Experimental high-dimensional Greenberger-Horne-Zeilinger entanglement with superconducting transmon qutrits

Alba Cervera-Lierta,1, 2, * Mario Krenn,1, 2, 3, † Álan Aspuru-Guzik,1, 2, 3, 4, † and Alexey Galda5, 6, §

1 Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Canada.
2 Department of Computer Science, University of Toronto, Canada.
3 Canadian Institute for Advanced Research (CIFAR) Leibnitz Fellow, Toronto, Canada.
4 Vector Institute for Artificial Intelligence, Toronto, Canada.
5 James Franck Institute, University of Chicago, Chicago, IL 60637, USA.
6 Computational Science Division, Argonne National Laboratory, Lemont, IL 60439, USA

Multiparticle entanglement is one of the core concepts in quantum information science with broad applications that span from condensed matter physics to quantum physics foundations tests. Although its most studied and tested forms encompass two-dimensional systems, current quantum platforms technically allow the manipulation of additional quantum levels. We report the first experimental demonstration of a high-dimensional multipartite entangled state in a superconducting quantum processor. We generate the three-qutrit Greenberger-Horne-Zeilinger state by designing the necessary pulses to perform high-dimensional quantum operations. We obtain the fidelity of 78 ± 1%, proving the generation of a genuine three-partite and three-dimensional entangled state. To this date, only photonic devices have been able to create and manipulate these high-dimensional states. Our work demonstrates that another platform, superconducting systems, is ready to exploit high-dimensional physics phenomena and that a programmable quantum device accessed on the cloud can be used to design and execute experiments beyond binary quantum computation.

I. INTRODUCTION

Entanglement is one of the most striking properties of quantum mechanics that has fascinated the physics community for the last 80 years. In the words of Schrödinger, he “would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought” [1]. The pioneer studies of quantum entanglement were restricted to two-particle phenomena, the well-known Einstein–Podolsky–Rosen (EPR) pairs [2]. However, entanglement can be a much richer marvel when present in a multipartite system. The keynote representative of a genuine multipartite entangled state was proposed by Greenberger-Horne-Zeilinger (GHZ) to show that local realism can be violated by quantum mechanics with deterministic experiments [3, 4]. Since then, multipartite entanglement has become a key topic in quantum information science for its rich and challenging mathematical structure [5], its implications in the study of local-realism theories [6], or the study of critical phases in condensed matter models [7, 8].

The vast majority of experimental entanglement studies have focused on two-dimensional systems. However, high-dimensional entanglement is broadly present in nature and can experimentally be generated in our labs. High-dimensional states have been extensively used in photonic systems [9–11]. Other experimental platforms such as trapped ions have simulated the dynamics of spin-1 systems [12], while ensembles of cold atoms can be used as quantum memories to store high-dimensional states [13, 14]. Nevertheless, to our knowledge, only photonic systems have been able to generate and perform quantum information tasks that require high-dimensional multipartite entanglement [10].

Quantum control beyond the two-level system has been exploited in superconducting quantum processors since the very beginning of this technology. Starting from the use of the higher levels for qubit readout [15–17] and the explicit use of the third level for spin-1 quantum simulation [18], the first steps towards the realization of ternary quantum computation with superconducting transmon devices have been taken in the last 10 years [18–26]. More recently, these efforts have led to the implementation of high-fidelity single-qutrit gates [27, 28]. In addition to paving the way to extend quantum computation to the three- and higher-level systems, these experiments enable the study of high-dimensional quantum physics.

In this work, we present the first observation of a high-dimensional GHZ state in a superconducting quantum system. Our experiment is run in a programmable device at least 30000 times faster than the preceding photonic experiment [29]. The ability to generate the qutrit GHZ state in a programmable cloud-based quantum processor offers a new pathway to investigate the unexplored world of multiparticle high-dimensional entanglement, as shown pictorially in Fig. 1. The experiment is executed in the five-transmon IBM Quantum system ibmq_rome using Qiskit Pulse [30] to program and calibrate the single-qutrit gate set in the (12) subspace. We report 78 ± 1% fidelity, exceeding the entanglement witness threshold to demonstrate genuine high-dimensional multipartite entanglement. Our results, together with the recent results from Ref.[31], are the first non-photonic experiments that explored multi-level entanglement.

* a.cervera.lierta@gmail.com
† mario.krenn@univie.ac.at
‡ alan@aspuru.com
§ agalda@uchicago.edu
qudit gates acting in the (12) subspace (see 1a):
\[ P_\alpha^{(12)}(\theta) = e^{-i\frac{\lambda_\alpha}{2}(12)} , \]
where \( \alpha = x, y, \) and \( \sigma_\alpha^{(12)} \) are the Gell-Mann matrices \( \lambda_\alpha, \gamma \) and the \( SU(3) \) generalization of the Pauli matrices \( \sigma_\alpha \) acting in the (12) subspace. Default hardware implementation of the CNOT gate was used, as defined by IBM Quantum. Extended to qudits, it acts as a \( SU(9) \) gate with the truth table shown in Tab. 1b. For the control qutrit in the (01) subspace, it acts as a standard qubit CNOT, and adds a \( \pi/2 \) phase to the \( |2\rangle \) state of the target qutrit. When the control qutrit in the \( |2\rangle \) state, it generates an equal superposition in the \( (0, 1) \) space of the target qutrit, and adds a relative phase to the \( |2\rangle \) state. The full characterization of the action of a CNOT gate in the \( SU(9) \) space is provided in Ref.[39].

