

A stabilizer framework for Contextual Subspace VQE and the noncontextual projection ansatz

Tim J. Weaving¹, Alexis Ralli¹, William M. Kirby², Andrew Tranter³, Peter J. Love^{2,4}, and Peter V. Coveney^{1,5,6}

¹Centre for Computational Science, Department of Chemistry, University College London, WC1H 0AJ, United Kingdom

²Department of Physics and Astronomy, Tufts University, Medford, MA 02155, USA

³Cambridge Quantum Computing, 9a Bridge Street Cambridge, CB2 1UB, United Kingdom

⁴Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, USA

⁵Advanced Research Computing Centre, University College London, WC1H 0AJ, United Kingdom

⁶Informatics Institute, University of Amsterdam, Amsterdam, 1098 XH, Netherlands

Quantum chemistry is a promising application for noisy intermediate-scale quantum (NISQ) devices. However, quantum computers have thus far not succeeded in providing solutions to problems of real scientific significance, with algorithmic advances being necessary to fully utilise even the modest NISQ machines available today. We discuss a method of ground state energy estimation predicated on a partitioning the molecular Hamiltonian into two parts: one that is *noncontextual* and can be solved classically, supplemented by a *contextual* component that yields quantum corrections obtained via a Variational Quantum Eigensolver (VQE) routine. This approach has been termed *Contextual Subspace VQE* (CS-VQE), but there are obstacles to overcome before it can be deployed on NISQ devices. The problem we address here is that of the *ansatz* – a parametrized quantum state over which we optimize during VQE. It is not initially clear how a splitting of the Hamiltonian should be reflected in our CS-VQE *ansätze*. We propose a ‘noncontextual projection’ approach that is illuminated by a reformulation of CS-VQE in the stabilizer formalism. This defines an *ansatz* restriction from the full electronic structure problem to the contextual subspace and facilitates an implementation of CS-VQE that may be deployed on NISQ devices. We validate the noncontextual

projection ansatz using a quantum simulator, with results obtained herein for a suite of trial molecules.

1 Introduction

Quantum computers promise to yield solutions to complex problems that have previously been unattainable by classical means, yet experimental demonstration remains challenging. To date, the largest molecules simulated on noisy intermediate-scale quantum (NISQ) hardware are H_{12} , conducted by Google using just 12 of the 53 qubits available on their superconducting quantum processor *Sycamore* [1], and H_2O , performed independently by IonQ using 3 qubits of an unspecified proprietary trapped ion device [2] and IBM, using 5 of the 27 qubits on the now-decommissioned *ibmq_dublin* superconducting device [3].

Due to the limitations of shallow circuit depth and short coherence times that characterise the NISQ era, we are not able to harness the full state-space afforded to these machines. To circumvent the above issues, we turn to the class of variational quantum algorithms, of which the Variational Quantum Eigensolver (VQE) [4] is most widely studied. In contrast with eigenvalue-finding algorithms requiring fault-tolerant machines such as Quantum Phase Estimation (QPE) [5] – which necessitates state evolution over an extended period of coherence – VQE executes a large ensemble of comparatively shallow parametrized circuits to estimate energy expectation values, informing a classical optimizer that

Tim J. Weaving: timothy.weaving.20@ucl.ac.uk

updates the parameter settings before reinitialization of the quantum circuit. Its success is predicated on the variational principle, meaning the ground state energy of the system bounds expectation values from below [6].

However, VQE is not without its challenges. First of all, the parametrized quantum state mentioned above – known as an *ansatz* – needs to be constructed carefully; it must be sufficiently expressible so the subspace of quantum states it spans contains the true ground state. On the other hand, if the ansatz is too expressible, we run into the problem of barren plateaus [7] where we observe vanishing gradients. This is more often a symptom of ‘hardware efficient’ ansätze [1, 8–12], which aim to access the largest possible region of Hilbert space for the fewest number of native quantum gates.

To avoid barren plateaus, one must take into account some of the underlying problem structure to define ansatz circuits whose image is confined to a smaller, but more targeted, region of Hilbert space. Within this category are ‘chemically inspired’ ansätze that represent sequences of electronic excitation operators in circuit; unitary coupled cluster (UCC) [13, 14] is widely acknowledged as the gold standard for electronic structure simulations, however it is computationally expensive in practice.

More recently, we have seen the development of hybrid ansätze that bridge the gap between hardware efficiency and chemical motivation. For example, in [15] a compact circuit designed to conserve molecule symmetries such as particle number and spin was presented, while Adaptive Derivative-Assembled Pseudo-Trotter (ADAPT) VQE [16–19] describes a more complete approach to scalable quantum chemistry simulations by defining selection criteria of ansatz terms from a pool of excitation operators.

Secondly, the energy estimation procedure in VQE invokes the measurement problem – in order to mitigate statistical error, many prepare-and-measure cycles are necessary to achieve sufficient precision in the estimate. The advances made in recent years towards measurement reduction techniques are expansive [20–29] and range from classical pre/post-processing of the measurement information such as in classical shadow tomography [30, 31] to Hamiltonian term-grouping schemes and reductions in the number of Hamil-

tonian terms at a cost of coherent resource, such as in unitary partitioning [32–35]. Combined with techniques of error mitigation [36–43], one can optimize VQE with the objective of maximal NISQ resource utilization.

In this work we are concerned with Contextual Subspace VQE (CS-VQE) [44], which describes a method of partitioning the molecular Hamiltonian into disjoint parts so that an electronic structure problem may be simulated to some degree on the available quantum device, even if the dimension of the problem is too great to be encoded on the number of qubits available. This is supplemented by some classical overhead, but in many instances chemical accuracy (1.6 mHa $\approx$ 4 kJ/mol) may be achieved using fewer qubits than in a full VQE routine.

There has since been further research into the use of classical estimates of the electronic structure problem to reduce the resource requirement on quantum hardware. In particular, Classically-Boosted VQE (CB-VQE) [45] identifies classically tractable states and excludes them from the quantum simulation, alleviating some measurement and fidelity requirements of the VQE routine.

CS-VQE bears a resemblance to the qubit reduction technique of qubit tapering [46, 47], which exploits $\mathbb{Z}_2$ symmetries of the Hamiltonian; the differences and similarities are highlighted herein and in [44]. There are still a number of problems to address before CS-VQE may be successfully deployed on real quantum hardware – most notably with regard to the ansatz, which is the principal consideration of this work.

2 Preliminaries

The notation used throughout shall be to write operators in standard capital font ($A, B, C, \dots$) – with the exception of single-qubit Pauli operators being written

$$\sigma_p := \begin{pmatrix} \delta_{p,0} + \delta_{p,3} & \delta_{p,1} - i\delta_{p,2} \\ \delta_{p,1} + i\delta_{p,2} & \delta_{p,0} - \delta_{p,3} \end{pmatrix} \quad (1)$$

for $p \in \{0, 1, 2, 3\}$ – sets in calligraphic letters ($\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$) and vector spaces in script typeface ($\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$). The state space of N qubits may be identified with the 2^N dimensional Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$, with the space of (bounded) linear operators acting upon $\mathcal{H}$ denoted $\mathcal{B}(\mathcal{H})$.

We introduce the *Pauli group* $\mathcal{P}_N \subset \mathcal{B}(\mathcal{H})$ of operators $\boldsymbol{\sigma} = \bigotimes_{i=0}^{N-1} \sigma_{p_i}$ for $p_i \in \{0, 1, 2, 3\}$, up to multiplication by $\pm 1, \pm i$. Note the distinction between the bold font $\boldsymbol{\sigma}$ denoting tensor products and σ_p a single-qubit Pauli operator; we will sometimes write $\sigma_p^{(i)}$ to index explicitly the qubit position $i \in \mathbb{Z}_N$ on which it acts. We shall also make use of the *commutator* $[A, B] := AB - BA$ and *anticommutator* $\{A, B\} := AB + BA$, defined for operators $A, B \in \mathcal{B}(\mathcal{H})$, which are zero when A and B commute/anticommute, respectively.

An N -qubit Hamiltonian can be written in the form

$$H_{\mathcal{T}} = \sum_{\boldsymbol{\sigma} \in \mathcal{T}} h_{\boldsymbol{\sigma}} \boldsymbol{\sigma}, \quad h_{\boldsymbol{\sigma}} \in \mathbb{R}, \quad (2)$$

for a set of Pauli operators $\mathcal{T} \subset \mathcal{P}_N$ – specifying real coefficients ensures $H_{\mathcal{T}}$ is *Hermitian*. The objective of quantum chemistry simulations is to estimate the ground state energy

$$\epsilon_0 := \min_{|\psi\rangle \in \mathcal{H}} \langle H_{\mathcal{T}} \rangle_{\psi} \quad (3)$$

where $\langle H_{\mathcal{T}} \rangle_{\psi} := \langle \psi | H_{\mathcal{T}} | \psi \rangle$ is the expectation value of $H_{\mathcal{T}}$ with respect to some quantum state $|\psi\rangle \in \mathcal{H}$. Many physical properties of the target system are determined by the ground state, motivating this goal.

The Variational Quantum Eigensolver (VQE) quantum-classical hybrid algorithm [4] is the most widely studied means of achieving this on NISQ hardware. VQE requires a parametrized ansatz state $|\psi_{\text{anz}}(\boldsymbol{\theta})\rangle \in \mathcal{H}$ whose parameters $\boldsymbol{\theta}$ are manipulated within a classical optimization scheme that aims to minimize the energy expectation value

$$E(\boldsymbol{\theta}) := \langle H_{\mathcal{T}} \rangle_{\psi_{\text{anz}}(\boldsymbol{\theta})}, \quad (4)$$

evaluated via many prepare-and-measure cycles. The choice of ansatz restricts us to a subspace of quantum states and therefore must be carefully designed to be sufficiently expressible so as to capture the true ground state of the system.

A common form of ansatz state – particularly in relation to the electronic structure problem – is

$$|\psi_{\text{anz}}(\boldsymbol{\theta})\rangle = e^{iA(\boldsymbol{\theta})} |\psi_{\text{ref}}\rangle, \quad (5)$$

where $|\psi_{\text{ref}}\rangle \in \mathcal{H}$ is some fixed reference state in which the quantum circuit is initialized and $A(\boldsymbol{\theta}) = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} \theta_{\boldsymbol{\sigma}} \boldsymbol{\sigma}$ for parameters $\theta_{\boldsymbol{\sigma}} \in \mathbb{R}$ and Pauli operators $\boldsymbol{\sigma} \in \mathcal{A} \subset \mathcal{P}_N$; the unitary $e^{iA(\boldsymbol{\theta})}$

effects excitations above the reference state. Such ansätze as unitary coupled cluster (UCC) [13, 14] may be expressed by our choice of A (taking as reference the Hartree-Fock state), in addition to any others based on the theory of excitation operators such as ADAPT-VQE [16–19]. The quantum advantage in VQE stems from the ability to prepare classically intractable states from our parametrized ansatz circuits.

3 Projections onto stabilizer subspaces

Given an operator $\boldsymbol{\sigma} \in \mathcal{P}_N$, the space of quantum states $|\psi\rangle \in \mathcal{H}$ that it stabilizes are those satisfying $\boldsymbol{\sigma} |\psi\rangle = |\psi\rangle$, the $+1$ -eigenspace of $\boldsymbol{\sigma}$. Extending this notion to an abelian subgroup of Pauli operators $\mathcal{Q} \subset \mathcal{P}_N$, there is an induced vector space $\mathcal{V}_{\mathcal{Q}}$ of states stabilized by the elements of $\mathcal{Q}$.

A particularly useful definition is that of a Hamiltonian *symmetry*, taken here to mean a set $\mathcal{S} \subset \mathcal{P}_N$ of Pauli operators such that

$$[\boldsymbol{\sigma}, \boldsymbol{\sigma}'] = 0 \quad \forall \boldsymbol{\sigma} \in \mathcal{S}, \boldsymbol{\sigma}' \in \mathcal{T}. \quad (6)$$

In other words, a symmetry of $H_{\mathcal{T}}$ is any set of Pauli operators that commute universally among $\mathcal{T}$, which we may extend to an abelian group $\overline{\mathcal{S}} := \langle \mathcal{S} \rangle$ generated by $\mathcal{S}$ under operator multiplication, which we shall call a *symmetry group*.

Note the setting in which we present symmetries here is stricter than the conventional definition, which considers any operator S that commutes with the Hamiltonian, i.e. $[S, H_{\mathcal{T}}] = 0$, to be a symmetry. Such an operator need not commute with the individual terms as we require here. For example, in the fermionic picture, the number operator $\sum_i a_i^\dagger a_i$ (where a is the fermionic annihilation operator and its Hermitian conjugate $a^\dagger$ represents the creation operator) commutes with the full second-quantized molecular Hamiltonian, but not with an arbitrary excitation term.

The operators of $\mathcal{S}$ will in general be algebraically dependent, but the theory of stabilizers [48] ensures the existence of a set of independent generators $\mathcal{G}$ such that $\overline{\mathcal{S}} = \langle \mathcal{G} \rangle$. Now, recall the Clifford group consists of unitary operators $U \in \mathcal{B}(\mathcal{H})$ (meaning $UU^\dagger = U^\dagger U = \mathbb{1}$) with the property $U\boldsymbol{\sigma}U^\dagger \in \mathcal{P}_N \quad \forall \boldsymbol{\sigma} \in \mathcal{P}_N$, i.e., U *normalizes* the Pauli group. We may construct a Clifford operation U mapping each symmetry generator to distinct single-qubit Pauli operators σ_p ,

where we are free to choose $p \in \{1, 2, 3\}$. More precisely, there exists a subset of qubit positions $\mathcal{I}_{\text{stab}} \subset \mathbb{Z}_N$ satisfying $|\mathcal{I}_{\text{stab}}| = |\mathcal{G}|$ and a bijective map $f : \mathcal{G} \rightarrow \mathcal{I}_{\text{stab}}$ such that

$$UGU^\dagger = \sigma_p^{(f(G))} \quad \forall G \in \mathcal{G}. \quad (7)$$

This is a powerful concept that provides a mechanism for reducing the number of qubits in the Hamiltonian whilst preserving its energy spectrum. This is at the core of qubit tapering [46, 47], in which it is observed

$$[G, H_{\mathcal{T}}] = 0 \implies [\sigma_p^{(f(G))}, H'_{\mathcal{T}}] = 0 \quad \forall G \in \mathcal{G}, \quad (8)$$

implying the rotated Hamiltonian $H'_{\mathcal{T}} := UH_{\mathcal{T}}U^\dagger$ consists solely of identity or Pauli σ_p operators in the qubit positions indexed by $\mathcal{I}_{\text{stab}}$. Taking expectation values, one may replace the qubits $\mathcal{I}_{\text{stab}}$ by their eigenvalues $\nu_i = \pm 1$; each assignment

$$\boldsymbol{\nu} = (\nu_i)_{i \in \mathcal{I}_{\text{stab}}} \in \{\pm 1\}^{\times \mathcal{I}_{\text{stab}}} \quad (9)$$

defines a symmetry *sector* and at least one such sector will contain the true solution to the eigenvalue problem. Note the other sectors still have physical significance and may for example relate to solutions of different particle numbers or to excited states. In Tables 3 - 10 below we report the symmetry generators and corresponding sector for the Hamiltonians representing the molecular systems listed in Table 1.

A quantum state consistent with any such sector must be stabilized by the operators $\nu_i \sigma_p^{(i)}$ and we may define a projection onto the corresponding stabilizer subspace. In general, a projection is defined to be an *idempotent* operator $P \in \mathcal{B}(\mathcal{H})$, i.e. $P^2 = P$; the projection onto the ± 1 -eigenspace of a single-qubit Pauli operator σ_p for $p \in \{1, 2, 3\}$ may be written

$$P_p^\pm := \frac{1}{2}(I \pm \sigma_p). \quad (10)$$

States with no component inside the chosen eigenspace are mapped to zero and observe that

$$P_p^\pm \sigma_q P_p^\pm = \pm \delta_{p,q} P_p^\pm \quad (11)$$

for $q \in \{1, 2, 3\}$.

Let $\mathcal{H}_{\text{stab}}$ be the reduced Hilbert space supported by the stabilized qubits $\mathcal{I}_{\text{stab}}$ and $\mathcal{H}_{\text{red}}$ its complement such that $\mathcal{H} = \mathcal{H}_{\text{stab}} \otimes \mathcal{H}_{\text{red}}$. Given

an assignment of eigenvalues $\boldsymbol{\nu} \in \{\pm 1\}^{\times \mathcal{I}_{\text{stab}}}$, we may project onto the corresponding sector via

$$P_{\boldsymbol{\nu}} := \bigotimes_{i \in \mathcal{I}_{\text{stab}}} P_p^{\nu_i} \quad (12)$$

and subsequently perform a *partial trace* over the stabilized qubits $\mathcal{I}_{\text{stab}}$. This is effected by the unique linear map $\text{Tr}_{\text{stab}} : \mathcal{H} \rightarrow \mathcal{H}_{\text{red}}$ satisfying the property $\text{Tr}_{\text{stab}}(A \otimes B) = \text{Tr}(A)B$ for all $A \in \mathcal{B}(\mathcal{H}_{\text{stab}})$ and $B \in \mathcal{B}(\mathcal{H}_{\text{red}})$.

Finally, we may define the full stabilizer subspace projection map

$$\begin{aligned} \pi_{\boldsymbol{\nu}}^U : \mathcal{B}(\mathcal{H}) &\rightarrow \mathcal{B}(\mathcal{H}_{\text{red}}); \\ A &\mapsto \text{Tr}_{\text{stab}}(P_{\boldsymbol{\nu}} U A U^\dagger P_{\boldsymbol{\nu}}) \end{aligned} \quad (13)$$

which, using the linearity of Tr_{stab} , yields a reduced Hamiltonian

$$\begin{aligned} H_{\mathcal{T}}^{\text{red}} &:= \pi_{\boldsymbol{\nu}}^U(H_{\mathcal{T}}) \\ &= \sum_{\boldsymbol{\sigma} \in \mathcal{T}} h'_{\boldsymbol{\sigma}} \boldsymbol{\sigma}'_{\text{red}} \end{aligned} \quad (14)$$

where $\boldsymbol{\sigma}' = U \boldsymbol{\sigma} U^\dagger = \bigotimes_{i=0}^{N-1} \sigma_{q_i}$ and we have written $\boldsymbol{\sigma}' = \boldsymbol{\sigma}'_{\text{stab}} \otimes \boldsymbol{\sigma}'_{\text{red}}$. The new coefficients $h'_{\boldsymbol{\sigma}} := h_{\boldsymbol{\sigma}} \prod_{\substack{i \in \mathcal{I}_{\text{stab}} \\ q_i \neq 0}} \nu_i$ differ from $h_{\boldsymbol{\sigma}}$ by a sign dependent on the chosen symmetry sector.

In qubit tapering U is taken as (7), with the corresponding basis $\mathcal{G}$ a generating set for a full Hamiltonian symmetry [46, 47]. Assuming identification of the correct sector, the ground state energy of the $(N - |\mathcal{G}|)$ -qubit reduced Hamiltonian $H_{\mathcal{T}}^{\text{red}}$ will coincide with the true value of the full system $H_{\mathcal{T}}$.

This stabilizer projection procedure is straightforward with respect to the Hamiltonian, since the stabilized qubits contain only operators with non-zero image under conjugation with $P_{\boldsymbol{\nu}}$. However, suppose we were to take another observable $A \in \mathcal{B}(\mathcal{H})$ and wish to determine a reduced form on $\mathcal{B}(\mathcal{H}_{\text{red}})$ that is consistent with the reduced Hamiltonian $H_{\mathcal{T}}^{\text{red}}$. This may be achieved by following precisely the same process that was applied to $H_{\mathcal{T}}$, but the symmetry $\mathcal{S}$ will not in general be a symmetry of A and therefore the ‘symmetry-breaking’ terms (those which anticommute with the generators $\mathcal{G}$) will vanish under projection onto the stabilizer subspace, as per (11). Letting $\mathcal{A} \subset \mathcal{P}_N$ be the set of terms in

the Pauli-basis expansion of A , observe:

$$\begin{aligned}
A^{\text{red}} &:= \pi_{\nu}^U(A) \\
&= \sum_{\sigma \in \mathcal{A}} h_{\sigma} \text{Tr}(P_{\nu} \sigma'_{\text{stab}} P_{\nu}) \sigma'_{\text{red}} \\
&= \sum_{\sigma \in \mathcal{A}} h_{\sigma} \text{Tr} \left(\bigotimes_{\substack{i \in \mathcal{I}_{\text{stab}} \\ q_i=0}} P_p^{\nu_i} \otimes \right. \\
&\quad \left. \bigotimes_{\substack{i \in \mathcal{I}_{\text{stab}} \\ q_i \neq 0}} \underbrace{P_p^{\nu_i} \sigma_{q_i} P_p^{\nu_i}}_{=\nu_i \delta_{p,q_i} P_p^{\nu_i}} \right) \sigma'_{\text{red}} \\
&= \sum_{\sigma \in \mathcal{A}} h_{\sigma} \sigma'_{\text{red}} \underbrace{\text{Tr}(P_{\nu})}_{=1} \prod_{\substack{i \in \mathcal{I}_{\text{stab}} \\ q_i \neq 0}} \nu_i \delta_{p,q_i} \\
&= \sum_{\substack{\sigma \in \mathcal{A} \\ q_i \in \{0,p\} \\ \forall i \in \mathcal{I}_{\text{stab}}}} h'_{\sigma} \sigma'_{\text{red}}, \tag{15}
\end{aligned}$$

recalling that q_i indicates the type of single-qubit Pauli acting on qubit position $i \in \mathbb{Z}_N$ in some tensor product σ , defined in Section 2.

The resulting form is identical to (14), except we are explicit that the terms surviving projection are only those whose qubit positions indexed by $\mathcal{I}_{\text{stab}}$ consist exclusively of identity and Pauli σ_p operators; this is trivially true for the Hamiltonian by construction. Most importantly, this extends the stabilizer subspace projection to ansätze defined on the full system for use in variational algorithms. It should be noted that the above operations are classically tractable and can be implemented efficiently in the symplectic representation of Pauli operators.

We would be remiss not to draw attention to the likeness of (13) with Positive Operator-Valued Measures (POVM) [49]; indeed, the projectors (12) define a complete set of *Kraus* operators [50]. The stabilizer subspace projection procedure is reduced to a matter of enforcing a partial measurement over some subsystem of the full problem, for which the relevant outcomes have been determined via an auxiliary method. For example, this could involve identifying a quantum state with a known non-zero overlap with the true ground state; measuring the symmetry generators $\mathcal{G}$ in this state will yield the correct sector.

Hartree-Fock often provides such a state for electronic structure problems, although it is not immune to failure – particularly in the strongly-correlated regime. In these cases, we should defer to more effective reference states such as those ob-

tained from the Møller–Plesset perturbation theory (MP2), coupled-cluster (CC) and so on. One can imagine a hierarchy of increasingly precise ground state approximations, for which we should hope to obtain at some point a non-zero overlap with the true ground state.

