

Cross-Verification of Independent Quantum Devices

C. Greganti^{1,2,*} T. F. Demarie,^{3–5,*} M. Ringbauer^{6,†} J. A. Jones⁷ V. Saggio,¹ I. Alonso Calafell¹
 L. A. Rozema,¹ A. Erhard,⁶ M. Meth,⁶ L. Postler,⁶ R. Stricker,⁶ P. Schindler,⁶ R. Blatt,^{6,8} T. Monz,^{6,9}
 P. Walther^{1,10} and J. F. Fitzsimons^{4,5,11,‡}

¹*Faculty of Physics, Vienna Center for Quantum Science and Technology (VCQ), University of Vienna, Boltzmannngasse 5, 1090 Vienna, Austria*

²*VitreaLab GmbH, Gutheil-Schoder-Gasse 17, 1230 Vienna, Austria*

³*Entropica Labs, 186b Telok Ayer Street, Singapore 068632*

⁴*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*

⁵*Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372*

⁶*Institut für Experimentalphysik, Universität Innsbruck, Technikerstrasse 25, 6020 Innsbruck, Austria*

⁷*Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*

⁸*Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, Otto-Hittmair-Platz 1, 6020 Innsbruck, Austria*

⁹*Alpine Quantum Technologies GmbH, 6020 Innsbruck, Austria*

¹⁰*Christian Doppler Laboratory for Photonic Quantum Computing, Faculty of Physics, University of Vienna, 1090 Vienna, Austria*

¹¹*Horizon Quantum Computing, 79 Ayer Rajah Crescent, #03-01 BASH, Singapore 139955*



(Received 17 November 2020; revised 22 April 2021; accepted 6 July 2021; published 2 September 2021)

Quantum computers are on the brink of surpassing the capabilities of even the most powerful classical computers, which naturally raises the question of how one can trust the results of a quantum computer when they cannot be compared to classical simulation. Here, we present a cross-verification technique that exploits the principles of measurement-based quantum computation to link quantum circuits of different input size, depth, and structure. Our technique enables consistency checks of quantum computations between independent devices, as well as within a single device. We showcase our protocol by applying it to five state-of-the-art quantum processors, based on four distinct physical architectures: nuclear magnetic resonance, superconducting circuits, trapped ions, and photonics, with up to six qubits and up to 200 distinct circuits.

DOI: [10.1103/PhysRevX.11.031049](https://doi.org/10.1103/PhysRevX.11.031049)

Subject Areas: Quantum Physics, Quantum Information

I. INTRODUCTION

Quantum computers represent a fundamental shift in the way we think about computation. By harnessing quantum interference effects between different possible branches of a computation, quantum processors have the potential to drastically outperform conventional computers for a range of tasks [1–6]. Potential applications of quantum computation range from cryptanalysis to the simulation of physical systems, and even to machine learning.

Extraordinary experimental efforts in recent years have enabled demonstrations of the technology's potential in a growing number of physical systems [7–10]. For certain simulation [11,12] and sampling [13] tasks, these devices are already pushing the limits of classical supercomputers, and it is foreseeable that the next generation of quantum processors will vastly outperform their classical counterparts.

Building such devices, however, remains challenging, with environmental interactions inducing noise that leads to potentially unreliable results for complex computations. This naturally raises the question of how we can trust the output of a quantum computation, and, more concretely, to what extent we can verify that the output is likely correct. The current standard approach is to benchmark the individual quantum gates that make up the computation [14,15] to obtain an indication for how well the full system could perform. In practice, however, such an extrapolation is typically unreliable because of effects such as

*The author contribute equally to this work.

†martin.ringbauer@uibk.ac.at

‡joe@horizonquantum.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

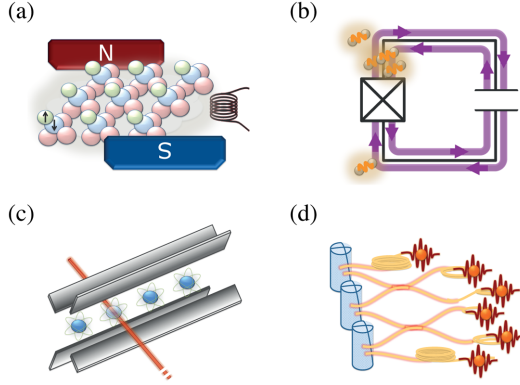


FIG. 1. Cartoon representation of the quantum processing architectures used in this work. (a) NMR device at the University of Oxford [7], (b) superconducting circuits at IBM [31] and Rigetti Computing [32], (c) a trapped-ion quantum processor at the University of Innsbruck [8], and (d) a photonic quantum processor at the University of Vienna [33].

non-Markovian behavior, spatially and temporally correlated noise, or unmodeled stray interactions [16], which highlights the need for a complementary technique to establish full system performance.

In order to fill this gap, significant work has been devoted to the development of cryptographically secure verification protocols [17–24], with one such technique having been experimentally demonstrated [25]. However, existing, provably secure, verification protocols require either quantum communication [17–20] or shared entanglement between devices [21–24], leaving them out of reach for existing quantum processors, which are typically unable to exchange quantum information with one another. One promising way of getting around this limitation and allowing for verification of a single, isolated, quantum processor is to rely on computational hardness assumptions [26]. However, because of large key sizes, implementing this approach would require an extremely sophisticated processor.

Here, we follow a different path by asking whether one noisy device can be used to efficiently check another noisy device, without relying on quantum communication or entanglement between the devices. We introduce a cross-check technique that is inherently agnostic to the underlying hardware, sensitive to systematic errors in the implementation, and applicable to any digital quantum computation. In contrast to related work on comparing output states from different quantum devices [27], our technique aims to verify the device rather than a certain output of it. The protocol is built on the framework of measurement-based quantum computation (MBQC) [28,29], which has proven to be a powerful tool for blind and verifiable computing protocols [30]. By exploiting the intrinsic symmetries of quantum circuits when mapped to an MBQC, our technique allows us to quantitatively relate the outputs of quantum circuits with different size and

structure, performed on independent physical devices in any architecture. While our approach does not aim to solve the full problem of verification that is secure against any dishonest behavior of the devices, it enables us to build a high level of trust in the outputs of honest quantum devices by observing the MBQC relationships upheld when sampling from different realizations of seemingly unrelated circuits.

We demonstrate this technique by running up to 200 distinct circuits of different width and depth on five state-of-the-art quantum processors, using four primary technologies for digital quantum computation: (1) a nuclear magnetic resonance (NMR) device [7] at the University of Oxford, (2) cloud-accessible superconducting systems from IBM [31] and Rigetti [32], (3) a trapped-ion quantum processor [8] at the University of Innsbruck, and (4) a photonic cluster-state quantum device [33] at the University of Vienna (see Fig. 1).

II. FROM MEASUREMENT-BASED QUANTUM COMPUTING TO CIRCUIT CROSS-CHECKS

In order to cross-check quantum processors, we make use of independent runs of several different yet related sampling problems, derived from a measurement-based implementation of a target computation. In contrast to the standard circuit model of quantum computing, where a unitary operation is described by a sequence of gates applied to a reference input, a MBQC is realized as a sequence of single-qubit measurements performed on highly entangled multiqubit states. These states are also known as graph states for their one-to-one correspondence with simple graphs $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$. Graph states are represented by a set of vertices \mathcal{V} , corresponding to single qubits initialized in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, and a set of edges \mathcal{E} , corresponding to pairwise controlled-Z entangling gates applied to the respective vertices (see Fig. 2 and Appendix B for details). The MBQC model is computationally equivalent to the circuit model for appropriate families of graphs [28], even when measurements are restricted to the xy plane of the Bloch sphere [34], as considered here.

One way to visualize a quantum computation as a MBQC pattern is to select, on a given graph, an (arbitrary) set of input vertices, which represents the initial state of the computation, and an equal-sized output set, which will contain the final state. Fixing these sets determines a unique set of paths from each input qubit to an output qubit, giving rise to the notions of (generalized) flow [35,36]. A deterministic computation then proceeds by sequentially subjecting each nonoutput qubit along the flow to a projective measurement $\Pi_\alpha = \frac{1}{\sqrt{2}}(|0\rangle \pm e^{-i\alpha}|1\rangle)$, at an angle α in the xy plane of the Bloch sphere, and applying outcome-dependent corrections to the neighboring qubits. By convention, a measurement outcome of zero requires no correction, whereas a measurement outcome of one

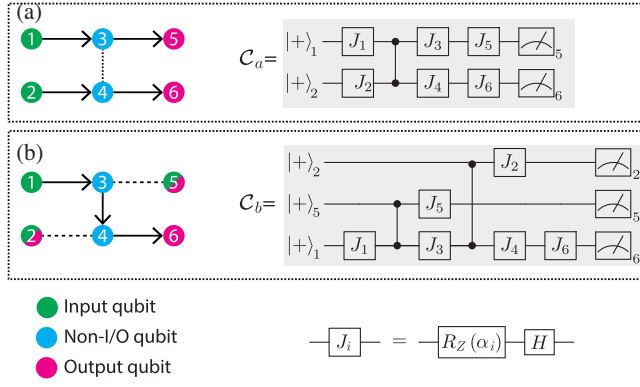


FIG. 2. Schematic representation of equivalent computations in MBQC (left) and in the circuit model (right). The underlying MBQC is based on a six-qubit H-shaped graph state with two different g-flows, consisting of (a) two and (b) three input/output qubits (other choices are possible). The direction of the flow is indicated by the arrows on the graph, with edges not involved in the flow shown as dashed lines. The qubits are measured according to the order of the labeling numbers. In panel (a), the input state of the circuit C_a is $|+\rangle_{12}$, associated with qubits 1 and 2 of the cluster state, whereas in panel (b), the input state of the circuit C_b is $|+\rangle_{251}$, associated with qubits 1, 2, and 5 of the cluster state. The qubit ordering in the circuit is chosen according to the output labels, and the detailed procedure for going from the graph and flow to the circuit can be found in Appendix B. Note that both quantum circuits on the right correspond to the same MBQC graph state on the left, albeit with a different flow. The basic single-qubit gate $\hat{J}(\alpha_i) = \hat{H}\hat{R}_z(\alpha_i)$ [for brevity, \hat{J}_i with $i = (1, \dots, 6)$] can be decomposed into a Hadamard gate \hat{H} and a rotation $\hat{R}_z(\alpha)$ around the z axis of the Bloch sphere.

requires an update of the measurement angles for subsequent measurements. Different choices of flow structure thus identify different possible circuits over a graph by determining the appropriate corrections for nonzero measurement outcomes and their order. The choice of measurement angles $\alpha = \{\alpha_1 \dots \alpha_n\}$ specifies an instance on this circuit.

