Simple Mitigation of Global Depolarizing Errors in Quantum Simulations

Joseph Vovrosh*, 1 Kiran E. Khosla, 1 Sean Greenaway, 1 Christopher Self, 1 M. S. Kim, 1 and Johannes Knolle 2, 3, 1

1Blackett Laboratory, Imperial College London, London SW7 2AZ, United Kingdom
2Department of Physics TQM, Technische Universität München, James-Franck-Straße 1, D-85748 Garching, Germany
3Munich Center for Quantum Science and Technology (MCQST), 80799 Munich, Germany

To get the best possible results from current quantum devices error mitigation is essential. In this work we present a simple but effective error mitigation technique based on the assumption that noise in a deep quantum circuit is well described by global depolarizing error channels. By measuring the errors directly on the device, we use an error model ansatz to infer error-free results from noisy data. We highlight the effectiveness of our mitigation via two examples of recent interest in quantum many-body physics: entanglement measurements and real time dynamics of confinement in quantum spin chains. Our technique enables us to get quantitative results from the IBM quantum computers showing signatures of confinement, i.e. we are able to extract the meson masses of the confined excitations which were previously out of reach. Additionally, we show the applicability of this mitigation protocol in a wider setting with numerical simulations of more general tasks using a realistic error model. Our protocol is device-independent, simply implementable and leads to large improvements in results if the global errors are well described by depolarization.

INTRODUCTION

Quantum computers are becoming large enough (∼ 50 – 100 qubits [1]) to, in principle, allow demonstrations of their ‘Quantum Advantage’ [2, 3]. However, the actual amount of entanglement that can be generated in current devices is constrained by noise and errors, limiting their ability to solve complex problems such as quantum simulation. To address this, various error mitigation strategies have recently been developed to counteract noise and boost the fidelity of experimental results.

Error mitigation differs from fault tolerance. Fault tolerant quantum computers will eventually be able to suppress errors by encoding quantum information in a redundantly large number of qubits as error correcting codes [4, 5]. In this way they will be able to execute arbitrarily deep circuits by the repeated application of active error corrections. Unfortunately, these encodings cannot be used in current devices as they require smaller hardware errors and larger numbers of qubits than are currently available. In contrast, error mitigation strategies are applied to unencoded physical qubits. Rather than actively correcting errors, they aim to estimate what the effect of the error was and infer the error-free result. This current stage in the development of quantum computers has been dubbed the noisy intermediate-scale quantum (NISQ) era [1] and is expected to last for the foreseeable future.

In the past few years, error mitigation has been a thriving research direction as more and more quantum devices become available. Error mitigation strategies typically address measurement errors [7–15] or the algorithms and gates employed for digital quantum simulation [16–23]. While some of these techniques have proven useful for certain quantum algorithms, their general success is hampered by the fact that they either rely on a high level of control of the quantum device itself, or are specifically designed for a given quantum simulation problem, i.e. exploiting specific symmetries [24].

One promising direction is to employ machine learning algorithms [25–27] for error mitigation. Generally speaking, these methods train classical computers to predict the error seen in a quantum device and use the results to infer the error-free quantum simulations. While successful, these methods require a large increase in the classical computational overhead and are somewhat uncontrolled. The latter drawback is also true for popular protocols based on the idea of increas-
ing errors in the device systematically and then extrapolating back to the zero error case [21, 28–34]. In general, how to tune the error rates varies from device to device and reliable fitting requires expert knowledge of the specific hardware.

Here, we propose a new protocol for gate error mitigation combining a raft of desirable features: it is easily implementable on any quantum device with little increase in workload, it is well rooted in a mathematical description of the errors, and it is suitable for any quantum algorithm of interest. Furthermore, as we show via specific examples, it can lead to large improvements in the performance of quantum devices, for example see Fig. 1 for Rényi entropy results.

The paper is organised as follows: First, we derive an ansatz for the density matrix resulting from the action of a noisy quantum simulator with global depolarizing errors. Following this, we explain how this ansatz can be used for a general error mitigation protocol. We then demonstrate its effectiveness by studying real time dynamics of the transverse field Ising model (TFIM) with a longitudinal field. In that context, some of the authors recently showed that signatures of confinement and entanglement spreading can be observed on the IBM quantum computer [35]. With our new error mitigation protocol we are able to measure the meson masses of confined mesons, for example see Fig. 1 for Rényi entropy results. The paper is organised as follows: First, we derive an ansatz for the density matrix resulting from the action of a noisy quantum simulator with global depolarizing errors. Following this, we explain how this ansatz can be used for a general error mitigation protocol. We then demonstrate its effectiveness by studying real time dynamics of the transverse field Ising model (TFIM) with a longitudinal field. In that context, some of the authors recently showed that signatures of confinement and entanglement spreading can be observed on the IBM quantum computer [35]. With our new error mitigation protocol we are able to measure the meson masses of confined mesons, for example see Fig. 1 for Rényi entropy results.

