Quantum Information Processing (2020) 19:146
https://doi.org/10.1007/s11128-020-02642-4

Benchmarking quantum processors with a single qubit

Oktay Göktaş[1,2] · Weng Kian Tham[1] · Kent Bonsma-Fisher[1,3] · Aharon Brodutch[1,4]

Received: 12 November 2019 / Accepted: 9 March 2020 / Published online: 18 March 2020

Abstract
The first generation of small noisy quantum processors have recently become available to non-specialists who are not required to understand specifics of the physical platforms and, in particular, the types and sources of noise. As such, it is useful to benchmark the performance of such computers against specific tasks that may be of interest to users, ideally keeping both the circuit depth and width as free parameters. Here, we benchmark the IBM quantum experience using the deterministic quantum computing with 1 qubit (DQC1) algorithm originally proposed by Knill and Laflamme in the context of liquid-state NMR. In the first set of experiments, we use DQC1 as a trace estimation algorithm to benchmark performance with respect to circuit depth. In the second set, we use this trace estimation algorithm to distinguish between knots, a classically difficult task which is known to be complete for DQC1. Our results indicate that the main limiting factor is the depth of the circuit and that both random and systematic errors become an issue when the gate count increases. Surprisingly, we find that at the same gate count wider circuits perform better, probably due to randomization of coherent errors.

Keywords Quantum benchmarking · DQC1 · Jones polynomials · NISQ · IBM

✉ Oktay Göktaş
oktay@agnostiqlabs.com

1 Department of Physics, Centre for Quantum Information and Quantum Control, University of Toronto, 60 St George St, Toronto, ON M5S 1A7, Canada

2 Agnostiq Labs, 100 College St, Toronto, ON M5G 1L5, Canada

3 National Research Council of Canada, 100 Sussex Dr., Ottawa, ON K1A 0R6, Canada

4 The Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON M5S 3G4, Canada

1 Introduction

Small noisy quantum processors can now be implemented in various platforms and architectures including superconducting circuits [1–4], trapped ions [5], optics [6] and NMR [7]. These and other near-future processors are not expected to be universal for quantum computation [8] and need to be benchmarked in tasks that are suitable for noisy processors with little or no error correction. The deterministic quantum computing with one-qubit (DQC1) algorithm, which was originally developed for noisy NMR quantum processors, offers a good way to benchmark these processors. In this work, we benchmark two IBM quantum processors: first using simple DQC1 circuits to calculate the trace of a unitary and then, in a specific task, using DQC1 to distinguish between knots.

The experiments used between 3 and 8 qubits and were initially run on the IBM Q 16 Rüschlikon [9] and later on the IBM Q 14 Melbourne [10]. The first set of experiments (Sect. 4) involved the estimation of the normalized trace of 1 and 3 qubit unitaries. The results allow us to make some general statements about the noise in the circuit, in particular depolarizing noise and systematic (coherent) errors. Somewhat surprisingly, the performance of the 3 qubit algorithms as a function of the number of gates was better than the 1 qubit algorithms, most likely due to the reduction in correlated noise when the gates act on different qubits. In the second set of experiments, we used the DQC1 algorithm to evaluate various Jones polynomials (Sect. 5). The results show that while the evaluated Jones polynomials tend to be far from theoretical values, the errors are consistent for the different circuits. This implies that the processors can be used to distinguish between various knots made by closing a braid of up to 3 strands, as long as the evaluations are run at approximately the same time (i.e., not hours apart) using the same subset of qubits, such that systematic errors in gate operations remain approximately the same from run to run.

2 DQC1

The DQC1 model was originally proposed by Knill and Laflamme [11] in the context of room-temperature, liquid-state NMR quantum computing where the initial (thermal) state ρ_i is very noisy. As a consequence of the noise, the signal-to-noise ratios in the readout are small and the computation is done on an ensemble with ensemble readout, i.e., the result is an estimate of the expectation value of some observable. The model is further restricted by only allowing Pauli measurements on one of the qubits. Knill and Laflamme noted that in an $N+1$-qubit NMR processor, it is possible to prepare an initial state of the type $\rho_i = [\alpha\,|0\rangle\,\langle 0| + (1-\alpha) I_1] \otimes I_N$ (where $I_n = \frac{1}{2^n}\mathbb{I}_n$ is the n qubit maximally mixed state) efficiently.[1] Under the assumption that the evolution is given by a unitary operator V, the expectation of the final readout on the first qubit will be $\mathrm{Tr}(V\rho_i V^\dagger \sigma_k^{(1)}) = \alpha\,\mathrm{Tr}(V\,(|0\rangle\,\langle 0| \otimes I_N)\,V^\dagger \sigma_k^{(1)})$ where $\sigma_k^{(1)}$ is a Pauli operator on the first qubit and $k \in \{x, y, z\}$. Noting that the polarization parameter α is simply used

[1] The initialization procedure works for some fixed $\alpha \ll 1$ that depends on the parameters of the experiment, and does not scale badly with N.

to rescale the expectation value, it is often convenient to assume $\alpha = 1$, as we will do throughout this work. Under this assumption, the first qubit is initially pure, or "clean," while the other qubits are completely mixed. This model is therefore sometimes called the "one clean qubit" model [12].

In the DQC1 model, the classical input describes the unitary operator V which is assumed to have an efficient description, i.e., it can be decomposed into a (polynomial in N) sequence of one and two qubit gates. It is common to further restrict V to a Hadamard operator on the first qubit, followed by a controlled unitary from the first qubit, U, targeting all other qubits, i.e., $V = (|0\rangle\langle 0| \mathbb{I}_N + |1\rangle\langle 1| U) H^{(1)}$. Here, U is an N qubit unitary with an efficient classical description, and $H^{(1)}$ is a Hadamard operator on the first qubit. (This is sometimes called cDQC1 [13].) In this restricted model, the state of the first qubit at the end of the computation (before readout) is given by

$$\rho_f^{(1)} = \frac{1}{2}\left(I_1 + \frac{\mathrm{Tr}U}{2^N}|1\rangle\langle 0| + \frac{\mathrm{Tr}U^\dagger}{2^N}|0\rangle\langle 1|\right) \tag{1}$$

so that $\frac{1}{2^N}\mathrm{Tr}\,U = \langle\rho_f^{(1)}(\sigma_x + i\sigma_y)\rangle$. That is, the model can be used to estimate the normalized trace of the unitary U. A number of results suggest that the trace estimation algorithm cannot be simulated efficiently by a classical computer, with some recent results including the use of DQC1 for parity learning [14], a sampling version [15] which follows the standard definition above, but allows single shot readout and a method for verifying the computation [16].

Shor and Jordan [12] used the DQC1 model to define a computational complexity class. They then showed that the trace estimation algorithm is computationally equivalent to the full DQC1 model and furthermore showed that adding a small (at most logarithmic in N) number of pure qubits does not change the computational power of the model. They also showed that the estimation of the Jones polynomial for the trace closure of a braid at the fifth root of unity (a problem in knot theory, see Sect. 5) is DQC1 complete.

