



Universiteit  
Leiden  
The Netherlands

## Optimization of quantum algorithms for near-term quantum computers

Bonet Monroig, X.

### Citation

Bonet Monroig, X. (2022, November 2). *Optimization of quantum algorithms for near-term quantum computers*. *Casimir PhD Series*. Retrieved from <https://hdl.handle.net/1887/3485163>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3485163>

**Note:** To cite this publication please use the final published version (if applicable).

# 3. Experimental error mitigation via symmetry verification in a variational quantum eigensolver

## 3.1. Introduction

oisy intermediate-scale quantum (NISQ) devices [47], despite lacking layers of quantum error correction (QEC), may already be able to demonstrate quantum advantage over classical computers for select problems [18, 19]. In particular, the hybrid quantum-classical variational quantum eigensolver (VQE) [39, 87] may have sufficiently low experimental requirements to allow estimation of ground-state energies of quantum systems that are difficult to simulate purely classically [76–79]. To date, VQEs have been used to study small examples of the electronic structure problem, such as  $\text{H}_2$  [41, 42, 45, 80, 83, 98],  $\text{HeH}^+$  [39, 99],  $\text{LiH}$  [42, 80, 98], and  $\text{BeH}_2$  [42], as well as exciton systems [100], strongly correlated magnetic models [80], and the Schwinger model [101]. Although these experimental efforts have achieved impressive coherent control of up to 20 qubits, the error in the resulting estimations has remained relatively high due to performance limitations in the NISQ hardware. Consequently, much focus has recently been placed on developing error mitigation techniques that offer order-of-magnitude accuracy improvement without the costly overhead of full QEC. This may be achieved by using known properties of the target state, e.g., by checking known symmetries in a manner inspired by QEC stabilizer measurements [52, 53], or by expanding around the experimentally-obtained state via a linear (or higher-order) response framework [82]. The former, termed symmetry verification (SV), is of particular interest because it is comparatively low-cost in terms of required hardware and additional measurements. Other mitigation techniques require understanding the underlying error models of the quantum device, allowing for an extrapolation of the calculation to the zero-error

### 3. Experimental error mitigation via symmetry verification

limit [49, 50, 81], or the summing of multiple calculations to probabilistically cancel errors [50, 51, 102].

We experimentally demonstrate the use of SV to reduce the error of a VQE estimating the ground-state energy and the ground state of the  $\text{H}_2$  molecule by one order of magnitude on average across the bond-dissociation curve. Using two qubits in a circuit QED processor, we prepare a variational ansatz state via an exchange gate that finely controls the transfer of population within the single-excitation subspace while respecting the underlying symmetry of the problem (odd two-qubit parity). We show that SV improves the energy and state estimates by mitigating the effect of processes changing total excitation number, specifically qubit relaxation and residual qubit excitation. We do this through a full density-matrix simulation that matches the experimental energy and state errors with and without SV, and then using this simulation to dissect the contribution of each error source. Finally, we explore the limitations of SV arising from statistical measurement noise, and find that enforcing the positivity of the fermionic 2-reduced density matrix ties the improvement in energy estimation from SV to the improvement in ground-state fidelity (which was previously not the case).

## 3.2. Variational quantum eigensolvers for the Hydrogen molecule

A VQE algorithm [39, 87] approximates the ground state  $\rho^{(0)}$  of a Hamiltonian  $\hat{H}$  by a variational state  $\rho^{(\text{raw})}(\vec{\theta})$ , with  $\vec{\theta}$  a set of parameters that control the operation of a quantum device. These parameters are tuned by a classical optimization routine to minimize the variational energy  $E(\vec{\theta}) = \text{Tr}[\rho^{(\text{raw})}(\vec{\theta})\hat{H}]$ . In practice, this is calculated by expanding  $\rho^{(\text{raw})}(\vec{\theta})$  and  $\hat{H}$  over the  $N$ -qubit Pauli basis  $\mathbb{P}^N := \{I, X, Y, Z\}^{\otimes N}$ ,

$$\rho^{(\text{raw})}(\vec{\theta}) = \frac{1}{2^N} \sum_{\hat{P} \in \mathbb{P}^N} \rho_{\hat{P}}^{(\text{raw})}(\vec{\theta}) \hat{P}, \quad \hat{H} = \sum_{\hat{P} \in \mathbb{P}^N} h_{\hat{P}} \hat{P}, \quad (3.1)$$

where the Pauli coefficients are given by  $\rho_{\hat{P}}^{(\text{raw})}(\vec{\theta}) = \text{Tr}[\hat{P}\rho^{(\text{raw})}]$ . The variational energy may then be calculated as

$$E^{(\text{raw})}(\vec{\theta}) = \sum_{\hat{P} \in \mathbb{P}^N} \rho_{\hat{P}}^{(\text{raw})}(\vec{\theta}) h_{\hat{P}}. \quad (3.2)$$

For example, consider the  $\text{H}_2$  molecule studied in this chapter. Mapping the Hamiltonian of this system (in the STO-3G basis) onto four qubits

