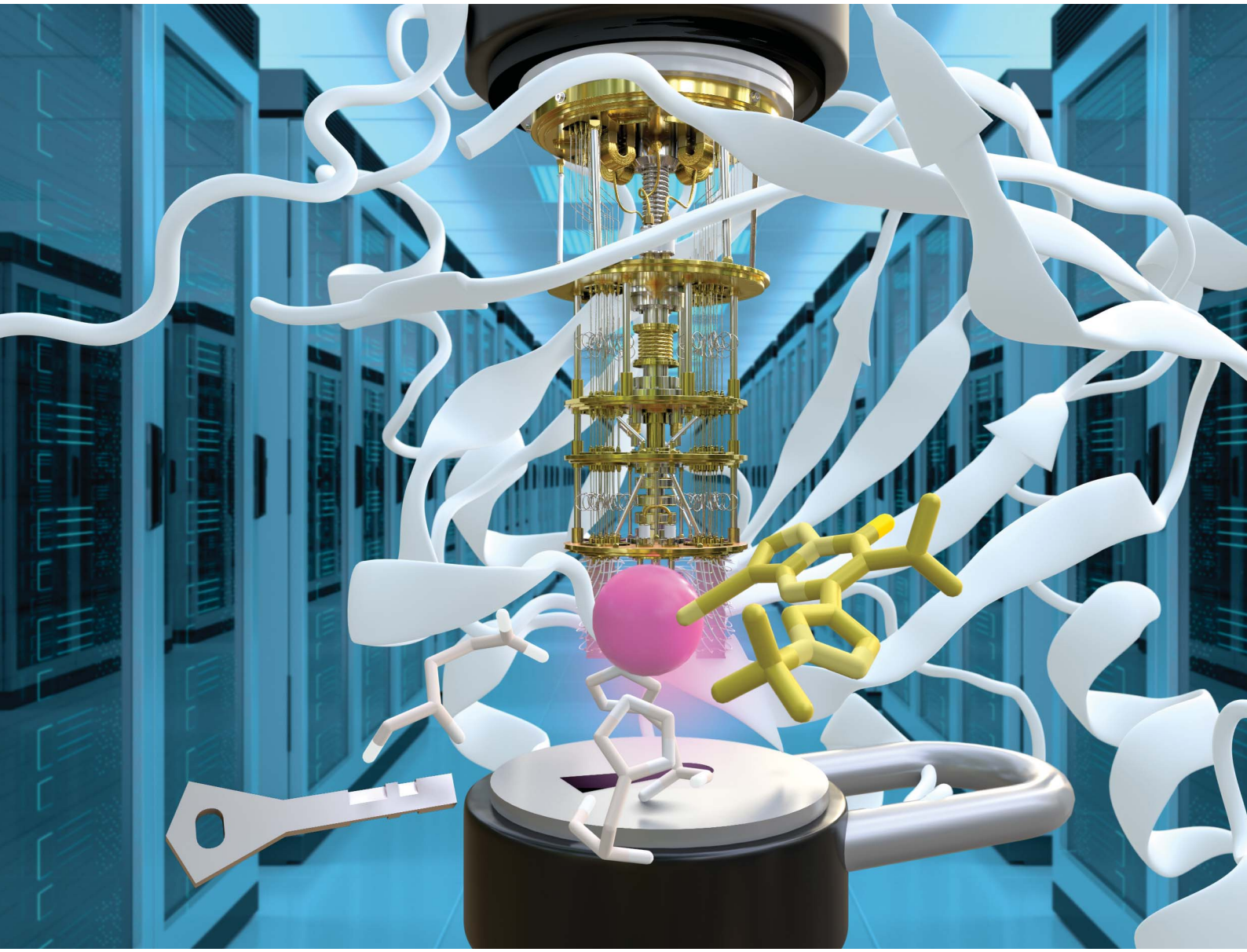
Volume 13
Number 11
21 March 2022
Pages 3047–3302

Chemical Science

rsc.li/chemical-science



ISSN 2041-6539

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

Robert M. Parrish, Nikolaj Moll *et al.*

Towards the simulation of large scale protein–ligand interactions on NISQ-era quantum computers

Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2022, **13**, 3094

All publication charges for this article have been paid for by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Towards the simulation of large scale protein–ligand interactions on NISQ-era quantum computers

Fionn D. Malone, [a] Robert M. Parrish, *[a] Alicia R. Welden,[a] Thomas Fox,[b] Matthias Degroote, [c] Elica Kyoseva,[c] Nikolaj Moll, *[c] Raffaele Santagati [c] and Michael Streif[c]

Received 14th October 2021
Accepted 14th January 2022

DOI: 10.1039/d1sc05691c

rsc.li/chemical-science

We explore the use of symmetry-adapted perturbation theory (SAPT) as a simple and efficient means to compute interaction energies between large molecular systems with a hybrid method combining NISQ-era quantum and classical computers. From the one- and two-particle reduced density matrices of the monomer wavefunctions obtained by the variational quantum eigensolver (VQE), we compute SAPT contributions to the interaction energy [SAPT(VQE)]. At first order, this energy yields the electrostatic and exchange contributions for non-covalently bound systems. We empirically find from ideal statevector simulations that the SAPT(VQE) interaction energy components display orders of magnitude lower absolute errors than the corresponding VQE total energies. Therefore, even with coarsely optimized low-depth VQE wavefunctions, we still obtain sub kcal mol^{-1} accuracy in the SAPT interaction energies. In SAPT(VQE), the quantum requirements, such as qubit count and circuit depth, are lowered by performing computations on the separate molecular systems. Furthermore, active spaces allow for large systems containing thousands of orbitals to be reduced to a small enough orbital set to perform the quantum portions of the computations. We benchmark SAPT(VQE) (with the VQE component simulated by ideal statevector simulators) against a handful of small multi-reference dimer systems and the iron center containing human cancer-relevant protein lysine-specific demethylase 5 (KDM5A).

Quantum chemistry has emerged as one of the most promising areas where a practical quantum advantage from near term quantum computers could be demonstrated.[1–3] Identifying industrially relevant applications that can practically benefit from quantum simulations is, however, a complicated task.[4,5] On the one hand, existing classical algorithms have benefited from decades of development, benchmarking and optimization so demonstrating a computational advantage over these is challenging given limitations on current quantum hardware.[6–8] On the other hand, current noisy intermediate-scale quantum (NISQ) hardware[9] suffers from relatively poor gate fidelity so that the resulting physical properties can often be biased and far from exact without error mitigation.[10] Thus it is important to design quantum algorithms that minimize the quantum resources required.

Coupled with these challenges is the problem of finding an industrially relevant application that can benefit from a quantum computer in the first place.[11] One such area that has been suggested as possibly benefiting from quantum computing is computer aided drug design (CADD).[12] CADD has a long history and has many components ranging from high-level optimization problems such as structure search and conformational sampling down to low-level quantum mechanical problems such as computing protein–ligand interaction energies, all of which could potentially benefit from a large scale quantum computer.[13,14] In this work we will focus on this final problem, namely computing the interaction energy and properties of large scale protein–ligand systems by approximately solving the electronic structure problem.

To date, most quantum algorithms aimed at solving the electronic structure problem directly are general and have mostly been applied to small molecules in small basis sets. Of course, applications are limited by current quantum resources so that reaching chemical accuracy, defined as calculating energy differences in the complete basis set limit to within 1 kcal mol^{-1} accuracy, is difficult to achieve. Nevertheless, relatively minor attention has been paid to the entire workflow required to solve an industrially relevant problem in drug design. This often includes highly tailored approaches that require classical pre- and post-processing, molecular dynamics and structure relaxation, active space selection and finally computation of interaction energies and related

[a]QC Ware Corporation, Palo Alto, CA, 94301, USA. E-mail: rob.parrish@qcware.com
[b]Medicinal Chemistry, Boehringer Ingelheim Pharma GmbH & Co. KG, Birkendorfer Straße 65, 88397, Biberach an der Riß, Germany
[c]Quantum Lab, Boehringer Ingelheim, 55218, Ingelheim am Rhein, Germany. E-mail: nikolaj.moll@boehringer-ingelheim.com

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

quantities all for systems containing hundreds or thousands of atoms. Thus, it is important to isolate potential application areas now and codify these workflows with quantum algorithms in mind as the pace of hardware improvement accelerates.

Conceptually, computing the interaction energy of a protein–ligand system is a straightforward task. One computes the ground state energy of the dimer and monomer systems separately and subtracts the two to determine the interaction energy. There are a number of issues with this approach, particularly when a quantum computer is involved. First, in finite Gaussian basis sets typically employed in quantum chemical computations, one has to account for basis set superposition error (BSSE) using the counterpoise correction.[15] This unnecessarily increases the qubit count requirements for the individual monomers and can potentially lead to convergence issues for hybrid quantum–classical algorithms like the variational quantum eigensolver (VQE).[16,17] A more concerning problem for NISQ computers is resolving total energies of individual monomers (typically on the order of 1000s of kcal mol^{-1}) to sufficient precision to subtractively resolve binding energies which are typically on the order of 5 kcal mol^{-1}. This is a major issue for NISQ approaches which typically evaluate total energy expectation values statistically, carrying very high measurement cost penalty for high-precision expectation values. This is also an increasingly challenging problem for heuristic algorithms like the VQE which would require very deep circuits with thousands of parameters to achieve the required precision (note that precision and accuracy are the same concern in subtracting total energies unless strict relative error cancellation can be ensured, which is not clear with methods like VQE). This may be practically impossible with the current general algorithms and available hardware although we note alternatives to the VQE may help to overcome this issue.[18]

In this work we propose using symmetry adapted perturbation theory (SAPT)[19,20] to directly compute the interaction energy through direct expectation values rather than differences, which overcomes some of these problems. Firstly, in principle SAPT does not suffer from BSSE as it directly computes the interaction energy as a perturbation series in the intermolecular potential (note that SAPT still suffers from basis set incompleteness error, as with all second-quantized methods). Secondly, monomer-centered basis sets can be used which can afford additional savings if the geometry of the monomers is fixed across the dissociation path, potentially reducing the number of different quantum computations that have to be performed. Moreover, as SAPT directly computes the interaction energy as a sum of expectation values (rather than differences of large expectation values), it often exhibits favorable error cancellation for errors inherent to the chosen wavefunction ansatz. We show that this observation can significantly reduce the resource requirements for circuit depth with only very coarse VQE wavefunctions required for sub kcal mol^{-1} accuracy in the interaction energy components. Finally, SAPT offers a physically motivated breakdown of the interaction energy components into electrostatic, exchange, induction and dispersion contributions which can offer valuable insight for medicinal chemists when designing protein inhibitors.[21]

Beyond suggesting SAPT as a useful approach for NISQ quantum computers we outline an efficient active space formulation of SAPT that can be applied for protein–ligand interactions for systems containing heavy metal centers and thousands of atoms. Key to this implementation is the GPU accelerated classical pre- and post-processing steps which practically help to run such simulations.[22] We will largely focus on the accurate qualitative description of systems with strong multi-reference character that can not easily be described by classical approaches and thus could offer a more transparent demonstration of a practical quantum advantage. In this paper we will only consider the first order contributions to the exchange energy leaving the second order terms, which require solving electronic response, for future work.

We begin by outlining in detail the active space formulation for first-order SAPT that can be coupled to any quantum simulation that can produce one- and two-particle reduced density matrices. Next we discuss the VQE ansatz used in this work, although the SAPT method itself is largely independent of the way in which the ground state properties are computed. Finally we benchmark our method using ideal quantum simulators and demonstrate a significant reduction in error found when poorly converged VQE wavefunctions are used for model multi-reference systems and for the human cancer relevant[23] lysine-specific demethylase 5 (KDM5A) protein with different ligand substitutions.

I. Methods

In this section we will describe the classical implementation of the density matrix formulation of SAPT followed by how this implementation can be efficiently adapted for NISQ devices.

A Notation

In this work we consider the interaction of two monomers A and B and will use the following notation for different real orbital types belonging to A and/or B.

- μ/ν – nonorthogonal atomic spatial orbital basis indices. Note that these could conceptually be either monomer-centered of dimer-centered bases as far as the theory is concerned. Unless otherwise noted, we use monomer-centered bases for all atomic spatial orbital bases encountered in practical test cases in this work.
- p/q – orthogonal molecular spatial orbital basis indices.
- i/j – orthogonal occupied spatial orbital basis indices.
- t/u – orthogonal active spatial orbital basis indices.
- a/b – orthogonal virtual spatial orbital basis indices.

Repeated indices within a monomer will be denoted with primes, *e.g.*, p, p', p'', p'''. Summation over repeated indices is assumed throughout. When dealing with spin-orbital quantities, we use the context specific notation of an "unbarred" orbital index to denote α and a "barred" orbital index to denote β, *i.e.*, $p^\dagger$ is an α spin-orbital creation operator on spatial orbital index p, while $\bar{p}^\dagger$ is a β spin-orbital creation operator on spatial orbital index p.

