over the same layout of qubits [55]. So to detect entanglement, the graph state is generally the preferred option. However the GHZ state has a convenient structure that can be utilised for determining the more strict GME property, by calculating its fidelity. The fidelity between a desired pure state \( \rho_{\text{ideal}} \) and the actual state \( \rho \) is a similarity metric which can be written as

\[
F(\rho, \rho_{\text{ideal}}) = \text{Tr}(\rho \rho_{\text{ideal}}),
\]

When \( \rho_{\text{ideal}} \) is a GHZ state, a fidelity value that is greater than or equal to 0.5 implies that \( \rho \) is GME. The fidelity of a GHZ state can be written as

\[
F = \frac{P + C}{2},
\]

where the population \( P = \rho_{00\ldots00\ldots0} + \rho_{11\ldots11\ldots1} \) can be directly measured as the sum of the \( |00\ldots0\rangle \) and \( |11\ldots1\rangle \) occupancies from the GHZ state and the coherence \( C = |\rho_{11\ldots10\ldots0}| + |\rho_{00\ldots01\ldots1}| \) can be indirectly measured through Multiple Quantum Coherences (MQC) [13] or parity oscillations [8, 22]. In this work, we use MQC since a refocusing \( \pi \)-pulse similar to a Hahn echo [45] can be used to refocus low frequency noise and reduce dephasing [13]. This is because MQC is more robust to noise and scales better with respect to measurement error mitigation. MQC is a technique that has been adapted from solid state nuclear magnetic resonance [42] to many-body correlations and quantum information scrambling in trapped ions [43, 44]. More recently, it has been used to detect 18-qubit GME in the IBM Q System One quantum device [13]. MQC works by utilising the amplified phase accumulated when phase rotating individual qubits. To help properly understand MQC and how it can be applied to calculate the fidelity, we will go through, in detail, steps to express the fidelity more concretely. The following working has been adapted from Wei et al [13] and Gärttner et al [44].

We begin by writing down the following expression for the density matrix

\[
\rho = \sum_{m,m'} \rho_{m,m'} |m\rangle \langle m'|,
\]

where the basis states \( |m\rangle \) are an equal superposition of each \( N \)-qubit state where there are \( m \) qubits in the \( |1\rangle \) state and \( N - m \) in the \( |0\rangle \) state. The state \( |m\rangle \) satisfies

![Figure 7. A comparison of measurement calibration matrices across samples of four over the device. We compare the full calibration matrix to that which is acquired by taking the tensor product of all local terms. In the matrix plots, we subtract off the identity matrix in order to better resolve finer details.](image-url)
\[
\sum_{j=1}^{N} \sigma_{j}^{2} |m\rangle = A_{m} |m\rangle,
\]

where the eigenvalues \(A_{m} = m - N/2\) form the set \{-\(N/2\), \(-N/2 + 1, \ldots, N/2 - 1, N/2\}\}. The density matrix can be expressed as a sum of different coherence sectors

\[
\rho = \sum_{q=-N}^{N} \rho_{q}
\]

where

\[
\rho_{q} = \sum_{m=0}^{N} \rho_{m,m-q} |m\rangle \langle m-q|,
\]

noting that \(\rho_{ij} = 0\) when \(i < 0\) or \(i > N\). The coherence sectors \(\rho_{q}\) account for the coherence between states \(|m\rangle\) and \(|m' = m - q\rangle\) such that \(A_{m} - A_{m'} = q\). Note that \(\rho\) being Hermitian implies that \(\rho_{q} = \rho_{-q}\) and each \(\rho_{q}\) being orthogonal to one another implies \(\text{Tr}(\rho_{q} \rho_{p}) = \delta_{q-p} \text{Tr}(\rho_{q} \rho_{q})\). Although the states \(\rho_{q}\) cannot be directly observed, we will show that the multiple quantum coherence amplitude defined as

\[
I_{q} := \text{Tr}(\rho_{q} \rho_{-q}) = \text{Tr}(\rho_{-q} \rho_{q}) = I_{-q},
\]

can. Since the density matrix of an ideal GHZ state only has non-zero values on each corner, the fidelity can be expanded as follows

\[
F = \text{Tr}(\rho \rho_{\text{GHZ}}) = \text{Tr}(\rho_{0} \rho_{0,0}^{\text{GHZ}}) + \text{Tr}(\rho_{N} \rho_{N,0}^{\text{GHZ}}) + \text{Tr}(\rho_{0} \rho_{N,0}^{\text{GHZ}}) + \text{Tr}(\rho_{N} \rho_{0,0}^{\text{GHZ}}).
\]

