Can shallow quantum circuits scramble local noise into global white noise?

Foldager, Jonathan; Koczor, Bálint

Published in:
Journal of Physics A: Mathematical and Theoretical

Link to article, DOI:
10.1088/1751-8121/ad0ac7

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

Can shallow quantum circuits scramble local noise into global white noise?

OPEN ACCESS
IOP Publishing
Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **57** (2024) 015306 (22pp)
https://doi.org/10.1088/1751-8121/ad0ac7

Can shallow quantum circuits scramble local noise into global white noise?

Jonathan Foldager[1,2] **and Bálint Koczor**[2,3,*]

[1] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark
[2] Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom
[3] Quantum Motion, 9 Sterling Way, London N7 9HJ, United Kingdom

E-mail: balint.koczor@materials.ox.ac.uk

Received 28 July 2023; revised 18 October 2023
Accepted for publication 8 November 2023
Published 6 December 2023

Abstract

Shallow quantum circuits are believed to be the most promising candidates for achieving early practical quantum advantage—this has motivated the development of a broad range of error mitigation techniques whose performance generally improves when the quantum state is well approximated by a global depolarising (white) noise model. While it has been crucial for demonstrating quantum supremacy that random circuits scramble local noise into global white noise—a property that has been proved rigorously—we investigate to what degree practical shallow quantum circuits scramble local noise into global white noise. We define two key metrics as (a) density matrix eigenvalue uniformity and (b) commutator norm that quantifies stability of the dominant eigenvector. While the former determines the distance from white noise, the latter determines the performance of purification based error mitigation. We derive analytical approximate bounds on their scaling and find in most cases they nicely match numerical results. On the other hand, we simulate a broad class of practical quantum circuits and find that white noise is in certain cases a bad approximation posing significant limitations on the performance of some of the simpler error mitigation schemes. On a positive note, we find in all cases that the commutator norm is sufficiently small guaranteeing a very good performance of purification-based error mitigation. Lastly, we identify techniques

* Author to whom any correspondence should be addressed.

that may decrease both metrics, such as increasing the dimensionality of the dynamical Lie algebra by gate insertions or randomised compiling.

Keywords: noisy quantum circuits, quantum error mitigation, quantum computing

1. Introduction

Current generations of quantum hardware can already significantly outperform classical computers in random sampling tasks [1, 2] and hopefully future hardware developments will enable powerful applications in quantum machine learning [3], fundamental physics [4, 5] and in developing novel drugs and materials [6–9]. The scale and precision of the technology today is, however, still below what is required for fully fault-tolerant quantum computation: Due to noise accumulation in the noisy intermediate-scale quantum (NISQ) era [10], one is thus limited to only shallow-depth quantum circuits which led to the development of a broad range of hybrid quantum–classical protocols and quantum machine learning algorithms [11–13].

The aim in this paradigm is to prevent excessive error buildup via a parameterised, shallow-depth quantum circuit and then perform a series of repeated measurements in order to extract expected values. These expected values are then post processed on a classical computer in order to update the parameters of the circuits, e.g. as part of a training procedure. A major challenge is the potential need for an excessive number of circuit repetitions which, however, can be significantly suppressed by the use of advanced training algorithms [14–16] or via classical-shadows-based protocols [17–20]. As such, the primary limitation of near-term applications is the damaging effect of gate noise on the estimated expected values which can only be reduced by advanced error mitigation techniques [12, 21].

Error mitigation comprises a broad collection of diverse techniques that generally aim to estimate precise expected values by suppressing the effect of hardware imperfections [12, 21]. Due to the diversity of techniques and due to the significant differences in the range of applicability, the need for performance metrics was recently emphasised [21]. This motivates the present work to characterise noise in typical practical circuits, e.g. in quantum simulation or in quantum machine learning, and define two key metrics that determine the performance of a broad class of error mitigation techniques: (a) eigenvalue uniformity W as a closeness to global depolarising (white) noise and (b) norm of the commutator C between the ideal and noisy quantum states. While (b) quantifies the stability of the quantum state's dominant eigenvector and thus determines the performance of purification based error mitigation techniques [22, 23], (a) implies a good performance of all error mitigation techniques.

Our primary motivation is that gate errors in complex quantum circuits are scrambled into global white noise [1, 24]. This property has been proved for random circuits by establishing exponentially decreasing error bounds; surprisingly, in our numerical simulations we find that in many practical scenarios the same bounds apply relatively well. In particular, we find that both our metrics, (a) the distance from global-depolarising noise and (b) the commutator norm, are approximated with a model function as

$$f(\nu) = \alpha \frac{e^{-\xi}\xi}{(1-e^{-\xi})\sqrt{\nu}} = \frac{\alpha}{\sqrt{\nu}} + O(\xi), \tag{1}$$

where ν is the number of gates in the quantum circuit, ξ is the number of expected errors in the entire circuit and α is a constant. As such, if one keeps the error rate small $\xi \ll 1$ but increases the number of gates in a circuit then both (a) and (b) are expected to decrease. This is a highly

desirable property in practice, e.g. white noise does not introduce a bias to the expected-value measurement but only a trivial, global scaling as we detail in the rest of this introduction.

Main contributions of the present work include the following.

(a) We simulate a broad range of quantum circuits often used in practice and numerically analyse the above two key metrics, W and C, that determine the performance of different error mitigation techniques.
(b) We identify scenarios where the above white-noise approximation holds well (i.e. small W), e.g. when gate parameters and circuit structures are sufficiently random.
(c) We identify strategies to reduce W, e.g. by improving scrambling local gate noise into global white noise through inserting additional gates into a circuit to increase the dimensionality of its Lie algebra [25].
(d) In most cases, however, we conclude that white noise is not necessarily a good approximation due to the large prefactor α in equation (1). Thus the performance of some error mitigation techniques that rely on a global-depolarising noise assumption is limited.
(e) On the other hand, we find that in all cases the commutator norm C, our other key metric, is smaller than W by at least 1-2 orders of magnitude guaranteeing a very good performance of purification-based error mitigation techniques even in the worst practical scenarios.

Our work is structured as follows. In the rest of this introduction we briefly review global depolarising noise and how it can be exploited in error mitigation, and then briefly review purification-based error mitigation techniques and their performance. In section 2 we introduce theoretical notions and finally in section 3 we present our simulation results.

1.1. Global depolarisation and error mitigation

In the NISQ-era, we do not have comprehensive solutions to error correction, which has led the field to develop error mitigation techniques. These techniques aim to extract expected values $\langle O\rangle_{\text{ideal}} := \text{tr}[O\rho_{id}]$ of observables—typically some Hamiltonian of interest—with respect to an ideal noiseless quantum state ρ_{id}.

A very simple error model, the global depolarising noise channel, has been very often considered as a relatively good approximation to complex quantum circuits. For qubit states, the channel mixes the ideal, noise-free state with the maximally mixed state Id/d of dimension $d = 2^N$ as

$$\rho_{wn} := \eta\rho_{id} + (1-\eta)\,\text{Id}/d. \tag{2}$$

Here $\eta \approx F$ is a probability that approximates the fidelity as $F = \eta + (1-\eta)/d$. The white noise channel has been commonly used in the literature for modelling errors in near-term quantum computers [26] and, in particular, it has been shown to be a very good approximation to noise in random circuits [1, 24]. White noise is extremely convenient as it lets the user extract, after rescaling by η, the ideal expected value of any traceless Hermitian observable O via

$$\langle O\rangle_{\text{ideal}} = \text{tr}\left[O\rho_{wn}\right]/\eta. \tag{3}$$

Of course, for small fidelities $\eta \ll 1$ the expected value $\text{tr}[O\rho_{wn}]$ requires a significantly increased sampling to suppress shot noise. In this model, the scaling factor η is a global property and can be estimated experimentally [27], e.g. via randomised measurements [26], via extrapolation [28] or via learning-based techniques [29].

Global depolarisation, however, may not be a sufficiently accurate model to capture more subtle effects of gate noise and thus rescaling an experimentally estimated expected value yields a biased estimate of the ideal one as $\langle O\rangle_{\text{bias}} := \text{tr}[O\rho]/\eta - \langle O\rangle_{\text{ideal}}$. The bias here $\langle O\rangle_{\text{bias}}$ is not a global property, i.e. it is specific to each observable, and requires the use of more advanced error mitigation techniques to suppress.

Intuitively, one expects the bias is small for quantum states that are well approximated by a global depolarising model as $\rho \approx \rho_{wn}$ and, indeed, we find a general upper bound in terms of the trace distance as

$$|\langle O\rangle_{\text{bias}}| = \frac{|\text{tr}[O\rho] - \text{tr}[O\rho_{wn}]|}{\eta} \leqslant \frac{||O||_\infty ||\rho - \rho_{wn}||_1}{\eta}. \tag{4}$$

Here $||\cdot||_1$ is the trace norm while $||O||_\infty$ is the operator norm as the absolute largest eigenvalue of the observable O, which we above assumed to be traceless, refer to [30] for a proof. As such, a small trace distance guarantees a small bias and thus indirectly determines the performance of all error mitigation techniques—and further protocols [19, 31].

In this work, we characterise how close noisy quantum states ρ in practical applications approach white noise states ρ_{wn} and consider various types of variational quantum circuits that are typical for NISQ applications. When the above trace distance is small then it guarantees a small bias in expected values which allows us to nearly trivially mitigate the effect of gate noise, i.e. via a simple rescaling.

