**Process Tomography on a 7-Qubit Quantum Processor via Tensor Network Contraction Path Finding**

Aidan Dang,¹ Gregory A. L. White,¹ Lloyd C. L. Hollenberg,¹ and Charles D. Hill¹,²

¹School of Physics, The University of Melbourne, Parkville, Victoria 3010, Australia
²School of Mathematics and Statistics, The University of Melbourne, Parkville, Victoria 3010, Australia

(Dated: 2021-12-13)

Quantum process tomography (QPT), where a quantum channel is reconstructed through the analysis of repeated quantum measurements, is an important tool for validating the operation of a quantum processor. We detail the combined use of an existing QPT approach based on tensor networks (TNs) and unsupervised learning with TN contraction path finding algorithms in order to use TNs of arbitrary topologies for reconstruction. Experiments were conducted on the 7-qubit IBM Quantum Falcon Processor `ibmq_casablanca`, where we demonstrate this technique by matching the topology of the tensor networks used for reconstruction with the topology of the processor, allowing us to extend past the characterisation of linear nearest neighbour circuits. Furthermore, we conduct single-qubit gate set tomography (GST) on each individual qubit to correct for separable errors during the state preparation and measurement phases of QPT, which are separate from the channel under consideration but may negatively impact the quality of its reconstruction. We are able to report a fidelity of 0.89 against the ideal unitary channel of a single-cycle random quantum circuit performed on `ibmq_casablanca`, after obtaining just $1.1 \times 10^5$ measurements for the reconstruction of this 7-qubit process. This represents more than five orders of magnitude fewer total measurements than the number needed to conduct full, traditional QPT on a 7-qubit process.

**I. INTRODUCTION**

As the development of quantum computing hardware continues to progress ¹², there is an increasing need for scalable techniques to characterise imperfect qubit processes. The standard approach for a full characterisation of the process for a physical instance of a quantum circuit is quantum process tomography (QPT) ³⁴. In QPT, the process for a circuit on $n$ qubits is reconstructed from a sufficiently large set of measurements made from an informationally complete set of circuit input and output settings. As such, estimating the $O(2^{4n})$-sized Choi matrix ⁵⁶ which uniquely describes the process on $n$ qubits through the standard QPT procedure scales exponentially in time. This prohibitive time cost has limited the use of the standard QPT procedure to general processes on up to 2 or 3 qubits ⁷¹³. When constrained by the limited connectivity, available basis gate operations and error rates of near-term quantum devices, only circuits with particular structure can actually be performed while maintaining a high level of coherence. The use of tensor network (TN) based algorithms ¹⁵⁻¹⁷ has proven extremely successful in exploiting such structure in order to reduce classical costs in problems such as quantum circuit simulation ¹⁸⁻²¹, calculating Hamiltonian ground states ²² and simulating time evolution ²³⁻²⁴ whenever such structure permits. In these algorithms, the classical time costs ²⁰ will typically scale as a function of entanglement induced by the circuit ²⁷ and the error rates of the target device ²⁻²⁸, as opposed to just the system’s size in qubits.

Recently, the use of tensor networks has been applied to the approximate characterisation of sufficiently noiseless instances of large quantum circuits with one-dimensional topologies ¹⁴. This TN-based approach for QPT also requires a smaller set of measurement data, which is then used to derive the final estimate of the process through an optimisation loop inspired by those typical in unsupervised learning algorithms. A schematic of the procedure is shown in Fig. ¹. We also note other approaches for QPT which operate on measurement data sets with restricted size such as ansatz bootstrapping ⁸ and self-guided QPT ⁵¹, as well as compressed sensing ³⁸⁻³⁹ and shadow tomography ⁴¹⁻⁵⁵ for quantum state tomography (QST).

In this work, we develop the tensor network QPT approach in two ways. First, by employing algorithms for finding optimal (or heuristically efficient) orderings for the pairwise contractions of tensors within a TN ³⁶⁻³⁷, we alleviate the limitation that the original TN-based procedure had with characterising circuits not easily described in one dimension. Second, we recognise that the preparation of the states for input to the circuit and the subsequent qubit measurements made during the data collection are themselves affected by noise but not characterised by QPT as part of the circuit’s process, so we apply single-qubit corrections derived from gate set tomography ³⁸⁻³⁹ to mitigate separable errors outside of the main process.

