Experimental non-Markovian process characterisation and control on a quantum processor

G. A. L. White,1,* C. D. Hill,1,2,† F. A. Pollock,3,‡ L. C. L. Hollenberg,1,§ and K. Modi4,¶

1School of Physics, University of Melbourne, Parkville, VIC 3010, Australia
2School of Mathematics and Statistics, University of Melbourne, Parkville, VIC, 3010, Australia
3School of Physics and Astronomy, Monash University, Clayton, VIC 3800, Australia

As experimentally available quantum devices increase in precision and accessibility, attention is turning to understanding and eliminating correlated – or non-Markovian – noise. Here, we develop and experimentally test a framework for characterising non-Markovian dynamics in quantum systems. Applying this technique over four different IBM Q superconducting quantum devices, we achieve a consistent reconstruction infidelity of $10^{-3}$. Our approach is based on the recently proposed process tensor. With this, we infer the non-Markovian process by measuring the system’s response to a basis of control operations. As a consequence, the effects of any control operation in the span are discernible, independent of the interaction with the environment. With our technique, we demonstrate several applications: We first estimate a statistically significant lower-bound on memory size for all of the devices. We demonstrate that the dynamical characterisation remains high in fidelity where conventional Markovian characterisation models suffer from an appreciable reduction in quality. Finally, we turn these high fidelity predictions into an adaptive control technique for non-Markovian systems and demonstrate decoupling of a qubit interacting with another qubit with an unknown Hamiltonian. We further show how the coupling can be manipulated to implement a non-unitary gate of our choosing. The results and methods have widespread application both to the study of fully general dynamics, and to optimal device control.

I. INTRODUCTION

The theoretical machinery for open quantum system dynamics is well-oiled in low-coupling cases, but strong environmental interactions can lead to non-trivial dynamical memory effects that are difficult to understand, much less control. The recent advent of high performance quantum information processors (QIPs) has precipitated greater sensitivity to complex dynamical effects. In particular, it is clear that device behaviour must be understood under a relaxed Markov assumption [1–3]. The resulting non-Markovian dynamics includes more general errors that may be temporally correlated or dependent on broader environmental context [4–6]. Characterisation techniques of quantum devices such as randomised benchmarking (RB) and gate set tomography (GST) have so far represented the front line in understanding and addressing noise [7–11]. However, constructing a digestible picture of non-Markovian behaviour has proven difficult, and violates the error model assumed in these methods. Chiefly, this is because quantum correlations can forbid the division of dynamical processes into arbitrary steps of completely positive (CP), linear maps [12]. If information back-flow from the environment can occur, then noisy effects can be influenced by past factors; this detail can no longer be ‘forgotten’.

For device control, this is problematic. The circuit model of quantum computation is predicated on identical gates implemented at different times having identical actions. Markovian errors multiply out and propagate in predictable ways. However, non-Markovian noise gives rise to adverse effects that are much more challenging to tame. For example, correlated errors can spread across the device, and have been shown to lower thresholds of quantum error correcting codes [13, 14]. Similarly, context-dependent gates allow for poorly understood forms of dynamical errors not describable by a Markov model. This is one of the largest obstacles to near-term QIPs; non-Markovian noise must be either eliminated or, as some have suggested, harnessed into a resource [15–19]. Until recently, there has not been a clear operational definition for quantum non-Markovianity, nor consensus that one unifying measure could even be found.

Using the recent process tensor framework [20], we develop a robust device characterisation technique which is inclusive of non-Markovian dynamics. We keep discussion fully general, but demonstrate the capabilities of this method on four different IBM Q superconducting quantum devices. We then examine the robustness of the framework’s assumptions; address shortcomings; and demonstrate its functionality in process characterisation, memory detection, and application to adaptive quantum control. We find that we can characterise arbitrary processes down to an average infidelity of $10^{-3}$. This outperforms the characterisation given by the standard technique of GST in the presence of non-Markovian effect, which employs a comprehensive Markov models. With non-Markovian dynamics fully accounted for, we discuss applications of the process tensor generically to adaptive quantum control. As an example, we demonstrate how
two qubits can be decoupled without any a priori knowledge or assumption about their interactions, and how typically inaccessible user-designated non-unitary control operations can be realised. The efficacy of this framework over a range of devices showcases its consistency and broad-range of applicability. Our results represent significant progress towards the characterisation and optimal control of non-Markovian QIPs and other quantum devices.

A. Non-Markovian processes

To characterise non-Markovian device features, we employ the process tensor framework, which was recently developed to describe arbitrary quantum processes. Here, we briefly outline some relevant background before detailing our approach to the problem. Traditional approaches to quantum stochastic dynamics are concerned with tracking the state of the system (S) as a function of time: \( \rho_t = tr_E[U_{t:0} \rho_{0:E}] \), where \( U(\cdot) = U(\cdot) U^{\dagger} \) is a unitary map on system-environment (SE), initially in state \( \rho_{0:E} \) (often required to be uncorrelated). However, real experiments are driven by sequences of control operations, mathematically represented by trace non-increasing CP maps \( \{ A_0, \ldots, A_{k-1} \} =: A_{k-1:0} \). The process tensor is designed to account for the intermediate control operations and quantifies quantum correlations between past events and the final state of the system. In doing so, the process tensor formally generalises the notion of a stochastic process to the quantum domain [21] and reduces to a classical stochastic process in the correct limit [22, 23]. The formalism gives rise to a clear necessary and sufficient definition of quantum non-Markovianity [24], as well as other features of non-Markovian memory [25–27]. Figure 1a, top and bottom, illustrates respectively the traditional approach and the process tensor approach to describing a quantum process. In the top panel, a quantum state left to evolve in isolation can be reconstructed at \( t \) via quantum state tomography (QST). In the bottom panel, events come in the form of control operations applied to \( S \) at times \( t_1 \) and \( t_2 \); the future states of the \( S \) branch at time \( t \) are conditioned on the outcomes of the control operations.