ERROR MITIGATION PROTOCOL

Error channels can be conveniently modelled through Kraus operators [36] which define a completely positive map on the density matrix

$$\rho \rightarrow \sum_i K_i \rho K_i^\dagger \text{ such that } \sum_i K_i^\dagger K_i = 1.$$ (1)

For a single qubit, choosing four Kraus operators to be proportional to the Pauli operators (identity included) defines a depolarizing channel. The proportionality constants are related to the individual error probabilities and must satisfy the identity constraint to preserve the trace of $\rho$.

Here, we concentrate on such depolarizing errors which are inevitably present in any digital quantum simulator platform and can be treated without much specific knowledge about the device performance (which also fluctuates over time). Moreover, our focus on depolarizing errors has the advantage of being treatable mathematically in a controlled way as detailed below. In fact, given a deep enough quantum circuit (with sufficient qubits) for self-averaging of incoherent errors to occur, this depolarizing error model is a good approximation to the physical errors in a device (more details are given in the supplementary material). Last but not least, the large improvement in the quality of results on the IBM device motivate our choice of depolarizing errors a posteriori.

An n-single qubit depolarizing error channel can be modelled via

$$\mathcal{E}^{\otimes n}(\rho) = (1 - p)\rho + \sum_{\alpha \in \{x,y,z\}} \sum_{j=1}^n (1 - p)^{n-1} \frac{1}{3} \sigma_\alpha^j \rho \sigma_\alpha^j + \ldots$$ (2)

where $\mathcal{E}$ is the error channel, $p$ is the probability of an error occurring for each qubit (assumed to be equal for each qubit, and for each Pauli error) and ‘...’ indicates higher-order terms corresponding to errors on multiple qubits [36]. One important feature of this mathematical formulation is that the second term, which describes the depolarizing error, commutes with any unitary operator. Consequently, the error on the $i^{th}$ qubit in a quantum circuit with purely depolarizing errors is

$$\mathcal{E}_i(\rho) = (1 - p_i)\rho + p_i \text{Tr}_i[\rho] \otimes \frac{I_2}{2}$$ (3)

in which $\rho$ is the density matrix, $I_2/2$ is the maximally mixed (i.e. completely depolarized) state for the $i^{th}$ qubit, $\text{Tr}_i$ is the partial trace over the $i^{th}$ qubit and $p_i$ is the error on the $i^{th}$ qubit.

Instead of dealing with all combinations of single qubit errors, we approximate the total error channel of Eq. (2), under the assumption of symmetric depolarization Eq. (3), as an effective depolarizing channel on the entire quantum state

$$\rho = (1 - p_{\text{tot}})\rho_{\text{exact}} + p_{\text{tot}} \frac{\mathcal{I}^{\otimes n}}{2^n}$$ (4)

where the effective total error probability is $p_{\text{tot}}$. In principle $p_{\text{tot}}$ is well approximated by $\prod_i (1 - p_i)$, however, we do not make that identification here. We will soon show that $p_{\text{tot}}$ can be measured directly on the device. The simplicity of this ansatz allows it to be easily calculated. It has already been shown to be useful when mitigating measurement error [37], however we later demonstrate that it is powerful tool for mitigating global depolarizing errors which do not themselves originate from local depolarizing errors. Furthermore, we later numerically demonstrate that this ansatz does not rely on the assumption of single- and two-qubit depolarizing errors.

The many partial traces over single qubits, which would have conserved some coherence in the remaining qubits, have been replaced by the maximally mixed state $\mathcal{I}^{\otimes n}/2^n$ over the global quantum state, destroying all coherence. We stress that even though this may not be a good approximation for a single layer of qubit errors, it becomes a reasonable approximation for the error channel of a many layer, many qubit circuit. Eq. (4) is our basic ansatz for an effective error model after a many-layered unitary circuit and $\rho_{\text{exact}}$ is the exact density operator without noise (see Supplementary Material for more details). With this ansatz one can analytically calculate the effect of errors on a measured observable, $\hat{O}$, via

$$\langle \hat{O} \rangle = (1 - p_{\text{tot}}) \text{Tr}[\hat{O}\rho_{\text{exact}}] + \frac{p_{\text{tot}}}{2^n} \text{Tr}[\hat{O}]$$ (5)
Our approach estimates $p_{\text{tot}}$ in order to apply error mitigation. We propose and test two approaches for finding $p_{\text{tot}}$, the first based on estimating the purity of the final state and the second based on studying specific observables. To estimate the purity, we employ the recent protocol for obtaining the trace of the reduced density matrix squared $\text{Tr}[\rho^2]$ via randomised measurements [38, 39], where $A$ is a subspace of the full density matrix. This randomised measurement scheme has been successfully implemented in trapped ion quantum simulators [40] and recently by some of us on the IBM quantum computer [35]. We stress the present mitigation is a far simpler task compared to inverting the quantum error channel to tomographically reconstruct the error free quantum state $\rho$.

