ARTICLE

Decoding quantum errors with subspace expansions

Jarrod R. McClean[1]*, Zhang Jiang[1], Nicholas C. Rubin[1], Ryan Babbush [1] & Hartmut Neven[1]

With rapid developments in quantum hardware comes a push towards the first practical applications. While fully fault-tolerant quantum computers are not yet realized, there may exist intermediate forms of error correction that enable practical applications. In this work, we consider the idea of post-processing error decoders using existing quantum codes, which mitigate errors on logical qubits using post-processing without explicit syndrome measurements or additional qubits beyond the encoding overhead. This greatly simplifies the experimental exploration of quantum codes on real, near-term devices, removing the need for locality of syndromes or fast feed-forward. We develop the theory of the method and demonstrate it on an example with the perfect [[5, 1, 3]] code, which exhibits a pseudo-threshold of $p \approx 0.50$ under a single qubit depolarizing channel applied to all qubits. We also provide a demonstration of improved performance on an unencoded hydrogen molecule.

[1] Google Inc., 340 Main Street, Venice, CA 90291, USA. *email: jmcclean@google.com

Rapid developments in both the theory and hardware for quantum computation push us closer than ever to the dream of practically useful quantum computing. However, while a key development in the road map of quantum computing was the concept of quantum error correction, the hardware requirements to implement fully fault-tolerant schemes for non-trivial algorithms may still be some years away. A natural question that arises from this realization is whether it will be possible to perform meaningful computations on non-fault tolerant or noisy intermediate scale quantum computers (NISQ)[1]. Experimental and theoretical proposals have explored the potential for performing a well-defined computational task faster than a classical computer on as few as 50 qubits, a task often referred to as "quantum supremacy"[2–4]. It remains an open question, however, if these results can be extended to applications of interest outside the domain of pure computation.

Many early proposals for practical applications have advocated the use of variational algorithms[5–18], which are known to experience a natural form of robustness against certain types of noise. In conjunction with this, much progress has been made in reducing the gate overhead required for practical applications, especially in the domain of quantum chemistry[19–22]. However, the impact of incoherent noise remains daunting for the accuracy thresholds specified[5,7,23]. Moreover, while there has been some success in implementing early quantum error correcting code experiments in various architectures[24,25], a full implementation remains daunting. As a result, in order to reach practical applications, it may be necessary to implement some form of partial error correction for NISQ computations. The exact form of this error correction could take to achieve success is yet unknown; however, it has been suggested that one of the best applications for early quantum computers is using them to study and optimize error correcting codes in real conditions[26]. Yet despite great theoretical progress, most quantum codes are difficult to study experimentally on NISQ devices due to the need for complicated syndrome measurements, fast feedback, and decoding capabilities.

An alternative approach that strays from traditional ideas of error correction and targets NISQ devices is "error mitigation". This term largely refers to techniques that reduce the influence of noise on a result using only batch measurements and offline classical processing as opposed to active measurement and fast feedback type corrections. While they are not believed to lead to scalable, fault-tolerant computation, it is hoped that sufficient mitigation may open the possibility of practical applications or inspire more near-term error correction ideas. A number of these techniques have been developed both within an application specific and general context[27–30]. If one specializes to the quantum structure of fermionic problems, notably the N-representability conditions, enforcing these as constraints alone can reduce the impact of noise in simulations[31]. More generally within quantum simulation[32], an error mitigation technique known as the quantum subspace expansion (QSE)[27] was predicted and experimentally confirmed to both approximate excited states and reduce errors through additional measurements and the solution of a small offline eigenvalue problem[14]. Since then there have been variations leveraging QSE that use both additional techniques from quantum chemistry for excited states[33] and imaginary time evolution[34].

In this work, we show that it is possible both to use existing quantum error correcting codes to mitigate errors on NISQ devices and to study the performance of these codes under experimental conditions using classical post-processing and additional measurements. We briefly review the theory of stabilizer codes[35] and post-processing in this framework, which we then generalize using quantum subspace expansions. Although a connection to symmetries was explored in the original work[27]

and this connection was extended in subsequent work[36,37] that has also been verified by experimental implementation[38], these papers have focused on application specific contexts. Here we generalize this to any circuit performed within a quantum code, and show how subspace expansions may be used to then correct some logical errors within the code space, as well as be applied to approximate symmetries of unencoded Hamiltonians. We provide a concrete example using the perfect [[5, 1, 3]] code to demonstrate post-processed quantum state recovery. When applied at the highest level, this recovery exhibits a $p \approx 0.50$ pseudo-threshold for an uncorrelated depolarizing channel applied to all qubits. An example of an unencoded hydrogen molecule is also demonstrated across the entire range of depolarizing errors. We close with an outlook and potential applications of this methodology.

Results

Correcting logical observables in post-processing. We begin by briefly reviewing and establishing notation for the relevant topics of quantum error correcting codes in the stabilizer formalism, and using this formalism to develop a set of projection operators. Consider a set of n physical qubits. Quantum error correcting codes utilize entanglement to encode a set of $k < n$ logical qubits, with the hope of improving robustness to probable errors. A code that requires at least a weight d Pauli operator to induce a logical error is said to have distance d. These three numbers are often used to define a quantum error correcting code, with the notation $[[n, k, d]]$.

The set of $2k$ logical operators formally written $\mathcal{L} = \{\overline{X}_i, \overline{Z}_i\}_{i=1,\ldots,k}$ perform the desired Pauli operation on states in the code space, which is the ground state subspace of the code Hamiltonian.

$$\mathcal{H}_c = -\sum_{M_i \in \mathcal{M}} M_i \tag{1}$$

where $\mathcal{M}$ is a set of check operators drawn from the stabilizer group that can be used to deduce error syndromes. More explicitly, $\mathcal{S}$ is the set of stabilizer generators and S is the full stabilizer group implying $S \subseteq \mathcal{M} \subseteq \mathcal{S}$, such that the minimal set is the stabilizer generators, but additional operators from the stabilizer group may be added, as in techniques where one uses redundancy in stabilizer operators to ameliorate the need for multiple measurements passes, also called single shot error correction[39,40].