IBM Quantum provides the device specifications after each calibration cycle, as well as a basic set of quantum gates that include single and two-qubit gates. The device specifications for ibmq_rome used in the experiment are shown in Tab. III. The qutrit GHZ state was experimentally achieved on transmons Q1, Q2, and Q3 connected in a linear chain.

As a qutrits readout protocol, we adopted the default 0–1 state discriminator, as implemented by IBM Quantum, to most accurately distinguish between the \( |000\rangle \) state and the rest of possible computational basis states. The main disadvantage of the default discriminator is that it is unable to correctly identify excitations to the \( |2\rangle \) state, misclassifying them as \( |1\rangle \). As a result, in order to measure other basis states, we perform local operations to lower the qutrits to the \( |000\rangle \) state. For instance, to measure the probability of obtaining the \( |010\rangle \) state, we apply \( X^{(01)} \) on the second qutrit and measure the probability of the \( |000\rangle \) state. Although this protocol increases the number of necessary experiments (one for each computational basis probability amplitude), it is aimed at reducing measurement errors by taking advantage of the high accuracy of the proprietary 0–1 discriminator offered by IBM Quantum, eliminating the need to calibrate a custom 0-1-2 discriminator, and only measuring the state with the lowest readout error [40].

In Fig. 4 from App. D, we present the measurement error mitigation matrix, commonly used to mitigate state preparation and measurement (SPAM) errors. It is obtained by preparing each computational basis state and collecting the probabilities of measuring each of these states.

The bottom panel of Fig. 2 shows the pulse sequence for the quantum circuit executed to create the three-dimensional GHZ state. Local qutrit gates were implemented by transmitting calibrated microwave pulses to manipulate qudits Q1, Q2, and Q3 through the drive channels D1, D2, and D3, correspondingly. Cross resonance (CNOT) gates between transmons Q2 and Q1, and between Q2 and Q3 were implemented by transmitting two-transmon control drives on channels u2 and u5, correspondingly, with rotary target corrections [41]. The total circuit duration is 3.16µs.
FIG. 2. Quantum circuit to generate the qutrit GHZ state, its diagrammatic representation through the GHZ − clock, and the pulse schedule. The single-qutrit gates act in the (01) and (12) subspaces, and the CNOT gates act as described in Tab. Ib. (a) First, a superposition in the (01) subspace is generated in the second qutrit using a $R_{y}^{(01)}$ gate with angle $\theta = 2 \arctan(1/\sqrt{2})$. (b) Next, two CNOT gates are applied to create two-dimensional entanglement across all three transmons. (c) The superposition state is shifted from the (01) to the (02) subspace by using $R_{x}^{(01)}$ and $R_{y}^{(01)}$ gates. (d) The third basis element is created by generating another (01) superposition, followed by repeating the application of CNOT gates controlled on the central qutrit to finish the qutrit GHZ state. At the bottom, the pulse schedule that executes the circuit. Drive channels d1, d2, and d3 are used to implement calibrated microwave pulses of qutrits Q1, Q2, and Q3 from the ibmq_rome device, respectively. The cross resonance gates utilized to implement the CNOT operations are executed on channels u2 and u5 by transmitting two-transmon control drives.

B. The GHZ circuit

The three-dimensional GHZ state in the computational basis can be written as

$$|GHZ\rangle = \frac{1}{\sqrt{3}} (|000\rangle + |111\rangle + |222\rangle).$$

This state is a maximally entangled state with Schmidt rank vector of (3, 3, 3) [42] and maximal entropy for all three bipartitions.

With the gates described in the previous subsection, it was not straightforward how to find the optimal circuit to generate the GHZ state in terms of the number of gates and circuit depth that also adapts to the chip connectivity. The superposition in the (01) subspace generated by the CNOT when the control qutrit is in the state $|2\rangle$ and the relative phases introduced by the $R_{x}^{(ij)}$ and $R_{y}^{(ij)}$ gates suggested that some extra gates may be necessary to obtain ideal operations like $X^{(ij)} = |i\rangle \langle j| + |j\rangle \langle i|$. For this reason, we relied on computers to find a suitable quantum circuit. The search for an optimal circuit could follow a standard procedure of creating a parameterized ansatz circuit that optimizes the real-valued angles of the individual gates via gradient descent or other classical optimization methods. This approach, however, may have two important disadvantages. On one side, it is not necessarily guaranteed that the obtained solution is indeed the minimal circuit. On the other side, the circuit obtained could contain an intricate parameter setting that is challenging to understand. As scientists, we aim to understand the solution so we can extract new general ideas and inspiration from it. For that reason, we followed a different approach. We restricted the angles of the rotation gates to $\theta = \pm \pi$ and $\theta = \pm \pi/2$ (except for only up to one free angle for the $R_{y}^{(01)}$ gate), which corresponding operations are easy to comprehend for humans. Then, we used the algorithm MELVIN [43, 44], adapted to operate with discrete digital qutrit gates, to find this circuit, where
The real and imaginary parts of these terms can be amplitudes of the computational basis states. The first sum corresponds to measuring the probability of the computational basis states.

