Purification-based quantum error mitigation of pair-correlated electron simulations

A list of authors and their affiliations appears at the end of the paper

An important measure of the development of quantum computing platforms has been the simulation of increasingly complex physical systems. Before fault-tolerant quantum computing, robust error-mitigation strategies were necessary to continue this growth. Here, we validate recently introduced error-mitigation strategies that exploit the expectation that the ideal output of a quantum algorithm would be a pure state. We consider the task of simulating electron systems in the seniority-zero subspace where all electrons are paired with their opposite spin. This affords a computational stepping stone to a fully correlated model. We compare the performance of error mitigations on the basis of doubling quantum resources in time or in space on up to 20 qubits of a superconducting qubit quantum processor. We observe a reduction of error by one to two orders of magnitude below less sophisticated techniques such as postselection. We study how the gain from error mitigation scales with the system size and observe a polynomial suppression of error with increased resources. Extrapolation of our results indicates that substantial hardware improvements will be required for classically intractable variational chemistry simulations.

The prospect of accurately simulating ground states of quantum systems on quantum hardware has motivated substantial theory and hardware developments over the last decade. With fault-tolerant quantum computing in its infancy[1] and many years from promised applications[2–6], attention has focused on algorithms requiring only short-depth quantum circuits, such as the variational quantum eigensolver (VQE)[7,8]. Theoretical developments in ansatz design[8–13] and measurement optimization[14–18] have enabled small to midscale VQE experiments[11,19–26]. A key target of variational quantum algorithms has been the electronic structure problem in chemistry[8,11,19–23,27]. Such simulations are challenging to implement on quantum hardware due to a long-range two-body fermionic Hamiltonian and stringent accuracy requirements. This makes it unclear whether a beyond-classical simulation of chemistry can be achieved without fault tolerance. Determining the requirements for such a simulation is a critical open problem.

The electronic structure problem can be expressed in models of varying complexity and realism. Quantum simulations of chemistry in the Hartree–Fock (mean-field) approximation were implemented for system sizes up to 12 qubits in ref. 23, and to our knowledge, this

study retains the record for the largest VQE calculation of a chemical ground state on quantum hardware. As a next step, one can consider working in the seniority-zero subspace of the entire Hilbert space, which assumes all electrons come in spin-up or spin-down pairs[12,28–33]. This has the advantage of projecting a local fermionic problem onto a local qubit problem[12]. The S_0 ground state is not a priori classically efficiently simulatable[12] (although good approximate methods are known to exist for many problems[33–35]). This makes it a good stepping stone beyond Hartree–Fock towards the full electronic structure problem.

Recent quantum experiments have relied on error-mitigation techniques[36], which are not scalable like error correction[1,37] but promise to substantially shrink experimental errors. Popular methods are based on postselection[38,39], rescaling[24,40,41], purification[23,42–44] and probabilistic cancellation[40,45]. Various schemes and combinations of error-mitigation techniques have been implemented in practice[20,22–26,46]. However, many of these methods do not promise to remove bias to the level of accuracy needed for useful simulation of chemistry, or remain untested beyond few-qubit experiments. Shifting from non-interacting fermions to correlated electronic structure, one loses two error-mitigation advantages

✉e-mail: teobrien@google.com; christian.gogolin@covestro.com; babbush@google.com; nickrubin@google.com

that were crucial to the success of ref. 23: efficient density-matrix purification by means of McWeeny iteration[47], and low-cost gradient estimation.

In this work, we compare the performance of three different error-mitigation techniques—postselection, echo verification (EV) and virtual distillation (VD)—on the problem of preparing ground states in the seniority-zero space, using up to 20 qubits of a superconducting quantum processor. Using either EV or postselected VD, we are able to accurately reproduce the ground-state energy and order parameter for an $N = 10$–qubit simulation of the Richardson–Gaudin (RG) model, improving over unmitigated estimates by one to two orders of magnitude. This demonstrates an improvement over classical pair coupled-cluster doubles (pCCD) and the non-interacting Bardeen–Cooper–Schrieffer (BCS) theory, neither of which are qualitatively correct over the entire range of coupling values considered. We study the scaling of the mean absolute error in energy and the order parameter with the system size and observe that EV and VD suppress errors by a factor N^α for values of α between 0.8 and 2.5. EV was further able to greatly mitigate error of six- and ten-qubit simulations of the ring opening of cyclobutene (CB). Although the stringent error requirements (<0.05 Hartree) to differentiate between mean-field and the exact solution could only be achieved for the six-qubit case, this still represents the largest VQE simulation of electronic structure for chemistry that we are aware of so far. From these data, we are able to estimate the minimum requirements for a beyond-classical quantum simulation of similar form: a 25× decrease in hardware error rates (from those observed in this work), a limit of $O(N)$ depth for future variational ansatzes and the need to pre-optimize ansatzes classically without intermediate calls to a device. Even if this list of requirements is achieved, meeting the high level of accuracy required for the electronic structure problem will pose a serious challenge, as chemical accuracy is around 60× less than our mean accuracy for the ten-qubit CB problem.

The RG model

We first prepare approximate ground states of the RG model on ten sites at half-filling across a range of coupling strengths g. This model is chosen as a well-known benchmark with well-understood exact and approximate solutions. We work in the seniority-zero space (where all electrons are paired with their opposite spin) and use a unitary pair coupled cluster doubles (UpCCD) variational ansatz[12] with parameters optimized in noiseless simulation. In Fig. 1a, we estimate the prepared states' energy using various error-mitigation techniques: postselection, EV and postselected virtual distillation (PS-VD). We compare these results to exact diagonalization in the S_0 subspace (also known as double occupied configuration interaction) and classical pCCD and BCS solutions. We see that using EV or PS-VD, we are able to reproduce the entire energy curve to high accuracy, which neither pCCD nor the non-interacting BCS theory can achieve. The experimental error in the result is the sum of the UpCCD model error and the experimental error. To disambiguate the effects of UpCCD model error, in Fig. 1b we plot the error between our experimental data and the UpCCD ground-state energy. The observed gain from error mitigating some estimate y can be quantified by the suppression factor

$$\eta_y = |y_{\text{raw}} - y_{\text{true}}| / |y_{\text{mit}} - y_{\text{true}}|. \quad (1)$$

We summarize this across our experimental data in Table 1. For the energy ($y = E$), we find PS-VQE consistently achieves $\eta_E \approx 2$, whereas EV achieves $\bar{\eta}_E = 85$ and $\max(\eta_E) = 460$ and VD achieves $\bar{\eta}_E = 60$ and $\max(\eta_E) = 140$. The residual error following EV or PS-VD fluctuates on a scale larger than statistical error bars, which we attribute to device drift.