B Symmetry adapted perturbation theory

Traditionally, the interaction energy between two monomers, E_{int}, is calculated in the supermolecular approach as

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$E_{\mathrm{int}} = E_{\mathrm{AB}} - E_{\mathrm{A}} - E_{\mathrm{B}}, \tag{1}$$

where E_{AB} is the ground state energy of the combined system and $E_{\mathrm{A/B}}$ are the energies of the individual fragments, evaluated at the frozen geometry of the dimer system (*i.e.*, no deformation energy contributions). In contrast, in SAPT, E_{int} is instead evaluated through a perturbation series in the intermonomer interaction potential

$$\hat{V} = \sum_{i=1}^{N_{\mathrm{A}}} \sum_{j=1}^{N_{\mathrm{B}}} \tilde{v}(\mathbf{r}_i, \mathbf{r}_j), \tag{2}$$

where $N_{\mathrm{A/B}}$ is the number of electrons in monomer A or B, and

$$\begin{aligned} \tilde{v}(\mathbf{r}_i, \mathbf{r}_j) = & \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ & - \frac{1}{N_{\mathrm{B}}} \sum_{\mathrm{J} \in \mathrm{B}} Z_{\mathrm{J}} \frac{1}{|\mathbf{R}_{\mathrm{J}} - \mathbf{r}_i|} - \frac{1}{N_{\mathrm{A}}} \sum_{\mathrm{I} \in \mathrm{A}} Z_{\mathrm{I}} \frac{1}{|\mathbf{R}_{\mathrm{I}} - \mathbf{r}_j|} \\ & + \frac{1}{N_{\mathrm{A}} N_{\mathrm{B}}} \sum_{\mathrm{I} \in \mathrm{A}} \sum_{\mathrm{I} \in \mathrm{B}} Z_{\mathrm{I}} Z_{\mathrm{J}} \frac{1}{|\mathbf{R}_{\mathrm{I}} - \mathbf{R}_{\mathrm{J}}|}, \end{aligned} \tag{3}$$

describes the interaction of the electrons and nuclei in monomer A with the electrons and nuclei in monomer B (and *vice versa*). Symmetrized Rayleigh–Schrödinger perturbation theory yields a perturbation series for the interaction energy

$$E_{\mathrm{int}} = E^{(1)}_{\mathrm{pol}} + E^{(1)}_{\mathrm{exch}} + E^{(2)}_{\mathrm{pol}} + E^{(2)}_{\mathrm{exch}} + \ldots \tag{4}$$

where, $E^{(1)}_{\mathrm{pol}}$ is the electrostatic energy, $E^{(2)}_{\mathrm{pol}}$ is a sum of dispersion and induction energies, while $E^{(n)}_{\mathrm{exch}}$ account for exchange interactions.

In order to evaluate the first two terms in the perturbation series in eqn (4) one first needs to solve for the ground state wavefunction, $|\Psi_{\mathrm{A}}\rangle$, of the individual monomer Hamiltonian, here given for monomer A

$$\begin{aligned} \hat{H}_{\mathrm{A}} = & \sum_{pp'} h_{pp'} \hat{E}_{pp'} \\ & + \frac{1}{2} \sum_{pp'p''p'''} v^{pp'}_{p''p'''} \left(\hat{E}_{pp''} \hat{E}_{p'p'''} - \delta_{p''p'} \hat{E}_{pp'''} \right), \end{aligned} \tag{5}$$

where h_{pq} and $v^{pp'}_{p''p'''} = (pp''|p'p''')$ are the usual one- and two-electron integrals and

$$\hat{E}_{pp'} = p^{\dagger} p' + \bar{p}^{\dagger} \bar{p}', \tag{6}$$

is the singlet-adapted one-particle substitution operator.

In the absence of the exact ground state monomer wavefunctions, the SAPT interaction energy is traditionally computed as a triple perturbation theory in the intramonomer fluctuation potentials (assuming a Møller–Plesset partitioning of the Hamiltonian). At the lowest order this gives rise to SAPT(HF) (or SAPT0), where each term is evaluated using Hartree–Fock density matrices.[20] Another popular approach is to use density functional theory SAPT(DFT)[24,25] to account for intramonomer electron correlation, but this approach is naturally limited by the performance of the chosen density functional.[26]

In this work we will assume that we can determine the exact (or at least very accurate) ground state properties of the individual monomers using a quantum computer. The first order SAPT expressions can then be evaluated using only the ground state unperturbed wavefunctions of the individual monomers, *i.e.*, $|\Psi_0\rangle = |\Psi_{\mathrm{A}}\rangle \otimes |\Psi_{\mathrm{B}}\rangle$. We will use the density matrix formulation of SAPT[27] systematized by Korona[28–30] and recently fully implemented for complete active space self consistent field (CASSCF) wavefunctions by Hapka and *et al.*[26] This formalism allows for the evaluation of the terms appearing in eqn (4) using just the ground state one- and two-particle reduced density matrices of the monomers with additional response terms for the second order terms. Detailed derivations of the density matrix formulation of SAPT are given elsewhere and here we will focus on the efficient implementation in terms of optimized chemistry primitives on quantum computers.

C Density matrix formulation of SAPT

The first order polarization energy is the electrostatic repulsion energy of monomer A and B, and is given by

$$E^{(1)}_{\mathrm{pol}} = \langle \Psi_{\mathrm{A}} \Psi_{\mathrm{B}} | \hat{V} | \Psi_{\mathrm{A}} \Psi_{\mathrm{B}} \rangle \tag{7}$$

$$= \bar{\gamma}_{pp'} \tilde{v}^{p'q'}_{pq} \bar{\gamma}_{qq'} \tag{8}$$

where, $\bar{\gamma}_{pp'}$ is the spin-summed one-particle reduced density matrix

$$\bar{\gamma}_{pp'} = \langle \Psi_{\mathrm{A}} | p^{\dagger} p' + \bar{p}^{\dagger} \bar{p}' | \Psi_{\mathrm{A}} \rangle, \tag{9}$$

and we have introduced the generalized two-electron repulsion integrals

$$\begin{aligned} \tilde{v}^{p'q'}_{pq} = & \left(pp' \middle| \hat{V} \middle| qq' \right) \\ = & (pp'|qq') \\ & + \frac{1}{N_{\mathrm{A}}} (V_{\mathrm{A}}|qq') S_{pp'} + \frac{1}{N_{\mathrm{B}}} (pp'|V_{\mathrm{B}}) S_{qq'} \\ & + \frac{V_{\mathrm{AB}}}{N_{\mathrm{A}} N_{\mathrm{B}}} S_{pp'} S_{qq'}, \end{aligned} \tag{10}$$

where $(pp'|qq')$ is a mixed two-electron electron repulsion integral between orbitals on monomer A and B, $S_{pp'} = (p|p')$ is an overlap integral and $(V_{\mathrm{X}}|qq')$ are matrix elements of the nuclear-attraction potential of monomer X in the orbital basis of the other monomer.

Similarly, the first order exchange energy is given as[27]

$$E^{(1)}_{\mathrm{exch}} = \frac{\langle \Psi_{\mathrm{A}} \Psi_{\mathrm{B}} | V \mathcal{A} | \Psi_{\mathrm{A}} \Psi_{\mathrm{B}} \rangle}{\langle \Psi_{\mathrm{A}} \Psi_{\mathrm{B}} | \mathcal{A} \Psi_{\mathrm{A}} \Psi_{\mathrm{B}} \rangle} - E^{(1)}_{\mathrm{pol}}, \tag{11}$$

where $\mathcal{A}$ is the antisymmetrizer operator. Using the S^2 approximation, and the density matrix formulation of SAPT,[27] it can be shown that eqn (11) can be written as

$$\begin{aligned} E^{(1)}_{\mathrm{exch}} = & -\frac{1}{2} (\bar{\gamma}_{pp'} \bar{\gamma}_{qq'} \tilde{v}^{q'p'}_{pq} \\ & + \bar{\gamma}_{pp'} \overline{\Gamma}^{qq'}_{q''q'''} S_{p'q'} \tilde{v}^{q'''q''}_{pq} \\ & + \bar{\gamma}_{qq'} \overline{\Gamma}^{pp'}_{p''p'''} S_{q'p'} \tilde{v}^{p''p'''}_{pq} \end{aligned} \tag{12}$$

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$+\overline{\Gamma}^{pp'}_{p''p'''}\overline{\Gamma}^{qq'}_{q''q'''}S_{p'q'''}S_{q'p'''}\tilde{v}^{p''q''}_{pq}$$

$$-E^{(1)}_{\mathrm{pol}}S_{pq'}\overline{\gamma}_{qq'}S_{qp'}\overline{\gamma}_{pp'}\Big), \tag{13}$$

$$= T_1 + T_2 + T_3 + T_4 + T_5 \tag{14}$$

where

$$\overline{\Gamma}^{pp'}_{p''p'''} = 2\Big(\Gamma^{pp'}_{p''p'''} + \Gamma^{p\overline{p}'}_{p''\overline{p}'''}\Big), \tag{15}$$

is the spin-summed two-particle reduced density matrix, where throughout this work we assume the monomers have singlet ground states, and

$$\Gamma^{p\overline{p}'}_{p''\overline{p}'''} = \left\langle \Psi_{\mathrm{A}}\middle| p^{\dagger}\overline{p}'^{\dagger}\overline{p}'''p''' \middle|\Psi_{\mathrm{A}}\right\rangle \tag{16}$$

As can be seen from eqn (8) and (14) the first order SAPT expressions require only the one- and two-particle density matrices to be evaluated.

D Efficient active space implementation

Given that NISQ-era devices are currently limited to tens of qubits (spin-orbitals) we will use an active space approach in order to tackle protein–ligand interactions. We will heavily leverage standard quantum chemistry primitives such as integral driven Coulomb and exchange matrix builds which exploit sparsity[31] and can efficiently be implemented on GPUs.[32–34] These considerations are important when simulating hundreds of atoms and thousands of basis functions.

In the active space approach we partition the one-electron orbital set into N_{c} core orbitals, N_{a} active orbitals and N_{i} virtual orbitals. This partitioning gives rise to modified monomer Hamiltonians given by (for example for monomer A)

$$\hat{H}_{\mathrm{A}}{}' = \sum_{tt'\sigma}\tilde{h}_{tt'}a^{\dagger}_{t\sigma}a_{t'\sigma} + \sum_{\sigma\sigma'}\sum_{tt't''t'''}(tt''|t't''')a^{\dagger}_{t\sigma}a^{\dagger}_{t'\sigma'}a_{t'''\sigma'}a_{t''\sigma}, \tag{17}$$

where the modified one-electron integrals $\tilde{h}_{tt'}$ now include core-active space interactions

$$\tilde{h}_{tt'} = h_{tt'} + \sum_{ii'}\left[(tt'|ii') - \frac{1}{2}(ti'|it')\right]\gamma_{ii'}. \tag{18}$$

For large scale applications N_{c} is typically large (100–1000) and eqn (18) can be efficiently evaluated in the AO basis

$$\tilde{h}_{\mu\mu'} = h_{\mu\mu'} + 2J^{\mathrm{c}}_{\mu\mu'} - K^{\mathrm{c}}_{\mu\mu'}, \tag{19}$$

before being transformed to the active space MO basis

$$\tilde{h}_{tt'} = \sum_{\mu\mu'}C_{\mu t}\tilde{h}_{\mu\mu'}C_{\mu't'} \tag{20}$$

where C are the molecular orbital coefficients. In eqn (19) we have introduced the usual (core) Coulomb and exchange matrices

$$J^{\mathrm{c}}_{\mu\mu'} \equiv J[\gamma^{\mathrm{c}}]_{\mu\mu'} = \sum_{\mu''\mu'''}(\mu\mu'|\mu''\mu''')\gamma^{\mathrm{c}}_{\mu''\mu'''} \tag{21}$$

$$K^{\mathrm{c}}_{\mu\mu'} \equiv K[\gamma^{\mathrm{c}}]_{\mu\mu'} = \sum_{\mu''\mu'''}(\mu\mu'''|\mu''\mu')\gamma^{\mathrm{c}}_{\mu''\mu'''}, \tag{22}$$

where the core density matrix is given by

$$\gamma^{\mathrm{c}}_{\mu\mu'} = \sum_{i}C_{\mu i}C_{\mu'i}. \tag{23}$$

Once the ground state in eqn (17) has been found and the active space one- and two-particle density matrices have been formed the first order SAPT contributions can be calculated as a classical post-processing step which can be implemented efficiently by considering the block structure of the one- and two-particle density matrices. Recall that in the MO basis we have

$$\bar{\gamma}^{\mathrm{c}} = \bar{\gamma}_{ii'} = 2\delta_{ii'} \tag{24}$$

$$\bar{\gamma}^{\mathrm{a}} = \bar{\gamma}_{tt'} \tag{25}$$

and

$$\overline{\Gamma}^{\mathrm{cc}} = \overline{\Gamma}^{ii'}_{i''i'''} = \overline{\gamma}_{ii''}\overline{\gamma}_{i'i'''} - \frac{1}{2}\overline{\gamma}_{ii'''}\overline{\gamma}_{i'i''} \tag{26}$$

$$\overline{\Gamma}^{\mathrm{ac}} = \begin{cases}\overline{\Gamma}^{ti}_{t'i'} = \overline{\Gamma}^{it}_{i't'} = \overline{\gamma}_{tt'}\overline{\gamma}_{ii'}\\ \overline{\Gamma}^{it}_{t'i'} = \overline{\Gamma}^{ti}_{i't'} = -\frac{1}{2}\overline{\gamma}_{tt'}\overline{\gamma}_{ii'}\end{cases} \tag{27}$$

$$\overline{\Gamma}^{\mathrm{aa}} = \overline{\Gamma}^{tt'}_{t''t'''} \tag{28}$$

where again i and t are occupied (core) and active orbital indices respectively. All other blocks of the density matrices are zero.