When not performing active error correction, one must rely on projection operations. Namely, members of the stabilizer group, $M_i \in \mathcal{S}$ have eigenvalues ± 1 and may each be used to construct a projector $P_i = (I + M_i)/2$ that removes components of the state outside the $+1$ eigenspace of the stabilizers or code space. One may use this to construct a projector that is a linear combination of projectors to remove desired errors outside the code space. We note that this clearly cannot remove logical errors made within the code space. This is related to the idea of error detection and post selection which has made recent progress both in theory and experiment[41–49]. Quantum error detection typically discards results based on syndrome measurements without using correction, however we will avoid the need for direct syndrome measurements, which can be cumbersome on geometrically local qubit layouts and are challenging to do in a fault tolerant fashion for complex codes. For a stabilizer group with generators S_i, the complete projector can be formed from $\prod_{S_i \in \mathcal{S}}(I + S_i)/2 = \prod_i P_i$. When taken over all the generators, the expression $\prod_i P_i$ is the sum of all elements of the stabilizer group with a constant coefficient

which we fix to $1/2^m$,

$$\overline{P} = \overline{P}^{\dagger} = \prod_i^m P_i = \frac{1}{2^m} \sum_{M_i \in \mathcal{S}} M_i \qquad (2)$$

where m is the number of stabilizer generators used. For the case of full projection, this will be the full stabilizer group which contains 2^m terms. While this is generally an exponential number of terms, it will be shown that the number of terms is not an explicit factor in the cost when a stochastic sampling scheme is used to apply the corrections. Rather the correction cost will depend on the volume of the state outside the code space. The group structure allows projective correction of the density matrix $\overline{P} \rho \overline{P}^{\dagger}$ on a NISQ device to be relatively straightforward.

It is important to emphasize here the reason this expansion of projectors is used here, as opposed to traditional measurement of stabilizer generators, which are not exponential in number. In particular, we are prescribing the use of transversal, destructive measurement to avoid the need for measurement syndrome qubits and associated fault-tolerant gadgets. As such, while stabilizer generators commute, transversal measurement of their components may not. For example, while $X_1 X_2$ and $Z_1 Z_2$ commute, measurement of Z_1 and Z_2 transversally to determine $Z_1 Z_2$ destroys the ability to recover $X_1 X_2$. Our scheme allows us to bypass this difficulty through stochastic operator sampling and works with codes of arbitrary structure. This transversal measurement scheme that avoids the need for ancilla is also at the core of the method's relatively high pseudothreshold for a given code.

Suppose that some logical Hermitian operator Γ is expressed as a sum of Pauli operators Γ_i as $\Gamma = \sum_i \gamma_i \Gamma_i$. Then the corrected value for the expectation of Γ may be computed from

$$\langle \Gamma \rangle = \frac{1}{c 2^m} \sum_{jk} \gamma_j \, \mathrm{Tr} \left[\rho \Gamma_j M_k \right] \qquad (3)$$

$$c = \mathrm{Tr} \left[\overline{P} \rho \overline{P}^{\dagger} \right] = \mathrm{Tr} \left[\overline{P} \rho \right] \qquad (4)$$

where we began with the standard expression for expectation values for an un-normalized quantum state, and used for reduction the properties of, Hermitian projectors, $\overline{P} = \overline{P}^{\dagger} = \overline{P}^{\dagger} \overline{P} = \overline{P} \overline{P}^{\dagger}$, and logical operators Γ commute with stabilizer group elements M_i, and

if $M_i^{\dagger} M_k$ is in the set of operators, we can rewrite it as a single sum over these operators which will be repeated. In the case that we use the operators built from the stabilizer generator projectors here, this will always be the case. As this expansion may contain a large number of terms, it is important to develop a scheme for sampling from it that maximizes efficiency. We discuss a simple stochastic scheme for sampling these corrections and the associated cost of doing so in the methods section.

We emphasize a distinction between this measurement scheme and traditional error correction/detection is that we do not need to measure the stabilizers in earnest. As this is a post-processing procedure, we are free to destroy the information in the state by measuring qubit-wise across Pauli operators. To be explicit, if one had the Pauli operator $X_1 Z_2 Z_3 X_4$ as a stabilizer, a true stabilizer measurement would require extracting only the ± 1 measurement using an ancilla. However in this scheme, we are free to use repeated preparations of the state and construct any unbiased estimator of $\langle X_1 Z_2 Z_3 X_4 \rangle$ we desire, including those which might destroy the encoded state. This dramatically simplifies the use of codes with non-local stabilizer measurements.

Relaxing projectors to subspace expansions. In the previous section, we showed how explicit projectors from quantum error correcting codes can be used to correct observables in post processing. We generally define the expansion of the problem around a reference quantum state into a small surrounding subspace as a quantum subspace expansion. Expanding the code projectors to the full group (as opposed to their stabilizer generators) is one restricted form of this, but the construction is much more general[27]. In this section, we show how these constructions can be relaxed for greater flexibility and power with simple relations to approximations of these projectors within a subspace. A schematic comparison of the deterministic subspace expansion and the stochastic variant we will use is shown in Fig. 1, which overviews the input and output quantities for each case. The conceptual overview for the goal of partitioning the space into good code space regions, and bad non-code space regions to be projected out is depicted in Fig. 2. We know that for an expansion built from a product of projection based on the stabilizer generators, the coefficients may be chosen to be uniform. However when one truncates terms from this series, this is no longer the
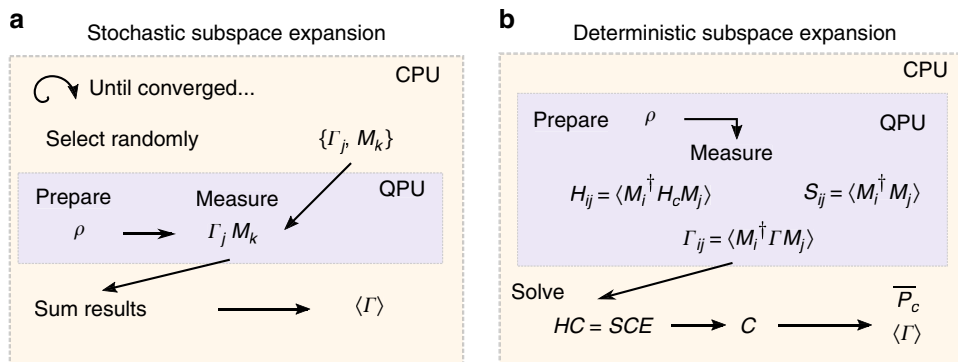