The RG Hamiltonian has a well-known phase transition in the attractive regime ($g \le 0$) in the thermodynamic limit, which appears in the BCS state at finite N but is not present in the true ground
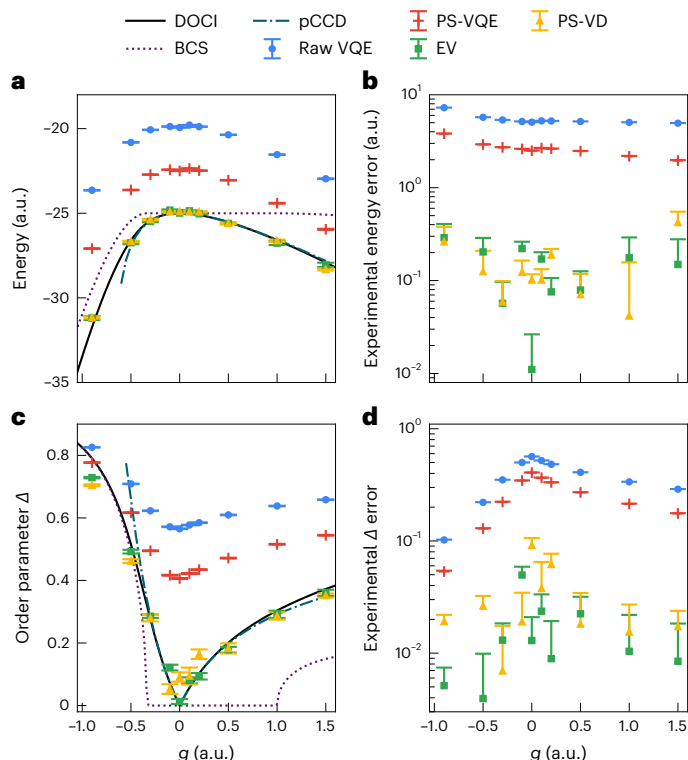

Fig. 1 | Digital quantum simulation of ground states of the RG model for ten spatial orbitals on a superconducting quantum device. a, Energy as a function of the coupling parameter g, estimated using various error-mitigation techniques (markers) and compared to classical models (lines). The classical pCCD results do not converge below a critical value, resulting in their cut-off. **b**, Log plot of experimental energy error (ignoring the model error from the UpCCD approximation). **c**, Many-body order parameter for the RG Hamiltonian (see text), again compared to classical models. **d**, Experimental error in estimating the superconducting order parameter versus the target state in the UpCCD approximation (again ignoring model error). Error bars show 1 standard deviation uncertainty from sampling noise, estimated by propagating variance (raw VQE, PS-VQE) or bootstrapping (EV, PS-VD); see Supplementary Information Section III for details. a.u., arbitrary units.

state due to finite size effects[48–50]. This gives us a second physically relevant target to simulate and a qualitative feature (the cusp at $g = 0$) to resolve. The traditional order parameter for the BCS state, $\Delta_{\text{BCS}} = \frac{1}{2N}\sum_j(\langle a_{j\uparrow} a_{j\downarrow}\rangle + \langle a_{j\downarrow}^\dagger a_{j\uparrow}^\dagger\rangle)$, is zero on the RG Hamiltonian ground state due to number conservation. However, $\Delta = \frac{1}{N}\sum_{j,\sigma}\sqrt{\langle n_{j\sigma}^2\rangle - \langle n_{j\sigma}\rangle^2}$ satisfies $\Delta = \Delta_{\text{BCS}}$ for the BCS ground state of the Hamiltonian, giving a many-body-order parameter[49]. In Fig. 1c, we plot experimental estimates of Δ across the range of g values considered. In the absence of error mitigation, although the order parameter dips around $g = 0$, the true cusp is not reproduced. Both EV and PS-VD clearly improve over the BCS approximation for $g > 0.5$, with EV particularly able to reproduce the cusp at $g = 0$. We measure the performance of the error-mitigation techniques by plotting the error in Δ against the noise-free UpCCD result in Fig. 1d. The suppression (equation (1), $y = \Delta$) from EV and PS-VD during order-parameter estimation is slightly less than for energy estimation: for EV, we find $\bar{\eta}_\Delta = 32$ and $\max(\eta_\Delta) = 56$; and for PS-VD, we find $\bar{\eta}_\Delta = 18$ and $\max(\eta_\Delta) = 51$.

We now study the scaling of the studied error-mitigation techniques for the RG model over system sizes $N = 4, 6, 8, 10$. In Fig. 2a, we plot the absolute error in the energy estimates, averaged across all points in Fig. 1, and repeat this for different system sizes N. (Plots equivalent to Fig. 1 for different system sizes can be found in Supplementary Information Section VI.D.) We observe that the energy

Table 1 | Observed mean and maximum suppression factors $\bar{\eta}$ and $\max(\eta)$ (equation (1)) across the RG problems studied in Fig. 1 and the CB problems studied in Fig. 3, for all error-mitigation methods studied

System	Metric	PS-VQE	EV	PS-VD
RG (ten qubits, energy)	$\bar{\eta}_E$	2.1	86	54
	$\max(\eta_E)$	2.5	460	120
RG (ten qubits, order parameter)	$\bar{\eta}_\Delta$	1.6	32	18
	$\max(\eta_\Delta)$	1.9	56	51
CB (six qubits, energy)	$\bar{\eta}_E$	5.0	54	–
	$\max(\eta_E)$	5.2	140	–
CB (ten qubits, energy)	$\bar{\eta}_E$	2.9	38	–
	$\max(\eta_E)$	3.0	190	–

error scales sublinearly after applying EV or VD, demonstrating an asymptotic difference from raw VQE or PS-VQE of a factor of $\sim N^2$. We similarly observe an $\sim N - N^2$ gap between EV/VD and VQE/PS-VQE in the order-parameter error (Fig. 2b); the difference in the suppression factor is not statistically significant. Next, we extract various fidelity metrics from the experimental data in Fig. 2c, which approximate the (inverse) increased variance of our error-mitigated estimators[36]. The data for PS-VQE (red plus), EV (green square) and PS-VD (yellow triangle) directly correspond to the circuit overhead required by different error-mitigation techniques. These fidelity metrics allow us to estimate the required number of experiment repetitions to reduce statistical noise in the energy estimation to <0.1 a.u.(arbitrary units) in Fig. 2d, assuming ten-qubit fidelity is maintained. We observe a substantial gap between the experimental and theoretical sampling cost of EV (green squares). We attribute this to known overheads of EV that are not in the theoretical model (Supplementary Information Section IV.A).

CB ring opening

We further validated scalable error-mitigation protocols by simulating the conrotatory ring opening pathway for CB in an active space of six orbital and six electrons and ten orbitals and ten electrons corresponding to a six- and ten-qubit simulation of the electronic structure Hamiltonian. The mechanism of this ring opening is described by the Woodward–Hoffmann rules for pericyclic ring openings corresponding to the in-phase combination of the two carbon $2p$ orbitals when brought together to form the four-member carbon ring. Similar to the RG model, this was chosen as a well-known benchmark system with well-understood chemistry.