Not much is gained by exploiting this block structure for first order terms that contain the one-particle density matrix only. Thus, we evaluate them in the AO basis directly as

$$E^{(1)}_{\mathrm{pol}} = \sum_{\mu\mu'}\overline{\gamma}_{\mu\mu'}\tilde{J}\left[\overline{\gamma}^{\mathrm{B}}\right]_{\mu\mu'}, \tag{29}$$

where the generalized Coulomb ($\tilde{J}$) and exchange ($\tilde{K}$) matrices are analogues of eqn (21) and (22) with the standard electron repulsion integrals replaced with their generalized counterparts (see eqn (10)) and the AO density matrices are given by

$$\overline{\gamma}_{\mu\mu'} = \sum_{ii'}C_{\mu i}\overline{\gamma}^{\mathrm{c}}_{ii'}C_{\mu'i'} + \sum_{tt'}C_{\mu t}\overline{\gamma}^{\mathrm{a}}_{tt'}C_{\mu't'}. \tag{30}$$

Similarly, the first term in the exchange matrix is evaluated as

$$T_1 = -\frac{1}{2}\sum_{\mu\mu'}\overline{\gamma}_{\mu\mu'}\tilde{K}\left[\overline{\gamma}^{\mathrm{B}}\right]_{\mu\mu'}, \tag{31}$$

while T_5 is a simple trace of a matrix product. The other three terms are a bit more complicated. For example, for T_2, there will be in total six terms arising from the different combinations of

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

core and active orbitals sets. Inserting the expressions for γ^{c} and I^{cc}, I^{ac} from eqn (28) we have

$$T_2\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}},\overline{I}_{\mathrm{B}}^{\mathrm{ac}}\right] = 4\overline{\gamma}_{uu'}\left[S_{ij}\left(\tilde{v}_{iu}^{ju'} - \frac{1}{2}\tilde{v}_{iu}^{u'j}\right) + S_{iu}\left(\tilde{v}_{ij}^{u'j} - \frac{1}{2}\tilde{v}_{ij}^{ju'}\right)\right] \quad (32)$$

This can be simplified by using the definitions of the generalized Coulomb and exchange matrices defined earlier to be written as

$$T_2\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}},\overline{I}_{\mathrm{B}}^{\mathrm{ac}}\right] = 4S_{ij}C_{\mu i}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}\right]_{\mu\nu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}\right]_{\mu\nu''}\right)C_{\nu''j} + 2\gamma_{uu'}S_{iu}C_{\mu i}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu''}\right)C_{\nu''u'}. \quad (33)$$

Similarly we have

$$T_2\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}},\overline{I}_{\mathrm{B}}^{\mathrm{cc}}\right] = 2\gamma_{tt'}S_{t'j'}C_{\mu t}\left(\tilde{J}\left[\gamma_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu'''} - \frac{1}{2}\tilde{K}\left[\gamma_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu'''}\right)C_{\nu'''j'}, \quad (34)$$

and

$$T_2\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}},\overline{I}_{\mathrm{B}}^{\mathrm{ac}}\right] = 2\overline{\gamma}_{tt'}S_{t'j}C_{\mu t}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}\right]_{\mu\nu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}\right]_{\mu\nu''}\right)C_{\nu''j} + \overline{\gamma}_{tt'}\overline{\gamma}_{uu'}S_{t'u}C_{\mu t}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu''}\right)C_{\nu''u'}. \quad (35)$$

For the core only contribution we find

$$T_2\left[\gamma_{\mathrm{A}}^{\mathrm{c}},I_{\mathrm{B}}^{\mathrm{cc}}\right] = \overline{\gamma}_{\mu\mu'}^{\mathrm{c}}S_{\mu'\nu'}\overline{\gamma}_{\nu'\nu'''}^{\mathrm{c}}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu'''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{\mu\nu'''}\right). \quad (36)$$

Terms involving I^{aa} typically cannot be simplified much, *e.g.*,

$$T_2\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}},\overline{I}_{\mathrm{B}}^{\mathrm{aa}}\right] = 2\overline{I}_{u''u'''}^{uu'}S_{iu'}\tilde{v}_{iu}^{u'''u''}. \quad (37)$$

To avoid the formation of the $O(N_{\mathrm{a}}^{3}N_{\mathrm{c}})$ generalized electron repulsion integrals it is helpful to note that

$$\tilde{v}_{iu}^{u'''u''} = \sum_{\mu}C_{\mu i}\tilde{v}_{\mu u}^{u'''u''}, \quad (38)$$

so that

$$\sum_{i}S_{iu'}\tilde{v}_{iu}^{u'''u''} = \sum_{\mu}\left(\sum_{i}C_{\mu i}S_{iu'}\right)\tilde{v}_{\mu u}^{u'''u''} \quad (39)$$

$$= \sum_{\mu}Q_{\mu u'}\tilde{v}_{\mu u}^{u'''u''} \quad (40)$$

$$\equiv \tilde{W}_{u'u}^{u'''u''}, \quad (41)$$

where we have defined

$$Q_{\mu u'} \equiv \sum_{i}C_{\mu i}S_{iu'}. \quad (42)$$

Recall that S is an intermolecular overlap integral matrix and will decay to zero as the monomers are separated. This allows us to write

$$T_2\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}},\overline{I}_{\mathrm{B}}^{\mathrm{aa}}\right] = 2\overline{I}_{u''u'''}^{uu'}\tilde{W}_{u'u}^{u'''u''}. \quad (43)$$

Finally, we have

$$T_2\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}},\overline{I}_{\mathrm{B}}^{\mathrm{aa}}\right] = \overline{\gamma}_{tt'}\overline{I}_{u''u'''}^{uu'}S_{t'u'}\tilde{v}_{tu}^{u'''u''}, \quad (44)$$

which does not simplify further. Note that the generalized two-electron integrals only need to be constructed for the last two terms and require at most $\mathcal{O}(N_{\mathrm{a}}^{4})$ storage. Although not a concern for the system sizes considered here, further reduction in computational cost and memory can be achieved through density fitting and related approaches.[35]

T_3 is analogous to T_2 while T_4 is quite verbose and contains sixteen terms. Full expressions for these are given in Appendix A. The above expressions are completely general and do not depend on the method used to evaluate the one- and two-particle density matrices.

E Variational quantum eigensolver

Up to this point we have assumed that the ground state one- and two-particle reduced density matrices of the monomers could be determined by some means. In this subsection we will give further details of the VQE algorithm used in this work. As described, SAPT is essentially a post-processing step that relies only on the availability of the one- and two-particle reduced density matrices. Therefore, it is not tied to any particular quantum algorithm, however, in this work we will focus on using the VQE.

In the SAPT(VQE) approach, one or both of the monomer active space wavefunctions are generated by VQE-type quantum circuits

$$|\Psi_{\mathrm{VQE}}\rangle \equiv \hat{U}_{\mathrm{VQE}}|\Phi_{\mathrm{I}}\rangle \quad (45)$$

where $|\Phi_{\mathrm{I}}\rangle$ is some initial state (typically the Hartree–Fock determinant). Note that with the VQE ansatz adopted for this paper, the active space wavefunction $|\Psi_{\mathrm{VQE}}\rangle$ will be taken to be real, and will be a definite eigenfunction of the $\hat{N}_{\alpha}$, $\hat{N}_{\beta}$, and $\hat{S}_{2}$ operators.

In the Jordan–Wigner representation used in this paper, the creation/annihilation operators are defined as,

$$p^{\pm} = \bigotimes_{p'=0}^{p'=p-1}\hat{Z}_{p'}\left(\hat{X}_{p} \mp i\hat{Y}_{p}\right)\Big/2, \quad (46)$$

$$\overline{p}^{\pm} = \bigotimes_{p'=0}^{p'=N_{\alpha}-1}\hat{Z}_{p'}\bigotimes_{\overline{p}'=0}^{\overline{p}'=\overline{p}-1}\hat{Z}_{\overline{p}'}\left(\hat{X}_{\overline{p}} \mp i\hat{Y}_{\overline{p}}\right)\Big/2, \quad (47)$$

where $p^{+} = p^{\dagger}$ and $p^{-} = p$ and we order the Jordan-Wigner strings in α- then β-order and $\hat{Z}$, $\hat{Y}$ and $\hat{X}$ are the usual Pauli operators.

In this work we use a modified version of the unitary cluster Jastrow wavefunction[36] (k-uCJ) which takes the form

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$|\Psi_{\mathrm{VQE}}\rangle = \prod_k \exp\left(-\hat{K}^{(k)}\right)\exp\left(\hat{T}^{(k)}\right)\exp\left(+\hat{K}^{(k)}\right)|\Phi_I\rangle, \quad (48)$$

where $\hat{K}(k)$ and $\hat{T}(k)$ are one- and two-body operators, and k is a parameter that controls the depth of the circuit and as a result its variational freedom. Our modified k-uCJ ansatz differs from ref. 36 in the choice of two-body operator and we restrict ourselves to real anti-symmetric matrices. For the one-body rotations we use spin-restricted orbital transformations,

$$\hat{K}^{(k)} \equiv \sum_{pp'} \kappa_{pp'}^{(k)}[(p^\dagger p' - p'^\dagger p) + (\overline{p}^\dagger \overline{p}' - \overline{p}'^\dagger \overline{p})] \quad (49)$$

where $\kappa_{pp'}^{(k)} = -\kappa_{p'p}^{(k)}$ is a real, antisymmetric $N_\mathrm{a} \times N_\mathrm{a}$ matrix of orbital rotation generators. The restricted orbital transformation operator is equivalent to a 1-particle spin-restricted orbital transformation *via*,

$$U_{pp'}^{(k)} \equiv \left[\exp\left(\kappa^{(k)}\right)\right]_{pp'} \quad (50)$$

This spin-restricted orbital rotation can be efficiently implemented in quantum circuits *via* a fabric of Givens rotations.[37]

For the two-particle operator we use a modified diagonal double-substitution operator,

$$\hat{T}^{(k)} \equiv \sum_{p=0}^{M-1} \sum_{\substack{p'=p \bmod 2 \\ p+=2}}^{M-2} \tau_{pp'}^{(k)} \Big[(p'+1)^\dagger \overline{(p'+1)}^\dagger p'\overline{p}' - p'^\dagger \overline{p}'^\dagger (p'+1)\overline{(p'+1)}\Big] \quad (51)$$

which is similar to the unitary pair coupled-cluster generalized singles and doubles expression (k-UpCCGSD).[38] The summation over the qubit index in eqn (51) is rather verbose, but can be understood as alternating between all pairs of nearest-neighbor spatial orbitals, with alternating 0 (even) and 1 (odd) starting spatial orbital. This pattern of summation indices is visually indicated in the P_X gates in Fig. 1. Note that this variant of the uCJ ansatz is similar in structure to the quantum number preserving fabric circuit[39] or the fermionic SWAP network implementation of k-UpCCGSD.[40] However, as our choice is not quite any of these ansatzes, for the remainder of this work we will name it k-muCJ for clarity, with the 'm' standing for modified. We stress again that the choice of VQE ansatz is largely irrelevant from a SAPT perspective and is not a major point in this paper.

The diagonal doubles operator in eqn (51) can be implemented as a product of four-qubit pair-exchange gates, $\hat{P}_\mathrm{X}(\theta)$, which have the action,

$$\hat{P}_\mathrm{X}(\theta) \equiv \begin{bmatrix} 1 & & & & & & & & & & & & & & & \\ & 1 & & & & & & & & & & & & & & \\ & & 1 & & & & & & & & & & & & & \\ & & & c & & & & & & & & & & -s & & \\ & & & & 1 & & & & & & & & & & & \\ & & & & & 1 & & & & & & & & & & \\ & & & & & & 1 & & & & & & & & & \\ & & & & & & & 1 & & & & & & & & \\ & & & & & & & & 1 & & & & & & & \\ & & & & & & & & & 1 & & & & & & \\ & & & & & & & & & & 1 & & & & & \\ & & & & & & & & & & & 1 & & & & \\ & & & +s & & & & & & & & & c & & & \\ & & & & & & & & & & & & & 1 & & \\ & & & & & & & & & & & & & & 1 & \\ & & & & & & & & & & & & & & & 1 \end{bmatrix} \quad (52)$$

in the four-qubit Hilbert space and in the above we use the shorthand $s = \sin(\theta)$ and $c = \cos(\theta)$. Note that, this operator implements a partial double-substitution in the closed-shell space of the four-qubit Hilbert space.[39] The decomposition of eqn (52) into more standard gates is given in ref. 39. An example of one layer of the muCJ circuit ansatz is given in Fig. 1.