2.1 Noise in DQC1

The DQC1 model is designed to handle noisy initial states, but, to the best of our knowledge, its performance under noisy dynamics has not been analyzed. An N qubit unitary V can generally be decomposed into a sequence of fundamental unitaries $\{W_k\}$ such that $V = \prod_k W_k$. Ideally, these fundamental unitaries correspond to gates that are physically implementable on the processor. But in practice, the gates are imperfect and errors that are often difficult to characterize degrade the computation [17,18].

One fairly simple model is to assume depolarizing noise, where each gate is a probabilistic mixture of the desired unitary W_k and a completely depolarizing channel. The ideal transformation $\rho \to W_k \rho W_k^\dagger$ of the state is replaced by $\rho \to \alpha_k W_k \rho W_k^\dagger + (1-\alpha_k)\mathbb{I}$ where α_k is the purity of the channel. All subsequent unitary operations and depolarizing channels leave the identity unchanged, so the state after the full sequence will be $\alpha V \rho_i V^\dagger + (1-\alpha)\mathbb{I}$ with $\alpha = \prod_k \alpha_k$.

The fact that the purity falls exponentially with the number of gates does not bode well for the computation. Standard quantum error correction methods rely on a supply of pure qubits so they are not suitable for DQC1. However, for our purposes, focusing on small or intermediate size processors, the issue of exponential noise might not be debilitating. Moreover, in a real situation, it is also possible to deviate slightly from the model and use some clean physical qubits and single shot measurements to combat errors.

The relatively simple behavior of the DQC1 algorithm in the presence of depolarizing noise makes it a good tool for benchmarking against a depolarizing noise model. In our results below, we used the R^2 of a fit to the depolarizing noise model as a benchmark. The behavior of the circuits as a function of the number of gates provided evidence that the most significant source of error was a systematic error in the CNOT gates.

3 Implementation of the algorithm

We implemented DQC1 on IBM superconducting qubit quantum processing units (QPU) via a Web-based application programming interface (API). The code is available online [19], and a technical descriptions of the processors can be found in Ref. [2]. All results described in the present work are limited to data obtained via the Web API, and not direct physical access to IBM hardware. Basic tests to benchmark DQC1 performance on gate-based machines, described in Sect. 4, were executed on the 16-qubit "Rüschlikon" QPU. Application of DQC1 to a useful task (evaluation of Jones polynomials), described in Sect. 5, was executed on the 14-qubit "Melbourne" QPU.

A single qubit (shown in red in Fig. 1) was designated the "clean" qubit, whereas another disjoint subset of qubits (shown in blue in Fig. 1) was chosen to be "noisy." A gate-based QPU, however, is usually designed to operate with pure states under unitary evolution as much as possible. To prepare these "noisy" qubits in the $(\mathbb{I}/2)^{\otimes N}$ state, we used two techniques. In the first set of experiments, we first entangled each qubit with an adjacent qubit to produce the (pure) Bell state $\left|\Phi^+\right\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ and then ignored (or traced away) that adjacent qubit. This approach introduces a 2X overhead in the number of qubits required for state preparation of these "noisy" qubits. For the estimation of the Jones polynomial, we used random bit flips on the "noisy" qubit for half the experiments and averaged over the results. This method can be modified for a multiple qubit scenario by randomly flipping all "noisy" qubits and averaging over the results. Note that preparing the input state with random bit-flips is operationally identical to tracing out one qubit of $|\Phi^+\rangle$, per above, over a finite number of shots.

4 Trace estimation on the quantum processor

We implemented the $N = 1$ version of the trace estimation algorithm (Fig. 2) with

$$U^{(l)}_{N=1}(\theta) = U_1(\theta)(U_1(\theta)^\dagger U_1(\theta))^{l-1},$$

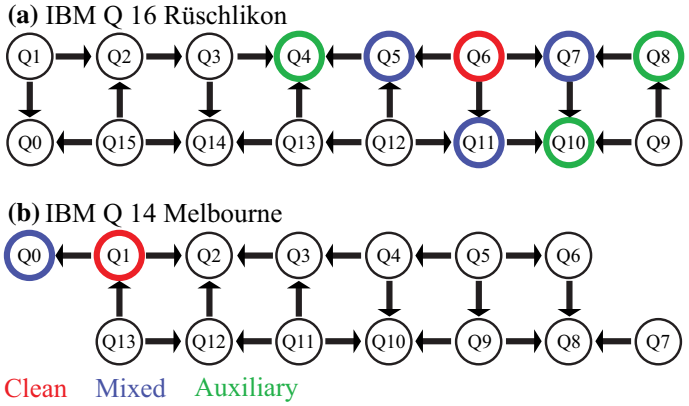

Fig. 1 **a** Qubits used on the IBM Q 16 Rüschlikon chip. Qubit 6 (red) was the clean qubit. Qubits 5, 7, and 11 (blue) were the mixed qubits. Mixed states were generated by first performing an entangling operation with auxiliary qubits 4, 8, and 10 (green), respectively. Black arrows show the control-target relationship for coupled qubits. **b** The IBM Q 14 Melbourne was used for the knot experiments. Qubits 1 and 0 were used for the pure and mixed states, respectively, in the first set of knot experiments. Subsequent experiments used all 18 pairs of connections between qubits (Color figure online)

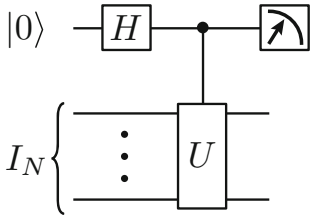

Fig. 2 A DQC1 circuit for estimating the trace of a unitary U with the first qubit initialized in a pure state. The same computation is run a large number of times with the final measurement cycled between σ_x and σ_y to get an estimate of the real and imaginary parts, respectively (Color figure online)

where $l \geq 1$ is the number of repetitions and $U_1(\theta) = e^{-i\theta/2}\,|0\rangle\langle 0| + e^{i\theta/2}\,|1\rangle\langle 1|$ (see inset in top row of Fig. 3). We also implemented the $N = 3$ version, replacing $U_1^{(l)}(\theta)$ with $U_3^{(l)}(\theta) = U_1^{(l)}(\theta)^{\otimes 3}$ so

$$U_{N=3}^{(l)}(\theta) = U_3(\theta)(U_3(\theta)^{\dagger}U_3(\theta))^{l-1}$$

(see inset in bottom row of Fig. 3).

The circuit was chosen to maximize contrast with respect to θ (i.e., $\mathrm{Tr}[U_1^{(l)}(0)] = 1$ and $\mathrm{Tr}[U_1^{(l)}(\pi)] = -1$) for *any* value of l. Increasing l merely introduces repetition of a gate sequence that should, logically, be equivalent to the identity. In practice, however, gate errors and noise means increasing l yields noisier outputs. Results for a final measurement of σ_x, σ_y and σ_z are shown in Fig. 3.