The geometries along the reaction path are determined from a nudged elastic-band calculation using density functional theory (B3LYP) to evaluate forces. The final structures use a minimal basis set (STO-3G) to generate the active-space Hamiltonians to project into the seniority-zero sector. The Woodward–Hoffmann rules are a type of molecular orbital theory, and thus we expect this reaction to be qualitatively described in mean-field theory. This is verified numerically for our seniority-zero model where the largest configuration interaction coefficient has an average value of 0.974(9) for six orbitals and 0.973(9) for ten orbitals, indicating a single-reference system. As such, our unitary pCCDs ansatz targets the dynamic correlation corrections to the mean field.

The average PS-VQE absolute error is 0.058 ± 0.006 and 0.395 ± 0.023 Hartree, and the mean suppression factor is $\bar{\eta}_E = 5.0$ and 2.9 for the six- and ten-orbital systems, respectively. The average echo-verified absolute error is 0.011 ± 0.005 and 0.064 ± 0.035 Hartree and the mean suppression factor (equation (1)) is $\bar{\eta}_E = 54$ and 38 for the six- and ten-orbital system respectively. Although there is notable improvement in energy across the reaction pathway for the
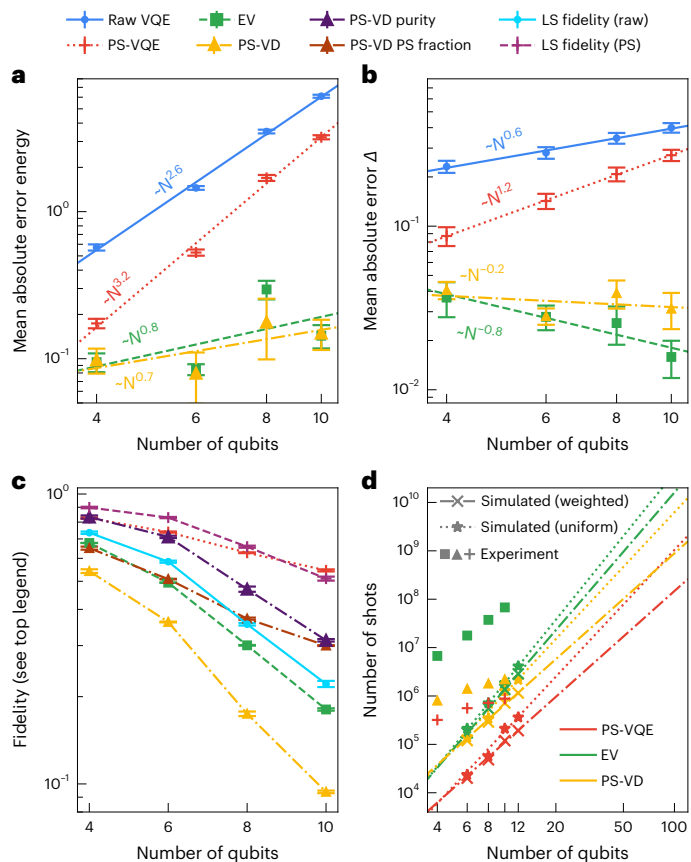

Fig. 2 | Scaling the simulation of the RG model to larger qubit counts.
a, Experimental energy error (versus the UpCCD ground state) averaged over all points studied of the RG model. Error bars show sample standard deviation and lines a power-law fit. **b**, Experimental error in order parameter (versus the UpCCD ground state) averaged over all points studied of the RG model. Error bars and lines same as **a**. **c**, Different fidelity metrics for PS-VQE, EV, PS-VD and Loschmidt (LS) echo (see legend) averaged over all points studied of the RG model. **d**, Number of shots required for convergence at $g = -0.9$. Crosses and pluses give simulated estimations with two types of term grouping (see text for details) using observed fidelities of a ten-qubit experiment from **c**. Other symbols give experimental shots used.

ten-orbital system, the magnitude of the errors is larger than the 0.037 Hartree energy difference between CB and 1,3-butadiene. Furthermore, a visual inspection of Fig. 3 indicates high parallelity errors in the ten-orbital system. Given that the error bars on EV are smaller than the parallelity error (point scatter), we attribute the main source of error to device drift.

Discussion and outlook

Many of the features of our experimental data can be understood physically. The similar performance of EV and PS-VD can be understood following the equivalence of the two techniques shown in refs. 46,51, which treat EV as purification of two states separated in time. However, this equivalence is broken by the further circuitry of the two techniques: EV requires Greenberger-Horne-Zeilinger (GHZ) state preparation, whereas VD requires bell-basis measurements and more routing. This doubles the susceptibility of the VD fidelity to measurement noise, to which we attribute the gap between this and the EV fidelity in Fig. 2c. Despite the fidelity gap, we observe PS-VD has a substantially smaller sampling cost than EV, which comes from known measurement bottlenecks[52] and overhead to reduce on-chip magnetic field fluctuations (Supplementary Information Section II.D). We can attribute the gap in performance between VD with and without
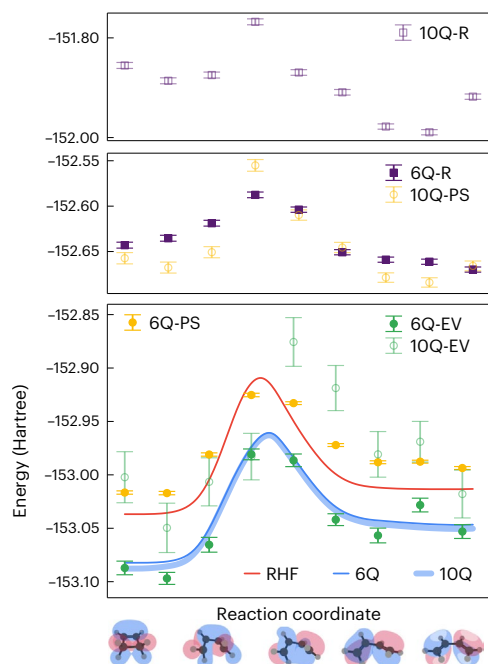
Fig. 3 | Conrotatory CB ring opening pathway simulated in the seniority-zero subspace. Comparison of raw VQE (purple, 'R'), PS-VQE (yellow, 'PS') and EV (green, 'EV') on an optimized unitary pair coupled-cluster ansatz. Darker (lighter) points correspond to six-qubit (6Q) (ten-qubit (10Q)) simulations. Error bars denote 1 s.d. calculated by either propagating error (VQE) or bootstrapping (EV). From left to right, the reaction path corresponds to the ring opening reaction. For the ten-orbital case, the unitary pair coupled-cluster ansatz (evaluated in simulation) has less than 1.8×10^{-4} energy difference from exact diagonalization in the seniority-zero space. The blue curves correspond to the exact diagonalization of the seniority-zero active space Hamiltonian spanning ten orbitals (broad, lighter-blue line) and six orbitals (narrow, darker-blue line). The red curve is the RHF mean-field energy. Some of the data are plotted on a discontinuous and different scale to preserve the visual scale of the reaction energy along the reaction coordinate.

postselection to this further measurement circuitry (Supplementary Information Section VI.A): we suggest that the bell-basis measurements used are not robust to noise that does not conserve particle number and that the primary effect of postselection is to mitigate this noise source. The performance improvement of EV/VD over postselection is expected from numerical predictions in refs. 42,44 and analytic results of ref. 53 (although we show in Supplementary Information Section VI.B that this cannot be guaranteed). Finally, we can explain the qualitative feature in Fig. 1 around $g = 0$, where the energy error drops but the order parameter error increases. This comes from two coincidental effects: Δ is more sensitive to error around $\Delta = 0$, but at this point the energy becomes less dependent on off-diagonal terms, reducing its sensitivity to phase noise.