1.2. Purification-based error mitigation and the commutator norm

Another core metric we will consider is the commutator norm between the ideal and noisy quantum states as $\mathcal{E}_C := ||[\rho_{id}, \rho]||_1$, which determines the performance of purification based error mitigation techniques [30]—a small commutator norm has significant practical implications as it guarantees that one can accurately determine expected values using the ESD/VD [22, 23] error mitigation techniques. In particular, independently preparing n copies of the noisy quantum state and applying a derangement circuit to entangle the copies, allows one to estimate the expected value

$$\frac{\text{tr}[\rho^n O]}{\text{tr}[\rho^n]} = \langle O\rangle_{\text{ideal}} + \mathcal{E}_{\text{ESD}}.$$

The approach is very NISQ-friendly [32, 33] and its approximation error $\mathcal{E}_{\text{ESD}}$ approaches in exponential order a noise floor as we increase the number of copies n [22]; This noise floor is determined generally by the commutator norm $\mathcal{E}_C$ but in the most typical applications of preparing eigenstates, the noise floor is quadratically smaller as $\mathcal{E}_C^2$ [30].

Note that this commutator can vanish even if the quantum state is very far from a white noise state, in fact it generally vanishes when ρ_{id} approximates an eigenvector of ρ. When a state is close to the white noise approximation then a small commutator norm is guaranteed, however, we demonstrate that the latter is a much less stringent condition and a much better approximation in practice than the former: in all instances we find that the commutator norm is significantly smaller than the trace distance from white noise states.

2. Theoretical background

In this section we introduce the main theoretical notations and recapitulate the most relevant results from the literature.

2.1. General properties of noisy quantum states

Recall that any quantum state of dimension d can be represented via its density matrix ρ that generally admits the spectral decomposition as

$$\rho = \sum_{k=1}^{d} \lambda_k |\psi_k\rangle\langle\psi_k|, \tag{5}$$

where we focus on N-qubit systems of dimension $d = 2^N$. Here λ_k are non-negative eigenvalues and $|\psi_k\rangle$ are eigenvectors. Since $\sum_k \lambda_k = 1$, the spectrum $\underline{\lambda}$ is also interpreted as a probability distribution.

If ρ is prepared by a perfect, noise-free unitary circuit, only one eigenvalue is different from zero and the corresponding eigenvector is the ideal quantum state as $|\psi_{id}\rangle$. In contrast, an imperfect circuit prepares a ρ that has more than one non-zero eigenvalues and is thus a probabilistic mixture of the pure quantum states $|\psi_k\rangle$, e.g. due to interactions with a surrounding environment. In fact, noisy quantum circuits that we typically encounter in practice produce quite particular structure of the eigenvalue distribution: one dominant component that approximates the ideal quantum state $|\psi_1\rangle \approx |\psi_{id}\rangle$ mixed with an exponentially growing number of 'error' eigenvectors that have small eigenvalues. White noise is the limiting case where non-dominant eigenvalues are exponentially small $\propto 1/d$ and the dominant eigenvector is identical to the noise-free state as $|\psi_1\rangle = |\psi_{id}\rangle$.

The quality of the noisy quantum state is then defined by the probability of the ideal quantum state as the fidelity $F := \langle\psi_{id}|\rho|\psi_{id}\rangle$; We show in appendix A that for any quantum state it approaches the dominant eigenvalue λ_1 as

$$\lambda_1 = F + O(\mathcal{E}_C), \tag{6}$$

where we compute the error term analytically in terms of the commutator norm $\mathcal{E}_C = ||[\rho_{id}, \rho]||_1$ from section 1.1. This property is actually completely general and applies to any density matrix.

2.2. Practically motivated noise models

Most typical noise models used in practice, such as local depolarising or dephasing noise, admit the following probabilistic interpretation: a noisy gate operation $\Phi(\rho)$ can be interpreted as a mixture of the noise-free operation U that happens with probability $1-\epsilon$ and an error contribution as

$$\Phi_k(\rho) = (1-\epsilon) U_k \rho U_k^\dagger + \epsilon \Phi_{\text{err}}\left(U_k \rho U_k^\dagger\right). \tag{7}$$

Here U_k is the k^{th} ideal quantum gate and the completely positive trace-preserving map Φ_{err} happens with probability ϵ and accounts for all error events during the execution of a gate. A quantum circuit is then a composition of a series of ν such quantum gates which prepares the convex combination as

$$\rho = \eta \rho_{id} + (1-\eta)\rho_{\text{err}}. \tag{8}$$

Here $\rho_{id} := |\psi_{id}\rangle\langle\psi_{id}|$ is the ideal noise-free state, ρ_{err} is an error density matrix and $\eta = (1-\epsilon)^\nu$ is the probability that none of the gates have undergone errors. This probability actually [24, 30] approximates the fidelity $F := \langle\psi_{id}|\rho|\psi_{id}\rangle$ given the noise model in equation (7) as

$$F = (1-\epsilon)^\nu + \mathcal{E}_F = e^{-\xi} + \mathcal{E}_F + O\left(\epsilon^2/\nu\right). \tag{9}$$

Here we approximate $(1-\xi/\nu)^{\nu} = e^{-\xi} + O(\epsilon^2/\nu)$ for small ϵ and large ν where $\xi := \epsilon\nu$ is the circuit error rate as the expected number of errors in a circuit. The approximation error can be defined as $\mathcal{E}_F = \langle\psi_{id}|\rho_{\text{err}}|\psi_{id}\rangle$ and in practice it is typically small while in the limiting case of white noise it decreases exponentially as $\mathcal{E}_F = 1/d$ due to $\rho_{\text{err}} = \text{Id}/d$.

Assuming sufficiently deep, complex circuits, [30] obtained an approximate bound for the commutator between the ideal and noisy quantum states as

$$|| \, [\rho_{id}, \rho] \, ||_1 \lessapprox \text{const} \times e^{-\xi} \, \xi / \sqrt{\nu}. \tag{10}$$

This bound confirms that as we increase the number of quantum gates ν in a circuit but keeping the circuit error rate ξ constant, the commutator norm decreases as $\propto 1/\sqrt{\nu}$ [30]. Furthermore, this function closely resembles to equation (1) which is a central aim of this work to explore.

2.3. White noise in random circuits

Random circuits have enabled quantum supremacy experiments using noisy quantum computers for two primary reasons: (a) the outputs of these circuits are hard to simulate classically and (b) they render local noise into global white noise [1], hence introducing only a trivial bias to the ideal probability distribution similarly as in section 1.1.

Dalzell *et al* [24] considered random circuits consisting of s two-qubit gates, each of which undergoes two single qubit (depolarising) errors each with probability $\tilde{\epsilon}$ (assuming single-qubit gates are noiseless). We can relate this to our model by identifying the local noise after each two-qubit gate with the error event in equation (7) via the probability $\epsilon = 1-(1-\tilde{\epsilon})^2 = 2\tilde{\epsilon} - \tilde{\epsilon}^2$. Dalzell *et al* [24] then established the fidelity $\tilde{F}$ of the quantum state which one obtains from a noisy cross-entropy score as

$$\tilde{F} = e^{-2s\tilde{\epsilon} \pm O\left(s\tilde{\epsilon}^2\right)} = e^{-\xi \pm O(\epsilon\xi)}.$$

This coincides with our approximation from equation (9) up to an additive error in the exponent which, however, diminishes for low gate error rates. In the following we will thus assume $F \equiv \tilde{F}$.

Measuring these noisy states in the standard measurement basis $\{|j\rangle\}_{j=1}^{d}$ produces a noisy probability distribution $\tilde{p}_{\text{noisy}}(j) = \langle j|\rho|j\rangle$. Dalzell *et al* [24] established that this probability distribution rapidly approaches the white noise approximation $\tilde{p}_{wn} = F\tilde{p}_{id} + (1-F)\tilde{p}_{\text{unif}}$ where $\tilde{p}_{id}(j) = \langle j|\rho_{id}|j\rangle$ is the ideal probability distribution while we define $\tilde{p}_{\text{unif}}(j) = 1/d$ and $d = 2^N$. In particular, the total variation distance (via the l_1 norm $||x||_1 = \sum_i |x_i|$) between the two probability distributions is upper bounded as

$$\frac{1}{2}||\tilde{p}_{\text{noisy}} - \tilde{p}_{wn}||_1 \leqslant O\left(F\epsilon\sqrt{\nu}\right) = O\left(e^{-\xi}\xi/\sqrt{\nu}\right). \tag{11}$$

This expression is formally identical to the bound on the commutator norm in equation (10); Indeed if the noise in the quantum state approaches a white noise approximation, it implies that the commutator norm must also vanish in the same order.

On the other hand, the reverse is not necessarily true as equation (11) is a stronger condition than equation (10) as the latter only guarantees that the dominant eigenvector approaches $|\psi_1\rangle \approx |\psi_{id}\rangle$ but does not imply anything about the eigenvalue distribution of ρ or ρ_{err}.

3. Numerical simulations

3.1. Target metrics

In the NISQ-era comprehensive error correction will not be feasible and thus hope is primarily based on variational quantum algorithms [11–13, 35, 36]. In this paradigm a shallow, parametrised quantum circuit is used to prepare a parametrised quantum state that aims to approximate the solution to a given problem, typically the ground state of a problem Hamiltonian. Due to its shallow depth the ansatz circuit is believed to be error robust and its tractable parametrisation allows to explore the Hilbert space near the solution. On the other hand, such circuits are structurally quite different than random quantum circuits and it was already raised in [24] whether error bounds on the white noise approximation extend to these shallow quantum circuits.