4 CS-VQE in the stabilizer formalism

We now describe the *Contextual Subspace VQE* (CS-VQE) method in the stabilizer setting introduced in Section 3. CS-VQE partitions the Hamiltonian (2) into two disjoint components – one that is noncontextual and another that is contextual which provides quantum corrections to the former via VQE [44]. Explicitly, this allows us to write

$$H_{\mathcal{T}} = H_{\mathcal{T}_{\text{nc}}} + H_{\mathcal{T}_{\text{c}}} \tag{16}$$

where $\mathcal{T}_{\text{nc}}$ is a noncontextual set of Pauli operators and $\mathcal{T}_{\text{c}} := \mathcal{T} \setminus \mathcal{T}_{\text{nc}}$ is what remains, which will in general be contextual.

CS-VQE differs from qubit tapering (described in section 3) in the following way: the latter exploits existing (i.e. physical) symmetries of the Hamiltonian, whereas in CS-VQE we impose additional ‘pseudo-symmetries’ derived from the noncontextual Hamiltonian. This results in a loss of information, since any terms of $\mathcal{T}$ not commuting with the symmetry generators will vanish under projection.

4.1 The noncontextual problem

The notion of contextuality goes back to the Bell-Kochen-Specker theorem [51–53]. Here we use an explicit condition for the noncontextuality of a set of Pauli operators, developed in [54] and independently in [55]. Strictly speaking, this condition tests for strong measurement contextuality. In this setting, a set $\mathcal{T}_{\text{nc}}$ is understood to be noncontextual if and only if commutation forms an equivalence relation on $\mathcal{T}_{\text{nc}} \setminus \mathcal{S}$, where we have defined the sub-Hamiltonian symmetry $\mathcal{S} := \{\sigma \in \mathcal{T}_{\text{nc}} \mid [\sigma, \sigma'] = 0 \ \forall \sigma' \in \mathcal{T}_{\text{nc}}\}$. There is an implied structure

$$\mathcal{T}_{\text{nc}} = \mathcal{S} \cup \mathcal{C}_1 \cup \dots \cup \mathcal{C}_M, \tag{17}$$

where the $\mathcal{C}_i$ are equivalence classes with respect to commutation – in other words, elements of the same class commute and across classes anticommute. Conversely, such a set of Pauli operators

is contextual if and only if commutation fails to be transitive on $\mathcal{T}_{\text{nc}} \setminus \mathcal{S}$.

The symmetry $\mathcal{S}$ can be expanded by taking pairwise products within equivalence classes – since $\{C_i, C_j\} = 0$ for $C_i \in \mathcal{C}_i, C_j \in \mathcal{C}_j$ with $i \neq j$, it is the case that $[C_i C'_i, C_j C'_j] = 0$ and we may define $\mathcal{S}' = \mathcal{S} \cup \bigcup_{i=1}^M \{C_i C'_i \mid C_i, C'_i \in \mathcal{C}_i\}$. As before, in Section 3, $\mathcal{S}'$ induces a symmetry group for which one may define independent generators $\mathcal{G}$ and a Clifford operation $U_{\mathcal{G}}$ mapping the generators to single-qubit Pauli operators; the expectation value over these qubits will again be determined by an assignment $\boldsymbol{\nu} \in \{\pm 1\}^{\times |\mathcal{G}|}$ of eigenvalues, analogous to the selection of a symmetry sector in qubit tapering.

From each class $\mathcal{C}_i$ we select a representative C_i and construct an observable $C(\mathbf{r}) := \sum_{i=1}^M r_i C_i$ where $\mathbf{r} \in \mathbb{R}^M$ and $|\mathbf{r}| = 1$. In [56] it was found that quantum states $|\psi_{(\boldsymbol{\nu}, \mathbf{r})}\rangle \in \mathcal{H}$ stabilized by the operators $\{\nu_{f(G)} G \mid G \in \mathcal{G}\} \cup \{C(\mathbf{r})\}$ are consistent with a classical objective function $\eta(\boldsymbol{\nu}, \mathbf{r})$ (derived in Appendix B), in the sense that $\eta(\boldsymbol{\nu}, \mathbf{r})$ coincides with the noncontextual energy expectation value $\langle H_{\mathcal{T}_{\text{nc}}} \rangle_{\psi_{(\boldsymbol{\nu}, \mathbf{r})}}$ for all parametrizations $(\boldsymbol{\nu}, \mathbf{r})$. This is a consequence of the joint probability distribution (41) chosen over the phase-space points of their (epistricted) model [56, 57].

The noncontextual energy spectrum is therefore parametrized by two vectors: the ± 1 eigenvalue assignments $\boldsymbol{\nu}$, determining the contribution of the universally commuting terms, and $\mathbf{r}$, encapsulating the remaining pairwise anticommuting classes. In this sense, we may refer to $(\boldsymbol{\nu}, \mathbf{r})$ as a state of the noncontextual Hamiltonian itself, abstracted from quantum states of the corresponding stabilizer subspace. Optimizing over these parameters, we obtain the noncontextual ground state energy

$$\epsilon_0^{\text{nc}} := \min_{\substack{\boldsymbol{\nu} \in \{\pm 1\}^{\times |\mathcal{G}|} \\ \mathbf{r} \in \mathbb{R}^M : |\mathbf{r}| = 1}} \eta(\boldsymbol{\nu}, \mathbf{r}) \quad (18)$$

and call an element $(\boldsymbol{\nu}, \mathbf{r})$ of the preimage $\eta^{-1}(\epsilon_0^{\text{nc}})$ a noncontextual ground state of $H_{\mathcal{T}_{\text{nc}}}$. Denote by $\Delta_{\text{nc}} := |\epsilon_0^{\text{nc}} - \epsilon_0|$ the absolute error with respect to the true ground state energy.

As a classical estimate to the ground state energy of the full Hamiltonian $H_{\mathcal{T}}$, in Section 5 we found the difference between the noncontextual ground state and Hartree-Fock energy to be negligible for each of the molecules simulated, since

the heuristic used to choose $H_{\mathcal{T}_{\text{nc}}}$ prioritizes diagonal Hamiltonian terms. In principle, it may be an improvement upon Hartree-Fock as the non-contextual set can also take into account an off-diagonal contribution within the anticommuting classes. This is highly dependent on the chosen form of noncontextual set and the effect on CS-VQE errors is not clear – some consideration of the resulting quantum corrections must inform this choice. We will discuss how one derives quantum corrections in the following section.

4.2 Quantum corrections

Our simulation approach has thus far been strictly classical – now we arrive at the quantum element of CS-VQE. We have derived a classical estimate of the ground state energy from the non-contextual part of the Hamiltonian $H_{\mathcal{T}_{\text{nc}}}$, however the contextual component $H_{\mathcal{T}_{\text{c}}}$ has so far been neglected.

While $C(\mathbf{r})$ is not a stabilizer in the strict sense (it is not an element of the Pauli group), it is unitarily equivalent to one as a linear combination of anticommuting Pauli elements. Similar to the symmetry generators $\mathcal{G}$, it is possible to define a unitary operation U_C mapping $C(\mathbf{r})$ onto a single-qubit Pauli operator, following the approach of unitary partitioning [32, 33, 35]. However, unlike the $U_{\mathcal{G}}$ rotation, U_C is not Clifford as it collapses M terms onto a single Pauli operator and can therefore introduce additional terms to the Hamiltonian. Appending $C(\mathbf{r})$ to our set of generators $\tilde{\mathcal{G}} := \mathcal{G} \cup \{C(\mathbf{r})\}$ and defining $U := U_C U_{\mathcal{G}}$, there exists a subset of qubit indices $\mathcal{I}_{\text{stab}}$ satisfying $|\mathcal{I}_{\text{stab}}| = |\tilde{\mathcal{G}}|$ and a bijective map $f : \tilde{\mathcal{G}} \rightarrow \mathcal{I}_{\text{stab}}$ such that $UGU^\dagger = \sigma_p^{(f(G))}$ for each $G \in \tilde{\mathcal{G}}$. We reiterate that $p \in \{1, 2, 3\}$ may be chosen at will – the approach taken in [44] is to select $p = 3$ to enforce diagonal generators.

Suppose we have a quantum state $|\psi_{(\boldsymbol{\nu}, \mathbf{r})}\rangle$ that is consistent with ϵ_0^{nc} ; since the rotated state $|\psi'_{(\boldsymbol{\nu}, \mathbf{r})}\rangle = U |\psi_{(\boldsymbol{\nu}, \mathbf{r})}\rangle$ must be stabilized by $\sigma_p^{(i)} \forall i \in \mathcal{I}_{\text{stab}}$, the qubit positions $\mathcal{I}_{\text{stab}}$ must be fixed. This implies a decomposition

$$|\psi'_{(\boldsymbol{\nu}, \mathbf{r})}\rangle = |b_{(\boldsymbol{\nu}, \mathbf{r})}\rangle_{\text{stab}} \otimes |\varphi\rangle_{\text{red}} \quad (19)$$

where $|b_{(\boldsymbol{\nu}, \mathbf{r})}\rangle$ represents a single basis state of $\mathcal{H}_{\text{stab}}$ and $|\varphi\rangle \in \mathcal{H}_{\text{red}}$ is independent of the parameters $(\boldsymbol{\nu}, \mathbf{r})$. Therefore, the expectation value

of the full Hamiltonian may be expressed as

$$\langle H_{\mathcal{T}} \rangle_{\psi_{(\nu, r)}} = \epsilon_0^{\text{nc}} + \langle \pi_{\nu}^U(H_{\mathcal{T}_c}) \rangle_{\varphi}, \quad (20)$$

where $\pi_{\nu}^U(H_{\mathcal{T}_c})$ contains only the terms of the contextual Hamiltonian that commute with all the noncontextual generators, just as in (15) – it was observed in [44] that any term which anticommutes with at least one noncontextual generator must have zero expectation value and our stabilizer subspace projection captures this fact.

Inspecting (20), we may optimize freely over quantum states φ , i.e., we are not constrained by the noncontextual ground state within $\mathcal{H}_{\text{red}}$. In fact, we may absorb the noncontextual ground state energy into the reduced contextual Hamiltonian, defining the *contextual subspace Hamiltonian*

$$\tilde{H}_{\mathcal{T}_c} := \epsilon_0^{\text{nc}} \cdot \mathbb{1} + \pi_{\nu}^U(H_{\mathcal{T}_c}); \quad (21)$$

this form is obtained naturally when applying the stabilizer subspace projection to the full Hamiltonian, which automatically includes the noncontextual energy by fixing the corresponding eigenvalue assignments.

Now, we may perform unconstrained VQE to obtain a quantum-corrected estimate

$$\epsilon_0^c := \min_{|\varphi\rangle \in \mathcal{H}_{\text{red}}} \langle \tilde{H}_{\mathcal{T}_c} \rangle_{\varphi} \quad (22)$$

of the true ground state energy with absolute error $\Delta_c := |\epsilon_0^c - \epsilon_0| \leq \Delta_{\text{nc}}$. We have equality when the stabilizers span every qubit position, which is the case when $|\tilde{\mathcal{G}}| = N$ since the generators must be algebraically independent – this means the initial quantum correction is trivial as the noncontextual part determines the entire system.

For instances of the electronic structure problem there is no guarantee that ϵ_0^c will achieve chemical accuracy ($\Delta_c < 1.6\text{mHa} \approx 4\text{kJ/mol}$) and, indeed, it might not improve upon the noncontextual estimate (although it will never be worse, due to the variational principle applying in this case). However, one can easily define a subset of $\mathcal{T}_{\text{nc}}$ that is again noncontextual – this is achieved by discarding one of the noncontextual generators $G \in \tilde{\mathcal{G}}$, along with the operators that it generates. We now append the discarded operators to the contextual Hamiltonian, relaxing the stabilizer constraint on the qubit position $f(G)$ and permitting a search over its Hilbert space. This process may be iterated until the noncontextual set is exhausted and we recover full VQE.

This means that, unless the ground state energy of $H_{\mathcal{T}_{\text{nc}}}$ and H coincides, CS-VQE will improve upon the noncontextual energy using less quantum resource than full VQE – this is more rigorously defined in the next section.

What we have described here is a technique of scaling the relative sizes of the noncontextual (read classical) and contextual (read quantum) simulations in a reciprocal manner. We can therefore trade-off quantum and classical workloads in CS-VQE.

4.3 Expanding the contextual subspace

Now we describe the process of growing the contextual subspace more rigorously. We select a subset of noncontextual generators $\mathcal{F} \subset \tilde{\mathcal{G}}$ whose stabilizer constraints we mean to enforce and construct a new noncontextual set $\mathcal{T}'_{\text{nc}} := \mathcal{T}_{\text{nc}} \cap \overline{\mathcal{F}}$; the contextual set is expanded accordingly by appending the terms not generated by $\mathcal{F}$, i.e., $\mathcal{T}'_c := \mathcal{T}_c \cup (\mathcal{T}_{\text{nc}} \setminus \overline{\mathcal{F}})$. As before, there exists a unitary operation $U_{\mathcal{F}}$, a subset of qubit indices $\mathcal{I}_{\text{fix}} \subset \mathcal{I}_{\text{stab}}$ and a bijective map $f : \mathcal{F} \rightarrow \mathcal{I}_{\text{fix}}$ satisfying $U_{\mathcal{F}} G U_{\mathcal{F}}^{\dagger} = \sigma_p^{(f(G))} \quad \forall G \in \mathcal{F}$ (the rotation $U_{\mathcal{F}}$ may or may not be Clifford depending on whether $C(\mathbf{r})$ is among the stabilizers we wish to fix).

Denote by $\epsilon_0^{\text{nc}}(\mathcal{F})$ the ground state energy of the new noncontextual Hamiltonian $\mathcal{T}'_{\text{nc}}$ with absolute error $\Delta_{\text{nc}}(\mathcal{F}) \geq \Delta_{\text{nc}}$. While this is weaker as an estimate of the true ground state energy of the full system, at the very least we are guaranteed to recover the initial noncontextual ground state energy from performing a simulation of the expanded contextual subspace [44], which we describe below.

The stabilizer constraints of $\mathcal{F}$ are enforced over the Hilbert space $\mathcal{H}_{\text{fix}} = (\mathbb{C}^2)^{\otimes \mathcal{I}_{\text{fix}}}$ of qubits indexed by $\mathcal{I}_{\text{fix}}$, whereas we are free to perform a VQE simulation over $\mathcal{H}_{\text{sim}} = (\mathbb{C}^2)^{\otimes \mathcal{I}_{\text{sim}}}$, the Hilbert space of the remaining $N - |\mathcal{F}|$ qubits indexed by $\mathcal{I}_{\text{sim}} = \mathbb{Z}_N \setminus \mathcal{I}_{\text{fix}}$. Invoking the stabilizer subspace projection map $\pi_{\nu'}^{U_{\mathcal{F}}}$ with the eigenvalue assignments $\nu' = (\nu_i)_{i \in \mathcal{I}_{\text{fix}}}$ yields an expanded contextual subspace Hamiltonian

$$\tilde{H}_{\mathcal{T}'_c} := \epsilon_0^{\text{nc}}(\mathcal{F}) \cdot \mathbb{1} + \pi_{\nu'}^{U_{\mathcal{F}}}(H_{\mathcal{T}'_c}). \quad (23)$$

Performing an $|\mathcal{I}_{\text{sim}}|$ -qubit VQE simulation over the contextual subspace we obtain a new

quantum-corrected estimate

$$\epsilon_0^c(\mathcal{F}) := \min_{|\varphi\rangle \in \mathcal{H}_{\text{sim}}} \langle \tilde{H}_{\mathcal{T}_c'} \rangle_\varphi \quad (24)$$

with an error satisfying $\Delta_c(\mathcal{F}) \leq \Delta_c$. Recall that $\Delta_c = \Delta_c(\tilde{\mathcal{G}})$ corresponds with the contextual error when we enforce the full set of noncontextual stabilizers.

Observe that, when $|\mathcal{I}_{\text{sim}}| = N$, we are simply performing full VQE over the entire system – this occurs when we do not enforce the stabilizer constraint for any of the noncontextual generators, i.e. $\mathcal{F} = \emptyset$. Therefore, it must be the case that

$$\Delta_c(\mathcal{F}) \rightarrow 0 \text{ as } |\mathcal{F}| \rightarrow 0. \quad (25)$$

Furthermore, given a nested sequence of generator subsets $(\mathcal{F}_i)_i$ with $\mathcal{F}_{i+1} \subset \mathcal{F}_i$, then $\Delta_c(\mathcal{F}_{i+1}) \leq \Delta_c(\mathcal{F}_i)$ and the convergence is monotonic; this was illustrated empirically in [44] for a suite of tapered test molecules of up to 18 qubits.

Suppose we wish to construct the optimal contextual subspace Hamiltonian of size $N' < N$. The problem reduces to minimizing the error $\Delta_c(\mathcal{F})$ over the $\binom{|\tilde{\mathcal{G}}|}{N-N'}$ generator subsets $\mathcal{F} \subset \tilde{\mathcal{G}}$ satisfying $|\mathcal{F}| = N - N'$. CS-VQE is highly sensitive to this choice and remains a vital open question for the continued success of the technique. For chemistry applications, we grow the contextual subspace until the CS-VQE error attains chemical accuracy, which means finding the minimal $\mathcal{F}$ such that $\Delta_c(\mathcal{F}) < 1.6\text{mHa}$. In general, we will not have access to a target energy and so will not necessarily know when the desired precision is achieved; instead, we might iterate until the VQE convergence is within some fixed bound.

The best stabilizer relaxation ordering heuristic found in [44] functions by greedily selecting combinations of $d \leq N$ generators that produce the greatest reduction in error, which necessitates $\sum_{k=0}^{\lfloor N/d \rfloor} \binom{N-dk}{d} = \mathcal{O}(N^{d+1})$ CS-VQE simulations; $d = 2$ was chosen in that work. However, there is room for more targeted approaches that exploit some structure of the underlying problem. For example, in quantum chemistry problems it could be that one should relax the stabilizers that have non-trivial action near the Fermi level – between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO). Excitations clustered around this gap are more likely to appear in the true ground

state and should therefore not be assigned definite values under the noncontextual projection. This idea comes from the theory of pseudopotential approximations [58], in which it is observed that chemically relevant electrons are predominantly those of the valence space, whereas the core may be ‘frozen’, thus reducing the electronic complexity.

4.4 The noncontextual projection ansatz

CS-VQE has thus far not been applied to systems exceeding 18 qubits and the resulting reduced Hamiltonians (23) have been solved by direct diagonalization [44] – clearly, this will not scale to larger systems, with the required classical memory increasing exponentially. Instead, they must be simulated by performing VQE routines, but defining an ansatz for the contextual subspace provided an obstacle to achieving this in practice.

However, having now placed the problem within the stabilizer formalism described in Section 3, we have already introduced (in Sections 4.1 - 4.3) the tools necessary to restrict an ansatz of the form (5) – defined over the full system – to the contextual subspace (23). The approach adopted here is equivalent to that which we defined for qubit tapering in (15). To restrict a parametrized ansatz operator

$$A(\boldsymbol{\theta}) = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} \theta_{\boldsymbol{\sigma}} \boldsymbol{\sigma} \mapsto \tilde{A}(\boldsymbol{\theta}) \in \mathcal{B}(\mathcal{H}_{\text{sim}}) \quad (26)$$

in line with the stabilizer constraints $\mathcal{F} \subset \tilde{\mathcal{G}}$ we may simply call upon the stabilizer subspace projection map $\pi_{\boldsymbol{\nu}'}^{U_{\mathcal{F}}}$ once more, which yields a restricted ansatz state

$$|\tilde{\psi}_{\text{anz}}(\boldsymbol{\theta})\rangle := e^{i\tilde{A}(\boldsymbol{\theta})} |\tilde{\psi}_{\text{ref}}\rangle \in \mathcal{H}_{\text{sim}} \quad (27)$$

where

$$\tilde{A}(\boldsymbol{\theta}) := \pi_{\boldsymbol{\nu}'}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta})). \quad (28)$$

Any rotated ansatz term $U_{\mathcal{F}} \boldsymbol{\sigma} U_{\mathcal{F}}^\dagger$ that is not identity or a Pauli σ_p on some subset of the qubit positions indexed by $\mathcal{I}_{\text{fix}}$ will vanish.

The restricted reference state $|\tilde{\psi}_{\text{ref}}\rangle$ is obtained from a partial projective measurement of $U |\psi_{\text{ref}}\rangle$ (see the discussion on POVMs in Section 3) with outcomes defined by $\boldsymbol{\nu}'$, which yields a product state

$$\frac{P_{\boldsymbol{\nu}'} U |\psi_{\text{ref}}\rangle}{\sqrt{\langle \psi_{\text{ref}} | U^\dagger P_{\boldsymbol{\nu}'} U | \psi_{\text{ref}} \rangle}} = |b_{(\boldsymbol{\nu}', r)}\rangle_{\text{fix}} \otimes |\tilde{\psi}_{\text{ref}}\rangle_{\text{sim}}. \quad (29)$$

The post-measurement state $|b_{(\nu, \mathbf{r})}\rangle \in \mathcal{H}_{\text{fix}}$ on the noncontextual subspace represents a single basis vector and can therefore be disregarded, leaving just the state of the contextual subspace – this we take as reference for our restricted ansatz. We stress this ‘measurement’ is not performed in circuit but is instead to be evaluated classically when constructing the restricted ansatz circuit.

We may now define the contextual subspace energy expectation function

$$\tilde{E}(\boldsymbol{\theta}) := \langle \tilde{H}_{\mathcal{T}_c} \rangle_{\tilde{\psi}_{\text{anz}}(\boldsymbol{\theta})} \quad (30)$$

with $\tilde{H}_{\mathcal{T}_c}$ as in (23), at which point we have reduced the problem to standard VQE, performed over a subspace of the full problem.

In order to prepare the projected ansatz state (27), we first initialize the $|\mathcal{I}_{\text{sim}}|$ -qubit quantum circuit in the noncontextual ground state, achieved by applying a Pauli σ_1 operator in each of the qubit positions $i \in \mathcal{I}_{\text{sim}}$ such that $\nu_i = -1$. This is visible in Figure 2, in which the VQE routine is initiated with the optimization parameters zeroed, i.e. $\boldsymbol{\theta} = \mathbf{0}$, and since $e^{i\tilde{A}(\mathbf{0})} = \mathbb{1}$ optimization begins at the noncontextual ground state energy.

It is not in general possible to implement the unitary operation $e^{i\tilde{A}(\boldsymbol{\theta})}$ exactly as a quantum circuit (except for in the case of completely commuting terms $\mathcal{A}$ of $A(\mathbf{r})$), however one may do so approximately via the commonly used technique of Trotterization (see Appendix A.1 for further details). With the Trotter number $n_T = 2$, this yields

$$e^{i\tilde{A}(\boldsymbol{\theta})} \approx \left(\prod_{\sigma \in \tilde{\mathcal{A}}} e^{i\frac{\theta_\sigma}{2} \sigma} \right)^2 \quad (31)$$

where the exponentiated Pauli operator $e^{i\frac{\theta_\sigma}{2} \sigma}$ may be performed in-circuit as described in Appendix A.2 using $\mathcal{O}(|\mathcal{I}_{\text{sim}}|)$ native quantum gates.

5 Simulation results

The molecular systems that were simulated to benchmark the noncontextual projection ansatz for CS-VQE are given in Table 1. The molecule geometries were obtained from the Computational Chemistry Comparison and Benchmark Database (CCCBDB) [59] and their Hamiltonians constructed using IBM’s Qiskit Nature [60] with PySCF the underlying quantum chemistry package [61].