As current quantum devices initialise systems in pure states that are then manipulated with unitary transformations, $\text{Tr}[\rho^2]$ over the full Hilbert space should lead to a result that is identically one. However, since the noisy implementation of quantum circuits will in general deviate from unitarity, after a given quantum circuit is run on a quantum processor this will generally not be the case. Instead, with Eq. (4) we expect that

$$\text{Tr}[\rho^2] = (1 - p_{\text{tot}}^2) + \frac{p_{\text{tot}}(1 - p_{\text{tot}})}{2^{n-1}} + \frac{p_{\text{tot}}^2}{2^n} \tag{6}$$

Using the fact that $\rho_{\text{exact}}$ is pure. Now, given that the left hand side, $\text{Tr}[\rho^2]$, can be measured directly on the device [35, 40], this quadratic equation can be solved to obtain the total error $p_{\text{tot}}$ [37, 41]. We stress that correlated errors in quantum circuits do not increase entropy and thus $p_{\text{tot}}$ obtained via this method should really be understood as the global depolarizing error probability. While this mitigation approach does not address coherent errors, it could additionally be combined with other techniques such as twirling [42–45].

Therefore, with $p_{\text{tot}}$ extracted and $\langle O \rangle$ measured the only unknown quantity in Eq. (5) is the desired error-free observable $\text{Tr}[O_{\text{exact}}]$. Note, we assume that $\text{Tr}[O]$ can be calculated which for most practical cases should be the case, e.g. see our application examples below.

Putting all steps together, we finally obtain our general error mitigation protocol:

1. Prepare the quantum state of interest by running a quantum circuit and measure $\text{Tr}[\rho^2]$, e.g. via randomized measurements [35, 39, 40].
2. Use the results to obtain values for $p_{\text{tot}}$ via Eq. (6).
3. Prepare the quantum state again and measure the desired observable $\langle O \rangle$. [46]
4. Use Eq. (5) with the measured value of $p_{\text{tot}}$ to obtain the desired $\text{Tr}[O_{\text{exact}}]$.

An alternative approach to estimating $p_{\text{tot}}$ is to consider specific observables whose expectation values are known. For example, consider the time dynamics of a system in which our quantum circuit approximates the time evolution operator $U(t) = \exp\{-iHt\}$ where we wish to measure $\langle O(t) \rangle$ for a range of times. We can tune the circuit such that $t \cdot (E_{\text{max}}) = \epsilon \ll 1$ where $E_{\text{max}}$ is the largest eigenvalue of the Hamiltonian (shifted so $E_{\text{min}} = 0$), so that our quantum circuit now approximates the identity operation. Assuming that $\langle O(t = 0) \rangle$ is known, we can use the measurements from the quantum device and Eq. (5) to obtain $p_{\text{tot}}$. Note, for infinite dimensional Hamiltonian with unbounded eigenvalues, the approximation is slightly more subtle, but does not apply to qubits. We show that, while this method is more efficient, it does not discriminate between coherent and incoherent errors in the purity.

Our protocol can be applied to essentially any quantum circuit and quantum simulation device. In the following, we choose a representative example from condensed matter physics as a first application. We showcase the effectiveness of our technique by presenting previously unobtainable results for confinement and entanglement dynamics in spin chains.

**APPLICATION TO SPIN CHAIN CONFINEMENT**

An ideal testing ground for NISQ devices is that of quench dynamics in spin-$\frac{1}{2}$ systems. A global quantum quench is a sudden change to the systems Hamiltonian, which induces non-equilibrium dynamics. Already one dimensional spin chains can show a wide variety of physical phenomena of interest, for example confinement of domain wall excitations [35, 47, 48], quantum many-body scars [49–53], or novel fracton excitations [54–56]. All of these show up in the time evolution which is challenging to simulate on classical computers as the Hilbert space grows exponentially with the number of spins $n$. As spin-$\frac{1}{2}$ systems directly map onto physical qubits, quantum computers are ideally suited for studying the rich physics of spin chains. Recently, first digital quantum simulation results have been reported [24, 35, 57–60] but in order to obtain results which are out of reach by classical simulations and to probe non-trivial quantum many body physics better error mitigation techniques are needed.

We concentrate on the one-dimensional TFIM with an additional longitudinal field given by the following Hamiltonian

$$H = -J \left[ \sum_i \sigma_i^z \sigma_{i+1}^z + h_x \sum_i \sigma_i^x + h_z \sum_i \sigma_i^z \right], \tag{7}$$

where $J$ is the Ising exchange of nearest neighbour spins $\sigma_i$ and $h_{x/z}$ are the relative strengths of the transverse and longitudinal fields respectively. For $h_z = 0$ the TFIM can be solved exactly via Jordan-Wigner transformation and its fermionic excitations are related to free domain wall motion. When turning on the longitudinal field, $h_z \neq 0$, a confining potential between these fermions is introduced. The attraction between fermions grows linearly with their separation, reminiscent of quark confinement in QCD. The result is the formation of ‘mesonic’ bound states of domain wall excitations.
FIG. 2. Results from the IBM machine before and after mitigation. Quench dynamics of local magnetization from Trotterized time evolution of the TFIM with longitudinal field before and after error mitigation. Here, $J = 1$, $h_x = 0.5$, $h_z = 0.75$ and $n = 7$. Data is shown for, $N_T$, the number of trotter steps $N_T = 5, 6$. Results go from qualitative agreement to quantitative agreement. After fitting a cosine function (dashed green) to the mitigated data (green dots) the dominant frequency is clearly captured by the IBM device, for details see Supplementary Material. Note, more details of the circuit composition of the evolution operator can be found in [35]

The confining potential between fermions induces non-ergodic behaviour [61], which manifests itself through persistent oscillations in the local magnetization, $\langle \sigma_z^\alpha \rangle$ ($\alpha \in \{ x, y, z \}$) and a slowing down of the entanglement spreading.