We simulate such quantum circuits under the effect of local depolarising noise—while note that a broad class of local coherent and incoherent error models can effectively be transformed into local depolarising noise using, e.g. twirling techniques or randomised compiling [37–40]. We analyse the resulting noisy density matrix ρ by calculating the following two quantities. First, we quantify 'closeness' to a white noise state from equation (2) by computing the uniformity measure W as the l_1-distance between the uniform distribution and the non-dominant eigenvalues of the output state as

$$W := \frac{1}{2}||p_{\text{err}} - p_{\text{unif}}||_1 = \frac{1}{2}\sum_{k=2}^{d}\left|\frac{\lambda_k}{1-\lambda_1} - \frac{1}{d-1}\right|, \tag{12}$$

where $||\cdot||_1$ refers to the l1-norm of the vector as in equation (11). In analogy to section 2.3, here we consider the probability distribution when measuring the density matrix in its eigenbasis, rather than in the standard basis, and thus define $p_{\text{err}} := (\lambda_2, \lambda_3, \ldots, \lambda_d)/(1-\lambda_1)$ via the non-dominant eigenvalues of the density matrix in statement 2 while $p_{\text{unif}}(j) = 1/(d-1)$. Thus, W depends only on spectral properties of the quantum state and can be computed straightforwardly in numerical simulations. We show in statement 2 that W is proportional to the trace distance from a white noise quantum state as

$$||\rho - \rho_{wn}||_1 = (1-\lambda_1)\, W + \mathcal{E}_w, \tag{13}$$

uo to a bounded error $\mathcal{E}_w$. The uniformity measure W thus determines the bias in estimating any traceless expected value as discussed in section 1.1.

Second, we calculate the commutator norm $\mathcal{E}_C$ from section 1.1 relative to $1-\lambda_1$ as

$$C := \frac{||\,[\rho_{id}, \rho]\,||_1}{1-\lambda_1} = ||\,[\rho_{id}, \rho_{\text{err}}]\,||_1 + \mathcal{O}(\mathcal{E}_q), \tag{14}$$

which we relate to the commutator norm between the 'error part' of the state ρ_{err} and the ideal quantum state ρ_{id} in lemma 1. In the following, we will refer to C as the commutator norm – and recall that it determines the ultimate performance of purification-based error mitigation as discussed in section 1.1.

While both metrics W and C have clear practical meaning as they determine the efficacy of error mitigation techniques, they carry rather orthogonal information about the nature of noise in the quantum state. Specifically, W informs about the distribution of eigenvalues of the density matrix while C quantifies the overlap between the ideal state and the noise in the density matrix. This difference is best illustrated via the following family of extremal states: there exist quantum states for which C vanishes but W is large, i.e. low-rank errors where

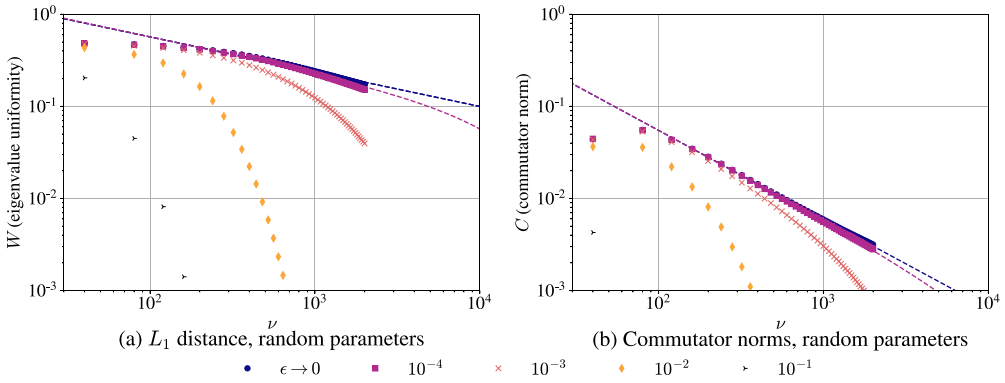


Figure 1. Simulating families of 10-qubit Strong entangling layer (SEL) ansatz circuits [34] at random gate parameters for an increasing number ν of gates and per-gate depolarising error rates ϵ. (a) the uniformity measure $W(\nu)$ of the error eigenvalues of the density matrix from equation (12) closely match the theoretical model (dashed lines) for random circuits and confirm that increasing the number of gates in random circuits scrambles local noise into global white noise. (b) the commutator norm $C(\nu)$ from equation (14) is significantly smaller in absolute value and decreases with a larger polynomial degree (steeper slope of the dashed lines) than the uniformity measure—this suggests that the dominant eigenvector of the density matrix ρ approximately commutes with ρ even when noise is not well described by white noise. The $\epsilon \to 0$ simulations were approximated using $\epsilon = 10^{-8}$ ($\epsilon = 10^{-7}$) when calculating W (C).

$\rho = \eta\rho_{id} + (1-\eta)|\psi_{\text{err}}\rangle\langle\psi_{\text{err}}|$ via an orthogonal noise component $\langle\psi_{\text{err}}|\psi_{id}\rangle = 0$. These states are indeed states for which purification-based techniques are expected to outperform white-noise based ones most significantly. Conversely, families of worst-case scenario states have been explicitly constructed in [30], highlighting that states exist for which C is maximal while W may be small. Furthermore, the relation between W and C is non-trivial and we leave a detailed analysis to future work.

3.2. *Random states via Strong Entangling ansätze*

We first consider a Strong Entangling Ansatz (SEA): it is built of alternating layers of parametrised single-qubit rotations followed by a series of nearest-neighbour CNOT gates as illustrated in figure A1—and assume a local depolarising noise with probability ϵ. We simulate random quantum circuits by randomly generating parameters $|\theta_k| \leqslant 2\pi$—note that these circuits are not necessarily Haar-random distributed and thus results in section 2.3 do not necessarily apply.

We simulate 10-qubit circuits and in figure 1(a) we plot the eigenvalue uniformity $W(\nu)$ while in figure 1(b) we plot the commutator norm $C(\nu)$ for an increasing number ν of quantum gates—all datapoints are averages over ten random seeds. These results surprisingly well recover the expected behaviour of random quantum circuits as for small error rates $\epsilon \to 0$ both quantities $W(\nu)$ and $C(\nu)$ can be approximated using the model function from equation (1) as we now discuss.

In section 2.3 we stated bounds of [24] on the distance between $\tilde{p}_{\text{noisy}}$ and $\tilde{p}_{wn}$. Based on the assumption that these bounds can be extended beyond the standard basis and also apply to

the probability distributions $p_{\text{noisy}} = \langle\psi_k|\rho|\psi_k\rangle$ and $p_{wn} := \langle\psi_k|\rho_{wn}|\psi_k\rangle$ we derive in statement 4 the approximate bound on the eigenvalue uniformity as

$$W(\nu) = O\left(\frac{e^{-\xi}\xi/\sqrt{\nu}}{1-e^{-\xi}}\right).$$

Furthermore, by combining equation (14) and the bound in equation (10) we find that the commutator norm C is similarly bounded by the same function. On the other hand, figure 1(b) suggests that the commutator norm decreases in a higher polynomial order and thus we approximate both $W(\nu)$ and $C(\nu)$ using the function

$$f(\nu) = \alpha\frac{\xi\, e^{-\xi}}{\nu^{\beta}\,(1-e^{-\xi})} = \alpha/\nu^{\beta} + \mathcal{O}(\xi) \tag{15}$$

where we fit the two parameters α and β to our simulated dataset. The second equation above is an expansion for small circuit error rates ξ as detailed in appendix A.2.1. Indeed, in figure 1(blue circles) for small $\epsilon \to 0$ we observe a nearly linear function in the log-log plot in figure 1 and thus remarkably well recover the theoretical bounds with the polynomial power approaching $b \to 1/2$.

Furthermore, comparing figure 1((b), blue circles) and figure 1((a), blue circles) suggest that the commutator norm has both a significantly smaller absolute value (smaller α) and decreases at a faster polynomial rate (larger beta) than the uniformity measure. In fact, the commutator norm is more than two orders of magnitude smaller than the uniformity measure which suggests that even when ρ_{err} is not approximated well by a white noise state it, nevertheless, almost commutes with the ideal pure state ρ_{id}.

We finally consider how the absolute factor α depends on the number of qubits: we perform simulations at a small error rate $\epsilon \to 0$ and fit our model function $\alpha\nu^{\beta}$ to extract $\alpha(N)$ for an increasing number of qubits. The results are plotted in figure A3(e) and suggest that the prefactor $\alpha(N)$ initially grows slowly but then saturates — this is consistent with a polylogarithmic depth being sufficient to reach anticoncentration [24].

3.3. *Variational Hamiltonian Ansatz*

Theoretical results guarantee that the SEL ansatz initialised at random parameters approach for an increasing depth unitary 2-designs thereby reproducing properties of random quantum circuits [41, 42]. It is thus not surprising that the model introduced in section 2.3 gives a remarkably good agreement between the SEL ansatz (dots on in figure 1) and genuine random circuits (fits as continuous lines in figure 1).

Here we consider the Hamiltonian Variational Ansatz (HVA) [43, 44] at more practical parameter settings: The HVA has the advantage that we can efficiently obtain parameters that increasingly better (as we increase the ansatz depth) approximate the ground state of a problem Hamiltonian—we will refer to these as VQE parameters. We also want to compare this circuit against random circuits and thus also simulate the HVA such that every gate receives a random parameter as detailed in appendix B.1.