### 3.2. Variational quantum eigensolvers for the Hydrogen molecule

via the Bravyi-Kitaev transformation [35] and then further reducing dimensions by projecting out two non-interacting qubits [41] gives

$$\begin{aligned}\hat{H}_{\text{H}_2} = & h_{II}II + h_{ZI}ZI + h_{IZ}IZ \\ & + h_{XX}XX + h_{YY}YY + h_{ZZ}ZZ,\end{aligned}\quad (3.3)$$

where coefficients  $h_{\hat{P}}$  depend on the interatomic distance  $R$ . These coefficients may be determined classically using the OpenFermion [90] and psi4 [89] packages. The Pauli coefficients  $\rho_{\hat{P}}^{(\text{raw})}$  of the density matrix  $\rho^{(\text{raw})}$  are extracted by repeated preparation and (partial) tomographic measurements of the ansatz state. As one only needs those Pauli coefficients  $\rho_{\hat{P}}^{(\text{raw})}$  with non-zero corresponding Hamiltonian coefficients  $h_{\hat{P}}$ , one need not perform full tomography of  $\rho^{(\text{raw})}$ . However, in a small-scale experiment, full state tomography of  $\rho^{(\text{raw})}$  may still be feasible, and may provide useful information for the purposes of benchmarking. In particular, the fidelity of  $\rho^{(\text{raw})}$  to  $\rho^{(0)}$ ,

$$F^{(\text{raw})} = \text{Tr}[\rho^{(\text{raw})}\rho^{(0)}], \quad (3.4)$$

is a more rigorous measure of the ability to prepare the ground state than the energy error,

$$\Delta E^{(\text{raw})} = \text{Tr}\left[\left(\rho^{(\text{raw})} - \rho^{(0)}\right)\hat{H}\right]. \quad (3.5)$$

Error mechanisms such as decoherence pull  $\rho^{(\text{raw})}$  away from  $\rho^{(0)}$ , decreasing  $F$  and increasing  $\Delta E$ .

These errors may be mitigated by using internal symmetries  $\hat{S} \in \mathbb{P}^{N-1}$  of the target problem, such as parity checks [52, 53]. These checks project  $\rho^{(\text{raw})}$  to a symmetry verified matrix  $\rho^{(\text{SV})}$  that lies in the  $\langle \hat{S} \rangle = s$  subspace of the symmetry. This projection could be performed via direct measurement of  $\hat{S}$  on the quantum device, but one may instead extract the relevant terms of the density matrix  $\rho^{(\text{SV})}$  in post-processing:

$$\rho_{\hat{P}}^{(\text{SV})} = \frac{\rho_{\hat{P}}^{(\text{raw})} + s\rho_{\hat{S}\hat{P}}^{(\text{raw})}}{1 + s\rho_{\hat{S}}^{(\text{raw})}}, \quad (3.6)$$

The right-hand side may be obtained by partial tomographic measurement of the ansatz state, with at most twice the number of Pauli coefficients

---

<sup>1</sup>As described in Refs. [52, 53], one does not require  $\hat{S}$  to be a Pauli operator, however this makes the SV procedure significantly simpler.

### 3. Experimental error mitigation via symmetry verification

that need to be measured. This upper bound is not always achieved. For example, the  $\hat{H}_{\text{H}_2}$  Hamiltonian has a  $\hat{S} = ZZ$  symmetry, which maps the non-zero Pauli terms in  $\hat{H}_{\text{H}_2}$  to other non-zero Pauli terms in  $\hat{H}_{\text{H}_2}$ . Symmetry verification in this problem then does not require any additional measurements to estimate  $E^{(\text{SV})}$  beyond those already required to estimate  $E^{(\text{raw})}$ . Even when it does require additional measurements, SV remains attractive because it does not require additional quantum hardware or knowledge of the underlying error model. One can show that the SV state  $\rho^{(\text{SV})}$  may be equivalently obtained via a variant of the quantum subspace expansion (QSE) [82], suggesting an alternative name of S-QSE [52].

One may further minimize the error in a quantum algorithm by tailoring the quantum circuit or the gates within. In a VQE, one wishes to choose a variational ansatz motivated by the problem itself [41, 103] while minimizing the required quantum hardware [42]. To balance these considerations, we suggest constructing an ansatz from an initial gate-set that is relevant to the problem at hand. For example, in the electronic structure problem, the quantum state is generally an eigenstate of the fermion number. When mapped onto qubits, this often corresponds to a conservation of the total qubit excitation number. Gates such as single-qubit  $Z$  rotations, two-qubit C-Phase [104], and two-qubit iSWAP [105] gates preserve this number, making these gates a good universal gate set (within the target subspace [106]) for quantum simulation of electronic structure. In the example of  $\text{H}_2$ , the total two-qubit parity ( $ZZ$ ) is indeed conserved and the ground state at any  $R$  may be generated by applying to  $|01\rangle$  or  $|10\rangle$  an exchange gate

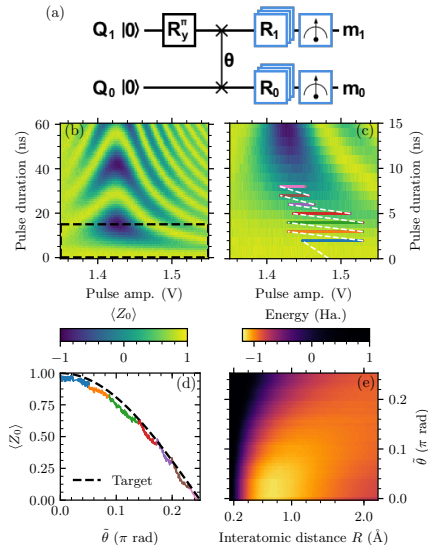
$$U_\theta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & i \sin \theta & 0 \\ 0 & i \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.7)$$

with  $R$ -dependent optimal exchange angle  $\theta$  and a follow-up phase correction on one qubit.