Molecular systems			Number of qubits		
Name	Charge	Mult.	Full	Taper	CS-VQE
Be	0	1	10	5	3
B	0	2	10	5	3
LiH	0	1	12	8	4
BeH	+1	1	12	8	6
HF	0	1	12	8	4
BeH ₂	0	1	14	9	7
H ₂ O	0	1	14	10	7
F ₂	0	1	20	16	10
HCl	0	1	20	17	4

Table 1: The systems investigated to benchmark the noncontextual projection ansatz (all in the STO-3G basis). The CS-VQE column indicates the fewest number of qubits required to achieve chemical accuracy.

Before we evaluate the efficacy of our noncontextual projection ansatz, there are a few features of (27) that should be highlighted. First of all, in (29) we are applying the operation U in-circuit, introducing further gates that will contribute additional noise. However, when the reference state is taken to be that of Hartree-Fock, we observed $U\psi_{\text{ref}}$ to coincide with the noncontextual ground state. This is an artifact of the noncontextual set construction heuristic – used within both this work and [44] – prioritizing diagonal entries. This need not always be the case, but for the molecular systems investigated this allows us to avoid performing U in-circuit and instead take the noncontextual ground state as our reference.

Secondly, application of the unitary partitioning rotations U_C to the ansatz operator $A(\boldsymbol{\theta})$ may introduce additional terms. The scaling was first proposed to be $\mathcal{O}(2^M)$ in [44], where M is the number of equivalence classes in (17). We obtained $M = 2$ for all of the molecules tested, although for a general Hamiltonian this need not be the case and is also dependent on the form of the noncontextual set $\mathcal{T}_{\text{nc}}$ – here we prioritize universally commuting terms, but it is equally valid to maximize the anticommuting contribution.

Despite this, upon the subsequent projection of $A(\boldsymbol{\theta})$, it is possible that a significant number of terms will vanish. This is highly dependent on the quality of the initial ansatz and how heavily it is supported on the stabilized qubit positions $\mathcal{I}_{\text{fix}}$. Figure 1 presents circuit depths of the noncontextual projection ansatz as a proportion

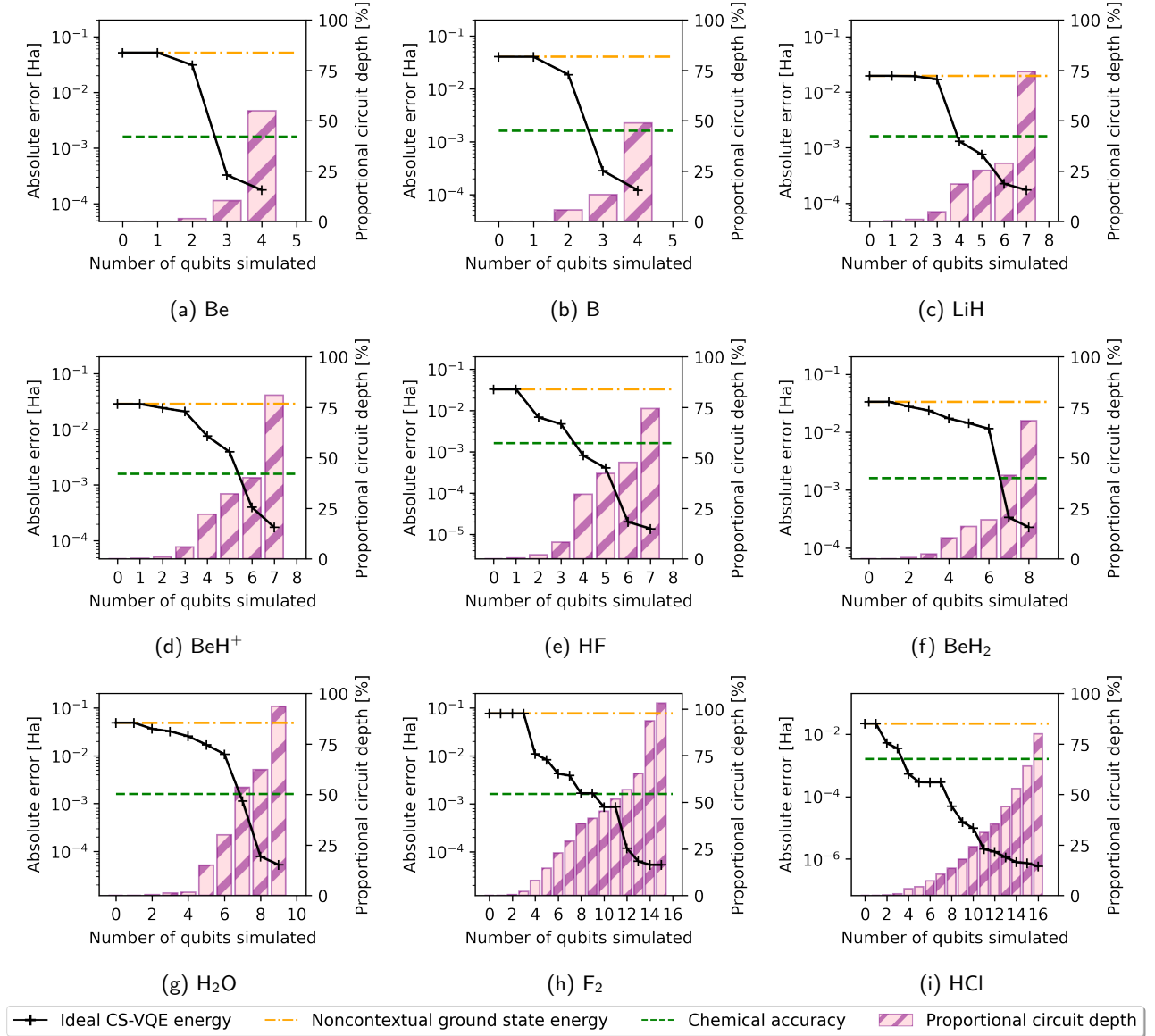


Figure 1: Ideal CS-VQE errors (left-hand axis) and corresponding noncontextual projection ansatz circuit depths as a proportion of the full UCCSD operator from which it is derived (right-hand axis) against the number of qubits simulated.

of the base ansatz from which it is derived – in this case the unitary coupled-cluster singles and doubles (UCCSD) operator. A net reduction in circuit depth is observed, which is quite dramatic up to the point of reaching chemical accuracy in the CS-VQE routine.

On the other hand, if we are able to define an ansatz that closely captures the underlying chemistry, in the sense that each of the excitations it encodes carry significant weight, then we can in principle observe an increase in circuit depth. The latter is more likely to be encountered in implementations of ADAPT-VQE [16–19] and similar approaches to ansatz construction, where ex-

citation terms are selected iteratively in line with some criteria to minimize ansatz redundancy.

For each of the molecules in Table 1, working in the fermionic picture we determined the most significant terms from the pool of single, double and triple (SDT) excitations via brute force, similar to the approach taken in [2]. This was achieved by diagonalizing the untapered Hamiltonian to obtain a true ground state, before assessing which of the resulting basis states exceeding a probability threshold of 10^{-t} for $t \in \mathbb{N}$ exhibit SDT excitations above the Hartree-Fock state. This allowed us to describe a reduced set of excitation operators that includes only the most impor-

tant determinants, yielding an artificial ‘ansatz’ operator that we shall refer to as excite_t – we stress that this approach is not scalable since it requires knowledge of the ground state *a priori*. The excite_4 operator was found to be sufficient for achieving chemical accuracy in the molecular systems of Table 1 when taken as an ansatz circuit, which is seen to be significantly smaller than UCCSD(T) in Table 2.

Molecule	Ansatz operator			
	UCCSDT	UCCSD	excite_4	minimal
Be	48	48	12	5
B	48	32	8	2
LiH	704	192	44	5
BeH ⁺	646	166	86	10
HF	92	92	44	4
BeH ₂	1312	224	72	10
H ₂ O	1892	324	168	20
F ₂	176	176	60	7
HCl	348	348	52	4

Table 2: The number of Pauli terms $|\mathcal{A}|$ for a selection of (tapered) ansätze, each of which is sufficiently expressible to achieve chemical accuracy.

The Jordan-Wigner transformation [62] maps a single fermionic annihilation operator onto two Pauli operators

$$a_i \mapsto \frac{1}{2}(\sigma_1^{(i)} + i\sigma_2^{(i)}) \otimes \bigotimes_{j<i} \sigma_3^{(j)}, \quad (32)$$

with the creation operator given by its Hermitian conjugate $a_i^\dagger$. Therefore, an excitation on $s \in \mathbb{N}$ spin orbitals of the form

$$\mathbf{a} = a_{i_1}^\dagger \dots a_{i_s}^\dagger a_{j_1} \dots a_{j_s} \quad (33)$$

is represented by 2^{2s} Pauli operators under this encoding. In the unitary coupled cluster theory, we are interested rather in the operator $\mathbf{a} - \mathbf{a}^\dagger$ to ensure unitarity upon exponentiation – this may be expressed by 2^{2s-1} Pauli terms.

As such, after a mapping onto qubits via the Jordan-Wigner transformation, single, double and triple excitations account for 2, 8 and 32 Pauli operator terms of $A(\boldsymbol{\theta})$ respectively, but not all of these are required to reach chemical accuracy. To determine only the necessary terms for the purposes of benchmarking our noncontextual projection ansatz, we start with a blank

ansatz and loop over $\mathcal{A}$, running a CS-VQE simulation for each term therein. After every cycle we append the Pauli operator that minimizes the energy to our ansatz until chemical accuracy is achieved and the process terminates – the resulting operator is referred to as *minimal* in Table 2. This procedure was conducted on an Atos Quantum Learning Machine (QLM) provided by the Leibniz Supercomputing Centre (LRZ).

The calculated minimal artificial ansatz operators are reported in Tables 3 - 10, along with a description of the smallest CS-VQE problem permitting chemical accuracy. This includes the optimal noncontextual generator subset $\mathcal{F}$, the resulting noncontextual projection ansatz (27), restricted reference state $|\tilde{\psi}_{\text{ref}}\rangle$ (29), the target error $\Delta_c(\mathcal{F})$ (25) and that which was actually achieved in our minimal ansatz simulations (Figure 2). We also include the corresponding contextual subspace Hamiltonians for reproducibility.

In Figure 2, we present the CS-VQE convergence data obtained from an exact wavefunction simulation of our noncontextual projection ansatz (27), which successfully reaches chemical accuracy in each instance. We used the Broyden–Fletcher–Goldfarb–Shanno (BFGS) [63] classical optimizer and computed the partial derivatives with respect to each ansatz parameter as per [64].

Extracting the optimal parameter configuration $\boldsymbol{\theta}_{\text{min}}$ – i.e. that which minimizes (30) – from the wavefunction simulations in Figure 2, we subsequently assess the effect of sampling noise on the simulation error with our ansatz circuit preparing the optimal quantum state $|\tilde{\psi}_{\text{anz}}(\boldsymbol{\theta}_{\text{min}})\rangle$.

To achieve an absolute error of $\Delta > 0$, one should expect to perform $\mathcal{O}(\frac{1}{\Delta^2})$ shots (for each term of the Hamiltonian) [4]. Conversely, suppose we are allocated a quantity $S \in \mathbb{N}$ of shots – the obtained error should be of the order $\mathcal{O}(\frac{1}{\sqrt{S}})$. In order to increase estimate accuracy, one should collect the Pauli terms into qubit-wise commuting (QWC) groups [25] that can be measured simultaneously.

In Figure 3, the number of shots $S = 2^n$ for $n = 0, \dots, 20$ carried out per QWC group is varied and we observe the root mean-square error (RMSE) over ten realizations of the ground state energy estimate, plotted on a log-log scale. For clarity, note the **only** source of noise here is that

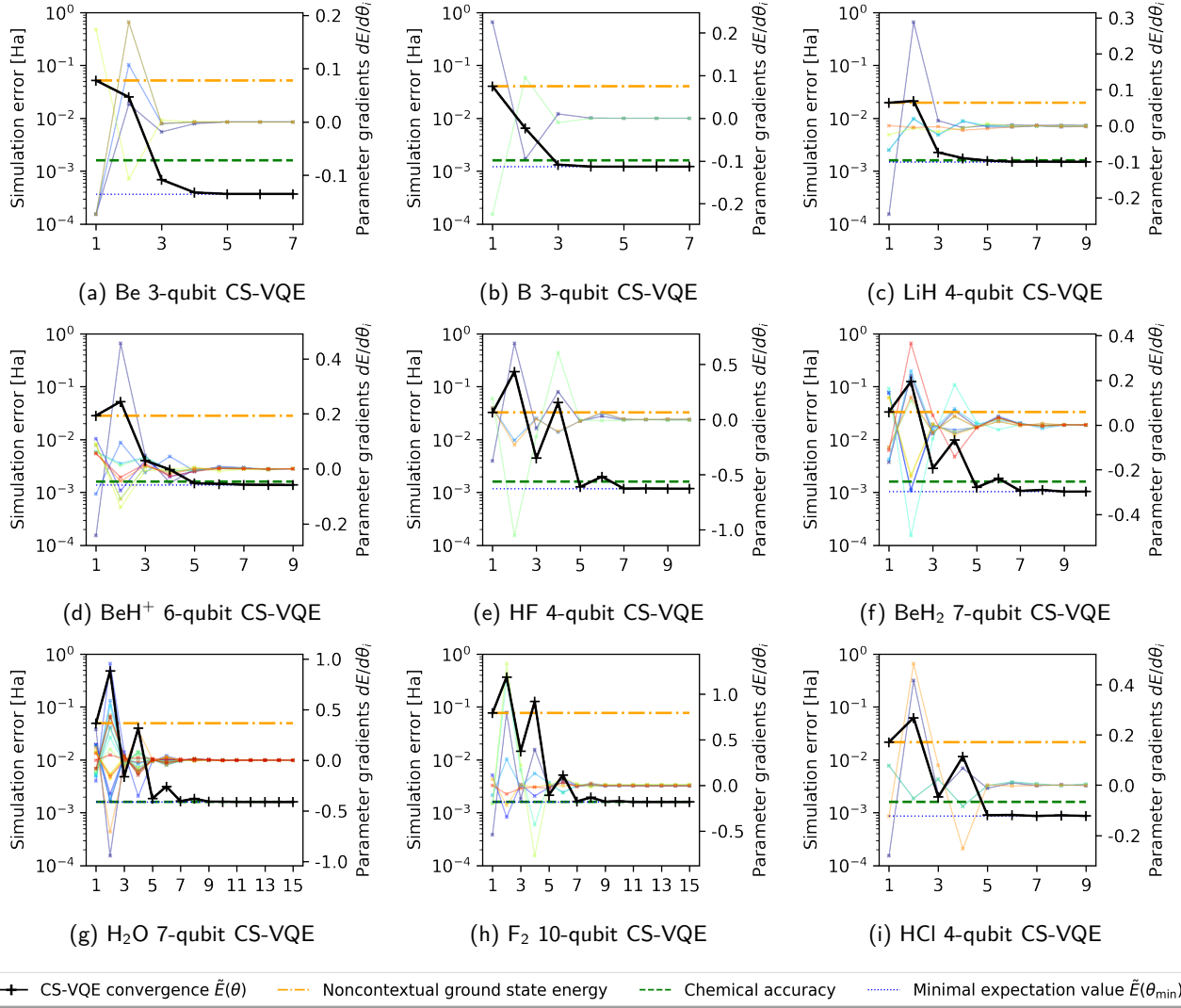


Figure 2: Empirical validation of the noncontextual projection ansatz for CS-VQE (27). We plot (on a $\log_{10}$ scale) the absolute error of a wavefunction simulation run on an Atos QLM for the suite of trial molecules outlined in Table 1, each shown to achieve chemical accuracy; the horizontal axis indicates the number of function evaluations (nfev). Broyden–Fletcher–Goldfarb–Shanno (BFGS) [63] is the classical optimizer taken in the VQE routine performed over the contextual subspace; we plot on the second (right-hand) axis the parameter gradients $\partial \tilde{E}(\theta)/\partial \theta_i$, computed using the parameter shift rule [64].

which arises from statistical variation of the quantum circuit sampling – we have **not** introduced hardware noise in the form of imperfect quantum gates or decoherence over time.

Two error regimes are observed, one of which is quite trivial: at high shot-counts we see a plateau resulting from the optimal error $|\tilde{E}(\theta_{\min}) - \epsilon_0|$ being recovered. To assess the convergence properties outside of this limiting region, we plot a line of best fit $m \cdot \log_{10}(S) + c$ among the data not exhibiting such behaviour; since the data is represented on a log-log scale, this corresponds with a decay in error of $\mathcal{O}(S^m)$. In each plot of Figure 3 we obtain $m \approx -0.5$, meaning the RMSE

follows the predicted decay of $\mathcal{O}(\frac{1}{\sqrt{S}})$.

In every simulation, chemical accuracy was achieved within $S = 2^{20} \approx 10^6$ shots per QWC group. However, our shot budget could be reduced by implementing more advanced allocation strategies, for example according to the magnitude of Hamiltonian term coefficients [65] or a classical shadow tomography approach [30, 31].

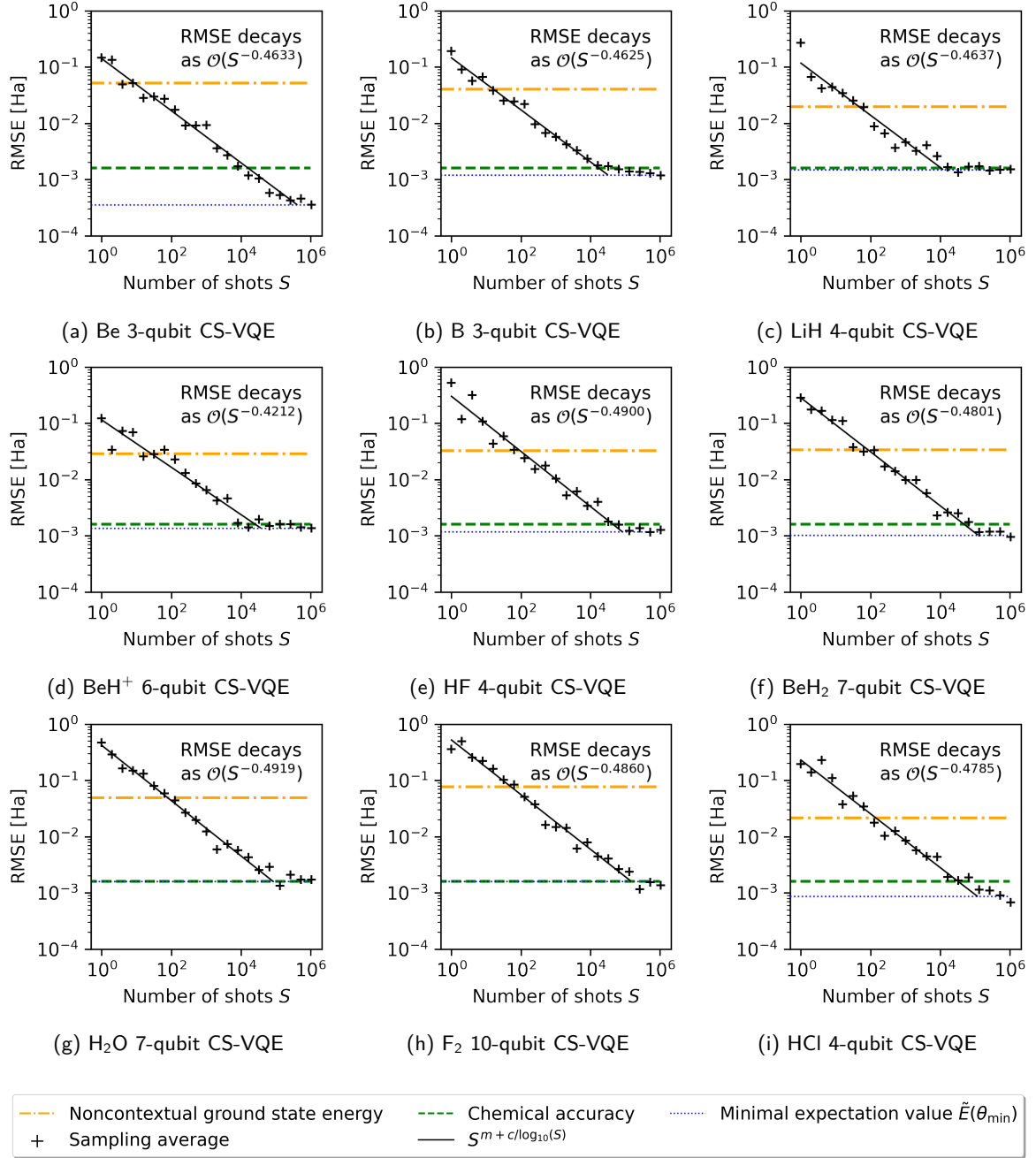


Figure 3: Each of the plots 3a - 3i correspond with 2a - 2i above and illustrate the statistical effect of sampling noise at the optimal parametrization $\theta_{\min}$ determined from the statevector simulations in Figure 2; for each molecule, $\theta_{\min}$ is given explicitly in Tables 3 - 10 of Appendix C. We plot the root mean-square error (RMSE) for ten ‘realizations’ of the ground state energy estimate with $S \leq 10^6$ shots executed via IBM’s QASM simulator; determining the line of best fit $m \cdot \log_{10}(S) + c$ with respect to the log-log data indicates a decay in error of $\mathcal{O}(S^m)$.

6 Discussion and further research

Given a general qubit Hamiltonian of the form (2), Contextual Subspace VQE allows one to scale the problem to the available quantum resource – regardless of the underlying physical system it might represent. In this way, it describes an interpolation between a purely classical estimate of the ground state energy and a full VQE simulation of the Hamiltonian performed over some ansatz space. In the context of electronic structure calculations, this often permits one to achieve chemical accuracy at a saving of qubits, as indicated in [44] and Figure 1.

We have placed CS-VQE on the theoretical footing of stabilizer subspace projections, which allows one to compare it against other qubit reduction techniques such as qubit tapering [46, 47]. Tapering defines a projection dependent on a symmetry of the full Hamiltonian and preserves the ground state energy exactly, whereas CS-VQE is approximate and projects onto a contextual subspace consistent with a symmetry of the noncontextual sub-Hamiltonian, augmented by an additional anticommuting part. Used in combination, the two techniques can effect a significant reduction in quantum resource requirements, as illustrated in Figures 1 and 2.

Previously, the only obstacle to building a CS-VQE framework that would be faithful to deployment on quantum devices was that of the ansatz, which we have addressed within this work. We tested our noncontextual projection ansatz for a suite of trial molecules (Table 1) using an Atos Quantum Learning Machine based at the Leibniz Supercomputing Centre; the results are found in Figure 2. The natural next step is to execute this method on a NISQ computer, challenging the current best-in-class electronic structure simulations from Google, IonQ and IBM [1–3]. To achieve this goal, CS-VQE could be combined with techniques of measurement reduction [20–35] and error mitigation [36–43].