With an ansatz of the form of eqn (48) we can write the VQE objective function as

$$E_{\mathrm{VQE}}(\kappa_{pq}^k, \tau_{pq}^k) \equiv \langle \Psi_{\mathrm{VQE}}(\kappa_{pq}^k, \tau_{pq}^k)|\hat{H}|\Psi_{\mathrm{VQE}}(\kappa_{pq}^k, \tau_{pq}^k)\rangle \quad (53)$$

$$= \langle \Phi_\mathrm{I}|\hat{U}^\dagger(\kappa_{pq}^k, \tau_{pq}^k)\hat{H}\hat{U}(\kappa_{pq}^k, \tau_{pq}^k)|\Phi_\mathrm{I}\rangle. \quad (54)$$

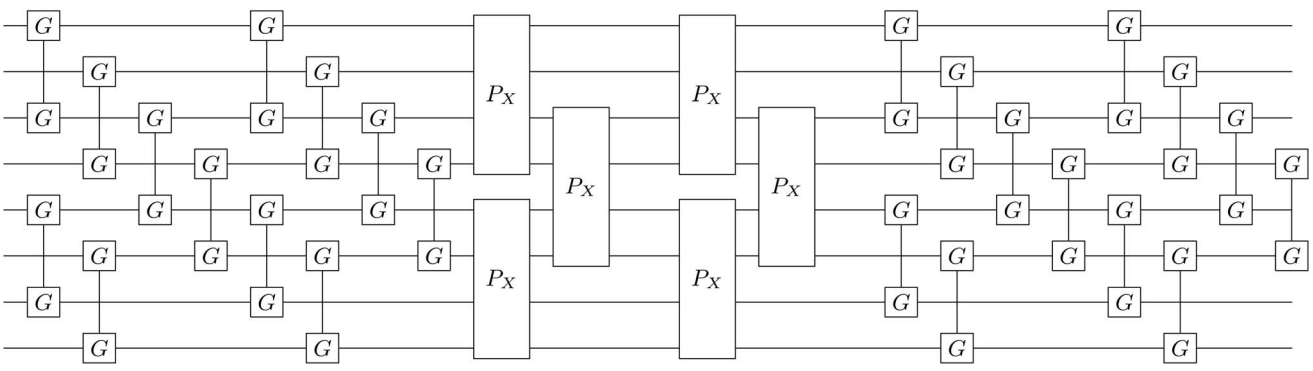


Fig. 1 Single layer ($k = 1$) k-muCJ VQE entangler circuit used in this work sketched for $M = 4$ spatial orbitals or $N = 8$ qubits. Even (odd) qubits label α (β) spin-orbitals. The circuit consists of a single layer of orbital rotations which are implemented as a series of two-qubit Givens rotations, followed by a layer of diagonal double substitution operators implemented *via* the four-qubit pair-exchange gate given in eqn (52), followed by another layer of orbital rotations. Note that for the spin-restricted ansatz used here the angles in the Givens gates for α and β spin-orbitals are the same.

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

The VQE algorithm then proceeds in hybrid form by using the quantum computer to evaluate eqn (54) before the variational parameters $\{\kappa_{pq}^k, \tau_{pq}^k\}$ are updated using a classical optimization algorithm.

To estimate the one- and two-particle reduced density matrices we write the expectation values in eqn (9) and (16) in terms of the Jordan-Wigner strings. The number of measurements scales like $\mathcal{O}(N_a^4)$ which can be reduced to an extent by accounting for symmetries in the one- and two-particle density matrices. Although not a focus of this paper, efficiently estimating density matrices is an active area of research and techniques exist to increase the parallelism of the measurements[41] and for error mitigation.[42]

II. Results

A Computational details

For the idealized experiments presented here we use the L-BFGS-B algorithm provided by scipy[43] to optimize the VQE objective function and used the RHF state as the initial wavefunction. Classical SCF computations and integral generation was performed using Terachem[32–34] interfaced through Lightspeed.[44] The GPU accelerated ideal VQE simulations were implemented using the quasar/vulcan codes. Double factorization was used for evaluating the total energy of the VQE ansatz.[45,46] VMD[47] was used for visualizing molecular orbitals and molecular structures with the exception of the KDM5A system which used MOE.[48] A sketch of the workflow is given in Fig. 2.

B Multi reference benchmarks

To assess the accuracy of SAPT(VQE) we will apply the method to investigate the intermolecular electrostatic and exchange energy to a selection of dimers from the S22 benchmark set.[49] We will modify these systems to ensure one of the monomers in question has some multi reference character that cannot be well described by SAPT(RHF). Note that previous studies of SAPT(CASSCF) have found little benefit of the method for single-reference problems in the S22 benchmark set,[26] so we only focus on more challenging multi-reference problems. In all examples we will benchmark our results against classical SAPT(CASCI) that employs the same active space as the SAPT(VQE) computations.

For our first test case we will investigate a benzene–*p*-benzyne dimer arranged in the T-shape configuration visualized in Fig. 3(a). This is a variation on the classic benzene dimer SAPT benchmark, except *p*-benzyne has a biradical ground state and thus benefits from a multi-reference approach. To describe *p*-benzyne we construct a (6e, 6o) active space from the HOMO-2 to LUMO+2 RHF/cc-pVDZ MOs and treat the benzene monomer at the RHF level of theory. Previous results suggest that *p*-benzyne represents a challenge for VQE with the number of parameters required to reach chemical accuracy being roughly half the number of configuration state functions in the exact solution.[39] Thus, it represents an interesting test case to see how errors from VQE propagate through to the resultant SAPT energy components.

As can be seen from Fig. 3, the *p*-benzyne molecule does represent a challenge for the *k*-muCJ ansatz, with a $k > 5$ required to reach an error in the total energy below 1 kcal mol^{-1}. In contrast we see that the errors in the SAPT(VQE) electrostatic and exchange energies are 2–4 orders of magnitude smaller than the corresponding VQE total energy and that a shallow depth $k = 1$ VQE ansatz would be sufficient for sub kcal mol^{-1} accuracy in these terms. We see this holds across the dissociation curve, which is expected as the VQE solution is fixed when working in a monomer-centered basis set, which is another advantage of SAPT(VQE). It should be noted that the SAPT(RHF) binding energy for this dimer system is on the order of 2 kcal mol^{-1} for this basis set.

For the next test case we investigate the ability of SAPT(VQE) to address intermolecular interactions involving bond dissociation in one of the monomers. We calculate the SAPT electrostatic and exchange energies for two interacting water molecules (displayed in Fig. 4(a)), one in its equilibrium geometry and in the other we symmetrically stretch the two OH bonds towards dissociation. This is again a challenging case for VQE to capture a double bond breaking. We used the 6-31G basis set and chose a (6e, 6o) active space for water from HOMO−2 to LUMO+2. Again, we see in Fig. 4 that RHF qualitatively fails to capture either the total energy of the stretched water monomer or the SAPT energy components of the dimer system. On the other hand, while sizeable errors are present in the VQE total energy (>10 kcal mol^{-1}) in the stretched system for

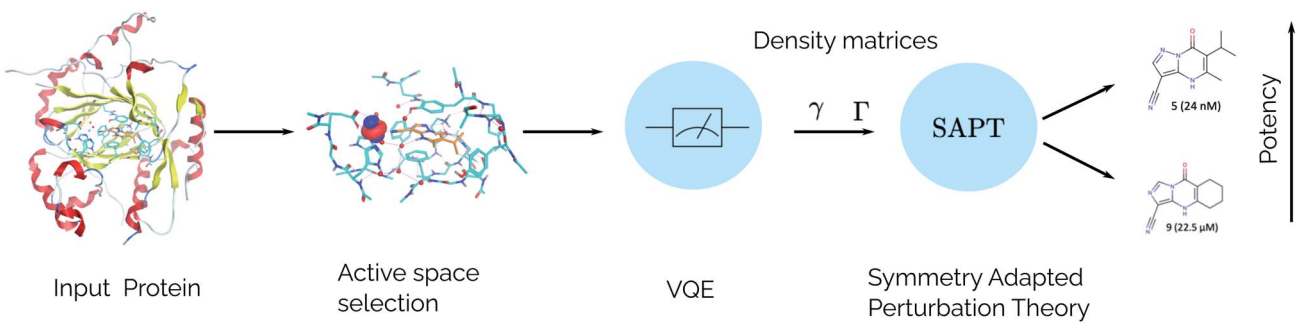


Fig. 2 SAPT(VQE) workflow for computing intermolecular interaction energies. In the first stage the input structures for the protein and ligand are provided. Following this, the structures are typically relaxed using classical molecular dynamics and DFT and a model system of the binding site is cutout from the full structure. At this stage an active space for the protein and/or ligand can be determined. Next, a quantum algorithm like the VQE is used to determine the one- and two-particle reduced density matrices which serve as input for SAPT. The SAPT interaction energy is finally computed as a classical post-processing step. Note that only the determination of the one- and two-particle density matrices (γ and Γ respectively) should be performed on a quantum computer.

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

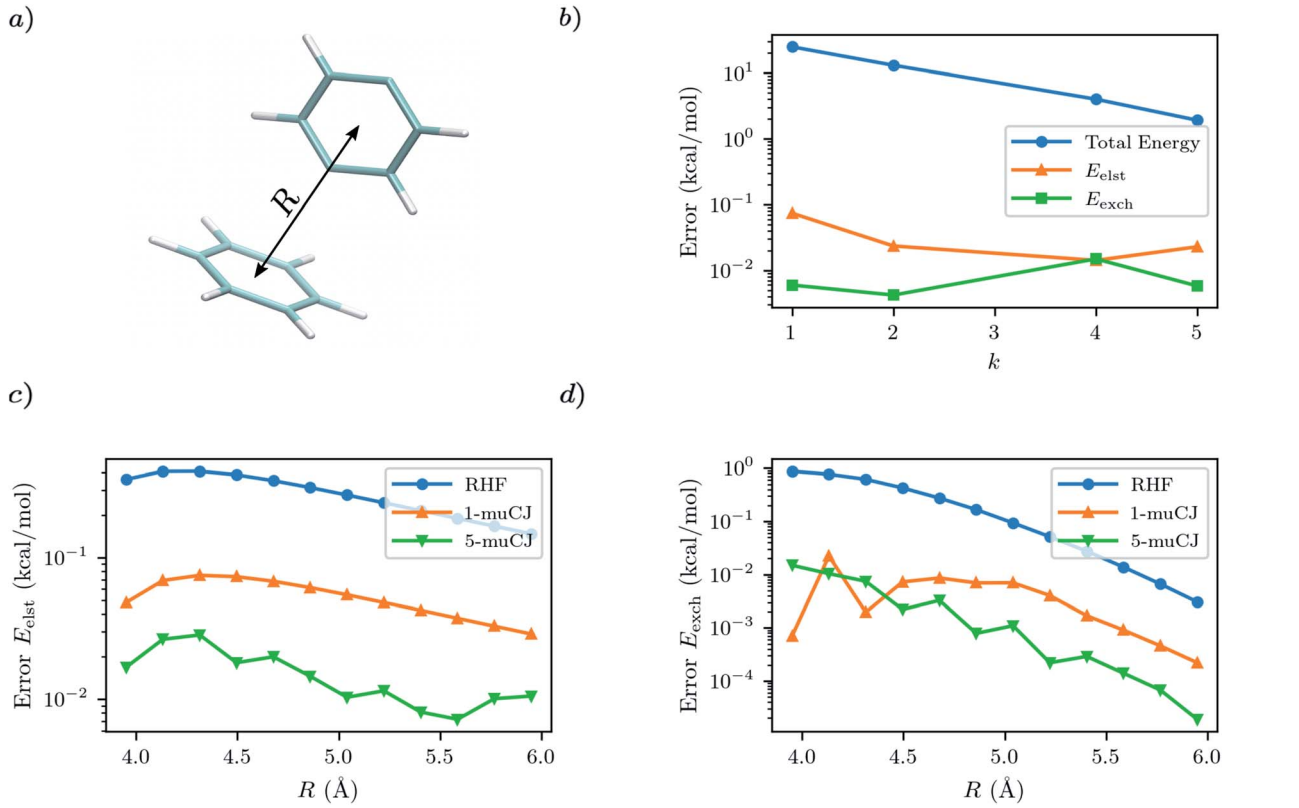


Fig. 3 (a) T-shaped configuration of benzene–*p*-benzyne dimer with the center-to-center intermonomer distance labelled as *R*. (b) Absolute error in *p*-benzyne monomer VQE total energy compared to errors in the SAPT(VQE) electrostatic and exchange energies at $R = 4.45$ Å intermonomer separation as a function of the circuit repetition factor *k* in the *k*-muCJ ansatz. (c and d) Absolute errors relative to SAPT(CASCI) in the electrostatic and exchange energies calculated using different monomer wavefunctions for *p*-benzyne as a function of the intermonomer separation.

$k = 1$, we find sub 0.1 kcal mol^{-1} accuracy in the SAPT energy components.

It is interesting to note that while the quality of the total energy necessarily improves when the circuit repetition factor (k) is increased (assuming a local minima is not arrived at during optimization), some non-monotonic behaviour is observed in the individual SAPT energy components. This is probably due to the difficulty in tightly converging the VQE total energy in log scale in these figures.

These two examples provide strong evidence that SAPT(VQE) yields significantly lower absolute errors in the target interaction energy contributions *vs.* those of the corresponding monomer total energies. This helps to reduce the depth of circuits necessary to achieve accurate interaction energy components.