Finally, we present the results of experiments using our TN-based QPT on circuits run on the `ibmq_casablanca` quantum processor ⁴⁰. This device has 7 qubits in an l-beam topology, making it a suitable test for our approach. We were able to provide a reconstruction for a 7-qubit random quantum circuit through $1.1 \times 10^5$ single-shot measurement data points; significantly less than the
approx 7 \times 10^{10} \text{ measurements required when performing the full, traditional QPT on 7 qubits.}

II. REVIEW OF TENSOR NETWORK BASED QPT WITH UNSUPERVISED LEARNING

We review the procedure for tensor network based QPT introduced by Torlai et al. [14]. For a quantum channel on \( N \) qubits described by a completely positive trace-preserving (CPTP) map \( \mathcal{E} \), a reconstruction \( \Lambda_\theta \) is generated for the corresponding Choi matrix \( \Lambda_\mathcal{E} \) [6]. The reconstruction \( \Lambda_\theta \) is represented by a TN structure known as a locally purified density operator (LPDO) [11]. In this TN representation, each qubit \( q_j \) acted on by the reconstructed process \( \Lambda_\theta \) is associated with a tensor \( [A_j] \) and its elementwise conjugate \( [A_j^*] \). Each of the tensors \( [A_j] \) are then connected to each other according to some topology (linear nearest neighbour (LNN) in the case of [14]), and the tensors \( [A_j^*] \) are connected similarly. Finally, a Kraus bond is formed between each \( [A_j] \) with its corresponding \( [A_j^*] \). For the LNN LPDO, the elements \( \langle \sigma, \tau | \Lambda_\theta | \sigma', \tau' \rangle \) of the Choi matrix \( \Lambda_\theta \) with input basis \( \{ |\sigma\rangle \} \) and output basis \( \{ |\tau\rangle \} \) are given by the contraction of the entire LPDO:

\[
\langle \sigma, \tau | \Lambda_\theta | \sigma', \tau' \rangle = \sum_{\{\mu, \mu', \nu, \nu'\}} \prod_{j=1}^{N} [A_j]_{\mu_j-1, \nu_j, \mu_j} [A_j^*]_{\nu_j', \nu_j, \mu_j'} \tag{1}
\]

where the Kraus bond between \( [A_j] \) and \( [A_j^*] \) is indexed by \( \nu_j \), and the LNN bonds within the non-conjugate layer \( \{ [A_j] \} \) and conjugate layer \( \{ [A_j^*] \} \) are indexed by \( \{ \mu_j \} \) and \( \{ \mu_j' \} \) respectively. Examples of these LPDO tensor networks are shown in Fig. 2.

In order for the corresponding map of the reconstructed Choi matrix \( \Lambda_\theta \) to be CP, \( \Lambda_\theta \) must be positive semidefinite, which is enforced by the LPDO tensor network structure through the conjugate and non-conjugate layers. For this same map to be TP, the partial trace over the outputs of the Choi matrix should result in the identity over the inputs, so that \( \text{Tr}_{\tau} \Lambda_\theta = \mathbb{I}_{\sigma} \). The normalisation \( Z = \text{Tr}_{\sigma, \tau} \Lambda_\theta \) of the Choi matrix should be set at \( d^N \), where \( d = 2 \) for qubits.

The complex-valued tensors in the non-conjugate layer \( \{ [A_j] \} \) form the variational parameters \( \vartheta \) and are iteratively trained via the following procedure. Data from the considered channel \( \mathcal{E} \) is gathered by applying it to the input product state \( \rho_\alpha = \bigotimes_{j=1}^{N} \rho_{\alpha_j} \) of single-qubit states \( \rho_{\alpha_j} \) each chosen from a set of size \( K_\rho \). Following the channel, a measurement \( M_\beta = \bigotimes_{j=1}^{N} M_{\beta_j} \) is obtained by applying a positive operator-valued measure (POVM) \( \{ M_{\beta_j} \}_{\beta_j=1}^{K_\mu} \) with \( K_\mu \) outcomes to each qubit. A single data point is therefore distinguished by its input tuple \( \alpha = (\alpha_1, \ldots, \alpha_N) \) and output tuple \( \beta = (\beta_1, \ldots, \beta_N) \) containing the indices of the elements within the set of