As expected, the results deviate further from the ideal as we go to higher gate counts. The reduction of the absolute values in the σ_x plots can be attributed to depolarizing noise; however, the fact that the shape changes in all three plots (and in particular the

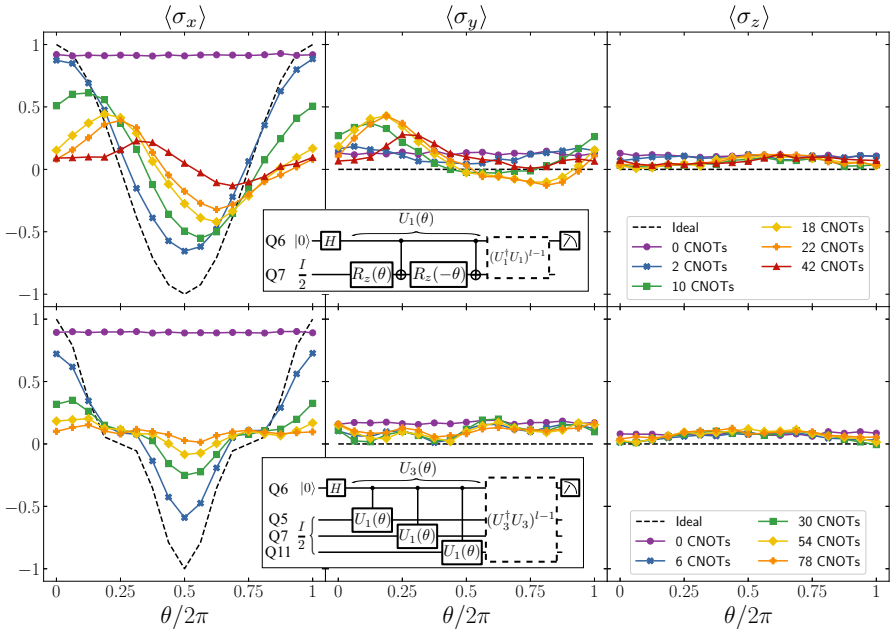

Fig. 3 Expectation values $\langle\sigma_x\rangle$, $\langle\sigma_y\rangle$ and $\langle\sigma_z\rangle$ when applying the c$-U(\theta)$ from the control qubit to one mixed qubit (Q7, top row), and three mixed qubits (Q5, Q7, Q11, bottom row). Colored curves represent applying the c$-U(\theta)$ multiple times. (The data for 0 CNOTs mean no operation applied at all.) As expected, the deviations from the theoretical curve get larger as the number of gates increases; however, the gate errors are not random and the contribution of non-depolarizing noise is significant. This is especially apparent in the σ_y plot which is expected to be near 0 at all times. Note that for large numbers of CNOT gates the deviation from 0 is far above the statistical uncertainty which is upper bounded by $1/2^{7.5} < 0.01$. The coherent errors are partially suppressed in the $1 + 3$ qubit circuit, probably due to coherent errors being averaged out over the different qubits (Color figure online)

deviations from 0 in σ_y), indicates a coherent error, probably as a result of a systematic error in the CNOT gates. For a quantitative indication of the coherent errors (more precisely the deviation from depolarizing noise), we defined the visibility

$$\text{Vis} = \max_{\theta} \langle\sigma_x\rangle \,.$$

This function decays exponentially with the number of gates when imperfections are due to depolarizing noise. A plot of visibility as a function of the number of gates (Fig. 4) shows that this is clearly not the case for qubit 11 (paired with 5) where there is a spike in visibility around 20 CNOT gates. The other couplings show the expected qualitative behavior, but a fit to a depolarizing noise model shows some deviations (see Table 1).

Generally, it is possible to convert biased noise channels into a depolarizing channels by adding some randomness and averaging. This is apparent when comparing the $1+1$ and $1 + 3$ qubit results. In Table 1, we see that fit for the $1 + 3$ qubit result is better than each of the individual results. This is most likely a result of averaging over

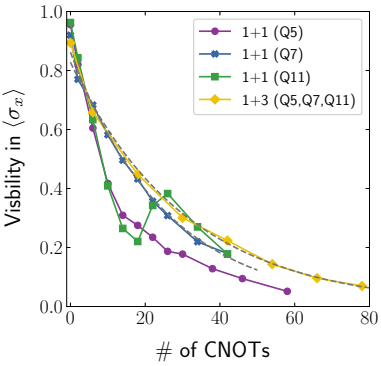

Fig. 4 Visibility in $\langle\sigma_x\rangle$ as a function of circuit depth. Decay in visibility is indicative of noise in the system. Assuming purely depolarizing noise decay is exponential. Gray dashed lines are fits of $f(x) = ae^{-x/\tau}$ to the $1+1$ (Q7) and $1+3$ (Q5, Q7, Q11) data

Table 1 Assuming only depolarizing noise, we fit a decay model $f(x) = ae^{-x/\tau}$ to the data in Fig. 4 (linear fit to the logarithm of the data)

Qubits	τ	$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}$
Q5	24.50	0.933
Q7	25.81	0.995
Q11	28.91	0.695
Q5, Q7, Q11	30.75	0.997

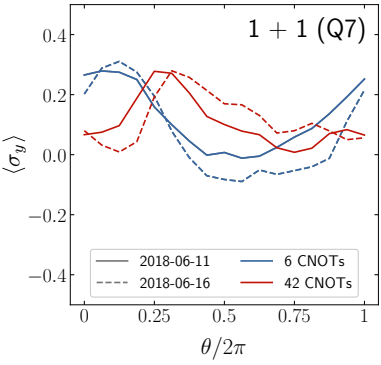

Fig. 5 σ_y expectation value of $1+1$ data (qubit 7) taken on different dates, 5 days apart. The variation between results at different times is far greater than the statistical uncertainty which is upper bounded by $1/2^{7.5} < 0.01$ and an indication that systematic calibration errors change significantly over time

different coherent errors for the 3 different pairs of interacting qubits. We note that in general this is not an indication of better performance overall. Though the exponential depolarizing rate appears slower in the $1+3$ qubit case, the circuit will have to be 3 times longer to perform a similar task.

Running the experiment at different times produced different results; in particular, the systematic (coherent) errors were not consistent over long periods of time. Results for the expectation value of σ_y taken on the same pair of qubits 5 days apart are plotted in Fig. 5. Since the theoretical expectation value should be constant ($\langle\sigma_y\rangle = 0$), the plots are a good indication of coherent errors which, even at a circuit depth of 6 CNOT gates, produce a visibly different plot.