### 3.3. Experimental error mitigation via symmetry verification

We now experimentally investigate the benefits of SV in the VQE of  $\text{H}_2$  using two of three transmon qubits in a circuit QED quantum processor

### 3.3. Experimental error mitigation via symmetry verification



**Figure 3.1.:** Quantum circuit and energy landscape of the variational eigensolver. (a) Quantum circuit for generating and measuring the variational ansatz state. (b) Coherent excitation exchange, produced as  $Q_0$  is fluxed into resonance with  $Q_1$  by a square flux pulse. Pulse amplitude ( $x$  axis) parametrizes the frequency to which  $Q_0$  is flux pulsed ( $\sim 1.428$  V bringing it on resonance with  $Q_1$ ). (c) Zoom-in of (b) into the region used in the experiment to control the exchange of population between  $Q_0$  and  $Q_1$ . Colored lines illustrate the hybrid path in pulse duration and amplitude that maps out a finely-adjustable  $\tilde{\theta}$  range. (d) Excitation of  $Q_0$  along the hybrid path, showing the matching of the experimentally-defined  $\tilde{\theta}$  to the target  $\theta$  defined in Eq. (3.7) (black dashed curve). Colors [matching (c)] illustrate different pulse durations used in each segment. (e) Landscape of energies  $E^{(\text{raw})}(\tilde{\theta}, R)$  as function of the experimentally-defined  $\tilde{\theta}$  angle and the interatomic distance  $R$ .

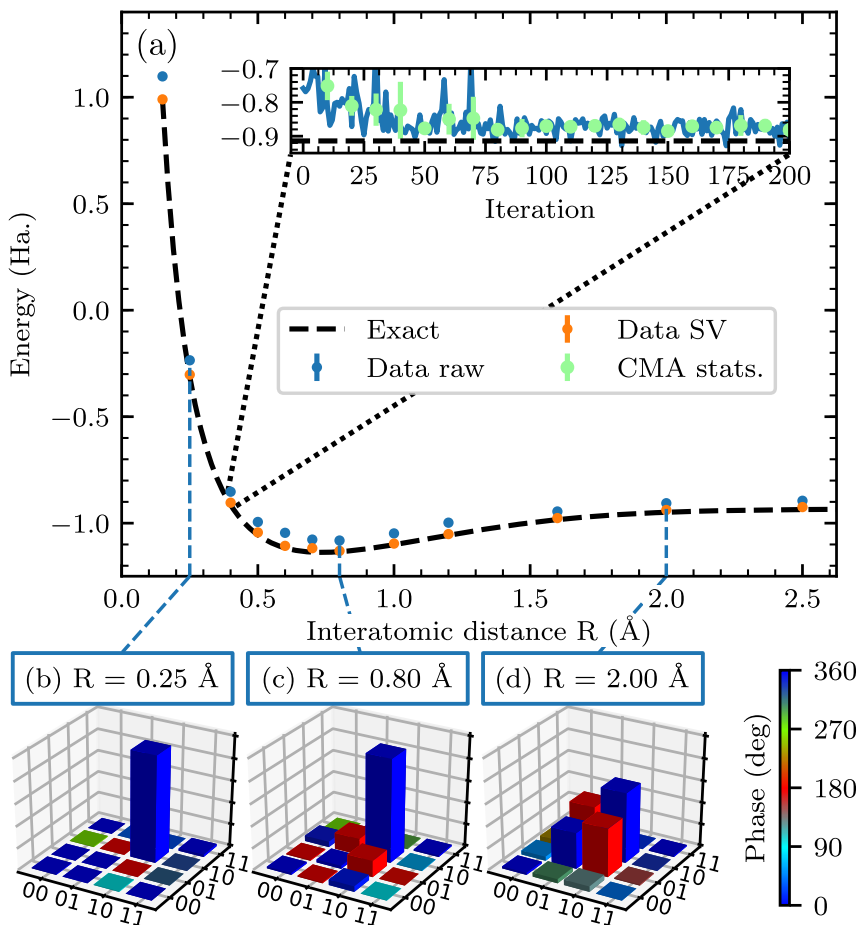
### 3. Experimental error mitigation via symmetry verification

(see details in [44]). The two qubits ( $Q_0$  and  $Q_1$ ) are coupled by a common bus resonator, and have dedicated microwave lines for single-qubit gating, flux bias lines for local and ns-scale control of their frequency, and dedicated readout resonators coupling to a common feedline for independent readout by frequency multiplexing. We prepare the ansatz state with an efficient circuit [Fig. 3.1(a)] that first excites  $Q_1$  with a  $\pi$  pulse to produce the state  $|10\rangle$ , and then flux pulses  $Q_0$  into resonance with  $Q_1$  to coherently exchange the excitation population. A sweep of flux-pulse amplitude and duration [Fig. 3.1(b)] reveals the expected chevron pattern that is the hallmark of coherent population exchange between the two qubits, albeit with some asymmetry arising from imperfect compensation of linear distortion in the flux-bias line. To finely control population exchange without being limited by the 1 ns resolution in pulse duration, we stitch together a hybrid path in pulse duration and amplitude. This results in a fine experimental knob  $\tilde{\theta}$  (1500 possible settings) that controls population exchange like  $\theta$  in Eq. (3.7) [Fig. 3.1(c)], although with additional single-qubit phases. The circuit concludes with simultaneous pre-rotation gates on both qubits followed by simultaneous measurement of both qubits, in order to perform tomography of the prepared ansatz state. To fully reconstruct the state, we use an overcomplete set of 36 pre-rotation pairs and extract estimates of the average measurement for each qubit as well as their shot-to-shot correlation using  $N_{\text{meas}}$  measurements per pre-rotation. Note that single-qubit phase corrections are not required immediately following the exchange gate, as phase rotations can be performed virtually from the fully-reconstructed state.