The key insight that we make use of is that, although MBQC performs a deterministic computation between a specific choice of input and output sets, there are always multiple such choices for a given graph. Consequently, there are multiple possible information flows, which is a concept known as flow ambiguity [37]. These alternate flows give rise to computations with different structure, different number of logical qubits, and different flow-dependent corrections, which effectively insert random Z gates into the circuit; see Appendix B for details. Thus, a MBQC implementing a specific computation can also be seen as providing the outcomes of a random set of other computations, each related to a unique computation in the circuit model, which provides a natural means for testing different devices against each other in a reliable fashion.

We now use this insight to generate MBQC-related sampling problems by converting a given MBQC with angles α into the circuit model for different choices of output sets and flows. To illustrate this approach, consider the six-qubit H-shaped graph of Fig. 2, where one choice of flow gives rise to the $n_O = 2$ qubit circuit C_a [Fig. 2(a)], while a different choice gives rise to the $n_O = 3$ qubit circuit C_b [Fig. 2(b)]. The angles for the elementary single-qubit gates are determined by the measurement angles of the underlying MBQC and are common to both circuits. This commonality can, however, be hidden by subjecting each set of angles to random π flips and sign flips according to individual cryptographic one-time pads for each circuit; see Appendix B for details. Our goal now is to relate the outputs of these very different computations.

In the MBQC picture, the measurement outcomes over nonoutput qubits occur with equal probability of $2^{-(n-n_O)}$ [38], while the probabilities for the output qubits (Pr) depend on the choice of measurement angles α . Hence, there is no bias, and we can, without loss of generality, focus on the case of all-zero outcomes for the nonoutput qubits, where no flow-dependent corrections are necessary. We emphasize that this choice is arbitrary and merely affects the relation between the measurement angles of the two circuits but has no effect on the success probability in the circuit picture, where these qubits do not exist. Hence, in the MBQC picture, the probability for obtaining all-zero outcomes using the flow corresponding to C_a is $2^{-4} \text{Pr}(0,0)_{C_a}$, where $\text{Pr}(0,0)_{C_a}$ is the probability of obtaining zero outcomes when measuring only the two output qubits. Similarly, the probability for all-zero outcomes using the flow of C_b is $2^{-3} \text{Pr}(0,0,0)_{C_b}$. Crucially, since these probabilities are obtained from the same graph, they must agree on those qubits that are in the output of both circuits.

Using this fact, we find that the outcome probabilities obtained from the two circuits are related as $\text{Pr}(x,y)_{C_a} = 2 \text{Pr}(0,x,y)_{C_b}$ for $x,y \in \{0,1\}^2$. The case where the first output qubit of circuit C_b is in state 1 can also be used for verification, but it is related to a computation with different angles in circuit C_a ; see Appendix B for more details. Moreover, the output strings can be randomized by adding a random multiple of π to the measurement angles for the qubits in the output set. This approach allows us to create distinct quantum circuits, such that the independently sampled probabilities of obtaining particular (nonzero) strings as outputs are related.

The central observation here is that we can use this technique to establish a connection between the outcome probabilities from two quantum circuits with different width, depth, and structure, but with MBQC-related angles for the single-qubit gates. These structural differences and their MBQC relation ensure that realistic errors in honest devices have distinct effects in the two circuits. Consistent results can then only be obtained when the computations

are close to correct; see Appendix D for details. Implementing these circuits (\mathcal{C}_a and \mathcal{C}_b in the case of the presented experiments) on a single device thus provides a means for self-verification of the device, while implementing them on different devices provides a pathway to cross-validate the two devices. The only requirements for our technique is that the devices are honest and uncorrelated and that they can sample from the output distribution of the respective circuits.

III. CROSS-VERIFICATION

In order to formally turn this into a test of consistency between quantum devices, we now consider two quantum processors, implementing computations derived from the same MBQC but with different output sets. Note that one could also start from a given circuit and derive the corresponding MBQC pattern and, from there, an alternate circuit for a different flow. Hence, in practice, the two quantum processors will typically implement quantum circuits, with the MBQC picture merely supplying the relation between the otherwise random outcomes. The underlying hardware and native gate sets of the two quantum processors are irrelevant for the protocol. When processors 1 and 2 have output sets of size n_{O_1} and n_{O_2} , respectively, with n_c qubits in common, there are a total of $n_v = n_{O_1} + n_{O_2} - n_c$ variable qubits in the combined output of the two circuits. Hence, there are 2^{n_v} different measurement strings that can be obtained from the combined output of the two experiments. We now denote by \mathbf{p}_j the vector of probabilities for obtaining each of these 2^{n_v} possible strings from the quantum circuit \mathcal{C}_j performed on the j th device (normalized as above). In order to construct this vector, despite the fact that each device only measures a subset of the n_v qubits, we note that all unmeasured qubits take uniformly random values by definition. A nonzero outcome of such an unmeasured qubit effectively translates into a change in the single-qubit angles $\tilde{\alpha}$ due to the corresponding flow-dependent corrections in the underlying MBQC. Hence, the full vector \mathbf{p}_j can be constructed from the output of the j th device by measuring the circuit \mathcal{C}_j for a set of different angles α according to the unmeasured qubits.

For the circuits in Fig. 2, for example, the vector $\mathbf{p}_{a/b}$ would have eight elements. Here, device b obtains the full vector from its output, while device a only obtains the first four elements (corresponding to outcome “0” on the unmeasured qubit 2). To obtain the remaining four elements, device a measures the same circuit but with a different set of angles α , which are obtained from the original set α by absorbing the flow-dependent corrections according to a “1” outcome on qubit 2 in the underlying MBQC.

Since the output sets of both quantum processors are overlapping subsets of these n_v qubits, this allows for a

Algorithm 1. Cross-check protocol.

- (1) Select a MBQC pattern and derive two circuits with different, but overlapping output sets from it.
 - (1') Alternatively to (1), start with a circuit to be verified, generate the corresponding MBQC pattern, and then derive a second circuit with a different, but overlapping output set.
 - (2) For each device, generate a set of angles α , as well as a family of such sets by taking into account the flow-dependent corrections required for nonzero outcomes on all qubits in the combined output set (see text for details).
 - (3) (optional) For each device, generate random bit strings \mathbf{k}_j and \mathbf{r}_j to mask the real measurement angles and outcomes. See Appendix B for details.
 - (4) Sample from both devices randomly over all specified angle sets.
 - (5) From the samples, estimate the squared ℓ^2 distance in Eq. (1) as prescribed in the text. Optionally, also perform the sanity checks outlined below, using the *same* data.
-

direct comparison of the devices on this larger space. Specifically, we aim to quantify the similarity of the two outputs by estimating the squared ℓ^2 distance between the vectors \mathbf{p}_j (see Appendix C for details),

$$\|\mathbf{p}_1 - \mathbf{p}_2\|^2 = \mathbf{p}_1 \cdot \mathbf{p}_1 - 2\mathbf{p}_1 \cdot \mathbf{p}_2 + \mathbf{p}_2 \cdot \mathbf{p}_2. \quad (1)$$

The terms $\mathbf{p}_j \cdot \mathbf{p}_j$ in Eq. (1) quantify the probabilities of obtaining the same result twice when sampling from the j th device. These probabilities can be estimated from the minimum number of runs required to obtain a collision among output strings (i.e., obtain the same string twice). Notably, since the outcomes of unmeasured qubits are independent and uniformly random by definition, it is sufficient to estimate these terms from the output of the devices directly (rather than the full vector \mathbf{p}_j); see Appendix C for details. Consequently, estimating $\mathbf{p}_j \cdot \mathbf{p}_j$ requires at most $O(2^{n_{O_j}/2})$ runs, independent of the probability distribution, because of a generalization of the birthday paradox [39]. The term $\mathbf{p}_1 \cdot \mathbf{p}_2$ in Eq. (1) can be estimated in a similar fashion as the minimum number of runs to obtain a collision between devices. This estimate is achieved by randomly fixing the value of nonoutput qubits in each shot while sampling from the two devices, which, by the same argument as above, requires at most $O(2^{(n_{O_1} + n_{O_2} - n_c)/2})$ runs. The full protocol is outlined in Algorithm 1.

Hence, while this upper bound remains exponential in the number of output qubits, it remains feasible for at least twice as many qubits as conventional classical simulation techniques, which typically scale as $2^{n_{O_j}}$ [13,40]. In practice, the scaling is expected to be even better in cases where there is a significant number of output qubits in common between the instances ($n_c \sim n_{O_1}, n_{O_2}$) or where

the output distribution for either computation is far from uniform ($\max \mathbf{p}_j \gg 2^{-n_{o_j}}$). Here, the quantity $\|\mathbf{p}_1 - \mathbf{p}_2\|^2$ provides a measure of similarity between the outputs of the devices, which can be estimated with exponentially fewer resources than circuit simulation methods.

IV. EXPERIMENTAL RESULTS

We experimentally perform MBQC-related two- and three-qubit circuits with different depths on five independent state-of-the-art quantum processors, covering four of the major quantum computing architectures (see Fig. 1). On these devices, we implement sampling instances for the six-qubit H-shaped graph shown in Fig. 2 for up to 200 random sets of angles $\{\alpha_i\}_{i=1}^6$, with $\alpha_i \in \{0, \pi/4, 2\pi/4, 3\pi/4, 4\pi/4, 5\pi/4, 6\pi/4, 7\pi/4\}$. We run two-qubit circuits of type \mathcal{C}_a on the Oxford and Innsbruck systems; three-qubit circuits of type \mathcal{C}_b on the Innsbruck, IBM, and Rigetti systems; and 34 instances of the six-qubit H-shaped MBQC on the Vienna system.

After taking into account the relation between the outputs of the different implementations and additional randomization (see Appendix B for details), we estimate the squared ℓ^2 distances (1) between pairs of devices, implementing different size computations for cross-verification. Because of the small size of the computations here, the full probability distribution is easily obtained, such that we do not have to work with collisions between output strings as would be required for larger circuits. The squared ℓ^2 values are then averaged over 34 random instances, which were performed on all devices. These results are consistent with the full data set (see Appendix C). Additionally, we compute the squared ℓ^2 distance between the two- and three-qubit circuits implemented on the Innsbruck device for self-verification. Crucially, in either case, we do not compare the output distributions to some “ideal theory” (e.g., from simulations), which would not be possible for future devices.