The remaining research questions surrounding CS-VQE are concerned with scaling the method to larger molecular systems. To this end, there are a few obstacles that must be overcome.

Firstly, the success of CS-VQE is sensitive to the generator subset $\mathcal{F}$ one chooses to constrain in the stabilizer subspace projection. To date, the most effective method for choosing this subset has been a greedy-search heuristic that scales as

$\mathcal{O}(N^{d+1})$ where $d \leq N$ is the search depth. Previously, this involved diagonalizing the contextual subspace Hamiltonians and choosing those which minimized the energy for each qubit number, but with the noncontextual projection ansatz in place the linear algebra (requiring exponential classical memory overhead) can be replaced with true CS-VQE routines. While this approach might be scalable, it is expensive – there is room for more targeted heuristics.

For example, we may draw on chemical intuition to inform the selection of a contextual subspace that captures optimal information about the underlying electronic structure problem for a fixed number of qubits. In Section 4.3 we discussed a potential heuristic that aims to relax stabilizers acting non-trivially around the HOMO-LUMO gap, in line with active space reduction techniques – alternatively, one might define a Hamiltonian term-importance metric that considers coefficient magnitudes [66] or second-order response with respect to a perturbation of the Hartree-Fock state [67]. In relation to this, it is also not clear which features of a molecular system mean that it might be more or less amenable to CS-VQE; additional insight here would allow one to predict how many qubits will be required to simulate a given problem to chemical accuracy.

The second obstacle to the scaling of CS-VQE concerns the selection of a noncontextual subset from terms of the full Hamiltonian. A reformulation in terms of graphs – e.g. representing Pauli operators as nodes with (non)adjacency indicating (anti)commutation – will allow one to identify what the equivalent problem(s) are in computer science and therefore draw upon the vast body of existing research and select the best algorithms designed to solve such computational problems of graph theory. It should be noted the ‘optimal’ noncontextual subset will not necessarily be that which minimizes the noncontextual ground state energy – some consideration of the resulting contextual subspaces must come into the construction of the noncontextual problem.

We leave the remainder of these and other problems to further work.

Acknowledgements

T.J.W. and A.R. acknowledge support from the Engineering and Physical Sciences Research

Council (EP/S021582/1 and EP/L015242/1, respectively). T.J.W. also acknowledges support from CBKSciCon Ltd., Atos, Intel and Zapata. W.K. and P.J.L. acknowledge support by the NSF STAQ project (PHY-1818914). W. K. acknowledges support from the National Science Foundation, Grant No. DGE-1842474. P.V.C. is grateful for funding from the European Commission for VECMA (800925) and EPSRC for SEAVEA (EP/W007711/1). We would like to thank both Atos and the Leibniz Supercomputing Centre (LRZ), who each provided access to separate Atos Quantum Learning Machine (QLM) simulators that aided with the computational workload.

References

- [1] Frank Arute et al. “Hartree-Fock on a superconducting qubit quantum computer”. In: *Science* 369.6507 (2020), pp. 1084–1089. DOI: [10.1126/science.abb9811](https://doi.org/10.1126/science.abb9811).
- [2] Yunseong Nam et al. “Ground-state energy estimation of the water molecule on a trapped-ion quantum computer”. In: *npj Quantum Information* 6.1 (2020), pp. 1–6. DOI: [10.1038/s41534-020-0259-3](https://doi.org/10.1038/s41534-020-0259-3).
- [3] Andrew Eddins et al. “Doubling the size of quantum simulators by entanglement forging”. In: *arXiv preprint* (2021). arXiv: [2104.10220](https://arxiv.org/abs/2104.10220).
- [4] Alberto Peruzzo et al. “A variational eigenvalue solver on a photonic quantum processor”. In: *Nature communications* 5.1 (2014), pp. 1–7. DOI: doi.org/10.1038/ncomms5213.
- [5] A Yu Kitaev. “Quantum measurements and the Abelian stabilizer problem”. In: *arXiv preprint* (1995). arXiv: [quant-ph/9511026](https://arxiv.org/abs/quant-ph/9511026).
- [6] David J Griffiths and Darrell F Schroeter. *Introduction to quantum mechanics*. Cambridge university press, 2018. DOI: [10.1017/9781316995433](https://doi.org/10.1017/9781316995433).
- [7] Jarrod R McClean et al. “Barren plateaus in quantum neural network training landscapes”. In: *Nature communications* 9.1 (2018), pp. 1–6. DOI: [10.1038/s41467-018-07090-4](https://doi.org/10.1038/s41467-018-07090-4).
- [8] Dave Wecker, Matthew B Hastings, and Matthias Troyer. “Progress towards practical quantum variational algorithms”. In: *Physical Review A* 92.4 (2015), p. 042303. DOI: [10.1103/PhysRevA.92.042303](https://doi.org/10.1103/PhysRevA.92.042303).
- [9] Abhinav Kandala et al. “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets”. In: *Nature* 549.7671 (2017), pp. 242–246. DOI: [10.1038/nature23879](https://doi.org/10.1038/nature23879).
- [10] Arthur G Rattew et al. “A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver”. In: *arXiv preprint* (2019). arXiv: [1910.09694](https://arxiv.org/abs/1910.09694).
- [11] Roeland Wiersema et al. “Exploring entanglement and optimization within the hamiltonian variational ansatz”. In: *PRX Quantum* 1.2 (2020), p. 020319. DOI: [10.1103/PRXQuantum.1.020319](https://doi.org/10.1103/PRXQuantum.1.020319).
- [12] Matthew P Harrigan et al. “Quantum approximate optimization of non-planar graph problems on a planar superconducting processor”. In: *Nature Physics* 17.3 (2021), pp. 332–336. DOI: [10.1038/s41567-020-01105-y](https://doi.org/10.1038/s41567-020-01105-y).
- [13] Andrew G Taube and Rodney J Bartlett. “New perspectives on unitary coupled-cluster theory”. In: *International journal of quantum chemistry* 106.15 (2006), pp. 3393–3401. DOI: [10.1002/qua.21198](https://doi.org/10.1002/qua.21198).
- [14] Jonathan Romero et al. “Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz”. In: *Quantum Science and Technology* 4.1 (2018), p. 014008. DOI: [10.1088/2058-9565/aad3e4](https://doi.org/10.1088/2058-9565/aad3e4).
- [15] Bryan T Gard et al. “Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm”. In: *npj Quantum Information* 6.1 (2020), pp. 1–9. DOI: [10.1038/s41534-019-0240-1](https://doi.org/10.1038/s41534-019-0240-1).
- [16] Harper R Grimsley et al. “An adaptive variational algorithm for exact molecular simulations on a quantum computer”. In: *Nature communications* 10.1 (2019), pp. 1–9. DOI: [10.1038/s41467-019-10988-2](https://doi.org/10.1038/s41467-019-10988-2).

- [17] Ho Lun Tang et al. “qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor”. In: *PRX Quantum* 2.2 (2021), p. 020310. DOI: [10.1103/PRXQuantum.2.020310](https://doi.org/10.1103/PRXQuantum.2.020310).
- [18] VO Shkolnikov et al. “Avoiding symmetry roadblocks and minimizing the measurement overhead of adaptive variational quantum eigensolvers”. In: *arXiv preprint* (2021). arXiv: [2109.05340](https://arxiv.org/abs/2109.05340).
- [19] Dmitry A Fedorov et al. “Unitary Selective Coupled-Cluster Method”. In: *arXiv preprint* (2021). arXiv: [2109.12652](https://arxiv.org/abs/2109.12652).
- [20] Ryan Babbush et al. “Low-depth quantum simulation of materials”. In: *Physical Review X* 8.1 (2018), p. 011044. DOI: [10.1103/PhysRevX.8.011044](https://doi.org/10.1103/PhysRevX.8.011044).
- [21] Daochen Wang, Oscar Higgott, and Stephen Brierley. “Accelerated variational quantum eigensolver”. In: *Physical review letters* 122.14 (2019), p. 140504. DOI: [10.1103/PhysRevLett.122.140504](https://doi.org/10.1103/PhysRevLett.122.140504).
- [22] Vladyslav Verteletskyi, Tzu-Ching Yen, and Artur F Izmaylov. “Measurement optimization in the variational quantum eigensolver using a minimum clique cover”. In: *The Journal of chemical physics* 152.12 (2020), p. 124114. DOI: [10.1063/1.5141458](https://doi.org/10.1063/1.5141458).
- [23] Andrew Jena, Scott Genin, and Michele Mosca. “Pauli partitioning with respect to gate sets”. In: *arXiv preprint* (2019). arXiv: [1907.07859](https://arxiv.org/abs/1907.07859).
- [24] Pranav Gokhale et al. “Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families”. In: *arXiv preprint* (2019). arXiv: [1907.13623](https://arxiv.org/abs/1907.13623).
- [25] Tzu-Ching Yen, Vladyslav Verteletskyi, and Artur F Izmaylov. “Measuring all compatible operators in one series of single-qubit measurements using unitary transformations”. In: *Journal of chemical theory and computation* 16.4 (2020), pp. 2400–2409. DOI: [10.1021/acs.jctc.0c00008](https://doi.org/10.1021/acs.jctc.0c00008).
- [26] William J Huggins et al. “Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers”. In: *npj Quantum Information* 7.1 (2021), pp. 1–9. DOI: [10.1038/s41534-020-00341-7](https://doi.org/10.1038/s41534-020-00341-7).
- [27] Pranav Gokhale et al. “ $O(N^3)$ measurement Cost for Variational Quantum Eigensolver on Molecular Hamiltonians”. In: *IEEE Transactions on Quantum Engineering* 1 (2020), pp. 1–24. DOI: [10.1109/TQE.2020.3035814](https://doi.org/10.1109/TQE.2020.3035814).
- [28] Giacomo Torlai et al. “Precise measurement of quantum observables with neural-network estimators”. In: *Physical Review Research* 2.2 (2020), p. 022060. DOI: [10.1103/PhysRevResearch.2.022060](https://doi.org/10.1103/PhysRevResearch.2.022060).
- [29] Ophelia Crawford et al. “Efficient quantum measurement of Pauli operators in the presence of finite sampling error”. In: *Quantum* 5 (2021), p. 385. DOI: [10.22331/q-2021-01-20-385](https://doi.org/10.22331/q-2021-01-20-385).
- [30] Hsin-Yuan Huang, Richard Kueng, and John Preskill. “Predicting many properties of a quantum system from very few measurements”. In: *Nature Physics* 16.10 (2020), pp. 1050–1057. DOI: [10.1038/s41567-020-0932-7](https://doi.org/10.1038/s41567-020-0932-7).
- [31] Charles Hadfield et al. “Measurements of quantum Hamiltonians with locally-biased classical shadows”. In: *arXiv preprint* (2020). arXiv: [2006.15788](https://arxiv.org/abs/2006.15788).
- [32] Artur F Izmaylov et al. “Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method”. In: *Journal of chemical theory and computation* 16.1 (2019), pp. 190–195. DOI: [10.1021/acs.jctc.9b00791](https://doi.org/10.1021/acs.jctc.9b00791).
- [33] Andrew Zhao et al. “Measurement reduction in variational quantum algorithms”. In: *Physical Review A* 101.6 (2020), p. 062322. DOI: [10.1103/PhysRevA.101.062322](https://doi.org/10.1103/PhysRevA.101.062322).
- [34] Xavier Bonet-Monroig, Ryan Babbush, and Thomas E O’Brien. “Nearly optimal measurement scheduling for partial tomography of quantum states”. In: *Physical Review X* 10.3 (2020), p. 031064. DOI: [10.1103/PhysRevX.10.031064](https://doi.org/10.1103/PhysRevX.10.031064).

- [35] Alexis Ralli et al. “Implementation of measurement reduction for the variational quantum eigensolver”. In: *Physical Review Research* 3.3 (2021), p. 033195. DOI: [10.1103/PhysRevResearch.3.033195](https://doi.org/10.1103/PhysRevResearch.3.033195).
- [36] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. “Error mitigation for short-depth quantum circuits”. In: *Physical review letters* 119.18 (2017), p. 180509. DOI: [10.1103/PhysRevLett.119.180509](https://doi.org/10.1103/PhysRevLett.119.180509).
- [37] Ying Li and Simon C Benjamin. “Efficient variational quantum simulator incorporating active error minimization”. In: *Physical Review X* 7.2 (2017), p. 021050. DOI: [10.1103/PhysRevX.7.021050](https://doi.org/10.1103/PhysRevX.7.021050).
- [38] Suguru Endo, Simon C Benjamin, and Ying Li. “Practical quantum error mitigation for near-future applications”. In: *Physical Review X* 8.3 (2018), p. 031027. DOI: [10.1103/PhysRevX.8.031027](https://doi.org/10.1103/PhysRevX.8.031027).
- [39] Abhinav Kandala et al. “Error mitigation extends the computational reach of a noisy quantum processor”. In: *Nature* 567.7749 (2019), pp. 491–495. DOI: [10.1038/s41586-019-1040-7](https://doi.org/10.1038/s41586-019-1040-7).
- [40] Tudor Giurgica-Tiron et al. “Digital zero noise extrapolation for quantum error mitigation”. In: *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE. 2020, pp. 306–316. DOI: [10.1109/QCE49297.2020.00045](https://doi.org/10.1109/QCE49297.2020.00045).
- [41] Andre He et al. “Zero-noise extrapolation for quantum-gate error mitigation with identity insertions”. In: *Physical Review A* 102.1 (2020), p. 012426. DOI: [10.1103/PhysRevA.102.012426](https://doi.org/10.1103/PhysRevA.102.012426).
- [42] Suguru Endo et al. “Hybrid quantum-classical algorithms and quantum error mitigation”. In: *Journal of the Physical Society of Japan* 90.3 (2021), p. 032001. DOI: [10.7566/JPSJ.90.032001](https://doi.org/10.7566/JPSJ.90.032001).
- [43] William J Huggins et al. “Virtual distillation for quantum error mitigation”. In: *Physical Review X* 11.4 (2021), p. 041036. DOI: [10.1103/PhysRevX.11.041036](https://doi.org/10.1103/PhysRevX.11.041036).
- [44] William M Kirby, Andrew Tranter, and Peter J Love. “Contextual Subspace Variational Quantum Eigensolver”. In: *Quantum* 5 (2021), p. 456. DOI: [10.22331/q-2021-05-14-456](https://doi.org/10.22331/q-2021-05-14-456).
- [45] Maxwell D Radin and Peter Johnson. “Classically-Boosted Variational Quantum Eigensolver”. In: *arXiv preprint* (2021). arXiv: [2106.04755](https://arxiv.org/abs/2106.04755).
- [46] Sergey Bravyi et al. “Tapering off qubits to simulate fermionic Hamiltonians”. In: *arXiv preprint* (2017). arXiv: [1701.08213](https://arxiv.org/abs/1701.08213).
- [47] Kanav Setia et al. “Reducing qubit requirements for quantum simulations using molecular point group symmetries”. In: *Journal of Chemical Theory and Computation* 16.10 (2020), pp. 6091–6097. DOI: [10.1021/acs.jctc.0c00113](https://doi.org/10.1021/acs.jctc.0c00113).
- [48] Daniel Gottesman. *Stabilizer codes and quantum error correction*. California Institute of Technology, 1997. arXiv: [quant-ph/9705052](https://arxiv.org/abs/quant-ph/9705052).
- [49] Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: [10.1017/CBO9780511976667](https://doi.org/10.1017/CBO9780511976667).
- [50] Karl Kraus et al. “States, effects, and operations: fundamental notions of quantum theory”. In: *Lectures in mathematical physics at the University of Texas at Austin* 190 (1983). DOI: [10.1007/3-540-12732-1](https://doi.org/10.1007/3-540-12732-1).
- [51] John S Bell. “On the einstein podolsky rosen paradox”. In: *Physique Physique Fizika* 1.3 (1964), p. 195. DOI: [10.1103/PhysicsPhysiqueFizika.1.195](https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195).
- [52] John S Bell. “On the problem of hidden variables in quantum mechanics”. In: *Reviews of Modern Physics* 38.3 (1966), p. 447. DOI: [10.1103/RevModPhys.38.447](https://doi.org/10.1103/RevModPhys.38.447).
- [53] Simon Kochen and Ernst P Specker. “The problem of hidden variables in quantum mechanics”. In: *The logico-algebraic approach to quantum mechanics*. Springer, 1975, pp. 293–328. DOI: [10.1007/978-94-010-1795-4_17](https://doi.org/10.1007/978-94-010-1795-4_17).

- [54] William M Kirby and Peter J Love. “Contextuality test of the nonclassicality of variational quantum eigensolvers”. In: *Physical review letters* 123.20 (2019), p. 200501. DOI: [10.1103/PhysRevLett.123.200501](https://doi.org/10.1103/PhysRevLett.123.200501).
- [55] Robert Raussendorf et al. “Phase-space-simulation method for quantum computation with magic states on qubits”. In: *Physical Review A* 101.1 (2020), p. 012350. DOI: [10.1103/PhysRevA.101.012350](https://doi.org/10.1103/PhysRevA.101.012350).
- [56] William M Kirby and Peter J Love. “Classical simulation of noncontextual Pauli Hamiltonians”. In: *Physical Review A* 102.3 (2020), p. 032418. DOI: [10.1103/PhysRevA.102.032418](https://doi.org/10.1103/PhysRevA.102.032418).
- [57] Robert W. Spekkens. “Quasi-Quantization: Classical Statistical Theories with an Epistemic Restriction”. In: *Quantum Theory: Informational Foundations and Foils*. Dordrecht: Springer Netherlands, 2016, pp. 83–135. ISBN: 978-94-017-7303-4. DOI: [10.1007/978-94-017-7303-4_4](https://doi.org/10.1007/978-94-017-7303-4_4). URL: https://doi.org/10.1007/978-94-017-7303-4_4.
- [58] Peter Schwerdtfeger. “The pseudopotential approximation in electronic structure theory”. In: *ChemPhysChem* 12.17 (2011), pp. 3143–3155. DOI: [10.1002/cphc.201100387](https://doi.org/10.1002/cphc.201100387).
- [59] Russell D. Johnson III. *NIST Computational Chemistry Comparison and Benchmark Database*. NIST Standard Reference Database Number 101. 2020. DOI: [10.18434/T47C7Z](https://doi.org/10.18434/T47C7Z).
- [60] MD SAJID ANIS et al. *Qiskit: An Open-source Framework for Quantum Computing*. 2021. DOI: [10.5281/zenodo.2573505](https://doi.org/10.5281/zenodo.2573505).
- [61] Qiming Sun et al. “PySCF: the Python-based simulations of chemistry framework”. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 8.1 (2018), e1340. DOI: [10.1002/wcms.1340](https://doi.org/10.1002/wcms.1340).
- [62] Pascual Jordan and Eugene Paul Wigner. “über das paulische äquivalenzverbot”. In: *The Collected Works of Eugene Paul Wigner*. Springer, 1993, pp. 109–129. DOI: [10.1007/BF01331938](https://doi.org/10.1007/BF01331938).
- [63] Roger Fletcher. *Practical methods of optimization*. John Wiley & Sons, 2013. DOI: [10.1002/9781118723203](https://doi.org/10.1002/9781118723203).
- [64] Robert M Parrish et al. “Hybrid quantum/classical derivative theory: Analytical gradients and excited-state dynamics for the multistate contracted variational quantum eigensolver”. In: *arXiv preprint* (2019). arXiv: [1906.08728](https://arxiv.org/abs/1906.08728).
- [65] Andrew Arrasmith et al. “Operator sampling for shot-frugal optimization in variational algorithms”. In: *arXiv preprint* (2020). arXiv: [2004.06252](https://arxiv.org/abs/2004.06252).
- [66] Dave Wecker et al. “Gate-count estimates for performing quantum chemistry on small quantum computers”. In: *Physical Review A* 90.2 (2014), p. 022305. DOI: [10.1103/PhysRevA.90.022305](https://doi.org/10.1103/PhysRevA.90.022305).
- [67] David Poulin et al. “The Trotter step size required for accurate quantum simulation of quantum chemistry”. In: *arXiv preprint* (2014). arXiv: [1406.4920](https://arxiv.org/abs/1406.4920).
- [68] Andrew Tranter et al. “Ordering of Trotterization: Impact on errors in quantum simulation of electronic structure”. In: *Entropy* 21.12 (2019), p. 1218. DOI: [10.3390/e21121218](https://doi.org/10.3390/e21121218).
- [69] Harper R Grimsley et al. “Is the Trotterized UCCSD Ansatz chemically well-defined?” In: *Journal of chemical theory and computation* 16.1 (2019), pp. 1–6. DOI: [10.1021/acs.jctc.9b01083](https://doi.org/10.1021/acs.jctc.9b01083).

A Circuit implementation

Here we introduce the key concepts that are required to implement our CS-VQE ansatz as a quantum circuit, withholding a thorough introduction to the topic (the reader is referred to Nielsen and Chuang for this [49]).

A.1 Trotterization

For operators $A, B \in \mathcal{B}(\mathcal{H})$, we have $e^{A+B} = e^A e^B$ if and only if A commutes with B (i.e. $[A, B] = 0$), contrary to the familiar rules of exponentiation for numbers; in the more general case, the Lie product formula states

$$e^{A+B} = \lim_{n \rightarrow \infty} \left(e^{A/n} e^{B/n} \right)^n. \quad (34)$$

The technique of *Suzuki-Trotter* truncates the above limit at some $n_T \in \mathbb{N}$ (the *Trotter number*) to yield an approximation of the exponentiated operator sum. Given a sum of operators $\sum_i \theta_i A_i$, we may write the first-order Trotter expansion:

$$e^{\sum_i \theta_i A_i} \approx \left(\prod_i e^{\frac{\theta_i}{n_T} A_i} \right)^{n_T}. \quad (35)$$

Used in combination with VQE, $n_T = 2$ is usually sufficient to reach chemical accuracy since we expect the optimizer to produce different coefficients that counteract the Trotter error. This is what we used when constructing our ansätze to keep the circuit depths at workable levels. It has also been observed that the ordering of Trotter terms has an impact on errors [68] – this will not be explored here, however it is a possible consideration for future research. There is contention over Trotterized UCC and whether it is well-defined as an ansatz, since the ordering of terms can induce variations in error beyond chemical accuracy, particularly for strongly correlated systems [69].