We now optimize the VQE to approximate the ground-state energy and ground state of  $\text{H}_2$ . At each chosen  $R$ , we employ the covariance matrix adaptation evolution strategy (CMA-ES) optimization algorithm [107], using  $E^{(\text{raw})}$  as cost function and  $\tilde{\theta}$  as single variational parameter. The evolutionary strategy optimizes  $\tilde{\theta}$  over repeated generations of  $N_{\text{pop}} = 10$  samples of  $E^{(\text{raw})}(\tilde{\theta})$ , each calculated from a raw density matrix  $\rho^{(\text{raw})}$  using linear inversion of  $N_{\text{meas}} = 10^3$ . A typical optimization [Fig. 3.2(a) inset] converges after  $\sim 20$  generations ( $\sim 2$  hours). The converged state is finally reconstructed with greater precision, using  $N_{\text{meas}} = 10^5$ . Figure 2 shows the resulting energy estimate for twelve values of  $R$  and the reconstructed optimized state at three such distances. These tomographs show that the optimal solutions are concentrated in the single-excitation subspace of the two qubits, with two-qubit entanglement increasing as a function of  $R$ .

Performing the described symmetry verification procedure on the con-

### 3.3. Experimental error mitigation via symmetry verification



**Figure 3.2.:** Convergence of the VQE algorithm. (a) Experimental VQE estimate of  $\text{H}_2$  ground-state energy as a function of interatomic distance  $R$ . At each chosen  $R$ , we minimize the raw energy  $E^{(\text{raw})}$  (blue data points) over the variational parameter  $\theta$  using the CMA-ES evolutionary algorithm [107]. Applying SV to the converged solution (orange data points) lowers the energy estimate towards the exact solution (dashed curve). Inset: A typical optimization trace for the convergence of the energy estimate. (b-d) The reconstructed density matrices of the converged states at (b)  $R = 0.25$  Å, (c)  $R = 0.80$  Å, and (d)  $R = 2.00$  Å, showing that the converged states lie mostly in the single-excitation subspace, and that entanglement increases with the interatomic distance  $R$ .