![Figure 1. The general procedure for QPT of a circuit \( U \) based on tensor networks with unsupervised learning [14]. Here, \( M \) instances of \( U \) are constructed with input states and output measurement bases drawn from an informationally complete set, configured by the choice of \( i_j \) and \( o_j \) rotations respectively. With single-shot measurement outcomes from each configuration, an optimisation loop based on stochastic gradient descent (SGD) with adaptive learning rate \( \eta \) and cost \( C \) can be performed in order to update the tensors \( [A_j] \) of a tensor network representation of the reconstructed process’ Choi matrix. For processes of sufficiently small tensor network bond and Kraus dimensions, the total number of required measurements \( M \) may be significantly less than the \( \mathcal{O}(12^n) \) required for full QPT via the traditional approach.](image)

![Figure 2. Example LPDO tensor networks for 3-qubit processes. In both cases, the Kraus bond indexed by \( \nu_j \) connects tensor \( [A_j] \) with its conjugate \( [A_j^*] \), and indices \( \{ \sigma_j, \sigma_j' \} \) and \( \{ \tau_j, \tau_j' \} \) are used for inputs and outputs to the process respectively. (a) An LPDO in the linear nearest neighbour configuration used in [14], where bonds indexed by \( \{ \mu_j \} \) connect pairs of consecutive tensors from the non-conjugate layer \( \{ [A_j] \} \). (b) A fully connected configuration, where each pair of tensors within the non-conjugate layer is joined by some bond.](image)
input states used and within the POVM measured. The chosen set of single-qubit input states and POVM should both be informationally complete to ensure unique characterisation of $\mathcal{E}$.

The probability $P_{\theta}(\beta \mid \alpha)$ for the properly normalised reconstructed process $\Lambda_\theta$ to produce the measurements $M_{\beta}$ given the input state $\rho_\alpha$ is calculated via the tensor network contraction of

$$P_{\theta}(\beta \mid \alpha) = \text{Tr}_{\sigma, \tau} \left[ (\rho_\alpha^T \otimes M_{\beta}) \Lambda_\theta \right], \quad (2)$$

an example of which is shown in Fig. 3. With a mini-batch of $M$ data points consisting of uniformly sampled input index tuples $\{\alpha_k\}_{k=1}^M$ with their respective measurement tuples $\{\beta_k\}_{k=1}^M$, some form of gradient descent may be applied to optimise for the parameters $\theta$ with respect to the mean negative log-likelihood $-\frac{1}{M} \sum_{k=1}^M \log P_{\theta}(\beta_k \mid \alpha_k)$. To control convergence toward the TP condition $\text{Tr}_{\tau} \Lambda_\theta = I_\sigma$, the regularisation term

$$\Gamma_{\theta} = \sqrt{d^{-N} \text{Tr}_{\sigma} (\Delta_\theta \Delta_\theta^T)} \quad (3)$$

where $\Delta_\theta = \text{Tr}_{\tau} \Lambda_\theta - I_\sigma$ is calculated through tensor network operations (see [14] Fig. 10), giving a final cost function of

$$C(\theta) = -\frac{1}{M} \sum_{k=1}^M \log P_{\theta}(\beta_k \mid \alpha_k) + \kappa \Gamma_{\theta} \quad (4)$$

over the mini-batch, where $\kappa$ is a tunable hyperparameter. The appropriate Wirtinger derivatives [42] for complex-valued gradient descent may be calculated through automatic differentiation.

III. EXTENSIONS TO THE QPT PROCEDURE

We introduce two extensions to the original QPT procedure based on tensor networks with unsupervised learning detailed in the previous section. For our implementation, we used the JAX library [13] for GPU acceleration, automatic differentiation and automatic vectorisation, and the TensorNetwork library [44] for tensor network semantics.