A.2 Exponentiating Pauli strings

Given a Pauli operator $P \in \mathcal{P}_N$ and some angle $\theta \in \mathbb{R}$, we would like to implement the exponential $e^{i\theta P}$ as a quantum circuit; this can be achieved with $\mathcal{O}(N)$ gates. We shall first assume that P consists of just Pauli I, Z operators, with the qubit positions acted upon by Z indexed with the set $\mathcal{I}_Z$. Observe that

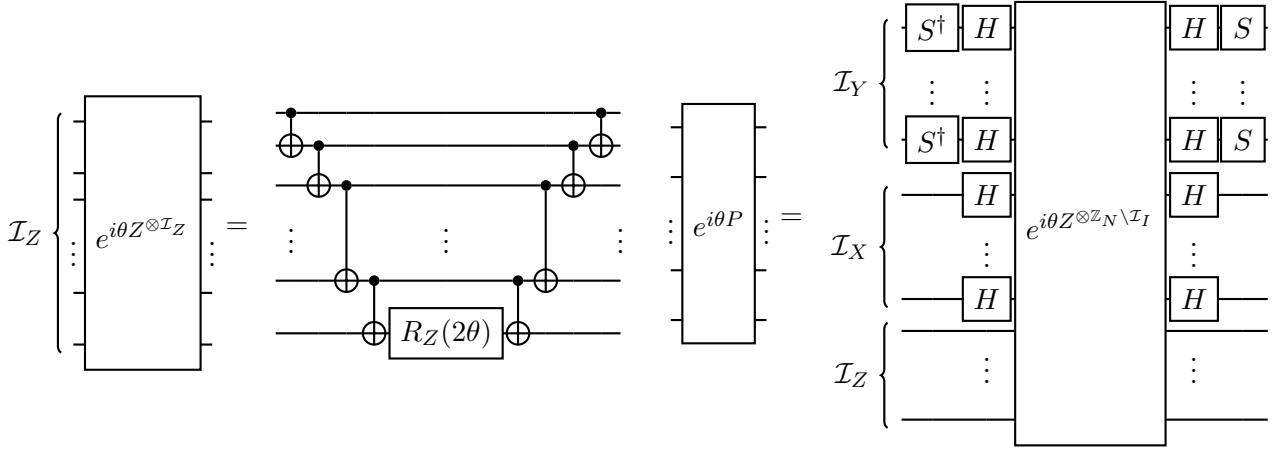
$$\begin{aligned} e^{i\theta Z^{\otimes \mathcal{I}_Z}} |\mathbf{b}\rangle &= \left(\cos \theta + i(-1)^{\rho_{\mathcal{I}_Z}(\mathbf{b})} \sin \theta \right) |\mathbf{b}\rangle \\ &= e^{i\theta(-1)^{\rho_{\mathcal{I}_Z}(\mathbf{b})}} |\mathbf{b}\rangle. \end{aligned} \quad (36)$$

where we have omitted the qubit positions that are identity. Therefore, we may realise this operation by storing $\rho_{\mathcal{I}_Z}(\mathbf{b})$, the parity of $|\mathbf{b}\rangle$ over $\mathcal{I}_Z$, in one of the qubits and applying to it an R_Z gate, defined by

$$R_Z(\phi) = \begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix}. \quad (37)$$

In (36), even parity results in a phase $e^{i\theta}$ whereas odd parity yields $e^{-i\theta}$ – these are obtained by application of $R_Z(2\theta)$ to each of $|0\rangle, |1\rangle$, respectively. The parity computation is accomplished via a ‘cascade’ of CNOT gates between adjacent qubits. We are now in a position to explicitly write down a quantum circuit that effects the exponentiation of a Pauli string consisting of I, Z operators (Figure A.1a).

This is the basic building block for exponentiating an arbitrary Pauli string, as we are able to make a change of basis from Pauli X, Y operators to a Pauli Z . Note that $X = HZH$ and $Y = SHZHS^\dagger$ where $S = R_Z(\pi/2)$. Once this transformation has been applied, the problem is reduced to that of before and we perform the same parity computation and Z rotation. We end with the reverse transformation taking us back into the original basis – the complete circuit is presented in Figure



(a) Circuit to realise the exponential $e^{i\theta P}$ where P consists of just Pauli I, Z operators. The qubits set to identity are omitted from the diagram. (b) Circuit to realise the exponential $e^{i\theta P}$ where P is an arbitrary Pauli string. The qubits set to identity are omitted from the diagram.

Figure A.1: Quantum circuit implementations for exponentiated Pauli operators.

A.1b. The total number of gates is at most $6N - 1$, achieved only when P consists solely of Pauli Y operators acting upon all of the N qubit positions.

It is noted that in the Trotterized circuit there will be many blocks of exponentiated Pauli strings of the form displayed in Figure A.1b, and it is possible that gate cancellation will occur between adjacent blocks where we have $SS^\dagger = H^2 = \mathbb{1}$. Furthermore, the implementation presented here is not unique – different circuits yielding the same quantum states are possible.

B Classical objective function

The CS-VQE method relies on solving a classical objective function that defines the energy spectrum of the noncontextual Hamiltonian

$$H_{\mathcal{T}_{\text{nc}}} = \sum_{\sigma \in \mathcal{T}_{\text{nc}}} h_{\sigma} \sigma. \quad (38)$$

The objective function is constructed from the decomposition 17, in which one observes that the sum over $\mathcal{T}_{\text{nc}}$ may be separated into individual summations over the symmetry terms $\mathcal{S}$ and the anticommuting classes $\mathcal{C}_i$ for $i \in \{1, \dots, M\}$, i.e. $\sum_{\sigma \in \mathcal{T}_{\text{nc}}} = \sum_{\sigma \in \mathcal{S}} + \sum_{i=1}^M \sum_{\sigma \in \mathcal{C}_i}$. Furthermore, since $\mathcal{G}$ generates the symmetry group $\mathcal{S}'$ and for any $C'_i \in \mathcal{C}_i$ we have $C_i C'_i \in \mathcal{S}'$ by construction, it holds that $\mathcal{S} \subset \overline{\mathcal{G}}$ and $\mathcal{C}_i \subset C_i \overline{\mathcal{G}} := \{C_i \sigma | \sigma \in \overline{\mathcal{G}}\}$. This means we may instead sum over the completion of $\mathcal{G}$, noting that any summand not appearing in $\mathcal{T}_{\text{nc}}$ will have zero coefficient:

$$H_{\mathcal{T}_{\text{nc}}} = \sum_{\sigma \in \overline{\mathcal{G}}} \left(h'_{\sigma} + \sum_{i=1}^M h_{\sigma, i} C_i \right) \sigma, \quad (39)$$

where

$$h'_{\sigma} = \begin{cases} h_{\sigma}, & \sigma \in \mathcal{S} \\ 0, & \text{otherwise} \end{cases}, \quad h_{\sigma, i} = \begin{cases} h_{C_i \sigma}, & C_i \sigma \in \mathcal{C}_i \\ 0, & \text{otherwise} \end{cases}. \quad (40)$$

We now wish to use this reformulation of the noncontextual Hamiltonian to evaluate expectation values in terms of the parameters $\nu \in \{\pm 1\}^{\times |\mathcal{G}|}$, denoting eigenvalue assignments to the generators $\mathcal{G}$ (sectors), and $\mathbf{r} \in \mathbb{R}^M$, a unit vector weighting the class representatives – these were introduced in Section 4. We define expectation values with respect to a joint probability distribution

$$P(\nu', \mathbf{r}' | \nu, \mathbf{r}) := \frac{\delta_{\nu, \nu'}}{2^M} \prod_{i=1}^M |r_i + r'_i|, \quad (41)$$

which is a special-case of the phase-space distribution given in [57]. In [56] it was proved the generators take expectation values

$$\langle G \rangle_{(\boldsymbol{\nu}, \mathbf{r})} = \nu_{f(G)} \quad \forall G \in \mathcal{G}, \quad (42)$$

recalling from Section 3 that $f : \mathcal{G} \rightarrow \mathcal{I}_{\text{stab}}$ is a mapping onto distinct qubit indices, and class representatives

$$\langle C_i \rangle_{(\boldsymbol{\nu}, \mathbf{r})} = r_i \quad \forall i \in \{1, \dots, M\}. \quad (43)$$

Observe that

$$\langle C(\mathbf{r}) \rangle_{(\boldsymbol{\nu}, \mathbf{r})} = \sum_{i=1}^M r_i^2 = |\mathbf{r}| = +1, \quad (44)$$

implying any quantum state consistent with the noncontextual state $(\boldsymbol{\nu}, \mathbf{r})$ must be stabilized by $C(\mathbf{r})$.

Putting everything together, we can express the expectation value of our noncontextual Hamiltonian as

$$\begin{aligned} \langle H_{\mathcal{T}_{\text{nc}}} \rangle_{(\boldsymbol{\nu}, \mathbf{r})} &= \sum_{\boldsymbol{\sigma} \in \overline{\mathcal{G}}} \left(h'_{\boldsymbol{\sigma}} + \sum_{i=1}^M h_{\boldsymbol{\sigma}, i} \langle C_i \rangle_{(\boldsymbol{\nu}, \mathbf{r})} \right) \langle \boldsymbol{\sigma} \rangle_{(\boldsymbol{\nu}, \mathbf{r})} \\ &= \sum_{\boldsymbol{\sigma} \in \overline{\mathcal{G}}} \left(h'_{\boldsymbol{\sigma}} + \sum_{i=1}^M h_{\boldsymbol{\sigma}, i} r_i \right) \prod_{G \in \mathcal{G}_{\boldsymbol{\sigma}}} \nu_{f(G)}, \end{aligned} \quad (45)$$

where $\mathcal{G}_{\boldsymbol{\sigma}} \subset \mathcal{G}$ satisfies $\boldsymbol{\sigma} = \prod_{G \in \mathcal{G}_{\boldsymbol{\sigma}}} G$. Note also that we have used the fact $\langle AB \rangle = \langle A \rangle \langle B \rangle$ when $[A, B] = 0$.

Taken as a classical optimization problem, minimizing the objective function

$$\eta(\boldsymbol{\nu}, \mathbf{r}) := \langle H_{\mathcal{T}_{\text{nc}}} \rangle_{(\boldsymbol{\nu}, \mathbf{r})} \quad (46)$$

is NP-complete in general. Despite this, we expect typical instances to be heuristically solvable by classical methods [56]. If one fixes the ± 1 eigenvalue assignments $\boldsymbol{\nu}$ – a case of identifying the correct symmetry sector, discussed in section 3 – this becomes a convex optimization problem over points of the unit $(M - 1)$ -sphere. In Tables 3 – 10 of Appendix C we list the symmetry generators $\mathcal{G}$, class representatives C_i and optimal parameters $\boldsymbol{\nu}, \mathbf{r}$ for the molecules in Table 1.

C Hamiltonian data

In the interest of reproducibility, we report in the following Tables 3 – 10 the tapering parameters, CS-VQE model data and noncontextual projection ansätze which permit chemical accuracy for the fewest number of qubits with respect to the molecular systems listed in Table 1. This should provide sufficient information for the reader to reproduce the simulation data in Figures 2, 3.

Table 3: Be qubit tapering and 3-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(7)}\sigma_3^{(2)}, \sigma_3^{(8)}\sigma_3^{(3)}, \sigma_3^{(9)}\sigma_3^{(4)}, \sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}$
Sector ν	$\{1, 1, 1, 1, 1\}$
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(0)}, \sigma_3^{(3)}\sigma_3^{(0)}, \sigma_3^{(2)}\sigma_3^{(0)}, \sigma_3^{(0)}$
Sector ν	$\{-1, -1, -1, -1\}$
Class observable $C(\mathbf{r})$	$1.722957938005497 \times 10^{-8} \cdot \sigma_1^{(1)}\sigma_3^{(0)} +$ $-0.9999999999999999 \cdot \sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}$
3-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(0)}, C(\mathbf{r})$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	$ 001\rangle$
Minimal ansatz $A(\boldsymbol{\theta})$	$\theta_0 \cdot \sigma_2^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} + \theta_1 \cdot \sigma_1^{(2)}\sigma_2^{(1)}\sigma_1^{(0)} + \theta_2 \cdot \sigma_2^{(4)}\sigma_2^{(1)}\sigma_2^{(0)} + \theta_3 \cdot \sigma_1^{(4)}\sigma_1^{(1)}\sigma_2^{(0)} +$ $\theta_4 \cdot \sigma_2^{(2)}\sigma_1^{(1)}\sigma_1^{(0)}$
Projected ansatz $\pi_{\nu'}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta}))$	$\theta_0 \cdot \sigma_2^{(1)}\sigma_1^{(0)} + \theta_1 \cdot \sigma_2^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + \theta_2 \cdot \sigma_1^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} + \theta_3 \cdot \sigma_2^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + \theta_4 \cdot \sigma_2^{(2)}\sigma_1^{(0)}$
Optimal parameters $\boldsymbol{\theta}_{\min}$	$(0.18855924109162348, 0.12262744298420095, 0.06888096831655995,$ $-0.12257006173804036, 0.06870565320643347)$
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.000327
VQE simulation absolute error $ \tilde{E}(\boldsymbol{\theta}_{\min}) - \epsilon_0 $ (Ha)	0.000354
Contextual subspace Hamiltonian	
<i>Diagonal terms</i>	
$-12.488598 \cdot \mathbb{1} + 0.858294 \cdot \sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.858294 \cdot \sigma_3^{(2)}\sigma_3^{(0)} + 0.023043 \cdot \sigma_3^{(1)} + 0.858294 \cdot \sigma_3^{(1)}\sigma_3^{(0)} + 0.023043 \cdot$ $\sigma_3^{(2)} + 0.023043 \cdot \sigma_3^{(2)}\sigma_3^{(1)} - 0.642471 \cdot \sigma_3^{(0)}$	
<i>Off-diagonal terms</i>	
$-0.012125 \cdot \sigma_1^{(2)}\sigma_3^{(0)} + 0.012125 \cdot \sigma_1^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.012125 \cdot \sigma_1^{(1)}\sigma_3^{(0)} + 0.012125 \cdot \sigma_3^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.012125 \cdot \sigma_1^{(2)}\sigma_1^{(1)} +$ $0.012125 \cdot \sigma_2^{(2)}\sigma_2^{(1)} + 0.043404 \cdot \sigma_2^{(2)}\sigma_2^{(1)}\sigma_1^{(0)} - 0.043404 \cdot \sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.043404 \cdot \sigma_3^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.043404 \cdot \sigma_2^{(1)}\sigma_2^{(0)} +$ $0.043404 \cdot \sigma_2^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} + 0.043404 \cdot \sigma_2^{(2)}\sigma_2^{(0)}$	

Table 4: B qubit tapering and 3-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(7)}\sigma_3^{(2)}, \sigma_3^{(8)}\sigma_3^{(3)}, \sigma_3^{(9)}\sigma_3^{(4)}, \sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}$
Sector ν	$\{-1, -1, 1, 1, -1\}$
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(0)}, \sigma_3^{(3)}, \sigma_3^{(2)}\sigma_3^{(0)}, \sigma_3^{(0)}$
Sector ν	$\{1, 1, 1, -1\}$
Class observable $C(\mathbf{r})$	$1.0 \cdot \sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} + 1.633162674519949 \times 10^{-11} \cdot \sigma_1^{(1)}\sigma_3^{(0)}$
3-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(0)}, C(\mathbf{r})$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	$ 011\rangle$
Minimal ansatz $A(\boldsymbol{\theta})$	$\theta_0 \cdot \sigma_2^{(4)}\sigma_2^{(1)}\sigma_2^{(0)} + \theta_1 \cdot \sigma_1^{(3)}\sigma_1^{(1)}\sigma_2^{(0)}$
Projected ansatz $\pi_{\nu'}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta}))$	$\theta_0 \cdot \sigma_3^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} + \theta_1 \cdot \sigma_1^{(2)}\sigma_2^{(1)}\sigma_1^{(0)}$
Optimal parameters $\boldsymbol{\theta}_{\text{min}}$	$(-0.16682258491581778, 0.16681318977165974)$
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.000281
VQE simulation absolute error $ \tilde{E}(\boldsymbol{\theta}_{\text{min}}) - \epsilon_0 $ (Ha)	0.001209
Contextual subspace Hamiltonian	
Diagonal terms	
$-22.759707 \cdot \mathbb{1} + 0.630938 \cdot \sigma_3^{(2)}\sigma_3^{(1)} - 0.630938 \cdot \sigma_3^{(2)} - 0.507699 \cdot \sigma_3^{(1)} + 0.477327 \cdot \sigma_3^{(2)}\sigma_3^{(0)} - 0.477327 \cdot \sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.319551 \cdot \sigma_3^{(0)}$	
Off-diagonal terms	
$-0.015815 \cdot \sigma_1^{(2)} - 0.015815 \cdot \sigma_1^{(2)}\sigma_3^{(1)} - 0.056284 \cdot \sigma_2^{(1)}\sigma_2^{(0)} + 0.056284 \cdot \sigma_3^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.056284 \cdot \sigma_1^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} - 0.056284 \cdot \sigma_2^{(2)}\sigma_2^{(1)}\sigma_1^{(0)}$	

Table 5: LiH qubit tapering and 4-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(10)}\sigma_3^{(9)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(9)}\sigma_3^{(3)}, \sigma_3^{(10)}\sigma_3^{(4)}, \sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}$
Sector ν	$\{1, 1, 1, 1\}$
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(7)}, \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(3)}, \sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(0)}$
Sector ν	$\{1, -1, -1, -1, -1, -1, -1\}$
Class observable $C(\mathbf{r})$	$-6.022245751909695 \times 10^{-9} \cdot \sigma_3^{(7)}\sigma_1^{(4)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + -1.0 \cdot \sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}$
4-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, C(\mathbf{r})$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	$ 0101\rangle$
Minimal ansatz $A(\boldsymbol{\theta})$	$\theta_0 \cdot \sigma_2^{(7)}\sigma_3^{(4)}\sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + \theta_1 \cdot \sigma_2^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_1^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} +$ $\theta_2 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_2^{(4)}\sigma_2^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} + \theta_3 \cdot \sigma_2^{(4)}\sigma_2^{(3)}\sigma_1^{(1)}\sigma_2^{(0)} + \theta_4 \cdot \sigma_2^{(1)}\sigma_1^{(0)}$
Projected ansatz $\pi_{\nu}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta}))$	$\theta_0 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(0)} + \theta_1 \cdot \sigma_2^{(3)}\sigma_1^{(2)}\sigma_1^{(0)} + \theta_2 \cdot \sigma_1^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} + \theta_3 \cdot \sigma_1^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} + \theta_4 \cdot \sigma_2^{(0)}$
Optimal parameters $\boldsymbol{\theta}_{\min}$	$(0.11286818003486217, 0.05757915581275913, 0.051018669392352556, 0.0321827636108793, 0.03234586900271787)$
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.001301
VQE simulation absolute error $ \tilde{E}(\boldsymbol{\theta}_{\min}) - \epsilon_0 $ (Ha)	0.001500
Contextual subspace Hamiltonian	
Diagonal terms	
$-8.021893 \cdot \mathbb{1} - 0.365272 \cdot \sigma_3^{(3)} + 0.21189 \cdot \sigma_3^{(3)}\sigma_3^{(2)} - 0.092758 \cdot \sigma_3^{(2)} - 0.365272 \cdot \sigma_3^{(1)} + 0.113844 \cdot \sigma_3^{(3)}\sigma_3^{(1)} - 0.06044 \cdot$ $\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} + 0.113953 \cdot \sigma_3^{(2)}\sigma_3^{(1)} + 0.21189 \cdot \sigma_3^{(1)}\sigma_3^{(0)} - 0.06044 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.084602 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.05629 \cdot$ $\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.092758 \cdot \sigma_3^{(0)} + 0.113953 \cdot \sigma_3^{(3)}\sigma_3^{(0)} - 0.05629 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(0)} + 0.122744 \cdot \sigma_3^{(2)}\sigma_3^{(0)}$	
Off-diagonal terms	
$0.01939 \cdot \sigma_1^{(3)} + 0.01939 \cdot \sigma_1^{(3)}\sigma_3^{(2)} + 0.010953 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} + 0.009003 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.012779 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} +$ $0.010953 \cdot \sigma_1^{(3)}\sigma_3^{(1)} + 0.009003 \cdot \sigma_1^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.012779 \cdot \sigma_1^{(3)}\sigma_3^{(0)} - 0.002941 \cdot \sigma_1^{(2)} - 0.002941 \cdot \sigma_1^{(2)}\sigma_3^{(1)} - 0.010682 \cdot$ $\sigma_1^{(2)}\sigma_3^{(1)} - 0.001697 \cdot \sigma_1^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.011925 \cdot \sigma_1^{(2)}\sigma_3^{(0)} + 0.000742 \cdot \sigma_2^{(3)}\sigma_2^{(2)} + 0.000742 \cdot \sigma_1^{(3)}\sigma_1^{(2)} + 0.034391 \cdot$ $\sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(1)} + 0.034391 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)} + 0.002737 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.002737 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.032397 \cdot$ $\sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(0)} + 0.032397 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(0)} - 0.010682 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)} - 0.001697 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.011925 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(0)} -$ $0.01939 \cdot \sigma_1^{(1)}\sigma_3^{(0)} - 0.010953 \cdot \sigma_3^{(3)}\sigma_1^{(1)}\sigma_3^{(0)} - 0.009003 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} - 0.012779 \cdot \sigma_3^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} - 0.01939 \cdot \sigma_1^{(1)} -$ $0.010953 \cdot \sigma_3^{(3)}\sigma_1^{(1)} - 0.009003 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(1)} - 0.012779 \cdot \sigma_3^{(2)}\sigma_1^{(1)} - 0.006588 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} - 0.006588 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(1)} -$ $0.006588 \cdot \sigma_1^{(3)}\sigma_1^{(1)}\sigma_3^{(0)} - 0.006588 \cdot \sigma_1^{(3)}\sigma_1^{(1)} + 0.002221 \cdot \sigma_1^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.002221 \cdot \sigma_1^{(2)}\sigma_1^{(1)} - 0.007859 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} -$ $0.007859 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} - 0.007859 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(1)} - 0.007859 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(1)} + 0.002221 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.002221 \cdot$ $\sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(1)} + 0.000742 \cdot \sigma_2^{(1)}\sigma_2^{(0)} + 0.000742 \cdot \sigma_1^{(1)}\sigma_1^{(0)} + 0.034391 \cdot \sigma_3^{(3)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.002737 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} +$ $0.032397 \cdot \sigma_3^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.034391 \cdot \sigma_3^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.002737 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.032397 \cdot \sigma_3^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.007859 \cdot$ $\sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.007859 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} + 0.007859 \cdot \sigma_1^{(3)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.007859 \cdot \sigma_1^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} - 0.008499 \cdot$ $\sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} - 0.008499 \cdot \sigma_1^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.030846 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.030846 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.030846 \cdot$ $\sigma_1^{(3)}\sigma_1^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.030846 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} - 0.008499 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} - 0.008499 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.002941 \cdot$ $\sigma_3^{(3)}\sigma_1^{(0)} + 0.010682 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.001697 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.011925 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.002941 \cdot \sigma_1^{(1)} + 0.010682 \cdot$ $\sigma_3^{(3)}\sigma_1^{(0)} + 0.001697 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.011925 \cdot \sigma_3^{(3)}\sigma_1^{(0)} + 0.002221 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.002221 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(0)} + 0.002221 \cdot$ $\sigma_1^{(3)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.002221 \cdot \sigma_1^{(3)}\sigma_1^{(0)} - 0.003139 \cdot \sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} - 0.003139 \cdot \sigma_1^{(2)}\sigma_1^{(0)} + 0.008499 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.008499 \cdot$ $\sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.008499 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(0)} + 0.008499 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(0)} - 0.003139 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} - 0.003139 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(0)}$	