Table I. Ternary quantum gates used to generate the GHZ state. We use the default pulse calibration for the qubit gates $R_y^{(q)}$ and CNOT and design the pulse sequence for the qutrit gates $R_y^{(y)}$ and $R_z^{(y)}$. For the CNOT gate, one needs to take into account its effect when the control is in the $|2\rangle$ state. This effect is reproduced in the CNOT truth table shown (more details can be found in Ref. [39]).

<table>
<thead>
<tr>
<th>Gate</th>
<th>Matrix</th>
<th>Control Target</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_y^{(q)}(\theta)$</td>
<td>$\begin{pmatrix} \cos(\theta/2) &amp; -\sin(\theta/2) &amp; 0 \ \sin(\theta/2) &amp; \cos(\theta/2) &amp; 0 \ 0 &amp; 0 &amp; 1 \end{pmatrix}$</td>
<td>$</td>
<td>0\rangle$</td>
</tr>
<tr>
<td>$R_y^{(y)}(\theta)$</td>
<td>$\begin{pmatrix} 1 &amp; 0 &amp; 0 \ 0 &amp; \cos(\theta/2) &amp; -\sin(\theta/2) \ 0 &amp; \sin(\theta/2) &amp; \cos(\theta/2) \end{pmatrix}$</td>
<td>$</td>
<td>1\rangle$</td>
</tr>
<tr>
<td>$R_z^{(y)}(\theta)$</td>
<td>$\begin{pmatrix} 1 &amp; 0 &amp; 0 \ 0 &amp; \cos(\theta/2) &amp; -i\sin(\theta/2) \ 0 &amp; -i\sin(\theta/2) &amp; \cos(\theta/2) \end{pmatrix}$</td>
<td>$</td>
<td>2\rangle$</td>
</tr>
</tbody>
</table>

(a) Single-qutrit gates.

(b) Truth table IBM default CNOT gate.

Table I. Ternary quantum gates used to generate the GHZ state. We use the default pulse calibration for the qubit gates $R_y^{(0)}$ and CNOT and design the pulse sequence for the qutrit gates $R_y^{(1)}$ and $R_z^{(1)}$. For the CNOT gate, one needs to take into account its effect when the control is in the $|2\rangle$ state. This effect is reproduced in the CNOT truth table shown (more details can be found in Ref. [39]).

we also include the connectivity constraints imposed by the chip design. Indeed, we quickly discovered a feasible solution that is optimal in terms of control operations and single-qutrit gates, and which we can interpret and explain straightforwardly.

The circuit found to generate the GHZ state is shown in Fig. 2. It requires four CNOT gates, two single-qubit gates and three single-qutrit gates. Notice that the CNOT gates employed, although for qutrits they act as shown in Tab. Ib, they behave as standard CNOT gates. The first part of the circuit only acts in the (01) subspace while in the second part both control and target qutrits are in the $|2\rangle$ state and, thus, they do not generate any superposition. We can only expect a relative phase $2\varphi$ to appear in the $|222\rangle$ state, which does not change the entanglement properties of the GHZ state generated. To illustrate how the quantum circuit produces the GHZ state, we introduce the GHZ-clock representation, briefly described in the App. A.

C. Fidelity and entanglement witness protocol

To obtain the fidelity of the generated state with respect to the theoretical GHZ state from Eq. (2), we follow the steps presented in Ref. [29]. Notice that we do not require to perform a full tomography protocol to obtain the fidelity. Denoting $\rho$ as the density matrix from the state generated in the chip, the fidelity with respect to the GHZ state can be computed as $\text{Tr}(\rho |GHZ\rangle \langle GHZ|)$, i.e.

$$F_{\text{exp}} = \frac{1}{3} \left( \sum_{i=0}^{2} \langle iii|\rho|iii\rangle + \sum_{i,j=0}^{2} \text{Re}(\langle ijj|\rho|jjj\rangle) \right).$$

The first sum corresponds to measure the probability amplitudes of the computational basis states $|000\rangle$, $|111\rangle$ and $|222\rangle$. Due to the Hermiticity of the density matrix, the second sum contains only three independent terms, $|000\rangle|111\rangle$, $|000\rangle|222\rangle$ and $|111\rangle|222\rangle$. The real and imaginary parts of these terms can be measured by computing the expectation value of certain combinations of $\sigma_x$ and $\sigma_y$ operators [29] (see App. B for details), although we will only need the real part to obtain the fidelity.

The final step is to certify the generation of three-partite entanglement. To do so, we follow the entanglement witness protocol described in Ref. [45]. This method establishes that the maximum possible fidelity between a GHZ state and an immediately lower-dimensional entangled state (i.e. a state with Schmidt rank vector of $(3, 3, 2)$, $(2, 3, 3)$ or $(3, 2, 3)$) is $2/3$. Thus, the condition $F_{\text{exp}} > 2/3$ certifies the generation of a three-dimensional three-partite entangled state. More details about this protocol are presented in App. B.