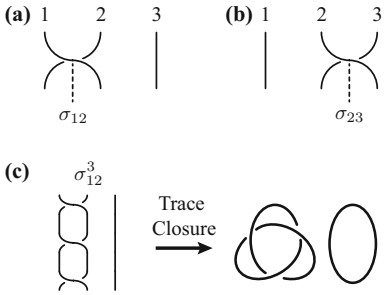

Fig. 6 Visualization of σ_{12} (**a**) and σ_{23} (**b**) crossing operations on three strands. **c** The braid word σ_{12}^3, three consecutive σ_{12} crossings. By taking the trace closure, connecting the bottom of a strand to its top, the first two strands form the trefoil knot, while the third strand forms the unknot. The braid closures of **a** and **b** are topologically equivalent, but the different braid words lead to different circuit implementations in the experiment

5 Distinguishing knots with Jones polynomials

The task of identifying whether two knots (smooth closed curves in $\mathbb{R}^3$) are topologically equivalent has implications beyond mathematics, reaching into statistical mechanics, quantum field theory and quantum gravity [20,21]. Knots can be faithfully represented in two-dimensional pictures, and so the task can be recast as determining whether two pictures of knots can be made equal using transformations called Reidemeister moves. This task is computationally intensive, and even in the simplest case, identifying a knot as the *unknot*, there is no known efficient solution [22]. Here, we consider a particular type of knot, the trace closure of a 3-strand braid, which can be represented by drawing a two-dimensional braid and closing each end at the bottom with the associated strand at the top (right-most to right-most etc., see Fig. 6c). Another type of knot, the plat closure, is constructed by connecting adjacent strand ends at the top and at the bottom.

The Jones polynomial is a complex function invariant for oriented knots [23,24], which allows one to distinguish one knot from another. However, constructing the polynomial for a given two-dimensional picture of a knot is not trivial. The number of terms in the polynomial (before simplification) scales exponentially with the number of strand crossings. Evaluating the Jones polynomial at a single point would give sufficient information to tell if two knots are different. Equivalent knots must have the same Jones polynomial value at any given point, whereas knots that are not the same might not. Approximating the value of the Jones polynomial at $e^{2\pi i/5}$ is a particularly interesting task for quantum computers. If one takes the plat closure of a braid, rather than the trace closure, the task is known to be BQP-complete [25].

Shor and Jordan [12] showed that approximating the polynomial at $e^{2\pi i/5}$ for the *trace* closure of a braid is a complete task for DQC1. Passante et al. [26] demonstrated this task in a four-qubit liquid-state NMR processor, studying knots with four strands and multiple crossings. Here, we use DQC1 to approximate the Jones polynomial for knots of three strands for the purpose of benchmarking the IBM Q 14 Melbourne quantum processor. We consider knots which are constructed by taking the trace closure

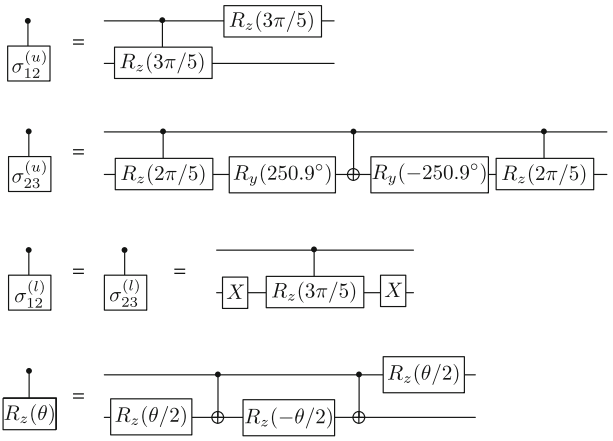

Fig. 7 Circuits for the controlled $\sigma_{12}^{(u)}, \sigma_{23}^{(u)}, \sigma_{12}^{(l)}$ and $\sigma_{23}^{(l)}$ unitaries. Here, $R_y(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, and $R_z(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$

of braid words up to 9 crossings (see Fig. 6). We study the same knots constructed by multiple iterations of either the σ_{12} crossing (first strand over the second), or the σ_{23} crossing (second over the third). Since approximating the values of Jones polynomials with a noisy machine is difficult, we are content with the ability to classify knots as different when they are indeed different. As in the previous sections, we do not perform any type of error mitigation during the computation or in post-processing, apart from simplifying the circuit to require fewer gates.

Following the treatment in Ref. [26], the unitary used in the DQC1 protocol is related to the braid through the Fibonacci representation. For a three-strand braid, the σ_{12} and σ_{23} unitaries are given by

$$\sigma_{12} = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \sigma_{23} = \begin{pmatrix} e & d & 0 & 0 \\ d & c & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2}$$

where $a = e^{3\pi i/5}$, $b = e^{-4\pi i/5}$, $c = \frac{b}{\phi^2} + \frac{a}{\phi}$, $d = \frac{b-a}{\phi^{3/2}}$, $e = \frac{b}{\phi} + \frac{a}{\phi^2}$, and $\phi = (1+\sqrt{5})/2$. For braids with n strands, we require $m \times m$ sized unitaries where m is the n^{th} number in the Fibonacci sequence. For 3 strands, the unitaries map between 3 states, requiring 2 qubits (Fig. 7). This results in an unused portion of the Hilbert space—the $|11\rangle$ state is not used for the approximation and adds a constant term to the trace which is straightforwardly dealt with.

To implement DQC1, we must construct controlled versions of the braid word unitaries. However, a single controlled-σ_{23} operation requires approximately 50 CNOT gates, meaning any braid word would be prohibitively long on current quantum processors. Indeed, Sect. 4 shows that $\langle\sigma_x\rangle$ and $\langle\sigma_y\rangle$ on the clean qubit would decay

substantially after just one controlled-σ_{23} gate. To simplify the problem, we use the fact that both σ_{12} and σ_{23} are block diagonal, and so any braid word will also be. We perform a controlled version of each block of a braid word in separate experiments, measure their traces via the clean qubit and combine them afterward. The controlled implementations of the blocks require substantially fewer CNOT gates: $\sigma_{12}^{(u)}$, $\sigma_{23}^{(u)}$ and $\sigma_{12}^{(l)} = \sigma_{23}^{(l)}$, where u (l) refers to the upper (lower) block of the unitary, can be performed with 2, 5 and 2 CNOT gates, respectively. The braid word for the trefoil knot outlined in Fig. 6, σ_{12}^3, requires 6 CNOT gates for each of the upper and lower blocks. Performing the braid word σ_{23}^3, which gives the same knot upon taking the trace closure, requires 15 CNOT gates for the upper block and 6 for the lower.