Table 6: BeH⁺ qubit tapering and 6-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(10)} \sigma_3^{(9)} \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)}, \sigma_3^{(9)} \sigma_3^{(3)}, \sigma_3^{(10)} \sigma_3^{(4)}, \sigma_3^{(11)} \sigma_3^{(10)} \sigma_3^{(9)} \sigma_3^{(8)} \sigma_3^{(7)} \sigma_3^{(6)}$
Sector ν	$\{-1, 1, 1, 1\}$
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(7)}, \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)}, \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)}, \sigma_3^{(3)}, \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)}, \sigma_3^{(1)} \sigma_3^{(0)}, \sigma_3^{(0)}$
Sector ν	$\{1, 1, 1, -1, 1, 1, -1\}$
Class observable $C(\mathbf{r})$	$0.002132853189680987 \cdot \sigma_3^{(7)} \sigma_1^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} +$ $-0.9999977254660489 \cdot \sigma_3^{(7)} \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(3)}$
6-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)}, C(\mathbf{r})$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	$ 011111\rangle$
Minimal ansatz $A(\boldsymbol{\theta})$	$\theta_0 \cdot \sigma_2^{(7)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} + \theta_1 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} +$ $\theta_2 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_2^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_2^{(0)} + \theta_3 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_1^{(1)} + \theta_4 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_2^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} +$ $\theta_5 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_2^{(1)} + \theta_6 \cdot \sigma_1^{(7)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_2^{(1)} + \theta_7 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(1)} \sigma_2^{(0)} +$ $\theta_8 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_2^{(1)} + \theta_9 \cdot \sigma_2^{(4)} \sigma_1^{(3)}$
Projected ansatz $\pi_{\nu}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta}))$	$\theta_0 \cdot \sigma_2^{(5)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + \theta_1 \cdot \sigma_3^{(2)} \sigma_2^{(1)} \sigma_1^{(0)} + \theta_2 \cdot \sigma_1^{(3)} \sigma_1^{(1)} \sigma_2^{(0)} + \theta_3 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_2^{(1)} \sigma_3^{(0)} +$ $\theta_4 \cdot \sigma_3^{(4)} \sigma_2^{(3)} \sigma_2^{(1)} \sigma_1^{(0)} + \theta_5 \cdot \sigma_2^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} + \theta_6 \cdot \sigma_1^{(5)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} +$ $\theta_7 \cdot \sigma_2^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_2^{(0)} + \theta_8 \cdot \sigma_3^{(5)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} + \theta_9 \cdot \sigma_2^{(3)}$
Optimal parameters $\boldsymbol{\theta}_{\min}$	$(0.08371449607143293, -0.06731041226523458, 0.04057176605563251,$ $-0.06669446825280333, -0.04286798205858595, -0.06491571226470562,$ $-0.03555738912871124, -0.04236314034136127, -0.03391807633816377,$ $-0.02729133620630072)$
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.000398
VQE simulation absolute error $ \tilde{E}(\boldsymbol{\theta}_{\min}) - \epsilon_0 $ (Ha)	0.001369
Contextual subspace Hamiltonian	
Diagonal terms	
$-14.14372 \cdot \mathbb{1} - 0.606505 \cdot \sigma_3^{(5)} + 0.85606 \cdot \sigma_3^{(4)} - 0.180517 \cdot \sigma_3^{(5)} \sigma_3^{(4)} - 0.85606 \cdot \sigma_3^{(4)} \sigma_3^{(2)} + 0.180517 \cdot \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} -$ $0.389236 \cdot \sigma_3^{(2)} - 0.370128 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.075857 \cdot \sigma_3^{(3)} \sigma_3^{(2)} - 0.186586 \cdot \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.186586 \cdot \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} -$ $0.231058 \cdot \sigma_3^{(3)} + 0.101635 \cdot \sigma_3^{(5)} \sigma_3^{(3)} - 0.174306 \cdot \sigma_3^{(4)} \sigma_3^{(3)} + 0.174306 \cdot \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.071726 \cdot \sigma_3^{(5)} \sigma_3^{(2)} - 0.606501 \cdot$ $\sigma_3^{(2)} \sigma_3^{(1)} + 0.141107 \cdot \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} - 0.180517 \cdot \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} + 0.180517 \cdot \sigma_3^{(4)} \sigma_3^{(1)} + 0.087626 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_3^{(1)} + 0.132008 \cdot$ $\sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} + 0.370155 \cdot \sigma_3^{(1)} \sigma_3^{(0)} - 0.087634 \cdot \sigma_3^{(5)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.1866 \cdot \sigma_3^{(4)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.1866 \cdot \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.117288 \cdot$ $\sigma_3^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.080114 \cdot \sigma_3^{(3)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.075869 \cdot \sigma_3^{(2)} \sigma_3^{(0)} - 0.231061 \cdot \sigma_3^{(0)} + 0.132008 \cdot \sigma_3^{(5)} \sigma_3^{(0)} - 0.174306 \cdot$ $\sigma_3^{(4)} \sigma_3^{(0)} + 0.174306 \cdot \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.080104 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.133449 \cdot \sigma_3^{(3)} \sigma_3^{(0)} + 0.101635 \cdot \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} -$ $0.071734 \cdot \sigma_3^{(1)}$	
Off-diagonal terms	
$-0.041367 \cdot \sigma_1^{(5)} + 0.002593 \cdot \sigma_1^{(5)} \sigma_3^{(2)} - 0.012067 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} + 0.012067 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.041401 \cdot \sigma_1^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} +$ $0.013714 \cdot \sigma_1^{(5)} \sigma_3^{(3)} \sigma_3^{(1)} + 0.014339 \cdot \sigma_1^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.014043 \cdot \sigma_1^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.012125 \cdot \sigma_1^{(4)} \sigma_3^{(2)} - 0.012125 \cdot$ $\sigma_1^{(4)} + 0.01206 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} - 0.01206 \cdot \sigma_1^{(5)} \sigma_3^{(4)} - 0.002586 \cdot \sigma_1^{(5)} \sigma_3^{(3)} - 0.01371 \cdot \sigma_1^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} - 0.014329 \cdot \sigma_1^{(5)} \sigma_3^{(1)} \sigma_3^{(0)} -$ $0.014037 \cdot \sigma_1^{(5)} \sigma_3^{(0)} + 0.005481 \cdot \sigma_1^{(3)} + 0.004615 \cdot \sigma_3^{(5)} \sigma_1^{(3)} + 0.001345 \cdot \sigma_3^{(4)} \sigma_1^{(3)} - 0.001345 \cdot \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} - 0.005518 \cdot$ $\sigma_3^{(5)} \sigma_1^{(3)} \sigma_3^{(2)} + 0.016128 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} + 0.006916 \cdot \sigma_1^{(3)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.016627 \cdot \sigma_1^{(3)} \sigma_3^{(0)} - 0.051034 \cdot \sigma_2^{(5)} \sigma_2^{(3)} \sigma_3^{(2)} - 0.013717 \cdot$ $\sigma_2^{(5)} \sigma_2^{(3)} + 0.018978 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} - 0.018978 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} - 0.051039 \cdot \sigma_1^{(5)} \sigma_1^{(3)} \sigma_3^{(2)} - 0.013712 \cdot \sigma_1^{(5)} \sigma_1^{(3)} + \dots$	

Table 7: HF qubit tapering and 4-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(10)}\sigma_3^{(9)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(9)}\sigma_3^{(3)}, \sigma_3^{(10)}\sigma_3^{(4)}, \sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}$
Sector ν	$\{-1, 1, 1, -1\}$
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}, \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(3)}, \sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(1)}, \sigma_3^{(0)}$
Sector ν	$\{1, 1, -1, -1, 1, -1, -1\}$
Class observable $C(\mathbf{r})$	$-2.0406484584883426 \times 10^{-7} \cdot \sigma_1^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_1^{(4)} + -0.9999999999999791 \cdot \sigma_3^{(4)}$
4-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}, \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	$ 0111\rangle$
Minimal ansatz $A(\boldsymbol{\theta})$	$\theta_0 \cdot \sigma_2^{(7)}\sigma_1^{(4)}\sigma_1^{(2)}\sigma_1^{(1)} + \theta_1 \cdot \sigma_2^{(7)}\sigma_3^{(4)}\sigma_1^{(3)}\sigma_2^{(2)}\sigma_2^{(1)} + \theta_2 \cdot \sigma_2^{(7)}\sigma_3^{(4)}\sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + \theta_3 \cdot \sigma_2^{(7)}\sigma_1^{(4)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_2^{(0)}$
Projected ansatz $\pi_{\nu}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta}))$	$\theta_0 \cdot \sigma_2^{(3)}\sigma_3^{(2)}\sigma_1^{(1)} + \theta_1 \cdot \sigma_2^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + \theta_2 \cdot \sigma_1^{(2)}\sigma_2^{(0)} + \theta_3 \cdot \sigma_2^{(3)}\sigma_3^{(2)}\sigma_1^{(0)}$
Optimal parameters $\boldsymbol{\theta}_{\min}$	$(0.13371906159805624, -0.030698667705248102, -0.025990818646901812, -0.030230480241272343)$
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.000809
VQE simulation absolute error $ \tilde{E}(\boldsymbol{\theta}_{\min}) - \epsilon_0 $ (Ha)	0.001175
Contextual subspace Hamiltonian	
Diagonal terms	
$-100.750933 \cdot 1 - 0.293887 \cdot \sigma_3^{(3)}\sigma_3^{(2)} - 0.65827 \cdot \sigma_3^{(3)} + 1.079643 \cdot \sigma_3^{(2)} + 0.293887 \cdot \sigma_3^{(1)}\sigma_3^{(0)} - 0.176986 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} -$ $0.166912 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.168352 \cdot \sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.65827 \cdot \sigma_3^{(1)} - 0.166912 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} - 0.186642 \cdot \sigma_3^{(3)}\sigma_3^{(1)} +$ $0.185838 \cdot \sigma_3^{(2)}\sigma_3^{(1)} + 1.079643 \cdot \sigma_3^{(0)} - 0.168352 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(0)} - 0.185838 \cdot \sigma_3^{(3)}\sigma_3^{(0)} + 0.214583 \cdot \sigma_3^{(2)}\sigma_3^{(0)}$	
Off-diagonal terms	
$-0.070333 \cdot \sigma_1^{(3)}\sigma_3^{(2)} + 0.070333 \cdot \sigma_1^{(3)} - 0.004181 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.004181 \cdot \sigma_1^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.029375 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} +$ $0.029375 \cdot \sigma_1^{(3)}\sigma_3^{(1)} - 0.04514 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(0)} + 0.045139 \cdot \sigma_1^{(3)}\sigma_3^{(0)} + 0.022413 \cdot \sigma_1^{(2)} - 0.039882 \cdot \sigma_1^{(3)}\sigma_1^{(2)} + 0.022413 \cdot$ $\sigma_3^{(3)}\sigma_1^{(2)} + 0.016504 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.016207 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.016207 \cdot \sigma_1^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.010177 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(1)} +$ $0.008341 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)} + 0.008341 \cdot \sigma_1^{(3)}\sigma_3^{(1)} - 0.006887 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(0)} + 0.030279 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(0)} + 0.030279 \cdot \sigma_1^{(3)}\sigma_3^{(0)} +$ $0.039882 \cdot \sigma_2^{(3)}\sigma_2^{(2)} - 0.016504 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.010177 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)} + 0.006887 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(0)} - 0.070333 \cdot$ $\sigma_1^{(1)} + 0.047032 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(1)} + 0.004181 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(1)} - 0.047032 \cdot \sigma_1^{(3)}\sigma_1^{(1)} + 0.029375 \cdot \sigma_3^{(3)}\sigma_1^{(1)} - 0.04514 \cdot$ $\sigma_3^{(2)}\sigma_1^{(1)} + 0.070333 \cdot \sigma_1^{(1)}\sigma_3^{(0)} - 0.047032 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} - 0.004181 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.047032 \cdot \sigma_1^{(3)}\sigma_1^{(1)}\sigma_3^{(0)} -$ $0.029375 \cdot \sigma_3^{(3)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.04514 \cdot \sigma_3^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.023462 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(1)} - 0.013348 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(1)} - 0.013348 \cdot \sigma_1^{(2)}\sigma_1^{(1)} -$ $0.023462 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.013348 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.013348 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_3^{(1)} - 0.023462 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(1)} + 0.023462 \cdot$ $\sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(1)}\sigma_3^{(0)} + 0.022413 \cdot \sigma_1^{(0)} - 0.022413 \cdot \sigma_3^{(1)}\sigma_1^{(0)} + 0.013348 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.016207 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} -$ $0.013348 \cdot \sigma_1^{(3)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.008341 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_1^{(0)} - 0.030279 \cdot \sigma_3^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} - 0.013348 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(0)} - 0.016207 \cdot$ $\sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(0)} + 0.013348 \cdot \sigma_1^{(3)}\sigma_1^{(0)} - 0.008341 \cdot \sigma_3^{(3)}\sigma_1^{(0)} + 0.030279 \cdot \sigma_3^{(2)}\sigma_1^{(0)} - 0.013221 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} - 0.023831 \cdot$ $\sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} - 0.023831 \cdot \sigma_1^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.013221 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(0)} + 0.023831 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(0)} + 0.023831 \cdot \sigma_1^{(2)}\sigma_1^{(0)} +$ $0.013221 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} - 0.013221 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(0)} + 0.039882 \cdot \sigma_1^{(1)}\sigma_1^{(0)} - 0.023462 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.016504 \cdot$ $\sigma_3^{(3)}\sigma_3^{(1)}\sigma_1^{(0)} + 0.023462 \cdot \sigma_1^{(3)}\sigma_1^{(1)}\sigma_1^{(0)} - 0.010177 \cdot \sigma_3^{(3)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.006887 \cdot \sigma_3^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.039882 \cdot \sigma_2^{(1)}\sigma_2^{(0)} -$ $0.023462 \cdot \sigma_1^{(3)}\sigma_3^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.016504 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + 0.023462 \cdot \sigma_1^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} - 0.010177 \cdot \sigma_3^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} +$ $0.006887 \cdot \sigma_2^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} - 0.038601 \cdot \sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} - 0.013221 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} - 0.013221 \cdot \sigma_1^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} - 0.038601 \cdot$ $\sigma_1^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} - 0.013221 \cdot \sigma_3^{(3)}\sigma_1^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} - 0.013221 \cdot \sigma_1^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} + 0.038601 \cdot \sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + 0.038601 \cdot$ $\sigma_2^{(3)}\sigma_2^{(2)}\sigma_2^{(1)}\sigma_2^{(0)}$	

Table 8: BeH₂ qubit tapering and 7-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(13)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(9)}\sigma_3^{(5)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(10)}\sigma_3^{(3)}, \sigma_3^{(11)}\sigma_3^{(4)}, \sigma_3^{(13)}\sigma_3^{(9)}\sigma_3^{(6)}\sigma_3^{(2)},$ $\sigma_3^{(13)}\sigma_3^{(12)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}$
Sector ν	$\{-1, 1, 1, 1, -1\}$
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(8)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(7)}, \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(1)}\sigma_3^{(0)},$ $\sigma_3^{(4)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(2)}, \sigma_3^{(1)}, \sigma_3^{(0)}$
Sector ν	$\{1, 1, 1, 1, 1, 1, -1\}$
Class observable $C(\mathbf{r})$	$-4.6564989580718305 \times 10^{-8} \cdot \sigma_3^{(7)}\sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} +$ $0.9999999999999989 \cdot \sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}$
7-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(8)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(1)}\sigma_3^{(0)}, C(\mathbf{r})$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	$ 0111001\rangle$
Minimal ansatz $A(\theta)$	$\theta_0 \cdot \sigma_1^{(7)}\sigma_1^{(4)}\sigma_2^{(1)} + \theta_1 \cdot \sigma_2^{(7)}\sigma_3^{(4)}\sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)} + \theta_2 \cdot \sigma_1^{(8)}\sigma_3^{(7)}\sigma_1^{(4)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_2^{(0)} +$ $\theta_3 \cdot \sigma_3^{(8)}\sigma_1^{(7)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_1^{(1)}\sigma_1^{(0)} + \theta_4 \cdot \sigma_2^{(8)}\sigma_3^{(7)}\sigma_2^{(4)}\sigma_2^{(2)}\sigma_3^{(1)} + \theta_5 \cdot \sigma_3^{(8)}\sigma_2^{(5)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_2^{(0)} +$ $\theta_6 \cdot \sigma_3^{(8)}\sigma_1^{(7)}\sigma_2^{(4)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} + \theta_7 \cdot \sigma_1^{(8)}\sigma_3^{(7)}\sigma_1^{(3)}\sigma_3^{(2)}\sigma_2^{(1)} + \theta_8 \cdot \sigma_3^{(8)}\sigma_1^{(6)}\sigma_1^{(3)}\sigma_3^{(2)}\sigma_2^{(0)} +$ $\theta_9 \cdot \sigma_2^{(8)}\sigma_3^{(7)}\sigma_1^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_2^{(0)}$
Projected ansatz $\pi_{\nu}^{U_{\mathcal{F}}}(A(\theta))$	$\theta_0 \cdot \sigma_1^{(6)}\sigma_3^{(5)}\sigma_2^{(1)} + \theta_1 \cdot \sigma_1^{(6)}\sigma_3^{(5)}\sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)} + \theta_2 \cdot \sigma_3^{(6)}\sigma_2^{(5)}\sigma_1^{(4)}\sigma_2^{(3)}\sigma_2^{(2)}\sigma_1^{(0)} +$ $\theta_3 \cdot \sigma_2^{(6)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_2^{(1)}\sigma_2^{(0)} + \theta_4 \cdot \sigma_3^{(6)}\sigma_1^{(5)}\sigma_1^{(4)}\sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)} + \theta_5 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_2^{(0)} +$ $\theta_6 \cdot \sigma_1^{(6)}\sigma_3^{(5)}\sigma_2^{(2)}\sigma_1^{(0)} + \theta_7 \cdot \sigma_3^{(6)}\sigma_2^{(5)}\sigma_1^{(4)}\sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(0)} + \theta_8 \cdot \sigma_2^{(4)}\sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} +$ $\theta_9 \cdot \sigma_3^{(6)}\sigma_1^{(5)}\sigma_1^{(4)}\sigma_1^{(3)}\sigma_2^{(2)}\sigma_1^{(0)}$
Optimal parameters $\theta_{\min}$	$(0.07287579044669423, -0.046785323621881605, -0.04636678903287742,$ $0.05672217087752035, -0.039440022198649054, 0.05305862174001008,$ $-0.04097855862486507, -0.040729032236758826, 0.05230541408425366,$ $0.027001642003059267)$
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.000336
VQE simulation absolute error $ \tilde{E}(\theta_{\min}) - \epsilon_0 $ (Ha)	0.001027
Contextual subspace Hamiltonian	
Diagonal terms	
$-15.761613 \cdot \mathbb{1} + 0.784453 \cdot \sigma_3^{(5)} - 0.506908 \cdot \sigma_3^{(6)} - 0.074747 \cdot \sigma_3^{(6)}\sigma_3^{(5)} - 0.854426 \cdot \sigma_3^{(5)}\sigma_3^{(4)} - 0.192618 \cdot \sigma_3^{(4)} +$ $0.172118 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)} - 0.854426 \cdot \sigma_3^{(4)}\sigma_3^{(3)} - 0.192618 \cdot \sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)} + 0.172118 \cdot \sigma_3^{(6)}\sigma_3^{(4)}\sigma_3^{(3)} + 0.389236 \cdot \sigma_3^{(5)}\sigma_3^{(3)} +$ $0.251859 \cdot \sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.185458 \cdot \sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.082112 \cdot \sigma_3^{(6)}\sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.176368 \cdot \sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} -$ $0.176368 \cdot \sigma_3^{(4)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.230909 \cdot \sigma_3^{(6)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.094215 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.161912 \cdot \sigma_3^{(1)}\sigma_3^{(0)} + 0.173625 \cdot$ $\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.173625 \cdot \sigma_3^{(6)}\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} - 0.062831 \cdot \sigma_3^{(6)}\sigma_3^{(3)} - 0.784453 \cdot \sigma_3^{(2)} - 0.233727 \cdot$ $\sigma_3^{(5)}\sigma_3^{(2)} + 0.110127 \cdot \sigma_3^{(6)}\sigma_3^{(2)} + 0.192618 \cdot \sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(2)} + 0.192618 \cdot \sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)} - 0.226253 \cdot \sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} -$ $0.251859 \cdot \sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.102917 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.176368 \cdot \sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.176368 \cdot$ $\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.10835 \cdot \sigma_3^{(6)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} + 0.104132 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(2)} - 0.506908 \cdot \sigma_3^{(1)} - 0.110127 \cdot \sigma_3^{(5)}\sigma_3^{(1)} +$ $0.204496 \cdot \sigma_3^{(6)}\sigma_3^{(1)} + 0.172118 \cdot \sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(1)} + 0.172118 \cdot \sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(1)} - 0.102917 \cdot \sigma_3^{(3)}\sigma_3^{(0)} - 0.230909 \cdot \sigma_3^{(0)} + 0.200334 \cdot$ $\sigma_3^{(6)}\sigma_3^{(0)} + 0.074747 \cdot \sigma_3^{(2)}\sigma_3^{(1)} + 0.082112 \cdot \sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(0)} - 0.10835 \cdot \sigma_3^{(5)}\sigma_3^{(0)} + 0.173625 \cdot \sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(0)} + 0.173625 \cdot$ $\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(0)} - 0.104132 \cdot \sigma_3^{(3)}\sigma_3^{(1)} + 0.094215 \cdot \sigma_3^{(2)}\sigma_3^{(0)} + 0.062831 \cdot \sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}$	
Off-diagonal terms	
$-0.017976 \cdot \sigma_1^{(6)}\sigma_3^{(5)} + 0.011524 \cdot \sigma_1^{(6)}\sigma_3^{(5)}\sigma_3^{(3)} - 0.012748 \cdot \sigma_1^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)} + \dots$	

$$\begin{aligned}
& \dots - 0.014341 \cdot \sigma_1^{(6)} \sigma_2^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.014341 \cdot \sigma_2^{(6)} \sigma_1^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} + 0.014341 \cdot \\
& \sigma_1^{(6)} \sigma_2^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_2^{(0)} + 0.014341 \cdot \sigma_2^{(6)} \sigma_1^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_2^{(0)} - 0.011348 \cdot \sigma_3^{(3)} \sigma_2^{(2)} \sigma_2^{(0)} - 0.011348 \cdot \\
& \sigma_3^{(3)} \sigma_1^{(2)} \sigma_1^{(0)} - 0.01528 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_3^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.01528 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_2^{(2)} \sigma_2^{(0)} + 0.01528 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_2^{(0)} - \\
& 0.01528 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(2)} \sigma_1^{(0)} + 0.014134 \cdot \sigma_2^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.014134 \cdot \sigma_2^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_2^{(0)} - 0.011348 \cdot \\
& \sigma_3^{(6)} \sigma_2^{(2)} \sigma_2^{(0)} - 0.011348 \cdot \sigma_3^{(6)} \sigma_1^{(2)} \sigma_1^{(0)} - 0.011524 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.004237 \cdot \sigma_2^{(1)} \sigma_2^{(0)} - 0.017976 \cdot \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - \\
& 0.012349 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.012748 \cdot \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.012748 \cdot \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.010104 \cdot \\
& \sigma_3^{(5)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_2^{(0)} - 0.017976 \cdot \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.002478 \cdot \sigma_3^{(6)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_2^{(0)} - 0.002478 \cdot \sigma_3^{(6)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - \\
& 0.011524 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.004237 \cdot \sigma_1^{(1)} \sigma_1^{(0)} - 0.012349 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.012748 \cdot \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + \\
& 0.012748 \cdot \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.010104 \cdot \sigma_3^{(5)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.038422 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.038422 \cdot \\
& \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.018018 \cdot \sigma_3^{(6)} \sigma_2^{(5)} \sigma_1^{(4)} \sigma_2^{(3)} \sigma_2^{(2)} \sigma_1^{(1)} \sigma_2^{(0)} + 0.018018 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.018018 \cdot \\
& \sigma_3^{(6)} \sigma_2^{(5)} \sigma_1^{(4)} \sigma_2^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.018018 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.020538 \cdot \sigma_1^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.0178 \cdot \sigma_2^{(2)} \sigma_2^{(1)} \sigma_1^{(0)} + \\
& 0.002738 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_2^{(2)} \sigma_1^{(1)} \sigma_2^{(0)} + 0.002738 \cdot \sigma_2^{(2)} \sigma_1^{(1)} \sigma_2^{(0)} + 0.0178 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_2^{(2)} \sigma_2^{(1)} \sigma_1^{(0)} + 0.020538 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} \sigma_1^{(0)}
\end{aligned}$$