Fig. 1 Algorithmic schematics of the stochastic and deterministic subspace expansions. The goal is to use an expansion in a subspace around a prepared quantum state, ρ, to improve the expected value of the logical observable $\langle \Gamma \rangle$, without requiring ancilla based syndrome measurements or feedforward. An observable in the logical space Γ is expressed as a sum of Pauli operators, Γ_k, while symmetries, M_k, either naturally dictated by a system or from the stabilizer group S are selected. In the stochastic case (**a**), the state ρ is re-prepared many times, and these measurements are used to assemble the corrected expectation value $\langle \Gamma \rangle$ by expanding the averaged result in the resulting subspace. In the deterministic case (**b**), we may expand the set of M_k to include non-symmetries, and the corresponding averages over ρ are evaluated with many repetitions to form the representations of the operators in the subspace around ρ. These matrices define an offline generalized eigenvalue problem whose solution, C, defines both an optimal projector in the basis of operators M_i, $\overline{P}_c$ and corrected expectation values $\langle \Gamma \rangle$ for desired observables. We note that a scheme for including recovery operations can be found in the methods section.
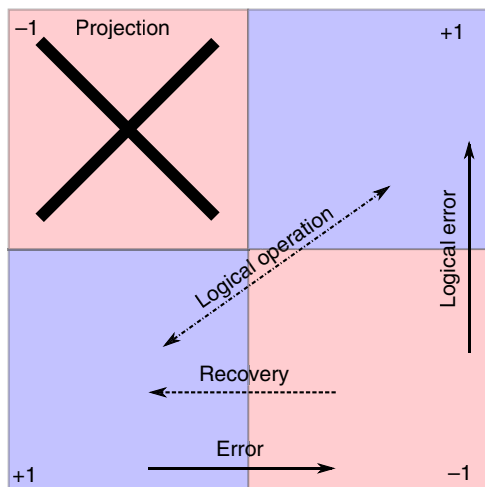
Fig. 2 Cartoon schematic of error correction vs error projection in a stabilizer code. We sketch the quantum space as divided into the blue code space, defined by $+1$ eigenvalues of stabilizers and the red non-code space as defined by having -1 eigenvalues for some of the stabilizers. In traditional error correction, the stabilizers are measured, the errors decoded, and recovery operations are applied to return one to the code space. In error projection, we use projectors based on stabilizers to remove sections of non-code space using only simple Pauli measurements and post-processing. We can also combine this technique with forms of recovery, but the effective difference is depicted by the discarding of large parts of errant Hilbert space.

case and we must consider a more general expression

$$\overline{P}_c = \sum_i^L c_i M_i \tag{5}$$

where L is the number of terms in the linear ansatz and the check operators, M_i, still come from the stabilizer group, however it no longer needs to be true that $\overline{P} \propto \prod_i (I + S_i)$. We choose a linear ansatz built from the stabilizer group to both facilitate the possible inclusion of symmetries as before, and to guarantee the existence of an exact solution for this problem. One could select more general operators than those in the stabilizer group, and indeed this was the approach to locating excited states in a previous work. If we choose $c_i = 1/2^m$ and let the sum run over the entire stabilizer group, we recover the expression from the projection technique. To find coefficients c_i, we formulate this problem as minimizing the distance to the code space subject to a normalization constraint. Using the Hamiltonian formulation of the code space, this is equivalent to approximating the ground state of the code space by

$$\min_{c_i} \operatorname{Tr}\left[\overline{P}_c \rho \overline{P}_c^\dagger H_c\right]$$
$$\text{such that } \operatorname{Tr}\left[\overline{P}_c \rho \overline{P}_c^\dagger\right] = 1 \tag{6}$$
$$\overline{P}_c = \sum_i c_i M_i.$$

This optimization is dependent both on the state ρ and choice of H_c in general. From the linear ansatz and normalization constraint, this problem is equivalent to minimization of a quadratic form on the surface of a sphere with a non-orthogonal metric. This general problem has a well-known solution in mathematics[50] and is commonly used within linear variational methods in chemistry[27] and physics derived from the Ritz method[51] or closely related Galerkin discretizations in applied

mathematics. The solution is given by the solution of the generalized eigenvalue problem

$$HC = SCE \tag{7}$$

$$H_{ij} = \operatorname{Tr}\left[M_i^\dagger H_c M_j \rho\right] \tag{8}$$

$$S_{ij} = \operatorname{Tr}\left[M_i^\dagger M_j \rho\right] \tag{9}$$

where H here forms a representation of the action of the code Hamiltonian in this stabilizer projector basis, the matrix S is the overlap or metric matrix defining the subspace geometry, C is the matrix of eigenvectors, and E is the diagonal matrix of eigenvalues. We note that for cases where the check operators M_i are built from projectors from generators, the solutions coincide with the previous formalism. When it is not the case that one builds the check operators from the product of stabilizer projectors, it provides an optimal solution that interpolates between different numbers of projectors in the subspace given. This type of expansion about a state is referred to as a quantum subspace expansion (QSE). The ground state eigenvector of this $N_M \times N_M$ eigenvalue problem, forms the optimal solution of the above problem within this subspace. In general, the optimal solution of this eigenvalue problem may not be a strict projector. However, this is not necessarily undesirable. In some cases, a lower energy state may be found for a specific problem Hamiltonian that corresponds to a rotation in the logical space.

When the solution of this eigenvalue problem is obtained, it may be used in a number of ways. In connection to the stochastic technique prescribed above, the values c_i may be used to obtain a corrected estimate of any desired logical observable. Alternatively, one may build a representation of an operator in the same basis defined by the expansion operator, and use it to perform further symmetry projections or improve the estimate of a logical observable as in ref. [27].

As in typical quantum error correction, the degeneracy of the ground state of the full code Hamiltonian prevents referencing a single state within the code space. This makes removing logical errors with the above procedure impossible. However, when considered in conjunction with a problem Hamiltonian such as that from a quantum physical system like an electronic system, it becomes possible to correct logical errors as well if the goal is to prepare an eigenstate of this Hamiltonian or minimize its energy. As a simple example, if one has a single encoded spin with a problem Hamiltonian $-\overline{Z}_i$ and the state is incorrectly found in $|\overline{1}\rangle$, by including as an expansion operator $\overline{X}_i$, one can correct an error in the logical space with this procedure, as it will detect the lower energy state, $|\overline{0}\rangle$, to be in the expanded subspace.

Corrections with unencoded systems
So far we have considered the case of decoding within an error correcting code that redundantly encodes quantum information via engineered symmetries. However, this strategy inevitably involves some overhead due to the encoding in the execution of gates, and in some near-term experiments, it will still be most practical to work directly in the space of a physical problem Hamiltonian H_p. Here we show how the machinery developed so far can be applied to this case.

In this case, the physical problem Hamiltonian H_p may have symmetries that are often known about the desired state ahead of time. For example in the case of an interacting fermion system, the total number of fermions, the total spin and S_z component, and symmetries related to spatial degrees of freedom in a system are often good candidates. The application of these symmetries has been explored previously, symmetries[27,36,38,52], and as these

are expected to be exact symmetries, it is always safe to apply them when the symmetry is known.