Having used the DQC1 protocol to measure the trace of each block of a braid word unitary, we then estimate the value of Jones polynomial at the fifth root of unity for each knot. To do this, we combine the measurements on each block of the unitary, e.g., $U = \sigma_{12}^3$, to find the weighted trace of the braid word. First, we subtract off the contribution to the trace of the lower block, $U^{(l)}$, from the $|11\rangle$ state. We then add the traces of the two blocks together while weighting the upper block by a factor of ϕ [26,27]:

$$\begin{aligned} \mathrm{WTr}\, U &= \phi \times \mathrm{Tr}\, U^{(u)} + \mathrm{Tr}\, U^{(l)} - 1 \\ &= \phi \times (\langle\sigma_x\rangle^{(u)} + i \left\langle\sigma_y\right\rangle^{(u)}) \\ &\quad + \langle\sigma_x\rangle^{(l)} + i \left\langle\sigma_y\right\rangle^{(l)} - 1. \end{aligned} \tag{3}$$

We then calculate the Jones polynomial value as

$$V_U(t = e^{2\pi i/5}) = (-(e^{2\pi i/5})^4)^{3w} \times \frac{1}{\phi} \mathrm{WTr}\, U, \tag{4}$$

where w is the writhe of the knot, defined as the difference between left-over-right crossings (σ_{12} and σ_{23}) and right-over-left crossings ($\sigma_{12}^\dagger$ and $\sigma_{23}^\dagger$). In this work, we only consider knots with $w > 0$.

In Fig. 8, we plot one set of results of the Jones polynomial, $V_U(e^{2\pi i/5})$, estimation for knots with 0 to 9 crossings, constructed solely by either σ_{12} or σ_{23} crossings. We find that as the numbers of crossings are increased, the estimated polynomials quickly deviate from theoretical values (see Fig. 9). This is to be expected from the studies in previous sections. As the circuit depth increases, the measured $\langle\sigma_x\rangle$ and $\left\langle\sigma_y\right\rangle$ on the clean qubit decay exponentially. This in turn causes the Jones polynomial values to tend to the origin on the complex plane with increasing circuit depths.

Though the deviation from the theoretical value is not ideal performance, the principle behind estimating the Jones polynomial is to distinguish between knots. In this spirit, we note that the distance between the two implementations of each knot (using either σ_{12} or σ_{23}) remains relatively low compared to the distance from their theoretical value. Differences between the two experimental implementations of the same knot are likely driven by the significantly higher circuit depth for each σ_{23} unitary (5 CNOTs vs. 2 CNOTs). Importantly, if we compare different knots that have the same circuit depth—e.g., σ_{12}^5 and σ_{23}^2 each use 10 CNOT gates for their upper blocks

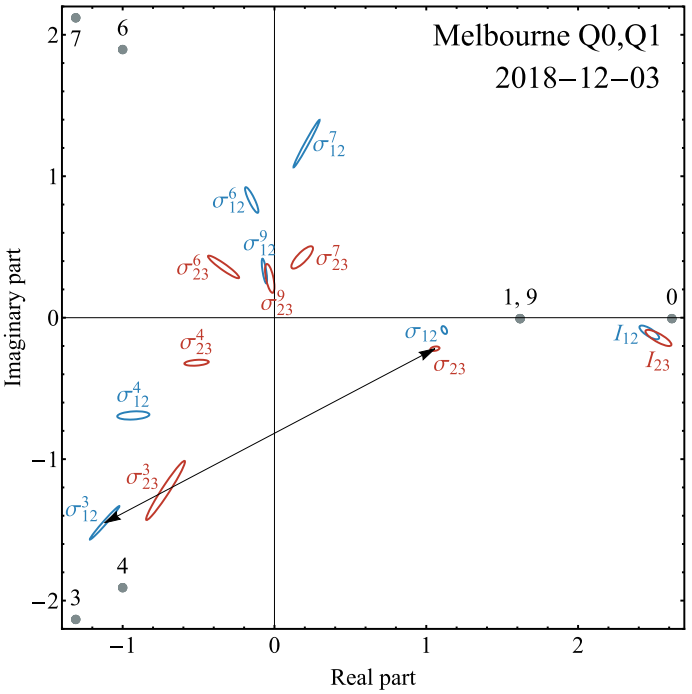

Fig. 8 Results of Jones polynomial evaluations on qubits Q0 and Q1 of the IBM Q 14 Melbourne. Knots are given by the trace closure of σ_{12}^k (blue points), and σ_{23}^k (red points), k runs from 0 to 9 crossings. Gray points mark the theoretical Jones polynomial values for each knot, with numbers representing the crossings in the associated braid word; for each k, the knots represented by σ_{12}^k and σ_{23}^k are equivalent. Ellipses represent the standard deviation of 12 trials (2^{12} shots each) for each knot. The distance between ellipses measures, in some sense, how well two knots can be distinguished on the IBM QPU. The black arrow between σ_{12}^3 and σ_{23} marks the distance between the two Jones polynomial estimates with a similar gate count (see also Table 2). For clarity, results from knots with 2, 5 and 8 crossings (which are closer to the center) are not plotted. Note that while the results generally get closer to the center as the gate count increases (a signature of depolarizing noise), results for 7 and 9 crossings appear in the wrong quadrant in both representations (a signature of systematic errors) (Color figure online)

and represent different knots—we see that they are largely distinguishable from one another when the gate count is low.

At higher gate counts, the values are not only closer to the origin (as expected), but also behave qualitatively different from the theoretical results; for example, in Fig. 8, the real part of the braid with 7 crossings should be more negative than that of 6 crossings, but in both implementations (σ_{12} and σ_{23}), the knot with 7 crossings is positive, while the knot with 6 is negative. Moreover, these types of errors, while fairly consistent on a single run of the experiment, appear to be very different when the experiment is repeated later and/or on different qubits. In Fig. 10a, we show the results for 0 and 3 crossings (σ_{12}) taken on all 18 different pairs of qubits. The results indicate that both the mean and the spread depend on the choice of qubits. In Fig. 10b, we compare the results of the braids with 0 and 3 crossings (σ_{12}) at different times, where we chose the qubits pairs Q5, Q6 and Q4, Q5 for best performance. Even at

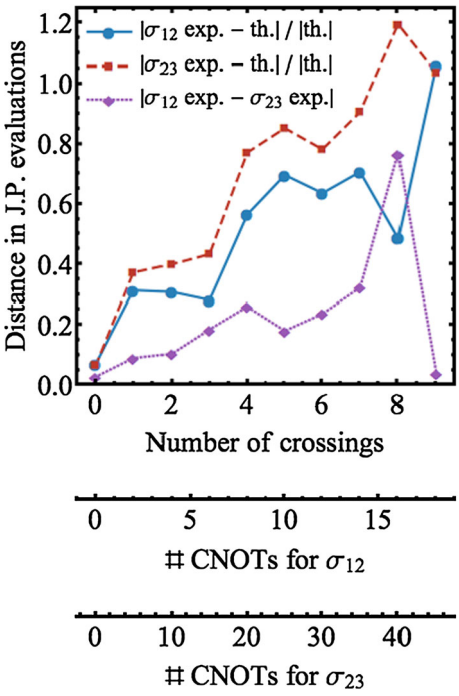

Fig. 9 Distance between the evaluated Jones polynomial from σ_{12} unitaries, σ_{23} unitaries and theory for knots with varying numbers of crossings. Distances are normalized by the theoretical values of the polynomials to account for values close to the origin. σ_{23} estimates deviate from the theoretical values for fewer crossings, likely due to the fact that a single σ_{12} unitary requires 2 CNOTs, while a σ_{23} requires 5 CNOTs. The distance between polynomial estimates for σ_{12} and σ_{23} implementations increases with number of crossings (with the exception of the final point where they both tend the origin), making two versions of the same knot distinguishable from one another