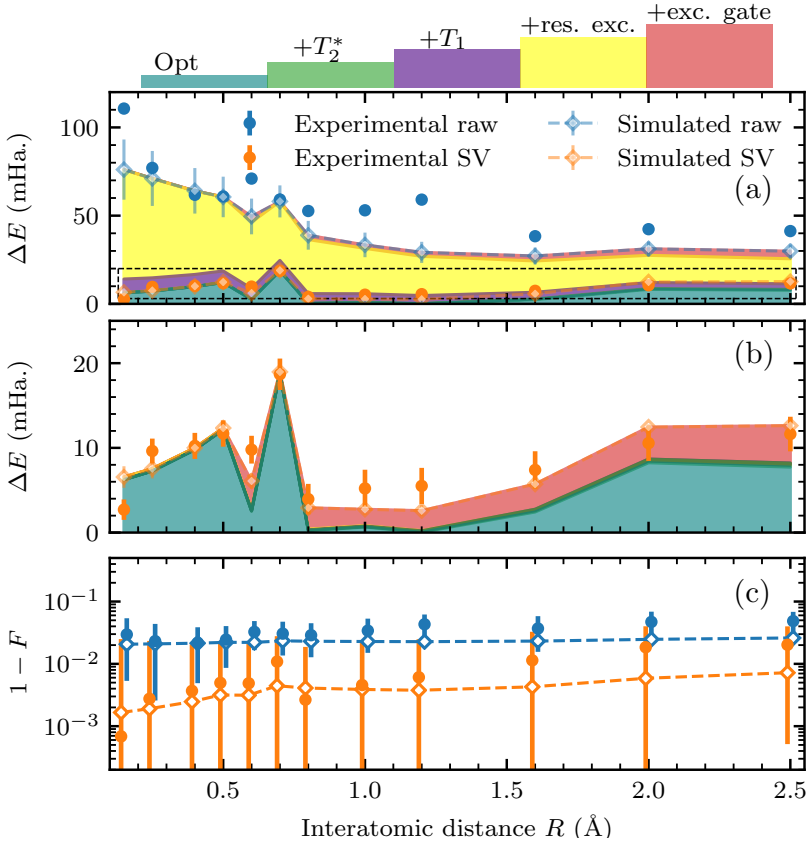


### 3. Experimental error mitigation via symmetry verification

verged states shows improvement across the entire bond-dissociation curve. To quantify the improvement, we focus on the energy error  $\Delta E$  and the infidelity  $1 - F$  to the true ground state, with and without SV (Fig. 3.3). SV reduces the energy error by an average factor  $\sim 10$  and reduces the infidelity by an average factor  $\sim 9$ . In order to quantitatively understand the limits of the VQE optimization, and to clearly pinpoint the origin of the SV improvement, we simulate the experiment via the density-matrix simulator *quantumsim* [92], using an error model built from independently measured experimental parameters [44]. We build the error model incrementally, progressively adding: optimization inaccuracy (the difference between the state ideally produced by the converged  $\theta$  and the true ground state); dephasing on both qubits (quantified by the measured Ramsey dephasing times  $T_2^*$ ); relaxation on both qubits (quantified by the measured relaxation times  $T_1$ ); residual qubit excitations (measured from single-shot histograms with each qubit prepared in  $|0\rangle$ ); and increased dephasing of  $Q_0$  during the exchange gate (quantified by its reduced  $T_2^*$  when tuned into the exchange interaction zone). By plotting the errors from each increment of the model, we are able to dissect the observed experimental error into its separate components without [Fig. 3.3(c)] and with [Fig. 3.3(b)] SV. Measured temporal fluctuations of dephasing, relaxation and residual excitation are used to obtain simulation error bars.

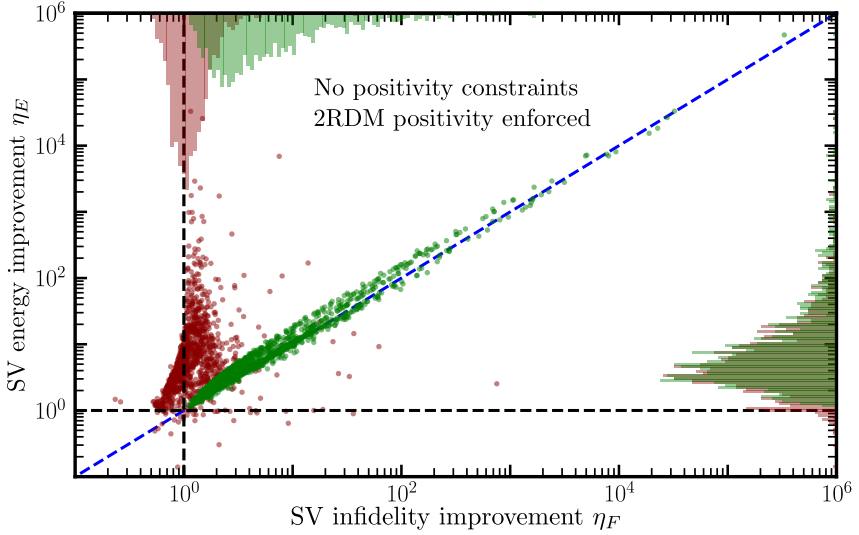
The simulation using the full error model shows fairly good matching with experiment for both the ground-state energy error [Figs. 3.3(a,b)] and the state infidelity [Fig. 3.3(c)], without and with SV. The error model dissection shows that the energy error when not using SV is dominated by residual qubit excitations. This is remarkable as the calibrated residual excitations are only 0.25% for  $Q_0$  and 1.34% for  $Q_1$  [44]. The improvement from SV results from the mitigation of errors arising from these residual excitations and from qubit relaxation. This is precisely as expected: these error mechanisms change total qubit excitation number and violate the underlying  $ZZ$  symmetry. Using SV changes the dominant error mechanism to optimization inaccuracy. This error could be reduced experimentally by increasing  $N_{\text{meas}}$  during the optimization, at the cost of increased convergence time. The improvement in state infidelity by SV can be explained along similar lines. We observe some increased deviations between the observed and simulated state infidelity at large  $R$ . We attribute these to limitations in our modeling of error during the exchange gate (whose duration increases with  $R$ ).

### 3.3. Experimental error mitigation via symmetry verification



**Figure 3.3.:** Impact of SV in ground-state energy and state fidelity, and dissected error budget. (a) Experimental (solid circles) energy error  $\Delta E$  without and with SV compared to the result (empty circles) of a full density-matrix simulation using the full error model. The contributions from optimizer inaccuracy, qubit dephasing, qubit relaxation, residual qubit excitations and increased  $Q_0$  dephasing during the exchange gate are shown as shaded regions for the case of no SV applied. Without SV,  $\Delta E$  is clearly dominated by residual qubit excitation. (b) Zoom-in on experimental and simulated  $\Delta E$  with SV and corresponding error budget. With SV, the effects of residual excitation and qubit relaxation are successfully mitigated, as predicted in Ref. 52. The remaining energy error is dominated by optimizer inaccuracy. Simulation error bars are obtained by modelling measured fluctuations of  $T_1$ ,  $T_2^*$ , and residual excitation. (c) Experimental (solid circles) infidelity to the true ground state without and with SV compared to simulation using the full error model (empty circles).

### 3. Experimental error mitigation via symmetry verification



**Figure 3.4.:** Constraining positivity with symmetry verification to mitigate the effect of sampling noise. The experimental data from Fig. 3.3 is split into 100 sample simulations for each  $R$ , increasing the sampling noise by a factor of 10 and making it comparable to other sources of experimental error. For each sample, we plot (red) the relative energy error and infidelity [Eq. (3.9)]. Values below 1 (dashed lines) indicate that SV has not provided an improvement, as may be the case when the density matrix has negative eigenvalues. We restore the improvement from SV by constraining the positivity of the 2-reduced density matrix [108] (green). Histograms on the top and right axes show the marginal distribution of the two scatter plots. When the density matrices are constrained to be positive, we observe the points fall along the line  $y = x$  (blue dashed line), indicating that SV improves both metrics by the same amount.