Mathematically, the controlled dynamics has the form

\[
\rho_k(A_{k-1:0}) = tr_E[U_{k:k-1}A_{k-1} \cdots U_{1:0}A_0(\rho_{0:E})],
\]

which can be rearranged, as depicted in Figure 1b, to define a mapping from past control to future states:

\[
\rho_k(A_{k-1:0}) = T^{k:0}[A_{k-1:0}].
\]

The process tensor \( T^{k:0} \) is a multi-linear map on the control operations, and includes all of the information hidden to the experimenter, including correlations in the initial state, and any intermediate interaction with the environment.

The set of possible sequences of CP maps \( A_{k-1:0} \) forms a product vector space, built up from the spaces of temporally local operations; in particular, \( A_{k-1:0} = \otimes_{j=0}^{k-1} A_j \) when the operations at each time are chosen independently. As such, the process tensor is completely characterised by its input-output relations on a complete basis of control operations, just as a quantum channel is unambiguously defined by its input-output relations on a complete basis of states. Let us denote the basis for CP maps at the \( j \)th time step as \( \{ \{ \alpha \}_{j=1}^{d_j^2} \} \) and the basis sequences as \( \{ B_{k-1:0} \}_{k=(1,1,\ldots,1)} \) such that an arbitrary sequence of operations can be written as \( A_{k-1:0} = \sum_{\mu} \alpha^{\mu} B_{k-1:0}^{\mu} \), see Figure 1c. Then the process tensor’s action is defined by

\[
\rho_k(A_{k-1:0}) = \sum_{\mu} \alpha^{\mu} \rho_k^{\mu} \quad \text{with} \quad \rho_k^{\mu} := T^{k:0}[B_{k-1:0}^{\mu}]. \tag{2}
\]

In other words, to reconstruct the process tensor we need to experimentally estimate \( \rho_k^{\mu} \) for all \( \mu \), this is depicted in Figure 1d. In the methods section, we detail explicitly the steps to go from \( \rho_k^{\mu} \) to constructing the process tensor. Once the process tensor is reconstructed, using Equation (2), one can predict the final density matrix corresponding to any choice of control sequence \( A_{k-1:0} \), as shown in Figure 1e. We use prediction fidelity of the final states, conditioned on controls, as a performance metric for our process characterisation.

II. PROCESS CHARACTERISATION

We look now to the practical determination of the process tensor in experiment. The experiments carried out in this work used cloud-based IBM Q superconducting quantum devices. We first evaluated predictive capabilities of process tensor over a host of different experiments on the IBM Q devices Johannesbur (shortened: ‘Jo’burg’), Boeblingen (‘Boeb.’), Poughkeepsie (‘PK’), and Valencia. Our main contribution is in demonstrating how this framework leads to high fidelity process characterisation and precise control over non-Markovian dynamics. Ideally, complete process tensor construction would be achieved with the full span of CP maps. Unfortunately, efficient measurement within the coherence time is beyond the scope of most current hardware. For now, this rules out non-unital and trace-decreasing maps, affording only unitary control, i.e., we do not have a complete basis of operations. With these limitations, processes can still be characterised in terms of restricted process tensors \( T^{k:0} \) [28], defined in a similar way to the full process tensor, but constrained to act only on the subspace of operations comprising the linear span of unitary maps. This reduces the control space to the CP maps [29]; this work deals only with single qubit process tensors, for which the dimension is 10.

We reconstruct and test the four-time restricted process tensor \( T^{4:0} \) for a single qubit process on IBM Q devices. To do so, we first reconstruct the final quantum state \( \rho_{ijk}^{ijk} \). This state depends on the past controls, i.e., the initial preparation \( P_0 \in \mathcal{P} \) and the subsequent
FIG. 1. An illustrative summary of process characterisation. a The state of an open system over time follows a trajectory through state space until some final time at which the state is probed (top). By applying control operations at times $t_1$ and $t_2$, an experimenter can anchor and change the trajectory, which can be inferred via a linear combination of trajectories corresponding to basis operations (bottom). b A sequence of operations $A_{k-1:0}$ can be expressed as a tensor product of independently chosen operations $A_j$ at each time step. These can then be individually decomposed into a chosen basis $\{B^\mu_{\mu j}\}$ together giving a basis of sequences $\{B^\mu_{\mu j-1:0}\}$. c A circuit model showing how a process can be fully characterised by measuring the output state for a complete set of basis operations at different times. Then, an arbitrary process can be expressed as a linear combination of each basis process; because of the linear construction, the intermediate evolution is completely preserved in the description of the arbitrary process. d The final state density matrix for the process $A_{k-1:0}$ can be expressed by tracing over all of the intermediate operations, contracting to a coefficient expansion for the measured density matrices in the basis processes.

units $\mathcal{U}_1^k \in \mathcal{U}$ and $\mathcal{U}_2^k \in \mathcal{U}$. The restricted process tensor is then obtained using Equation (2). The set $\mathcal{U}$ contains 28 random unitaries, where the first $n$ elements $\mathcal{U}^{(n)}$ are used to reconstruct $T_{r^{3:0}}$. The remaining $28 - n$ elements are contracted with the reconstructed $T_{r^{3:0}}$ to obtain predictions $\rho_{ijk}^j$. We then compute the reconstruction fidelity

$$F_{ijk} := \left[\text{tr}\sqrt{\rho_{ijk}^j \rho_{ijk}^j} \sqrt{\rho_{ijk}^j \rho_{ijk}^j}^2 \right]^2 (3)$$

to gauge the accuracy of the prediction.