Instead, the key results of these comparisons are the estimated squared ℓ^2 distance for each pair of devices performing computations of different size and depth, shown in Fig. 3(a). A value close to 0, for example, between Vienna and Oxford or Vienna and Innsbruck, indicates agreement between the devices, whereas any systematic error or statistical noise will lead to a larger value. For example, comparing an ideal two-qubit circuit to a fully depolarized circuit for the instances considered in our experiments would return a value of about 0.428. In the limit of large numbers of qubits, the output distribution of our MBQC-derived circuits is expected to follow a Porter-Thomas (PT) distribution [41]. Although, our technique works independently of the underlying distribution, using this observation, one finds that the squared ℓ^2 distance between an ideal and fully depolarized circuit tends to 2^{-n_v} , which implies that a minimal requirement for passing our

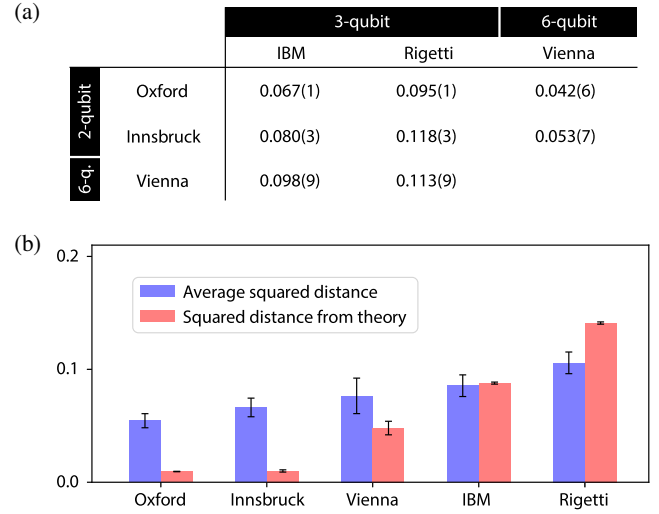


FIG. 3. Experimental comparison of quantum devices. (a) Experimental squared ℓ^2 distances between pairs of independent quantum devices, averaged over 34 instances. Only computations using different numbers of physical qubits are compared, so the implementations represent fundamentally different sampling problems. Uncertainties in parentheses correspond to 1 standard deviation of statistical noise. (b) Squared ℓ^2 distances for each quantum device averaged over comparisons to all other devices (blue). To illustrate that this metric qualitatively captures the performance of the individual devices, we also show the squared ℓ^2 distances of each device against the (nonscalable) theory prediction (red). These two quantities are not expected to coincide. However, arranging devices according to either metric yields the same order in our experiments.

cross-check protocol must be that $\ell^2 \ll 2^{-n_v}$. In addition, the values $\mathbf{p}_j \cdot \mathbf{p}_j$ from Eq. (1) provide some insights into the individual devices as an additional (optional) sanity check using the same data. In the extreme case of a fully depolarized output, this value would reach a global minimum of 2^{-n_v} , while a low-noise device is expected to lead to a value that is twice as large. Of course, the noise in real experiments is much more complicated, and the exact dependence of the squared ℓ^2 distance on such physical noise models remains an interesting question for future research.

When more than two devices are used, one can furthermore make a prediction about the performance of the individual devices by averaging the squared ℓ^2 distance with all other devices; see Fig. 3(b). To test whether these values indeed quantify the true accuracy of the respective devices, we compare them to the ℓ^2 distance with this ideal theory prediction, which is possible for the small circuits here. Note that this theory comparison is just for illustration and not required for the evaluation of our protocol on future devices. We find that the average squared ℓ^2 distance over all experimental comparisons qualitatively agrees with the theory expectation. This result indicates that these averages provide a good estimate of the true system performance,

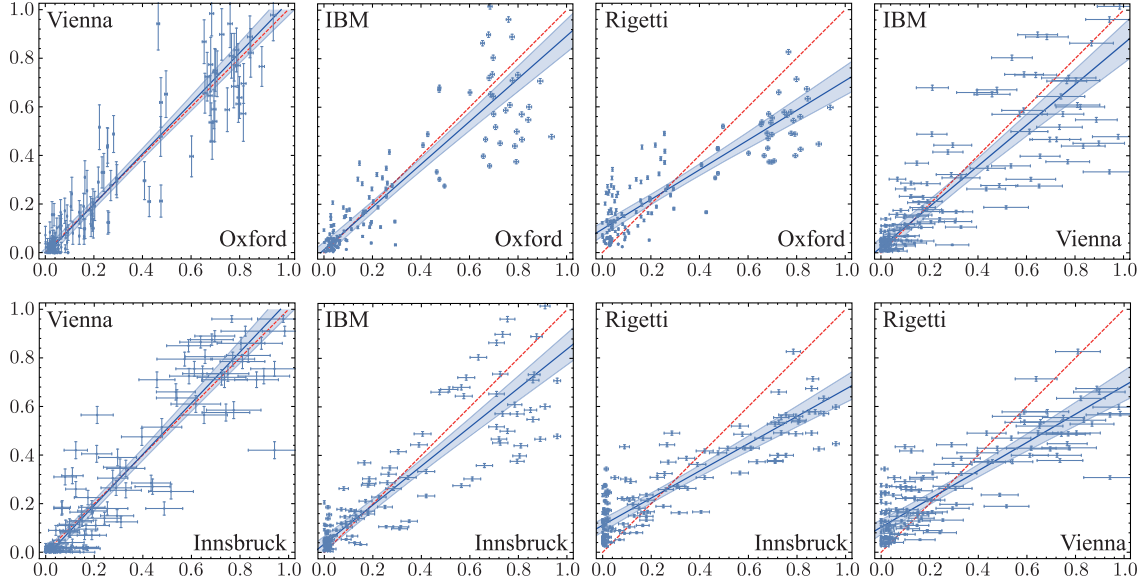


FIG. 4. Experimental comparison between outputs of computations for pairs of different devices. Data points represent the (rescaled) outcome probabilities of one device (horizontal axis, bottom label) versus another (vertical axis, top label), for MBQC-related instances. We compare all pairs of independent devices, implementing different-size circuits and based on different physical systems. In the ideal case, all points should, up to unavoidable shot noise, lie on the diagonal. For each data set containing 136 outcome probabilities (34 circuits and four outcome combinations each), we perform linear total least-squares regression (solid blue line) to quantify the deviation from the ideal correlation (red dashed line), yielding regression slopes with 1σ uncertainties of (top, left to right) 1.04(2), 0.88(3), 0.61(2), 0.84(3), (bottom, left to right) 1.05(2), 0.80(3), 0.56(2), and 0.58(2), respectively. Experimental error bars correspond to 1σ statistical uncertainty associated with the data points, and the blue shaded bands represent 3σ mean prediction intervals for the regression.

while the values from individual comparisons accurately capture the relative performance of the devices, thus enabling verification of the underlying quantum processors.

Besides cross-verification between dissimilar quantum devices, our technique also provides an intriguing pathway towards self-verification of a single device. Using the Innsbruck trapped-ion system, we implement 200 MBQC-related instances of the two- and three-qubit circuits C_a and C_b of Fig. 2 and estimate $\|\mathbf{p}_{C_a} - \mathbf{p}_{C_b}\|^2 = 0.033(1)$. This result indicates very good [relative to the results in Fig. 3(a)] agreement between the two circuits, which is confirmed by direct comparison to theory. Notably, by virtue of sampling from multiple instances of vastly different circuits, even systematic errors would be detected, as they manifest very differently in the two circuits (see Appendix D for details). This demonstrates that our technique can be used for independent verification of a single quantum processor.

V. FURTHER INSIGHTS INTO CROSS-VERIFICATION

In order to gain some insight into the measured ℓ^2 distances, correlation plots between the MBQC-related outcomes on all pairs of devices computing different size and depth circuits are shown in Fig. 4. We emphasize that it is not scalable to produce such plots for larger computations, as

they require measuring the full output distribution for multiple computations. However, for the devices and circuits used here, this task remains feasible and provides useful additional information to aid the interpretation of the ℓ^2 distances used for scalable cross-verification in Fig. 3.

While these correlation plots only provide a crude indication of the strength and direction of the correlations between the devices, there are some notable features. In the ideal case, one expects the two MBQC-related circuits to produce identical outcome probabilities, resulting in clustering around the 45° line. On the other hand, depolarizing noise affecting the device on the vertical (horizontal) axis would result in a mean correlation with slope smaller (larger) than 1. Large scattering around the mean further indicates higher levels of noise in one or both devices since each point corresponds to a different output or instance of the computation. In addition, Fig. 5 shows correlation plots of the experimental outcomes per individual device against the respective ideal theory prediction obtained via direct simulation of the corresponding circuits. Figure 6 shows correlation plots of experimental outcomes between the two- and three-qubit circuits C_a and C_b , performed on the Innsbruck device for self-verification.

VI. DISCUSSION

As quantum processors start to surpass their classical counterparts, verification by direct comparison to theory

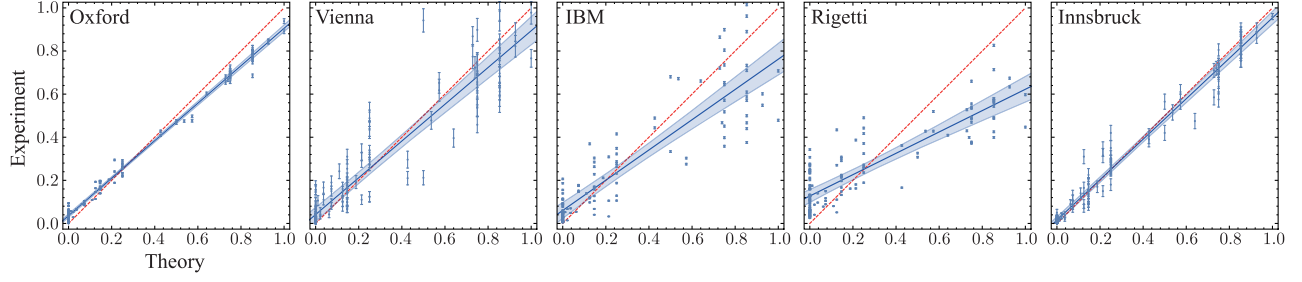


FIG. 5. Comparison between experimental outcome probabilities and theoretical expected values per single device. From left to right: Oxford, Vienna, IBM, Rigetti, and Innsbruck. As in Fig. 4, data points represent the (rescaled) outcome probabilities obtained from the respective device (Experiment), against the corresponding theory value (Theory) obtained from direct circuit simulation. For each data set, we perform linear least-squares regression (blue line) to quantify the deviation from the ideal correlation (red dashed line), yielding regression slopes with 1σ uncertainties of 0.869(6), 0.85(3), 0.70(3), 0.50(2), and 0.94(1), respectively. Experimental error bars correspond to 1σ statistical uncertainty, and the blue shaded bands represent 3σ mean prediction intervals for the regression.

will no longer be an option. The technique we present here provides a feasible alternative by validating quantum devices against each other. While not providing a complete toolkit for characterization of individual quantum processors, our technique takes a crucial step away from the dependence on classical methods. By sampling from different physical devices implementing circuits that differ in the number of qubits, depth, and structure, our technique is robust against systematic as well as statistical errors. While related work aims at estimating a fidelity between specific output states of two quantum processors [27], our method estimates the similarity of outputs for random computations, thereby verifying the full device performance. By implementing these dissimilar circuits on a single device, our technique also provides an avenue towards internal self-verification of single quantum devices.