We will now work towards modifying the expression to only consist of directly measurable observables. By substituting values 0 and \(N\) for \(q\) in equation (B6), we get

\[
\rho_{0}^{\text{GHZ}} = \frac{1}{2} |00 \ldots 0\rangle \langle 00 \ldots 0| + |11 \ldots 1\rangle \langle 11 \ldots 1|,
\]

\[
\rho_{N}^{\text{GHZ}} = \frac{1}{2} |11 \ldots 1\rangle \langle 00 \ldots 0|,
\]

\[
\rho_{0} = \rho_{0,0} |00 \ldots 0\rangle \langle 00 \ldots 0| + \rho_{0,N} |11 \ldots 1\rangle \langle 11 \ldots 1|
\]

\[
\rho_{N} = \rho_{N,0} |11 \ldots 1\rangle \langle 00 \ldots 0|.
\]

From equation (B10), due to orthogonality of the basis states, observe that

\[
\text{Tr}(\rho_{0} \rho_{0}^{\text{GHZ}}) = \frac{1}{2} (\rho_{0,0,0} + \rho_{11,11,11})
\]

which is the half population term shown in equation (B2). The remaining two terms in equation (B9) will be expressed in terms of the amplitude \(I_{N}\) introduced in equation (B7). From equations (B11) and (B13) we can write \(\rho_{N} = \kappa \rho_{N}^{\text{GHZ}}\) for some complex constant \(\kappa\). By substituting this into equation (B7), it can be shown that

\[
\text{Tr}(\rho_{N} \rho_{N}^{\text{GHZ}}) = \kappa \text{Tr}(\rho_{N}^{\text{GHZ}}) = \frac{1}{2} \kappa I_{N},
\]

\[
\text{Tr}(\rho_{0} \rho_{N}^{\text{GHZ}}) = \frac{1}{2} \kappa I_{N},
\]

\[
\text{Tr}(\rho_{N} \rho_{0}^{\text{GHZ}}) = \frac{1}{2} \kappa I_{N},
\]

\[
= 2 \sqrt{I_{N}} (\text{Tr}(\rho_{N}^{\text{GHZ}} \rho_{N}^{\text{GHZ}}) + \text{Tr}(\rho_{0}^{\text{GHZ}} \rho_{0}^{\text{GHZ}}))
\]

\[
= \sqrt{I_{N}}.
\]

Finally, we can rewrite the fidelity as

\[
F = \frac{1}{2} (\rho_{0,0,0} + \rho_{11,11,11}) + \sqrt{I_{N}},
\]

with population \(P = \rho_{0,0,0} + \rho_{11,11,11}\) and coherence \(C = 2 \sqrt{I_{N}}\).

Now that we have expressed the coherence with respect to the quantum coherence amplitude \(I_{N}\), we will show how \(I_{N}\) can be measured. It can be calculated from the overlap signals \(S_{f} := \text{Tr}(\rho_{f} \rho)\), which is a directly measurable quantity where \(\rho_{f} := e^{-i \sum_{j} \phi_{j} |j\rangle \langle j|} e^{i \sum_{j} \phi_{j} |j\rangle \langle j|}\) is produced by rotating each qubit of the state \(\rho\) by \(\phi\) about the \(Z\)-basis. For an ideal GHZ state, this is equivalent to adding a phase of \(N\phi\) to the state, that is

\[
\frac{1}{\sqrt{N}} (|00 \ldots 0\rangle + e^{-iN\phi} |11 \ldots 1\rangle),
\]

which has the overlap signal reduce to
where the constant term corresponds to the diagonal elements of the density matrix and the cosine term corresponds to the off-diagonal corner elements. To see the behaviour of the overlap signal when more states are included, we can consider the case of general pure states and express $\rho$ as

$$\rho_{\text{pure}} := \sum_{m} a_{m} |m\rangle \langle m|,$$

(B20)

where $a_{m}$ is the amplitude for the state $|m\rangle$ where $|m\rangle$ is defined as in equation (B3). With the help of equation (B4), we can write

$$\rho_{\text{pure}} := e^{-i\frac{\pi}{2}\sum_{j}i_{j}^{2}} \sum_{m=0}^{N} a_{m} |m\rangle \langle m| e^{-i\frac{\pi}{2}\sum_{j}i_{j}^{2}} \sum_{m'=0}^{N} \langle m'| a_{m}^{*} e^{-i\frac{\pi}{2}\sum_{j}i_{j}^{2}} |m\rangle$$