C Protein–ligand interactions

For our final example we examine the ability of our implementation of SAPT(VQE) to tackle large protein–ligand interactions, specifically lysine-specific demethylase 5A (KDM5A) depicted in Fig. 5(a). KDM5A is believed to be relevant for human cancers[23] and contains a metal center [Fe(II)] which may pose a challenge for classical electronic structure theory methods. To make the problem tractable, we used a model system of the binding site which was cut out from the full binding domain of KDM5A and focuses on the immediate surrounding of metal ion and ligand. The detailed preparation of the model system is described in the Supporting Information. This final structures contained each 380 atoms and were studied with DFT using a 6-31G basis set and the ωB97X-D functional[53] using the Gaussian software package.[54] The ligand and all atoms in radius of 4.5 Å of the ligand were further relaxed with the oxygen atoms of the water and the iron atom been keeping fixed. The structures have an electronic size of (1482e, 2214o). The protein–ligand system is visualized in Fig. 5(b). We computed the first order SAPT(VQE) energy contributions for 5 different inhibitors from ref. 52 and shown in Fig. 5(c) to assess how important multi-reference effects were for describing the interaction energy with this protein. We treated each of the ligands at the RHF level and KDM5A using VQE.

To construct a potentially representative active space for KDM5A we used the AVAS procedure,[55] and first built a relatively large space by including the 3d orbitals from the iron center, the oxygen 2p orbitals from the nearest two water molecules and the 2p orbitals from the neighbouring oxygen atom associated with the glutamic acid, as well as two 2p orbitals from the neighbouring nitrogen atoms from two neighbouring histidines. A representative Fe 3d-like orbital from this procedure is depicted in Fig. 5(b). This leads to an active space size from the AVAS procedure of (36e, 27o), which would require a 54 qubit quantum computer, and is thus outside the reach of current simulators and hardware. To reduce the size of the active space we performed a loosely converged ($\varepsilon = 1 \times 10^{-4}$) selected heat bath configuration interaction computation

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

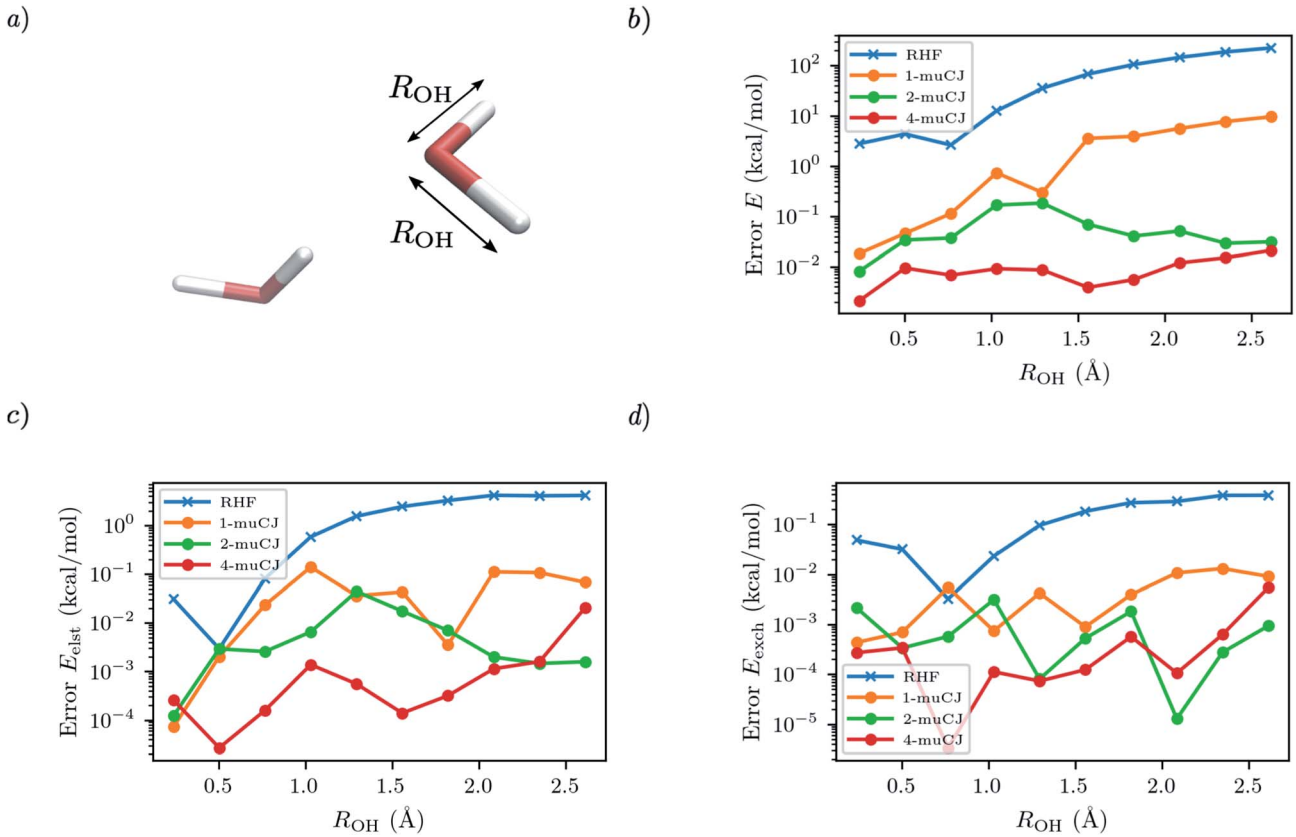


Fig. 4 (a) Geometry of the water dimer studied with the OH bond length of single monomer distance labelled as R_{OH}. The intermonomer separation is fixed throughout. (b) Absolute error in the stretched water monomer VQE total energy as a function of R_{OH} for different muCJ repetition factor k in the k-muCJ ansatz. (c and d) Absolute errors relative to SAPT(CASCI) in the electrostatic and exchange energies calculated as a function of R_{OH} for different muCJ repetition factor k in the k-muCJ ansatz.

(SHCI)[56,57] in this active space. We then constructed a smaller active space using the SHCI natural orbitals (NOs), only keeping those NOs with occupation $0.02 \leq n_i \leq 1.97$. This leads to a smaller active space of (8e, 8o), for the low-spin configuration, corresponding to a 16 qubits VQE simulation which is possible using simulators and within reach of current hardware. The natural orbital occupation structure and dominant single reference nature of SHCI ground state suggests that the KDM5A problem will not be challenging for VQE.

In order to connect with experiment we score the ligands based on the difference of their interaction energy relative to a reference ligand

$$\Delta E_{int}(x) = E_{int}(x_{ref}) - E_{int}(x) \quad (55)$$

where x are the ligand labels from Fig. 5(c) and we choose $x_{ref} = 5$ as the reference point as it is experimentally the most potent ligand from the subset of ligands taken from ref. 52. We then compare this energy difference to differences in the experimental free energy of binding[58]

$$\Delta\Delta G = \Delta G(x_{ref}) - \Delta G(x) = -RT \log\left(\frac{IC_{50}(x_{ref})}{IC_{50}(x)}\right), \quad (56)$$

where R is the gas constant, the temperature T is taken to be room temperature and we took the IC_{50} values from ref. 52. Note that we are comparing two different quantities, namely the theoretical interaction energy difference and the experimental free energy of binding and thus we cannot necessarily expect any quantitative relationship. For example, the interaction energies account for one leg of the thermodynamic cycle while the free energy of binding accounts for the whole cycle. In addition, the number of water molecules surrounding the binding site may depend on the specific ligand being studied, with larger ligands replacing more water molecules. While neither of these effects was considered here it is interesting to see what relationship if any these numbers have to one another.

The results of these computations are shown in Fig. 5(d). To get some insight into the accuracy of existing approaches we first present supermolecular DFT(ωB97X-D), RHF and SAPT$^{(2)}$(RHF) the results of which are shown in the upper panel of Fig. 5(d). Note that here SAPT$^{(2)}$(RHF) contains both first- and second-order SAPT energy contributions[22] and were performed in a dimer-centered basis set, this in contrast to the rest of the results in this paper which only considered the first-order SAPT contributions in a monomer-centered basis. The DFT results are not counterpoise corrected but we found this consideration unimportant for the present system. We find that supermolecular RHF incorrectly predicts ligand-4 to have the largest interaction energy suggesting that a more accurate method is required. This observation is

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

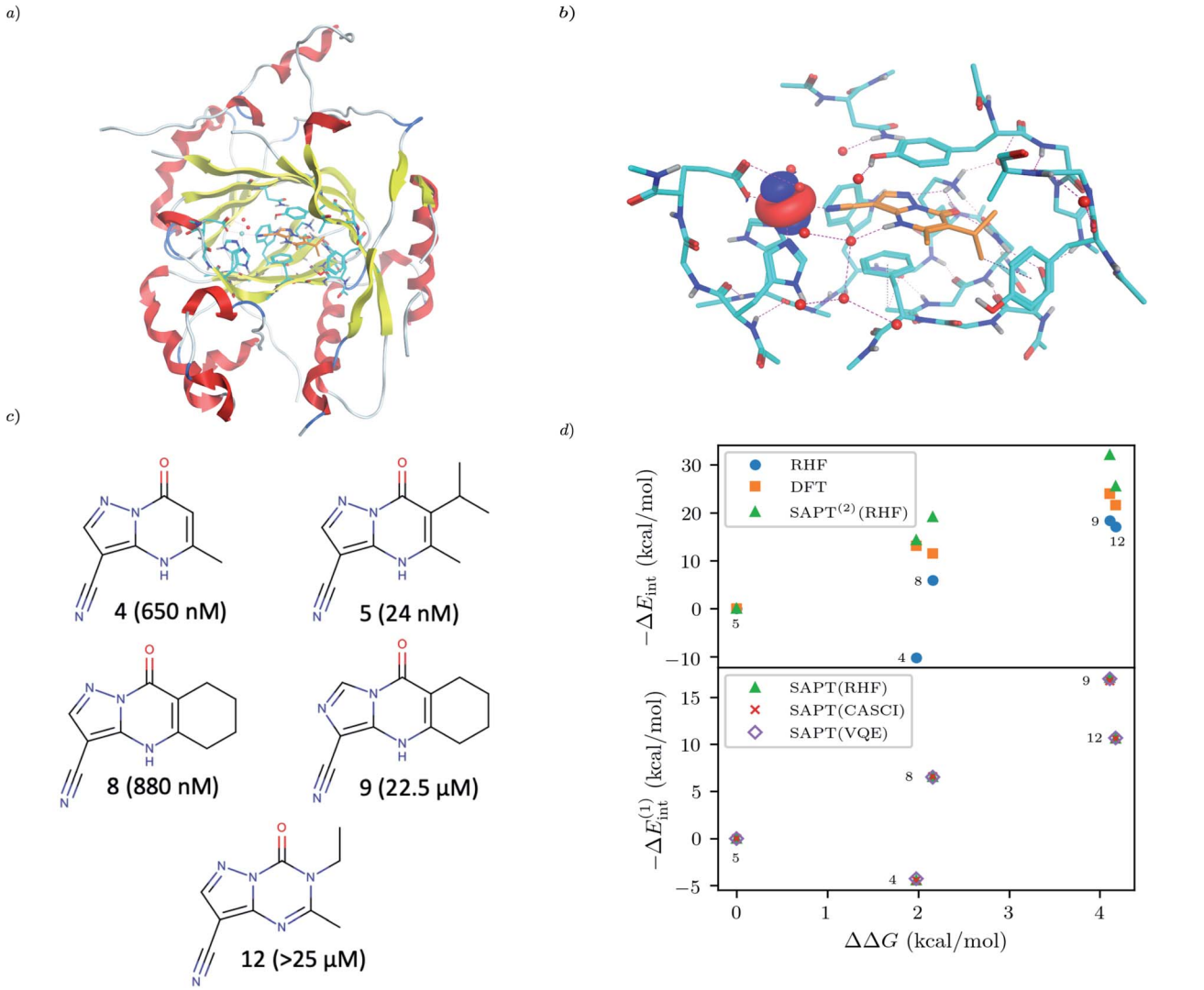


Fig. 5 (a) KDM5A protein structure[50,51] (b) representative minimal active space molecular orbitals produced by the AVAS procedure, for the protein cutout described in the main text with ligand-5 colored as orange (c) 5 KDM5A inhibitors from ref. 52, (d) Upper Panel: Interaction energy differences computed using the super molecular approach (DFT(ωB97X-D) and RHF) and SAPT$^{(2)}$(RHF) as a function of the difference in experimental free energies (see main text for definition). Lower Panel: First-order SAPT(VQE), SAPT(CASCI) and SAPT(RHF) first order interaction energy differences as a function of the differences in experimental free energies of binding. Note that $\Delta E_{\rm int}(x) = E_{\rm int}(x_{\rm ref}) - E_{\rm int}(x)$ so $-\Delta E_{\rm int}$ will be positive if the ligand is predicted to have a smaller interaction energy (*i.e.* less negative) than the reference ligand (here taken to be ligand-5).

confirmed by the SAPT(RHF) and supermolecular DFT results which correctly score ligand-5 as the most potent in this set.