A. LPDOs with Arbitrary Topologies

The numerical experiments presented in [14] were performed using LPDO tensor networks with 1-dimensional linear nearest neighbour (LNN) connectivity. These networks allow for an optimal contraction strategy for calculating derivatives and performing simulated QPT measurement sampling. However, they are not well suited for characterising processes on quantum systems without 1D LNN connectivity since any non-LNN interaction may result in an increase in the dimension of all bonds spanned by the interaction.

As such, we introduce the use of LPDOs with arbitrary topologies (which still follow the structure of an LPDO, with a conjugate and non-conjugate layer connected with a Kraus bond per site), an example of which is provided in Fig. 2(b). These tensor networks do not have a single strategy for performing optimised contractions, unlike the 1D LNN LPDOs. Instead, algorithms such as those implemented in the opt_einsum library [15] for finding contraction orderings with optimised floating point operations count were used [36]. We combine the use of these algorithms for finding contraction orderings with automatic vectorisation from JAX to ensure that efficient contractions of an entire mini-batch’s worth of probabilities $\{P_{\theta}(\beta_k \mid \alpha_k)\}$ could be performed. Finally, automatic differentiation [40] is applied in order to calculate gradients for the purpose of optimisation.

B. Gate Set Tomography

When gathering data points on a physical device, errors introduced during input state initialisation and final POVM measurement will result in a different actual input state and POVM than those specified during the classical reconstruction. As such, we perform single-qubit gate set tomography (GST) [38, 39] on each qubit to obtain its input state set and POVM corrected for single-qubit errors. These input state sets and POVMs specific for each qubit were then used in the relevant tensor network contractions, such as in the calculation of $P_{\theta}(\beta \mid \alpha)$. We performed GST using the pyGSTi library [47].

IV. RESULTS

We present results for QPT experiments performed on the ibmq_casablanca 7-qubit system, one of the IBM Falcon Quantum Processors, whose I-beam topology shown in Fig. 4 matches with the LPDO tensor network. Unless otherwise noted, we use a regularisation weight of...
\( \kappa = 0.1 \) and the Adam optimiser [48] with default parameters. We perform each experiment and gather the corresponding GST corrections in a contiguous reserved time slot in order to minimise the potential for large calibration changes in the device itself. The \( \sqrt{X} \) and CNOT error rates averaged over this time slot are also shown in Fig. 4.

For each of our experiments, we use the informationally complete set of single-qubit states containing

\[
X_+ = |+\rangle \langle +|, \quad X_- = |\rangle \langle -|,
\]
\[
Y_+ = |+i\rangle \langle +i|, \quad Y_- = |i\rangle \langle -i|,
\]
\[
Z_+ = |0\rangle \langle 0|, \quad Z_- = |1\rangle \langle 1|
\]

to select inputs \( \rho_\alpha \) to the channel for each data instance. We use the same set properly normalised (i.e. \( \{X_\pm/3, Y_\pm/3, Z_\pm/3\} \)) as the single-qubit POVM \( M_{\beta_j} \) for each qubit following the channel. The chosen inputs \( \rho_\alpha \) are prepared by performing the relevant single-qubit rotation from the default initial state \( |0\rangle \), and the POVM measurements are emulated by choosing either the \( X \) or \( Z \) basis uniformly at random to measure in (which consists of performing the necessary rotation and final computational basis measurement). We performed separate instances of GST on each qubit in parallel to better characterise the actual input states \( \rho_\alpha \) and output POVMs \( M_{\beta_j} \) used in our experiments. These corrections were then used during the numerical reconstruction procedure. We place barriers in between the state initialisation, process and final measurement to prevent compiler optimisation across these phases.