(B21)

$$= \sum_{m=0}^{N} a_{m} e^{-i\frac{\pi}{2}(m-N/2)} |m\rangle \sum_{m'=0}^{N} \langle m'| a_{m}^{*} e^{-i\frac{\pi}{2}(m'-N/2)} |m\rangle$$

(B22)

$$= \sum_{m=0}^{N} a_{m} e^{-i\frac{\pi}{2}m} |m\rangle \sum_{m'=0}^{N} \langle m'| a_{m}^{*} e^{-i\frac{\pi}{2}m'} |m\rangle.$$  

(B23)

The overlap signal can then be reduced to the magnitude square of a complex Fourier series as follows

$$S_{\phi}^{\text{pure}} = \text{Tr} [\rho_{\phi} \rho_{\text{pure}}]$$

(B24)

$$= \text{Tr} \left[ \sum_{j=0}^{N} a_{j} e^{-i\phi} |j\rangle \langle j| \sum_{k=0}^{N} \langle k| a_{k}^{*} e^{i\phi} \sum_{l=0}^{N} a_{l}|l\rangle \sum_{m=0}^{N} \langle m| a_{m}^{*} e^{-i\phi} |m\rangle \right]$$

(B25)

$$= \sum_{j=0}^{N} |a_{j}|^{2} e^{-i\phi} \sum_{k=0}^{N} \langle k| a_{k}^{*} e^{i\phi} \sum_{l=0}^{N} a_{l}|l\rangle$$

(B26)

$$= \left( \sum_{j=0}^{N} |a_{j}|^{2} e^{-i\phi} \right) \left( \sum_{k=0}^{N} |a_{k}|^{2} e^{i\phi} \right)$$

(B27)

$$= \left[ \sum_{j=0}^{N} |a_{j}|^{2} e^{-i\phi} \right]^{2}.$$  

(B28)

When we substitute $a_{0}$, $a_{N} = 1/\sqrt{N}$ and $a_{i}=0$ for $i \neq 0$, $N$, the expression reduces down to the ideal case for GHZ states shown in equation (B19). By including amplitudes of non-all-zero and non-all-one states, lower frequencies are introduced into the behaviour and the $N$ frequency component is damped.

The amplitude $I_{N}$ can be derived from $S_{\phi}$ as follows

$$S_{\phi} = \text{Tr} \left( \sum_{q} e^{-i\phi \sum_{j} i_{j}^{2}} \rho_{q} \sum_{p} \rho_{p} \right)$$

(B29)

$$= \text{Tr} \left( \sum_{q} e^{-i\phi (m-N/2) - i\phi (m-N/2)} \rho_{q} \sum_{p} \rho_{p} \right)$$

(B30)

$$= \sum_{q} e^{-i\phi q} \text{Tr} \left( \rho_{q} \sum_{p} \rho_{p} \right)$$

(B31)

$$= \sum_{q} e^{-i\phi q} \text{Tr} (\rho_{q} \rho_{-q})$$

(B32)

$$= \sum_{q} e^{-i\phi q} I_{q},$$

(B33)

where equations (B4) and (B6) are used to help obtain equation (B30). Fourier transforming this result gives $I_{q} = N^{-1} \sum_{\phi} e^{i\phi S_{\phi}}$ where $N$ is the number of angles $\phi$ within the summation. To ensure that the frequency $N$ is detectable, measure $S_{\phi}$ for $\phi = \frac{n}{N+1}$ where $j = 0, 1, \ldots, 2N + 1$ which can detect up to frequency $N + 1$.

ORCID iDs

Gary J Mooney © https://orcid.org/0000-0002-3253-9815
Gregory A L White © https://orcid.org/0000-0001-6673-6676
References