We now consider how the required number of experiment repetitions (shots) scales to classically intractable simulations, using the data in Fig. 2d. Taking a 50-qubit experiment as a proxy lower bound for a beyond-classical quantum computation, and assuming the ability to freely weight our shot distribution, we estimate that at least 10^8 or 10^9 shots would be required when using VD or EV, respectively. This is executable on current hardware in a wall-clock time of >1 hour or >10 hours, respectively. Including the difference between experiment and theory at ten qubits raises the cost of EV to 5×10^{10} shots, which would require several days to achieve. These numbers do not include the large multiplicative cost of variational optimization. Furthermore, the requirements for accurate electronic structure simulations may be lower than the 0.1 a.u. requirement considered here. Methods to

pre-optimize variational ansatzes classically and applications of VQE to problems simpler than electronic structure may thus be necessary for beyond-classical VQE experiments.

Next, we consider what device fidelity would be required to scale to a beyond-classical experiment. To maintain circuit fidelity F over a depth $3N/2$, fully parallel circuit as N scales from 10 to 50 requires all error rates to drop by a factor of 25 ($\sim N^2$). As any reduction in F incurs a poly(F^{-1}) sampling cost[42,44], and as F scales exponentially in the error rate, and as $F \approx 10\%$ for PS-VD (Fig. 2d), we see little room for negotiation on this 25× lower bound. To achieve a 25× decrease in error rate would require a 50-qubit device with fidelities on all two-qubit gates $\leq 3 \times 10^{-4}$. This analysis likely precludes a large number of ansatzes with depth much greater than linear. For instance, successfully implementing a 50-qubit VQE with ansatz depth $3N^2/2$ with EV or VD would require error rates to drop ~1,000×. However, the sublinear scaling in the energy error following the application of EV or VD (Fig. 2a) indicates that a 25× decrease in error rate would be sufficient to deal with the bias in the resulting experiment; the dominant cost of error-mitigated quantum experiments comes from the sampling and fidelity requirements. Confirming this point in studies on other systems is a key target for future work.

We have observed that EV and VD suppress errors by one to two orders of magnitude on a range of quantum simulation problems using up to 20 superconducting qubits. We further observe that the error-suppression performance of these mitigation techniques increases with the number of qubits. This allows us to conclude that variance, not bias, will be the limiting factor in purification-based error-mitigation techniques going forward. As our error-mitigation protocols target expectation values themselves and are blind to the outer loops combining these to estimate energies or order parameters, we expect these results to hold for a wider range of problems than the VQE examples studied here. Any algorithm using circuits with similar structure, such as time evolution on spin chains, or brick-wall parameterized quantum circuits for quantum machine learning, should see similar gains in performance. Testing whether this extends to more generic circuits (for example, higher-dimensional or sparse circuits, or primitive circuits for fault tolerance[54]) is a clear target for future work.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-023-02240-y.

References

1. Acharya, R. et al. Suppressing quantum errors by scaling a surface code logical qubit. *Nature* **614**, 676–681 (2023).
2. von Burg, V. et al. Quantum computing enhanced computational catalysis. *Phys. Rev. Res.* **3**, 033055 (2021).
3. Lee, J. et al. Even more efficient quantum computations of chemistry through tensor hypercontraction. *PRX Quantum* **2**, 030305 (2021).
4. Gidney, C. & Ekerå, M. How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *Quantum* **5**, 433 (2021).
5. Campbell, E. T. Early fault-tolerant simulations of the Hubbard model. *Quant. Sci. Technol.* **7**, 015007 (2021).
6. Berry, D. W. et al. Quantifying quantum advantage in topological data analysis. Preprint at https://arxiv.org/abs/2209.13581 (2022).
7. Peruzzo, A. et al. A variational eigenvalue solver on a quantum processor. *Nat. Commun.* **5**, 4213 (2014).
8. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 23023 (2016).