III. RESULTS AND DISCUSSION

We performed a total of 195 experiments to reconstruct the density matrix elements necessary to compute the fidelity with respect to the GHZ state (Eq. (3)): 3 to obtain the diagonal elements and 8 different $X$ and $Y$ projective measurements in the three possible subspaces, each requiring 8 circuits to obtain the result by measuring only the $|000\rangle$ state. Although we do not require the imaginary part of these elements to obtain the fidelity, we measure it to obtain a better estimation of the generated state. The number of shots for each experiment was $n = 512$.

The reconstruction of the density matrix is depicted in Fig. 3. It does not include all off-diagonal terms needed to reconstruct entirely the quantum state, but only those necessary to calculate the fidelity with respect to the GHZ state. The real and imaginary parts of each of these elements are shown in Eq. (C3) from App. C. We obtain a fidelity of

$$F_{\text{exp}} = 0.78 \pm 0.01,$$

after applying the readout error mitigation, and $F_{\text{raw}} = 0.69 \pm 0.02$ with the raw measured data, in both cases, exceeding the bound of 2/3 required to certify three-dimensional three-partite entanglement.
FIG. 3. Experimental results of the density matrix elements needed to compute the fidelity with respect to the GHZ state. Only six elements are required to reconstruct the fidelity: the three diagonal terms \(\langle 000|\rho|000\rangle\), \(\langle 111|\rho|111\rangle\) and \(\langle 222|\rho|222\rangle\), and the off-diagonal terms \(\langle 000|\rho|111\rangle\), \(\langle 000|\rho|222\rangle\) and \(\langle 111|\rho|222\rangle\). The plot shows the absolute value of these elements (in grey, all other elements not measured). The corresponding fidelity obtained is 78 ± 1%, exceeding the 2/3 threshold required to certify genuine three-partite entanglement. The real and imaginary parts of these elements are presented in App. C.

The non-zero imaginary parts of some of these elements suggest possible relative phases between the three GHZ computational basis states. Notice that the state \((|000\rangle + e^{i\phi_1}|111\rangle + e^{i\phi_2}|222\rangle)/\sqrt{3}\) with \(-\pi \leq \phi_1, \phi_2 \leq \pi\), has the same entanglement properties as the GHZ from Eq. (2). By comparing the theoretical density matrix elements with the ones measured in our experiment, we can extract the values of these phases. In particular, we obtain \(\phi_1 = -0.04 \pm 0.03\) and \(\phi_2 = 0.25 \pm 0.03\). The details of this analysis are provided in App. C.

Multiple sources of error can affect the fidelity of the GHZ state. Since we are dealing with higher-dimensional states, we have two relaxation processes due to spontaneous emission: the \(|1\rangle \rightarrow |0\rangle\) and the \(|2\rangle \rightarrow |1\rangle\) transitions. This effect is observed in a count dropping for the diagonal states \(|222\rangle\) and \(|111\rangle\) in comparison with the \(|000\rangle\) state, and similarly in the off-diagonal terms. We also expect that errors coming from transmon cross-talk, as it is also common in qubit devices, can be aggravated due to the inclusion of \(|1\rangle \rightarrow |2\rangle\) transition frequencies. Finally, the phase \(\varphi\) introduced by the CNOT gate in the \((12)\) subspace is translated to a relative phase in the GHZ \((222)\) basis state. Evidence of this phase is observed when computing relative phases of the state generated. The \(\phi_2\) phase is different from the \(\phi_1\) phase as it potentially includes a non-zero contribution of this \(\varphi\) phase.

We note that it is feasible to increase the GHZ fidelity of this experiment by obtaining a full characterization of the CNOT gate that will allow for introducing a phase shift that cancels the effect of the \(\varphi\) phase. Active error mitigation techniques that will allow compensating the dephasing errors can also be implemented, as shown in Ref. [31], where dynamical decoupling protocols were used. The statistical errors can also be reduced by increasing the number of shots of each experiment.

IV. CONCLUSIONS AND OUTLOOK

We report the first experimental three-qutrit maximal entangled state in a superconducting device, the GHZ state, with a fidelity of 78 ± 1%. To achieve that, we calibrate additional qutrit gates using the pulse-level programming model Qiskit Pulse via cloud access to the IBM Quantum device ibmq_rome, find an optimal circuit in terms of the number of entangling gates, and experimentally reconstruct the density matrix elements needed to obtain the fidelity with respect to the theoretical GHZ state.

We demonstrate that superconducting quantum devices are ready to escape the flatland represented by
binary quantum systems. From a quantum technological point of view, the capacity to store more quantum information in higher dimensions has direct implications to improve quantum error-correcting codes [46–49]. On top of that, the multi-level quantum computation can also introduce an advantage in comparison with the current binary computation [50, 51]. Besides applications, high-dimensional multipartite systems enable the study of the foundations of quantum physics. The non-trivial structure of the high-dimensional Hilbert space is illustrated by the fact that it took nearly 20 years to theoretically generalize the GHZ theorem beyond qubits [52–55]. The result required the introduction of non-Hermitian operations, also present in other tests of local realism like Bell Inequalities [56, 57].

