A state-averaged orbital-optimized hybrid quantum–classical algorithm for a democratic description of ground and excited states — Source link ↗

Saad Yalouz, Saad Yalouz, Bruno Senjean, Bruno Senjean ...+5 more authors

Institutions: Leiden University, VU University Amsterdam, Google

Topics: Quantum computer, Conical intersection, Degenerate energy levels, Electronic structure and Quantum

Related papers:

- A state-averaged orbital-optimized hybrid quantum-classical algorithm for a democratic description of ground and excited states
- A variational eigenvalue solver on a photonic quantum processor
- Quantum Computing in the NISQ era and beyond
- Variational Quantum Computation of Excited States
- Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets

PAPER

A state-averaged orbital-optimized hybrid quantum–classical algorithm for a democratic description of ground and excited states

To cite this article: Saad Yalouz *et al* 2021 *Quantum Sci. Technol.* **6** 024004

View the article online for updates and enhancements.

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

Quantum Science and Technology

PAPER

A state-averaged orbital-optimized hybrid quantum–classical algorithm for a democratic description of ground and excited states

Saad Yalouz[1,2,*] , Bruno Senjean[1,2], Jakob Günther[1], Francesco Buda[3], Thomas E O'Brien[2,4] and Lucas Visscher[1]

1   Theoretical Chemistry, Vrije Universiteit, De Boelelaan 1083, NL-1081 HV, Amsterdam, The Netherlands
2   Instituut-Lorentz, Universiteit Leiden, PO Box 9506, 2300 RA Leiden, The Netherlands
3   Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, PO Box 9502, 2300 RA Leiden, The Netherlands
4   Google Research, 80636 Munich, Germany
*   Authors to whom any correspondence should be addressed.

E-mail: yalouzsaad@gmail.com

Keywords: quantum computing, quantum chemistry, conical intersection

Abstract

In the noisy intermediate-scale quantum (NISQ) era, solving the electronic structure problem from chemistry is considered as the 'killer application' for near-term quantum devices. In spite of the success of variational hybrid quantum/classical algorithms in providing accurate energy profiles for small molecules, careful considerations are still required for the description of complicated features of potential energy surfaces. Because the current quantum resources are very limited, it is common to focus on a restricted part of the Hilbert space (determined by the set of active orbitals). While physically motivated, this approximation can severely impact the description of these complicated features. A perfect example is that of conical intersections (i.e. a singular point of degeneracy between electronic states), which are of primary importance to understand many prominent reactions. Designing active spaces so that the improved accuracy from a quantum computer is not rendered useless is key to finding useful applications of these promising devices within the field of chemistry. To answer this issue, we introduce a NISQ-friendly method called 'state-averaged orbital-optimized variational quantum eigensolver' which combines two algorithms: (1) a state-averaged orbital-optimizer, and (2) a state-averaged VQE. To demonstrate the success of the method, we classically simulate it on a minimal Schiff base model (namely the formaldimine molecule CH_2NH) relevant also for the photoisomerization in rhodopsin—a crucial step in the process of vision mediated by the presence of a conical intersection. We show that merging both algorithms fulfil the necessary condition to describe the molecule's conical intersection, i.e. the ability to treat degenerate (or quasi-degenerate) states on the same footing.

1. Introduction

Quantum computing promises to solve problems that are beyond the capacities of classical devices. Technological advances in quantum computers occur rapidly [1–8], and it is now of primary importance to develop quantum algorithms dedicated to specific tasks of high industrial and societal impact. The electronic structure problem in chemistry is considered as one of the killer applications of quantum computers [9–14], and it remains an open question whether these applications will be achieved in the noisy intermediate-scale quantum (NISQ) era [15]. While quantum phase estimation algorithms [16–18] can in principle solve the electronic structure problem in its entirety, they have received limited focus in the NISQ era due to their apparent longer circuit depth requirements. Instead, variational hybrid quantum/classical algorithms require much shallower circuits at the cost of additional measurements, with as a spearhead the

variational quantum eigensolver (VQE) [7, 19, 20]. Those NISQ-friendly algorithms are hoped to solve quantum chemical problems of moderate-large size in the near future.

While computing the ground state of a molecule gives information about the equilibrium geometries and transition states (essential to determine the energy barrier of a chemical reaction), several chemical reactions also depend on excited states (see for instance reference [21] and references therein). To date, most variational quantum algorithms have been specifically designed to prepare ground states [1, 22–37] while excited state calculations have received less attention until only recently [23, 38–49]. It should be noted that some excited states can also be obtained as the lowest energy solutions of a given set of symmetries [50–55].