9. Wecker, D., Hastings, M. B. & Troyer, M. Progress towards practical quantum variational algorithms. *Phys. Rev. A* **92**, 042303 (2015).
10. Grimsley, H. R., Economou, S. E., Barnes, E. & Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* **10**, 3007 (2019).
11. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).
12. Elfving, V. E., Millaruelo, M., Gámez, J. A. & Gogolin, C. Simulating quantum chemistry in the seniority-zero space on qubit-based quantum computers. *Phys. Rev. A* **103**, 032605 (2021).
13. Lee, J., Huggins, W. J., Head-Gordon, M. & Whaley, K. B. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.* **15**, 311 (2018).
14. Huggins, W. J. et al. Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers. *npj Quantum Inf.* **7**, 23 (2021).
15. Cotler, J. & Wilczek, F. Quantum overlapping tomography. *Phys. Rev. Lett.* **124**, 100401 (2020).
16. Bonet-Monroig, X., Babbush, R. & O'Brien, T. E. Nearly optimal measurement scheduling for partial tomography of quantum states. *Phys. Rev. X* **10**, 031064 (2020).
17. Verteletskyi, V., Yen, T.-C. & Izmaylov, A. F. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *J. Chem. Phys.* **152**, 124114 (2020).
18. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
19. O'Malley, P. J. J. et al. Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 31007 (2016).
20. Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495 (2019).
21. Hempel, C. et al. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* **8**, 031022 (2018).
22. Sagastizabal, R. et al. Error mitigation by symmetry verification on a variational quantum eigensolver. *Phys. Rev. A* **100**, 010302 (2019).
23. Arute, F. et al. Hartree-Fock on a superconducting qubit quantum computer. *Science* **369**, 1084–1089 (2020).
24. Stanisic, S. et al. Observing ground-state properties of the Fermi-Hubbard model using a scalable algorithm on a quantum computer. *Nat. Comm.* **13**, 5743 (2022).
25. Kim, Y. et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. *Nat. Phys.* **19**, 752–759 (2023).
26. van den Berg, E., Minev, Z. K., Kandala, A. & Temme, K. Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors. *Nat. Phys.* **19**, 1116–1121 (2023).
27. Motta, M. et al. Quantum chemistry simulation of ground- and excited-state properties of the sulfonium cation on a superconducting quantum processor. *Chem. Sci.* **14**, 2915–2927 (2023).
28. Limacher, P. A. et al. A new mean-field method suitable for strongly correlated electrons: computationally facile antisymmetric products of nonorthogonal geminals. *J. Chem. Theory Comput.* **9**, 1394 (2013).
29. Boguslawski, K. et al. Efficient description of strongly correlated electrons with mean-field cost. *Phys. Rev. B* **89**, 201106 (2014).
30. Johnson, P. A. et al. Richardson-Gaudin mean-field for strong correlation in quantum chemistry. *J. Chem. Phys.* **153**, 104110 (2020).
31. Gunst, K., Van Neck, D., Limacher, P. A. & De Baerdemacker, S. The seniority quantum number in tensor network states. *SciPost Chem.* **1**, 001 (2021).
32. Kossoski, F., Damour, Y. & Loos, P.-F. Hierarchy configuration interaction: combining seniority number and excitation degree. *J. Phys. Chem. Lett.* **13**, 4342–4349 (2022).
33. Fecteau, C.-E. et al. Near-exact treatment of seniority-zero ground and excited states with a Richardson-Gaudin mean field. *J. Chem. Phys.* **156**, 194103 (2022).
34. Dukelsky, J., Roman, J. M. & Sierra, G. Comment on 'polynomial-time simulation of pairing models on a quantum computer'. *Phys. Rev. Lett.* **90**, 249803 (2003).
35. Dukelsky, J. Integrable Richardson-Gaudin models in mesoscopic physics. *J. Phys. Conf. Ser.* **338**, 012023 (2012).
36. Cai, Z. et al. Quantum error mitigation. *Rev. Mod. Phys.* (in the press).
37. Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: towards practical large-scale quantum computation. *Phys. Rev. A* **86**, 032324 (2012).
38. McArdle, S., Yuan, X. & Benjamin, S. Error-mitigated digital quantum simulation. *Phys. Rev. Lett.* **122**, 180501 (2019).
39. Bonet-Monroig, X., Sagastizabal, R., Singh, M. & O'Brien, T. Low-cost error mitigation by symmetry verification. *Phys. Rev. A* **98**, 062339 (2018).
40. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
41. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
42. Huggins, W. J. et al. Virtual distillation for quantum error mitigation. *Phys. Rev. X* **11**, 041036 (2021).
43. Koczor, B. Exponential error suppression for near-term quantum devices. *Phys. Rev. X* **11**, 031057 (2021).
44. O'Brien, T. E. et al. Error mitigation via verified phase estimation. *PRX Quantum* **2**, 020317 (2021).
45. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).
46. Huo, M. & Li, Y. Dual-state purification for practical error mitigation. *Phys. Rev. A* **105**, 022427 (2022).
47. McWeeny, R. Some recent advances in density matrix theory. *Rev. Mod. Phys.* **35**, 668 (1963).
48. von Delft, J., Zaikin, A., Golubev, D. & Tichy, W. Parity-affected superconductivity in ultrasmall metallic grains. *Phys. Rev. Lett.* **77**, 3189 (1996).
49. Braun, F. & von Delft, J. Superconductivity in ultrasmall metallic grains. *Phys. Rev. B* **59**, 9527 (1999).
50. Dukelsky, J. & Sierra, G. The crossover from the bulk to the few-electron limit in ultrasmall metallic grains. *Phys. Rev. B* **61**, 12302 (1999).
51. Cai, Z. Resource-efficient purification-based quantum error mitigation. Preprint at https://arxiv.org/abs/2107.07279 (2021).
52. Polla, S., Anselmetti, G.-L. R. & O'Brien, T. E. Optimizing the information extracted by a single qubit measurement. *Phys. Rev. A*, **108**, 012403 (2023).
53. Koczor, B. The dominant eigenvector of a noisy quantum state. *New. J. Phys.* **23**, 123047 (2021).
54. Pivetaeu, C., Sutter, D., Bravyi, S., Gambetta, J. M. & Temme, K. Error mitigation for universal gates on encoded qubits. *Phys. Rev. Lett.* **127**, 200505 (2021).

T. E. O'Brien [1,2] ✉, G. Anselmetti[3], F. Gkritsis [3], V. E. Elfving [4], S. Polla[1,2], W. J. Huggins [1], O. Oumarou [3], K. Kechedzhi[1], D. Abanin[1], R. Acharya[1], I. Aleiner [1], R. Allen[1], T. I. Andersen[1], K. Anderson[1], M. Ansmann [1], F. Arute[1], K. Arya[1], A. Asfaw[1], J. Atalaya[1], J. C. Bardin [1,5], A. Bengtsson [1], G. Bortoli [1], A. Bourassa [1], J. Bovaird[1], L. Brill[1], M. Broughton[1], B. Buckley[1], D. A. Buell[1], T. Burger[1], B. Burkett [1], N. Bushnell[1], J. Campero[1], Z. Chen[1], B. Chiaro[1], D. Chik[1], J. Cogan[1], R. Collins[1], P. Conner[1], W. Courtney[1], A. L. Crook[1], B. Curtin [1], D. M. Debroy[1], S. Demura [1], I. Drozdov [1,6], A. Dunsworth[1], C. Erickson[1], L. Faoro[1], E. Farhi[1], R. Fatemi[1], V. S. Ferreira[1], L. Flores Burgos[1], E. Forati[1], A. G. Fowler[1], B. Foxen [1], W. Giang[1], C. Gidney[1], D. Gilboa[1], M. Giustina[1], R. Gosula [1], A. Grajales Dau[1], J. A. Gross [1], S. Habegger [1], M. C. Hamilton[1,7], M. Hansen[1], M. P. Harrigan [1], S. D. Harrington [1], P. Heu[1], M. R. Hoffmann[1], S. Hong [1], T. Huang[1], A. Huff[1], L. B. Ioffe[1], S. V. Isakov[1], J. Iveland[1], E. Jeffrey[1], Z. Jiang[1], C. Jones[1], P. Juhas [1], D. Kafri[1], T. Khattar[1], M. Khezri[1], M. Kieferová [1,8], S. Kim[1], P. V. Klimov[1], A. R. Klots[1], A. N. Korotkov[1,9], F. Kostritsa[1], J. M. Kreikebaum[1], D. Landhuis [1], P. Laptev[1], K.-M. Lau[1], L. Laws[1], J. Lee[1,10], K. Lee[1], B. J. Lester [1], A. T. Lill[1], W. Liu[1], W. P. Livingston [1], A. Locharla[1], F. D. Malone[1], S. Mandrà[1], O. Martin[1], S. Martin[1], J. R. McClean [1], T. McCourt[1], M. McEwen [1,11], X. Mi [1], A. Mieszala[1], K. C. Miao[1], M. Mohseni [1], S. Montazeri [1], A. Morvan [1], R. Movassagh[1], W. Mruczkiewicz [1], O. Naaman [1], M. Neeley [1], C. Neill [1], A. Nersisyan[1], M. Newman[1], J. H. Ng[1], A. Nguyen[1], M. Nguyen[1], M. Y. Niu[1], S. Omonije[1], A. Opremcak[1], A. Petukhov[1], R. Potter[1], L. P. Pryadko [1,12], C. Quintana[1], C. Rocque[1], P. Roushan [1], N. Saei[1], D. Sank [1], K. Sankaragomathi [1], K. J. Satzinger [1], H. F. Schurkus [1], C. Schuster[1], M. J. Shearn[1], A. Shorter[1], N. Shutty [1], V. Shvarts[1], J. Skruzny[1], W. C. Smith[1], R. D. Somma[1], G. Sterling[1], D. Strain[1], M. Szalay[1], D. Thor[1], A. Torres[1], G. Vidal[1], B. Villalonga[1], C. Vollgraff Heidweiller [1], T. White[1], B. W. K. Woo [1], C. Xing[1], Z. J. Yao[1], P. Yeh [1], J. Yoo[1], G. Young[1], A. Zalcman [1], Y. Zhang[1], N. Zhu [1], N. Zobrist [1], D. Bacon[1], S. Boixo [1], Y. Chen[1], J. Hilton[1], J. Kelly [1], E. Lucero[1], A. Megrant [1], H. Neven [1], V. Smelyanskiy[1], C. Gogolin [3] ✉, R. Babbush [1] ✉ & N. C. Rubin [1] ✉