**Example 1: Measuring Meson Masses**

The frequencies of the oscillating magnetization can be mapped directly to the energy of the domain wall bound states [48], which have a large overlap with chosen initial states. These so called meson masses are defined as the energy difference between the lowest excited states and the ground state.

Recent work using a trapped ion quantum simulator to simulate the long-range TFIM model showed how, by choosing a variety of initial states, the meson masses can be measured through the persistent oscillations of local magnetization [62]. Previous attempts to perform a similar measurement on a digital quantum computer have failed for the short ranged TFIM, Eq. (7), because the results are too noisy to resolve the smaller amplitude of oscillations [35]. As our main result, we show that our new error mitigation enables us to obtain the meson masses from the IBM device.

We are mainly interested in the time dependence of the local magnetization, which further simplifies with $\text{Tr}[\sigma^\alpha] = 0$ in Eq. (5) to

$$\langle \sigma_z^\alpha \rangle = (1 - p_{\text{tot}}) \langle \sigma_z^\alpha \rangle_{\text{exact}}.$$  

(8)

The time dependence can be calculated by applying a quantum circuit from a Trotterisation of the time evolution operator (for details of the implementation and quantum circuits see Refs. [24, 35]). Increasing the number of trotter steps, $N_T$, leads to a deeper circuit and the ensuing increase in errors results in a peculiar dampening of the magnetization dynamics which can be removed via our error mitigation. In Fig. 2 we display the results of the local magnetization dynamics. Our mitigation technique not only allows us to measure the main oscillation frequency on the IBM device but in fact to do so with quantitative agreement with the analytical results (the latter is explained in Ref. [35]).

We note that an additional simplification can be used for the local magnetization for added efficiency — no additional measurement of $\text{Tr}[\rho^2]$ was required because the result of $\langle \sigma_z^2 \rangle$ is known at $tJ = 0$ and thus $p_{\text{tot}}$ can be inferred from the corresponding measurement on the IBM device after running the time evolution circuit for $tJ \approx 0$. This big simplification avoids the costly randomized measurement scheme and should be generally applicable for observables whose value is known at $tJ = 0$. However, as this does not distinguish correlated noise from the uncorrelated noise assumed in our ansatz, it is possible for over-estimations in the mitigation to occur, resulting in non-physical results. An example of this is seen in Fig. 2 at time $tJ = 0.1$ in which the local magnetization is measured to be greater than unity.

We have obtained a range of results with different number of trotter steps. Under the assumption that $(1 - p_{\text{tot}})$ scales as $(1 - p_T)^{N_T}$, where $p_T$ is the error in one trotter step, we can extrapolate the results back to the error-free case, see green data points in Fig. 2. We are then in a position to suppress the noise to a level which enables us to extract different meson masses on the IBM device by a basic fit of the main oscillation frequency. In Fig. 3 we show the data for the first three masses obtainable by starting from different initial states, see insets. Here, 7 spins are mapped onto 5 qubits resulting in a circuit with $5N_T + 5$ single qubit gates and $8N_T$ CNOT gates with $N_T = 5, 6$. Remarkably, we find quantitative agreement between the mitigated results and the theoretical predictions for the scaling of the meson masses with the transverse and longitudinal fields [35].

**Example 2: Entanglement Spreading**

As a second example of we study the suppression of half chain entanglement entropy spreading after the spin chain quench. For $h_z = 0$ the entanglement entropy is expected to increase linearly from the ballistic spreading of free fermionic excitations [64]. However, with a non-zero longitudinal field this growth is suppressed in a characteristic fashion due to confinement [47]. To observe this we have implemented the
FIG. 3. Results from the IBM device Toronto [63] of meson masses. (a,b) The quench dynamics of the z-axis local magnetization is shown for different initial states. Here, $J = 1$, $h_x = 0.75$, $h_y = 0.75$ and $n = 7$. (a) Results calculated via exact diagonalization. (b) Mitigated results from the IBM device. Here, clear dominant oscillations are extracted that quantitatively agree with the analytically derived values. (c) An illustration of how the masses are extracted from the analytically derived energy levels [35]. (d) A comparison of the masses obtained from the IBM device and the analytically derived values for varying $h_z$ showing this quantitative agreement. Note, more details of the circuit composition of the evolution operator can be found in [35].

randomized measurement protocol of Ref. [40] on the IBM device for measuring the second order Rényi entanglement entropy. In Ref. [35] two of us obtained qualitative agreement for entanglement dynamics of six spins compared to an exact diagonalization (ED) calculation, but for a quantitative agreement a large shift of the results was needed.