In theory, a minimal complete basis ($n = 10$) is all that is required for a restricted process tensor. In practice, however, we find that sampling error and, to a lesser extent, gate error, introduces inconsistencies in the linear equations described in Equation (2), amplifying reconstruction errors. The Moore-Penrose pseudoinverse (discussed in the methods section) finds the coefficients minimising the least-squares error between overdetermined and inconsistent linear equations. Consequently, adding in new basis elements will suppress the noise in the fidelities of prediction. We find a surprisingly large improvement. To further minimise bias in the noise, we also order our basis from least to most overlap with the rest of the set, as determined by the Hilbert-Schmidt inner product. This basis re-ordering improved predictive fidelity by 20%.

We summarise the average reconstruction fidelity between prediction and experiment of each basis in Figure 2a. The ‘Johannesburg (extended)’ experiment refers to process tensor experiments with idle time increased by a factor of 32. Meanwhile, ‘Johannesburg (Bell)’ is the result of creating a Bell pair, and then acting the unitaries on one half. The results both demonstrate the effects of basis size on process tensor performance, and showcase its ability to characterise processes. Adding in new basis elements offers substantial improvement in comparison to
III. BOUNDING MEMORY AND COMPARISON WITH GST

The impetus of the previous section was to demonstrate an experimentally verifiable method of characterising arbitrary dynamics. We now show that the above processes are indeed non-Markovian by lower bounding the memory in QIPs. We will then show that process tensors make more accurate predictions than comparable Markov models constructed using gate set tomography.

A. Estimating memory size

To fully account for the non-Markovianity in a system requires in-situ measurements, which break all correlations between the system and its environment, and represent a clean barrier to any past-future dependence [24]. Barring access to these, a restricted process tensor can only infer aspects of the non-Markovianity. Here, we introduce one such method to extract a lower bound on non-Markovianity.

Because the maximally depolarising channel

$$\mathcal{R}[\rho] = \mathbb{I}, \forall \rho$$ (4)

lies within the span of unitary operations, we can use it as an information barrier between time steps. A non-zero mutual information between the input operation and final measurement suggests information has travelled into the environment and returned after $\mathcal{R}$ has been applied [27]. Figure 3a illustrates this idea for the processes we consider here, where $\mathcal{R}$ takes either the first operation position, the second, or both. This tests the timing and duration of different memory effects.

The utility of the process tensor here is that it enables us to numerically search for the encoding and decoding operations which give the largest lower bound to non-Markovianity along different paths. Respectively, these are sets $\mathcal{E}$ and $\mathcal{D}$, the first of which contains two unitary operations applied with equal likelihood, and the second contains two orthogonal measurement effects. The quantities we compute are the conditional mutual information (CMI) for each case:

$$\arg\max_{\mathcal{E}, \mathcal{D}} I(E : D | E, U_1, \mathcal{R}, \mathcal{D}),$$

$$\arg\max_{\mathcal{E}, \mathcal{D}} I(E : D | E, \mathcal{R}, U_2, \mathcal{D}),$$

$$\arg\max_{\mathcal{E}, \mathcal{D}} I(E : D | E, \mathcal{R}, \mathcal{R}, \mathcal{D}),$$

where:

$$I(E : D) = \sum_{e \in \mathcal{E}} \sum_{d \in \mathcal{D}} p(E, D)(e, d) \log \left( \frac{p(E, D)(e, d)}{p(E)p(D|d)} \right).$$ (8)

For each experiment, we summarise the memory lower bound in Figure 3b. Note that we include an extra experiment ‘Valencia (H env)’, in which the neighbouring
FIG. 3. Discernment of the presence of memory in the quantum device. a The circuit depicting the process tensor. Quantum information can travel in and out of the system across one or many operations. Each gate is a place-holder for a larger set. b The maximum CMI, which is a conservative lower bound for non-Markovian memory, through $\mathcal{R}$ for each process tensor experiment, with 95% confidence intervals. This shows statistically significant non-zero memory in the device, which shows consistency in the timescale and the environmental interactions present.

B. Comparison with a Markov model

We performed two experiments under two different scenarios on the IBM Q Valencia 5-qubit quantum device. The first is identical to the process tensor experiments of Section II using the set $\mathcal{U}$. In addition, using GST we characterised all 28 unitary operations in $\mathcal{U}$, the 4 preparations in $\mathcal{P}$, as well as the the initial state and the final measurement. The estimates for each map were multiplied out to produce a Markovian prediction for the final density matrix. Both the process tensor and GST experiments were conducted first with neighbouring qubits initialised in the $|0\rangle$ state, and then again initialised in the $|+\rangle$ state. Figure 4 shows the distribution of the reconstruction fidelities for both the process tensor and GST. With a coherent environment, GST performs about $1.2\%$ worse. The process tensor tends to perform better in cases where the final state density matrices are more mixed, because this necessarily suppresses any directional bias in the noise.