Table 9: H₂O qubit tapering and 7-qubit CS-VQE parameters

Tapering parameters			
Symmetry generators $\mathcal{G}$	$\sigma_3^{(13)}\sigma_3^{(11)}\sigma_3^{(9)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(1)}\sigma_3^{(0)},$ $\sigma_3^{(13)}\sigma_3^{(12)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}$	$\sigma_3^{(11)}\sigma_3^{(4)},$	$\sigma_3^{(13)}\sigma_3^{(9)}\sigma_3^{(6)}\sigma_3^{(2)},$
Sector ν	{-1, 1, 1, -1}		
Noncontextual model			
Symmetry generators $\mathcal{G}$	$\sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)},$ $\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(1)}, \sigma_3^{(0)}$	$\sigma_3^{(8)}\sigma_3^{(6)}\sigma_3^{(4)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)},$ $\sigma_3^{(7)}, \sigma_3^{(6)}, \sigma_3^{(5)}, \sigma_3^{(4)},$	
Sector ν	{1, 1, -1, -1, -1, -1, 1, -1, -1}		
Class observable $C(\mathbf{r})$	$9.459071266871024 \times 10^{-7} \cdot \sigma_3^{(7)}\sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} + -0.9999999999995526 \cdot \sigma_3^{(9)}\sigma_3^{(5)}\sigma_3^{(3)}$		
7-qubit CS-VQE parameters			
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(8)}\sigma_3^{(6)}\sigma_3^{(4)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_3^{(0)}, \sigma_3^{(7)}$		
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	1111111)		
Minimal ansatz $A(\boldsymbol{\theta})$	$\theta_0 \cdot \sigma_1^{(9)}\sigma_3^{(8)}\sigma_3^{(6)}\sigma_1^{(5)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} + \theta_1 \cdot \sigma_3^{(9)}\sigma_2^{(8)}\sigma_2^{(6)}\sigma_3^{(5)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_3^{(1)} +$ $\theta_2 \cdot \sigma_1^{(9)}\sigma_3^{(8)}\sigma_3^{(6)}\sigma_2^{(5)}\sigma_1^{(2)}\sigma_1^{(1)} + \theta_3 \cdot \sigma_3^{(9)}\sigma_2^{(8)}\sigma_3^{(6)}\sigma_2^{(4)}\sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} +$ $\theta_4 \cdot \sigma_2^{(9)}\sigma_3^{(8)}\sigma_2^{(6)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_1^{(1)} + \theta_5 \cdot \sigma_1^{(9)}\sigma_3^{(8)}\sigma_3^{(6)}\sigma_3^{(4)}\sigma_1^{(2)}\sigma_3^{(3)}\sigma_3^{(1)} +$ $\theta_6 \cdot \sigma_2^{(8)}\sigma_3^{(6)}\sigma_2^{(5)}\sigma_2^{(2)}\sigma_3^{(1)} + \theta_7 \cdot \sigma_1^{(9)}\sigma_3^{(8)}\sigma_3^{(6)}\sigma_2^{(5)}\sigma_3^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} +$ $\theta_8 \cdot \sigma_1^{(8)}\sigma_2^{(6)}\sigma_2^{(2)}\sigma_2^{(1)} + \theta_9 \cdot \sigma_3^{(9)}\sigma_2^{(8)}\sigma_1^{(6)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} +$ $\theta_{10} \cdot \sigma_1^{(8)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_2^{(4)}\sigma_2^{(2)}\sigma_2^{(1)} + \theta_{11} \cdot \sigma_2^{(9)}\sigma_3^{(8)}\sigma_2^{(6)}\sigma_2^{(2)}\sigma_3^{(1)} +$ $\theta_{12} \cdot \sigma_1^{(8)}\sigma_3^{(6)}\sigma_2^{(5)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_2^{(1)} + \theta_{13} \cdot \sigma_3^{(9)}\sigma_1^{(8)}\sigma_3^{(6)}\sigma_1^{(5)}\sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} +$ $\theta_{14} \cdot \sigma_1^{(9)}\sigma_3^{(8)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_2^{(4)}\sigma_1^{(2)}\sigma_3^{(1)} + \theta_{15} \cdot \sigma_1^{(9)}\sigma_3^{(8)}\sigma_3^{(6)}\sigma_1^{(4)}\sigma_1^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} +$ $\theta_{16} \cdot \sigma_3^{(9)}\sigma_3^{(5)}\sigma_1^{(3)}\sigma_1^{(2)}\sigma_2^{(0)} + \theta_{17} \cdot \sigma_2^{(9)}\sigma_1^{(8)}\sigma_1^{(6)}\sigma_1^{(5)} + \theta_{18} \cdot \sigma_2^{(8)}\sigma_3^{(7)}\sigma_1^{(6)} +$ $\theta_{19} \cdot \sigma_2^{(9)}\sigma_1^{(8)}\sigma_2^{(5)}\sigma_2^{(4)}$		
Projected ansatz $\pi_{\nu'}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta}))$	$\theta_0 \cdot \sigma_1^{(5)}\sigma_3^{(4)}\sigma_2^{(3)}\sigma_3^{(0)} + \theta_1 \cdot \sigma_1^{(6)}\sigma_3^{(4)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_3^{(0)} + \theta_2 \cdot \sigma_2^{(5)}\sigma_3^{(4)}\sigma_1^{(1)} +$ $\theta_3 \cdot \sigma_3^{(5)}\sigma_1^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_3^{(1)}\sigma_2^{(0)} + \theta_4 \cdot \sigma_1^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_2^{(1)}\sigma_3^{(0)} +$ $\theta_5 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_1^{(4)}\sigma_2^{(3)}\sigma_3^{(2)}\sigma_3^{(0)} + \theta_6 \cdot \sigma_2^{(5)}\sigma_3^{(4)}\sigma_2^{(3)}\sigma_2^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} +$ $\theta_7 \cdot \sigma_2^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(2)}\sigma_1^{(0)} + \theta_8 \cdot \sigma_2^{(6)}\sigma_1^{(1)}\sigma_3^{(0)} + \theta_9 \cdot \sigma_2^{(6)}\sigma_3^{(4)}\sigma_3^{(3)}\sigma_1^{(0)} +$ $\theta_{10} \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_2^{(4)}\sigma_1^{(1)}\sigma_3^{(0)} + \theta_{11} \cdot \sigma_1^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_2^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_3^{(0)} +$ $\theta_{12} \cdot \sigma_3^{(6)}\sigma_2^{(5)}\sigma_2^{(2)}\sigma_2^{(1)}\sigma_3^{(0)} + \theta_{13} \cdot \sigma_3^{(6)}\sigma_2^{(5)}\sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + \theta_{14} \cdot \sigma_3^{(5)}\sigma_1^{(4)}\sigma_2^{(3)}\sigma_1^{(2)} +$ $\theta_{15} \cdot \sigma_2^{(4)}\sigma_3^{(3)}\sigma_1^{(2)}\sigma_3^{(1)}\sigma_1^{(0)} + \theta_{16} \cdot \sigma_1^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} + \theta_{17} \cdot \sigma_1^{(6)}\sigma_2^{(5)}\sigma_3^{(2)} + \theta_{18} \cdot \sigma_2^{(6)}\sigma_3^{(4)}\sigma_3^{(2)} +$ $\theta_{19} \cdot \sigma_1^{(5)}\sigma_2^{(4)}\sigma_3^{(2)}$		
Optimal parameters $\boldsymbol{\theta}_{\min}$	(-0.07853011888280402, -0.04683529332762384, -0.046619063390598, 0.03246200088459131, 0.04188972292860196, 0.031409352807453134, -0.04191755758497103, 0.03149572165094569, 0.045258021849913996, 0.03102875503041955, 0.03197117303639109, 0.03440982346866502, -0.032543555879139346, -0.0203806172329641, -0.020521970546116004, -0.016734487124074264, 0.010826771365931482, -0.01247856503788453, -0.014772732750533521, 0.009661953139303407)		
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.001138		
VQE simulation absolute error $ \tilde{E}(\boldsymbol{\theta}_{\min}) - \epsilon_0 $ (Ha)	0.001593		
Contextual subspace Hamiltonian			
Diagonal terms			
$-80.703433 \cdot \mathbb{1} + 0.068284 \cdot \sigma_3^{(5)}\sigma_3^{(2)} - 0.099028 \cdot \sigma_3^{(6)}\sigma_3^{(4)}\sigma_3^{(2)} - 0.11275 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)} + 0.472957 \cdot \sigma_3^{(6)} + 0.135218 \cdot$ $\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(2)} - 0.120353 \cdot \sigma_3^{(4)}\sigma_3^{(2)} + 0.508045 \cdot \sigma_3^{(5)} + 0.229846 \cdot \sigma_3^{(2)} - 0.125217 \cdot \sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(2)} + \dots$			

$$\begin{aligned}
& \cdots + 0.138205 \cdot \sigma_3^{(6)} \sigma_3^{(5)} + 0.832573 \cdot \sigma_3^{(4)} + 0.140809 \cdot \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} - 0.128083 \cdot \sigma_3^{(6)} \sigma_3^{(2)} + 0.138156 \cdot \sigma_3^{(6)} \sigma_3^{(4)} + 0.12528 \cdot \\
& \sigma_3^{(5)} \sigma_3^{(4)} + 0.068284 \cdot \sigma_3^{(3)} \sigma_3^{(2)} + 0.313723 \cdot \sigma_3^{(5)} \sigma_3^{(3)} - 0.141657 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} + 0.152393 \cdot \sigma_3^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.305894 \cdot \\
& \sigma_3^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.508045 \cdot \sigma_3^{(3)} - 0.142986 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.149934 \cdot \sigma_3^{(6)} \sigma_3^{(3)} + 0.15624 \cdot \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.161517 \cdot \\
& \sigma_3^{(4)} \sigma_3^{(3)} - 0.099028 \cdot \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.141657 \cdot \sigma_3^{(5)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.149265 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.137214 \cdot \sigma_3^{(6)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - \\
& 0.142986 \cdot \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.153593 \cdot \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.11275 \cdot \sigma_3^{(3)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.125217 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.472957 \cdot \\
& \sigma_3^{(1)} + 0.152393 \cdot \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} - 0.137214 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} + 0.196529 \cdot \sigma_3^{(6)} \sigma_3^{(1)} + 0.149934 \cdot \sigma_3^{(5)} \sigma_3^{(1)} + 0.16911 \cdot \\
& \sigma_3^{(4)} \sigma_3^{(1)} + 0.135218 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} + 0.138205 \cdot \sigma_3^{(3)} \sigma_3^{(1)} - 0.120353 \cdot \sigma_3^{(2)} \sigma_3^{(0)} + 0.832573 \cdot \sigma_3^{(0)} + 0.15624 \cdot \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(0)} - \\
& 0.153593 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.16911 \cdot \sigma_3^{(6)} \sigma_3^{(0)} + 0.161517 \cdot \sigma_3^{(5)} \sigma_3^{(0)} + 0.182113 \cdot \sigma_3^{(4)} \sigma_3^{(0)} + 0.140809 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + \\
& 0.12528 \cdot \sigma_3^{(3)} \sigma_3^{(0)} - 0.128083 \cdot \sigma_3^{(2)} \sigma_3^{(1)} + 0.138156 \cdot \sigma_3^{(1)} \sigma_3^{(0)}
\end{aligned}$$

Off-diagonal terms

$$\begin{aligned}
& 0.102823 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.000753 \cdot \sigma_3^{(5)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.009429 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} + 0.040233 \cdot \sigma_3^{(6)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - \\
& 0.000752 \cdot \sigma_3^{(5)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.102822 \cdot \sigma_1^{(3)} \sigma_3^{(0)} + 0.009429 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.040232 \cdot \sigma_3^{(6)} \sigma_1^{(3)} \sigma_3^{(0)} + 0.034436 \cdot \\
& \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.034436 \cdot \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.025624 \cdot \sigma_1^{(3)} \sigma_3^{(1)} + 0.025624 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.039505 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - \\
& 0.039505 \cdot \sigma_1^{(3)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.024454 \cdot \sigma_1^{(3)} \sigma_3^{(2)} - 0.024454 \cdot \sigma_1^{(3)} - 0.054563 \cdot \sigma_1^{(6)} - 0.024745 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(2)} - 0.054563 \cdot \\
& \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(2)} - 0.024745 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} - 0.017815 \cdot \sigma_1^{(6)} \sigma_3^{(5)} - 0.017815 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} - 0.018426 \cdot \sigma_1^{(6)} \sigma_3^{(4)} - 0.018426 \cdot \\
& \sigma_1^{(6)} \sigma_3^{(2)} - 0.019338 \cdot \sigma_1^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.010311 \cdot \sigma_1^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} - 0.019338 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.010311 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} + \\
& 0.019338 \cdot \sigma_1^{(6)} \sigma_3^{(3)} \sigma_3^{(0)} - 0.010736 \cdot \sigma_1^{(6)} \sigma_3^{(3)} + 0.019338 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.010736 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + \\
& 0.011074 \cdot \sigma_1^{(6)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.011074 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.030101 \cdot \sigma_1^{(6)} \sigma_3^{(1)} - 0.030101 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} - 0.023615 \cdot \\
& \sigma_1^{(6)} \sigma_3^{(0)} - 0.023615 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.039505 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} + 0.102823 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} - 0.025624 \cdot \sigma_3^{(6)} \sigma_1^{(5)} - \\
& 0.039505 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_3^{(4)} - 0.102823 \cdot \sigma_1^{(5)} \sigma_3^{(4)} + 0.025624 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_3^{(2)} + 0.024454 \cdot \sigma_1^{(5)} \sigma_3^{(2)} - 0.024454 \cdot \sigma_1^{(5)} + \\
& 0.076048 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(0)} + 0.000752 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} - 0.076048 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.000752 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} - \\
& 0.009429 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.009429 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.040232 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} - 0.040232 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(1)} + \\
& 0.034436 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.034436 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(0)} - 0.00757 \cdot \sigma_1^{(6)} \sigma_1^{(5)} - 0.000737 \cdot \sigma_2^{(6)} \sigma_2^{(5)} \sigma_3^{(4)} - 0.008307 \cdot \\
& \sigma_1^{(6)} \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(2)} + 0.008307 \cdot \sigma_1^{(6)} \sigma_1^{(5)} \sigma_3^{(2)} + 0.000737 \cdot \sigma_2^{(6)} \sigma_2^{(5)} + 0.00757 \cdot \sigma_1^{(6)} \sigma_1^{(5)} \sigma_3^{(4)} - 0.074819 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} - \\
& 0.029127 \cdot \sigma_3^{(6)} \sigma_1^{(4)} - 0.029889 \cdot \sigma_3^{(5)} \sigma_1^{(4)} - 0.074819 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} - 0.019845 \cdot \sigma_3^{(6)} \sigma_1^{(4)} \sigma_3^{(2)} - 0.029127 \cdot \sigma_1^{(4)} \sigma_3^{(2)} - \\
& 0.029889 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} - 0.019845 \cdot \sigma_1^{(4)} + 0.019261 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.018995 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} - \\
& 0.019261 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.017346 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.019261 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} - 0.018995 \cdot \\
& \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} - 0.019261 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.017346 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} + 0.024135 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} \sigma_1^{(3)} \sigma_3^{(0)} + \\
& 0.024135 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.022362 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(1)} - 0.022362 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} \sigma_3^{(1)} - 0.035773 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(0)} - \\
& 0.035773 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.003167 \cdot \sigma_1^{(6)} \sigma_1^{(4)} + 0.037857 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} + 0.037857 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} - 0.000729 \cdot \\
& \sigma_1^{(6)} \sigma_1^{(4)} \sigma_3^{(2)} + 0.000491 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(2)} + 0.003167 \cdot \sigma_2^{(6)} \sigma_2^{(4)} - 0.000729 \cdot \sigma_2^{(6)} \sigma_2^{(4)} \sigma_3^{(2)} + 0.000491 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} - \\
& 0.01895 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.001756 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} + 0.01895 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.003445 \cdot \\
& \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} - 0.01895 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_1^{(3)} \sigma_3^{(0)} - 0.001756 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} + 0.01895 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + \\
& 0.003445 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} + 0.014761 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} \sigma_1^{(3)} \sigma_3^{(0)} + 0.014761 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(2)} \sigma_1^{(3)} \sigma_3^{(0)} + 0.02616 \cdot \\
& \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(1)} + 0.02616 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(1)} - 0.001269 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(0)} - 0.001269 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(0)} + 0.015938 \cdot \\
& \sigma_1^{(5)} \sigma_1^{(4)} - 0.005541 \cdot \sigma_3^{(6)} \sigma_2^{(5)} \sigma_2^{(4)} - 0.010398 \cdot \sigma_3^{(6)} \sigma_2^{(5)} \sigma_2^{(4)} \sigma_3^{(2)} + 0.010398 \cdot \sigma_2^{(5)} \sigma_2^{(4)} \sigma_3^{(2)} + 0.005541 \cdot \sigma_2^{(5)} \sigma_2^{(4)} - \\
& 0.015938 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_1^{(4)} - 0.010751 \cdot \sigma_2^{(6)} \sigma_1^{(5)} \sigma_2^{(4)} \sigma_3^{(2)} + 0.000302 \cdot \sigma_1^{(6)} \sigma_2^{(5)} \sigma_2^{(4)} + 0.01045 \cdot \sigma_1^{(6)} \sigma_2^{(5)} \sigma_2^{(4)} \sigma_3^{(2)} + \\
& 0.01045 \cdot \sigma_2^{(6)} \sigma_2^{(5)} \sigma_1^{(4)} + 0.000302 \cdot \sigma_2^{(6)} \sigma_2^{(5)} \sigma_1^{(4)} \sigma_3^{(2)} + 0.010751 \cdot \sigma_1^{(6)} \sigma_1^{(5)} \sigma_1^{(4)} - 0.016195 \cdot \sigma_2^{(3)} \sigma_2^{(2)} - 0.016195 \cdot \\
& \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_2^{(3)} \sigma_2^{(2)} - 0.028907 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} + 0.016195 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.028907 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} + \\
& 0.016195 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.011031 \cdot \sigma_2^{(6)} \sigma_3^{(4)} \sigma_2^{(3)} \sigma_1^{(2)} + 0.011031 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_2^{(2)} + \\
& 0.014434 \cdot \sigma_2^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} - 0.011031 \cdot \sigma_2^{(6)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.014434 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} - \\
& 0.011031 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.017769 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_2^{(3)} \sigma_2^{(2)} + 0.017769 \cdot \sigma_2^{(5)} \sigma_3^{(4)} \sigma_2^{(3)} \sigma_1^{(2)} - 0.016195 \cdot \\
& \sigma_2^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} + 0.017769 \cdot \sigma_2^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.016195 \cdot \sigma_3^{(6)} \sigma_1^{(5)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} - 0.017769 \cdot \\
& \sigma_3^{(6)} \sigma_1^{(5)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(3)} \sigma_3^{(0)} - 0.007079 \cdot \sigma_1^{(6)} \sigma_1^{(5)} \sigma_2^{(3)} \sigma_2^{(2)} - 0.007079 \cdot \sigma_2^{(6)} \sigma_2^{(5)} \sigma_2^{(3)} \sigma_2^{(2)} - 0.011768 \cdot \sigma_2^{(6)} \sigma_2^{(5)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} + \\
& 0.007079 \cdot \sigma_2^{(6)} \sigma_2^{(5)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.011768 \cdot \sigma_1^{(6)} \sigma_1^{(5)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} + 0.007079 \cdot \sigma_1^{(6)} \sigma_1^{(5)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.008863 \cdot \\
& \sigma_3^{(5)} \sigma_2^{(4)} \sigma_2^{(3)} \sigma_1^{(2)} + 0.008863 \cdot \sigma_1^{(4)} \sigma_2^{(3)} \sigma_2^{(2)} + 0.002499 \cdot \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} - 0.008863 \cdot \sigma_3^{(5)} \sigma_2^{(4)} \sigma_1^{(3)} \sigma_2^{(2)} \sigma_1^{(3)} \sigma_3^{(0)} + \\
& 0.002499 \cdot \sigma_1^{(4)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} - 0.008863 \cdot \sigma_1^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.010132 \cdot \sigma_1^{(5)} \sigma_1^{(4)} \sigma_2^{(3)} \sigma_2^{(2)} - 0.010132 \cdot \sigma_2^{(5)} \sigma_2^{(4)} \sigma_2^{(3)} \sigma_2^{(2)} - \\
& 0.024802 \cdot \sigma_1^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} + 0.010131 \cdot \sigma_1^{(5)} \sigma_1^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.024802 \cdot \sigma_2^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} + 0.010131 \cdot \\
& \sigma_2^{(5)} \sigma_2^{(4)} \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.028907 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} \sigma_3^{(2)} + 0.014434 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(2)} + \dots
\end{aligned}$$

$$\begin{aligned}
& \dots - 0.015252 \cdot \sigma_1^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - 0.015252 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_1^{(0)} - 0.019261 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} + 0.019261 \cdot \\
& \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - 0.019261 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} + 0.019261 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_1^{(0)} - 0.02551 \cdot \\
& \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.02551 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.02551 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - 0.02551 \cdot \\
& \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_1^{(0)} - 0.005188 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.005188 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.005188 \cdot \\
& \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - 0.005188 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} + 0.015938 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - 0.015938 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - \\
& 0.01045 \cdot \sigma_2^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_2^{(0)} + 0.003167 \cdot \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.000302 \cdot \sigma_2^{(3)} \sigma_1^{(1)} \sigma_2^{(0)} - 0.000729 \cdot \sigma_1^{(1)} \sigma_1^{(0)} - 0.001756 \cdot \\
& \sigma_3^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.003445 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.037857 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.014761 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} + \\
& 0.02616 \cdot \sigma_3^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.001269 \cdot \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.037857 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.000491 \cdot \\
& \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.000302 \cdot \sigma_2^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_1^{(0)} + 0.003167 \cdot \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.01045 \cdot \sigma_2^{(3)} \sigma_2^{(1)} \sigma_1^{(0)} - 0.000729 \cdot \sigma_2^{(1)} \sigma_2^{(0)} - \\
& 0.001756 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.003445 \cdot \sigma_3^{(5)} \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.014761 \cdot \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.02616 \cdot \\
& \sigma_3^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.001269 \cdot \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.000491 \cdot \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.007527 \cdot \sigma_1^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - \\
& 0.007527 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.007527 \cdot \sigma_1^{(6)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.007527 \cdot \sigma_1^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.01895 \cdot \\
& \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.01895 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.01895 \cdot \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.01895 \cdot \\
& \sigma_1^{(5)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.005188 \cdot \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.005188 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.005188 \cdot \\
& \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.005188 \cdot \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.030954 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.030954 \cdot \\
& \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.030954 \cdot \sigma_1^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.030954 \cdot \sigma_2^{(6)} \sigma_3^{(5)} \sigma_2^{(4)} \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + \\
& 0.010751 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.010751 \cdot \sigma_1^{(3)} \sigma_1^{(1)} \sigma_1^{(0)}
\end{aligned}$$