While the VQE parameter settings capture the relevant behaviour in practice as one approaches a solution, the random parameters are more relevant, e.g. at the early stages of a VQE parameter optimisation. Furthermore, as the circuit is entirely composed of the Pauli

terms of the problem Hamiltonian, the dimensionality of its dynamical Lie algebra is entirely determined by the problem Hamiltonian in contrast to an exponentially growing algebra of the SEL ansatz [25].

3.4. Heisenberg XXX spin model

We first consider a VQE problem of finding the ground state of the 1-dimensional XXX spin-chain model. We construct the HVA ansatz from section 3.3 for this problem Hamiltonian as a sum $\mathcal{H}_{\mathrm{XXX}} = \mathcal{H}_0 + \mathcal{H}_1$ as

$$\mathcal{H}_0 = \sum_{k=1}^{N} \Delta_k Z_k, \qquad \mathcal{H}_1 = \sum_{k=1}^{N} \left[X_k X_{k+1} + Y_k Y_{k+1} + Z_k Z_{k+1}\right].$$

The Pauli operators *XX*, *YY* and *ZZ* determine couplings between nearest neighbour spins in a 1D chain and we choose them to be of unit strength. Furthermore, Z_k are local on-site interactions $|\Delta_k| \leqslant 1$ that were generated uniformly randomly such that the Hamiltonian has a non-trivial ground state.

First, we simulate the HVA ansatz for $N = 10$ qubits with randomly generated circuit parameters as $|\theta_k| \leqslant 2\pi$ and plot results for an increasing number of quantum gates in figures 2(a) and (c). We find a similar behaviour for the eigenvalue uniformity $W(\nu)$ as with random SEL circuits in figure 1(a) and obtain a reasonably good fit for $\epsilon \to 0$ using our model function from equation (15). The commutator norm in figure 2(c) is again significantly smaller in magnitude than the uniformity measure and decreases faster, in a higher polynomial order similarly to the random SEL ansatz in figure 1(b) .

Second, in figures 2(b) and (d) we simulate the ansatz at the VQE parameters that approximate the ground state. Since the ansatz parameters become very small as one approaches an adiabatic evolution, it is not surprising that the output density matrix is not well-approximated by a white noise state: the uniformity measure is very large in figure 2(b). The commutator norm in figure 2(d) again, is significantly smaller than $W(\nu)$ and although it appears to slowly grow with ν, it appears to decrease for $\nu \to \infty$. This agrees with observations of [30] that the circuits need not be random for the commutator to be sufficiently small in practice.

Furthermore, in figures A3(a) and (b) we investigate the dependence on N and find that the prefactor α grows slowly and appears to saturate for $N \geqslant 10$.

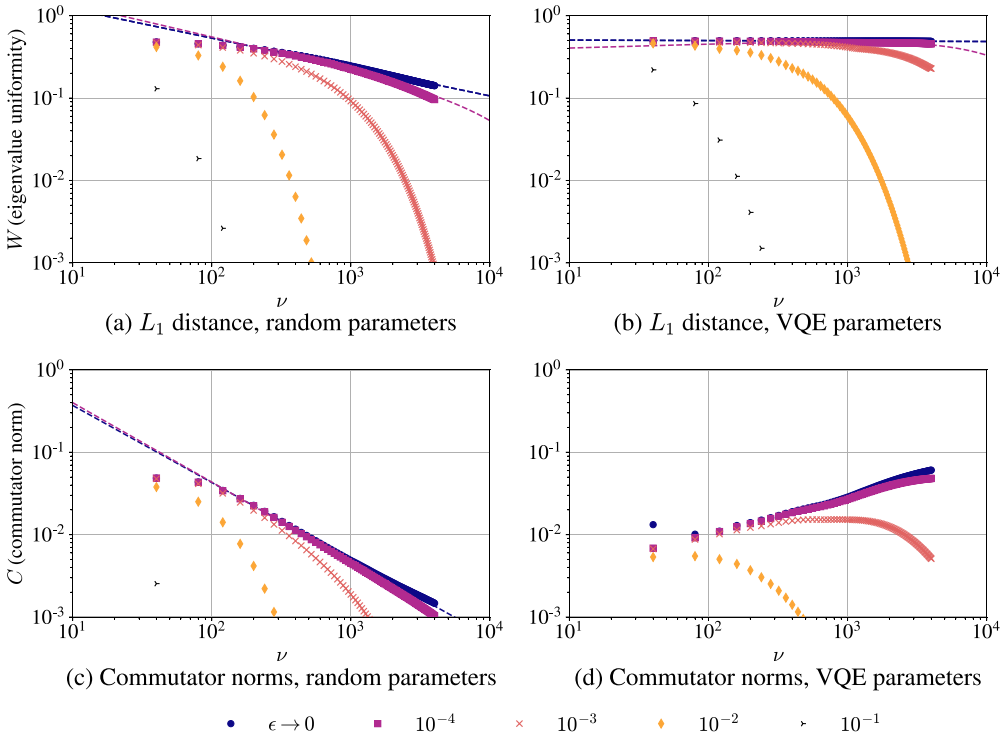
3.5. TFI

In the next example we consider the transverse-field Ising (TFI) model $\mathcal{H}_{\mathrm{TFI}} = \mathcal{H}_0 + \mathcal{H}_1$ using constant on-site interactions $h_i = 1$ and randomly generated coupling strengths $|J_i| \leqslant 1$ as

$$\mathcal{H}_0 = -\sum_i h_i X_i, \qquad \mathcal{H}_1 = -\sum_i J_i Z_i Z_{i+1}. \tag{16}$$

We first simulate the HVA ansatz with random variational parameters in figures 3(a) and (c). While at small error rates $\epsilon \to 0$ figure 3((a), blue) can be fitted well with our polynomial approximation form equation (15), we observe that the eigenvalue uniformity $W(\nu)$ in figure 3((a), blue) decreases in a small polynomial degree.

Indeed, as the HVA ansatz is specific to a particular Hamiltonian, its dynamical Lie algebra may have a low dimensionality [25] resulting in a limited ability to scramble local noise into



(a) L_1 distance, random parameters (b) L_1 distance, VQE parameters

(c) Commutator norms, random parameters (d) Commutator norms, VQE parameters

Figure 2. XXX Hamiltonian: same simulations as in figure 1 but using 10-qubit HVA quantum circuits constructed for the XXX spin problem Hamiltonian. (a), (c) at randomly chosen circuit parameters of the HVA we find the same conclusions as for random circuits in figure 1. (b) when the HVA circuit approximates the ground state of the Hamiltonian (VQE parameters) we find the noise in the circuit is not approximated well by white noise, i.e. the uniformity measure $W(\nu)$ is large and does not decrease as we increase ν. (d) On the other hand, the commutator norm $C(\nu)$ is significantly smaller than $W(\nu)$ confirming that the ideal quantum state approximately commutes with the noisy one. The $\epsilon \to 0$ simulations were approximated using $\epsilon = 10^{-8}$ ($\epsilon = 10^{-7}$) when calculating W (C).

white noise; this explains why in figure 3(a) the uniformity measure decreases more slowly, i.e. in a smaller polynomial order, than random circuits. For this reason, we additionally simulate in figure A2 the TFI-HVA ansatz but with adding R_z gates in each layer whose generator is not contained in the problem Hamiltonian. The increased dimensionality of the dynamic Lie algebra, indeed, improves scrambling as the white noise approximation is clearly better in figure A2–while note that the increased dimensionality may also lead to exponential inefficiencies in training the circuit [25].

In stark contrast to the case of the uniformity measure $W(\nu)$, we find that the commutator norm in figure 3((c), blue) decreases substantially for an increasing ν despite the low dimensionality of the Lie algebra. This nicely demonstrates that a small commutator norm is a much more relaxed condition than white noise as the latter requires that the noise is fully scrambled

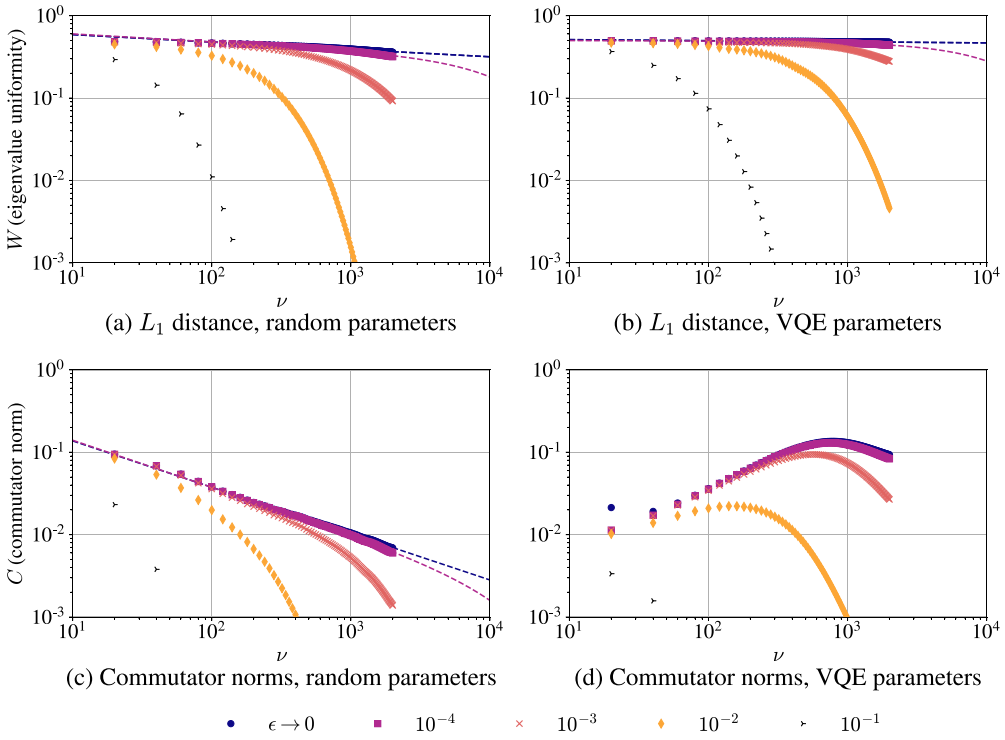


Figure 3. TFI same simulations as in figure 1 but using 10-qubit HVA quantum circuits constructed for the TFI spin problem Hamiltonian. (a), (c) at randomly chosen circuit parameters $W(\nu)$ decreases more slowly, in smaller polynomial order than random circuits—see text and see simulations with added layers of R_z gates in figure A2. (b) at the VQE parameters white noise is again not a good approximation, i.e. the uniformity measure $W(\nu)$ is large and does not decrease as we increase ν. (d) the commutator norm $C(\nu)$ is smaller than $W(\nu)$ in absolute value by an order of magnitude. The $\epsilon \to 0$ simulations were approximated using $\epsilon = 10^{-8}$ ($\epsilon = 10^{-7}$) when calculating W (C).