We emphasise that our comparison of the outcomes of the two techniques is not framed competitively. Indeed, they are qualitatively different: while GST estimates the stationary maps of a given (presumed composable) gateset, the process tensor characterises all possible outcomes in a set process. Figure 4 observes the breakdown of a Markov model, and benchmarks the process tensor against the state-of-the-art as a complementary tool to describing processes.
IV. CONTROL IN THE PRESENCE OF MEMORY

In addition to non-Markovian characterisation and diagnostics, we now show that the process tensor can be a useful tool for quantum control. With a direct map from control operations to experimental outcomes, the data can be used to find which gates optimally output a desired state in a parametrised circuit. This outcome could harness external couplings to that end, using only local operations to manipulate them. Having already captured the process, the need for hybrid quantum-classical optimisation is eliminated. The desired result could be the most entangled state, highest fidelity equal superposition, or some member of a decoherence-free subspace. The procedure naturally accounts for any mitigating background, such as environmental noise or crosstalk. It is a matter of simple numerical optimisation to find the sequence of operations achieving the closest possible state to the one we desire: (i) Select an objective function $L$ which computes some quantum on the output density matrix, subject to the sequence $A_{k-1:0}$ of operations performed. (ii) Find:

$$\arg\min_{A_{k-1:0}} L \left( T_k^{k:0}[A_{k-1:0}] \right).$$

(9)

For unitaries, this is a straightforward minimisation over three parameters per time-step.

As an example, we first consider two neighbouring qubits initialised in the $|\pm\rangle$ state. Figure 5a shows the consequences of their natural coupling, extracted from the reconstructed two-qubit density matrix after some idle time. The results, which summarise negativity, mutual information, and state purities, show genuine entanglement between the two qubits. This form of dynamical behaviour will give rise to correlated errors in devices. After detection of a non-trivial interactions, we can use Equation (9) to decouple the qubits. So-called ‘bang-bang’ decoupling approaches have been thoroughly studied in the literature, but usually require a priori knowledge of the system-environment interaction Hamiltonian [32]. Using a one-step process tensor to form outcomes, our objective function is $2 - \gamma_1 - \gamma_2$, where $\gamma_i$ is the purity of the reduced state: $\gamma_i = \text{tr}(\rho_i^2)$. Performing the minimisation in Equation (9), we find the best decoupling operation. We then repeat the experiment of Figure 5a, but periodically apply the decoupling operation. This yields the results in Figure 5b, wherein the purities of each qubit have been significantly increased, and the entanglement over time suppressed.

We apply this same technique to exploit non-Markovianity for enhanced quantum control, inspired by the ideas in [15]. Arguments for the use of non-Markovianity as a resource are founded upon accessing Hilbert space trajectories otherwise unavailable with system control. We broaden our control set by using the process tensor to include non-unitarity, limited only by the strength and duration of the underlying interaction. We

FIG. 5. Entanglement, mutual information, and purities extracted from the two-qubit density matrix after being initialised in the $|++\rangle$ state. (a) The two neighbouring qubits are left idle. (b) As a simple demonstrative application of the process tensor, we use the construction from Equation (9) to find the optimal decoupling pulse. We periodically apply this gate to qubit 1. We see greatly improved coherences and almost complete elimination of entanglement between the two qubits, without actually characterising the nature of the interaction. (c) We use the process tensor to implement specific non-unitary gates. We plot the process fidelity as a function of the unitarity for two randomly chosen operations, according to the measure given in [31].
achieve this by constructing a single-step process tensor on one half of a pair of coupled qubits for a set of four preparation operations. Then, we use Equation (9) to find the parameters that produce final states closest to the ideal outputs of a randomly selected non-unitary operation, before applying the corresponding gate and performing quantum process tomography on it. The process fidelity of these non-unitary maps compared to their targets is plotted as a function of unitarity in Figure 5c. It reaches up to 97%, showing that we can extend the control capabilities of the device by using the process tensor and a nearby coupled qubit, which could be used to reconstruct the full process tensor. Critically, for this to work, we do not need to perform control operations on the neighbouring qubit beyond its initialisation.

This simple framework is widely applicable to many forms of quantum control. In particular, it allows for either mitigating or controlling non-Markovian noise without first understanding it at a microscopic level. Broadly, the user need only specify a desired outcome, without studying the means to achieve it.

V. CONCLUSION AND OUTLOOK

In this paper, we have bridged the gap from a theoretical framework of non-Markovian dynamics to an experimental method which verifiably offers non-Markovian diagnostics and control. First, we demonstrated a high fidelity non-Markovian characterisation technique over a range of devices. We used this to bound the non-Markovian memory present. Then, using the reconstructed process tensor, we demonstrated operationally tractable control techniques to decouple the system qubit from its neighbour, as well as applying well-characterised intermediary non-unitary operations on the system. These methods pave the way to mitigate non-Markovian noise and streamline the performance of quantum devices.

Like many tomographic techniques, the construction of the process tensor scales unfavourably in both the number of time-steps and number of qubits. However, for processes with finite Markov order it is possible to reconstruct a primitive building block, from which the whole process can be inferred [27]. One immediate future avenue is complete process characterisation, as suggested in the previous section, which will offer better benchmark for the length of the memory. Although we found success with the use of an overcomplete basis, it would likely be fruitful to explore coupling smaller bases with conventional denoising techniques, the use of a mutually unbiased unitary basis [33], or machine learning reconstruction methods [34]. Much like with the study of many-body entanglement, there is ample room to reduce experimental overhead with some well-placed physical assumptions.