[1]Google Research, Mountain View, CA, USA. [2]Instituut-Lorentz, Universiteit Leiden, Leiden, the Netherlands. [3]Covestro Deutschland AG, Leverkusen, Germany. [4]PASQAL SAS, Massy, France. [5]Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, MA, USA. [6]Department of Physics, University of Connecticut, Storrs, CT, USA. [7]Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. [8]QSI, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, New South Wales, Australia. [9]Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA. [10]Department of Chemistry, Columbia University, New York, NY, USA. [11]Department of Physics, University of California, Santa Barbara, CA, USA. [12]Department of Physics and Astronomy, University of California, Riverside, CA, USA. ✉e-mail: teobrien@google.com; christian.gogolin@covestro.com; babbush@google.com; nickrubin@google.com

Methods

Simulating the seniority-zero subspace

The seniority of a Slater determinant is the number of unpaired electrons; thus, the seniority-zero (S_0) sector of Hilbert space for an N-electron system in M orbitals is the space of $\binom{M}{N/2}$ determinants leaving no electrons unpaired given a particular pairing of the spin orbitals. Seniority is not a global symmetry of the electronic structure Hamiltonian, and it is basis dependent; it has been used as a way to classify determinant subspaces to generate better approximations for solving the Schrödinger equation[28,29,31,32] and as a starting point for modelling strong correlations from electron-pair states[55].

Supported by the S_0 subspace, there exists a set of operators satisfying the su(2) algebra constructed from pairs of fermion ladder operators and the spatial orbital number operator[56]

$$P_p^\dagger = a_{p\alpha}^\dagger a_{p\beta}^\dagger, \quad N = \sum_{p,\sigma} a_{p\sigma}^\dagger a_{p\sigma},$$

$$[P_p, P_q^\dagger] = (1 - N_p)\delta_{p,q}, \quad [N_p, P_q] = -2P_q\delta_{p,q}, \tag{2}$$

where p, q and α, β are orbital and spin indices, respectively, and δ is the Kronecker delta function. These operators form a basis for Hamiltonians projected into the S_0 subspace. The equivalence to an su(2) algebra means seniority-zero models resemble Heisenberg spin $-1/2$ models, which are easily expressed as Pauli operators.

In this work, we focus on two Hamiltonians to validate purification-based error-mitigation strategies. The first is the RG or pairing model

$$\hat{H} = \sum_{p=1}^N \epsilon_p N_p + g \sum_{p\neq q=1}^N P_p^\dagger P_q, \tag{3}$$

with ϵ_p the single-particle energy for site p, and g the coupling strength. At small N this corresponds to a single strength all-to-all coupling of fermion pairs. In this Article, we work at half-filling by adding a chemical potential μ to the single-particle energies:

$$\epsilon_p = p - \mu, \quad \mu = \frac{1}{2}(N + 1) \tag{4}$$

This is a model for a small superconducting grain when $g < 0$ (refs. 48–50), but with a g-dependent Debye frequency[49]. The second model is the electronic structure Hamiltonian (H_{elec}) projected into the S_0 subspace:

$$H_{S_0} = P_{S_0} H_{\text{elec}} P_{S_0} = \sum_p (h_{p,p}) N_p$$
$$+ \frac{1}{4} \sum_{p\neq q} (2V_{pqpq} - V_{pqqp}) N_p N_q + \sum_{pq} (V_{ppqq}) P_p^\dagger P_q, \tag{5}$$

where h and V are the one- and two-body term coefficients. The all-to-all connected Heisenberg spin Hamiltonian is not known to be classically solvable in general, but good approximate methods exist. This is especially true for the RG model, which is often integrable[35], well-approximated by density-matrix renormalization group[34] and pair coupled-cluster techniques in the repulsive regime and solvable by quantum Monte Carlo in the attractive regime (where it has no sign problem). Pair coupled-cluster theory is also known to work well for the electronic structure problem in the S_0 subspace[29,57–59] while full configuration interaction quantum Monte Carlo shows a reduced sign problem[60]. As such, although we have strong evidence that the quantum circuits used in this text are not classically simulatable (Supplementary Information Section II.A), we do not believe directly scaling S_0 simulations represents the easiest path to a quantum advantage in chemistry; this is instead a stepping stone between a mean-field solution and the full electronic structure problem.

The unitary pair coupled-cluster ansatz and energy estimation

In this work, we use a Trotterized UpCCD ansatz[12] compiled into a set of qubits in a ladder geometry with nearest-neighbour coupling. The ladder ansatz (instead of a generic ring) allows us to efficiently measure terms in the Hamiltonian corresponding qubits that are not physically adjacent after encoding with a minimal number of SWAP operations. When mapped from fermions to qubits, the UpCCD ansatz has the form

$$U(\boldsymbol{\theta}) = \prod_\ell^{N_\ell} U_e(\theta^{e,\ell}) U_o(\theta^{o,\ell}) \tag{6}$$

$$U_o(\theta^{o,\ell}) = \prod_{n=0}^{N/2-1} \text{GS}_{2n+1,(2n+2)\%N} \left(\theta_{2n+1,(2n+2)\%N}^{o,l}\right) \tag{7}$$

$$U_e(\theta^{e,\ell}) = \prod_{n=0}^{N/2-1} \text{GS}_{2n,2n+1} \left(\theta_{2n,2n+1}^{e,l}\right) \tag{8}$$

where each $\text{GS}_{ij}(\theta)$ is a Givens SWAP gate corresponding to the product of a Givens rotation gate by an angle θ on a pair labelled by qubits i and j followed by a SWAP operation[12]

$$\text{GS}(\phi) = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\text{SWAP}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) & 0 \\ 0 & \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\text{Givens}}. \tag{9}$$