In the lower panel of Fig. 5(d) we plot the first order interaction energy difference, $\Delta E^{(1)}_{\rm int}$ computed through SAPT(RHF), SAPT(CASCI) and SAPT(VQE) all computed using a monomer-centered basis. The first thing we see is the excellent agreement between SAPT(CASCI) and SAPT(VQE) for each ligand considered. Note that we used the 1-muCJ ansatz for the VQE computations which offers further evidence of the error reduction capabilities of SAPT. Apart from this excellent agreement we also see that there is little difference between SAPT(RHF) and SAPT(CASCI). We found that the maximum difference between SAPT(CASCI) and SAPT(RHF) was roughly 1–2 kcal mol^{-1}, however this error nearly cancels when looking at the $\Delta E_{\rm int}$ and is thus not visible on the scale of the plot. Ultimately this is not surprising given that the problem was found to be not strongly correlated.

Finally, we see that first order SAPT is not accurate enough to correctly score the respective ligands with the missing induction and dispersion components proving critical to achieving this. Indeed the first order interaction energy for ligand-5 is positive and thus unbound at this level of theory. In particular, from the full SAPT$^{(2)}$(RHF) energy decomposition (see Supplementary Information) we find that the induction components for ligand-4 and ligand-5 are similar in magnitude while the dispersion component for ligand-5 is roughly 1.4 times that of ligand-4. This large dispersion energy ensures that ligand-5 has the larger (absolute) interaction energy. Nevertheless, we see that SAPT(VQE) can potentially tackle industrially relevant drug design problems, particularly in cases where it is not known *a priori* how strongly correlated the ground state is.

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

III. Conclusion

We described the theory and implementation of SAPT on a NISQ-era quantum computer, focusing on the efficient implementation and classical workflows necessary to tackle industrially relevant problems in drug design. We derived in detail the equations necessary for an active space formulation of SAPT(VQE) that requires only the one- and two-particle reduced density matrices measured on a quantum computer. The SAPT(VQE) components are computed as an efficient classical post-processing step. This classical post-processing step, written in terms of optimized quantum chemical primitives, was shown capable of simulating systems with hundreds of atoms and potentially hundreds of basis functions in the active space.

Beyond an efficient classical implementation, we found (through ideal VQE simulations) that SAPT naturally reduces the error incurred by approximately solving the Schrödinger equation on a quantum computer, which we attribute to the theory directly computing energy differences. This fact coupled with a monomer basis formulation will substantially reduce the resource requirements for computing binding energies of large protein–ligand interactions on NISQ-era quantum computers. In particular, sub-kcal mol^{-1} accuracy in the energy components for the electrostatic and exchange energies can be computed using coarse VQE wavefunctions, which otherwise exhibit gross errors (>10 kcal mol^{-1}) in the total energy. Given the practical challenges associated with optimizing VQE wavefunctions on current hardware we believe SAPT may help to extend the scope of the method due to the apparent ability to use low depth VQE ansatzes like 1-uCJ. SAPT(VQE) appears to offer, then, a reduction in the quantum resources required compared to a simple supermolecular VQE for the computation of protein–ligand interaction energies. This reduction in qubit count and circuit depth naturally yields a lower precision requirement on the quantum circuit and thus has the potential to enable the simulation of larger systems on NISQ hardware than what is possible with supermolecular VQE implementations.

In the future, it will be critical to determine how robust SAPT(VQE) is to noise channels either through modeling or real hardware experiments. Another important question will be how to reduce the measurement overhead of accumulating the one- and two-particle reduced density matrices, a tentative solution to which is sketched in Appendix B. A natural extension will be to determine the second-order SAPT terms that would allow for accurate interaction energies and induction and dispersion energy components to be computed. Other interesting questions are how to increase the quantitative accuracy of the method by including correlation out of the active space[18,59] and to gather more challenging drug–protein systems potentially containing multiple metal centers.

IV. Appendix A: full active space SAPT exchange expressions

In this appendix we give full expressions for the remaining exchange energy contributions (T_3 and T_4) not given in Section ID. Note that we will often perform optimizations similar to those provided by eqn (41) with the general strategy to be the removal of core orbital indices from the generalized ERI expressions through the formation of appropriate intermediate matrices.

T_3 is very similar to T_2 and we have

$$\begin{aligned} T_3 = {} & T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{ac}}\right] + T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{cc}}\right] \\ & + T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{ac}}\right] + T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{aa}}\right] \\ & + T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{aa}}\right] + T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{cc}}\right], \end{aligned} \tag{A1}$$

where

$$\begin{aligned} T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{ac}}\right] = {} & 4 S_{ji} C_{\nu j}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}}\right]_{\nu\mu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}}\right]_{\nu\mu''}\right) C_{\mu'' i} \\ & + 2\overline{\gamma}_{tt'} S_{jt} C_{\nu j}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{\nu\mu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{\nu\mu''}\right) C_{\mu'' t'}, \end{aligned} \tag{A2}$$

$$T_3\left[\gamma_{\mathrm{B}}^{c}, \Gamma_{\mathrm{A}}^{\mathrm{cc}}\right] = \overline{\gamma}_{\nu\nu'}^{\mathrm{c}} S_{\nu'\mu'} \overline{\gamma}_{\mu'\mu'''}^{\mathrm{c}}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{\nu\mu'''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{\nu\mu'''}\right), \tag{A3}$$

$$\begin{aligned} T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{ac}}\right] = {} & 2\overline{\gamma}_{uu'} S_{u'i} C_{\nu u}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}}\right]_{\nu\mu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}}\right]_{\nu\mu''}\right) C_{\mu'' i} \\ & + \overline{\gamma}_{uu'}\overline{\gamma}_{tt'} S_{u't} C_{\nu u}\left(\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{\nu\mu''} - \frac{1}{2}\tilde{K}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{\nu\mu''}\right) C_{\mu'' t'}, \end{aligned} \tag{A4}$$

$$T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{aa}}\right] = 2\overline{\Gamma}_{t''t'''}^{tt'} V_{tt'}^{t''t'''}, \tag{A5}$$

where as in eqn (41) we have defined

$$V_{tt'}^{t''t'''} = \sum_j S_{jt'} \tilde{v}_{tj}^{t''t'''}, \tag{A6}$$

and, lastly,

$$T_3\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}, \overline{\Gamma}_{\mathrm{A}}^{\mathrm{aa}}\right] = \overline{\gamma}_{uu'}\overline{\Gamma}_{t''t'''}^{tt'} S_{u't'} V_{tu}^{t''t'''}. \tag{A7}$$

For T_4 there are sixteen terms in total. First we have

$$T_4\left[\overline{\Gamma}_{\mathrm{A}}^{\mathrm{aa}}, \overline{\Gamma}_{\mathrm{B}}^{\mathrm{aa}}\right] = \overline{\Gamma}_{tt'}^{t''t'''}\overline{\Gamma}_{uu'}^{u''u'''} S_{t'''u'} S_{t'u'''} \tilde{v}_{tu}^{t''u''}, \tag{A8}$$

which can't be simplified further. Next we have

$$\begin{aligned} T_4\left[\overline{\Gamma}_{\mathrm{A}}^{\mathrm{aa}}, \overline{\Gamma}_{\mathrm{B}}^{\mathrm{ac}}\right] = {} & \overline{\Gamma}_{tt'}^{t''t'''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t'''j} S_{t'j''} \tilde{v}_{tu}^{t''u''} \\ & - \frac{1}{2}\overline{\Gamma}_{tt'}^{t''t'''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t'''j} S_{t'u''} \tilde{v}_{tu}^{t''j''} \\ & + \overline{\Gamma}_{tt'}^{t''t'''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t'''u} S_{t'u''} \tilde{v}_{tj}^{t''j''} \\ & - \frac{1}{2}\overline{\Gamma}_{tt'}^{t''t'''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t'''u} S_{t'j''} \tilde{v}_{tj}^{t''u''}, \end{aligned} \tag{A9}$$

which can be written as

$$\begin{aligned} T_4\left[\overline{\Gamma}_{\mathrm{A}}^{\mathrm{aa}}, \overline{\Gamma}_{\mathrm{B}}^{\mathrm{ac}}\right] = {} & 2\overline{\Gamma}_{tt'}^{t''t'''} S_{t'''j} S_{t'j} \tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}\right]_{tt''} \\ & - \overline{\Gamma}_{tt'}^{t''t'''}\overline{\gamma}_{uu''} S_{t'u''} X_{tu}^{t''t'''} \\ & + \overline{\Gamma}_{tt'}^{t''t'''}\overline{\gamma}_{uu''} S_{t'''u} S_{t'u''} \tilde{J}\left[\overline{\gamma}_{B}^{\mathrm{c}}\right]_{tt''} \\ & - \overline{\Gamma}_{tt'}^{t''t'''}\overline{\gamma}_{uu''} S_{t'''u} Y_{tt'}^{t''u''}, \end{aligned} \tag{A10}$$

where

© 2022 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$X^{t''t'''}_{tu} \equiv \sum_{j''} S_{t'''j''} \tilde{v}^{t'''j''}_{tu} \tag{A11}$$

$$Y^{t''u''}_{tt'} \equiv \sum_{j} S_{t'j} \tilde{v}^{t''u''}_{tj}. \tag{A12}$$

Similarly

$$\begin{aligned} T_4\left[\overline{\Gamma}^{\mathrm{ac}}_{\mathrm{A}}, \overline{\Gamma}^{\mathrm{aa}}_{\mathrm{B}}\right] = {} & \overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\Gamma}^{u''u'''}_{uu'} S_{i''u'} S_{i'u'''} \tilde{v}^{t''u''}_{tu} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\Gamma}^{u''u'''}_{uu'} S_{t''u'} S_{iu'''} \tilde{v}^{i''u''}_{tu} \\ & +\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\Gamma}^{u''u'''}_{uu'} S_{t''u'} S_{tu'''} \tilde{v}^{i''u''}_{iu} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\Gamma}^{u''u'''}_{uu'} S_{i''u'} S_{tu'''} \tilde{v}^{t''u''}_{iu}, \end{aligned} \tag{A13}$$

which can be written as

$$\begin{aligned} T_4\left[\overline{\Gamma}^{\mathrm{ac}}_{\mathrm{A}}, \overline{\Gamma}^{\mathrm{aa}}_{\mathrm{B}}\right] = {} & 2\overline{\Gamma}^{u''u'''}_{uu'} S_{iu'} S_{iu'''} \tilde{J}\left[\overline{\gamma}^{\mathrm{a}}_{\mathrm{A}}\right]_{uu''} \\ & -\overline{\gamma}_{tt''}\overline{\Gamma}^{u''u'''}_{uu'} S_{t''u'} Z^{u'''u''}_{tu} \\ & +\overline{\gamma}_{tt''}\overline{\Gamma}^{u''u'''}_{uu'} S_{t''u'} S_{tu'''} \tilde{J}\left[\overline{\gamma}^{c}_{A}\right]_{uu''} \\ & -\overline{\gamma}_{tt''}\overline{\Gamma}^{u''u'''}_{uu'} S_{tu'''} U^{t''u''}_{u'u}, \end{aligned} \tag{A14}$$

where

$$Z^{u'''u''}_{tu} \equiv \sum_{i''} S_{i''u'''} \tilde{v}^{t''u''}_{tu} \tag{A15}$$

$$U^{t''u''}_{u'u} \equiv \sum_{i} S_{iu'} \tilde{v}^{t''u''}_{iu}. \tag{A16}$$

Next

$$\begin{aligned} T_4\left[\overline{\Gamma}^{\mathrm{aa}}_{\mathrm{A}}, \overline{\Gamma}^{\mathrm{cc}}_{\mathrm{B}}\right] = {} & \overline{\Gamma}^{t''t'''}_{tt'} \overline{\gamma}_{jj''}\overline{\gamma}_{j'j'''} S_{t'''j'} S_{t'j'''} \tilde{v}^{t''j''}_{tj} \\ & -\frac{1}{2}\overline{\Gamma}^{t''t'''}_{tt'} \overline{\gamma}_{jj'''}\overline{\gamma}_{j'j''} S_{t'''j'} S_{t'j''} \tilde{v}^{t''j''}_{tj}, \end{aligned} \tag{A17}$$

which can be written as

$$T_4\left[\overline{\Gamma}^{\mathrm{aa}}_{\mathrm{A}}, \overline{\Gamma}^{\mathrm{cc}}_{\mathrm{B}}\right] = 2\overline{\Gamma}^{t''t'''}_{tt'} S_{t'''j'} S_{t'j'} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{B}}\right]_{tt''} - 2\overline{\Gamma}^{t''t'''}_{tt'} R^{t''t'''}_{tt'}, \tag{A18}$$

where

$$R^{t''t'''}_{tt'} \equiv \sum_{jj''} S_{t'''j''} S_{t'j} \tilde{v}^{t''j''}_{tj}, \tag{A19}$$

and similarly

$$\begin{aligned} T_4\left[\overline{\Gamma}^{\mathrm{cc}}_{\mathrm{A}}, \overline{\Gamma}^{\mathrm{aa}}_{\mathrm{B}}\right] = {} & \overline{\gamma}_{ii''}\overline{\gamma}_{i'i'''}\overline{\Gamma}^{u''u'''}_{uu'} S_{i'''u'} S_{i'u'''} \tilde{v}^{i''u''}_{iu} \\ & -\frac{1}{2}\overline{\gamma}_{ii'''}\overline{\gamma}_{i'i''}\overline{\Gamma}^{u''u'''}_{uu'} S_{i'''u'} S_{i'u'''} \tilde{v}^{i'''u''}_{iu}, \end{aligned} \tag{A20}$$

which can be written as

$$T_4\left[\overline{\Gamma}^{\mathrm{cc}}_{\mathrm{A}}, \overline{\Gamma}^{\mathrm{aa}}_{\mathrm{B}}\right] = 2\overline{\Gamma}^{u''u'''}_{uu'} S_{i'u'} S_{i'u'''} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{A}}\right]_{uu''} - 2\overline{\Gamma}^{u''u'''}_{uu'} A^{u'''u''}_{u'u}. \tag{A21}$$

$$A^{u'''u''}_{u'u} \equiv \sum_{ii''} S_{iu'} S_{i''u'''} \tilde{v}^{t''u''}_{iu}. \tag{A22}$$