VI. ACKNOWLEDGMENTS

This work was supported by the University of Melbourne through the establishment of an IBM Q Network Hub at the University. GALW is supported by an Australian Government Research Training Program Scholarship. CDH is supported through a Laby Foundation grant at The University of Melbourne. KM is supported through Australian Research Council Future Fellowship FT160100073.

REFERENCES

[1] J. F. Haase, P. J. Vetter, T. Unden, A. Smirne, J. Rosskopf, B. Naydenov, F. Jelezko, M. B. Plenio, and S. F. Huelga, Phys. Rev. Lett. 121, 060401 (2018), arXiv:1802.00819.
[2] B. Pokharel, N. Anand, B. Fortman, and D. A. Lidar, Phys. Rev. Lett. 121, 220502 (2018), arXiv:1807.08768.
[3] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout, (2019), arXiv:1908.09855.
[4] A. Rivas, S. F. Huelga, and M. B. Plenio, Rep. Prog. Phys. 77, 094001 (2014), arXiv:1405.0303.
[5] L. Li, M. J. Hall, and H. M. Wiseman, Phys. Rep. 759, 1 (2018), arXiv:1712.08879.
[6] F. Sakuldee, S. Milz, F. A. Pollock, and K. Modi, J.Phys. A 51, 414014 (2018), arXiv:1802.03190.
[7] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. 78, 390 (1997), arXiv:quant-ph/9611013.
[8] R. Blume-Kohout, J. Gamble, E. Nielsen, J. Mizrahi, J. Sterk, and P. Maunz, (2013), arXiv:1310.4492.
[9] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, Nat. Commun. 8, 1 (2017), arXiv:1605.07674.
[10] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, Phys. Rev. A 87, 062119 (2013), arXiv:1211.0322.
[11] G. A. L. White, C. D. Hill, and L. C. L. Hollenberg, (2019), arXiv:1911.12096.
[12] P. Pechukas, Phys. Rev. Lett. 73, 1060 (1994).
[13] N. H. Nickerson and B. J. Brown, Quantum 3, 131 (2019), arXiv:1712.00502.
[14] N. Maskara, A. Kubica, and T. Joehym-O’Connor, Phys. Rev. A 99, 52351 (2019), arXiv:1802.08680.
[15] D. M. Reich, N. Katz, and C. P. Koch, Sci. Rep. 5, 1 (2015), arXiv:1409.7497.
[16] C. Pineda, T. Gorin, D. Davalos, D. A. Wisniacki, and I. García-Mata, Phys. Rev. A 93, 1 (2016), arXiv:1505.03505.
[17] N. P. Kumar, S. Banerjee, and C. M. Chandrashekar, Sci. Rep. 8, 1 (2018), arXiv:1802.05478.
VII. METHODS

Process Tensor Experiments

Here, we discuss the construction of a multi-time process tensor both in particular to the experiments conducted in this work, and more generally with respect to a greater set of controls. The process tensor constructed was over three time-steps of varying sizes. The experimental steps for this are as follows:

1. Initialise the qubit in state $|0\rangle$
2. Apply $P^k \in \mathcal{P} = \{H, S \cdot H, \mathbb{I}, X\}$
3. Apply $U^j \in \mathcal{U}$
4. Leave some amount of time.
5. Apply $U^k \in \mathcal{U}$.
6. Leave idle.
7. Repeat this sequence three times for the three QST basis measurements required.
8. Store this density matrix as $\rho_{3}^{ijk}$
9. Repeat this for all combinations of the elements of $\mathcal{P}$ and $\mathcal{U}$ in each slot.

For our experiments, this is a total of $(4 \times 28 \times 28) \times 3 = 9408$ experiments. Interleaved between each operation is idle time equivalent to a single gate. The circuit diagram for these experiments is given in Figure 6. We ran these at 1600 shots each with the exception of ‘Boebblingen’, which had 4096 shots. This data was then partitioned into process tensor construction, and experimental verification. The former consists of the construction of a basis-$n$ process tensor, which used the first $4 \times n \times n$ control sequences to form a basis. We then used the remaining $4 \times (28-n) \times (28-n)$ sequences which lie outside the basis set as verification density matrices for the process tensor predictions. It is worth noting that action of the process tensor is insensitive to state preparation and measurement (SPAM) errors. Any initial state or final measurement error channel are absorbed into the definition of the process tensor, and the expansion remains the same.