By using the ansatz in Eq. (4) we can see that the effect of global depolarizing errors on measurements of $\text{Tr}[\rho_A^2]$, where $A$ is a subsystem in consideration, is

$$\text{Tr}[\rho_A^2] = (1 - p_{tot})^2 \text{Tr}[\rho_{A,exact}^2] + \frac{p_{tot}(1 - p_{tot})}{2^n - 1} + \frac{p_{tot}^2}{2^n}.$$ (9)

If $p_{tot}$ is known, $\text{Tr}[\rho_{A,exact}^2]$ can be extracted and the second order Rényi entropy measurement is calculated via

$$S^{(2)}(\rho_{exact}) = - \log_2(\text{Tr}[\rho_{A,exact}^2]).$$ (10)

Fig. 1 shows how this mitigation protocol eliminates the error in the second order Rényi entropy results. With our error mitigation protocol and Eq. (9) we now obtain quantitative agreement with ED results for six spins and can account for the large shift of the results.

APPLICATIONS TO MORE GENERAL EXAMPLES

To demonstrate that our method works more generally, we simulate example tasks of measuring the expectation value of different operators. We choose a layered, brickwork circuit consisting of a layer of single qubit rotations followed by a layer of entangling CNOTs — the same circuit structure as for Trotterized evolution (see Supplementary material). However, instead of time parameterized Trotter evolution, we simply choose random single qubit angles, corresponding to different states, and therefore different expectation values.

In order to obtain a larger amount of data to support our mitigation technique, we numerically simulate the layered circuits with noisy gates. Qiskit’s circuit simulator allows one to simulate arbitrary gate-based noise models by specifying independent Kraus operators for each single- and two-qubit gates. Using this noisy simulator, we can directly test how gate-level non-depolarizing errors can result in an effective global circuit-level depolarizing error, and how well our mitigation technique works in the presence of local non-depolarizing errors. To simulate a realistic noise model, we take the Kraus operators directly from the ibmq_santiago backend [65]. This error model goes beyond the single gate depolarizing assumption, by including (asymmetric) thermal relaxation.
In order to test local, non-local, single, and many Pauli-string operators, we have chosen the following operators: single $Z$ operator (local, single term), TFIM Hamiltonian (local, many terms), Random Pauli string (non-local, single term), and the molecular Hamiltonian of a $H_4$ (non-local, many terms). Here many ‘terms’ refers to the number of non-commuting Pauli strings that must be measured to construct the expectation value, and locality refers to only containing single and two-qubit Pauli strings. Note the four hydrogen Hamiltonian (see Supplementary for details) is chosen as a simple molecular test case with a large Hilbert space, and where we can ignore complications from freezing out orbitals.

Figure 4 shows unmitigated and mitigated expectation values for each operator, and clearly demonstrates the added value of purity measurements (see Supplementary for simulation details). For expectation value mitigation (i.e. using a single expectation value to calibrate $p_{\text{tot}}$), a single extra circuit is needed. For trace mitigation (using Eq. (9) to calibrate $p_{\text{tot}}$) uses 500 extra circuits to find $p_{\text{tot}}$; these are split into five groups to reduce the bias. Once $p_{\text{tot}}$ is calibrated for a given circuit depth, this single value is used for error mitigation of the same circuit structure, but for different single qubit parameter angles.

As the main results we find that our global depolarising assumption is an effective error channel for the whole circuit, and, crucially, it does not require each single- and two-qubit errors to be themselves purely depolarising.

**DISCUSSION**

In this work we have proposed an error mitigation technique which is simple to implement but which retains the mathematical rigour of more complicated techniques. Our protocol is directly applicable to any quantum simulation whose measurements are basic expectation values, as exemplified by our results for the time evolution of the local magnetization in spin chain dynamics. Though our error mitigation method may not replace full state tomography, we expect it to be useful in measuring more complicated quantities beyond simple observables, as corroborated by our results for the entanglement entropy. An interesting avenue of future research will be an application to variational quantum eigensolver (VQE) problems [66] and other quantum circuits.

The computational overhead of implementing our protocol is dominated by the cost of evaluating $\text{Tr}(\rho^2)$ on the quantum device. The number of corresponding measurements is $N_u N_m$ where $N_u$ is the number of randomised unitaries and $N_m$ is the number of random measurements. Within the randomised scheme of Ref. [39] it grows exponentially with the system size [40], which poses a potential problem for large quantum computers but is easily feasible for currently available NISQ devices. Alas, in some cases the costly randomized measurement scheme can be avoided entirely, i.e. in our benchmark example of spin chain dynamics a single measurement of a local observable whose exact value is known was sufficient. This should be true generally for quantum circuits which can be tuned to be close to the identity. However, this scalable simplification for obtaining $p_{\text{tot}}$ potentially faces the problem that it includes correlated errors which makes the randomised measurements preferable as long as it is feasible.