in the entirety of the exponentially large Hilbert space. Finally, we simulate the TFI circuits at VQE parameters and find qualitatively the same behaviour as in the case of the XXX problem.

3.6. Quantum Chemistry: LiH

We consider a 6-qubit Lithium Hydride (LiH) Hamiltonian in the Jordan-Wigner encoding as a linear combination of non-local Pauli strings $P_k \in \{\mathrm{Id}, X, Y, Z\}^{\otimes N}$ as

$$\mathcal{H}_{LiH} = \sum_{k=1}^{r_h} h_k P_k. \tag{17}$$

We construct the HVA ansatz by splitting this Hamiltonian into two parts with $\mathcal{H}_0$ being composed of the diagonal Pauli terms in equation (17) while $\mathcal{H}_1$ composed of non-diagonal Pauli strings.

Such chemical Hamiltonians typically have a very large number of terms with $r_h \gg 1$ but a significant fraction only have small weights h_k thus the HVA would have a large number of gates with only very small rotation angles [47]. For these reasons we construct a more

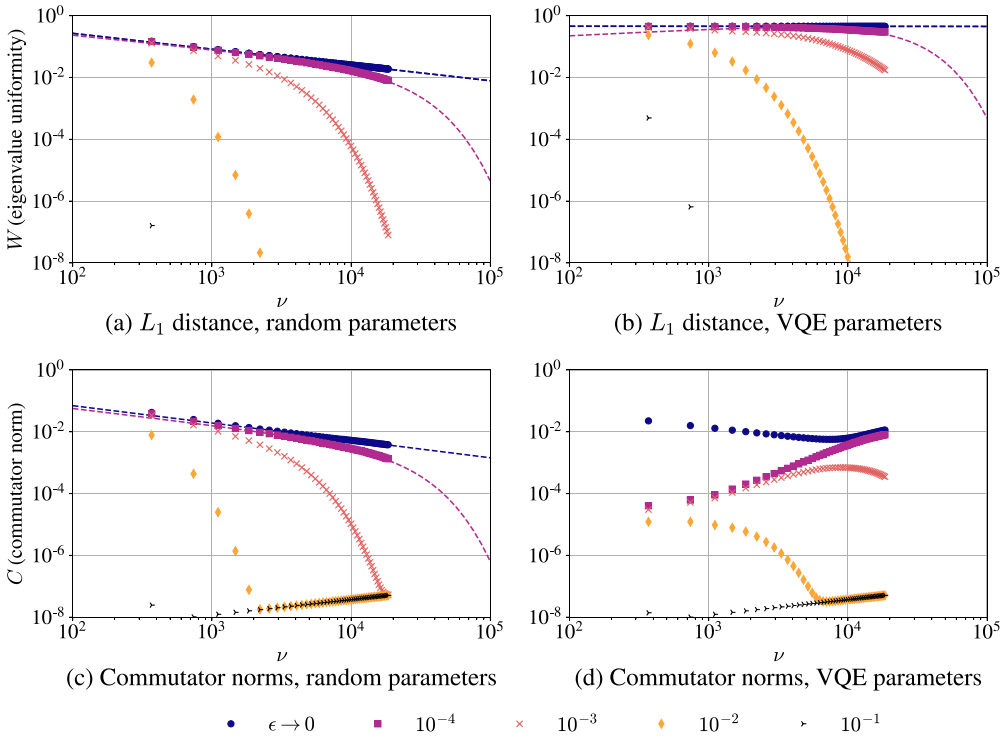


Figure 4. LiH same simulations as in figure 1 but using 6-qubit HVA quantum circuits constructed for a LiH molecular Hamiltonian. (a), (c) at randomly chosen circuit parameters both $W(\nu)$ and $C(\nu)$ decrease as expected for random circuits due our randomised compiling strategy [45, 46]. (b) at the VQE parameters white noise is an increasingly bad approximation, i.e. the uniformity measure $W(\nu)$ increases as we increase ν. (d) the commutator norm $C(\nu)$ is smaller than $W(\nu)$ in absolute value by 2 orders of magnitude. The $\epsilon \to 0$ simulations were approximated using $\epsilon = 10^{-8}$ ($\epsilon = 10^{-7}$) when calculating W (C).

efficient circuit whose basic building blocks are constructed using sparse compilation techniques [45, 46]: Each single layer in the HVA ansatz consists of gates that correspond to 100 randomly selected terms of the Hamiltonian with sampling probabilities $p_k \propto |h_k|$ proportional to the Pauli coefficients. This approach has the added benefit that it makes the circuit structure random as opposed to the fixed structures in sections 3.4 and 3.5.

Results shown in figures 4(a) and (c) agree with our findings from the previous sections: at randomly chosen circuit parameters the uniformity measure decreases according to equation (15); the commutator norm similarly decreases but in a higher polynomial order while its absolute value is smaller by at least an order of magnitude. In contrast, figure 4(b) suggests that the errors are not well approximated by white noise with a large and non-decreasing $W(\nu) \approx 0.5$. Furthermore, figure 4(b) again confirms that despite white noise is not a good approximation, the commutator norm is small in absolute value, i.e. $\approx 10^{-3}$ in the practically relevant region. This guarantees a very good performance of the ESD/VD error mitigation techniques sufficient for nearly all practical purposes.

4. Discussion

Random quantum circuits—instrumental for demonstrating quantum advantage—are known to scramble local gate noise into global white noise for sufficiently long circuit depths [1]: general bounds have been proved on the approximation error which decrease as $\nu^{-1/2}$ as we increase the number ν of gates in the random circuit [24].

In this work we consider shallow-depth, variational quantum circuits that are typical in practical applications of near-term quantum computers and answer the question: can variational quantum circuits scramble local gate noise into global depolarising noise? While the answer to this question is relevant for the fundamental understanding of noise processes in near-term quantum devices, it has significant implications in practice: the degree to which local noise is scrambled into white noise determines the performance of a broad class of error mitigation techniques that are of key importance to achieving value with near-term devices [21]. Our approach in this work is to analytically derive key metrics and numerically explore them in circuits most relevant in practical applications of quantum computers which are, however, beyond the reach of analytical techniques. In particular, we derive two simple metrics that bound performance guarantees: first, the uniformity measure W characterises the performance of error mitigation techniques that assume global depolarising (white) noise [26]; second, the norm C of the commutator between the ideal and noisy quantum states determines the performance of purification-based error mitigation techniques [22, 23] via bounds of [30].

We perform a comprehensive set of numerical experiments to simulate typical applications of near-term quantum computers and analyse characteristics of noise based on the aforementioned two metrics. In all experiments in which we *randomly initialise parameters of the variational circuits* we semiquantitatively find the same conclusions. First, both metrics, the eigenvalue uniformity W and the commutator norm C are well described by our polynomial approximation from equation (15) for small gate error rates. Second, this confirms that, similarly to genuine random circuits, local errors get scrambled into global white noise with a polynomially decreasing approximation error as we increase the number of gates. Third, the commutator C decreases at a higher polynomial rate and has a significantly, by 1–2 orders of magnitude, smaller absolute value in the practically relevant region than the eigenvalue uniformity W. This confirms that purification based techniques are expected to have a superior performance compared to error mitigation techniques that, e.g. assume a global depolarising noise.

We then investigate the practically more relevant case when the ansatz circuits are initialised near the ground state of a problem Hamiltonian; in all cases we semiquantitatively find the same conclusions. First, the errors do not get scrambled into white noise and the approximation errors are large thus effectively prohibiting or at least significantly limiting the use of error mitigation techniques that assume global depolarising noise. We identify reasons why the white noise approximation may fail, i.e. when the dimensionality of the Lie algebra generated by the problem Hamiltonian is not sufficiently (exponentially) large. In such scenarios the circuit structure built out of Hamiltonian terms is not computationally universal and cannot reproduce Haar-random circuits, not even in the limit of infinite-depth. Thus, analysing in detail the relationship between the Lie algebra generated by the problem Hamiltonian and the ability of the circuit to scramble local noise motivates future research. Second, the commutator norm is quite small in absolute value, i.e. $\approx 10^{-2}$–10^{-4} in the practically relevant region; Since the ansatz circuit prepares the ground state, the square of the commutator norm determines the performance of ESD/VD thus for all applications we simulated we expect a very good performance of the ESD/VD approach. Third, we identify strategies to improve scrambling of

local noise into global white noise as we increase circuit depth: We find that inserting additional gates to a HVA that is otherwise not contained in the problem Hamiltonian increases the dimensionality of the dynamic Lie algebra and thus leads to a reduction of both metrics. We find that applying randomised compiling to these non-random, practical circuits also reduces both metrics.