The unitary basis was constructed with a randomly generated set of 28 ordered unitary matrices using the scipy.stats.unitary_group.rvs() function. We parametrise these gates using the standard qiskit unitary parametrisation:

$$U(\theta, \phi, \lambda) = \begin{pmatrix} \cos(\theta/2) & -e^{i\lambda} \sin(\theta/2) \\ e^{i\phi} \sin(\theta/2) & e^{i\lambda+i\phi} \cos(\theta/2) \end{pmatrix}. \quad (10)$$

![FIG. 6. Circuit diagram depicting the generic experiment required to construct the two step process tensor. Each gate represents an element from either the preparation set $\mathcal{P}$ or the more general unitary basis set $\mathcal{U}$. The identity gates represent idle time which we allow to vary. Finally, measurements in three bases are made for QST.](image-url)
On the IBM superconducting devices, these so-called u3 gates are implemented in two physical pulses corresponding to rotations around the $x-$axis, and three frame shifts corresponding to rotations around the $z-$axis [35, 36]. Explicitly,

$$U(\theta, \phi, \lambda) = R_x(\phi + 3\pi)R_x(\pi/2)R_z(\theta + \pi)R_x(\pi/2)R_z(\lambda).$$  

(11)

Consequently, the physical duration of each u3 gate is independent of the $\theta, \phi, \lambda$ parameters – approximately 72 ns. We then leave the system idle for a duration of one u3 gate. Following this is one more u3 gate, an identical wait time, and then each of three basis measurements in $X, Y,$ and $Z$ Pauli bases required to reconstruct the output density matrix. The maximum likelihood method introduced in [37] is then used to find the closest physical output density matrix. The maximum likelihood method corresponds to the two-step process tensor using a basis of $n$ operations is given as

$$\mathcal{T}^{3:0} = \sum_{i=1}^{4} \sum_{j=1}^{n} \sum_{k=1}^{n} \left( D_{ij}^0 \otimes \Delta_k^0 \otimes \Delta_k^0 \right) \otimes \rho_{jik}^{(3:0)}$$  

(15)

Where the $\{D_{ij}^0\}$ are dual to the preparation operations $\mathcal{P}$, and the $\{\Delta_k^0\} = \{\Delta_k^0\}$ are dual to the circuit operations $U^{(n)}$. Sampling error in the final state density matrix, as well as error in the gates themselves will collectively introduce inconsistencies in the set of linear equations described by Equation (2). The error becomes significant if the basis is biased in a particular direction of superoperator space. Originally, our minimal complete basis – which had been randomly selected – produced a reconstruction fidelity of around 70%. To mitigate this error, we re-ordered our basis according to the least to most overlap with the remainder of the set according to the Hilbert-Schmidt inner product. For the first ten elements, this overlap was $[0.0336, 0.0400, 0.0438, 0.0489, 0.0505, 0.0518, 0.0594, 0.0600, 0.0619, 0.0621]$. After re-ordering, the reconstruction fidelity of the minimal complete basis improved to around 95%. This effect was only discovered after the completion of the experiments, at which point it was too late to change the experimental basis itself. In future, a better course of action would be to examine the selection of a set of mutually unbiased unitary operators.

In general, the only positive dual operators are entanglement-breaking channels. With a restricted basis, the process tensor constructed here is not unit trace, nor is it a positive operator. Physically meaningful quantities can only be extracted from its action on the restricted basis, rather than from the explicit form given in Equation (15). For this reason, we keep the emphasis of the process tensor in this work on its ‘actions’ rather than on information that can be gleaned from the object itself. Note that the expansion coefficients are calculated in the contraction of the operation with the process tensor. We discuss this explicitly below.

### Control basis and process reconstruction

An arbitrary $A_j$, at time step $j$, on a system of dimension $d_S$ may be decomposed into a linear expansion of some ordered basis $\{B_{ij}^j\}$ such that

$$A_j = \sum_{\mu_j=1}^{d_S} \alpha_{ij}^{\mu_j} B_{ij}^j.$$  

(12)

A sequence of (independently chosen) control operations may be written with a tensor product structure $A_{k-1:0} = \bigotimes_{j=0}^{k-1} A_j$, for which each constituent map can be further decomposed into the chosen basis. The complete spatio-temporal basis of operations is then given by

$$\begin{align*}
\{B_{\vec{\mu}_{k-1:0}}\} = \bigotimes_{j=0}^{k-1} \{B_{ij}^j\}_{\vec{\mu}=(1,1,\ldots,1)},
\end{align*}$$  

(13)

where $\vec{\mu} = (\mu_0, \mu_1, \cdots, \mu_{k-1})$ is a $k-$dimensional vector of index elements, each taking values between 1 and $d_S^j$. That is, it is the set with cardinality $d_S^k$ of all combinations of the $k$ tensor products of each member of $\{B_{ij}^j\}$ at each time step. Measuring the output state $\rho_{\vec{\mu}}$, for each of these basis operations is sufficient to construct the process tensor. We signify the matrix form of the process tensor $\mathcal{T}$ with a caret: $\hat{T}$

$$\hat{T}_{k:0} = \sum_{\vec{\mu}} \Delta_{k-1:0}^{\vec{\mu}} \otimes \rho_{\vec{\mu}}^{(k:0)},$$  

(14)

where the set $\{\Delta_{k-1:0}^{\vec{\mu}}\}$ is known as the dual set to $\{B_{k-1:0}^{\vec{\mu}}\}$, satisfying $\text{tr} \left[ \hat{B}_{k-1:0}^{\vec{\mu}} (\Delta_{k-1:0}^{\vec{\mu}})^T \right] = \delta_{\vec{\mu}\vec{\mu}}$. This dual set can be easily computed for any linearly independent set of vectors. To be explicit, the matrix form for