Our first experiment involved characterising the process of a single layer of Hadamard gates performed on \textit{ibmq\_casablanca}. In Fig. 5(a), we show the fidelity \( F(\Lambda_\epsilon, \Lambda_{\theta_j}) \) between our reconstruction \( \Lambda_{\theta_j} \) against the ideal unitary Hadamard layer process \( \Lambda_\epsilon \) and (b) the loss over the training data set of size \( M = 1.2 \times 10^4 \) and the validation data set of size 0.25M at each epoch. No mini-batching was performed since the training set here was sufficiently small. The fidelity reported at the lowest validation loss was 0.963. (c) The fidelity calculated at the point of lowest validation loss over 1000 training epochs with various data set sizes \( M \), where the validation data set has size 0.25M in each instance.

\[
F(\Lambda_\epsilon, \Lambda_{\theta_j}) = d^{-2N} \left( \text{Tr} \sqrt{\Lambda_\epsilon \Lambda_{\theta_j} \sqrt{\Lambda_\epsilon}} \right)^2
\]

when \( \Lambda_\epsilon \) and \( \Lambda_{\theta_j} \) are properly normalised. In the general case this calculation is not efficient, but when \( \Lambda_\epsilon \) is a unitary process such as our Hadamard layer, its LPDO tensor network will have trivial Kraus bonds of dimen-
sion 1, essentially disconnecting the non-conjugate and conjugate layers (which we denote as $|\Lambda_\mathcal{E}\rangle$ and $\langle\Lambda_\mathcal{E}|$ respectively, so that $\Lambda_\mathcal{E} = |\Lambda_\mathcal{E}\rangle\langle\Lambda_\mathcal{E}|$). In this case, the fidelity calculation reduces to the efficient tensor network contraction $F(\Lambda_\mathcal{E}, \Lambda_\varphi) = d^{-2N} \langle \Lambda_\mathcal{E} | \Lambda_\varphi | \Lambda_\mathcal{E} \rangle$.

In Fig. 5(b), we present the loss $\mathcal{C}(\varphi)$ over the training data set of size $M = 1.2 \times 10^4$ over the duration of the training period. To prevent overfitting when selecting the most appropriate model parameters $\vartheta_i$ at some epoch $t$, we perform cross-validation by calculating the loss over a distinct validation set of size $0.25M$ at each epoch. We then select the parameters $\vartheta_i$ which result in the lowest validation loss across all training epochs. With these parameters, our reconstruction suggests that $ibmq\_casablanca$ performed the 7-qubit Hadamard layer with a fidelity of 0.963. Furthermore, in Fig. 5(c), we show the infidelity with respect to training set size $M$. Convergence here suggests that a sufficiently large training set was indeed used to derive an accurate reconstruction.

Upon transpilation of the 7-qubit Hadamard layer to the basis gate set of $ibmq\_casablanca$, the resulting circuit contains one $\sqrt{X}$ gate on each qubit. By compoundung individual $\sqrt{X}$ gate errors $r_g$ from the rates shown in Fig. 3, we derive a baseline fidelity $F_{\text{base}} = \prod_g (1 - r_g) = 0.9957$ for the 7-qubit Hadamard layer. The presence of crosstalk and other multi-qubit errors are not captured by this value, so our reconstruction’s fidelity of 0.963 still appears reasonable.

Next, we performed experiments to characterise single-cycle random quantum circuits (RQCs) [2] with CZ gates as interactions. In these circuits, the 2-qubit interactions are applied two at a time, to pairs of qubits

1. $(q_0, q_1)$ and $(q_3, q_5)$
2. $(q_1, q_2)$ and $(q_4, q_5)$
3. $(q_1, q_3)$ and $(q_5, q_6)$.

Just before each interaction, a randomly selected single-qubit rotation $\pi$ was applied to each of the qubits being
interacted, as seen in Fig. 6. We present results for the QPT procedure in Fig. 7 for a single instance of chosen random unitaries. An ideal realisation of the single-cycle RQC with CZ interactions would be perfectly characterisable by the topologically matching LPDO with bond dimension 2 (equal to the operator Schmidt rank of the CZ gate) and Kraus dimension 1 (due to unitarity). With our reconstruction of all bond and Kraus dimensions set to 5, we report a fidelity of 0.892 against the ideal unitary RQC process at the lowest validation loss. It is worth noting that this reconstruction was performed with just $M = 1.1 \times 10^5$ data instances, each itself a single-shot measurement performed on $\text{ibmq\_casablanca}$. By compounding the error rates in Fig. 4 with the relevant CNOT and $\sqrt{X}$ gate counts in Fig. 6, the fidelity derived from $\text{ibmq\_casablanca}$’s own calibration data is extremely close with $F_{\text{base}} = 0.897$. Convergence of the reconstruction’s fidelity against the ideal unitary process shown in Fig. 7(c) again suggests that this choice of $M$ is sufficient.