While these symmetries are exact and effective to apply, they are often more expensive to implement than one might desire. For example, the problem of number symmetry in a fermion Hamiltonian can take eigenvalues that range from 0 to the number of spin orbitals in the system. Thus to select just the correct particle number, one may have to construct a projector which removes all the components except the desired particle number N_p, or $\propto \prod_{n \neq N_p} (n - \hat{N})$ where $\hat{N}$ is the number operator on all the fermionic modes of the system. This may result not only in many terms, but a sum of terms that do not individually represent symmetries of the system, which can complicate sampling.

As a result, it is much simpler and effective to start with symmetries that have only two distinct eigenvalues, also referred to as $\mathbb{Z}_2$ symmetries of the problem Hamiltonian. An example of this is the number symmetry operator, which in the Jordan-Wigner representation takes the simple form $\prod_i Z_i$. An extension of this, is to use both the up-spin (α) and down-spin (β) number parities, which generate the full number parity, and offer additional power in their projection. These are simply given by $\prod_{i \in \alpha} Z_i$ and $\prod_{i \in \beta} Z_i$ respectively. These simple parity symmetries have been utilized before to reduce the number of qubits[52], however just as in subsystem error correcting codes, retaining these redundancies can sometimes be beneficial for gate depth or efficiency of representation. That work also contained a general algorithm for searching for unknown $\mathbb{Z}_2$ symmetries in these Hamiltonians that can be used here. One may also use this to consider beneficial projectors that are derived from approximate symmetries of the Hamiltonian or more generally operators that do not commute with the Hamiltonian, but have known structure with regards to the problem.

Example demonstrations
Here we both exhibit some of the performance of the presented techniques and clarify their construction through the use of simple examples. Both a general error correcting code and specific problem Hamiltonian systems with symmetries are studied. We note that in our numerical studies we use a single qubit depolarizing channel defined by

$$\mathcal{E}_p(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z) \quad (10)$$

which corresponds to the convention that the totally mixed state is achieved at $p = 3/4$. Although a single qubit depolarizing channel is not a perfect error model for any system, experimental data suggests that correlated errors are much weaker in many architectures than independent, single-qubit errors. In addition, it is expected that in the gate model, use of randomized compiling is advisable to decohere errors, which tends to map other error channels closer to this model on average and allows one to treat phenomena like measurement errors as bit-flip errors. Application of these techniques and analysis to true devices is, of course, the eventual goal in order to both understand their applicability and optimize existing quantum codes.

To see how the general recovery process using stabilizer codes can work in practice, let us consider the concrete example of the perfect [[5, 1, 3]] code, which is a distance 3 code that encodes 1 logical qubit in 5 physical qubits. To evaluate the performance in practice, we perform the following numerical experiment. A logical state $|\overline{\Psi}\rangle$ in the [[5, 1, 3]] code is prepared, then subjected to an uncorrelated depolarizing channel on all qubits with probability p. The logical state is selected at random within the space to not exhibit any special properties with regards to errors. In connection with the formalism above, we evaluate the

expectation value of the logical operator $A = |\overline{\Psi}\rangle\langle\overline{\Psi}|$. This operator does not generally have a simple Pauli expansion as other observables typically would, but gives a stringent test for the performance of the method for all observables on the state of interest. The subspace expansion is then performed with $\mathcal{S}^{(l)}$ as expansion operators, and the fidelity $F_L = |\overline{\Psi}\rangle\langle\overline{\Psi}|$, of the resulting state is evaluated with $|\overline{\Psi}\rangle$.

This code has logical operators and stabilizer generators

$$\overline{X} = XXXXX \quad (11)$$

$$\overline{Z} = ZZZZZ \quad (12)$$

$$S = \{XZZXI, IXZZX, \\ XIXZZ, ZXIXZ\}. \quad (13)$$

We denote the two states of the logical qubit as $|\overline{0}\rangle$ and $|\overline{1}\rangle$, and an arbitrary code space state that is a superposition of these two states as $|\overline{\Psi}\rangle$ or the pure state density matrix $\overline{\rho}$.

We denote the hierarchy of check operators as the elements in the sum generated by $S^{(l)} = \prod_{i=1}^{l}(I + S_i)$, where the ordering of stabilizer generators has been fixed. To see how this hierarchy works in practice, consider an uncorrelated depolarizing channel acting on all 5 physical qubits with probability p. In this situation, we have up to 5 qubit errors, which we do not expect the code can recover from without introduction of a problem Hamiltonian, however they occur with probability p^5, which can be quite small for modest p.

The cross lines in Fig. 3 show the performance using fixed projectors at those level, which exactly coincide with the QSE relaxation. The starred lines show the result of removing 2 check operators at random and re-performing the QSE expansion to show the performance smoothly interpolates between those limits. We plot the logical infidelity $1 - F_L$ where the physical line denotes the trivial encoding into one qubit and compare the two for a range of values of p. The infidelity is calculated with respect to the correct logical state, which is a single state in the physical Hilbert space. This is done as we are seeking to discard all non-codeword states, but cannot do so when encoded yet uncorrected. Hence for the encoded but uncorrected state, fully depolarizing each qubit individually will lead to a fidelity of $1/2^n$, where n is the number of physical qubits. We define the pseudo-threshold to be the value of p for which the logical infidelity in the encoded space is lower than the physical infidelity for the unencoded system, and we see that at both levels there is a pseudo-threshold in this model. For $S^{(4)}$, the pseudo-threshold is numerically found to be $p = 0.50$ for this code and symmetric depolarizing channel. We see here that the pseudo-threshold generally depends on the truncation level, however it converges to the pseudo-threshold of the full code for the given operation and error model. If we are sufficiently below the corresponding pseudo-threshold, in the case where no recovery operations are used, we may obtain up to a p^{d-1} suppression of the error rate. A topic of interest is to investigate how this pseudothreshold may relate to the distance in the more general case using our technique which avoids the need for ancilla measurement qubits.

We now look at an example of an unencoded Hamiltonian, which has become a canonical test case for quantum computing in quantum chemistry. This is the second quantized hydrogen molecule in a minimal, 4 qubit basis. In this case, the α and β number parity operators are an efficient choice of projectors. In the Jordan-Wigner representation, these are given by $\prod_{i \in \text{even}} Z_i$ and $\prod_{i \in \text{odd}} Z_i$ when an even-odd ordering of orbitals are used.