The fact that our basic assumption of global depolarizing errors leads to such large improvements in results is in itself remarkable. The basic conclusion is that the total error on the IBM device is close to a global depolarizing error at least for our choice of problems. We note our error ansatz, Eq. (4) is an approximation even for single qubit depolarizing errors, let alone the more complex channels that are no-doubt present in physical devices. Nevertheless the ansatz works remarkably well for correcting physical errors of large/deep circuits and is a no-lose addition to quantum simulation protocols. We suggest this is because the depolarizing channel is a good approximation per gate in a many gate quantum circuit, even if this approximation breaks down for single layers of gates. Of course, our mitigation falls short of accounting for large coherent or correlated errors and it will be a worthwhile endeavour to think about a controlled extension of our basic ansatz Eq. (4) for the density matrix of a NISQ device. In particular, whether we can extend it to incorporate other aspects of error channels [33].
As an application of our error mitigation we have presented previously unobtainable quantitative results for confinement and entanglement dynamics of a quantum spin chain. We have been able to extract the first meson masses of confinement induced bound states and observed the corresponding halting of entanglement spreading. In that context, an ambitious next step would be to use the error mitigation to extend times that can be simulated on a quantum computer to probe the confinement induced slow-thermalisation [48] or to see meson scattering events that have recently been predicted [67–69].

In general, we expect that our simple error mitigation brings us closer to exploiting the quantum advantage of available NISQ devices for real world practical applications.

Acknowledgements.— We are grateful for discussions with Hongzheng Zhao, Adam Smith and Peter Haynes. We acknowledge the Samsung Advanced Institute of Technology Global Research Partnership, travel support via the Imperial-TUM flagship partnership and the use of IBM Quantum services 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 Quantum team. This work is also supported by the UK Hub in Quantum Computing and Simulation, part of the UK National Quantum Technologies Programme with funding from UKRI EPSRC grant EP/T001062/1.

[1] J. Preskill, Quantum 2, 79 (2018).
[2] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Nature Physics 14, 595 (2018).
[3] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, et al., Nature 574, 505 (2019).
[4] A. M. Steane, Physical Review Letters 77, 793 (1996).
[5] P. W. Shor, Physical review A 52, R2493 (1995).
[6] “ibmq_paris v1.6.23, ibm quantum team. retrieved from https://quantum-computing.ibm.com,” (2020).
[7] “Qiskit: An open-source framework for quantum computing,” (2019).
[8] M. S. Jattana, F. Jin, H. De Raedt, and K. Michielsen, Quantum Information Processing 19, 1 (2020).
[9] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Quantum 4, 257 (2020).
[10] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, arXiv preprint arXiv:2006.14044 (2020).
[11] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei, Physical Review A 100, 052315 (2019).
[12] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven, Nature Communications 11, 1 (2020).
[13] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. O’Brien, Physical Review A 98, 062339 (2018).
[14] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. Bultink, X. Fu, C. Price, V. Ostroukh, N. Muthusubramanian, A. Bruno, et al., Physical Review A 100, 010302 (2019).
[15] M. C. Tran, Y. Su, D. Carney, and J. M. Taylor, arXiv preprint arXiv:2006.16248 (2020).
[16] W. J. Huggins, J. McClean, N. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush, arXiv preprint arXiv:1907.13117 (2019).
[17] S. McArdle, X. Yuan, and S. Benjamin, Physical review letters 122, 180501 (2019).
[18] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Physical review letters 120, 210501 (2018).
[19] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. De Jong, Physical Review A 95, 042308 (2017).
[20] J. L. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. Kimchi-Schwartz, J. McClean, J. Carter, W. De Jong, and I. Siddiqi, Physical Review X 8, 011021 (2018).
[21] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature 567, 491 (2019).
[22] M. Otten and S. K. Gray, npj Quantum Information 5, 1 (2019).
[23] M. Otten and S. K. Gray, Physical Review A 99, 012338 (2019).
[24] A. Smith, M. Kim, F. Pollmann, and J. Knolle, npj Quantum Information 5, 1 (2019).
[25] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, arXiv preprint arXiv:2005.07601 (2020).
[26] A. Zlokapa and A. Gheorghiu, arXiv preprint arXiv:2005.10811 (2020).
[27] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, arXiv preprint arXiv:2005.10189 (2020).
[28] Y. Li and S. C. Benjamin, Physical Review X 7, 021050 (2017).
[29] S. Endo, S. C. Benjamin, and Y. Li, Physical Review X 8, 031027 (2018).
[30] K. Temme, S. Bravyi, and J. M. Gambetta, Physical review letters 119, 180509 (2017).
[31] E. Gustafson, Y. Meurice, and J. Unmuth-Yockey, Physical Review D 99, 094503 (2019).
[32] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, and W. J. Zeng, arXiv preprint arXiv:2005.10921 (2020).
[33] Z. Cai, arXiv preprint arXiv:2007.01265 (2020).
[34] A. He, B. Nachman, W. A. de Jong, and C. W. Bauer, arXiv preprint arXiv:2003.04941 (2020).
[35] J. Vovrosh and J. Knolle, arXiv:2001.03044 (2020).
[36] M. A. Nielsen and I. Chuang, “Quantum computation and quantum information.” (2002).
[37] B. Vermersch, A. Elben, M. Dalmonte, J. I. Cirac, and P. Zoller, Physical Review A 97, 023604 (2018).
[38] S. van Enk and C. Beenakker, Physical review letters 108, 110503 (2012).
[39] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, Physical review letters 120, 050406 (2018).
[40] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Science 364, 260 (2019).
[41] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, Physical Review Letters 124, 010504 (2020).
[42] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Physical Review A 54, 3824 (1996).
[43] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Physical review letters 76, 722 (1996).
[44] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Physical Review A 77, 012307 (2008).
[45] J. Emerson, M.结尾
assuming many qubits, but not necessarily all qubits have gates applied. We don’t set $(1 - p’) \neq \prod_i (1 - p_i)$, but rather introduce an effective $p’$ to account for components in the total sum that have some reasonable overlap with $U \rho U^\dagger$. E.g., for each $j$, the state $\text{Tr}_j[U \rho U^\dagger] \otimes I_j/2$ may have a large overlap with $U \rho U^\dagger$, and to a lesser extent $\text{Tr}_{ij}$ for each $i, j$, and we’d like to include such contributions without worrying about combinatorial multiplicity, or the state-dependent details of all the partial traces. We can follow the same steps to arrive at Eq. (11) for entangling gates $U_{ij}$ once we note that the 2-qubit depolarizing error channel be written as $\mathcal{E}_{ij} [\rho] = (1 - p_{ij}) U_{ij} \rho U_{ij}^\dagger + p_{ij} \text{Tr}_{ij} [\rho] \otimes I_j/2$, where $I_j/4 = \text{diag}[1, 1, 1, 1]/4$ is the maximally mixed two qubit state.