[1] Vidal G 2003 Efficient classical simulation of slightly entangled quantum computations Phys. Rev. Lett. 91 147902
[2] Verstraete F and Cirac J I 2004 Renormalization algorithms for quantum-many body systems in two and higher dimensions arXiv: cond-mat/0407006
[3] Verstraete F, Murg V and Cirac J I 2008 Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems Adv. Phys. 57 143–224
[4] Cross A W, Bishop L S, Sheldon S, Nation P D and Gambetta J M 2019 Validating quantum computers using randomized model circuits Phys. Rev. A 100 032328
[5] Tölli G and Gühne O 2005 Entanglement detection in the stabilizer formalism Phys. Rev. A 72 022340
[6] Lewenstein M, Kraus B, Cirac J I and Horodecki P 2000 Optimization of entanglement witnesses Phys. Rev. A 62 052310
[7] Terhal B M 2000 Bell inequalities and the separability criterion Phys. Lett. A 271 319–26
[8] Leibfried D et al 2005 Creation of a six-atom ‘Schrödinger cat’ state Nature 438 639–42
[9] Koch J, Terri M Y, Gambetta J, Houck A A, Schuster D I, Majer J, Blais A, Devoret M H, Girvin S M and Schoelkopf R J 2007 Charge-insensitive qubit design derived from the Cooper pair box Phys. Rev. A 76 042319
[10] Jurcevic P et al 2021 Demonstration of quantum volume 64 on a superconducting quantum computing system Quantum Science and Technology 6 (2) 025020
[11] Lundeen J S, Feito A, Coldenstrodt-Ronge H, Pregnell K L, Silberhorn C, Ralph T C, Eisert J, Plenio M B and Walmsley I A 2009 Tomography of quantum detectors Nat. Phys. 5 27–30
[12] Maciejewski P B, Zimbros Z and Osmani M 2020 Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography Quantum 4 237
[13] Wei K K, Lau I, Srivivasan S, Sundaresan N, McClure D T, Toyli D, McKay D C, Gambetta J M and Sheldon S 2020 Verifying multiparticle entangled Greenberger–Horne–Zeilinger states via multiple quantum coherences Phys. Rev. A 101 032343
[14] Neeley M et al 2010 Generation of three-qubit entangled states using superconducting phase qubits Nature 467 570–3
[15] DiCarlo L, Reed M D, Sun L, Johnson R B, Chow J M, Gambetta J M, Frunzio L, Girvin S M, Devoret M H and Schoelkopf R J 2010 Preparation and measurement of three-qubit entanglement in a superconducting circuit Nature 467 574–8
[16] Barends R et al 2014 Superconducting quantum circuits at the surface code threshold for fault tolerance Nature 508 500–3
[17] Song C et al 2017 7-qubit entanglement and parallel logic operations with a superconducting circuit Phys. Rev. Lett. 119 180511
[18] Gong M et al 2019 Genuine 12-qubit entanglement on a superconducting quantum processor Phys. Rev. Lett. 122 110501
[19] Song C et al 2019 Generation of multicomponent atomic Schrödinger cat states of up to 20 qubits Science 365 574–7
[20] Sackett C A et al 2000 Experimental entanglement of four particles Nature 404 256–9
[21] Häffner H et al 2005 Scalable multipartite entanglement of trapped ions Nature 438 643–6
[22] Monz T, Schindler P, Barreiro J T, Chwalla M, Nigg D, Coish W A, Harlander M, Hänsel W, Hennrich M and Blatt R 2011 14-qubit entanglement: Creation and coherence Phys. Rev. Lett. 106 110506
[23] Pogorelov I et al 2021 Compact ion-trap quantum computing demonstrator PRX Quantum 2 020343
[24] Bouwmeester D, Pan J W, Daniell M, Weinfurter H and Zeilinger A 1999 Observation of three-photon greenberger-horne-zeilinger entanglement Phys. Rev. Lett. 82 1345
[25] Pan J-W, Daniell M, Gasparoni S, Wehls G and Zeilinger A 2001 Experimental demonstration of four-photon entanglement and high-fidelity teleportation Phys. Rev. 86 4435
[26] Zhao Z, Chen Y-A, Zhang A-N, Yang T, Briegel H J and Pan J-W 2004 Experimental demonstration of four-photon entanglement and open-destination teleportation Nature 430 54–8
[27] Lu C-Y, Zhou X-Q, Gühne O, Gao W-B, Zhang J, Yuan Z-S, Goebel A, Yang T and Pan J-W 2007 Experimental entanglement of six photons in graph states Nat. Phys. 3 91–5