Finally, we compare the reduction of the reconstruction obtained from our TN-based approach with that of the traditional, full QPT approach on the subset of qubits $\{q_1, q_3, q_5\}$. Here, the process on $\{q_1, q_3, q_5\}$ is the same single-cycle RQC over all seven qubits, but where the remaining qubits $\{q_0, q_2, q_4, q_6\}$ are each initialised to $Z_+$. Importantly, this process is not a unitary operation on qubits $\{q_1, q_3, q_5\}$.

Using the default settings in Qiskit [50] for performing traditional QPT on IBM quantum systems, we use the informationally complete set of input qubit states $\{Z_+, Z_-, X_+, Y_+\}$ (instead of the six possible input states used during the TN-based experiments) and measure along an axis chosen uniformly randomly from $\{X, Y, Z\}$ (resulting in an identical POVM set used in the TN-based experiments). Therefore, for this three-qubit process we need to perform sufficiently many measurements to obtain 1272 probability distributions. Using a typical 2000 shots per selection of input states and measurement axes gives a total of $12^3 \times 2000 = 3.46 \times 10^6$ individual measurements that we performed in order to derive a reconstruction through the traditional QPT approach. We report a fidelity of 0.80 between the reconstruction derived from the traditional, full QPT approach with the ideal (non-unitary) process on the three qubits.

For the TN-based approach, we again perform tomography of the RQC across all seven qubits, but then reduce it to the three-qubit process by applying the tensor network of the reconstruction $A_0$ to the GST-corrected input states corresponding to $Z_+$ for qubits $\{q_0, q_2, q_4, q_6\}$ and tracing the output indices associated with these qubits. An example of reducing an LPDO down to a subset of qubits is shown in Fig. 8. Using $4 \times 10^5$ single-shot measurements, we report a fidelity of 0.96 between the reconstruction derived from the TN-based approach against the ideal process, and a fidelity of 0.81 between the reconstructions of the TN-based and traditional approaches.

While the agreement between the two approaches is decent, we justify the discrepancy by noting that the traditional QPT performed through Qiskit does not include any form of GST correction, whereas our TN-based approach makes use of it to correct for single-qubit errors in the initialisation of all qubits, including qubits $\{q_0, q_2, q_4, q_6\}$, and the POVM measurements. Furthermore, despite using an order of magnitude fewer total measurements, the total time taken to collect the single-shot measurements for the TN-based approach was significantly longer than the time taken to perform all measurements for the traditional approach. It is possible that changes in the device’s calibration during this prolonged data collection period may have contributed to this discrepancy between the reconstructions from each approach. This particular issue is due to shortcomings of the classical queueing and processing of circuit requests on the IBM Quantum systems which we hope can be rectified in a future update [51].

V. CONCLUSION

We presented several extensions to an existing scheme for quantum process tomography based on tensor networks and unsupervised learning. Namely, we combined the use of algorithms for finding good tensor network contraction orderings with the typical machine learning procedures of batching and automatic differentiation to construct tensor networks with topologies matched toward a particular device architecture. This can allow for a more efficient tensor network representation of the reconstructed process when compared to the equivalent 1D tensor network of identical bond and Kraus dimensions. We also suggest the use of gate set tomography during the collection of data points in order to correct for single-qubit initialisation and measurement errors.

Experiments were conducted on the $\text{ibmq\_casablanca}$ 7-qubit quantum processor to demonstrate the use of our QPT procedure with tensor networks matched to the device’s I-beam topology. For the single Hadamard layer, we report a fidelity of 0.963 against the ideal unitary determined from $M = 1.2 \times 10^4$ single-shot measurements, and similarly a fidelity of 0.892 for the random quantum circuit of one cycle determined from $M = 1.1 \times 10^5$ measurements. Both of these measurement data sets are significantly smaller than the $2000 \times 12^7 \approx 7.2 \times 10^{10}$
measurements needed for performing full QPT on a 7-qubit process via the traditional approach.