The landscape of ground- and excited-state potential energy surfaces (PES) can be quite complicated, making even qualitative numerical description challenging. As atomic nuclei move, ground- and excited-state energies may tend towards each other, leading to either avoided crossings or true inversions in the ordering of the electronic states. These inversions are referred to as conical intersections and are ubiquitous in photochemistry and photobiology [56, 57]. Their presence mediates important reactions, for instance the photoisomerisation of the chromophore of the visual photoreceptor rhodopsin, ensuing absorption of light. Photoisomerisation is a crucial step in the process of vision [58–62]. Conical intersections also find their importance in photosynthesis [63], photostabilization of DNA [64–67] and excitation-energy transfer [68] such as in supramolecular light-harvesting antennae [69]. Despite their indisputable importance, their description remains a challenge in the chemistry community [56, 57, 70]. Diagonalization of the electronic Hamiltonian matrix being intractable on classical devices, it is of common use to reduce the Hilbert space size by considering only a subspace of (so-called active) orbitals of the full problem. However, such a truncated Hilbert space is not guaranteed to be accurate enough to represent a conical intersection anymore. A possible solution is provided by state-averaged orbital-optimization, which consists in modifying the truncated Hilbert space (without modifying its size) to hopefully recover the conical intersection. On classical computers, this orbital-optimization is part of the so-called state-averaged complete active space self-consistent field (SA-CASSCF) method [71, 72], known for its democratic treatment of multiple (possibly degenerate) eigenstates.

For quantum computers in the NISQ era, the concept of active space also represents an important issue. Indeed, the limitations of quantum resources (e.g. maximum number of qubits) constrain algorithms to focus on small-sized active spaces, thus preventing any calculation on large systems. Designing Hilbert spaces so that the improved accuracy from a quantum computer is not rendered useless by the active space approximation itself is key to finding useful applications of these promising devices within the field of chemistry. In this context, getting access to complicated chemical features such as avoided crossing or conical intersections can appear as a real challenge for near-term quantum computers.

Motivated by this issue, we introduce in this paper a NISQ-friendly variational hybrid quantum/classical algorithm inspired by the SA-CASSCF method, so-called state-averaged orbital-optimized VQE (SA-OO-VQE). The ability of SA-OO-VQE to capture a conical intersection is demonstrated on a minimal Schiff base model of the rhodopsin chromophore, i.e. the formaldimine molecule.

The paper is organized as follows. After a brief description of the electronic structure problem in section 2.1, the concept of conical intersection is introduced in section 2.2. A particular attention is paid to the efficiency of classical algorithms to represent this characteristic in chemistry (section 2.3), as shown on the formaldimine molecule in section 2.4. Then, the SA-OO-VQE hybrid quantum/classical algorithm is detailed in section 3, where each building blocks of the algorithm is described separately. Following a brief introduction on quantum variational algorithms in section 3.1, we detail how SA-OO-VQE mimics its classical analog SA-CASSCF by combining a state-averaged VQE (section 3.2.1) with a state-averaged orbital-optimization procedure (section 3.2.2). The importance of choosing a flexible wavefunction ansatz such as the generalized unitary coupled cluster (with double excitations only) is discussed in section 3.2.3. Following the computational details in section 4, we demonstrate the efficiency of our method on the formaldimine molecule in section 5. Finally, conclusions and perspectives are given in section 6.

2. Theory

2.1. Electronic structure problem

In quantum chemistry, the Born–Oppenheimer approximation assumes that molecular structure and dynamics can be treated in two parts: considering first the atomic nuclei, and then the motion of the electrons around them. Atomic nuclei, being many times heavier than their electronic counterparts, are often treated as classical point particles with definite position $\mathbf{R}_A$ and momentum $\mathbf{P}_A$ (with A the index of the nucleus). As classical particles have no zero-point energy, the ground state of the nuclei is motionless, i.e. $\mathbf{P}_A = 0$. By contrast, the light electrons behave as fundamentally quantum particles, and form a

collective many-body wavefunction $\Psi(\mathbf{r})$, where $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e})$ is the $3N_e$-dimensional global position vector of the N_e electrons. This wavefunction evolves by the Schrödinger equation with the fixed nuclei providing a background potential field. The quantum operator describing this dynamics is the so-called electronic structure Hamiltonian (in atomic units)