A particularly intriguing feature of our technique is the way in which it allows us to compare devices using

radically different implementations. Recently, a detailed comparison of a trapped-ion system and a superconducting processor highlighted the advantages of each system for certain identical problems [42], concluding that each processor was well suited to different tasks. In this work, we overcome the heterogeneity of quantum physical systems, introducing a verification technique that links computational circuits with different sizes and depths and, consequently, can be run on the many types of quantum computers. The building block of our cross-check protocol is represented by measurement-based quantum computation, which was already proven to be essential for quantum computation security [43], quantum error correction [44], and quantum simulation [45]. This technique will prove useful in providing consistent benchmarks across the increasingly diverse range of quantum processors.

ACKNOWLEDGMENTS

The authors thank Sam Ferracin for helpful comments on the manuscript. T.M., R.B., P.S., M.R. gratefully acknowledge support by the Austrian Science Fund (FWF), through the SFB FoQuS (FWF Project No. F40) and SFB BeyondC (FWF Project No. F7109), by the EU H2020-FETFLAG-2018-03 under Grant Agreement No. 820495, by the U.S. Army Research Office (ARO) through Grant No. W911NF-14-1-0103, and by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via the U.S. ARO Grant No. W911NF-16-1-0070. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 801110 and the Austrian Federal Ministry of Education, Science and Research (BMBWF). P.W. acknowledges support from the research platform TURIS, the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS (No. W1210-4), BeyondC (F7113) and NaMuG (P30067-N36), the U.S. Air Force

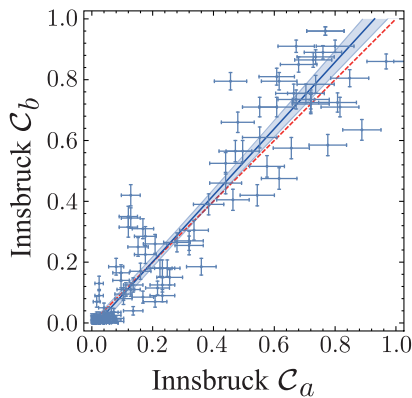


FIG. 6. Comparison between experimental outcome probabilities for the circuits C_a and C_b on the Innsbruck device. Linear total least-squares regression (solid blue line) yields a regression slope of 1.09(2) compared to the ideal value of 1 (dashed red line). Experimental error bars correspond to 1σ statistical uncertainty, and the blue shaded bands represent 3σ mean prediction intervals for the regression.

Office of Scientific Research via QAT4SECOMP (FA2386-17-1-4011), the European Commission via High Dimensional Quantum Photonic Platform Projects (HiPhoP) (No. 731473), and Red Bull GmbH. J. F. F. acknowledges support from Singapore's Ministry of Education, National Research Foundation (ANR-NRF Grant No. NRF2017-NRF-ANR004), and the U.S. Air Force Office of Scientific Research (AOARD Grant No. FA2386-18-1-4003).

APPENDIX A: EXPERIMENTAL METHODS

NMR (Oxford) experiments [7] were performed on a Varian Unity Inova spectrometer with a nominal ^1H frequency of 600 MHz using a $\text{H}\{\text{CN}\}$ probe with a single pulsed field gradient. The NMR sample comprised ^{13}C -labeled sodium formate dissolved in D_2O at 25 °C, providing a heteronuclear two-spin system. With both spins on resonance, the Hamiltonian took the form of a spin-spin ZZ coupling of 194.7 Hz, and the B_1 field strengths were measured to give nutation rates of approximately 25 kHz for ^1H and 17 kHz for ^{13}C .

Pseudopure two-qubit states were prepared using the method of Ref. [46]. Single-qubit rotations in the XY plane were implemented using simple pulses, while two-qubit rotations were implemented as delays. Fixed Z rotations were implemented as frame rotations [47], which were propagated through the pulse sequence [48] to points where they could be dropped. The variable small-angle Z rotations were implemented using a pair of π pulses with phases separated by $\theta/2$, with the phase of the first pulse chosen to partially cancel with the preceding Hadamard gate.

At the end of the algorithm a crush gradient was applied to project the density matrix onto the computational basis and the ^1H NMR spectrum observed after a $\pi/2$ pulse. NMR spectra were processed using custom software, and the intensity of the two components of the ^1H doublet were determined by integration and normalized to a reference spectrum. Corresponding measurements on the second qubit were performed by repeating the experiment with the reverse assignment of qubits to physical spins. From the collection of these measurements, the populations of the computational basis states can be estimated. Because of the imperfect calibration, these populations do not quite sum to one, and some can be slightly negative. This issue was resolved by subtracting the most negative population found in any group of experiments from all the populations in that group and by then normalizing the populations for each experiment.

Photons (Vienna) experiments were based on the generation of the maximally entangled six-qubit H-shaped cluster state. Three polarization-entangled pairs of photons were produced via three identical Sagnac-PPKTP pulsed downconversion sources and later entangled by using partial fusion gates at polarizing beam splitters [49,50].

The qubits were encoded by the polarization of the six photons.

The laser repetition rate was set to 152 MHz, by doubling the original rate with a passive multiplexing scheme [51], to reduce multiphoton noise for an average power of 220 mW per source. The pump photons had a wavelength of 772.9 nm and a pulse width of 2.1 ps. The crystals' temperature was stabilized at 24 °C. The single-qubit measurements were implemented with an optics tomographic unit of three motorized wave plates and a polarizing beam splitter per photon. Twelve, multi-element, superconducting-nanowire, single-photon detectors, composed of four channels each and kept at $T = 0.9$ K, enabled a pseudonumber resolving detection, with an average quantum efficiency of 0.87. Because of technical problems, two of the multi-element detectors had to be replaced with two single-element detectors. A customized time tagging and logic module for 48 input channels counted the sixfold photon events. After postselection, we obtained a total sixfold coincidence rate of 0.08 Hz. The purity of the single photons, measured with fourfold HOM interference, corresponded to 0.94 [33]. We characterized the six-photon cluster state by using a subset of stabilizer operators, the so-called identity product [52], giving a lower bound on the state fidelity of $F_{\text{exp}} \geq 0.64 \pm 0.04$, and by using a technique based on a probabilistic protocol for entanglement detection [50,53], estimating a fidelity of 0.75 ± 0.06 .

Superconducting (IBM and Rigetti) qubits were used independently via the two cloud-accessible quantum processors: the ibmqx2 (also known as IBM Q 5 Yorktown) from IBM [31,54] and the Rigetti 19Q from Rigetti [32,55]. Both apparatuses used transmon qubits—charge qubits that show insensitivity to charge noise thanks to an additional large capacitor in the circuit. Variations of the two devices can be found on the circuit wiring and reading, and the fabrication materials, e.g., the ibmqx2, had a star-shaped connected circuit, based on fixed-frequency transmons [56], with three qubits available as control qubits, whereas the Rigetti 19Q had tunable-frequency transmon qubits [57,58], each coupled to three fixed-frequency transmon qubits.

The ibmqx2 processor was calibrated twice during our experimental runs and kept at a temperature of 17.5 mK. We selected qubits 2, 3, and 4 with frequencies of [5.2, 5.0, 5.3] GHz, single-qubit gate errors of $[3.4, 3.6, 3.3] \times 10^{-3}$, and readout errors of $[3.5, 1.5, 1.6] \times 10^{-2}$, respectively. The two-qubit gate errors consisted of $[6.7, 3.7] \times 10^{-2}$ for the controlled gate between qubits 2 and 3 and between qubits 2 and 4, respectively. The coherence times were [48.5, 51.7, 39.4] μs for depolarization and [28.9, 75.6, 49.9] μs for spin dephasing, whereas the gate time was about 250 ns.

On the Rigetti 19Q, we exploited the qubits labeled 2, 8, and 13, as they showed reduced noise. The chip was maintained at a temperature of 10 mK. From Ref. [58], single-qubit readout fidelities were equal to 0.97, 0.947,

and 0.921; single-qubit gate fidelities corresponded to 0.981, 0.987, and 0.993 for qubits 2, 8, and 13, respectively; and two-qubit gate fidelities were 0.906 (between qubits 2 and 8) and 0.881 (between qubits 8 and 13). The qubits coherence time was about 20 μ s, whereas the entangling gate time was 100–250 ns.

Trapped-ion (Innsbruck) experiments were performed with qubits encoded in the electronic states of a string of $^{40}\text{Ca}^+$ ions confined in a linear Paul trap [8]. Each ion encoded a qubit in the ground state $S_{1/2}(m = -1/2) = |1\rangle$ and the metastable state $D_{5/2}(m = -1/2) = |0\rangle$, which determined the qubit lifetime of about 1 s. Single-qubit Z rotations were implemented via Stark shifts induced by tightly focused laser beams, while collective rotations around any equatorial axis of the Bloch sphere were achieved by resonant illumination of the whole ion string. Entangling operations were implemented via global Mølmer-Sørensen interactions using a bichromatic laser field [59]. Local gates, as well as two-qubit entangling gates, achieved fidelities greater than 99% and operated on a timescale of 20–30 μ s (80 μ s for entangling gates), much faster than the coherence time, which is on the order of about 100 ms and dominated by laser phase noise. Every run of the experiment consisted of Doppler and sideband cooling of the ion string, followed by a gate sequence, and finally projection onto the computational subspace via fluorescence detection on the $P_{1/2}$ – $S_{1/2}$ transition with a CCD camera. One such run took about 15 ms, and each experiment was repeated at least 100 times to gather statistics.