Table 2 Comparison of Jones polynomial estimates for implementations of different knots using the similar circuit depths

CNOTs in σ_{23}^k	Unitaries	J.P. dist. (Exp.)	J.P. dist. (Theory)
5	$\lvert\sigma_{23} - \sigma_{12}^2\rvert$	1.42 ± 0.05	2.15
5	$\lvert\sigma_{23} - \sigma_{12}^3\rvert$	2.50 ± 0.08	3.62
10	$\lvert\sigma_{23}^2 - \sigma_{12}^5\rvert$	0.66 ± 0.05	1
15	$\lvert\sigma_{23}^3 - \sigma_{12}^7\rvert$	2.6 ± 0.2	4.25
15	$\lvert\sigma_{23}^3 - \sigma_{12}^8\rvert$	2.0 ± 0.2	3.24

A large distance, relative to the error, indicates that two knots can be distinguished using the IBM QPU

relatively low gate counts, we observe deviations from one run of the experiment to the next. One consequence of these results is that there is no simple way to correct for errors in post-processing. Such corrections would have been possible if the dominant source of error was depolarization, in which case we could multiply the results by a factor that depends on the number of gates.

6 DQC1 as a benchmark

Benchmarking protocols have usually been designed with the experimental physicist in mind, often to quantify performance in terms of noise per gate, and usually with

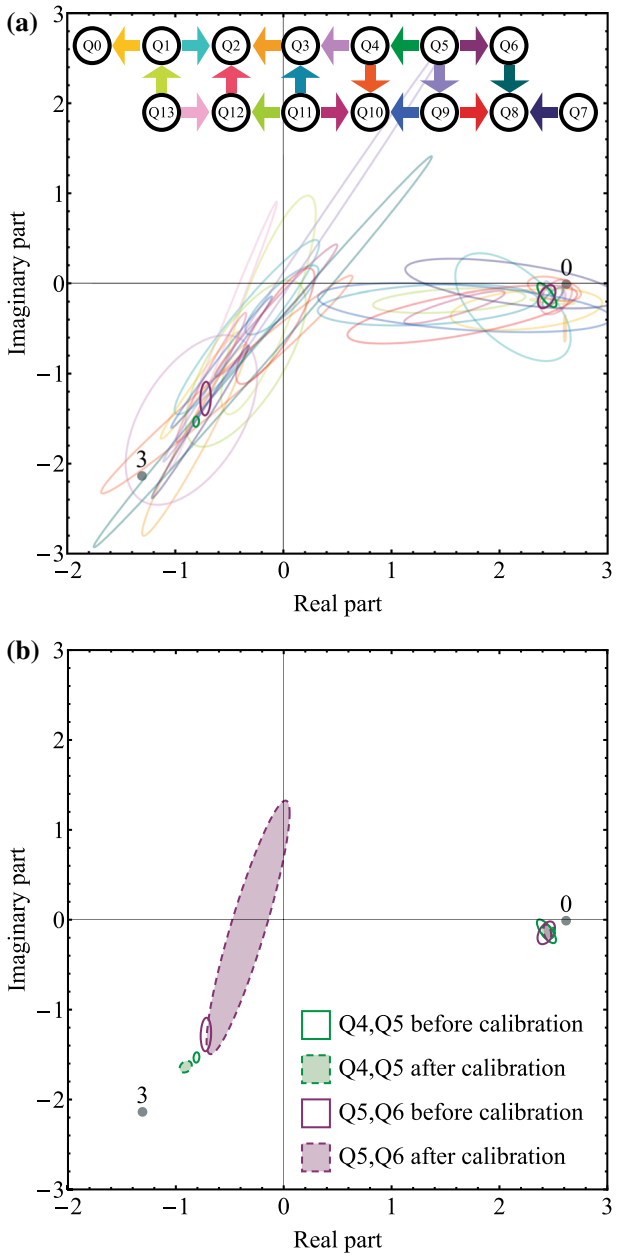

Fig. 10 **a** Jones polynomial results of the σ_{12}^{0} (right) σ_{12}^{3} (left) unitaries using different qubit combinations on the Melbourne processor. Shown in bold are the Q4–Q5 (dark green) and Q5–Q6 (dark purple) pairs which gave the best performance. Computations were performed on February 22, 2019 (before calibration). **b** Comparing qubit pairs Q4–Q5 (dark green) and Q5–Q6 (dark purple) before (solid) and after (dashed, filled) the February 25, 2019 re-calibration. The second set of computations was performed on February 26, 2019 (after calibration). A routine calibration process was performed on the machine by the IBM team between these two dates. For both **a** and **b** ellipses represent the standard deviation of estimating Jones polynomials over 10 trials (1024 shots each) (Color figure online)

an outlook toward fault tolerance [28–33]. These protocols can be used to assess performance and provide feedback for the development of hardware as well as guides for the development of software. They are, however, computationally intensive, require a high level of expertise to understand and are often difficult to put into the context of specific tasks. The trace estimation algorithm in Sect. 4 provides some of the same features at a qualitative level, using a fairly simple protocol. While the characterization of errors is not precise, it is sufficient for making an educated guess as to the types of errors and their relative significance. In the case of the results presented in Sect. 4 (see Figs. 3, 4, 5), we see an indication of coherent noise which is not consistent over time. In principle, had the dominant source of error been more consistent, the results would have provided some means to mitigate the errors, either by modifying the circuit or in post-processing. This would have been particularly useful if the main issue was depolarizing noise, which could be mitigated by multiplying the expectation values by a constant. (Note that this is not a scalable method.) We note that the analysis of the results is particularly easy in the DQC1 circuit since the output is a set of two numbers related to measurements on a single qubit, regardless of the input size. Moreover, the inputs on all except the first qubit are in a mixed state and can be treated as random (a property which is essential for many randomized benchmarking protocols).