With this work, we show that, at the present day, anyone with internet access can experimentally study high-dimensional multipartite entanglement that heretofore was limited to a handful of laboratories worldwide. We anticipate an interest increase in this field with its most exciting applications yet to be imagined.

DATA AVAILABILITY

Experimental data and its analysis can be found at https://github.com/alexgalda/qutrit-GHZ.

ACKNOWLEDGEMENTS

We acknowledge the use of IBM Quantum services for this work. We acknowledge the access to advanced services provided by the IBM Quantum Researchers Program, and also the IBM Quantum Researcher Access Award Program. M.K. and A.C.L appreciate Manuel Erhard’s insights on high-dimensional entanglement. A.A.-G. also acknowledges support from the Canada Industrial Research Chairs Program and the Canada 150 Research Chairs Program. M.K. acknowledges support from the FWF (Austrian Science Fund) via the Erwin Schrödinger fellowship No. J4309.

[38] Xiao-Min Hu, Wen-Bo Xing, Chao Zhang, Bi-Heng Liu,Matej Pivolouska, Marcus Huber, Yun-Feng Huang, Chuan-Feng Li, and Guang-Can Guo, “Experimental creation of multi-photon high-dimensional layered quantum states,” NPJ Quantum Inf. 6, 1–5 (2020).
Appendix A: GHZ-clock

We introduce a new graphical representation to illustrate the logic of how the GHZ circuit works. The GHZ-clock is a circle divided into three sections, each corresponding to one of the two-dimensional subspaces: (01), (12), and (02). We draw an arrow pointing at each basis state with a length proportional to the corresponding probability amplitude. Therefore, this representation can be used to depict qutrit pure states that do not contain the three levels in one of the basis states. One may consider introducing a color or thickness dimension to the arrows to represent relative phases, although this will not be necessary for the GHZ circuit under discussion.

As shown in Fig. 2, we divide each subsection into eight segments representing the computational basis states of each subspace. The circuit starts with a single arrow of unit length pointing at 12 o’clock, i.e. the |000⟩ state. The first operation generates the superposition \( \sqrt{2/3}|0⟩ + \sqrt{1/3}|1⟩ \) on the second qutrit, producing the state \( \sqrt{2/3}|000⟩ + \sqrt{1/3}|010⟩ \), which is represented with a 2/3 arrow pointing at 12 o’clock and a 1/3 arrow pointing at the |010⟩ state (approximately, 1 o’clock). The next two CNOT gates move this second arrow to the |111⟩ state located in the fourth quadrant of the circle. The three single-qutrit operations that come after moving the state from the (01) to the (02) subspace, i.e. the |111⟩ arrow moves to the third quadrant, pointing at the |222⟩ state. The last part of the circuit repeats the process of generating a |000⟩ and |111⟩ superposition while keeping the superposition with the |222⟩. This is performed by generating an equal superposition on the second qutrit, the |(|0⟩ + |1⟩)/√2⟩ state, thus dividing into two the previous |000⟩ arrow. This produces a third arrow and, therefore, equals the length of the arrows to 1/3. After applying again the two CNOT gates, this third arrow points at the |111⟩ state, completing the GHZ state generation.

As stated above, this representation can be used to depict any pure state that does not contain any basis element with the three levels. It can be used for more than three particles by dividing each section into more segments. An extension that includes the remaining basis states or higher dimensions can be proposed by extending this representation to the 3D space. Similarly, one can propose a clever basis state ordering in each section that allows the derivation of simple rules that represent the application of qutrit quantum gates. These extensions are out of the scope of this work and we leave them for a future project.

Appendix B: Measurement of density matrix elements and entanglement witness

To obtain the elements of the density matrix needed to compute the fidelity with respect to the GHZ state, we will follow the protocol introduced in Ref. [29]. Given a three-particle density matrix of the form

\[
\rho = \sum_{i,j,k=0}^{2} a_{ijk} a_{imn}^* \langle ijk | lmn \rangle, \tag{B1}
\]

it can be shown that the real and imaginary part of an element of \( \rho \) can be expressed as a combination of expectation values of \( \sigma_x^{(ab)} \) and \( \sigma_y^{(ab)} \) in the subspaces.
\((ab) = (01), (02), (12)\) as follows:

\[
\text{Re}\left((ijk)|\rho|lmm\right) = \frac{1}{8} \left(\sigma_x^{(i)} \sigma_y^{(j)} \sigma_z^{(k)} - \sigma_y^{(i)} \sigma_x^{(j)} \sigma_z^{(k)} - \sigma_z^{(i)} \sigma_y^{(j)} \sigma_x^{(k)} \right),
\]

\[
\text{Im}\left((ijk)|\rho|lmm\right) = \frac{1}{8} \left(\sigma_x^{(i)} \sigma_y^{(j)} \sigma_z^{(k)} + \sigma_y^{(i)} \sigma_x^{(j)} \sigma_z^{(k)} + \sigma_z^{(i)} \sigma_y^{(j)} \sigma_x^{(k)} \right),
\]