1. Photonic H-shaped cluster-state characterization

We first characterized the six-photon cluster state using a technique based on subsets of stabilizer operators, referred to as identity products (ID) [52]. The method exploited the entanglement of the operators to obtain a lower bound on the fidelity of the state and a proof of a Bell-type inequality with a minimal number of measurement settings. There existed a large, unquantified number of equivalent minimal subsets of stabilizers for the six-qubit states. Here, we repeated the characterization procedure with two equivalent ID sets, composed of seven measurements:

ZZIII	ZZIII
ZZIXXX	ZZIXXX
IZIXYY	IZIYYX
XYXYYY	YYYXYX
YYYYXX	XXYYYY
XXYYXX	XYXXYX
YXXYXX	YXXXYX
(a)	(b)

where we omitted the tensor product symbols for compactness. From the ID measurements, we extracted an ID-Bell

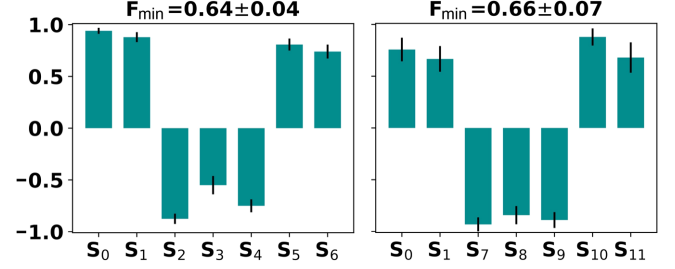


FIG. 7. Expectation values of the measured stabilizer operators for the two identity products related to the six-qubit H-shaped cluster state.

parameter $\langle\alpha_{\text{ID}}\rangle_{\text{exp}}$, where $\alpha_{\text{ID}} = \sum_i^M \lambda_i O_i$ and M is the number of measurement settings in the ID and O_i is the i th stabilizer operator of the ID. A Bell-type violation, in this case, was obtained if $\langle\alpha_{\text{ID}}\rangle_{\text{exp}} > M - 2$. The experimental results showed violations of the Bell-type inequality of 3.4σ and 2.4σ , respectively. A minimum value of the fidelity can be calculated as $F_{\text{min}} = (\langle\alpha_{\text{ID}}\rangle_{\text{exp}} - M + 4)/4$, providing $F_{\text{min}} = 0.64 \pm 0.04$ for the first ID and $F_{\text{min}} = 0.66 \pm 0.07$ for the second ID. The error bars were reduced for the first set because of a longer acquisition time: 1.5 h and 0.7 h for the first and second sets, respectively. The experimental expectation values for the two IDs are reported in Fig. 7. In both cases, the results of the six-qubit H-shaped cluster state show a violation of the ID-Bell inequality and high minimal fidelity. The nonideal results are mainly due to the unbalanced losses present at the polarizing-beam-splitter stage, the imperfect polarization compensation along the single-mode fibers connecting the three sources, and the nonunity purity of the single photons.

Furthermore, we followed a probabilistic protocol for entanglement detection [50,53] in order to estimate the experimental fidelity of the state. This method entailed a significant reduction of resources; that is, in our case, only a very low number of detection events (around 100) was needed to verify the presence of entanglement in our cluster state with more than 99% confidence. We obtain a fidelity of 0.75 ± 0.06 , which is comparable to fidelities obtained in state-of-the-art photonic experiments [60]. More details about the cluster-state characterization can be found in Ref. [50].

APPENDIX B: CONVERTING BETWEEN CIRCUIT MODEL AND MBQC

The reference gate in MBQC was $\hat{J}(\alpha) = \hat{H}\hat{R}_z(\alpha)$, which follows from the basics of the one-qubit teleportation scheme [34,61]. Single-qubit universality was obtained by realizing that $\hat{J}(\alpha)\hat{J}(0) = \hat{R}_x(\alpha)$.

The underlying graph for the MBQC pattern can be constructed by decomposing a generic unitary computation on a fixed initial state, $|+\rangle^{\otimes N}$, in terms of $\hat{J}(\alpha)$ gates and \hat{C}_Z

entangling gates. For each $\hat{J}(\alpha)$ gate, we added a vertex and drew an edge to connect this vertex to the vertex that represents the preceding $\hat{J}(\alpha)$ gate as dictated by the circuit. This process was carried out recursively, hence creating N wires, which represent the unitary evolution of each initial qubit state. The last step was drawing an edge for each \hat{C}_Z gate, by connecting the two vertices representing the \hat{J} gates that immediately follow the \hat{C}_Z gate in the quantum circuit representation. These few steps gave us the adjacency matrix of a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, with vertex set \mathcal{V} and edge set \mathcal{E} . The cardinality of the vertex set was $|\mathcal{V}| = N + M$, where M is the total number of $\hat{J}(\alpha)$ gates in the circuit.

1. Circuits for the six-qubit H-shaped cluster

We considered the two quantum circuits shown in Fig. 2 of the main text, associated with the six-qubit H-shaped cluster state. The unitary evolution of two g-flows corresponded to

$$\hat{\mathbf{J}}(\alpha_5, \alpha_6) \hat{\mathbf{J}}(\alpha_3, \alpha_4) C_Z \hat{\mathbf{J}}(\alpha_1, \alpha_2) |++\rangle, \quad (\text{B1})$$

$$\hat{\mathbf{J}}(\alpha_6)_3 \hat{\mathbf{J}}(\alpha_2, \alpha_4)_{1,3} C_{Z(1,3)} \hat{\mathbf{J}}(\alpha_5, \alpha_3)_{2,3} C_{Z(2,3)} \hat{\mathbf{J}}(\alpha_1)_3 |++\rangle, \quad (\text{B2})$$

where $\hat{\mathbf{J}}(\alpha_i, \alpha_j) = J(\alpha_i)_1 \otimes J(\alpha_j)_2$ and $\hat{\mathbf{J}}(\alpha_i, \alpha_j, \alpha_k) = J(\alpha_i)_1 \otimes J(\alpha_j)_2 \otimes J(\alpha_k)_3$. When using $\hat{\mathbf{J}}(\alpha_i)$ or $\hat{\mathbf{J}}(\alpha_i, \alpha_j)$ in a circuit with more qubits, this is implicitly understood as acting on the first set of qubits, unless a subscript indicate which qubits were acted on. The angles α can be randomly chosen within a specific set. The relationships between MBQC-related outcomes of the circuits \mathcal{C}_a and \mathcal{C}_b were

$$\begin{aligned} \Pr(0, 0)_{\mathcal{C}_a} &= 2 \cdot \Pr(0, 0, 0)_{\mathcal{C}_b}, \\ \Pr(0, 1)_{\mathcal{C}_a} &= 2 \cdot \Pr(0, 0, 1)_{\mathcal{C}_b}, \\ \Pr(1, 0)_{\mathcal{C}_a} &= 2 \cdot \Pr(0, 1, 0)_{\mathcal{C}_b}, \\ \Pr(1, 1)_{\mathcal{C}_a} &= 2 \cdot \Pr(0, 1, 1)_{\mathcal{C}_b}, \end{aligned} \quad (\text{B3})$$

where the labels of the outcomes are $\Pr(b_5, b_6)_{\mathcal{C}_a}$ and $\Pr(b_2, b_5, b_6)_{\mathcal{C}_b}$.

The MBQC protocols can be expressed in terms of the stabilizer formalism [62]. A graph state is invariant under stabilizer operations: Given a graph state on n qubits $|\mathcal{G}\rangle = (\prod_{\mathcal{G}} \hat{C}_Z) |+\rangle^{\otimes n}$, we have

$$\hat{K}_v |\mathcal{G}\rangle = |\mathcal{G}\rangle, \quad \forall v \in \mathcal{V}. \quad (\text{B4})$$

We can rewrite the computation by applying a stabilizer operator on each vertex of the graph state [28, 29]. We considered the stabilizers in their most general form, not restricting to the Pauli group, and a random bit string

$\mathbf{k} = \{k_i\}_{i=1}^6$, $k_i \in \mathbb{Z}_2$ associated with the six stabilizers. Then, the measurement angles α can be rewritten as

$$\text{angles} = \begin{cases} \tilde{\alpha}_1 = (-1)^{k_1} \alpha_1 + k_3 \pi \\ \tilde{\alpha}_2 = (-1)^{k_2} \alpha_2 + (k_4 + r_1) \pi \\ \tilde{\alpha}_3 = (-1)^{k_3} \alpha_3 + (k_1 + k_4 + k_5) \pi \\ \tilde{\alpha}_4 = (-1)^{k_4} \alpha_4 + (k_2 + k_3 + k_6) \pi \\ \tilde{\alpha}_5 = (-1)^{k_5} \alpha_5 + (k_3 + r_2) \pi \\ \tilde{\alpha}_6 = (-1)^{k_6} \alpha_6 + (k_4 + r_3) \pi \end{cases} \quad (\text{B5})$$

where $\mathbf{r} = \{r_i\}_{i=1}^3$ and can be used to mask the real outcomes of the computation. Note that finding these relations, and thus identifying MBQC-related sampling problems, is computationally efficient because of the graph structure of the problem.