For the AC–AC contribution we get sixteen terms

$$\begin{aligned} T_4\left[\overline{\Gamma}^{ac}_{\mathrm{A}}, \overline{\Gamma}^{ac}_{\mathrm{B}}\right] = {} & \overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''j} S_{ij''} \tilde{v}^{t''u''}_{tu} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''j} S_{iu''} \tilde{v}^{t''j''}_{tu} + \overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''u} S_{iu''} \tilde{v}^{t''j''}_{tj} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''u} S_{ij''} \tilde{v}^{t''u''}_{tj} - \frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''j} S_{ij''} \tilde{v}^{i''u''}_{tu} \\ & +\frac{1}{4}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''j} S_{iu''} \tilde{v}^{i''j''}_{tu} - \frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''u} S_{iu''} \tilde{v}^{i''j''}_{tj} \\ & +\frac{1}{4}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''u} S_{ij''} \tilde{v}^{i''u''}_{tj} + \overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''j} S_{tj''} \tilde{v}^{i''u''}_{iu} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''j} S_{tu''} \tilde{v}^{i''j''}_{iu} + \overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''u} S_{tu''} \tilde{v}^{i''j''}_{ij} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{t''u} S_{tj''} \tilde{v}^{i''u''}_{ij} - \frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''j} S_{tj''} \tilde{v}^{t''u''}_{iu} \\ & +\frac{1}{4}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''j} S_{tu''} \tilde{v}^{t''j''}_{iu} - \frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''u} S_{tu''} \tilde{v}^{t''j''}_{ij} \\ & +\frac{1}{4}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''} S_{i''u} S_{tj''} \tilde{v}^{i''u''}_{ij}, \end{aligned} \tag{A23}$$

which can be simplified to

$$\begin{aligned} T_4\left[\overline{\Gamma}^{ac}_{\mathrm{A}}, \overline{\Gamma}^{ac}_{\mathrm{B}}\right] = {} & 4\overline{\gamma}_{tt''} S_{ij} S_{ij} \tilde{J}\left[\overline{\gamma}^{\mathrm{a}}_{\mathrm{B}}\right]_{tt''} - 2\overline{\gamma}_{uu''} S_{ij} S_{iu''} \tilde{J}\left[\overline{\gamma}^{\mathrm{a}}_{\mathrm{A}}\right]_{uj} \\ & +2\overline{\gamma}_{tt''}\overline{\gamma}_{uu''} S_{iu} S_{iu''} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{B}}\right]_{tt''} - 2\overline{\gamma}_{uu''} S_{iu} S_{ij} \tilde{J}\left[\overline{\gamma}^{\mathrm{a}}_{\mathrm{A}}\right]_{ju''} \\ & -2\overline{\gamma}_{tt''} S_{t''j} S_{ij} \tilde{J}\left[\overline{\gamma}^{\mathrm{a}}_{\mathrm{B}}\right]_{ti} + \overline{\gamma}_{uu''} S_{iu''} \tilde{K}\left[\overline{\gamma}_{tt''} S_{t''j}\right]_{iu} \\ & -\overline{\gamma}_{tt''}\overline{\gamma}_{uu''} S_{t''u} S_{iu''} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{B}}\right]_{ti} + S_{ij} \tilde{K}[\overline{\gamma}_{tt''} S_{t''u}\overline{\gamma}_{uu''}]_{ij} \\ & +4\overline{\gamma}_{tt''} S_{t''j} S_{tj} \tilde{J}\left[\overline{\gamma}^{\mathrm{a}}_{\mathrm{B}}\right]_{ii} - \overline{\gamma}_{tt''}\overline{\gamma}_{uu''} S_{t''j} S_{tu''} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{A}}\right]_{uj} \\ & +2\overline{\gamma}_{tt''}\overline{\gamma}_{uu''} S_{t''u} S_{tu''} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{B}}\right]_{ii} - \overline{\gamma}_{tt''}\overline{\gamma}_{uu''} S_{t''u} S_{tj} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{A}}\right]_{ju''} \\ & -2\overline{\gamma}_{tt''} S_{ij} S_{tj} \tilde{J}\left[\overline{\gamma}^{\mathrm{a}}_{\mathrm{B}}\right]_{it''} + \overline{\gamma}_{tt''}\overline{\gamma}_{uu''} S_{tu''} \tilde{K}\left[S_{ij}\right]_{t''u} \\ & -\overline{\gamma}_{tt''}\overline{\gamma}_{uu''} S_{iu} S_{tu''} \tilde{J}\left[\overline{\gamma}^{\mathrm{c}}_{\mathrm{B}}\right]_{it''} + \overline{\gamma}_{tt''} S_{ij} \tilde{K}[S_{iu}\overline{\gamma}_{uu''}]_{t''j}. \end{aligned} \tag{A24}$$

Next we have

$$\begin{aligned} T_4\left[\overline{\Gamma}^{\mathrm{ac}}_{\mathrm{A}}, \overline{\Gamma}^{\mathrm{cc}}_{\mathrm{B}}\right] = {} & \overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj''}\overline{\gamma}_{j'j'''} S_{i''j'} S_{ij'''} \tilde{v}^{t''j''}_{tj} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj'''}\overline{\gamma}_{j'j''} S_{i''j'} S_{ij'''} \tilde{v}^{t''j''}_{tj} \\ & -\frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj''}\overline{\gamma}_{j'j'''} S_{t''j'} S_{ij'''} \tilde{v}^{t''j''}_{tj} \\ & +\frac{1}{4}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj'''}\overline{\gamma}_{j'j''} S_{t''j'} S_{ij'''} \tilde{v}^{t''j''}_{tj} \\ & +\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj''}\overline{\gamma}_{j'j'''} S_{t''j'} S_{tj'''} \tilde{v}^{i''j''}_{ij} \end{aligned}$$

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

$$- \frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj'''}\overline{\gamma}_{j'j''}S_{t''j'}S_{tj'''}\tilde{v}_{ij}^{i'''j''}$$

$$- \frac{1}{2}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj''}\overline{\gamma}_{j'j'''}S_{i''j'}S_{tj'''}\tilde{v}_{ij}^{t''j'''}$$

$$+ \frac{1}{4}\overline{\gamma}_{tt''}\overline{\gamma}_{ii''}\overline{\gamma}_{jj'''}\overline{\gamma}_{j'jj''}S_{i''j'}S_{tj'''}\tilde{v}_{ij}^{t''j''}, \quad \text{(A25)}$$

which can be written as

$$T_4\left[\overline{\Gamma}_{\mathrm{A}}^{\mathrm{ac}}, \overline{\Gamma}_{\mathrm{B}}^{\mathrm{cc}}\right] = 8S_{ij'}S_{ij'}\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}}\right]_{jj} - 4S_{ij'}S_{ij}\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{a}}\right]_{jj'}$$

$$- 2\overline{\gamma}_{tt''}S_{t''j'}S_{ij'}\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{ti} + 2S_{ij}\tilde{K}\left[\overline{\gamma}_{tt''}S_{t''j'}\right]_{ij}$$

$$+ 4\overline{\gamma}_{tt''}S_{t''j'}S_{tj'}\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{ii} - 2\overline{\gamma}_{tt''}S_{t''j'}S_{tj}\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{jj'}$$

$$- 2\overline{\gamma}_{tt''}S_{ij'}S_{tj'}\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{it''} + 2\overline{\gamma}_{tt''}S_{tj}\tilde{K}\left[S_{ij'}\right]_{t''j}, \quad \text{(A26)}$$

and similarly

$$T_4\left[\overline{\Gamma}_{\mathrm{A}}^{\mathrm{cc}}, \overline{\Gamma}_{\mathrm{B}}^{\mathrm{ac}}\right] = \overline{\gamma}_{ii''}\overline{\gamma}_{i'i'''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''j}S_{i'j''}\tilde{v}_{iu}^{i''u''}$$

$$- \frac{1}{2}\overline{\gamma}_{ii''}\overline{\gamma}_{i'i''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''j}S_{i'u''}\tilde{v}_{iu}^{i''j''}$$

$$+ \overline{\gamma}_{ii''}\overline{\gamma}_{i'i'''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''u}S_{i'u''}\tilde{v}_{ij}^{i''j''}$$

$$- \frac{1}{2}\overline{\gamma}_{ii''}\overline{\gamma}_{i'i'''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''u}S_{i'j''}\tilde{v}_{ij}^{i''u''}$$

$$- \frac{1}{2}\overline{\gamma}_{ii'''}\overline{\gamma}_{i'i''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''j}S_{i'j''}\tilde{v}_{iu}^{i''u''}$$

$$+ \frac{1}{4}\overline{\gamma}_{ii''}\overline{\gamma}_{i'i''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''j}S_{i'u''}\tilde{v}_{iu}^{i''j''}$$

$$- \frac{1}{2}\overline{\gamma}_{ii'''}\overline{\gamma}_{i'i''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''u}S_{i'u''}\tilde{v}_{ij}^{i'''j''}$$

$$+ \frac{1}{4}\overline{\gamma}_{ii'''}\overline{\gamma}_{i'i''}\overline{\gamma}_{uu''}\overline{\gamma}_{jj''}S_{i'''u}S_{i'j''}\tilde{v}_{ij}^{i''u''} \quad \text{(A27)}$$

which can be written as

$$T_4\left[\overline{\Gamma}_{\mathrm{A}}^{\mathrm{cc}}, \overline{\Gamma}_{\mathrm{B}}^{\mathrm{ac}}\right] = 8S_{i'j}S_{i'j}\tilde{J}[\overline{\gamma}_{uu''}]_{ii} - 2\overline{\gamma}_{uu''}S_{i'j}S_{i'u''}\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{uj}$$

$$+ 4\overline{\gamma}_{uu''}S_{i'u}S_{i'u''}\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{ii} - 2\overline{\gamma}_{uu''}S_{i'u}S_{i'j}\tilde{J}\left[\overline{\gamma}_{\mathrm{A}}^{\mathrm{c}}\right]_{ju''}$$

$$- 4S_{ij}S_{i'j}\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{a}}\right]_{ii'} + 2\overline{\gamma}_{uu''}S_{i'u''}\tilde{K}\left[S_{ij}\right]_{i'u}$$

$$- 2\overline{\gamma}_{uu''}S_{iu}S_{i'u''}\tilde{J}\left[\overline{\gamma}_{\mathrm{B}}^{\mathrm{c}}\right]_{ii'} + 2S_{i'j}\tilde{K}[S_{iu}\overline{\gamma}_{uu''}]_{i'j}. \quad \text{(A28)}$$

Finally, we have

$$T_4\left[\overline{\Gamma}_{\mathrm{A}}^{\mathrm{cc}}, \overline{\Gamma}_{\mathrm{B}}^{\mathrm{cc}}\right] = \overline{\gamma}_{ii''}\overline{\gamma}_{i'i'''}\overline{\gamma}_{jj''}\overline{\gamma}_{j'j'''}S_{i'''j'}S_{i'j'''}\tilde{v}_{ii''jj''}$$

$$- \frac{1}{2}\overline{\gamma}_{ii''}\overline{\gamma}_{i'i'''}\overline{\gamma}_{jj'''}\overline{\gamma}_{j'j''}S_{i'''j'}S_{i'j'''}\tilde{v}_{ii''jj''}$$

$$- \frac{1}{2}\overline{\gamma}_{ii'''}\overline{\gamma}_{i'i''}\overline{\gamma}_{jj''}\overline{\gamma}_{j'j'''}S_{i'''j'}S_{i'j'''}\tilde{v}_{ii''jj''}$$

$$+ \frac{1}{4}\overline{\gamma}_{ii'''}\overline{\gamma}_{i'i'''}\overline{\gamma}_{jj''}\overline{\gamma}_{j'j''}S_{i'''j'}S_{i'j'''}\tilde{v}_{ii''jj''}, \quad \text{(A29)}$$

which can be written in terms of $\tilde{J}$ and $\tilde{K}$ matrices as was done for the SAPT(HF).