A different approach to benchmarking is based on computational tasks of specific interest. In the early days of quantum computing, the standard task was factoring [34], but recently, developments have shifted to algorithms better suited to noisy intermediate-scale quantum (NISQ) devices [35–38]. Benchmarks built around these algorithms can be used to compare between vastly different quantum platforms as well as between quantum and classical platforms. Additionally, these task-specific protocols have the advantage that the analysis of the results requires minimal understanding of the underlying physics. However, the results tend to be very specific to the task at hand, and little can be learned about performance more generally [39–41]. The protocol for distinguishing between knots (or estimating the Jones polynomial) in Sect. 5 can be used for this type of benchmarking with the advantage that, unlike many other protocols, it can be linked to a problem for which a quantum computer is expected to exponentially outperform classical computers.

Finally, a large effort is currently being directed at proving quantum supremacy [1], usually by outperforming classical computers on a problem which is expected to be hard (usually sampling problems). In the context of DQC1, this can be approached in two ways: First, estimation of some Jones polynomials (Sect. 5) at specific points is expected to be a hard problem for classical computers and has historically been one of the interesting problems for quantum computers. Second, the DQC1 protocol can be modified slightly into a sampling problem (by allowing single qubit outcomes) which is known to be hard under the same type of assumptions used for other quantum supremacy protocols [42].

7 Discussion

As quantum computers become more available to non-specialists, there is need for tools that can be used and understood by users who are not interested in the inner

workings of the machine. The benchmarking experiments we used in this work were designed with the end user of a small noisy quantum processor in mind. Using two protocols based on the DQC1 trace estimation algorithm, we benchmarked IBM 14- and 16-qubit processors and showed how performance degrades with circuit depth, in one case for the task of distinguishing knots.

In the first set of experiments, we looked at how visibility drops as a function of circuit depth and showed that coherent errors can be particularly harmful. For example, we noted a spike in visibility (Fig. 4) for the DQC1 experiment with Qubit 11 on the "Rüschlikon". We also observed an undesired buildup of an imaginary component (coherence in the σ_y axis) as the gate count increased. This was evident as early as 10 CNOT gates in a 2-qubit experiment (Fig. 3). Surprisingly, we noted that the performance is not reduced (and was even enhanced on average) in a 4-qubit experiment compared to the 2-qubit experiments when results with a similar CNOT gate count were compared (Figs. 3 and 4). Finally, we saw that the results were not consistent when the same experiment was performed at different times (Fig. 5). However, the qualitative results regarding performance at different gate counts remained the same.

In the second set of experiments, we used 2 qubits on the IBM "Melbourne" to estimate the Jones polynomial at the fifth root of unity for various knots. While the results deviated from theory (Fig. 9), it was possible to compare knots at low gate counts (Fig. 8). However, at higher gate counts, both depolarizing and systematic errors start to dominate the results. While depolarizing noise can be countered by repeating the experiment more times and normalizing,[2] the systematic errors (which are not constant in time) are difficult to deal with even in a small circuit. We note that these errors prevented us from outperforming the 4-qubit liquid-state NMR experiment [26].[3]

The relatively simple benchmarking procedure leads us to the conclusion that, at least from the end user's perspective, a major issue with current small noisy quantum computers is the constantly changing environment that leads to frequently changing systematic errors. While it is clear that this is a major engineering challenge for superconducting architectures due to the sensitivity to environmental conditions, it is worthwhile considering alternative architectures that may be more stable. A different approach might be a method to reduce systematic errors on the software side, for example, by using wider and shallower circuits or to turn these into statistical errors by using various randomization techniques.

Acknowledgements We acknowledge use of the IBM Q for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Q team. This work was partially supported by the CIFAR "Quantum Information Science" program and an NSERC grant "Experimental Quantum Information, Quantum Measurement, and Quantum Foundations With Entangled Photons and Ultracold Atoms" via Aephraim Steinberg's research group. KBF acknowledges the NSERC PDF program for funding.

[2] Repeating experiments is generally not a scalable technique.

[3] However, this is not an apples-to-apples comparison since (without fault tolerance) a 4-qubit machine is expected to outperform a 14-qubit machine in a $\leq$ 4-qubit experiment.

References

1. Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J.C., Barends, R., Biswas, R., Boixo, S., Brandao, F.G.S.L., Buell, D.A., Burkett, B., Chen, Y., Chen, Z., Chiaro, B., Collins, R., Courtney, W., Dunsworth, A., Farhi, E., Foxen, B., Fowler, A., Gidney, C., Giustina, M., Graff, R., Guerin, K., Habegger, S., Harrigan, M.P., Hartmann, M.J., Ho, A., Hoffmann, M., Huang, T., Humble, T.S., Isakov, S.V., Jeffrey, E., Jiang, Z., Kafri, D., Kechedzhi, K., Kelly, J., Klimov, P.V., Knysh, S., Korotkov, A., Kostritsa, F., Landhuis, D., Lindmark, M., Lucero, E., Lyakh, D., Mandrà, S., McClean, J.R., McEwen, M., Megrant, A., Mi, X., Michielsen, K., Mohseni, M., Mutus, J., Naaman, O., Neeley, M., Neill, C., Niu, M.Y., Ostby, E., Petukhov, A., Platt, J.C., Quintana, C., Rieffel, E.G., Roushan, P., Rubin, N.C., Sank, D., Satzinger, K.J., Smelyanskiy, V., Sung, K.J., Trevithick, M.D., Vainsencher, A., Villalonga, B., White, T., Yao, Z.J., Yeh, P., Zalcman, A., Neven, H., Martinis, J.M.: Quantum supremacy using a programmable superconducting processor. Nature **574**, 505 (2019)
2. IBM Quantum Experience. https://www.research.ibm.com/ibm-q
3. Takita, M., Cross, A.W., Córcoles, A.D., Chow, J.M., Gambetta, J.M.: Experimental demonstration of fault-tolerant state preparation with superconducting qubits. Phys. Rev. Lett. **119**, 180501 (2017)
4. Pokharel, B., Anand, N., Fortman, B., Lidar, D.: Demonstration of fidelity improvement using dynamical decoupling with superconducting qubits. Phys. Rev. Lett. **121**, 220502 (2018)
5. Figgatt, C., Maslov, D., Landsman, K.A., Linke, N.M., Debnath, S., Monroe, C.: Complete 3-qubit grover search on a programmable quantum computer. Nat. Commun. **8**, 1918 (2017)
6. Rudolph, T.: Why I am optimistic about the silicon-photonic route to quantum computing (2016). arXiv:1607.08535
7. Lu, D., Li, K., Li, J., Katiyar, H., Park, A.J., Feng, G., Xin, T., Li, H., Long, G., Brodutch, A., Baugh, J., Zeng, B., Laflamme, R.: Enhancing quantum control by bootstrapping a quantum processor of 12 qubits. npj Quantum Inf. **3**, 45 (2017)
8. Preskill, J.: Quantum computing in the NISQ era and beyond. Quantum **2**, 79 (2018)
9. 16-Qubit Backend: IBM Q team, IBM Q 16 Rüschlikon backend specification V1.1.0 (2018). https://github.com/Qiskit/ibmq-device-information/blob/master/backends/rueschlikon/. Accessed 12 Nov 2019
10. 14-Qubit Backend: IBM Q team, IBM Q 14 Melbourne backend specification V1.1.0 (2018). https://github.com/Qiskit/ibmq-device-information/tree/master/backends/melbourne/. Accessed 12 Nov 2019
11. Knill, E., Laflamme, R.: Power of one bit of quantum information. Phys. Rev. Lett. **81**, 5672 (1998)
12. Shor, P.W., Jordan, S.P.: Estimating Jones polynomials is a complete problem for one clean qubit. Quantum Inf. Comput. **8**, 681 (2007)
13. Boyer, M., Brodutch, A., Mor, T.: Entanglement and deterministic quantum computing with one qubit. Phys. Rev. A **95**, 022330 (2017)
14. Park, D.K., Rhee, J.K.K., Lee, S.: Noise-tolerant parity learning with one quantum bit. Phys. Rev. A **97**, 032327 (2018)
15. Morimae, T., Fujii, K., Nishimura, H.: Power of one non-clean qubit. Phys. Rev. A **95**, 042336 (2017)
16. Kapourniotis, T., Kashefi, E., Datta, A.: In: 9th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2014), Leibniz International Proceedings in Informatics (LIPIcs), vol. 27, ed. by S.T. Flammia, A.W. Harrow (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2014), Leibniz International Proceedings in Informatics (LIPIcs), vol. 27, pp. 176–204. https://doi.org/10.4230/LIPIcs.TQC.2014.176. http://drops.dagstuhl.de/opus/volltexte/2014/4815
17. Emerson, J., Silva, M., Moussa, O., Ryan, C., Laforest, M., Baugh, J., Cory, D.G., Laflamme, R.: Symmetrized characterization of noisy quantum processes. Science **317**, 1893 (2007)
18. Wallman, J., Granade, C., Harper, R., Flammia, S.T.: Estimating the coherence of noise. New J. Phys. **17**, 113020 (2015)
19. The code used for running the algorithms is available online via GitHub. https://github.com/agnostiQ/DQC1-knots. Accessed 12 Nov 2019
20. Baez, J., Muniain, J.P.: Gauge Fields, Knots and Gravity. World Scientific Publishing Company, Singapore (1994)
21. Pullin, J.: Knot theory and quantum gravity in loop space: a primer. AIP Conf. Proc. **317**, 141 (1994)
22. Lackenby, M.: A polynomial upper bound on Reidemeister moves. Ann. Math. **182**, 491 (2015)