We can construct the quantum channel for the full circuit. The non-unitary part of Eq. (11) commutes with itself over different layers meaning the a full circuit error channel is approximated as

$$\mathcal{E}_{\text{circ}} [\rho] = (1 - p_{\text{tot}}) U_{\text{circ}} \rho U_{\text{circ}}^\dagger + p_{\text{tot}} T_2^{\otimes n}/2^n$$

(12) where $(1 - p_{\text{tot}}) = \prod_i (1 - p_i)$ (product over effective layer $p’$s) holds due to commuting non-unitary terms. While this is not equal to $\prod_i (1 - p_i)$ (product over gate error probabilities), the latter is a good zeroth-order estimate of this term.

Finally we note the total map Eq. (12) unravels to an effective error model for single (i.e. one or two qubit) gates, given by: $\mathcal{E}_{i} [\rho] = (1 - p’)[U_i \rho U_i^\dagger + p’I_i/2^n]$. Note the all-or-nothing description of the map: either the gate is implemented perfectly, or the entire quantum state is destroyed, not just part of the state as is the case for $\text{Tr}_i [\rho] \otimes I_i/2$ etc. While just part of the state as is the case for $\text{Tr}_i [\rho] \otimes I_i/2$ etc. While
this seems quite an unphysical model for a single gate, it is a
good model of the effective gate error per gate in a multi-gate
circuit.

**Cosine Fitting**

In order to extract the frequencies from the local magneti-
ization results presented in Fig. 2 and Fig. 3, a function that
contained a singular cosine was used, namely,

\[
\langle \sigma^z(t) \rangle \sim Ae^{-dt} \cos(\omega t) + c_1 t + c_2.
\]

Here, \( A \) is the amplitude of the oscillations, \( D \) allows for de-
cay of the amplitude, \( \omega \) is the oscillation frequency and \( c_1 \) and
\( c_2 \) give a time dependent shift. While this function can not
capture long term behaviour of \( \langle \sigma^z \rangle \), it is able to extract out
the dominant oscillations for short time dynamics.

**Statistical Errors Introduced by the Error Mitigation**

The error when calculating \( p_{\text{tot}} \) from \( \text{Tr}(\rho^2) \) is given by

\[
\delta p_{\text{tot}} = \pm (1 - p_{\text{tot}}) \frac{\delta \text{Tr}(\rho^2)}{2(2n - \text{Tr}(\rho^2))},
\]

where \( \delta p_{\text{tot}} \) is the error in \( p_{\text{tot}} \) and \( \delta \text{Tr}(\rho^2) \) is the error in the
measurement of \( \text{Tr}(\rho^2) \). Note that \( \delta \text{Tr}(\rho^2) \) can be calculated by resampling e.g. Jackknife or Bootstrap [35, 40]. Thus,
when using the proposed error mitigation on an observable \( \hat{O} \)
this results additional uncertainty on top of the measurement
errors. More precisely,

\[
\delta(\hat{O}) = \pm \left( \frac{\delta \text{Tr}(\hat{O})}{1 - p_{\text{tot}}} + \frac{\delta p_{\text{tot}} (\langle \hat{O} \rangle - 2^{-n} \text{Tr}(\hat{O}))}{(1 - p_{\text{tot}})^2} \right),
\]

where \( \delta(\hat{O}) \) is the error in the mitigated value of \( \langle \hat{O} \rangle \) and
\( \delta \text{meas}(\hat{O}) \) is the error in the measurement of \( \langle \hat{O} \rangle \). Note, as the
local magnetisation results presented in this paper bypassed
the measurement of \( T\text{r}(\rho^2) \), the error contributing to \( \delta p_{\text{tot}} \) is
just shot noise. Thus, \( \delta(\sigma^z) \) is of the on the order \( \sim O(10^{-3}) \)
and is not displayed in Fig. 2 or Fig. 3.