[28] Yao X-C, Wang T-X, Xu P, Lu H, Pan G-S, Bao X-H, Peng C-Z, Lu C-Y, Chen Y-A and Pan J-W 2012 Observation of eight-photon entanglement Nat. Photonics 6 225–8
[29] Huang Y-F, Liu B-H, Peng L, Li Y-H, Li L, Li C-F and Gao G-C 2011 Experimental generation of an eight-photon greenberger-horne-zeilinger state Nat. Commun. 2 1–6
[30] Wang X-L et al 2016 Experimental ten-photon entanglement Phys. Rev. Lett. 117 210502
[31] Zhong H-S et al 2018 12-photon entanglement and scalable scattered boson sampling with entangled optimal-photon pairs from parametric down-conversion Phys. Rev. Lett. 121 250505
[32] Gao W-B, Lu C-Y, Yao X-C, Lu P, Gühne O, Goebel A, Chen Y-A, Peng C-Z, Chen Z-B and Pan J-W 2010 Experimental demonstration of a hyper-entangled ten-qubit Schrödinger cat state Nat. Phys. 6 351–6
[33] Wang X-L et al 2018 18-qubit entanglement with six photons' three degrees of freedom Phys. Rev. Lett. 120 260502
[34] Neumann P, Mizuochi N, Rempp F, Hemmer P, Watanabe H, Yasumasa S, Jacques V, Gaebel T, Jelezko F and Wrachtrup J 2008 Multiparticle entanglement among single spins in diamond Science 320 1326–9
[35] Rauschenbeutel A, Nogues G, Osnaghi S, Bertet P, Brunet M, Raimond J-M and Haroche S 2000 Step-by-step engineered multiparticle entanglement Science 288 2024–8
[36] Omran A et al 2019 Generation and manipulation of Schrödinger cat states in Rydberg atom arrays Science 365 570–4
[37] Takeda K, Noiri M, Nakajima T, Yoneda J, Kobayashi T and Tarucha S 2021 Quantum tomography of an entangled three-spin state in silicon Nature Nanotechnology 1 1–5
[38] Wang X-L, Pan J-W, Yao X-C and Zeng B 2018 16-qubit IBM universal quantum computer can be fully entangled Npj Quantum Information 4 1–6
[39] Friis N et al 2018 Observation of entangled states of a fully controlled 20-qubit system Phys. Rev. X 8 021012
[40] Mooney G J, Hill C D and Hollenberg L C J 2019 Entanglement in a 20-qubit superconducting quantum computer Sci. Rep. 9 1–8
[41] Pu Y, Wu Y, Jiang N, Zhang W, Li C, Zhang S and Duan L 2018 Experimental entanglement of 25 individually accessible atomic quantum interfaces Science Advances 4 eaar3931
[42] Baumann M, Munowitz M, Garroway A N and Pines A 1985 Multiple-quantum dynamics in solid state NMR J. Chem. Phys. 83 2015–25
[43] Wei K K, Ramamathan C and Cappelletto P 2018 Exploring localization in nuclear spin chains Phys. Rev. Lett. 120 070501
[44] Gärttner M, Bohnet J G, Safavi-Naini A, Wall M L, Bollinger J J and Rey A M 2017 Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet Nat. Phys. 13 781–6
[45] Hahn E L 1950 Spin echoes Phys. Rev. 80 580
[46] Blume-Kohout R, Gamble J K, Nielsen E, Rudinger K, Mizrahi J, Fortier K and Maunz P 2017 Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography Nat. Commun. 8 1–13
[47] Smolin J A, Gambetta J M and Smith G 2012 Efficient method for computing the maximum-likelihood quantum state from measurements with additive gaussian noise Phys. Rev. Lett. 108 070502
[48] Gottesman D 2016 Quantum fault tolerance in small experiments arXiv:1610.03507
[49] Meade B, Lafayette L, Sauter G and Tosello D 2017 Spartan HPC-Cloud Hybrid: Delivering Performance and Flexibility (The University of Melbourne) (https://doi.org/10.4225/49/58ead90dceaaa)
[50] White G A L, Hill C D and Hollenberg L C L 2021 Performance optimization for drift-robust fidelity improvement of two-qubit gates Phys. Rev. Applied 15 014023
[51] Greenbaum D 2015 Introduction to quantum gate set tomography arXiv:1509.02921
[52] Nielsen E, Rudinger K, Proctor T, Russo A, Young K and Blume-Kohout R 2020 Probing quantum processor performance with pyGSTi Quantum Science and Technology 5 044002
[53] Geller M R 2020 Rigorous measurement error correction Quantum Science and Technology 5 03LT01
[54] Geller M R and Sun M 2020 Towards efficient correction of multiqubit measurement errors: Pair correlation method Quantum Science and Technology 6 025009
[55] Briegel H J and Raussendorf R 2001 Persistent entanglement in arrays of interacting particles Phys. Rev. Lett. 86 910