The effectiveness of the technique on this system is evaluated numerically by preparing the exact ground state of the hydrogen molecule and subjecting it to an independent depolarizing
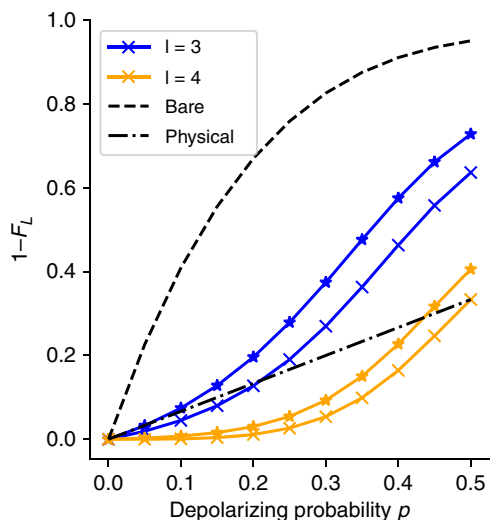
Fig. 3 Pseudo-threshold crossover for recovery using the [[5, 1, 3]] code.
The model examines the impact of errors under an uncorrelated depolarizing channel showing a $p \approx 0.50$ pseudo threshold for the full correction procedure. We plot the logical infidelity $1 - F_L$ where F_L is the logical fidelity of a selected state in the code space of the [[5, 1, 3]] code as a function of the depolarizing probability for each qubit p. The label l denotes the number of products from the stabilizer generators used in the expansion operator set. The physical line depicts the same error if the logical state is encoded in a single qubit in the standard way. The "bare'' line indicates the logical error rate with no recovery procedure applied. The starred lines close to each level of the hierarchy show an approximation to that level of projection using the QSE projection with 2 less check operators to demonstrate the smooth performance of the subspace procedure.

Fig. 4 Error suppression using natural molecular symmetries. Errors are plotted as a function of depolarizing probability and included projectors for an H_2 molecule at a bond length of 1.50 $\mathring{A}$ with a physical encoding. F_L here is the logical fidelity, which is the same as physical for this encoding. At this bond length, the ground state wavefunction requires considerable entanglement to be qualitatively correct. The stabilizer elements used to perform the projection here are the α- and β- number parity as well as the total X operator, which need not be an exact symmetry, given in the Jordan-Wigner encoding as $S = \{Z_0 Z_2, Z_1 Z_3, X_0 X_1 X_2 X_3\}$. The depolarizing channel is applied individually to each qubit on the exact ground state. As there is no error correction encoding overhead in this case, the expected improvement over the uncorrected solution is always positive. Hence it is always advantageous in these cases where simple symmetries are known to include these measurements and corrections.

channel on all 4 qubits. We choose as the generating operators $S = \{Z_0 Z_2, Z_1 Z_3, X_0 X_1 X_2 X_3\}$ which up- and down-spin number parity operators as well a non-local operator which need not be an exact symmetry of the Hamiltonian. The ordering of l matches that given here. The logical infidelity as a function of the depolarizing probability is plotted in Fig. 4. In contrast to the encoded case, we see an improvement over the whole range of depolarizing strengths. The stretched geometry of the molecule ensures that a high degree of entanglement is required to achieve a low logical fidelity, making this a sensitive test of performance. We see that in some cases an improvement of up to 3x in the logical infidelity. As the number of operators to measure here is quite modest and the improvement is universal, it suggests this will present an advantageous correction to include in almost all near-term implementations.

Discussion
As has been conjectured before, one of the best uses of early quantum devices may be to tune quantum error correcting codes under actual device conditions[26]. The modeling of true noise within the device is incredibly difficult as the system size grows, and studying which codes excel under natural conditions and how to optimize them may lead to progress towards fully fault tolerant computation. Indeed, knowledge of biased noise sources can vastly increase the threshold of a given code[53]. The tool we have provided here gives a method to experimentally study the encoding through post-processing while removing the complication of fault-tolerant syndrome measurement or fast feedback. This allows one to explore a wider variety of codes experimentally before worrying about these final details. We propose that one can run simple gate sequences in the logical space with known
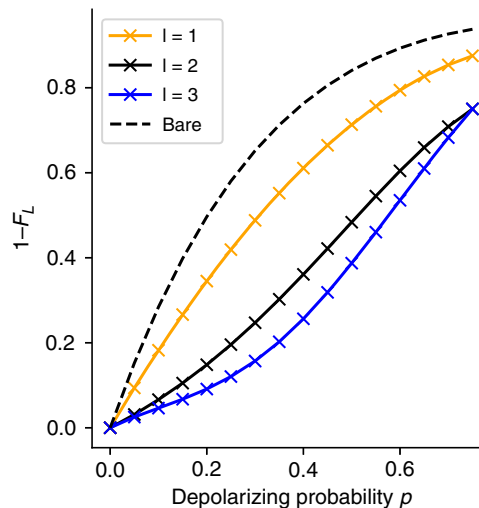
results, and use the post processing decoder here to study the decay of errors as stabilizers are added. This limit will inform the propagation of logical errors in the system and allow one to make code optimizations before full fault tolerant protocols are available.

We note that this type of decoder benefits greatly from the fact that check measurements need not be geometrically local to be implementable in a realistic setting. This allows one to explore and utilize codes that are not geometrically local on the architecture in use, which may have nicer properties with respect to distance and rate than geometrically local codes. Moreover, they naturally allow implementation of recent fermionic based codes, such as Majorana loop stabilizer codes[54] or variations of Bravyi-Kitaev superfast[55] thought to be good candidates for near-term simulations, without the need for complicated decoding circuits or ancilla for syndrome measurements.

Moreover, as this method is a post-processing method, it is entirely compatible with the extrapolation techniques introduced for error mitigation[28–30]. In these techniques, one artificially introduces additional noise to extrapolate to a lower noise limit. To apply this technique to quantum subspace expansions, one simply needs to perform the extrapolation on each of the desired matrix elements, then proceed as normal.

In this work, we have introduced a method for mitigating errors and studying error correcting codes using a post processing technique based on quantum subspace expansions. We showed that in implementations of this method, one can achieve a pseudo-threshold of $p \approx 0.50$ under a depolarizing channel acting on single qubits in the [[5, 1, 3]] code and made connections to the traditional theory of stabilizer codes. We believe this method has the potential to play a role in the development and

optimization of quantum codes under realistic noise conditions as well as the ability to remove errors from early application initiatives.