23. Jones, V.F.R.: A polynomial invariant for knots via Von Neumann algebras. Bull. Am. Math. Soc. **12**, 103 (1985)
24. Jones, V.F.R.: Hecke algebra representations of braid groups and link polynomials. Ann. Math. **126**, 335 (1987)
25. Aharonov, D., Jones, V., Landau, Z.: A polynomial quantum algorithm for approximating the Jones polynomial. Algorithmica **55**, 395 (2009)
26. Passante, G., Moussa, O., Ryan, C.A., Laflamme, R.: Experimental approximation of the Jones polynomial with one quantum bit. Phys. Rev. Lett. **103**, 250501 (2009)
27. Passante, G.: On experimental deterministic quantum computation with one quantum bit (DQC1). Ph.D. thesis, University of Waterloo (2012)
28. Lu, D., Li, H., Trottier, D.A., Li, J., Brodutch, A., Krismanich, A.P., Ghavami, A., Dmitrienko, G.I., Long, G., Baugh, J., et al.: Experimental estimation of average fidelity of a clifford gate on a 7-qubit quantum processor. Phys. Rev. Lett. **114**, 140505 (2015)
29. Knill, E., Leibfried, D., Reichle, R., Britton, J., Blakestad, R.B., Jost, J.D., Langer, C., Ozeri, R., Seidelin, S., Wineland, D.J.: Randomized benchmarking of quantum gates. Phys. Rev. A **77**, 012307 (2008)
30. Emerson, J., Alicki, R., Życzkowski, K.: Scalable noise estimation with random unitary operators. J. Opt. B: Quantum Semiclass. Opt. **7**, S347–S352 (2005)
31. Boixo, S., Isakov, S.V., Smelyanskiy, V.N., Babbush, R., Ding, N., Jiang, Z., Bremner, M.J., Martinis, J.M., Neven, H.: Characterizing quantum supremacy in near-term devices. Nat. Phys. **14**, 595–600 (2018)
32. McKay, D.C., Sheldon, S., Smolin, J.A., Chow, J.M., Gambetta, J.M.: Three-qubit randomized benchmarking. Phys. Rev. Lett. **122**, 200502 (2019)
33. Proctor, T.J., Carignan-Dugas, A., Rudinger, K., Nielsen, E., Blume-Kohout, R., Young, K.: Direct randomized benchmarking for multiqubit devices. Phys. Rev. Lett. **123**, 030503 (2019)
34. Vandersypen, L.M.K., Steffen, M., Breyta, G., Yannoni, C.S., Sherwood, M.H., Chuang, I.L.: Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. Nature **414**, 883–887 (2001)
35. Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.H., Zhou, X.Q., Love, P.J., Aspuru-Guzik, A., O'Brien, J.L.: A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. **5**, 4213 (2014)
36. Farhi, E., Goldstone, J., Gutmann, S.: A quantum approximate optimization algorithm (2014). arXiv:1411.4028
37. Benedetti, M., Garcia-Pintos, D., Perdomo, O., Leyton-Ortega, V., Nam, Y., Perdomo-Ortiz, A.: A generative modeling approach for benchmarking and training shallow quantum circuits. npj Quantum Inf. **5**, 45 (2019)
38. Verdon, G., Broughton, M., McClean, J.R., Sung, K.J., Babbush, R., Jiang, Z., Neven, H., Mohseni, M.: Learning to learn with quantum neural networks via classical neural networks (2019). arXiv:1907.05415
39. Cross, A.W., Bishop, L.S., Sheldon, S., Nation, P.D., Gambetta, J.M.: Validating quantum computers using randomized model circuits. Phys. Rev. A **100**, 032328 (2019)
40. Shin, S.W., Smith, G., Smolin, J.A., Vazirani, U.: How "quantum" is the D-wave machine? (2014). arXiv:1401.7087
41. Hamerly, R., Inagaki, T., McMahon, P.L., Venturelli, D., Marandi, A., Onodera, T., Ng, E., Rieffel, E., Fejer, M.M., Utsunomiya, S., Takesue, H., Yamamoto, Y.: In: Conference on Lasers and Electro-Optics, p. FTu4A.2. Optical Society of America (2018)
42. Morimae, T., Fujii, K., Fitzsimons, J.F.: Hardness of classically simulating the one-clean-qubit model. Phys. Rev. Lett. **112**, 130502 (2014)