For example, we selected the original angle set—randomly generated—to be $\alpha = \{\frac{3}{4}\pi, \frac{7}{3}\pi, (\pi/3), 0, \frac{2}{3}\pi, \pi\}$ and the random strings to be $\mathbf{k} = \{1, 0, 0, 0, 1, 0\}$, $\mathbf{r} = \{0, 1, 1\}$. Using the relations above, we obtained $\tilde{\alpha} = \{\frac{5}{4}\pi, \frac{7}{3}\pi, \frac{7}{3}\pi, 0, (\pi/3), 0\}$. Simulating the two circuits, we obtained

$$\Pr(b_5, b_6) = \begin{cases} 0.207 & \text{if } b_5, b_6 = (0, 0), \\ 0.393 & \text{if } b_5, b_6 = (0, 1), \\ 0.043 & \text{if } b_5, b_6 = (1, 0), \\ 0.357 & \text{if } b_5, b_6 = (1, 1), \end{cases} \quad (\text{B6})$$

and

$$\Pr(b_6, b_2, b_5) = \begin{cases} 0.179 & \text{if } b_2, b_5, b_6 = (0, 0, 0), \\ 0.021 & \text{if } b_2, b_5, b_6 = (0, 0, 1), \\ 0.196 & \text{if } b_2, b_5, b_6 = (0, 1, 0), \\ 0.104 & \text{if } b_2, b_5, b_6 = (0, 1, 1), \\ 0.060 & \text{if } b_2, b_5, b_6 = (1, 0, 0), \\ 0.064 & \text{if } b_2, b_5, b_6 = (1, 0, 1), \\ 0.065 & \text{if } b_2, b_5, b_6 = (1, 1, 0), \\ 0.311 & \text{if } b_2, b_5, b_6 = (1, 1, 1). \end{cases} \quad (\text{B7})$$

We compared the following strings:

$$\begin{aligned} \Pr(0, 0)_{\mathcal{C}_a} &= 2 \cdot \Pr(0 \oplus r_1, 0 \oplus r_2, 0 \oplus r_3)_{\mathcal{C}_b} \\ &= 2 \cdot \Pr(0, 1, 1)_{\mathcal{C}_b}, \end{aligned} \quad (\text{B8})$$

$$\begin{aligned} \Pr(0, 1)_{\mathcal{C}_a} &= 2 \cdot \Pr(0 \oplus r_1, 0 \oplus r_2, 1 \oplus r_3)_{\mathcal{C}_b} \\ &= 2 \cdot \Pr(0, 1, 0)_{\mathcal{C}_b}, \end{aligned} \quad (\text{B9})$$

$$\begin{aligned} \Pr(1, 0)_{\mathcal{C}_a} &= 2 \cdot \Pr(0 \oplus r_1, 1 \oplus r_2, 0 \oplus r_3)_{\mathcal{C}_b} \\ &= 2 \cdot \Pr(0, 0, 1)_{\mathcal{C}_b}, \end{aligned} \quad (\text{B10})$$

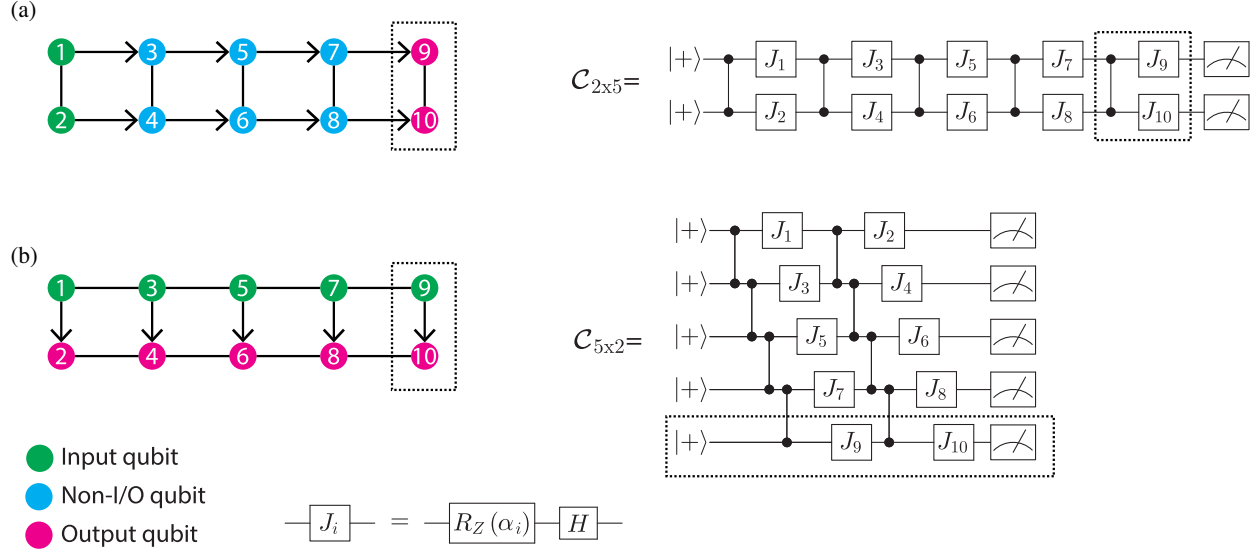


FIG. 8. Box cluster 2×5 with two choices of g-flow. In diagram (a), the “left-to-right” flow maps to the two-qubit, depth-5 circuit $C_{2 \times 5}$, whereas in diagram (b), “top-to-bottom” flow maps to a five-qubit, depth-2 circuit $C_{5 \times 2}$. The construction for the box cluster 2×4 is equivalent to the elements in dashed borders being removed.

$$\begin{aligned} \Pr(1, 1)_{C_a} &= 2 \cdot \Pr(0 \oplus r_1, 1 \oplus r_2, 1 \oplus r_3)_{C_b} \\ &= 2 \cdot \Pr(0, 0, 0)_{C_b}. \end{aligned} \quad (\text{B11})$$

By checking the outcomes above, we confirmed the correctness of the relations.

APPENDIX C: SQUARED ℓ^2 DISTANCE

The central figure of merit for scalably comparing MBQC-related circuits is the squared ℓ^2 distance introduced in the main text, Eq. (1). In general, the number of output qubits n_{O_1} and n_{O_2} for the two circuits will differ, and a subset of n_c of these qubits will be in the output set of both computations. For example, the circuits in Fig. 8 have $n_{O_a} = 2$ and $n_{O_b} = 5$ output qubits, respectively, and an $n_c = 1$ qubit that is in both output sets. Therefore, there are $n_v = n_{O_a} + n_{O_b} - n_c = 6$ qubits (in the underlying MBQC) that must be considered for estimating the squared ℓ^2 distance in Eq. (1) in the main text.

A crucial observation is that, because of a generalization of the birthday paradox [39], the term $\mathbf{p}_j \cdot \mathbf{p}_j$ is related to the probability of obtaining the same output string from the j th device twice. Recall that the vectors \mathbf{p}_j consist of the probabilities of obtaining each of the 2^{n_v} measurement strings, while the experiment only produces strings of length $n_{O_j} \leq n_v$. Hence, $\mathbf{p}_j \cdot \mathbf{p}_j$ turns out to be the weighted average over multiple instances of the j th circuit. Using the fact that in the MBQC picture the outcomes of nonoutput qubits are uniformly random, it is thus trivially sufficient to

estimate the term $\mathbf{p}_j \cdot \mathbf{p}_j$ from the output of only one instance of the circuit, requiring a number of runs that scales as $O(2^{n_{O_j}/2})$.

Estimating the term $\mathbf{p}_1 \cdot \mathbf{p}_2$, on the other hand, requires us to consider collisions among the strings of length n_v , which is achieved by randomly fixing the values of non-output qubits for either circuit (which corresponds to running different computations according to MBQC corrections for nonzero outcomes) for each sampling run. In the example of Fig. 8, circuit *a* requires randomly fixing four of the six combined output qubits and measuring the others, while circuit *b* requires fixing one of the six combined output qubits and measuring the rest. Using this technique, one can estimate the probability for a collision among the 2^6 possible strings m in the underlying MBQC between the two devices with a number of runs that scales as $O(2^{(n_{O_1} + n_{O_2} - n_c)/2})$.

1. Subsampling

The data presented in the main text were obtained by averaging the squared ℓ^2 distances obtained from 34 out of 200 randomly chosen instances over the related circuits. Averaging over more instances would increase the confidence in the final estimate; however, this comes with an additional resource overhead. To investigate this trade-off, we estimate the squared ℓ^2 distances from subsets of varying sizes up to the full data set. The results, shown in Fig. 9, indicate quick convergence to the mean value over the full data set. In particular, the choice of 34 instances is sufficient to clearly distinguish the different devices.

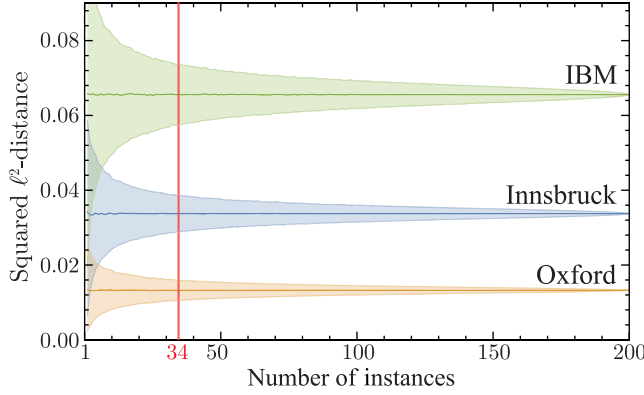


FIG. 9. Averaged squared ℓ^2 distances for different numbers of averaged instances. We show data for the Innsbruck self-verification (blue), the cross-verification between Innsbruck and Oxford (orange), and the cross-verification between Innsbruck and IBM (green). The solid lines correspond to the mean values over all samples of a fixed size, and the shaded regions depict the 1σ spread of these values. Statistical uncertainties are not taken into account in this analysis.

APPENDIX D: FROM CONSISTENCY TO VERIFICATION

Verification is often taken to consist of three requirements: consistency, completeness, and correctness. The protocol introduced in the main text quantifies *consistency* by the estimated ℓ^2 distance between computational outputs. In this context, we emphasize again that, although different computations could, in principle, lead to the same measurement outcomes in a single, fixed measurement, our figure of merit $\|\mathbf{p}_1 - \mathbf{p}_2\|^2$ is not susceptible to such a problem since it involves sampling over the single-qubit gate angles, thereby realizing a family of distinct

computational instances. Completeness is also satisfied by randomly sampling from the possible computations on a given cluster state. Correctness, while not immediately obvious, is addressed under the additional assumptions that devices act honestly and errors are not correlated between devices.

Recall that the protocol in the main text establishes consistency between two (or more) circuits corresponding to different flows in an underlying MBQC pattern. The resulting structural differences between the two circuits ensure that local single- or two-qubit errors in one circuit produce effects that can only be mimicked by correlated errors at different locations in the other circuit (see Fig. 10). Hence, for the two circuits to produce consistent yet incorrect results, errors would have to be increasingly fine-tuned and correlated between devices. On the other hand, for honest devices where errors accumulate independently, the output distributions will quickly diverge in terms of their ℓ^2 distance.

APPENDIX E: COMPLEMENTARY RESULTS

Here, we report data related to the evaluation of quantum circuits equivalent to closed lattice cluster states, without performing the respective measurement-based quantum computation. Specifically, we consider the closed 2D cluster states involving eight and ten qubits shown in Fig. 8. We refer to those as box cluster 2×4 and box cluster 2×5 , respectively, where $2 \times j$ ($j = [4, 5]$) labels the height and width of the cluster.