### 3.4. Effect of symmetry verification on positivity constrains

VQEs rely on variational bounding to ensure that the obtained approximation to the ground-state energy is accurate, but this is only guaranteed when the experimental results correspond to a physical state. Our method for calculating the ground-state energy [Eq. (3.1)] independently estimates each Pauli coefficient of the density matrix with error  $\propto N_{\text{meas}}^{-1/2}$ . Such estimation cannot guarantee a set of Pauli coefficients that could have come from a positive density matrix. This in turn breaks the variational lower bound on the energy estimate, and increases the error in estimates of other properties of the true ground state [108, 109]. As experimental error is reduced,  $\rho^{(\text{raw})}$  tends towards a rank-1 density matrix, increasing its chance of being unphysical [109]. Moreover,  $\rho^{(\text{SV})}$  is a lower-rank density matrix than  $\rho^{(\text{raw})}$  (being projected onto a subspace of the Hilbert space), which implies that unphysicality may be enhanced by SV. The variance in a given term  $\rho_{\hat{P}}$  post-SV can be calculated as

$$\text{Var}[\rho_{\hat{P}}^{(\text{SV})}] \approx \frac{3N_{\text{meas}}}{N_{\text{meas}}(1 + \text{Tr}[\rho^{(\text{raw})}\hat{S}])}. \quad (3.8)$$

SV has maximal impact on the quantum state precisely when this denominator is small, so this represents a natural bound for the power of SV as an error mitigation strategy.

The effect of sampling noise may be mitigated somewhat by restricting the fermionic 2-reduced density matrix to be positive (which may be completed in polynomial time) [108]. To investigate the effect of such mitigation, we bin the data used for final tomography of converged states to construct 100 density matrices with  $N_{\text{meas}} = 10^3$  at each  $R$ , thus increasing the sampling noise by a factor of 10. We wish to study the relative improvement of SV in the two figures of merit, which we quantify as

$$\eta_E = \frac{|\Delta E^{(\text{raw})}|}{|\Delta E^{(\text{SV})}|} \text{ and } \eta_F = \frac{|1 - F^{(\text{raw})}|}{|1 - F^{(\text{SV})}|}, \quad (3.9)$$

when physicality of the raw density matrices is enforced and not. To enforce physicality, we employ a convex optimization routine to find the closest positive semidefinite matrix to the experimentally measured  $\rho^{(\text{raw})}$  (closest in the  $L^2$  norm sense on the space induced by the the Pauli basis). We then apply symmetry verification to the post-processed density matrix. Figure 3.4 shows a scatter plot of  $\eta_E$  and  $\eta_F$ , and relative histograms

### 3. Experimental error mitigation via symmetry verification

of each. Without enforcing physicality, SV makes no significant improvement to the state fidelity, although it almost always improves the energy error. However, when positivity is enforced, SV greatly improves the overlap with the true ground state. We also find that the improvement in the energy from SV is equal to the improvement in fidelity when the starting state is physical, but is relatively uncorrelated when the starting state is not. This makes sense, as the energy gain from SV given a physical matrix comes directly from substituting higher energy states with density on the ground state. It is unclear whether such a strong trend will continue in larger systems without requiring too stringent a positivity constraint. As this is a four-orbital two-electron system, enforcing the positivity of the 2-reduced density matrix enforces positivity on the entire density matrix (which is exponentially difficult in the system size [110]). Testing this scalability is a clear direction for future research <sup>2</sup>.

## 3.5. Conclusion

In summary, we have experimentally demonstrated the use of SV to mitigate errors in the VQE of  $H_2$  with two transmon qubits. We implemented an efficient variational ansatz based on an exchange gate producing finely adjustable population transfer in the single-excitation subspace, respecting the  $ZZ$  symmetry of the  $H_2$  Hamiltonian. Verification of this symmetry reduced the error of the estimated ground-state energy and the ground state by one order of magnitude on average over the full dissociation curve. A full density-matrix simulation of our system allowed us to budget the contributions from known experimental error mechanisms. We observe that SV mitigates the effect of processes that affect total qubit excitation number, specifically qubit relaxation and residual excitation. Finally, we have investigated the effect of reconstructing density matrices via linear tomographic inversion in the presence of sampling, which voids the guarantee of positivity and in turn the guarantee that SV improves estimation of the ground state. Intriguingly, we observe that when physicality is enforced, the reduction in energy error from SV is directly linked to the increase in fidelity to the ground state. If this observation extends to larger systems, a user can be confident that symmetry-verified Pauli coefficients are accurate for calculations beyond the ground-state energy.

---

<sup>2</sup>Note that, for this system, enforcing positivity of the 1-reduced density matrix corresponds to ensuring that all expectation values are bounded between  $-1$  and  $1$ , and so this does not provide any additional data.

# Appendix

## 3.A. Appendix: Tomographic reconstruction and limitations

Tomographic reconstruction was performed with the same technique described in [111]. We provide a brief description here for completeness. For each measurement channel (measurement of  $Q_1$ , measurement of  $Q_0$ , and their correlation), the average measurement outcome is given by  $\langle m_i \rangle = \text{Tr}(\hat{M}_i \rho)$ , with operator

$$\hat{M}_i = \beta_{II}^i \hat{I}I + \beta_{IZ}^i \hat{I}Z + \beta_{ZI}^i \hat{Z}I + \beta_{ZZ}^i \hat{Z}Z, \quad (3.10)$$

and real-valued coefficients  $\beta_j^i$ . Single-shot measurements of  $Q_0$  and  $Q_1$  are 1-bit digitized before correlation and before averaging each of the three channels.