To compute these expectation values, we will project into the \(X\) and \(Y\) basis and measure in the computational basis using the gates \(H_x\) and \(H_y\) respectively. These gates can be decomposed into the native gate set as

\[
H^{(01)} = \bar{R}_x^{(01)} (-\pi) R_y^{(01)} (\pi/2),
\]

\[
H^{(02)} = \bar{R}_x^{(01)} (\pi) \bar{R}_y^{(12)} (\pi) R_x^{(12)} (-\pi/2) R_y^{(01)} (-\pi),
\]

\[
H^{(12)} = R_y^{(12)} (-\pi/2) R_x^{(12)} (-\pi),
\]

\[
H^{(01)} = \bar{R}_y^{(02)} (-\pi/2) \bar{R}_y^{(01)} (\pi),
\]

\[
H^{(02)} = \bar{R}_x^{(12)} (\pi) \bar{R}_y^{(01)} (-\pi/2) \bar{R}_y^{(01)} (\pi) R_y^{(12)} (-\pi),
\]

\[
H^{(12)} = R_y^{(12)} (-\pi/2) \bar{R}_y^{(12)} (\pi).
\]

With the density matrix elements, we can compute the fidelity with respect to the GHZ state \(F_{\text{exp}}\) following Eq. (3). With that fidelity, we can apply the entanglement witness protocol described in [45]. We have to show that the state generated has a multipartite entanglement structure of the form \((3,3,3)\), where \((x,y,z)\) corresponds to the Schmidt rank vector, i.e. the rank of each reduced density matrix for the three possible bipartitions. To prove that, we have to demonstrate that this state cannot be decomposed into states of lower dimensionality structure. The protocol consists of comparing the fidelity of this state with a state with entanglement structure \((3,3,3)\), i.e. a three-dimensional and three-partite state.

with Schmidt rank \(\xi \leq \chi\) for the same bipartition is [58]

\[
F_{\text{max}} = \max_{\sigma, \rho} \text{Tr}(\sigma |\psi\rangle\langle\psi|) = \sum_{i=1}^{\lambda_2} \lambda_i^2.
\]

The GHZ Schmidt coefficients for any bipartition are \((1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\). Thus, the maximal fidelity of state with Schmidt vector \((3,3,2)\) and a \((3,3,3)\) state like the GHZ is

\[
F_{\text{max}} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.
\]

Then, any fidelity \(F_{\text{exp}} > 2/3\) can only be generated by a state with entanglement structure \((3,3,3)\), i.e. a three-dimensional and three-partite state.

**Appendix C: Data and error analysis**

To reconstruct each density matrix element, we have to compute the expectation value of eight Pauli strings, see Eqs. (B2–B3). To measure each expectation value, we project into the corresponding Pauli basis and measure on the computational basis. As described in the main text, our measurement protocol discriminates between the state \(|000\rangle\) and the rest of the possible states. Thus, we will measure the amplitudes of the computational basis elements by local-transforming each of them to the \(|000\rangle\) state. For instance, to obtain the \(|011\rangle\) amplitude we will apply a \(X^{(01)}\) gate on the second and third qutrit, while for obtaining the amplitude of the \(|020\rangle\) state we will apply the \(X^{(01)}X^{(12)}\) transformation on the second qutrit. Since we are only interested in density matrix elements that involve up to two levels, each expectation value will only require a two-dimensional measurement, either between the \((01), (12)\) or \((02)\) levels, so we will have to compute the eight possible computational basis states for each level pair. This measurement protocol implies that we need to perform eight experiments to obtain the expectation value of each Pauli string.

Table II show the expectation values needed to reconstruct the density matrix elements \((000|\rho|111), (000|\rho|222)\) and \((111|\rho|222)\). The total number of shots for each experiment is \(n = 512\). We assume that each experiment is independent and it takes approximately the same amount of time since the only difference between them is the application of up to two single-qutrit gates. We assume the counts obtained follow a Binomial distribution with \(\mu = np\), \(\text{Var} = np(1-p)\) with \(p = \text{ncounts}/n\). Since \(n \gg 1\), this distribution can be approximated as a Normal distribution.

To obtain each expectation value, we assign the eigenvalue +1 or -1 (eigenvalues of \(\sigma_{x,y}\)) to each output. Then, we add them following Eq. (B2) and Eq. (B3) to reconstruct the density matrix element. We will compute and propagate the errors according to

\[
\sigma^{(ij)}_{abc} = \sum_{k=1}^{n} \text{Var}(\langle ab\rangle_k^{(ij)}),
\]
where $k$ represents the eight computational basis elements, $abc$ corresponds to the Pauli string (e.g. $xxy \equiv \sigma_x \sigma_y \sigma_y$) and $(ij)$ the two levels. Thus, the standard deviation of the real part of a density matrix element $(ij|\rho|jjj)$ will be

$$\sigma^{(ij)}_{\text{Re}} = \frac{1}{8} \sqrt{\left(\sigma_{vxx}^{(ij)}\right)^2 + \left(\sigma_{vyx}^{(ij)}\right)^2 + \left(\sigma_{vyy}^{(ij)}\right)^2 + \left(\sigma_{vzy}^{(ij)}\right)^2},$$

(C2)

and similarly for the imaginary part. For the diagonal matrix elements, the standard deviation is just the square root of the variance for each number of diagonal element counts.