Methods

Stochastic sampling for corrections. As suggested in the exposition on the correction formalism, it is key for efficiency to sample the terms for the correction in a way that reflects the state rather than number of terms. We outline and analyze a simple stochastic scheme for performing this sampling here.

Suppose that we want to measure the corrected expectation value of some logical operator Γ, which can be decomposed into Pauli operators Γ_j as $\Gamma = \tilde{\gamma} \sum_j \gamma_j \Gamma_j$, where $\sum_j \gamma_j = 1$ and $\gamma_j \geq 0$ from absorbing required signs into Γ_j. Projecting this into the code space of some selected code as before, we have

$$\mu_\Gamma = \mathrm{Tr}\left[\overline{P} \rho \overline{P} \Gamma\right]$$

$$= \frac{\tilde{\gamma}}{2^m} \sum_{ij} \gamma_j \,\mathrm{Tr}\left[\rho \Gamma_j M_i\right] \tag{14}$$

$$= \tilde{\gamma} \sum_{\chi \in \{0,1\}^m} \sum_j \frac{\gamma_j}{2^m} \Gamma_{\chi,j} \tag{15}$$

where $\chi = (\chi_1, \chi_2, \ldots, \chi_m)$ is a bit string that we use to conveniently enumerate the stabilizer group operators and we define

$$\Gamma_{\chi,j} = \mathrm{Tr}\left[\rho \Gamma_j S_\chi\right], \quad S_\chi = S_1^{\chi_1} S_2^{\chi_2} \cdots S_m^{\chi_m} \tag{16}$$

where we note that this also encompasses the measurement of the normalization correction c as well for $\Gamma = I$. We will also assume here that as in the case of stabilizer projectors and the $\mathbb{Z}_2$ construction of physical symmetries that each M_i is a symmetry of the encoding, commuting with the desired Hamiltonian and physical observables.

To sample the trace stochastically, we may use the coefficients of the terms as a normalized probability distribution. One may choose the distribution to depend on χ as in an importance sampling scheme below, however taking the uniform distribution is perhaps the most straightforward and $p_{\chi,j} = \frac{\gamma_j}{2^m}$ gives the mean

$$\mathbb{E}\left[\hat{\mu}_\Gamma\right] = \tilde{\gamma} \sum_{\chi,j} p_{\chi,j} \mathbb{E}\left[\Gamma_{\chi,j}\right]$$

$$= \tilde{\gamma} \sum_{\chi,j} p_{\chi,j} (q_{\chi,j}^1 - q_{\chi,j}^{-1}) \tag{17}$$

$$= \tilde{\gamma} \sum_{\chi,j,x \in \{-1,1\}} x \cdot p_{\chi,j,x}$$

where we used $\hat{\mu}_\Gamma$ to emphasize that this is an expected value for our estimator of μ_Γ, $q_{\chi,j}^x$ is the probability of getting a measurement result $x \in \{+1, -1\}$ from measuring the Pauli operator $\Gamma_{\chi,j}$, and we have lumped this into the probability distribution as $p_{\chi,j,x} = p_{\chi,j} q_{\chi,j}^x$. This has a simple construction for stochastic evaluation, which is to enumerate all the possible terms in the decomposition, draw N_s terms with the probabilities from this distribution which will yield either $+1$ or -1 from the Pauli measurements, add them together and divide by N_s. The variance in the estimate will be given by the variance of this estimator divided by N_s.

From our construction, we see that we can view the estimator as a binomial distribution with a probability for the two results $x \in \{-1, +1\}$ derived from marginalizing over the joint distribution for projector terms χ and Pauli decomposition terms j to find

$$p_x = \sum_{\chi,j} p_{\chi,j,x}. \tag{18}$$

As a result, one may write down a particularly simple form of the variance for the estimator given by

$$Var[\hat{\mu}_\Gamma] = \tilde{\gamma}^2 p_{+1}(1 - p_{+1}). \tag{19}$$

To understand how the state influences the variance, we consider a simple example case using the total depolarizing channel and a single Pauli operator Γ, with $\gamma_i = \tilde{\gamma} = 1$. For the total depolarizing channel with probability w, we have

$$\rho = (1 - w)|\overline{\Psi}\rangle\langle\overline{\Psi}| + \frac{w}{2^n} \mathbb{1}. \tag{20}$$

In the limit of $w = 0$ we have that all states are in the code space, and hence the sum over χ trivially collapses, and we have that $p_x = q^x$, which gives the same statistics as the original measurement of Γ. Thus in such a case one has no dependence on the number of stabilizer terms used in the expansion. Hence, adding this procedure to a perfect state is expected to incur no additional cost on average.

Considering the imperfect case $w > 0$, marginalizing over χ is equivalent to applying the code space projector and hence geometrically analogous to determining the volume of the state in the code space. This must correspond equate to a portion of the average being 0, however as we are only capable of measuring 1 and -1, it then must constitute an equal probability of being in $+1$

and -1 that is determined by the volume of non-code space the state occupies. More explicitly

$$\sum_\chi p_{\chi,j,x} = \frac{1}{2}(1 + x \,\mathrm{Tr}\left[\rho \overline{P}\right]) p_{j,x}. \tag{21}$$

As a single Pauli is traceless, in our simple example we find for the case of the totally mixed state that

$$p_{+1} = \frac{1}{2}(2 - w)q^{+1} \tag{22}$$

$$Var[\hat{\mu}_\Gamma] = \frac{1}{4}(2 - w)w(q^{+1})^2 \tag{23}$$

to further simplify, suppose we were measuring the $+1$ eigenstate of Γ, so that $q^{+1} = 1$, then

$$Var[\hat{\mu}_\Gamma] = \frac{1}{4}(2 - w)w \tag{24}$$

then we see as expected, that for a perfect state ($w = 0$) the variance is minimal and independent of the number of terms, and that the variance increases as the state quality degrades.

The general picture of viewing the sum over χ as reflecting the volume of space attached to the projector lets one easily reason about the generalization of this scheme to sampling with recovery. In that case, we simply attach one more probability which allows us to sample over the different selected errors E_i, with probability p_{E_i} and proceed as before. The key difference is that we see the marginalization over χ is now better conditioned, as it is determined by the ratio of the volume of recovered space to the volume of Hilbert space rather than the volume of the code space to the volume of Hilbert space. As a result, the sample variance for recovery may be lower than the sample variance for strict projection. However as discussed, the ultimate quality of recovery is expected to be superior for strict projection due to removing errors up to weight $d - 1$ instead of $(d - 1)/2$.