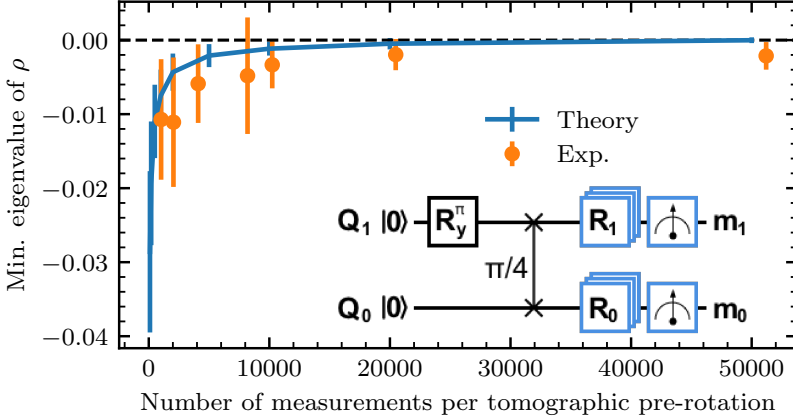
The simultaneously applied measurement pre-rotations  $R_0$  and  $R_1$  consist of the 36 pairs created by drawing each rotation separately from the set  $\{I, X_\pi, X_{\pi/2}, Y_{\pi/2}, X_{-\pi/2}, Y_{-\pi/2}\}$ . These measurement pre-rotations effectively change the measurement operator to

$$M_i^{k,l} = \text{Tr}\left(R^{k,l,\dagger} \hat{M}_i R^{k,l}\right).$$

There are thus 108 linear equations (36 per channel) linking the averaged measurement to the 15 nontrivial 2-qubit Pauli coefficients (we force  $\langle \hat{I}I \rangle = 1$ ). We then extract the Pauli coefficients by performing least-squares linear inversion. Prior to the linear inversion, the measurements are scaled to approximately match the noise in the three channels.

The coefficients  $\beta_j^i$  are obtained from standard calibration measurements. The two qubits are nominally prepared in the four computational states and measured. In total, we perform  $7 \times N_{\text{meas}}$  measurements per computational state. The matrix relating the four measurement averages of a channel to the coefficients has elements of the form  $\langle \hat{I}I \rangle$ ,  $\pm \langle \hat{I}Z \rangle$ ,  $\pm \langle \hat{Z}I \rangle$  and  $\pm \langle \hat{Z}Z \rangle$ . By taking into account the calibrated residual qubit excitations, which reduce the magnitude of  $\langle \hat{I}Z \rangle$ ,  $\langle \hat{Z}I \rangle$ , and  $\langle \hat{Z}Z \rangle$  from

### 3. Experimental error mitigation via symmetry verification



**Figure 3.5.:** Minimal eigenvalue of density matrices obtained from linear tomographic reconstruction with different  $N_{\text{meas}}$ . Here, the state preparation targets a Bell state using our variational ansatz (inset) with  $\theta = \pi/4$  (producing a  $\sqrt{\text{iSWAP}}$  gate). Physicality constraints on density matrices restrict their eigenvalues to be non-negative. We observe negative minimum eigenvalues over the entire range of  $N_{\text{meas}}$ . A *quantumsim* simulation produces a similar trend, asymptotically approaching a physical state by  $N_{\text{meas}} \sim 50,000$ .

unity, we ensure that the coefficients  $\beta_j^i$  and thus also the operator  $M_i$  are not corrupted by residual excitation [112].

Tomography by linear inversion does not ensure physicality of the reconstructed density matrix. We investigate this effect by performing tomography with variable  $N_{\text{meas}}$  on the state produced by our ansatz with  $\theta = \pi/4$  and extracting the minimum eigenvalue of the reconstructed density matrix (Fig. 3.5). A negative minimum eigenvalue manifests unphysicality over the  $N_{\text{meas}}$  range covered. Our *quantumsim* simulation produces a similar trend, asymptotically approaching a physical state by  $N_{\text{meas}} = 5 \times 10^4$ . These observations led us to choose  $N_{\text{meas}} = 10^5$  for the final state tomography post VQE convergence in Fig. 3, and to further investigate (in Fig. 4) how unphysicality can violate the variational principle, producing reductions in energy from imprecise state reconstruction rather than algorithmic precision.

### 3.B. Appendix: Constraining the positivity of reduced density matrices

Testing whether a  $N$ -qubit density matrix  $\rho$  is positive is in general QMA-hard [110]. However, if we trace out all but a polynomial number of degrees of freedom of  $\rho$ , testing positivity of the resulting reduced density matrix  $\rho^{(\text{red})}$  is tractable on classical hardware, and obtaining the closest nearby positive matrix is similarly so. This gives a set of necessary but insufficient physicality conditions for  $\rho$ , but enforcing  $k$ -local constraints (on a density matrix from a VQE) tends to be sufficient to variationally bound the resulting energies [108]. Following the reduction, we write  $\rho^{(\text{red})}$  as a vector over the Pauli basis,

$$\rho^{(\text{red})} = \sum_{\hat{P}} \rho_{\hat{P}}^{(\text{red})} \hat{P}. \quad (3.11)$$

Then, we attempt to find the density matrix  $\tilde{\rho}^{(\text{red})}$  closest to  $\rho^{(\text{red})}$  in the  $L^2$ -norm

$$\sum_{\hat{P}} \left( \rho_{\hat{P}}^{(\text{red})} - \tilde{\rho}_{\hat{P}}^{(\text{red})} \right)^2, \quad (3.12)$$

subject to the conditions  $\tilde{\rho}^{(\text{red})} \succ 0$ , and  $\tilde{\rho}_I^{(\text{red})} = 1$ . This gives a quadratic minimization problem with cone inequality and linear equality constraints, which we solve using interior point methods.