**Numerical simulations**

All numerical simulations are performed using qiskit [7]
with the santaiago error model (as of 12\text{th} March 2021).
For the hydrogen Hamiltonian, openfermion [71] was used
to compute the orbital integrals (for a row of four H-atoms
with 0.1 nm inter-atomic distance), and then map the prob-
lem to qubits using the symmetry conserving Bravyi-Kitaev
transformation [72], finally the \( Z_2 \) symmetries (corresponding
to particle spin) are removed, reducing the number of qubits
by two. For the TFIM, the Hamiltonian was reduced to two
non-commuting Pauli strings from which the total Hamilton-
ian can be evaluated. Finally the single Z, and \( XYZIXX \)
(chosen randomly) operators were measured trivially. We do
not directly account for State-Preparation and Measurement
(SPAM) errors, but we note that we can (and do) include
SPAM errors as contributing to \( p_{\text{tot}} \) in our channel ansatz.

We numerically simulate circuits consisting of layers of pa-
rameterized \( u^3 \) gates, generic single-qubit rotation gates with
3 Euler angles, followed by an entangling layer. The circuit
depth is defined by the number of entangling (CNOT) layers.
There is a final \( u^3 \) layer before measurement. The parameter-
ized circuit now creates a state \( \rho(\theta) \), and varying \( \theta \) changes the
state and corresponding operator expectation values. For the
numerical simulations we consider how our mitigation per-
forms for different depth circuits, and for the four tasks shown
in the main text. For each depth the calibration \( p_{\text{tot}} \) must be
done independently, as \( p_{\text{tot}} \) is an effective parameter for the
whole quantum channel and cannot simple be calculated a pri-
or for different depths.

**Calibration.** — Two independent methods of calibrating
\( p_{\text{tot}} \) are considered. Firstly, via single expectation value: di-
rect simulation of a circuit with fixed parameters, which is
then compared to the error free expectation value to find \( p_{\text{tot}} \).
Secondly, via trace: from directly measuring \( \text{Tr}[\rho^2] \) and esti-
mating \( p_{\text{tot}} \) via Eq. (9).

For single expectation calibration, a single fixed parameter
circuit was used to find \( \langle O \rangle \) (for each operator \( O \) considered).
This result was compared to the error free simulation of the
exact same circuit to estimate \( p_{\text{tot}} \). To evaluate these expecta-
tion values 40960 shots were used for each non-commuting
term in the decomposition of the given operator into Pauli
strings. For trace mitigation, an extra 500 circuits are used
to estimate \( p_{\text{tot}} \) using 8196 shots per circuit. This was broken
up into five lots of 100 circuits to estimate \( p_{\text{tot}} \) at five differ-
ent (random) circuit parameterizations. For each parameteri-
ization, the method of Ref. [39] is used to find five independent
estimates of \( p_{\text{tot}} \), which are then averaged to reduce the bias
in the estimate of \( p_{\text{tot}} \) for the circuit.

**Mitigation.** — Once \( p_{\text{tot}} \) is estimated either from the single
value or trace measurement, it is used to mitigate further simu-
lated expectation values. New expectation values are obtained
by varying the parameters in the circuit, while keeping the cir-
cuit structure fixed. For each depth and each task, thirty dif-
ferent (parameterized) circuits are simulated, and mitigation
applied. Figure 5 shows the mitigation results for different cir-
cuit depths. While the full trace-estimation is robust showing
a significant improvement in every case, the single expectation
value mitigation is more variable. This is due to the vanishing
expectation value of Pauli strings for random circuits, which
in turn increases the effect of shot noise when inverting Eq. 8.
We note this shot noise problem is absent when calibrating a
Pauli string (or single operator) with a (noise free) unit expecta-
tion value, as in the main text.
FIG. 5. Complete mitigation simulations with non-depolarizing errors. Scaling of our mitigation schemes (circles) with circuit depth for four different operator expectation values. Single operator (red circles) show some improvement, while trace mitigated (green circles) show a robust improvement over the unmitigated results. As in the main text, numerical simulation uses the \textit{ibmq\_santiago} error model, and circuits are layers $u_3$ and CNOT, with the depth equal to the number of entangling layers. Note there are some combinations (e.g. TFIM at depth 14 etc) where the single expectation value mitigation found $p_{tot} = 1$, and therefore did not improve results. Likewise there are other combinations (e.g. TFIM depth 22) where the single expectation value mitigation is omitted because it gave an unphysical estimate of $p_{tot} < 0$. Inset numbers show $p_{tot}$ as estimated for the trace estimation.