We then compared the reconstruction of the single-cycle RQC on *ibmq_casablanca* reduced to a 3-qubit subset with that of the traditional QPT procedure performed on these three qubits, and report a fidelity of 0.81 between the reconstructions of these two approaches. Given that our TN-based approach for characterising this RQC on seven qubits used roughly an order of magnitude fewer total measurements than the traditional approach even on just three qubits, we can be confident in the viability of our approach for the characterisation of large quantum processes which can be represented as local purified density operators with manageable bond dimensions.

**ACKNOWLEDGMENTS**

We thank G. Torlai for useful discussion regarding particular details about [14]. This work was supported by the University of Melbourne through the establishment of an IBM Quantum Network Hub at the university. A.D. and G.A.L.W are supported by Australian Government Research Training Program Scholarships. C.D.H. is supported through a Laby Foundation grant at the University of Melbourne.

[1] J. Preskill, Quantum Computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018)
[2] F. Arute, K. Arya, R. Babbush, *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019)
[3] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, *J. Mod. Opt.* **44**, 2455 (1997)
[4] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Quantum-process tomography: Resource analysis of different strategies, *Phys. Rev. A* **77**, 032322 (2008)
[5] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10**, 285 (1975)
[6] C. J. Wood, J. D. Biamonte, and D. G. Cory, Tensor networks and graphical calculus for open quantum systems, *Quantum Inf. Comput.* **15**, 759 (2015)
[7] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Bouland, M. Saraceno, S. Lloyd, and D. G. Cory, Quantum process tomography of the quantum Fourier transform, *J. Chem. Phys.* **121**, 6117 (2004)
[8] L. C. G. Govia, G. J. Ribeill, D. Ristè, M. Ware, and H. Krovi, Bootstrapping quantum process tomography via a perturbative ansatz, *Nat. Commun.* **11**, 1084 (2020)
[9] A. M. Childs, I. L. Chuang, and D. W. Leung, Realization of quantum process tomography in NMR, *Phys. Rev. A* **64**, 012314 (2001)
[10] J. L. O’Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White, Quantum Process Tomography of a Controlled-NOT Gate, *Phys. Rev. Lett.* **93**, 080502 (2004)
[11] R. C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. D. O’Connell, D. Sank, H. Wang, J. Wenner, M. Steffen, A. N. Cleland, and J. M. Martinis, Quantum process tomography of a universal entangling gate implemented with Josephson phase qubits, *Nat. Phys.* **6**, 409 (2010)
[12] H. N. Tinkey, A. M. Meier, C. R. Clark, C. M. Seck, and K. R. Brown, Quantum process tomography of a Molmer-Sørensen gate via a global beam, *Quantum Sci. Technol.* **6**, 034013 (2021)
[13] T. Yamamoto, M. Neeley, E. Lucero, R. C. Bialczak, J. Kelly, M. Lenander, M. Mariantoni, A. D. O’Connell, D. Sank, H. Wang, M. Weides, J. Wenner, Y. Yin, A. N. Cleland, and J. M. Martinis, Quantum process tomography of two-qubit controlled-Z and controlled-NOT gates using superconducting phase qubits, *Phys. Rev. B* **82**, 184515 (2010)
[14] G. Torlai, C. J. Wood, A. Acharya, G. Carleo, J. Carrasquilla, and L. Aolita, Quantum process tomography with unsupervised learning and tensor networks (2020), arXiv:2006.02424 [quant-ph]
[15] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys. (N. Y.)* **349**, 117 (2014)
[16] R. Orús, Tensor networks for complex quantum systems, *Nat. Rev. Phys.* **1**, 538 (2019)
[17] J. Biamonte and V. Bergholm, Quantum Tensor Networks in a Nutshell (2017), arXiv:1708.00006 [quant-ph]
[18] A. Dang, C. D. Hill, and L. C. L. Hollenberg, Optimising Matrix Product State Simulations of Shor’s Algorithm, *Quantum* **3**, 116 (2019)
[19] J. Gray and S. Kourtis, Hyper-optimized tensor network contraction, *Quantum* **5**, 410 (2021)