While purification-based techniques [22, 23] have been shown to perform well on specific examples, the present systematic analysis of circuit noise puts these results into perspective and demonstrates the following: First, the superior performance of the ESD/VD technique is not necessarily due to randomness in the quantum circuits—albeit, in deep and random circuits its performance is further improved. Second, while some error mitigation techniques perform well on quantum circuits well-described by white noise [26, 28, 29], we identify various practical scenarios where a limited performance is expected.

The present work advances our understanding of the nature of noise in near-term quantum computers and helps making progress towards achieving value with noisy quantum machines in practical applications. As such, results of the present work will be instrumental for identifying design principles that lead to robust, error-tolerant quantum circuits in practical applications.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/jfold/shallow-circuits-noise.

Acknowledgments

The authors thank Simon Benjamin for his support. The authors thank Richard Meister for instructions on pyQuEST and Arthur Rattew for multiple discussions on quantum simulations. All simulations in this work were performed using the simulation tools QuEST [48] and its Python interface pyQuEST [49]. J F was supported by the William Demant Foundation [Grant Number 18-4438]. B K thanks the University of Oxford for a Glasstone Research Fellowship and Lady Margaret Hall, Oxford for a Research Fellowship. The authors also acknowledge funding from the EPSRC projects Robust and Reliable Quantum Computing (RoaRQ, EP/W032635/1) and Software Enabling Early Quantum Advantage (SEEQA, EP/Y004655/1). B K derived analytical results and contributed to writing the manuscript. J F performed numerical simulations and contributed to writing the manuscript.

Appendix A. Derivation of equation (6)

Recall that any quantum state can be transformed into a non-negative arrowhead matrix following Statement 1 from [30] as $\tilde{\rho} = F|\tilde{\psi}_{id}\rangle\langle\tilde{\psi}_{id}| + D + C$ with

$$\tilde{\rho} = \begin{pmatrix} F & C_2 & C_3 & \dots & C_d \\ C_2 & D_2 & & & \\ C_3 & & D_3 & & \\ \vdots & & & \ddots & \vdots \\ C_d & & & \dots & D_d \end{pmatrix}. \tag{A.1}$$

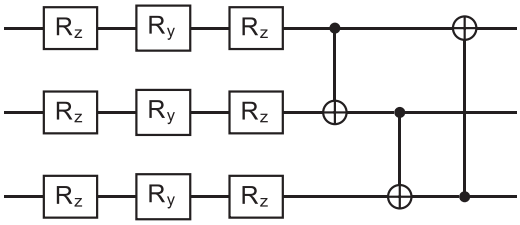


Figure A1. A single layer of the Strong Entangling Layers ansatz for three qubits: it first applies single-qubit gates *Ry*, *Rz* and *Ry* on all qubits which is then followed by nearest neighbour CNOT gates.

We obtain the above matrix by applying a suitable unitary transformation $\tilde{\rho} := U\rho U^\dagger$ such that $|\tilde{\psi}_{id}\rangle := U|\psi_{id}\rangle = (1,0,\dots 0)$ while $F, C_k, D_k \geqslant 0$ with $k \in \{2,3,\dots,d\}$ with d denoting the dimension, and all other matrix entries are zero. Given the above arrowhead representation of a quantum state, one can analytically compute eigenvalues of the density matrix as roots of the following secular equation [30, 50]

$$P(x) = x - F + \sum_{k=2}^{d} \frac{C_k^2}{(D_k - x)} = 0. \tag{A.2}$$

With this we compute the deviation between dominant eigenvalue λ_1 and the fidelity as

$$\begin{aligned} \lambda_1 - F &= \sum_{k=2}^{d} \frac{C_k^2}{(\lambda_1 - D_k)} \\ &\leqslant \max_k (\lambda_1 - D_k)^{-1} \sum_{k=2}^{d} C_k^2 \\ &\leqslant ||\,[\rho_{id}, \rho]\,||^2 \,(2\lambda_1 - 1)^{-1}, \end{aligned}$$

where we have used that $D_k \leqslant \lambda_1$ and that all summands are non-negative as $D_k, C_k, \lambda_1 \geqslant 0$, and in the second inequality we have used the series of matrix norms $\sum_{k=2} C_k^2 = ||C||_{HS}^2/2 = |||[\rho_{id}, \rho]|||_\infty^2$ as established in [30]. We have also introduced the abbreviation $|||[\rho_{id}, \rho]|||$ given all p-norms of the matrix $[\rho_{id}, \rho]$ are equivalent up to a constant factor. In particular, any p-norm of the commutator can be computed as $|||[\rho_{id}, \rho]|||_p = 2^{1/p}\sqrt{\mathrm{Var}[\rho]}$ where we used the quantum mechanical variance $\mathrm{Var}[\rho] := \langle\psi_{id}|\rho^2|\psi_{id}\rangle - F^2$ as established in [30]. Furthermore, in the second inequality in equation (A.2) we have used that $\max_k(\lambda_1 - D_k)^{-1} = (\lambda_1 - D_2)^{-1} \leqslant (\lambda_1 - \lambda_2)^{-1} \leqslant (2\lambda_1 - 1)^{-1}$ by substituting the general inequality $\lambda_2 \leqslant (1 - \lambda_1)$ due to the fact that $\mathrm{tr}[\rho] = 1$.

By denoting the commutator norm as $\mathcal{E}_C$, we can thus finally conclude that $\lambda_1 - F \in O(\mathcal{E}_C)$ as stated in equation (6).

A.1. Trace distance from white noise states

In this section we evaluate analytically the trace distance of any quantum state ρ from the corresponding white noise state in equation (2) in terms of a distance between probability distributions.

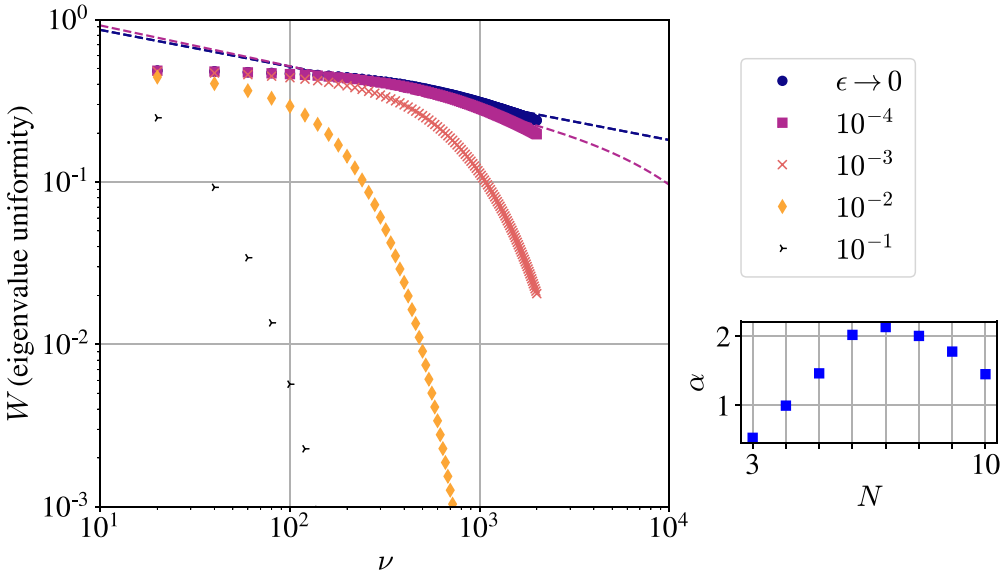


Figure A2. (left) TFI-HVA ansatz: same simulations as in figure 3(a) but with added parametrised R_z gates after each layer. The additional gates increase the dimensionality of the dynamic Lie algebra which leads to a faster scrambling of local gate noise into white noise, e.g. the $\epsilon \to 0$ curve is steeper than in figure 3(a). See appendix B for more details. (right) the dependence on the number of qubits shows a very similar trend as without the Rz gates, i.e. compare to figure A3(c).

Statement 1. We can approximate the white noise-state in equation (2) in terms of the dominant eigengvalue λ_1 and the dominant eigenvector $|\psi_1\rangle$ of the quantum state as

$$\rho_{wn} = \lambda|\psi_1\rangle\langle\psi_1| + (1-\lambda_1)\,\mathrm{Id}/d + \mathcal{E}_w, \tag{A.3}$$

up to an approximation error $\mathcal{E}_w$ that is bounded via equation (A.6).