### Construction of a dual set

The procedure to construct the dual operators is as follows: for a complete set of linearly independent operations $\{B^j\}$ whose matrix forms are $\{\mathcal{B}^j\}$, we can compile the basis into a single matrix $\mathcal{B}$. Write each $\mathcal{B}^j = \sum_{j} \mathcal{B}^j \Gamma_j$, where $\{\Gamma_j\}$ form a Hermitian, self-dual, linearly-independent basis satisfying $\text{tr}[\Gamma_i \Gamma_k] = \delta_{ik}$. In our case, we select $\{\Gamma_j\}$ to be the standard basis, meaning that the $k$th column of the matrix $\mathcal{B} = \sum_{ij} \mathcal{B}_{ij} |i\rangle \langle j|$ is $\mathcal{B}^k$ flattened into a 1D vector. Because the $\{\mathcal{B}^j\}$ are linearly independent, $\mathcal{B}$ is invertible. Let the matrix $\mathcal{F}^\dagger = \mathcal{B}^{-1}$ such that $\mathcal{B} \cdot \mathcal{F}^\dagger = I$. This means that the rows of $\mathcal{F}^\dagger$ are orthogonal to the rows of $\mathcal{B}$. The dual matrices can then be defined as $\Delta^j = \sum_{ij} f_{ij} \Gamma_j$, en-
suring that $\text{tr}[\hat{B}^i \Delta^j] = \delta_{ij}$. Note that in this work, our basis is restricted to the sub-manifold of unitary matrices. This means that the dimension $d$ of the space is less than the order $n$ of the matrices. Therefore we construct $\mathcal{F}^i$ as the Moore-Penrose or the right inverse of $\mathcal{B}$. We also primarily operate in an over-complete setting, where the number of basis operations is greater than the dimension of the space, meaning that they cannot all be linearly independent. Here, we relax the duality condition $\text{tr}[\hat{B}^i \Delta^j] = \delta_{ij}$, but retain $\sum_i \Delta^i = I$ to ensure that the expansion of any operation within the basis is complete. The over-completeness technique is necessary for a high fidelity reconstruction, owing to the sensitivity of the matrix pseudoinverse to shot-noise.

**Contracting an Operation**

The expansion coefficients discussed are useful in conceptual discussions of the process tensor, but in practice these are not directly computed. Instead, the action of the process tensor on a sequence of operations is found by projecting the process tensor onto the Choi state of this sequence (up to a transpose). Below, we explicitly step through this computation.

\[
\mathcal{T}^{k:0} [\mathbf{A}_{k-1:0}] = \text{tr}_{\rm in} \left[ \left( \bigotimes_{i=0}^{k-1} \hat{A}_i \otimes \mathbb{I} \right) \left( \sum_{l} (\Delta_{k-1:0}^l)^T \otimes \rho_{k}^l \right) \right] \\
= \text{tr}_{\rm in} \left[ \left( \bigotimes_{i=0}^{k-1} \hat{A}_i \otimes \mathbb{I} \right) \sum_{l} \left( \Delta_{k-1:0}^l \right)^T \otimes \rho_{k}^l \right] \\
= \text{tr}_{\rm in} \left[ \sum_{l} \alpha_l^\mu \left( \bigotimes_{i,j=0}^{k-1} \left( \Delta_{i,j}^l \right)^T \otimes \rho_{k}^l \right) \right] \\
= \sum_{l} \alpha_l^\mu \prod_{i,j=0}^{k-1} \text{tr} \left[ \left( \Delta_{i,j}^l \right)^T \rho_{k}^l \right] \\
= \sum_{l} \alpha_l^\mu \text{tr} \left[ \delta_{\mu} \rho_{k}^l \right] \\
= \frac{\text{tr}}{\rho_k(\mathbf{A}_{k-1:0})}.
\]

The direct calculation of each expansion coefficient is therefore given by

\[
\alpha_l^\mu = \text{tr} \left[ \hat{A}_{k-1:0} (\Delta_{k-1:0}^l)^T \right] \\
= \text{tr} \left[ \bigotimes_{i=0}^{k-1} \hat{A}_i (\Delta_{i:0}^l)^T \right] \\
= \prod_{i=0}^{k-1} \text{tr} \left[ \hat{A}_i (\Delta_{i:0}^l)^T \right] = \prod_{i=0}^{k-1} \alpha_i^\mu.
\]

**Bounding Memory**

In Section III A, we estimate a lower bound for the memory present in the devices. This is accomplished with the contraction of different encoding operations with the process tensor and forming predictions for the output in this way. For the case where $\mathcal{R}$ is contracted in position one, the explicit steps are as follows:

1. Pick $\mathcal{E}^0, \mathcal{E}^1 \in U(2)$
2. Pick $p_{e_0}$ and $p_{e_1}$ s.t. $p_{a_i} \in [0, 1]$ and $p_{e_0} + p_{e_1} = 1$ (in this experiment, we set $p_{e_0} = p_{e_1} = 0.5$).
3. Pick $\mathcal{D} \in U(2)$
4. Pick $\mathcal{V} \in U(2)$
5. Compute the 4 values of $p_{(E,D)}(e_i, d_j)$ by collecting the density matrix $\rho' = \mathcal{T}^{3:0}[\mathcal{E}'^i, \mathcal{V}, \mathcal{R}]$ and then setting $p_{(E,D)}(e_i, d_j) = p_{e_i} \cdot \text{Tr}(|j\rangle \langle j'| \cdot \mathcal{D} \rho_{i} \mathcal{D}^\dagger)$
6. Compute the marginal distributions: $p_E(e_i) = \sum_j p_{(E,D)}(e_i, d_j)$ and $p_D(d_j) = \sum_i p_{(E,D)}(e_i, d_j)$
7. Finally, compute $I(E : D)$