One potential way to suppress the numbers of samples required is to use variance reduction techniques such as importance sampling. This approach requires a priori knowledge of the values of $\Gamma_{\chi,j,x}$, and samples the high weight terms preferentially while applying a correction to the measured values to remain unbiased. One approach is to sample the bit string χ according to its Pauli-weight W_χ, i.e, the number of qubits that the stabilizer operator acts non-trivially on. For the single-qubit depolarizing channel with error probability p, one may sample the bit string χ with probability proportional to $(1 - p)^{W_\chi}$. This is based on the intuition that the quantum information stored in low-weight operators decays more slowly under local noisy channels.

To be more explicit in the construction for the random sampling method discussed here applied to the recovery procedure discussed in Section "Corrections with recovery operations". Suppose that the projector on the subspace corresponding to the error operator E_α is

$$\overline{P}_\alpha = \prod_j \frac{1}{2}(1 + (-1)^{s_{\alpha,j}} S_j), \tag{25}$$

where $s_{\alpha,j} = 0, 1$ and $s_\alpha = (s_{\alpha,1}, s_{\alpha,2}, \ldots, s_{\alpha,m})$ are the syndromes of the α-th error. The recovered state takes the form

$$\mathcal{R}(\rho) = \sum_\alpha R_\alpha \overline{P}_\alpha \rho \overline{P}_\alpha R_\alpha^\dagger \tag{26}$$

$$= \overline{P}\left(\sum_\alpha R_\alpha \rho R_\alpha^\dagger\right)\overline{P}. \tag{27}$$

The expectation value for the recovered state reads

$$\mathrm{Tr}\left[\mathcal{R}(\rho)\Gamma\right] = \sum_{\alpha,j} \mathrm{Tr}\left(R_\alpha \rho R_\alpha^\dagger \overline{P}\Gamma_j \overline{P}\right) \tag{28}$$

$$= \frac{1}{2^m} \sum_{\alpha,\chi,j} \mathrm{Tr}\left(R_\alpha \rho R_\alpha^\dagger \Gamma_j S_\chi\right) \tag{29}$$

where $\Gamma_\alpha = R_\alpha^\dagger \Gamma R_\alpha$ is a logical operator. Hence if we absorb the sign into the logical operator through a careful choice of recovery operator, we may generalize the stochastic sampling scheme to taking the expectation value as

$$\mathbb{E}[\hat{\mu}_\Gamma] = \tilde{\gamma} \sum_{\chi,j,\alpha,x \in \{-1,1\}} x \cdot p_{\chi,j,\alpha,x} \tag{30}$$

where for uniform sampling we choose

$$p_{\chi,j,\alpha} = \frac{\gamma_i b_\alpha}{2^m} \tag{31}$$

with $\sum_\alpha b_\alpha = 1$ and $b_\alpha > 0$ by choice of the recovery operators. In this case, the stochastic sampling algorithm is given by choosing a Pauli operator defined by $R_\alpha^\dagger \Gamma_j S_\chi R_\alpha$ with probability $p_{\chi,j,\alpha}$, recording the series of $+1$, -1 results, and finding their average as before. The calculation of the variance follows as in the previous case. To reduce variance, the frequency of sampling α maybe chosen to be

proportional to the error probability of E_α for the expected errors on the physical system of interest.

Corrections with recovery operations

The power of error correction extends beyond the simple identification of errors and includes recovery operations that restore some states to the original code space. The formalism here built on projectors and post-processing would seem at first glance unable to take advantage of such unitary projection operations; however, we will show how one can use these recovery operations to some advantage in sampling complexity over the unrecovered projections.

Consider a set of Pauli errors on the system of physical qubits $\{E_i\}$ which is known to be correctable within the chosen code. These errors will either commute or anti-commute with the stabilizers of the code to produce a syndrome of the error that has happened, which we denote s_j^i for the j'th syndrome measurement of the i'th error. We will assume the recovery operation for this error within the code is known, and is denoted as R_i.

The formalism presented here avoids direct stabilizer measurement by design to favor implementation on NISQ devices, hence we need to specify how one uses recovery operations within the projection formalism. Similar to a projector on the code space, we may formulate a projector onto the error subspace that corresponds to error E_i acting on the code space. This is given by

$$\overline{P}_{E_i} = \prod_j \frac{1}{2}(I + (-1)^{s_j^i} S_j) \tag{32}$$

where $s_j^i \in \{0, 1\}$ is the syndrome associated with the error E_i and stabilizer generator S_j. Once one has projected into this space, we can now use the recovery operation, R_i to map the state back into the code space before using it. If we take the set of all correctable errors, including no error as the identity, then we get an updated correction formula for projection with recovery as

$$\langle \Gamma \rangle = \frac{1}{c} \sum_i \mathrm{Tr}\left[R_i \overline{P}_{E_i} \rho \overline{P}_{E_i}^\dagger R_i^\dagger \Gamma\right] \tag{33}$$

$$c = \sum_i \mathrm{Tr}\left[\overline{P}_{E_i} \rho\right] \tag{34}$$

where now we have assumed the ability to apply the recovery operations R_i, however in many cases this again reduces to a simple sum over Pauli operators that may be stochastically sampled, where many of the same simplifications resulting from commutation of logical operators with stabilizers are possible.

The consequences of including recovery on top of projection are interesting. The immediate practical benefit in increasing the size of Hilbert space over which one attains signal. This is reflected in the estimation of the value c, and leads in practice to lower errors with small, finite samples. However a tradeoff is being made in including these values in that it may reduce the overall projection quality. Consider for example a distance d code. If one has a Pauli error with weight greater than $(d-1)/2$, a recovery operation may become a logical error which is not then removed by this procedure. In contrast, strict projection is capable of removing errors of all the way up to weight $d-1$, which is a significant boost in maximum potential. However, as mentioned, the tradeoff of finite sampling complexity with potential for correction must be carefully balanced in real implementations.

Data availability
Data used in the generation of figures is available upon request.

Code availability
Sample code for the deterministic subspace expansion available upon request.