Proof. We start by approximating the weight η in equation (2) as $\eta \approx F \approx \lambda_1$ via equation (9) as well as we approximate the dominant eigenvalue using equation (6) and then collect the approximation errors as

$$\rho_{wn} = \lambda|\psi_{id}\rangle\langle\psi_{id}| + (1-\lambda_1)\,\mathrm{Id}/d + \mathcal{E}_F + \mathcal{E}_C + O\left(\epsilon^2/\nu\right).$$

We now use results in [30] for bounding the distance between the ideal and noisy quantum states as

$$\left|\left||\psi_{id}\rangle\langle\psi_{id}| - |\psi_1\rangle\langle\psi_1|\right|\right|_1 = \sqrt{1-\langle\psi_{id}|\psi_1\rangle} = 1 - O\left(\frac{\mathcal{E}_C}{\lambda_1-\lambda_2}\right), \tag{A.4}$$

where $\mathcal{E}_C$ is the commutator norm from equation (6). We thus establish the approximation

$$\rho_{wn} = \lambda|\psi_1\rangle\langle\psi_1| + (1-\lambda_1)\,\mathrm{Id}/d + \mathcal{E}_w, \tag{A.5}$$

where we collect all approximation errors as

$$|\mathcal{E}_w| \leqslant |\mathcal{E}_F| + O\left(\epsilon^2/\nu\right) + O\left[\mathcal{E}_C\left(1 + \frac{1}{1-\lambda_2/\lambda_1}\right)\right]. \tag{A.6}$$

□

Statement 2. We define the eigenvalue uniformity as $W := \frac{1}{2}||p_{\text{err}} - p_{\text{unif}}||_1$ via the non-dominant eigenvalues of the density matrix $p_{\text{err}} := (\lambda_2, \lambda_3, \ldots, \lambda_d)/(1-\lambda_1)$. This metric is related to the trace distance from a white noise state (as in equation (4)) as

$$||\rho - \rho_{wn}||_1 = (1-\lambda_1)\, W + \mathcal{E}_w, \tag{A.7}$$

where the approximation error $\mathcal{E}_w$ is stated in statement 1.

Proof. We substitute the approximation of ρ_{wn} from equation (A.3) including the error term $\mathcal{E}_w$ and then we use the spectral decomposition of ρ to obtain the trace distance as

$$\begin{aligned} ||\rho - \rho_{wn}||_1 &= ||\sum_{k=2}^{d} \lambda_k |\psi_k\rangle\langle\psi_k| - (1-\lambda_1)\,\text{Id}/d||_1 + \mathcal{E}_w \\ &= \frac{1}{2}\sum_{k=2}^{d} \left|\lambda_k - \frac{1-\lambda_1}{d}\right| + \mathcal{E}_w \\ &= \frac{1-\lambda_1}{2}||p_{\text{err}} - p_{\text{unif}}||_1 + \mathcal{E}_w. \end{aligned}$$

In the second equation we analytically evaluated the trace distance and thus in the third equation we rewrite the result in terms of p_{err} which is our 'error probability' distribution as $p_{\text{err}} := (\lambda_2, \lambda_3, \ldots, \lambda_d)/(1-\lambda_1)$. □

Statement 3. Alternatively to statement 2, if a quantum state admits the decomposition in equation (8) then we can state the trace distance without approximation as

$$||\rho - \rho_{wn}||_1 = \frac{(1-\eta)}{2}||p_\mu - p_{\text{unif}}||_1. \tag{A.8}$$

This is directly analogous to the uniformity measure of the non-dominant eigenvalues of ρ in statement 2, however, this expression quantifies the uniformity of the probability distribution p_μ which are eigenvalues of the error density matrix ρ_{err}.

Let us assume the decomposition in equation (8). We find the following result via a direct calculation as

$$\begin{aligned} ||\rho - \rho_{wn}||_1 &= (1-\eta)\,||\rho_{\text{err}} - \text{Id}/d||_1 \\ &= (1-\eta)\,||\sum_{k=1}^{d} \mu_k |\phi_k\rangle\langle\phi_k| - \text{Id}/d||_1 \\ &= \frac{(1-\eta)}{2}||\sum_{k=1}^{d} |\mu_k - 1/d| \\ &= \frac{(1-\eta)}{2}||p_\mu - p_{\text{unif}}||_1 \end{aligned}$$

where we have used the spectral resolution of the error density matrix and then analytically evaluated the trace distance. Given ρ_{err} is a positive-semidefinite matrix with unit trace, its eigenvalues μ_k form a probability distribution that we denote as p_μ.

A.2. Upper bounding the uniformity measure

In this section we upper bound the uniformity measure based on the number of gates and error rates in a quantum circuit.

Statement 4. We adopt the bounds of [24] in equation (11) for the distance between probability distributions measured in the standard basis $\frac{1}{2}||\tilde{p}_{\text{noisy}} - \tilde{p}_{wn}||_1$ and assume the same bounds approximately apply to any measurement basis. Then, it follows that the uniformity measure from statement 2 is approximately bounded by the same bounds as

$$W = O\left(\frac{e^{-\xi}\xi/\sqrt{\nu}}{1-e^{-\xi}}\right) + O\left(\frac{\mathcal{E}_w}{1-\lambda_1}\right),$$

where the approximation error $\mathcal{E}_w$ is stated in statement 1.

Proof. Let us consider measurements performed in the basis as the eigenvectors of the density matrix which yield probabilities as the eigenvalues as

$$p_{\text{noisy}} = \langle\psi_k|\rho|\psi_k\rangle = (\lambda_1, \lambda_2 \ldots, \lambda_d).$$

Measuring the white noise state in the same basis yields the following approximation of the probabilities using the error term from equation (A.3) as

$$p_{wn} := \langle\psi_k|\rho_{wn}|\psi_k\rangle = \left(\lambda_1, \frac{1-\lambda_1}{d} \ldots, \frac{1-\lambda_1}{d}\right) + \mathcal{E}_w. \tag{A.9}$$

The distance of the above two measurement probability distributions is then

$$\frac{1}{2}||p_{\text{noisy}} - p_{wn}||_1 = (1-\lambda_1)\, W + \mathcal{E}_w,$$

where $W = \frac{1}{2}||p_{\text{err}} - p_{\text{unif}}||_1$ is our eigenvalue uniformity from statement 2. Under the assumption that the upper bound on the measurement probabilities $\frac{1}{2}||\tilde{p}_{\text{noisy}} - \tilde{p}_{wn}||_1$ from equation (11) approximately holds for any measurement basis we can bound the eigenvalue uniformity as

$$\begin{aligned} W &= \frac{1}{2(1-\lambda_1)}||p_{\text{noisy}} - p_{wn}||_1 + \frac{\mathcal{E}_w}{1-\lambda_1} \\ &\leqslant O\left(\frac{F}{1-\lambda_1}\epsilon\sqrt{\nu}\right) + \frac{\mathcal{E}_w}{1-\lambda_1} \\ &= O\left(\frac{e^{-\xi}\xi/\sqrt{\nu}}{1-e^{-\xi}}\right) + O\left(\frac{\mathcal{E}_w}{1-\lambda_1}\right). \end{aligned}$$

In the last equation we introduced the approximation of F from equation (9) as well as the approximate dominant eigenvalue from equation (6).

□

A.2.1. Expanding the upper bound. We now expand the upper bound from statement 4 for small ξ a.s. More specifically, we consider the parametrised fit function from equation (15) and substitute the Taylor expansion $e^{-\xi} = 1 - \xi + \xi^2 + \ldots$ as

$$\begin{aligned} \alpha\frac{e^{-\xi}\xi/\sqrt{\nu}^{\beta}}{1-e^{-\xi}} &= \alpha\frac{e^{-\xi}}{\nu^{\beta}}\frac{\xi}{\xi - \xi^2/2 + \ldots} \\ &= \alpha\frac{e^{-\xi}}{\nu^{\beta}}\frac{1}{1 - \xi/2 + \ldots} \\ &= \alpha\frac{1}{\nu^{\beta}}\frac{1-\xi+\ldots}{1-\xi/2+\ldots} = \frac{\alpha}{\nu^{\beta}} + O(\xi). \end{aligned}$$

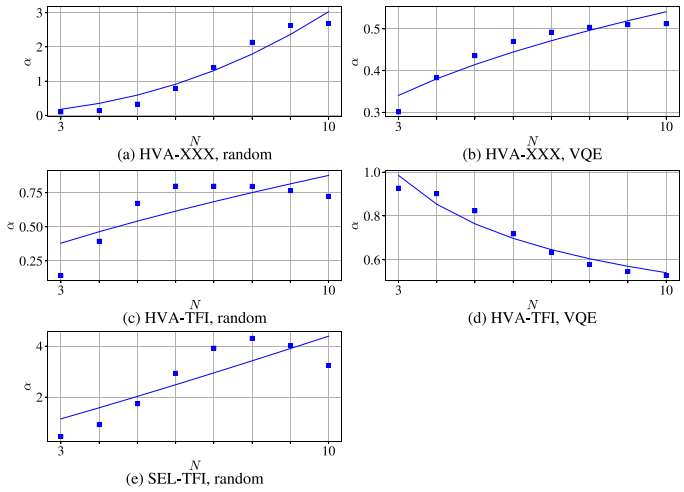


Figure A3. Fit parameters α from equation (15) for an increasing number of qubits: The circuits in figures 1–4 at $\epsilon \to 0$ were simulated for an increasing number of qubits and the curve from equation (15) was fitted.

A.3. Commutator norm

Lemma 1. *The commutators norms are approximately related as*

$$\frac{||\left[\rho_{id},\rho\right]||_1}{1-\lambda_1} = ||\left[\rho_{id},\rho_{\text{err}}\right]||_1 + \mathcal{E}_q, \tag{A.10}$$

up to the approximation error $\mathcal{E}_q$.