These steps are framed as an optimisation problem where $\mathcal{E}, \mathcal{D}$, and $\mathcal{V}$ are chosen such that $I(E : D)$ is maximised. Implicit in this exercise is the assumption that operations outside the preparation set achieve the same reconstruction fidelity as the latter steps shown in Figure 2. Although we did not examine this assumption for every machine, in Figure 7 we construct a four time process
tensor on Valencia using the basis $U^{(4)} \otimes U^{(n)} \otimes U^{(n)}$. We then compare the reconstruction infidelity from predictions made by the process tensor: firstly, compared to gate sequences where the preparation operation lay inside the basis set (with $U^1$ and $U^0$ outside), and secondly compared to gate sequences where the preparation operations were the next four elements of $U$. We find these two collections to be identical within error bars for all basis sizes 10 and above. Given that Valencia had the worst reconstruction fidelity of the machines, we view this as sufficient evidence that the assumption is valid across all machines.

Gate set tomography comparison

The GST experiments conducted in Section III B were completed using the pyGSTi quantum processor performance package [38, 39]. Following the procedures outlined in the documentation, with background given in [9, 40], we characterised the 28 random unitaries as well as the 4 preparation gates in 8 groups of 4 gates. The software package designates the circuits required, and carries out the maximum likelihood reconstruction of the gates with the constraint of complete positivity and trace preservation. The gate sequences were repeated in powers of 2: 1, 2, 4, 8, 16, and 32 times. The process tensor and GST experiments were conducted in the one calibration period for the device in a window of approximately 5 hours.

Adaptive control methods

Here, we more explicitly discuss our adaptive control methods using the process tensor. In each case, the system qubit and its neighbour were both initialised in the $|+\rangle$ state. We sought to use the process tensor to control the always-on interaction between the two qubits without actually learning it. The circuit diagrams describing both experiments are in Figure 8.

1. Decoupling

In the first scenario operations only on qubit 1, we construct a single step process tensor with a size-24 basis, 256 ns of idle time on either side, and two-qubit state tomography at the end. Altogether, this is $24 \times 9 = 216$ experiments. Strictly speaking only single qubit state tomography is required for the purpose of decoupling one qubit, however we created a mapping to the two-qubit output in order to specifically best show these two qubits decoupled. With the intermediate operation parametrised as in Equation (10), the minimisation performed was:

$$\arg \min_{\theta, \phi, \lambda} 2 - \gamma_1 - \gamma_2$$

(20)

where $\gamma_i$ is the purity of the $i$th reduced density matrix produced by the process tensor. The total density matrix is $T^{1:0}[U(\theta, \phi, \lambda)]$. The decoupling operation found was

$$\begin{pmatrix}
0.0051 \\
\sqrt[188]{e^{0.188}} \\
0.0051 \cdot e^{i(2.257)}
\end{pmatrix}$$

In Figure 5b, we apply this operation to the system

2. Non-unitary operations

For the purpose of implementing our own chosen non-unitary operations, we created a one-step basis-24 process tensor on a single qubit whose neighbour was in the $|+\rangle$ state: approximately 800 ns of idle time after $P$ preparations, followed by $U^{(24)}$, followed by another 800 ns and then QST. We then generated a set of random non-unitary operations with unitarity ranging from 1/3 to 1.0. These are denoted by $N(\alpha, \eta)$, where

$$N(\alpha, \eta) = \sqrt{\eta}E(\alpha) + \sqrt{1-\eta}Y E(\alpha),$$

and $E(\alpha) = (R_X(\alpha)R_Y(\alpha)R_Z(\alpha))$. (21)

The two operations shown in Figure 5c are two different randomly generated values for $\alpha$. The unitarity of the operations is then varied by varying $\eta$ from 0 to 0.5 in the above equation. Using these operations as a target map, we numerically found the gate parameters minimising the

FIG. 8. Circuit diagrams for each of the application experiments. a For the decoupling of two qubits, we allow evolution time before and after the process tensor. The distinction between here and other process tensor experiments that we conducted is that we map from the operation on one qubit to the two-qubit density matrix, rather than solely single qubits. b To enact non-unitary gates of our choosing, we conduct a similar experiment. This time, however, there are four basis preparation operations to begin with, and QST only on the single qubit. This is so that we can optimise the action of the gate over a complete basis of inputs.
trace distance between the target outputs of the non-unitary map and the process tensor predictions for a set of four inputs. That is, we applied the minimisation:

$$\arg\min_{\theta, \phi, \lambda} \frac{1}{2} (||\tau_X - \rho_X||_1 + ||\tau_Y - \rho_Y||_1 + ||\tau_Z - \rho_Z||_1 + ||\tau_{I-Z} - \rho_{I-Z}||_1),$$

(22)

where each $\rho_j$ is the ideal output of $\mathcal{N}(\alpha, \eta)$ acting on the $X, Y, Z,$ and $\mathbb{I} - Z$ eigenvectors, and each $\tau_j$ is the $T_{2:0}[\mathcal{P}_j, U(\theta, \phi, \lambda)]$ predicted density matrices. Then, using the optimal values of $\theta, \phi,$ and $\lambda$, we performed quantum process tomography and compared the process tensor of our implementation $\mathcal{N}'(\alpha, \eta)$ with the ideal $\mathcal{N}(\alpha, \eta)$.

All experimental data is available upon reasonable request.