The corresponding density matrix elements obtained are

$$\tilde{\rho} = \begin{pmatrix} 0.35 \pm 0.02 & (0.327 \pm 0.008) + i(0.014 \pm 0.008) & (0.231 \pm 0.008) + i(0.059 \pm 0.008) \\ (0.327 \pm 0.008) - i(0.014 \pm 0.008) & 0.29 \pm 0.02 & (0.186 \pm 0.007) - i(0.070 \pm 0.008) \\ (0.231 \pm 0.008) - i(0.059 \pm 0.008) & (0.186 \pm 0.007) + i(0.070 \pm 0.008) & 0.21 \pm 0.02 \end{pmatrix},$$

(C3)

where we used $\tilde{\rho}$ to distinguish it from the total density matrix $\rho$.

$$\sigma_F = \frac{1}{3} \sum_{i=0}^{2} (\sigma^{(ii)})^2 + \frac{1}{8} \left(4 \left(\sigma^{(01)}_{\text{Re}}\right)^2 + 4 \left(\sigma^{(02)}_{\text{Re}}\right)^2 + 4 \left(\sigma^{(12)}_{\text{Re}}\right)^2\right).$$

(C4)

Notice that due to the hermiticity of the density matrix and because the original GHZ does not contain complex coefficients, we can compute the fidelity only with the real part of the density matrix elements by using the identity $z + z^* = 2 \text{Re}(z)$.

Although the original GHZ state [3] does not contain relative phases between the computational basis states, any state of the form

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{3}} \left( |000\rangle + e^{i\phi_1} |111\rangle + e^{i\phi_2} |222\rangle \right),$$

(C5)

with $-\pi \leq \phi_1, \phi_2 \leq \pi$ share the same entanglement properties as the original GHZ state ($\phi_1 = \phi_2 = 0$), i.e. Schmidt vector of $(3,3,3)$ [42] and maximal entropy for all bipartitions. We can obtain the values of these relative phases by comparing the experimental density matrix with the theoretical one. The partial GHZ density matrix with phases is

$$\tilde{\rho}_{\text{GHZ}} (\phi_1, \phi_2) = \frac{1}{3} \begin{pmatrix} 1 & e^{-i\phi_1} & e^{-i\phi_2} \\ e^{i\phi_1} & 1 & e^{i(\phi_1 - \phi_2)} \\ e^{-i\phi_2} & e^{i(\phi_1 - \phi_2)} & 1 \end{pmatrix},$$

(C6)

where we used again $\tilde{\rho}$ to distinguish it from the total density matrix $\rho$. From this matrix we compute $\text{Arg}(000|\rho|111) = -\phi_1$ and $\text{Arg}(000|\rho|222) = -\phi_2$. Similarly, we can obtain the difference between the phases using the third off-diagonal term $\text{Arg}(111|\rho|222) = \phi_1 - \phi_2$ and compare it with the difference between the individually obtained phases.

The experimental $\tilde{\rho}$ matrix written in terms of the absolute and argument values of each measured density matrix element is

<table>
<thead>
<tr>
<th>Pauli string</th>
<th>(01) subspace</th>
<th>(12) subspace</th>
<th>(02) subspace</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{\rho_1 \rho_1}$</td>
<td>$0.69 \pm 0.03$</td>
<td>$0.41 \pm 0.03$</td>
<td>$0.46 \pm 0.03$</td>
</tr>
<tr>
<td>$\sigma_{\rho_1 \rho_2}$</td>
<td>$-0.68 \pm 0.03$</td>
<td>$-0.35 \pm 0.03$</td>
<td>$-0.47 \pm 0.03$</td>
</tr>
<tr>
<td>$\sigma_{\rho_2 \rho_2}$</td>
<td>$-0.58 \pm 0.03$</td>
<td>$-0.36 \pm 0.03$</td>
<td>$-0.46 \pm 0.03$</td>
</tr>
<tr>
<td>$\sigma_{\rho_2 \rho_3}$</td>
<td>$-0.67 \pm 0.03$</td>
<td>$-0.37 \pm 0.03$</td>
<td>$-0.46 \pm 0.03$</td>
</tr>
</tbody>
</table>

TABLE II. Expectation values of the Pauli strings necessary to reconstruct the real and imaginary parts of the density matrix elements following the tomography protocol presented in Eq. (B2) and (B3). The total number of shots per experiment is $n = 512$. To compute each expectation value, one needs to perform a total of 8 experiments in order to measure the probability amplitudes of each computational basis state following the measurement protocol presented in the main text. We apply the measurement error mitigation matrix from Fig. 4 as described in the main text.
The calibration matrix is given by

\[
\tilde{\rho} = \begin{pmatrix}
0.35 \pm 0.02 & (0.327 \pm 0.008)e^{-i(0.04 \pm 0.03)} & (0.29 \pm 0.02) & (0.239 \pm 0.008)e^{-i(0.25 \pm 0.03)} \\
(0.327 \pm 0.008)e^{-i(0.04 \pm 0.03)} & 0.35 \pm 0.02 & (0.199 \pm 0.007)e^{-i(0.36 \pm 0.04)} & (0.21 \pm 0.02) \\
(0.239 \pm 0.008)e^{-i(0.25 \pm 0.03)} & (0.199 \pm 0.007)e^{-i(0.36 \pm 0.04)} & 0.21 \pm 0.02 & (0.199 \pm 0.007)e^{-i(0.36 \pm 0.04)} \\
(0.21 \pm 0.02) & (0.199 \pm 0.007)e^{-i(0.36 \pm 0.04)} & (0.199 \pm 0.007)e^{-i(0.36 \pm 0.04)} & 0.21 \pm 0.02
\end{pmatrix}.
\]