V. Appendix B: quantum–classical optimization for electrostatic energy

When only monomer A is defined to be quantum, an interesting alternative approach exists to evaluating the one-particle density matrix on the quantum computer. We may first classically form the contributions from the nuclei of monomer A,

$$E_{\mathrm{elst,u}}^{\mathrm{VQE}} \leftarrow \sum_{\mathrm{A,B}}\frac{Z_{\mathrm{A}}Z_{\mathrm{B}}}{r_{\mathrm{AB}}} + \sum_{\mathrm{A}\nu\nu'}(A|\nu\nu')\gamma_{\nu\nu'}. \quad \text{(B1)}$$

We then classically form the image of the electrostatic potential of monomer B in the atomic orbital basis of monomer A as,

$$W_{\mu\mu'} \equiv \sum_{B}(\mu\mu'|B) + J_{\mu\mu'}[\gamma_{\nu\nu'}]. \quad \text{(B2)}$$

We can form the monomer A core ↔ monomer B contributions classically as,

$$E_{\mathrm{elst,u}}^{\mathrm{VQE}} \leftarrow \sum_{\mu\mu'}\gamma_{\mu\mu'}^{\mathrm{core}}W_{\mu\mu'}. \quad \text{(B3)}$$

We can then classically form the image of the electrostatic potential of monomer B in the active space molecular orbital basis of monomer A as,

$$W_{tt'} \equiv \sum_{\mu\mu'}C_{\mu t}W_{\mu\mu'}C_{\mu't'}. \quad \text{(B4)}$$

We can then diagonalize the $W_{tt'}$ operator to form,

$$W_{tt'} = \sum_{s}U_{ts}w_sU_{t's}. \quad \text{(B5)}$$

where U_{ts} is $\mathcal{SO}(N_{\mathrm{t}})$. In this "electrostatic potential natural orbital basis" the remaining monomer A active ↔ monomer B contributions can be evaluated by a single commuting group of simultaneous Z-basis measurements,

$$E_{\mathrm{elst,u}}^{\mathrm{VQE}} \leftarrow \sum_{\mu\mu'}\gamma_{\mu\mu'}^{\mathrm{a}}W_{\mu\mu'} = \sum_{tt'}\gamma_{tt'}^{\mathrm{a}}W_{tt'}. \quad \text{(B6)}$$

$$= \sum_{tt'}\langle\Omega_{\mathrm{A}}|t^{\dagger}t' + \overline{t}^{\dagger}\overline{t}'|\Omega_{\mathrm{A}}\rangle W_{tt'}. \quad \text{(B7)}$$

© 2022 The Author(s). Published by the Royal Society of Chemistry

$$= \sum_s \langle \Omega_A | \hat{V}^\dagger(U_{ts}) [s^\dagger s + \bar{s}^\dagger \bar{s}] \hat{V}(U_{ts}) | \Omega_A \rangle w_s. \quad \text{(B8)}$$

$$= \sum_s w_s - \frac{1}{2} \sum_s \langle \Omega_A | \hat{V}^\dagger(U_{ts}) \left[\hat{Z}_s + \hat{Z}_{\bar{s}} \right] \hat{V}(U_{ts}) \langle \Omega_A | w_s. \quad \text{(B9)}$$

This is in contrast to a naive implementation which would require $\mathcal{O}(N^2)$ circuit evaluations or an optimal method of $N/2$ circuits evaluations.[41] Similar ideas can be explored for the exchange expression. Given that its structure mirrors that of a total energy evaluation (a 4-index tensor contracted with the two-particle reduced density matrix) it seems possible that a double factorization[45,46] approach may be possible. These ideas will be explored at a later date.

Data availability

Molecular geometries, raw data and additional calculations are available in the supporting information.

Author contributions

R. M. P conceived and directed this project. R. M. P. and F. D. M derived and implemented the SAPT(VQE) method. F. D. M. ran and analyzed the SAPT(VQE) simulations. T. F. prepared the KDM5A structures, assisted with their visualization and helped analyze the data. N. M. optimized the KDM5A geometries and computed the DFT binding energies. A. R. W., M. D., E. K., R. S., N. M. and M. S. critiqued and helped improve SAPT(VQE). All authors discussed the results and wrote the manuscript.

Conflicts of interest

F. D. M., R. M. P., and A. R. W. own stock/options in QC Ware Corp.

Acknowledgements

QC Ware Corp. acknowledges generous funding from Boehringer Ingelheim for the undertaking of this project. We thank Clemens Utschig–Utschig and Christofer Tautermann for insightful discussions.

References

1 A. Aspuru-Guzik, A. D. Dutoi, P. J. Love and M. Head-Gordon, *Science*, 2005, **309**, 1704.
2 S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin and X. Yuan, *Rev. Mod. Phys.*, 2020, **92**, 015003.
3 Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis and A. Aspuru-Guzik, *Chem. Rev.*, 2019, **119**, 10856, pMID: 31469277.
4 B. Bauer, S. Bravyi, M. Motta and G. K.-L. Chan, *Chem. Rev.*, 2020, **120**, 12685.
5 V. von Burg, G. H. Low, T. Häner, D. S. Steiger, M. Reiher, M. Roetteler and M. Troyer, *Phys. Rev. Res.*, 2021, **3**, 033055.
6 M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. A. Jiménez-Hoyos, T. N. Lan, J. Li, F. Ma, A. J. Millis, N. V. Prokof'ev, U. Ray, G. E. Scuseria, S. Sorella, E. M. Stoudenmire, Q. Sun, I. S. Tupitsyn, S. R. White, D. Zgid and S. Zhang, Simons Collaboration on the Many-Electron Problem, *Phys. Rev. X*, 2017, **7**, 031059.
7 K. T. Williams, Y. Yao, J. Li, L. Chen, H. Shi, M. Motta, C. Niu, U. Ray, S. Guo, R. J. Anderson, J. Li, L. N. Tran, C.-N. Yeh, B. Mussard, S. Sharma, F. Bruneval, M. van Schilfgaarde, G. H. Booth, G. K.-L. Chan, S. Zhang, E. Gull, D. Zgid, A. Millis, C. J. Umrigar and L. K. Wagner, Simons Collaboration on the Many-Electron Problem, *Phys. Rev. X*, 2020, **10**, 011041.
8 J. J. Eriksen, T. A. Anderson, J. E. Deustua, K. Ghanem, D. Hait, M. R. Hoffmann, S. Lee, D. S. Levine, I. Magoulas, J. Shen, N. M. Tubman, K. B. Whaley, E. Xu, Y. Yao, N. Zhang, A. Alavi, G. K.-L. Chan, M. Head-Gordon, W. Liu, P. Piecuch, S. Sharma, S. L. Ten-no, C. J. Umrigar and J. Gauss, *J. Phys. Chem. Lett.*, 2020, **11**, 8922.
9 J. Preskill, *Quantum*, 2018, **2**, 79.
10 J. Cohn, M. Motta and R. M. Parrish, *PRX Quantum*, 2021, **2**(4), 040352.
11 V. E. Elfving, B. W. Broer, M. Webber, J. Gavartin, M. D. Halls, K. Patrick Lorton and A. D. Bochevarov, 2020, arXiv:2009.12472v1.
12 A. Heifetz, *Quantum mechanics in drug discovery*, Springer, 2020.
13 Y. Cao, J. Romero and A. Aspuru-Guzik, *IBM J. Res. Dev.*, 2018, **62**, 6.
14 C. Outeiral, M. Strahm, J. Shi, G. M. Morris, S. C. Benjamin and C. M. Deane, 2021, The prospects of quantum computing in computational molecular biology, arXiv:2005.12792.
15 S. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
16 A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik and J. L. O'Brien, *Nat. Commun.*, 2014, **51**(5), 1.
17 J. R. McClean, J. Romero, R. Babbush and A. Aspuru-Guzik, *New J. Phys.*, 2016, **18**, 023023.
18 W. J. Huggins, B. A. O'Gorman, N. C. Rubin, D. R. Reichman, R. Babbush and J. Lee, 2021, arXiv:2106.16235.
19 B. Jeziorski, M. Bulski and L. Piela, *Int. J. Quantum Chem.*, 1976, **10**, 281.
20 B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887.
21 R. M. Parrish, T. M. Parker and C. D. Sherrill, *J. Chem. Theory Comput.*, 2014, **10**, 4417.
22 R. M. Parrish, K. C. Thompson and T. J. Martínez, *J. Chem. Theory Comput.*, 2018, **14**, 1737.
23 G.-J. Yang, M.-H. Zhu, X.-J. Lu, Y.-J. Liu, J.-F. Lu, C.-H. Leung, D.-L. Ma and J. Chen, *J. Hematol. Oncol.*, 2021, **141**(14), 1.
24 H. L. Williams and C. F. Chabalowski, *J. Phys. Chem. A*, 2000, **105**, 646.
25 A. J. Misquitta and K. Szalewicz, *Chem. Phys. Lett.*, 2002, **357**, 301.

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Open Access Article. Published on 17 January 2022. Downloaded on 12/8/2025 11:58:09 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

26 M. Hapka, M. Przybytek, and K. Pernal, 2021, arXiv:2104.06891v1.
27 R. Moszynski, B. Jeziorski, S. Rybak, K. Szalewicz and H. L. Williams, *J. Chem. Phys.*, 1994, **100**, 5080.
28 T. Korona, *J. Chem. Phys.*, 2008, **128**, 224104.
29 T. Korona, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6509.
30 T. Korona, *J. Chem. Theory Comput.*, 2009, **5**, 2663.
31 J. Almlöf, K. Faegri and K. Korsell, *J. Comput. Chem.*, 1982, **3**, 385.
32 I. S. Ufimtsev and T. J. Martínez, *J. Chem. Theory Comput.*, 2008, **4**, 222.
33 I. S. Ufimtsev and T. J. Martinez, *J. Chem. Theory Comput.*, 2009, **5**, 1004.
34 I. S. Ufimtsev and T. J. Martinez, *J. Chem. Theory Comput.*, 2009, **5**, 2619.
35 E. G. Hohenstein and C. D. Sherrill, *J. Chem. Phys.*, 2010, **132**, 184111.
36 Y. Matsuzawa and Y. Kurashige, *J. Chem. Theory Comput.*, 2020, **16**, 944.
37 I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan and R. Babbush, *Phys. Rev. Lett.*, 2018, **120**, 110501.
38 J. Lee, W. J. Huggins, M. Head-Gordon and K. B. Whaley, *J. Chem. Theory Comput.*, 2018, **15**, 311.
39 G.-L. R. Anselmetti, D. Wierichs, C. Gogolin and R. M. Parrish, 2021, arXiv preprint arXiv:2104.05695.
40 B. O'Gorman, W. J. Huggins, E. G. Rieffel and K. B. Whaley, 2019, arXiv preprint arXiv:1905.05118.
41 Google AI and Collaborators, *Science*, 2020, **369**, 1084.
42 J. Tilly, P. V. Sriluckshmy, A. Patel, E. Fontana, I. Rungger, E. Grant, R. Anderson, J. Tennyson and G. H. Booth, *Phys. Rev. Res.*, 2021, **3**, 033230.
43 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, *et al.*, *Nat. Methods*, 2020, **17**, 261.
44 See https://github.com/robparrish/lightspeed for details on how to obtain the source code.
45 M. Motta, E. Ye, J. R. McClean, Z. Li, A. J. Minnich, R. Babbush and G. K.-L. Chan, *npj Quantum Information*, 2021, **71**(7), 1.
46 W. J. Huggins, J. R. McClean, N. C. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley and R. Babbush, *npj Quantum Information*, 2021, **71**(7), 1.
47 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33.
48 S. Vilar, G. Cozza and S. Moro, *Curr. Top. Med. Chem.*, 2008, **8**, 1555.
49 P. Jurečka, J. Šponer, J. Černý and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985.
50 S. S. Labadie, P. S. Dragovich, R. T. Cummings, G. Deshmukh, A. Gustafson, N. Han, J. C. Harmange, J. R. Kiefer, Y. Li, J. Liang, B. M. Liederer, Y. Liu, W. Manieri, W. Mao, L. Murray, D. F. Ortwine, P. Trojer, E. VanderPorten, M. Vinogradova and L. Wen, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 4492.
51 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235.
52 V. S. Gehling, S. F. Bellon, J.-C. Harmange, Y. LeBlanc, F. Poy, S. Odate, S. Buker, F. Lan, S. Arora, K. E. Williamson, P. Sandy, R. T. Cummings, C. M. Bailey, L. Bergeron, W. Mao, A. Gustafson, Y. Liu, E. VanderPorten, J. E. Audia, P. Trojer and B. K. Albrecht, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 4350.
53 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615.
54 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision D.01*, Gaussian Inc, Wallingford CT, 2013.
55 E. R. Sayfutyarova, Q. Sun, G. K.-L. Chan and G. Knizia, *J. Chem. Theory Comput.*, 2017, **13**, 4063.
56 A. A. Holmes, N. M. Tubman and C. J. Umrigar, *J. Chem. Theory Comput.*, 2016, **12**, 3674.
57 S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi and C. J. Umrigar, *J. Chem. Theory Comput.*, 2017, **13**, 1595.
58 C. Yung-Chi and W. H. Prusoff, *Biochem. Pharmacol.*, 1973, **22**, 3099.
59 T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush and J. R. McClean, *Phys. Rev. X*, 2020, **10**, 011004.

© 2022 The Author(s). Published by the Royal Society of Chemistry