Different types of circuits are performed on pairs of quantum devices. In the following table, all the MBQC-related devices with the implemented circuit specifications (input qubits and circuit depth) are reported. Different instances of 2×4 and 2×5 clusters realized on different devices.

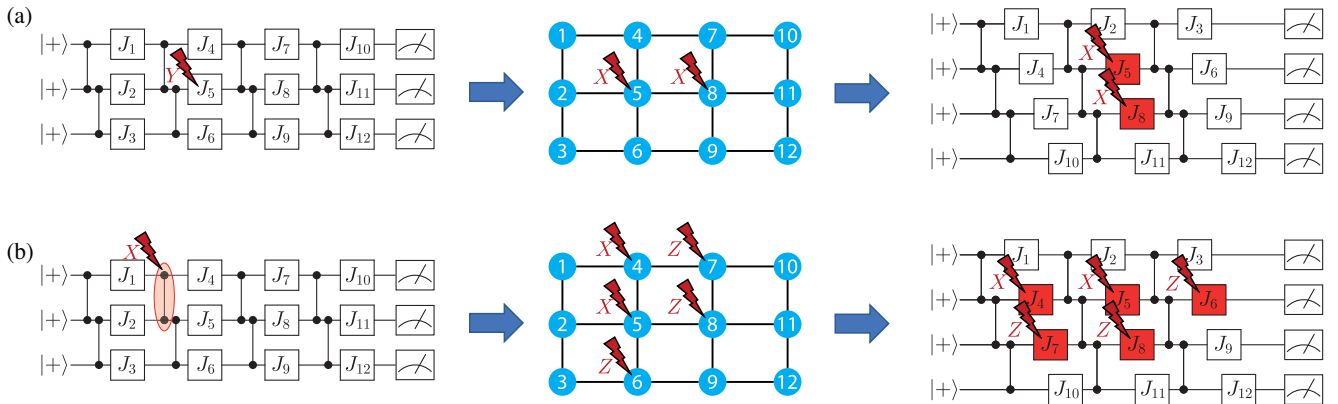


FIG. 10. Examples of the required error correlations between the two MBQC-derived circuits for the protocol to fail. The left circuit corresponds to a horizontal flow in a 3×4 lattice cluster, where either a (a) single-qubit or a (b) two-qubit error occurs. The effect of either error in the underlying MBQC pattern is shown in the middle. Translating to the circuit picture for a vertical flow, the right circuit then shows the error that would be required on the other device (correlated with the first) for the two circuits to still produce consistent results.

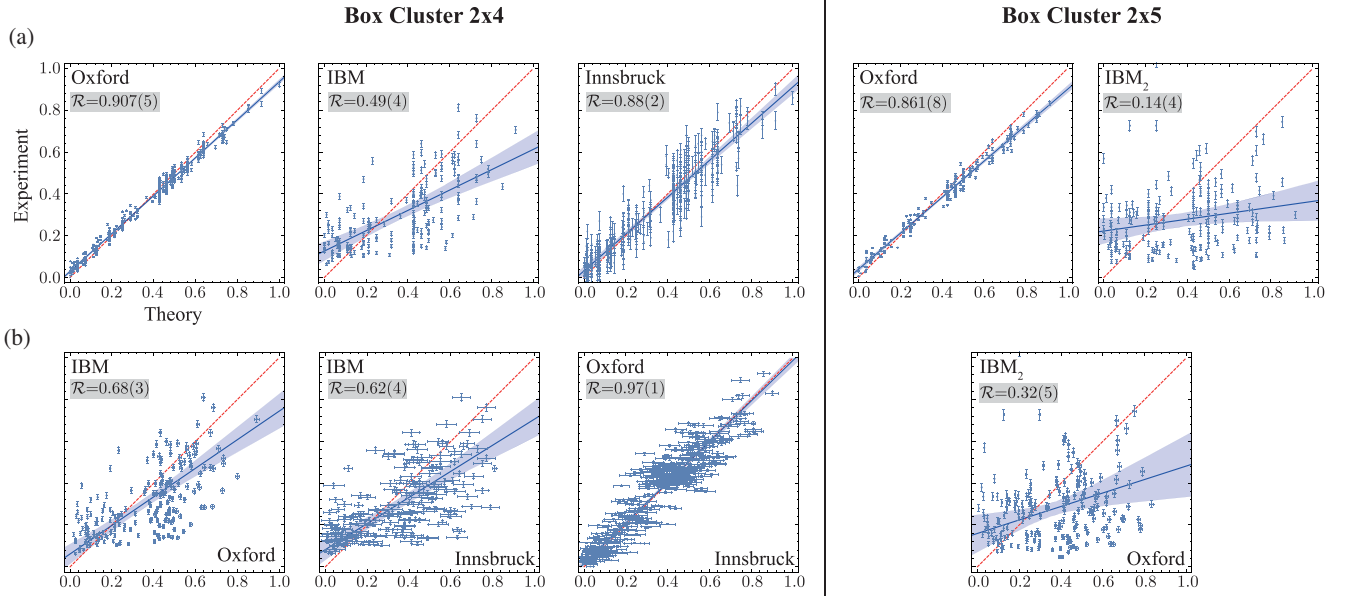


FIG. 11. Data for cross-verification using the box clusters 2×4 (left) and 2×5 (right). (a) Scatter plots of outcome probabilities compared to the theoretical expectation for the following (from left to right): Oxford, Innsbruck, IBM for the 2×4 -cluster, and Oxford, IBM₂ for the 2×5 cluster. The squared ℓ^2 distances $\|\mathbf{p}_1 - \mathbf{p}_2\|^2$ for these theory comparisons are 0.00200(3), 0.0578(5), 0.0123(6), 0.00400(8), 0.145(1). This is in good agreement with the trends seen from a linear least-squares regression (blue line) quantifying the deviation from the ideal correlation (red dashed line), with the resulting regression slopes \mathcal{R} with 1σ uncertainties given in the top-left corner of the respective figure panel. (b) Scatter plots of the outcome probabilities for 2×2 cross-check verification between Oxford-IBM, Innsbruck-IBM, and Innsbruck-Oxford for the 2×4 cluster, and Oxford-IBM₂ for the 2×5 cluster. The squared ℓ^2 distances $\|\mathbf{p}_1 - \mathbf{p}_2\|^2$ for the cross-validations are 0.0504(6), 0.060(2), 0.0115(6), 0.118(1). This is in good agreement with the regression coefficients \mathcal{R} obtained from linear, total, least-squares regression (blue line) given in the top-left corner of the respective figure panel. Experimental error bars correspond to 1σ statistical uncertainty, and the blue shaded bands represent 3σ mean prediction intervals for the regression.

Box cluster	Input	Depth	Q device
2×4	2	4	Oxford
	2	4	Innsbruck
	4	2	IBM
2×5	2	5	Oxford
	5	2	IBM ₂

Note that the 5×2 cluster was measured using a different IBM device, namely, the ibmqx3 [63] (IBM Q 16 Rueschlikon), which we refer to as IBM₂. This device has similar specifications as the first IBM quantum processor used here, but it allows for computations with up to 16 qubits.

1. Box cluster

As in the main text, the equivalences for the outcome probabilities obtained from the two circuits based on the box cluster 2×4 are

$$\begin{aligned}\Pr(0,0)_{\mathcal{C}_{2 \times 4}} &= 4 \cdot \Pr(0,0,0,0)_{\mathcal{C}_{4 \times 2}}, \\ \Pr(0,1)_{\mathcal{C}_{2 \times 4}} &= 4 \cdot \Pr(0,0,0,1)_{\mathcal{C}_{4 \times 2}}.\end{aligned}$$

Similarly, for the box cluster 2×5 , we obtain

$$\begin{aligned}\Pr(0,0)_{\mathcal{C}_{2 \times 5}} &= 8 \cdot \Pr(0,0,0,0,0)_{\mathcal{C}_{5 \times 2}}, \\ \Pr(0,1)_{\mathcal{C}_{2 \times 5}} &= 8 \cdot \Pr(0,0,0,0,1)_{\mathcal{C}_{5 \times 2}}.\end{aligned}$$

We ran 100 $\mathcal{C}_{2 \times 4}$ circuits on the Oxford and Innsbruck machines, with the 100 MBQC-related $\mathcal{C}_{4 \times 2}$ circuits run on the IBM processor. For the 2×5 case, we ran 100 $\mathcal{C}_{2 \times 5}$ circuits on the Oxford machine and the 100 MBQC-related $\mathcal{C}_{5 \times 2}$ circuits on the IBM processor. In each case, pairwise cross-check verification was performed between all devices, as well as individual comparisons to theory. Scatter plots of the outcome probabilities are shown in Fig. 11, and all relevant numerical values are given in the caption of that figure. As in the main text, we find that the squared ℓ^2 distance provides a very good estimate of the true performance of the devices, in agreement with the theory simulation.

[1] E. Bernstein and U. Vazirani, *Quantum Complexity Theory*, SIAM J. Comput. **26**, 1411 (1997).