### 3.C. Appendix: Theoretical modeling of the experiment

We use our full-density-matrix simulator *quantumsim* to model the experiment. The error model takes as input parameters the measured values of  $T_1$ ,  $T_2^*$  and residual excitation for both qubits, and  $T_2^{*,\text{red}}$  for  $Q_0$ . We also include the effect of fluctuations on the device parameters by Monte Carlo sampling.

#### 3.C.1. Numerical simulations

The simulations are performed by extracting the full-density-matrix  $\rho^{(\text{sim})}$  at the end of the circuit. We use the converged value of  $\tilde{\theta}$  at each  $R$  to generate the quantum state and extract the Pauli coefficients  $\rho_{\hat{P}}^{(\text{sim})}(\tilde{\theta}) =$



### 3. Experimental error mitigation via symmetry verification

$\text{Tr}[\hat{P}\rho^{(\text{sim})}]$ . We add sampling noise to each coefficient, drawn from a zero-mean Gaussian distribution with variance  $(1 + \rho_{\hat{P}}^{(\text{sim})})(1 - \rho_{\hat{P}}^{(\text{sim})})/N_{\text{meas}}$ , where  $N_{\text{meas}} = 4 \times 10^5$ . Note that this is greater than the number of measurements per tomographic prerotation in the experiment, as data from multiple tomographic prerotations is used to estimate each Pauli coefficient. To account for fluctuations on the device parameters  $T_1$ ,  $T_2^*$ , and residual excitations, we average over  $10^4$  simulations for every  $R$ . For each simulation, we draw parameters from independent normal distributions using values in Table I of the supplementary material of Ref. [44]. As the dephasing noise  $T_2^*$  depends on  $T_1$ ,

$$\frac{1}{T_2^*} = \frac{1}{T_\phi} + \frac{1}{2T_1}, \quad (3.13)$$

it is more appropriate to sample the pure dephasing rate  $T_\phi$  independently. We calculate the  $T_\phi$  mean and variance  $(\bar{T}_\phi, \text{Var}[\bar{T}_\phi])$  from  $T_2^*$ ,

$$\begin{aligned} \bar{T}_\phi &= \frac{1}{\frac{1}{T_2^*} - \frac{1}{2T_1}}, \\ \text{Var}[\bar{T}_\phi] &= \bar{T}_\phi^2 (T_2^*)^{-2} \left[ \text{Var}[T_2^*] - \frac{\text{Var}[T_1]}{2T_1^2 (T_2^*)^{-2}} \right]. \end{aligned} \quad (3.14)$$

From the  $10^4$  simulations we obtain 95%-confidence error bars for  $\Delta E$  and  $F$  as twice the population standard deviation.

#### 3.C.2. Exchange gate

Due to quasi-static flux noise, the angle of the unitary exchange gate (Eq. 7) differs between subsequent applications. Assuming that these fluctuations are fast on the scale of the 2 hour optimization, this may be simulated by integrating over the range of applied gates, resulting in an incoherent noise model. As the gate is not repeatedly applied during a single-shot experiment, this incoherent approximation does not lead to an error in the final result. To perform the integration, we convert our unitary  $U_\theta$  into a Pauli Transfer Matrix representation (PTM) [113, 114]:

$$[R_\theta]_{i,j} = \frac{1}{2} \text{Tr}[\hat{P}_i U_\theta \hat{P}_j U_\theta^\dagger], \quad \hat{P}_i, \hat{P}_j \in \mathbb{P}^N, \quad (3.15)$$

which may then be integrated over a probability distribution in the deviation  $\delta$  from the target angle  $\theta$ :

$$[\tilde{R}]_{i,j} = \int d\delta p(\delta) [R_{\theta+\delta}]_{i,j}. \quad (3.16)$$

### 3.C. Appendix: Theoretical modeling of the experiment

We choose for  $p(\delta)$  a Gaussian distribution:  $p(\delta) = e^{-\frac{\delta^2}{2\sigma^2}}$ . In order to obtain the distribution width  $\sigma^2$ , we note that the same effect causes single-qubit dephasing of  $Q_0$  when fluxed to the exchange point when  $Q_1$  is fluxed away. We may thus estimate  $\sigma$  as

$$\sigma^2 = 1 - e^{-\frac{t_{\text{int}}}{T_2^{*,\text{red}}}}, \quad (3.17)$$

where  $t_{\text{int}}$  is the exchange gate duration and  $T_2^{*,\text{red}}$  the dephasing time of  $Q_0$  at the exchange point (with  $Q_1$  fluxed away). The final gate simulation also includes the effect of amplitude damping on both qubits, and the dephasing of  $Q_1$  at the sweet spot as discrete error channels of duration  $\frac{t_{\text{int}}}{2}$  on either side of the gate  $\tilde{R}$ .