Table 10: F₂ qubit tapering and 10-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(16)}\sigma_3^{(15)}\sigma_3^{(14)}\sigma_3^{(12)}\sigma_3^{(11)}\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(3)}\sigma_3^{(0)},$ $\sigma_3^{(18)}\sigma_3^{(17)}\sigma_3^{(16)}\sigma_3^{(12)}\sigma_3^{(11)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(2)}\sigma_3^{(1)}, \sigma_3^{(18)}\sigma_3^{(17)}\sigma_3^{(15)}\sigma_3^{(14)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(4)},$ $\sigma_3^{(19)}\sigma_3^{(18)}\sigma_3^{(17)}\sigma_3^{(16)}\sigma_3^{(15)}\sigma_3^{(14)}\sigma_3^{(13)}\sigma_3^{(12)}\sigma_3^{(11)}\sigma_3^{(10)}$
Sector ν	{-1, 1, 1, -1}
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(15)}, \sigma_3^{(14)}\sigma_3^{(13)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(2)}, \sigma_3^{(13)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(2)}, \sigma_3^{(12)}, \sigma_3^{(11)},$ $\sigma_3^{(10)}, \sigma_3^{(9)}, \sigma_3^{(8)}, \sigma_3^{(6)}, \sigma_3^{(5)}, \sigma_3^{(4)}, \sigma_3^{(3)}, \sigma_3^{(2)}, \sigma_3^{(1)}, \sigma_3^{(0)}$
Sector ν	{1, -1, 1, -1, -1, -1, -1, -1, 1, -1, -1, -1, -1, -1}
Class observable $C(\mathbf{r})$	$-1.0 \cdot \sigma_3^{(7)} + -5.3197323983259087 \times 10^{-11} \cdot \sigma_1^{(7)}$
10-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(14)}\sigma_3^{(13)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(2)}, \sigma_3^{(11)}, \sigma_3^{(10)}, \sigma_3^{(5)}, \sigma_3^{(2)}, C(\mathbf{r})$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	0111101111⟩
Minimal ansatz $A(\theta)$	$\theta_0 \cdot \sigma_2^{(15)}\sigma_3^{(14)}\sigma_3^{(13)}\sigma_1^{(12)}\sigma_1^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_1^{(3)} +$ $\theta_1 \cdot \sigma_1^{(15)}\sigma_3^{(14)}\sigma_3^{(13)}\sigma_3^{(12)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_3^{(9)}\sigma_1^{(8)}\sigma_1^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_2^{(3)} +$ $\theta_2 \cdot \sigma_2^{(15)}\sigma_1^{(12)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_1^{(6)}\sigma_3^{(3)}\sigma_1^{(1)}\sigma_1^{(0)} + \theta_3 \cdot \sigma_2^{(15)}\sigma_3^{(12)}\sigma_1^{(9)}\sigma_2^{(6)}\sigma_3^{(3)}\sigma_2^{(1)} +$ $\theta_4 \cdot \sigma_2^{(15)}\sigma_3^{(12)}\sigma_3^{(9)}\sigma_1^{(8)}\sigma_2^{(6)}\sigma_3^{(3)}\sigma_2^{(1)}\sigma_2^{(0)} + \theta_5 \cdot \sigma_2^{(15)}\sigma_3^{(14)}\sigma_1^{(13)}\sigma_2^{(6)}\sigma_3^{(5)}\sigma_2^{(4)} +$ $\theta_6 \cdot \sigma_2^{(15)}\sigma_3^{(14)}\sigma_3^{(13)}\sigma_3^{(12)}\sigma_3^{(11)}\sigma_3^{(10)}\sigma_1^{(9)}$
Projected ansatz $\pi_{\nu}^{U_{\mathcal{F}}}(A(\theta))$	$\theta_0 \cdot \sigma_2^{(9)}\sigma_1^{(7)}\sigma_1^{(4)}\sigma_1^{(2)} + \theta_1 \cdot \sigma_1^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_1^{(5)}\sigma_1^{(4)}\sigma_2^{(2)} + \theta_2 \cdot \sigma_2^{(9)}\sigma_1^{(7)}\sigma_1^{(4)}\sigma_3^{(2)}\sigma_1^{(0)} +$ $\theta_3 \cdot \sigma_2^{(9)}\sigma_3^{(7)}\sigma_1^{(6)}\sigma_2^{(4)}\sigma_3^{(2)}\sigma_2^{(1)} + \theta_4 \cdot \sigma_2^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_1^{(5)}\sigma_2^{(4)}\sigma_3^{(2)}\sigma_1^{(0)} + \theta_5 \cdot \sigma_2^{(9)}\sigma_2^{(4)}\sigma_2^{(3)} +$ $\theta_6 \cdot \sigma_2^{(9)}\sigma_3^{(7)}\sigma_1^{(6)}\sigma_3^{(3)}$
Optimal parameters $\theta_{\min}$	(0.2353125578050386, -0.033111405385919786, 0.035257482489025, 0.02552471469582267, 0.02017574021677397, -0.03218643232070007, 0.017530009168854648)
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.000858
VQE simulation absolute error $ \tilde{E}(\theta_{\min}) - \epsilon_0 $ (Ha)	0.001593
Contextual subspace Hamiltonian	
Diagonal terms	
$-174.844449 \cdot \mathbb{1} + 1.690487 \cdot \sigma_3^{(9)} + 1.561318 \cdot \sigma_3^{(8)} + 0.158174 \cdot \sigma_3^{(9)}\sigma_3^{(8)} - 1.589989 \cdot \sigma_3^{(8)}\sigma_3^{(3)} - 0.158174 \cdot \sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(3)} +$ $1.60047 \cdot \sigma_3^{(3)} + 1.801314 \cdot \sigma_3^{(7)} + 0.108278 \cdot \sigma_3^{(9)}\sigma_3^{(7)} + 0.148899 \cdot \sigma_3^{(8)}\sigma_3^{(7)} - 0.148899 \cdot \sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(3)} + 2.09314 \cdot \sigma_3^{(6)} +$ $0.140245 \cdot \sigma_3^{(9)}\sigma_3^{(6)} + 0.132255 \cdot \sigma_3^{(8)}\sigma_3^{(6)} - 0.132255 \cdot \sigma_3^{(8)}\sigma_3^{(6)}\sigma_3^{(3)} + 0.124369 \cdot \sigma_3^{(7)}\sigma_3^{(6)} + 2.261138 \cdot \sigma_3^{(5)} + 0.15131 \cdot$ $\sigma_3^{(9)}\sigma_3^{(5)} + 0.141413 \cdot \sigma_3^{(8)}\sigma_3^{(5)} - 0.141413 \cdot \sigma_3^{(8)}\sigma_3^{(5)}\sigma_3^{(3)} + 0.133485 \cdot \sigma_3^{(7)}\sigma_3^{(5)} + 0.098873 \cdot \sigma_3^{(6)}\sigma_3^{(5)} - 15.824225 \cdot$ $\sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)} - 0.195553 \cdot \sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(3)} - 0.194514 \cdot \sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(3)} - 0.19781 \cdot \sigma_3^{(9)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)} + 0.202966 \cdot$ $\sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)} - 0.202966 \cdot \sigma_3^{(9)}\sigma_3^{(8)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)} - 0.219559 \cdot \sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)} + 1.69386 \cdot \sigma_3^{(4)} + 0.193504 \cdot$ $\sigma_3^{(9)}\sigma_3^{(4)} + 0.166728 \cdot \sigma_3^{(8)}\sigma_3^{(4)} - 0.166728 \cdot \sigma_3^{(8)}\sigma_3^{(4)}\sigma_3^{(3)} + 0.176053 \cdot \sigma_3^{(7)}\sigma_3^{(4)} + 0.165246 \cdot \sigma_3^{(6)}\sigma_3^{(4)} + 0.17606 \cdot \sigma_3^{(5)}\sigma_3^{(4)} -$ $0.224844 \cdot \sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)} + 0.166728 \cdot \sigma_3^{(9)}\sigma_3^{(3)} + 0.156437 \cdot \sigma_3^{(7)}\sigma_3^{(3)} + 0.159354 \cdot \sigma_3^{(6)}\sigma_3^{(3)} + 0.161979 \cdot \sigma_3^{(5)}\sigma_3^{(3)} -$ $0.206889 \cdot \sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)} + 0.158174 \cdot \sigma_3^{(4)}\sigma_3^{(3)} - 0.194514 \cdot \sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)} + 1.805098 \cdot \sigma_3^{(2)} + 0.176053 \cdot \sigma_3^{(9)}\sigma_3^{(2)} +$ $0.156437 \cdot \sigma_3^{(8)}\sigma_3^{(2)} - 0.156437 \cdot \sigma_3^{(8)}\sigma_3^{(2)}\sigma_3^{(3)} + 0.171221 \cdot \sigma_3^{(7)}\sigma_3^{(2)} + 0.159571 \cdot \sigma_3^{(6)}\sigma_3^{(2)} + 0.160439 \cdot \sigma_3^{(5)}\sigma_3^{(2)} -$ $0.201494 \cdot \sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(2)} + 0.108278 \cdot \sigma_3^{(4)}\sigma_3^{(2)} + 0.148899 \cdot \sigma_3^{(3)}\sigma_3^{(2)} - 0.195532 \cdot \sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(2)} + 2.109264 \cdot$ $\sigma_3^{(1)} + 0.165246 \cdot \sigma_3^{(9)}\sigma_3^{(1)} + 0.159354 \cdot \sigma_3^{(8)}\sigma_3^{(1)} - 0.159354 \cdot \sigma_3^{(8)}\sigma_3^{(1)}\sigma_3^{(3)} + 0.159571 \cdot \sigma_3^{(7)}\sigma_3^{(1)} + 0.158442 \cdot \sigma_3^{(6)}\sigma_3^{(1)} +$ $0.156601 \cdot \sigma_3^{(5)}\sigma_3^{(1)} - 0.205474 \cdot \sigma_3^{(9)}\sigma_3^{(7)}\sigma_3^{(6)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(1)} + 0.140245 \cdot \sigma_3^{(4)}\sigma_3^{(1)} - 0.202966 \cdot \sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(1)} + 0.132255 \cdot$ $\sigma_3^{(3)}\sigma_3^{(1)} - 15.205265 \cdot \sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(4)}\sigma_3^{(3)}\sigma_3^{(1)} - 0.219559 \cdot \sigma_3^{(7)}\sigma_3^{(5)}\sigma_3^{(3)}\sigma_3^{(1)} + 0.718005 \cdot \sigma_3^{(9)}\sigma_3^{(6)}\sigma_3^{(4)}\sigma_3^{(1)} + \dots$	

Table 11: HCl qubit tapering and 4-qubit CS-VQE parameters

Tapering parameters	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(18)} \sigma_3^{(17)} \sigma_3^{(14)} \sigma_3^{(13)} \sigma_3^{(9)} \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)}, \sigma_3^{(18)} \sigma_3^{(17)} \sigma_3^{(14)} \sigma_3^{(13)} \sigma_3^{(8)} \sigma_3^{(7)} \sigma_3^{(4)} \sigma_3^{(3)},$ $\sigma_3^{(19)} \sigma_3^{(18)} \sigma_3^{(17)} \sigma_3^{(16)} \sigma_3^{(15)} \sigma_3^{(14)} \sigma_3^{(13)} \sigma_3^{(12)} \sigma_3^{(11)} \sigma_3^{(10)}$
Sector ν	$\{-1, 1, -1\}$
Noncontextual model	
Symmetry generators $\mathcal{G}$	$\sigma_3^{(16)} \sigma_3^{(15)} \sigma_3^{(14)} \sigma_3^{(13)} \sigma_3^{(12)} \sigma_3^{(11)} \sigma_3^{(10)} \sigma_3^{(9)} \sigma_3^{(8)}, \sigma_3^{(15)} \sigma_3^{(14)} \sigma_3^{(11)} \sigma_3^{(10)} \sigma_3^{(7)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(1)} \sigma_3^{(0)},$ $\sigma_3^{(14)}, \sigma_3^{(13)}, \sigma_3^{(12)}, \sigma_3^{(11)}, \sigma_3^{(10)}, \sigma_3^{(9)}, \sigma_3^{(8)}, \sigma_3^{(7)} \sigma_3^{(4)} \sigma_3^{(3)} \sigma_3^{(1)} \sigma_3^{(0)}, \sigma_3^{(5)}, \sigma_3^{(4)},$ $\sigma_3^{(3)}, \sigma_3^{(2)}, \sigma_3^{(1)}, \sigma_3^{(0)}$
Sector ν	$\{1, 1, -1, -1, -1, -1, -1, -1, 1, -1, -1, -1, -1, -1, -1\}$
Class observable $C(\mathbf{r})$	$-1.9936299005523587 \times 10^{-7} \cdot \sigma_3^{(15)} \sigma_3^{(14)} \sigma_3^{(11)} \sigma_3^{(10)} \sigma_3^{(6)} \sigma_3^{(4)} \sigma_3^{(3)} +$ $-0.9999999999999801 \cdot \sigma_3^{(15)} \sigma_3^{(14)} \sigma_3^{(11)} \sigma_3^{(10)} \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(2)}$
4-qubit CS-VQE parameters	
Fixed stabilizers $\mathcal{F}$	$\sigma_3^{(14)}, \sigma_3^{(13)}, \sigma_3^{(12)}, \sigma_3^{(11)}, \sigma_3^{(10)}, \sigma_3^{(9)}, \sigma_3^{(8)}, \sigma_3^{(5)}, \sigma_3^{(4)}, \sigma_3^{(3)}, \sigma_3^{(2)}, \sigma_3^{(1)}, \sigma_3^{(0)}$
Reference state $ \tilde{\psi}_{\text{ref}}\rangle$	$ 1111\rangle$
Minimal ansatz $A(\boldsymbol{\theta})$	$\theta_0 \cdot \sigma_1^{(16)} \sigma_3^{(15)} \sigma_3^{(14)} \sigma_2^{(13)} \sigma_1^{(7)} \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} +$ $\theta_1 \cdot \sigma_1^{(16)} \sigma_3^{(15)} \sigma_3^{(14)} \sigma_3^{(13)} \sigma_2^{(12)} \sigma_1^{(7)} \sigma_3^{(6)} \sigma_3^{(5)} \sigma_1^{(4)} +$ $\theta_2 \cdot \sigma_1^{(16)} \sigma_3^{(15)} \sigma_3^{(14)} \sigma_2^{(13)} \sigma_1^{(7)} \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_1^{(3)} +$ $\theta_3 \cdot \sigma_1^{(16)} \sigma_3^{(15)} \sigma_3^{(14)} \sigma_3^{(13)} \sigma_2^{(12)} \sigma_1^{(7)} \sigma_3^{(6)} \sigma_3^{(5)} \sigma_3^{(4)} \sigma_2^{(3)}$
Projected ansatz $\pi_{\nu'}^{U_{\mathcal{F}}}(A(\boldsymbol{\theta}))$	$\theta_0 \cdot \sigma_2^{(3)} \sigma_1^{(1)} + \theta_1 \cdot \sigma_3^{(3)} \sigma_2^{(2)} \sigma_1^{(1)} + \theta_2 \cdot \sigma_2^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} + \theta_3 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_2^{(0)}$
Optimal parameters $\boldsymbol{\theta}_{\min}$	$(0.11734408126111578, \quad -0.0324191385451714, \quad -0.03239043238733037,$ $0.025611946691747454)$
Absolute error $\Delta_c(\mathcal{F})$ (Ha)	0.000537
VQE simulation absolute error $ \tilde{E}(\boldsymbol{\theta}_{\min}) - \epsilon_0 $ (Ha)	0.000865

Contextual subspace Hamiltonian

Diagonal terms

$$\begin{aligned} & -459.943201 \cdot \sigma_3^{(3)} \sigma_3^{(2)} + 0.503116 \cdot \sigma_3^{(3)} + 0.723375 \cdot \sigma_3^{(2)} + 0.196914 \cdot \sigma_3^{(1)} \sigma_3^{(0)} + 0.145853 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + \\ & 0.123346 \cdot \sigma_3^{(3)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.117638 \cdot \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.503116 \cdot \sigma_3^{(1)} + 0.123346 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} + 0.127488 \cdot \sigma_3^{(3)} \sigma_3^{(1)} + \\ & 0.119024 \cdot \sigma_3^{(2)} \sigma_3^{(1)} + 0.723375 \cdot \sigma_3^{(0)} + 0.117638 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.119024 \cdot \sigma_3^{(3)} \sigma_3^{(0)} + 0.130881 \cdot \sigma_3^{(2)} \sigma_3^{(0)} \end{aligned}$$

Off-diagonal terms

$$\begin{aligned}
& 0.036191 \cdot \sigma_1^{(3)} - 0.036191 \cdot \sigma_1^{(3)} \sigma_3^{(2)} - 0.021181 \cdot \sigma_1^{(3)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.021181 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.004633 \cdot \sigma_1^{(3)} \sigma_3^{(1)} - \\
& 0.004633 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(1)} + 0.010376 \cdot \sigma_1^{(3)} \sigma_3^{(0)} - 0.010376 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_3^{(0)} + 0.014092 \cdot \sigma_1^{(2)} - 0.014092 \cdot \sigma_3^{(3)} \sigma_1^{(2)} + \\
& 0.002325 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} - 0.002325 \cdot \sigma_1^{(2)} \sigma_3^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.002343 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} - 0.002343 \cdot \sigma_1^{(2)} \sigma_3^{(3)} - 0.01411 \cdot \\
& \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(0)} + 0.01411 \cdot \sigma_1^{(2)} \sigma_3^{(0)} - 0.040539 \cdot \sigma_1^{(3)} \sigma_1^{(2)} - 0.040539 \cdot \sigma_2^{(3)} \sigma_2^{(2)} + 0.021418 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_3^{(0)} + 0.021418 \cdot \\
& \sigma_2^{(3)} \sigma_2^{(2)} \sigma_1^{(3)} \sigma_3^{(0)} - 3.9e-05 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} - 3.9e-05 \cdot \sigma_2^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} + 0.004465 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(0)} + 0.004465 \cdot \sigma_2^{(3)} \sigma_2^{(2)} \sigma_3^{(0)} - \\
& 0.036191 \cdot \sigma_1^{(1)} + 0.021181 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} - 0.004633 \cdot \sigma_3^{(3)} \sigma_1^{(1)} - 0.010376 \cdot \sigma_3^{(2)} \sigma_1^{(1)} + 0.036191 \cdot \sigma_1^{(1)} \sigma_3^{(0)} - 0.021181 \cdot \\
& \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_3^{(0)} + 0.004633 \cdot \sigma_3^{(3)} \sigma_1^{(1)} \sigma_3^{(0)} + 0.010376 \cdot \sigma_3^{(2)} \sigma_1^{(1)} \sigma_3^{(0)} - 0.03522 \cdot \sigma_1^{(3)} \sigma_1^{(1)} + 0.03522 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} + 0.03522 \cdot \\
& \sigma_1^{(3)} \sigma_1^{(1)} \sigma_3^{(0)} - 0.03522 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_3^{(0)} + 0.009644 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} - 0.009644 \cdot \sigma_1^{(2)} \sigma_1^{(1)} - 0.009644 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} \sigma_3^{(0)} + \\
& 0.009644 \cdot \sigma_1^{(2)} \sigma_1^{(1)} \sigma_3^{(0)} + 0.022546 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} + 0.022546 \cdot \sigma_2^{(3)} \sigma_2^{(2)} \sigma_1^{(1)} - 0.022546 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} \sigma_3^{(0)} - 0.022546 \cdot \\
& \sigma_2^{(3)} \sigma_2^{(2)} \sigma_1^{(1)} \sigma_3^{(0)} - 0.014092 \cdot \sigma_1^{(0)} + 0.014092 \cdot \sigma_3^{(3)} \sigma_1^{(0)} - 0.002325 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.002343 \cdot \sigma_3^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} + \\
& 0.01411 \cdot \sigma_3^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} + 0.002325 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} + 0.002343 \cdot \sigma_3^{(3)} \sigma_1^{(0)} - 0.01411 \cdot \sigma_3^{(2)} \sigma_1^{(0)} + 0.009644 \cdot \sigma_1^{(3)} \sigma_3^{(1)} \sigma_1^{(0)} - \\
& 0.009644 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - 0.009644 \cdot \sigma_1^{(3)} \sigma_1^{(0)} + 0.009644 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(0)} - 0.015457 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} + \dots
\end{aligned}$$

$$\begin{aligned}
& \cdots + 0.015457 \cdot \sigma_1^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} + 0.015457 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_1^{(0)} - 0.015457 \cdot \sigma_1^{(2)} \sigma_1^{(0)} + 0.003375 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} + 0.003375 \cdot \\
& \sigma_2^{(3)} \sigma_2^{(2)} \sigma_3^{(1)} \sigma_1^{(0)} - 0.003375 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_1^{(0)} - 0.003375 \cdot \sigma_2^{(3)} \sigma_2^{(2)} \sigma_1^{(0)} - 0.040539 \cdot \sigma_1^{(1)} \sigma_1^{(0)} + 0.021418 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - \\
& 3.9e - 05 \cdot \sigma_3^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.004465 \cdot \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.040539 \cdot \sigma_2^{(1)} \sigma_2^{(0)} + 0.021418 \cdot \sigma_3^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 3.9e - 05 \cdot \\
& \sigma_3^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.004465 \cdot \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.022546 \cdot \sigma_1^{(3)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.022546 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.022546 \cdot \sigma_1^{(3)} \sigma_2^{(1)} \sigma_2^{(0)} + \\
& 0.022546 \cdot \sigma_1^{(3)} \sigma_3^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} - 0.003375 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.003375 \cdot \sigma_1^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} - 0.003375 \cdot \sigma_3^{(3)} \sigma_1^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + \\
& 0.003375 \cdot \sigma_1^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + 0.027799 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.027799 \cdot \sigma_2^{(3)} \sigma_2^{(2)} \sigma_1^{(1)} \sigma_1^{(0)} + 0.027799 \cdot \sigma_1^{(3)} \sigma_1^{(2)} \sigma_2^{(1)} \sigma_2^{(0)} + \\
& 0.027799 \cdot \sigma_2^{(3)} \sigma_2^{(2)} \sigma_2^{(1)} \sigma_2^{(0)}
\end{aligned}$$