The GS gate corresponds to a coherent partial pair excitation (by the angle ϕ) followed by a pair SWAP. Given several layers N_ℓ in equation (6) and total number of qubits N, there are a total of $N_\ell N/2$ free parameters in the ansatz. To minimize the amount of time qubits are idle, we order the spatial orbitals such that the Fermi vacuum is $|0101 \ldots 01\rangle$—for example, the restricted Hartree–Fock state—corresponding to an interleaved list of occupied and virtual orbital labels in ascending energy order. The Hamiltonian qubit ordering is then chosen such that when all $\theta = 0$, the Hartree–Fock state for each model is returned. The alternating SWAP gate arrangement allows us to couple each occupied pair with each unoccupied pair once in depth $N/2$ (Supplementary Information Section II.C). Thus, in this work, we set $N_\ell = N/2$ for all systems. Each GS(θ) gate is compiled into a product of three controlled-Z (CZ) gates interleaved with tunable single-qubit microwave gates (Extended Data Fig. 1, top; see Supplementary Information Section II.B for decompositions).

To perform energy estimation on our two S_0 models, expectation values with respect to nearest-neighbour and non-nearest-neighbour qubits are required. The expectation value $\langle X_i X_j + Y_i Y_j \rangle$ is estimated by performing a number-preserving diagonalization[16,23] mapping the expectation value to the difference of $\langle Z_i \rangle$ and $\langle Z_j \rangle$. The ladder geometry allows us to measure all non-nearest-neighbour pairs across the rungs of the ladder in a similar fashion at the further cost of at most one SWAP operation. The full measurement protocol is detailed further in Supplementary Information Section II.C. All-to-all coupling is achieved in N circuits, bringing the total number of different circuits to measure the Hamiltonian's expectation value to $N + 1$. Strategies with fewer circuits exist, but they do not allow for postselection on particle number.

EV and VD

EV, introduced in ref. 44, is an error-mitigation technique that uses two copies of a quantum state $|\psi\rangle$ reflected in time (preparation ↔ unpreparation) to estimate $\langle\psi|O|\psi\rangle$ for a unitary O (refs. 46,52). EV can be implemented without control gates, given a known reference eigenstate $|\phi\rangle$ of O orthogonal to $|\psi\rangle$ (here $|\phi\rangle = |00 \ldots\rangle$). To implement (control-free) EV, we act O on a prepared superposition of $|\psi\rangle$ and $|\phi\rangle$, generated by

acting our UpCCD ansatz on the cat state $|00 \dots 0\rangle + |0101 \dots 01\rangle$. Then, we estimate the expectation value of $|\phi\rangle\langle\psi|$ (the term $|\phi\rangle\langle\psi|$ is not Hermitian but may be written as a sum of the Hermitian operators $|\phi\rangle\langle\psi| + |\psi\rangle\langle\phi|$ and $i|\phi\rangle\langle\psi| - i|\psi\rangle\langle\phi|$) on the resulting state $|\Psi\rangle = O\frac{1}{\sqrt{2}}(|\psi\rangle + |\phi\rangle)$. The estimation is performed by inverting the preparation unitary. In total, adding control-free EV to a variational circuit requires doubling the circuit depth and adding circuitry to prepare and unprepare GHZ states. This will reduce the fidelity of the final circuit from F to $< F^2$. As experiment runtime scales inversely with circuit fidelity, this will cost us slightly in the experiment.

To estimate expectation values by means of EV, we use the fact that

$$\langle\Psi|\phi\rangle\langle\psi|\Psi\rangle = \frac{1}{2}\langle\psi|O|\psi\rangle e^{i\phi}, \tag{10}$$

where $O|\phi\rangle = e^{i\phi}|\phi\rangle$. The expectation value $\langle\psi|O|\psi\rangle$ can be recovered from equation (10) as the other terms are known. The largest effect of noise on the echo-verified estimator (equation (10)) is to dampen $\langle\Psi|\phi\rangle\langle\psi|\Psi\rangle \rightarrow F\langle\Psi|\phi\rangle\langle\psi|\Psi\rangle$ (ref. [44]). We can estimate F independently by removing O from the circuit, which yields a Loschmidt echo of the preparation unitary[61]. This is achieved in practice by removing a virtual Z rotation (Extended Data Fig. 1, bottom), making the estimated Loschmidt fidelity an accurate estimate of F. Further EV implementation details can be found in Supplementary Information Section II.D.

VD[42,43] is an error-mitigation technique that uses collective measurements of k copies of a state ρ to estimate expectation values with respect to $\rho^k/\text{Tr}[\rho^k]$, where Tr is the trace. VD schemes are based on the observation that the cyclic shift operator $S^{(k)}$ is easily diagonalized and therefore can be measured, which yields, for example, for $k = 2$,

$$\text{Tr}[\rho \otimes \rho S^{(2)}] = \text{Tr}[\rho^2], \quad \text{Tr}[\rho \otimes \rho S^{(2)} O_s] = \text{Tr}[\rho^2 O], \tag{11}$$

with $O_s = \frac{1}{2}(I \otimes O + O \otimes I)$, and I is the identity. $S^{(2)}$ can be simultaneously diagonalized with O_s when $O = Z_i$ by a GS$(\pi/4)$ rotation between pairs of identified qubits on the two registers. For two $N/2 \times 2$ ladders on a square lattice geometry, this requires one round of SWAP gates to shift identified qubits next to each other. Operators $O \neq Z_i$ are measured by rotating to Z_i (Section 4.2) and following the above procedure. The VD circuit is only six two-qubit gates deeper than PS-VQE. However, as we double the number of qubits used, we more than double the total circuit size and reduce the total circuit fidelity again from F to $< F^2$.

As the GS$(\pi/4)$ gate is number-conserving, VD can be combined with postselection: the global excitation number $\sum_j (Z_j \otimes I + I \otimes Z_j)$ is a good symmetry. This requires that the state before measurement also conserve number. This is true when estimating $\langle X_i X_j + Y_i Y_j\rangle$ but not when estimating $\langle Z_i Z_j\rangle$: when mapping $Z_i Z_j \rightarrow Z_i$, one can only preserve the parity of the total number of excitations. In the main text of this work, we present results showing VD with postselection only (PS-VD). We compare VD with and without postselection in Supplementary Information Section VI.A.

Data availability
Raw and processed experimental data can be found at https://doi.org/10.5281/zenodo.7225821.

Code availability
Code to generate quantum circuits and process raw data is available in ReCirq: https://github.com/quantumlib/ReCirq.