[20] C. Huang, F. Zhang, M. Newman, *et al.*, Classical Simulation of Quantum Supremacy Circuits (2020), arXiv:2005.06787 [quant-ph]
[21] F. Pan and F. Zhang, Simulating the Sycamore quantum supremacy circuits (2021), arXiv:2103.03074 [quant-ph]
[22] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys. (N. Y.)* **326**, 96 (2011)
[23] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, *Phys. Rev. Lett.* **93**, 040502 (2004)
[24] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems, *Phys. Rev. Lett.* **93**, 207204 (2004)
[25] S. Paecel, T. Köhler, A. Svoboda, S. R. Mannmana, U. Schollwöck, and C. Hubig, Time-evolution methods for matrix-product states, *Ann. Phys. (N. Y.)* **411**, 167998 (2019)
[26] B. O’Gorman, Parameterization of Tensor Network Contraction, in *14th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2019)*, Vol. 135, edited by W. van Dam and L. Mancinska (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2019) p. 10:1
[27] G. Vidal, Efficient Classical Simulation of Slightly En-
[28] A. Dang, Distributed Matrix Product State Simulations of Large-Scale Quantum Circuits (2017), hdl:11343/239081
[29] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What Limits the Simulation of Quantum Computers?, Phys. Rev. X 10, 041038 (2020)
[30] K. Noh, L. Jiang, and B. Fefferman, Efficient classical simulation of noisy random quantum circuits in one dimension, Quantum 4, 318 (2020)
[31] Z. Hou, J.-F. Tang, C. Ferré, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Experimental realization of self-guided quantum process tomography, Phys. Rev. A 101, 022317 (2020)
[32] D. L. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52, 1289 (2006)
[33] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography via Compressed Sensing, Phys. Rev. Lett. 105, 150401 (2010)
[34] S. Aaronson, Shadow tomography of quantum states, SIAM J. Comput. 49, STOC18-368 (2020)
[35] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. 16, 1050 (2020)
[36] R. N. C. Pfeifer, J. Haegeman, and F. Verstraete, Faster identification of optimal contraction sequences for tensor networks, Phys. Rev. E 90, 033315 (2014)
[37] F. Schindler and A. S. Jermyn, Algorithms for tensor network contraction ordering, Mach. Learn.: Sci. Technol. 1, 035001 (2020)
[38] D. Greenbaum, Introduction to Quantum Gate Set Tomography (2015), arXiv:1509.02921 [quant-ph]
[39] M. R. Geller, Conditionally Rigorous Mitigation of Multiqubit Measurement Errors, Phys. Rev. Lett. 127, 090502 (2021)
[40] IBM Quantum, https://quantum-computing.ibm.com/ (2021).
[41] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Positive Tensor Network Approach for Simulating Open Quantum Many-Body Systems, Phys. Rev. Lett. 116, 237201 (2016)
[42] W. Wirtinger, Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen, Math. Ann. 97, 357 (1927)
[43] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs, http://github.com/google/jax (2018).
[44] C. Roberts, A. Milsted, M. Ganahl, A. Zalcman, B. Fontaine, Y. Zou, J. Hidary, G. Vidal, and S. Leichenauer, TensorNetwork: A Library for Physics and Machine Learning (2019), arXiv:1905.01330 [physics.comp-ph]
[45] D. G. A. Smith and J. Gray, opt_einsum - A Python package for optimizing contraction order for einsum-like expressions, J. Open Source Softw. 3, 753 (2018)
[46] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, Differentiable Programming Tensor Networks, Phys. Rev. X 9, 031041 (2019)
[47] Erik, L. Saldyt, Rob, et al., pyGSTi: A python implementation of Gate Set Tomography (2021).
[48] D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization (2014), arXiv:1412.6980 [cs.LG]
[49] F. Mezzadri, How to generate random matrices from the classical compact groups, Not. Am. Math. Soc. 54, 592 (2007).
[50] M. S. ANIS, H. Abraham, AduOffei, et al., Qiskit: An Open-source Framework for Quantum Computing (2021).
[51] Qiskit Runtime - IBM Quantum, https://quantum-computing.ibm.com/lab/docs/iql/runtime/ (2021).