Proof. Using the decomposition from equation (8) we obtain

$$||\left[\rho_{id},\rho\right]||_1 = ||\left[\rho_{id},\eta\rho_{id}\right] + \left[\rho_{id},(1-\eta)\,\rho_{\text{err}}\right]||_1 = (1-\eta)\,||\left[\rho_{id},\rho_{\text{err}}\right]||_1 \tag{A.11}$$

We can approximate $\eta = \lambda_1 + \mathcal{O}(\mathcal{E}_F) + \mathcal{O}(\mathcal{E}_C)$ via equations (6) and (9) obtain that

$$\frac{||\left[\rho_{id},\rho\right]||_1}{1-\lambda_1} = ||\left[\rho_{id},\rho_{\text{err}}\right]||_1 + \mathcal{E}_q. \tag{A.12}$$

The error term can be obtained via the triangle inequality as

$$|\mathcal{E}_q| \leqslant \left[\mathcal{O}\left(\mathcal{E}_F\mathcal{E}_C\right) + \mathcal{O}\left(\mathcal{E}_C^2\right)\right]/\left(1-\lambda_1\right).$$

□

Appendix B. Further details of numerical simulations

B.1. The SEL and HVA ansätze

The circuit structure of the SEL ansatz used in figure 1 is illustrated in figure A1: it consists of alternating layers of parametrised single-qubit rotations and a ladder of nearest-neighbour CNOT gates.

Let us now define the HVA ansatz. In particular, recall that the HVA ansatz is a discretisation of the adiabatic evolution

$$U\left(\underline{\beta},\underline{\gamma}\right) = \prod_{k=1}^{\nu} e^{-i\gamma_k \mathcal{H}_1} e^{-i\beta_k \mathcal{H}_0},$$

which is applied to the initial state as the ground state of the trivial Hamiltonian $\mathcal{H}_0$.

The individual evolutions are then trotterised such that a piece of time evolution $e^{-i\gamma_k \mathcal{H}_1}$ is broken up into products of evolution operators under the individual Hamiltonian terms as

$$e^{-i\gamma_k \mathcal{H}_1} \to \prod_{l=1}^{r_h} e^{-i\gamma_k h_l P_l}.$$

Above we utilised the decomposition of the non-trivial part of the Hamiltonian $\mathcal{H}_1 = \sum_{l=1}^{r_h} h_l P_l$ into Pauli strings $P_l \in \{\mathrm{Id}, X, Y, Z\}^{\otimes N}$.

We set the circuit parameter as $\gamma_k = k/\nu$ and $\beta_k = 1 - k/\nu$, such that the circuit approximates a discretised adiabatic evolution between $\mathcal{H}_0$ and $\mathcal{H}_1$ – and we will refer to these as VQE parameters.

In the case of random parametrisation of the HVA ansatz, every gate implementing the evolution under a single Pauli string $e^{-i\gamma_k h_l P_l}$ is assigned a random parameter as $e^{-i\theta_q P_l}$ with $|\theta_q| \leqslant 2\pi$.

B.2. Inserting additional gates to the TFI ansatz

In figure A2 we repeated the same simulation as in figure 3(a), i.e. using a HVA ansatz for the TFI spin model at random circuit parameters, but we appended to each layer a series of parametrised Rz gates on each qubit. This guarantees that the dynamic Lie algebra generated by the Pauli terms of the TFI problem in equation (16) is expanded by the inclusion of Pauli Z operators. Increasing the circuit depth of the HVA ansatz thus leads to a faster increase of the dimensionality of the Lie algebra which demonstrably leads to a faster scrambling of local noise into global white noise, e.g. steeper slope of the $\epsilon \to 0$ fit in figure A2 than in figure 3.

B.3. Scaling with the number of qubits

In figure A3 we simulate the same circuits as in figures 1–4 at error rates $\epsilon \to 0$ and plot the fit parameter α—which is the prefactor in equation (15)—for an increasing number of qubits. The results appear to confirm an asymptotically non-increasing trend confirming theoretical expectations of [24] for random circuits whereby α is constant bounded in terms of the number of qubits.

ORCID iDs

Jonathan Foldager https://orcid.org/0000-0001-7470-0382
Bálint Koczor https://orcid.org/0000-0002-4319-6870

References

[1] Arute F *et al* 2019 *Nature* **574** 505–10
[2] Tillmann M, Dakić B, Heilmann R, Nolte S, Szameit A and Walther P 2013 *Nat. Photon.* **7** 540–4
[3] Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N and Lloyd S 2017 *Nature* **549** 195–202

[4] Jafferis D, Zlokapa A, Lykken J D, Kolchmeyer D K, Davis S I, Lauk N, Neven H and Spiropulu M 2022 *Nature* **612** 51–55
[5] Kokail C *et al* 2019 *Nature* **569** 355–60
[6] Cao Y *et al* 2019 *Chem. Rev.* **119** 10856–915
[7] McArdle S, Endo S, Aspuru-Guzik A, Benjamin S C and Yuan X 2020 *Rev. Mod. Phys.* **92** 015003
[8] Bauer B, Bravyi S, Motta M and Chan G K L 2020 *Chem. Rev.* **120** 12685–717
[9] Motta M and Rice J E 2022 *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **12** e1580
[10] Preskill J 2018 *Quantum* **2** 79
[11] Cerezo M *et al* 2021 *Nat. Rev. Phys.* **3** 625–44
[12] Endo S, Cai Z, Benjamin S C and Yuan X 2021 *J. Phys. Soc. Japan* **90** 032001
[13] Bharti K *et al* 2022 *Rev. Mod. Phys.* **94** 015004
[14] Cerezo M, Verdon G, Huang H Y, Cincio L and Coles P J 2022 *Nat. Comput. Sci.* **2** 567–76
[15] van Straaten B and Koczor B 2021 *PRX Quantum* **2** 030324
[16] Koczor B and Benjamin S C 2022 *Phys. Rev. Res.* **4** 023017
[17] Huang H Y, Kueng R and Preskill J 2020 *Nat. Phys.* **16** 1050–7
[18] Boyd G and Koczor B 2022 *Phys. Rev.* X **12** 041022
[19] Chan H H S, Meister R, Goh M L and Koczor B 2022 arXiv:2212.11036
[20] Jnane H, Steinberg J, Cai Z, Nguyen H C and Koczor B 2023 arXiv:2305.04956
[21] Cai Z, Babbush R, Benjamin S C, Endo S, Huggins W J, Li Y, McClean J R and O'Brien T E 2022 arXiv:2210.00921
[22] Koczor B 2021 *Phys. Rev.* X **11** 031057
[23] Huggins W J, McArdle S, O'Brien T E, Lee J, Rubin N C, Boixo S, Whaley K B, Babbush R and McClean J R 2021 *Phys. Rev.* X **11** 041036
[24] Dalzell A M, Hunter-Jones N and Brandão F G 2021 arXiv:2111.14907
[25] Larocca M, Czarnik P, Sharma K, Muraleedharan G, Coles P J and Cerezo M 2022 *Quantum* **6** 824
[26] Vovrosh J, Khosla K E, Greenaway S, Self C, Kim M S and Knolle J 2021 *Phys. Rev.* E **104** 035309
[27] Tsubouchi K, Sagawa T and Yoshioka N 2022 arXiv:2208.09385
[28] Endo S, Benjamin S C and Li Y 2018 *Phys. Rev.* X **8** 031027
[29] Strikis A, Qin D, Chen Y, Benjamin S C and Li Y 2021 *PRX Quantum* **2** 040330
[30] Koczor B 2021 *New J. Phys.* **23** 123047
[31] Koczor B and Benjamin S C 2022 *Phys. Rev.* A **106** 062416
[32] O'Brien T E *et al* 2022 arXiv:2210.10799
[33] Jnane H, Undseth B, Cai Z, Benjamin S C and Koczor B 2022 *Phys. Rev. Appl.* **18** 044064
[34] Schuld M, Bocharov A, Svore K M and Wiebe N 2020 *Phys. Rev.* A **101** 032308
[35] Koczor B, Endo S, Jones T, Matsuzaki Y and Benjamin S C 2020 *New J. Phys.* **22** 083038
[36] Foldager J, Pesah A and Hansen L K 2022 *Sci. Rep.* **12** 1–11
[37] Silva M, Magesan E, Kribs D W and Emerson J 2008 *Phys. Rev.* A **78** 012347
[38] Magesan E, Gambetta J M and Emerson J 2012 *Phys. Rev.* A **85** 042311
[39] Cai Z and Benjamin S C 2019 *Sci. Rep.* **9** 1–11
[40] Cai Z, Xu X and Benjamin S C 2020 *npj Quantum Inf.* **6** 1–9
[41] McClean J R, Boixo S, Smelyanskiy V N, Babbush R and Neven H 2018 *Nat. Commun.* **9** 1–6
[42] Cerezo M, Sone A, Volkoff T, Cincio L and Coles P J 2021 *Nat. Commun.* **12** 1–12
[43] Wecker D, Hastings M B and Troyer M 2015 *Phys. Rev.* A **92** 042303
[44] Wiersema R, Zhou C, de Sereville Y, Carrasquilla J F, Kim Y B and Yuen H 2020 *PRX Quantum* **1** 020319
[45] Campbell E 2019 *Phys. Rev. Lett.* **123** 070503
[46] Ouyang Y, White D R and Campbell E T 2020 *Quantum* **4** 235
[47] Koczor B, Morton J and Benjamin S 2023 arXiv:2305.19881
[48] Jones T, Brown A, Bush I and Benjamin S C 2019 *Sci. Rep.* **9** 1–11
[49] Meister R 2022 pyQuEST—a python interface for the quantum exact simulation toolkit (available at: https://github.com/rrmeister/pyQuEST)
[50] O'leary D and Stewart G 1990 *J. Comput. Phys.* **90** 497–505