References

55. Surján, P. R., Szabados, Á., Jeszenszki, P. & Zoboki, T. Strongly orthogonal geminals: size-extensive and variational reference states. *J. Math. Chem.* **50**, 534–551 (2012).
56. Ring, P. & Schuck, P. *The Nuclear Many-Body Problem* (Springer Science & Business Media, 2004).
57. Henderson, T. M., Bulik, I. W., Stein, T. & Scuseria, G. E. Seniority-based coupled cluster theory. *J. Chem. Phys.* **141**, 244104 (2014).
58. Stein, T., Henderson, T. M. & Scuseria, G. E. Seniority zero pair coupled cluster doubles theory. *J. Chem. Phys.* **140**, 214113 (2014).
59. Vu, N. & Eugene DePrince, A. Size-extensive seniority-zero energy functionals derived from configuration interaction with double excitations. *J. Chem. Phys.* **152**, 244103 (2020).
60. Shepherd, J. J., Henderson, T. M. & Scuseria, G. E. Using full configuration interaction quantum Monte Carlo in a seniority zero space to investigate the correlation energy equivalence of pair coupled cluster doubles and doubly occupied configuration interaction. *J. Chem. Phys.* **144**, 094112 (2016).
61. Mi, X. et al. Information scrambling in quantum circuits. *Science* **374**, 1479–1483 (2021).

Author contributions
T.E.O. calibrated the device and ran the experiments. V.E.E., G.A. and C.G. designed the UpCCD ansatz. V.E.E., G.A. and T.E.O. designed the scheduling of the ansatz onto a 2×N grid. F.G. and C.G. designed the conjugate model gradient descent algorithm and preoptimized ansatz parameters for the ansatz. N.C.R. wrote the pair coupled-cluster code, preoptimized ansatz parameters and performed the classical chemistry calculations for the CB model. T.E.O., W.J.H., S.P., K.K. and R.B. designed and optimized the error-mitigation and EV measurement strategies. O.O. and C.G. developed the BQP completeness proof for the UpCCD ansatz. T.E.O., N.C.R., C.G., F.G., V.E.E. and R.B. wrote the paper. T.E.O., C.G., R.B. and N.C. led and coordinated the project. All authors contributed to revising the manuscript and writing the Supplementary Information. All authors contributed to the experimental and theoretical infrastructure to enable the experiment.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41567-023-02240-y.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41567-023-02240-y.

Correspondence and requests for materials should be addressed to T. E. O'Brien, C. Gogolin, R. Babbush or N. C. Rubin.

Peer review information *Nature Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.
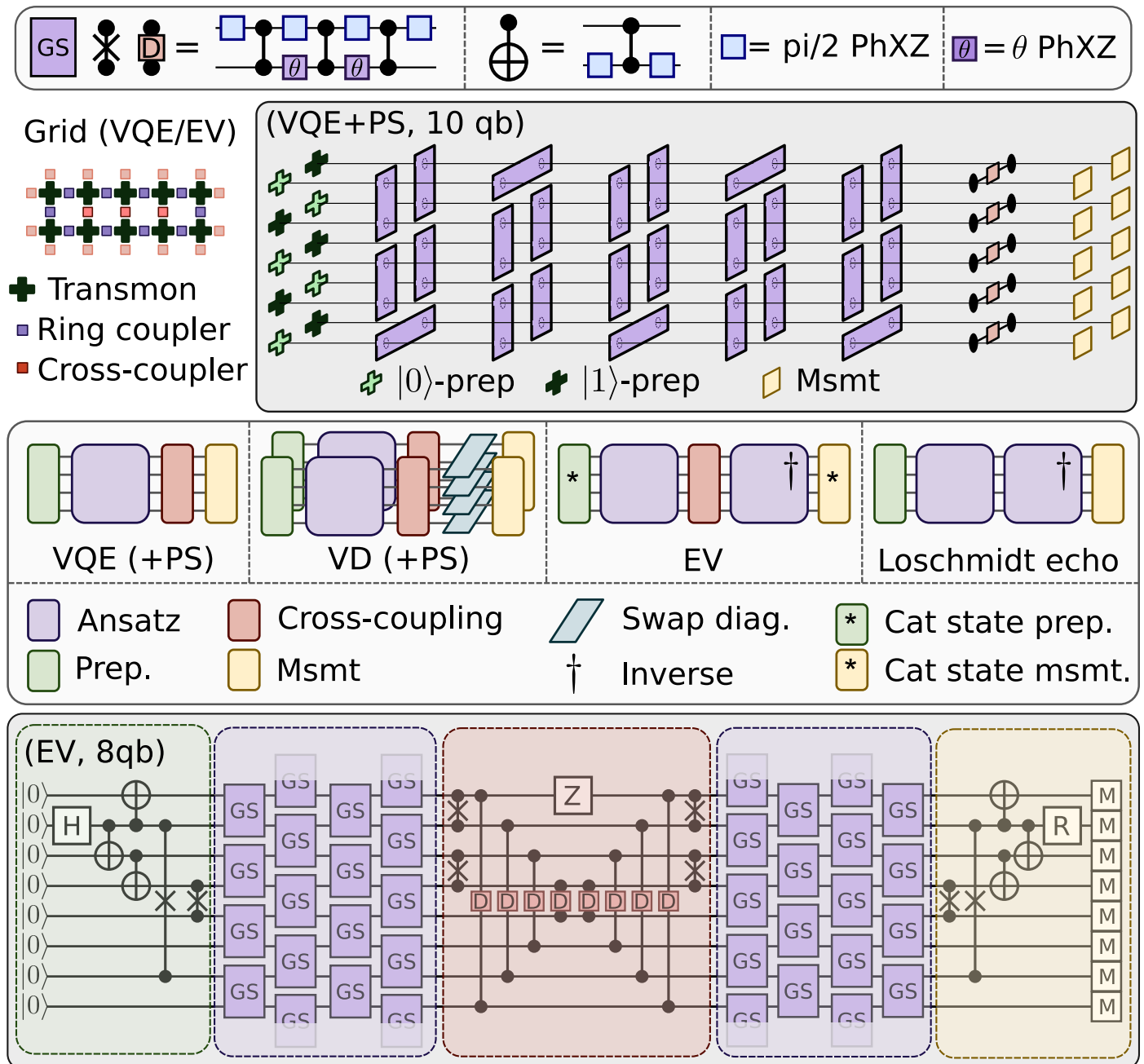
Extended Data Fig. 1 | The UpCCD ansatz and its compilation to a 2D superconducting transmon grid. (top) Decomposition of the gates used in this experiment to CZ and single-qubit gates. See Supplementary Information Section II.B for details. (second from top, left) 2 × 5 grid with couplers in a square lattice geometry, showing couplers used during the ansatz (ring coupler, purple) and those used only during measurement (cross-coupler, red). (second from top, right) 2+1D circuit cartoon of a combined ansatz and measurement on a 2 × 5 transmon qubit array. (third from top) Cartoon of error-mitigation techniques used in this experiment. Different circuit pieces are described in the legend. (bottom) an example 8-qubit echo verification circuit to measure the expectation value of $(X_1X_7 + Y_1Y_7 + Z_1 + Z_7)/2$. Shaded gates at the top and bottom of the qubit array wrap around the 2 × 4 ring.