FIG. 4. Calibration matrix for \(|0\rangle/|1\rangle\) readout on the \([Q3, Q2, Q1]\) subset of qutrits of the ibmq_rome chip.

To obtain the corresponding errors we apply the standard error propagation formula

\[
\sigma_f(x,y) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2},
\]

where \(x\) and \(y\) are the real and imaginary parts of each element, \(\sigma_x\) and \(\sigma_y\) their standard deviations as computed in Eq. C2, and \(f(x,y) = \sqrt{x^2 + y^2}\) for the absolute value, and \(f(x,y) = 2 \arctan(y/(\sqrt{x^2 + y^2} + x))\) for the argument.

From (C7), we obtain the relative phases

\[
\phi_1 = -0.04 \pm 0.03, \quad \phi_2 = 0.25 \pm 0.03,
\]

and \((\phi_1 - \phi_2) = -0.36 \pm 0.04\), estimated from the \(\langle 111|\rho|222\rangle\) term, and \((\phi_1 - \phi_2) = -0.29 \pm 0.04\), estimated from the above values of the phases.

Appendix D: Chip characterization and readout calibration matrix

In this paper we used ibmq_rome, which is one of the IBM Quantum Falcon Processors. Pulse-level control was implemented using Qiskit Pulse [30] to generate the pulses for the single qutrit gates \(R_x^{(12)}\) and \(R_y^{(12)}\) at the \(f_{12}\) transition frequency. The device specifications were provided by IBM Quantum [59] and are shown in Table III. We used qutrits Q1-Q2-Q3, also denoted as \([1, 2, 3]\) subset.

Measurement error mitigation was performed by correcting the average shot counts collected in the experiments by using the calibration matrix from Fig. 4 obtained using the Qiskit Ignis framework. The calibration matrix was generated by preparing 8 basis input states and computing the probabilities of measuring counts in all other basis states. By construction of the experiment, only the probability of the \(|000\rangle\) state was necessary to derive the fidelity of the GHZ state. As a result, only a single element of the calibration matrix, the readout fidelity of the \(|000\rangle\) state, equal to 0.96, was used to mitigate state preparation and measurement errors.
| Qutrit $|0\rangle \leftrightarrow |1\rangle$ frequency, $\omega_{01}/2\pi$ (GHz) | $Q_0$ | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ |
|---|---|---|---|---|
| | 4.969 | 4.770 | 5.015 | 5.259 | 4.998 |
| Qutrit $|1\rangle \leftrightarrow |2\rangle$ frequency, $\omega_{12}/2\pi$ (GHz) | | 4.631 | 4.443 | 4.677 | 4.926 | 4.658 |
| Lifetime $T_{1(0)}$ ($\mu$s) | 90 | 102 | 44 | 70 | 82 |
| Echo time $T_{2\text{echo}}$, $|1\rangle/|0\rangle$ ($\mu$s) | 71 | 83 | 86 | 159 | 138 |
| Readout error | 2.8e-2 | 3.0e-2 | 3.3e-2 | 3.2e-2 | 3.8e-2 |
| Prob. Prep. $|0\rangle$ Meas. $|1\rangle$ | 2.3e-2 | 2.3e-2 | 1.5e-2 | 1.3e-2 | 3.2e-2 |
| Prob. Prep. $|1\rangle$ Meas. $|0\rangle$ | 3.4e-2 | 3.8e-2 | 5.1e-2 | 5.1e-2 | 4.4e-2 |
| $X$ gate error | 2.5e-4 | 2.2e-4 | 4.3e-4 | 3.5e-4 | 3.5e-4 |
| $u_2$ gate duration (ns) | 36 | 36 | 36 | 36 | 36 |

<table>
<thead>
<tr>
<th>[0, 1]</th>
<th>[1, 0]</th>
<th>[1, 2]</th>
<th>[2, 1]</th>
<th>[2, 3]</th>
<th>[3, 2]</th>
<th>[3, 4]</th>
<th>[4, 3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNOT gate error</td>
<td>7.4e-3</td>
<td>7.4e-3</td>
<td>1.8e-2</td>
<td>1.8e-2</td>
<td>9.7e-3</td>
<td>9.7e-3</td>
<td>9.5e-3</td>
</tr>
<tr>
<td>CNOT gate duration (ns)</td>
<td>320</td>
<td>356</td>
<td>1109</td>
<td>1145</td>
<td>377</td>
<td>341</td>
<td>476</td>
</tr>
</tbody>
</table>

TABLE III. Calibration data for ibmq_rome.