- [2] L. K. Grover, *Quantum Mechanics Helps in Searching for a Needle in a Haystack*, *Phys. Rev. Lett.* **79**, 325 (1997).
- [3] D. R. Simon, *On the Power of Quantum Computation*, *SIAM J. Comp.* **26**, 1474 (1997).
- [4] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, *SIAM J. Comput.* **26**, 1484 (1997).
- [5] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
- [6] A. W. Harrow and A. Montanaro, *Quantum Computational Supremacy*, *Nature (London)* **549**, 203 (2017).
- [7] J. A. Jones, *Quantum Computing with NMR*, *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 91 (2011).
- [8] P. Schindler, D. Nigg, T. Monz, J. T. Barreiro, E. Martinez, S. X. Wang, S. Quint, M. F. Brandl, V. Nebendahl, C. F. Roos, M. Chwalla, M. Hennrich, and R. Blatt, *A Quantum Information Processor with Trapped Ions*, *New J. Phys.* **15**, 123012 (2013).
- [9] S. Barz, I. Kassal, M. Ringbauer, Y. O. Lipp, B. Dakić, A. Aspuru-Guzik, and P. Walther, *A Two-Qubit Photonic Quantum Processor and Its Application to Solving Systems of Linear Equations*, *Sci. Rep.* **4**, 6115 (2014).
- [10] M. H. Devoret and R. J. Schoelkopf, *Superconducting Circuits for Quantum Information: An Outlook*, *Science* **339**, 1169 (2013).
- [11] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, *Probing the Relaxation Towards Equilibrium in an Isolated Strongly Correlated One-Dimensional Bose Gas*, *Nat. Phys.* **8**, 325 (2012).
- [12] S. Braun, M. Friesdorf, S. S. Hodgman, M. Schreiber, J. P. Ronzheimer, A. Riera, M. del Rey, I. Bloch, J. Eisert, and U. Schneider, *Emergence of Coherence and the Dynamics of Quantum Phase Transitions*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3641 (2015).
- [13] F. Arute *et al.*, *Quantum Supremacy Using a Programmable Superconducting Processor*, *Nature (London)* **574**, 505 (2019).
- [14] E. Magesan, J. M. Gambetta, and J. Emerson, *Scalable and Robust Randomized Benchmarking of Quantum Processes*, *Phys. Rev. Lett.* **106**, 180504 (2011).
- [15] A. Erhard, J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt, *Characterizing Large-Scale Quantum Computers via Cycle Benchmarking*, *Nat. Commun.* **10**, 5347 (2019).
- [16] J. J. Wallman and J. Emerson, *Noise Tailoring for Scalable Quantum Computation via Randomized Compiling*, *Phys. Rev. A* **94**, 052325 (2016).
- [17] J. F. Fitzsimons and E. Kashefi, *Unconditionally Verifiable Blind Quantum Computation*, *Phys. Rev. A* **96**, 012303 (2017).
- [18] D. Aharonov, M. Ben-Or, and E. Eban, *Interactive Proofs for Quantum Computations*, .
- [19] T. Morimae, *Verification for Measurement-Only Blind Quantum Computing*, *Phys. Rev. A* **89**, 060302 (2014).
- [20] M. Hayashi and T. Morimae, *Verifiable Measurement-Only Blind Quantum Computing with Stabilizer Testing*, *Phys. Rev. Lett.* **115**, 220502 (2015).
- [21] B. W. Reichardt, F. Unger, and U. Vazirani, *Classical Command of Quantum Systems*, *Nature (London)* **496**, 456 (2013).
- [22] M. McKague, *Interactive Proofs for BQP via Self-Tested Graph States*, *Theory Comput.* **12**, 1 (2016).
- [23] J. F. Fitzsimons, M. Hajdušek, and T. Morimae, *Post hoc Verification of Quantum Computation*, *Phys. Rev. Lett.* **120**, 040501 (2018).
- [24] A. Coladangelo, A. B. Grilo, S. Jeffery, and T. Vidick, *Verifier-on-a-Leash: New Schemes for Verifiable Delegated Quantum Computation, with Quasilinear Resources*, *Adv. Crypt.* **11478**, 247 (2019).
- [25] S. Barz, J. F. Fitzsimons, E. Kashefi, and P. Walther, *Experimental Verification of Quantum Computation*, *Nat. Phys.* **9**, 727 (2013).
- [26] U. Mahadev, *Classical Verification of Quantum Computations*, in *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, 2018), pp. 259–267, <https://doi.org/10.1109/FOCS.2018.00033>.
- [27] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, *Cross-Platform Verification of Intermediate Scale Quantum Devices*, *Phys. Rev. Lett.* **124**, 010504 (2020).
- [28] R. Raussendorf and H. J. Briegel, *A One-Way Quantum Computer*, *Phys. Rev. Lett.* **86**, 5188 (2001).
- [29] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest, *Measurement-Based Quantum Computation*, *Nat. Phys.* **5**, 19 (2009).
- [30] J. F. Fitzsimons, *Private Quantum Computation: An Introduction to Blind Quantum Computing and Related Protocols*, *npj Quantum Inf.* **3**, 23 (2017).
- [31] IBM Quantum Experience, <http://www.research.ibm.com/quantum>.
- [32] Rigetti Computing, <https://www.rigetti.com/>.
- [33] C. Greganti, P. Schiansky, I. A. Calafell, L. M. Procopio, L. A. Rozema, and P. Walther, *Tuning Single-Photon Sources for Telecom Multi-Photon Experiments*, *Opt. Express* **26**, 3286 (2018).
- [34] A. Mantri, T. F. Demarie, and J. F. Fitzsimons, *Universality of Quantum Computation with Cluster States and (X,Y)-Plane Measurements*, *Sci. Rep.* **7**, 42861 (2017).
- [35] V. Danos and E. Kashefi, *Determinism in the One-Way Model*, *Phys. Rev. A* **74**, 052310 (2006).
- [36] D. E. Browne, E. Kashefi, M. Mhalla, and S. Perdrix, *Generalized Flow and Determinism in Measurement-Based Quantum Computation*, *New J. Phys.* **9**, 250 (2007).
- [37] A. Mantri, T. F. Demarie, N. C. Menicucci, and J. F. Fitzsimons, *Flow Ambiguity: A Path Towards Classically Driven Blind Quantum Computation*, *Phys. Rev. X* **7**, 031004 (2017).
- [38] M. Mhalla, M. Murao, S. Perdrix, M. Someya, and P. S. Turner, in *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, Vol. 6745 LNCS (2014) pp. 174–187.
- [39] W. Knight and D. M. Bloom, *A Birthday Problem*, *Am. Math. Mon.* **80**, 1141 (1973).
- [40] E. Pednault, J. A. Gunnels, G. Nannicini, L. Horesh, and R. Wisnieff, *Leveraging Secondary Storage to Simulate Deep 54-Qubit Sycamore Circuits*, [arXiv:1910.09534](https://arxiv.org/abs/1910.09534).

- [41] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Characterizing Quantum Supremacy in Near-Term Devices*, *Nat. Phys.* **14**, 595 (2018).
- [42] N. M. Linke, D. Maslov, M. Roetteler, S. Debnath, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, *Experimental Comparison of Two Quantum Computing Architectures*, *Proc Natl. Acad. Sci. U.S.A.* **114**, 3305 (2017).
- [43] S. Barz, E. Kashefi, A. Broadbent, J. Fitzsimons, A. Zeilinger, and P. Walther, *Demonstration of Blind Quantum Computing*, *Science* **335**, 303 (2012).
- [44] B. P. Lanyon, P. Jurcevic, M. Zwerger, C. Hempel, E. A. Martinez, W. Dür, H. J. Briegel, R. Blatt, and C. F. Roos, *Measurement-Based Quantum Computation with Trapped Ions*, *Phys. Rev. Lett.* **111**, 210501 (2013).
- [45] W. McCutcheon, A. McMillan, J. G. Rarity, and M. S. Tame, *Experimental Demonstration of a Measurement-Based Realisation of a Quantum Channel*, *New J. Phys.* **20**, 033019 (2018).
- [46] M. Kawamura, B. Rowland, and J. A. Jones, *Preparing Pseudopure States with Controlled-Transfer Gates*, *Phys. Rev. A* **82**, 032315 (2010).
- [47] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng, *An Algorithmic Benchmark for Quantum Information Processing*, *Nature (London)* **404**, 368 (2000).
- [48] M. D. Bowdrey, J. A. Jones, E. Knill, and R. Laflamme, *Compiling Gate Networks on an Ising Quantum Computer*, *Phys. Rev. A* **72**, 032315 (2005).
- [49] C.-Y. Lu, X.-Q. Zhou, O. Gühne, W.-B. Gao, J. Zhang, Z.-S. Yuan, A. Goebel, T. Yang, and J.-W. Pan, *Experimental Entanglement of Six Photons in Graph States*, *Nat. Phys.* **3**, 91 (2007).
- [50] V. Saggio, A. Dimić, C. Greganti, L. A. Rozema, P. Walther, and B. Dakić, *Experimental Few-Copy Multipartite Entanglement Detection*, *Nat. Phys.* **15**, 935 (2019).
- [51] M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, *Reducing Multi-Photon Rates in Pulsed Down-Conversion by Temporal Multiplexing*, *Opt. Express* **2011**, 22698 (2019).
- [52] C. Greganti, M.-C. Roehsner, S. Barz, M. Waegell, and P. Walther, *Practical Efficient Experimental Characterization of Multiqubit Stabilizer States*, *Phys. Rev. A* **91**, 022325 (2015).
- [53] A. Dimić and B. Dakić, *Single-Copy Entanglement Detection*, *npj Quantum Inf.* **4**, 11 (2018).
- [54] IBM Q Team, 5-qubit backend, “IBM Q 5 Yorktown Backend Specification V1.0.0,” <https://ibm.biz/qiskit-yorktown> (2017).
- [55] J. S. Otterbach *et al.*, *Unsupervised Machine Learning on a Hybrid Quantum Computer*, [arXiv:1712.05771](https://arxiv.org/abs/1712.05771).
- [56] G. Wendin, *Quantum Information Processing with Superconducting Circuits: A Review*, *Rep. Prog. Phys.* **80**, 106001 (2017).
- [57] S. A. Caldwell *et al.*, *Parametrically Activated Entangling Gates Using Transmon Qubits*, *Phys. Rev. Applied* **10**, 034050 (2018).
- [58] N. Didier, E. A. Sete, M. P. da Silva, and C. Rigetti, *Analytical Modeling of Parametrically-Modulated Transmon Qubits*, *Phys. Rev. A* **97**, 022330 (2018).
- [59] A. Sorensen and K. Molmer, *Entanglement and Quantum Computation with Ions in Thermal Motion*, *Phys. Rev. A* **62**, 022311 (2000).
- [60] X.-L. Wang, L.-K. Chen, W. Li, H.-L. Huang, C. Liu, C. Chen, Y.-H. Luo, Z.-E. Su, D. Wu, Z.-D. Li, H. Lu, Y. Hu, X. Jiang, C.-Z. Peng, L. Li, N.-L. Liu, Y.-A. Chen, C.-Y. Lu, and J.-W. Pan, *Experimental Ten-Photon Entanglement*, *Phys. Rev. Lett.* **117**, 210502 (2016).
- [61] D. Gottesman and I. L. Chuang, *Demonstrating the Viability of Universal Quantum Computation Using Teleportation and Single-Qubit Operations*, *Nature (London)* **402**, 390 (1999).
- [62] D. Gottesman, *Stabilizer Codes and Quantum Error Correction*, Ph.D. thesis, California Institute of Technology (1997).
- [63] IBM Q Team, 16-qubit backend, “IBM Q 16 Rueschlikon Backend Specification V1.0.0,” <https://ibm.biz/qiskit-rueschlikon> (2017).