References
1. Preskill, J. Quantum Computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
2. Boixo, S. et al. Characterizing quantum supremacy in near-term devices. *Nat. Phys.* **14**, 595–600 (2018).
3. Neill, C. et al. A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science* **360**, 195–199 (2018).
4. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
5. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
6. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 023023 (2016).
7. O'Malley, P. J. J. et al. Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 031007 (2016).
8. Farhi, E., Goldstone, J., & Gutmann, S., A Quantum Approximate Optimization Algorithm, ArXiv e-prints, http://arxiv.org/abs/1411.4028 (2014).
9. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).
10. Shen, Y. et al. Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. *Phys. Rev. A* **95**, 020501 (2017).
11. Santagati, R. et al. Witnessing eigenstates for quantum simulation of hamiltonian spectra. *Sci. Adv.* **4**, eaap9646 (2018).
12. Dumitrescu, E. F. et al. Cloud quantum computing of an atomic nucleus. *Phys. Rev. Lett.* **120**, 210501 (2018).
13. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
14. Colless, J. I. et al. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X* **8**, 011021 (2018).
15. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* **4**, 014008 (2018).
16. Hempel, C. et al. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* **8**, 031022 (2018).
17. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **9**, 4812 (2018).
18. Farhi, E. & Neven, H. Classification with quantum neural networks on near term processors, arXiv preprint arXiv:1802.06002 (2018)
19. Motta, M. et al. Low rank representations for quantum simulation of electronic structure, arXiv preprint arXiv:1808.02625 (2018)
20. Kivlichan, I. D. et al. Quantum simulation of electronic structure with linear depth and connectivity. *Phys. Rev. Lett.* **120**, 110501 (2018).
21. Babbush, R. et al. Encoding electronic spectra in quantum circuits with linear t complexity. *Phys. Rev. X* **8**, 041015 (2018a).
22. Babbush, R. et al. Low-depth quantum simulation of materials. *Phys. Rev. X* **8**, 011044 (2018b).
23. Sawaya, N. P. D., Smelyanskiy, M., McClean, J. R. & Aspuru-Guzik, A. Error sensitivity to environmental noise in quantum circuits for chemical state preparation. *J. Chem. Theory Comput.* **12**, 3097 (2016).
24. Nigg, D. et al. Quantum computations on a topologically encoded qubit. *Science* **345**, 302–305 (2014).
25. Barends, R. et al. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500–503 (2014).
26. Iyer, P. & Poulin, D. A small quantum computer is needed to optimize fault-tolerant protocols. *Quant. Sci. Technol.* **3**, 030504 (2018).
27. McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95**, 042308 (2017).
28. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).
29. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
30. Otten, M. & Gray, S. K. Accounting for errors in quantum algorithms via individual error reduction. *npj Quant. Inf.* **5**, 11 (2019).
31. Rubin, N. C., Babbush, R. & McClean, J. Application of fermionic marginal constraints to hybrid quantum algorithms. *New J. Phys.* **20**, 053020 (2018).
32. Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).

33. Parrish, R. M., Hohenstein, E. G., McMahon, P. L. & Martínez, T. J. Quantum computation of electronic transitions using a variational quantum eigensolver. *Phys. Rev. Lett.* **122**, 230401 (2019).

34. Motta, M., et al. Quantum imaginary time evolution, quantum lanczos, and quantum thermal averaging. *Nat. Phys.* https://doi.org/10.1038/s41567-019-0704-4 (2019)

35. Gottesman, D., Stabilizer codes and quantum error correction, arXiv preprint quant-ph/9705052 (1997)

36. Bonet-Monroig, X., Sagastizabal, R., Singh, M. & O'Brien, T. E. Low-cost error mitigation by symmetry verification. *Phys. Rev. A* **98**, 062339 (2018).

37. McArdle, S., Yuan, X. & Benjamin, S. Error-mitigated digital quantum simulation. *Phys. Rev. Lett.* **122**, 180501 (2019).

38. Sagastizabal, R. et al. Experimental error mitigation via symmetry verification in a variational quantum eigensolver. *Phys. Rev. A* **100**, 010302 (2019).

39. Bombín, H. Single-shot fault-tolerant quantum error correction. *Phys. Rev. X* **5**, 031043 (2015).

40. Campbell E. A theory of single-shot error correction for adversarial noise. *Quantum Sci. Technol.* http://iopscience.iop.org/10.1088/2058-9565/aafc8f (2019)

41. Galindo, A. & Martín-Delgado, M. A. Information and computation: classical and quantum aspects. *Rev. Mod. Phys.* **74**, 347 (2002).

42. Gottesman, D., Quantum fault tolerance in small experiments, arXiv preprint arXiv:1610.03507 (2016)

43. Li, M., Gutiérrez, M., David, S. E., Hernandez, A. & Brown, K. R. Fault tolerance with bare ancillary qubits for a [[7, 1, 3]] code. *Phys. Rev. A* **96**, 032341 (2017).

44. Linke, N. M. et al. Fault-tolerant quantum error detection. *Sci. Adv.* **3**, e1701074 (2017).

45. Trout, C. J. et al. Simulating the performance of a distance-3 surface code in a linear ion trap. *New J. Phys.* **20**, 043038 (2018).

46. Harper, R. & Flammia, S. T. Fault-tolerant logical gates in the IBM quantum experience. *Phys. Rev. Lett.* **122**, 080504 (2019).

47. Vuillot, C., Is error detection helpful on ibm 5q chips? arXiv preprint arXiv:1705.08957 (2017)

48. Rosenblum, S. et al. Fault-tolerant detection of a quantum error. *Science* **361**, 266–270 (2018).

49. Willsch, D., Willsch, M., Jin, F., De Raedt, H. & Michielsen, K. Testing quantum fault tolerance on small systems. *Phys. Rev. A* **98**, 052348 (2018).

50. Forsythe, G. E. & Golub, G. H. On the stationary values of a second-degree polynomial on the unit sphere. *J. Soc. Ind. Appl. Math.* **13**, 1050–1068 (1965).

51. Ritz, W. Uber eine neue methode zur losung gewisser variationsprobleme der mathematischen physik. *J. fur Mathematik* **135**, 1–61 (1909).

52. Bravyi, S., Gambetta, J. M., Mezzacapo A. & Temme K. Tapering off qubits to simulate fermionic hamiltonians. arXiv preprint arXiv:1701.08213 (2017)

53. Tuckett, D. K., Bartlett, S. D. & Flammia, S. T. Ultrahigh error threshold for surface codes with biased noise. *Phys. Rev. Lett.* **120**, 050505 (2018).

54. Jiang, Z., McClean, J., Babbush, R. & Neven, H., Majorana loop stabilizer codes for error correction of fermionic quantum simulations, arXiv preprint arXiv:1812.08190 (2018)

55. Setia, K. & Whitfield, J. D. Bravyi-kitaev superfast simulation of electronic structure on a quantum computer. *J. Chem. Phys.* **148**, 164104 (2018).

Author contributions
All authors contributed to the generation of ideas, analysis of results, and writing of the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to J.R.M.

Peer review information *Nature Communications* thanks Joydip Ghosh and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.