$$\hat{\mathcal{H}} = -\frac{1}{2}\sum_{i=1}^{N_e} \nabla_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j}^{N_e} \frac{1}{\mathbf{r}_{ij}} - \sum_{i=1}^{N_e}\sum_{A=1}^{N_a} \frac{Z_A}{\mathbf{r}_{iA}}, \tag{1}$$

where $\mathbf{r}_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, $\mathbf{r}_{iA} = |\mathbf{r}_i - \mathbf{R}_A|$ and Z_A are respectively the distance between electrons i and j, the distance between electron i and nuclei A, and the atomic number of atom A. In this context, solving the electronic structure problem means solving the time-independent Schrödinger equation

$$\hat{\mathcal{H}}|\Psi_\ell\rangle = E_\ell|\Psi_\ell\rangle \tag{2}$$

which yields a set of electronic eigenstates $|\Psi_\ell\rangle \equiv \Psi_\ell(\mathbf{r})$ with corresponding energies E_ℓ. These eigenstates and energies depend on the atomic positions themselves—$E_\ell = E_\ell(\mathbf{R})$ and $|\Psi_\ell\rangle = |\Psi_\ell(\mathbf{R})\rangle$, where $\mathbf{R}$ is the $3N_a$-dimensional position vector of the N_a nuclei. (Note the change of notation: $\Psi_\ell(\mathbf{r})$ is a single number from evaluating the function Ψ_ℓ at the fixed electronic co-ordinates $\mathbf{r}$, while $|\Psi_\ell(\mathbf{R})\rangle$ is an entire function defined by a fixed set of nuclear co-ordinates $\mathbf{R}$). In practice, it is convenient to work in a finite basis of N_o orthonormal molecular orbitals (MO) $\{\phi_p(\mathbf{r})\}$. The latter are usually solutions to the mean-field single-particle problem, determined by the Hartree–Fock (HF) method [72]. In this basis, the spin-free Hamiltonian reads

$$\hat{\mathcal{H}}(\mathbf{R}) = \sum_{pq}^{N_o} h_{pq}(\mathbf{R})\hat{E}_{pq} + \frac{1}{2}\sum_{pqrs}^{N_o} g_{pqrs}(\mathbf{R})\hat{e}_{pqrs}, \tag{3}$$

where (we drop the dependence on $\mathbf{R}$ for convenience),

$$h_{pq} = \int \phi_p^*(\mathbf{r}_1)\left(-\frac{1}{2}\nabla_{\mathbf{r}_1} - \sum_{A=1}^{N_a}\frac{Z_A}{\mathbf{r}_{1A}}\right)\phi_q(\mathbf{r}_1)\mathrm{d}\mathbf{r}_1 \tag{4}$$

and

$$g_{pqrs} = \iint \phi_p^*(\mathbf{r}_1)\phi_r^*(\mathbf{r}_2)\frac{1}{\mathbf{r}_{12}}\phi_q(\mathbf{r}_1)\phi_s(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{5}$$

are the one- and two-electron integrals, and $\hat{E}_{pq} = \sum_\sigma \hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma}$ and $\hat{e}_{pqrs} = \sum_{\sigma,\tau} \hat{a}_{p\sigma}^\dagger \hat{a}_{r\tau}^\dagger \hat{a}_{s\tau} \hat{a}_{q\sigma}$ are the one- and two-body spin-free operators. Inverting the problem, the set $\{E_\ell(\mathbf{R})\}$ (now basis-dependent) may be considered as a set of potential energy manifolds for the classical motion of the nuclei.

2.2. Conical intersections in nature

A naive study of quantum mechanics would suggest that the two lowest PESs of a molecule $E_0(\mathbf{R})$ and $E_1(\mathbf{R})$ would never cross. Perturbation theory suggests that such a crossing point or degeneracy would be prone to any disturbance: if $E_0(\mathbf{R}) = E_1(\mathbf{R})$, we would expect a perturbation V to induce a gap $E_1(\mathbf{R}) - E_0(\mathbf{R}) \sim |V|$. However, some degeneracies are more robust than others. Consider the effective block Hamiltonian obtained by projecting the full Hamiltonian $\hat{\mathcal{H}}(\mathbf{R})$ in equation (3) onto the two lowest states decoupled from other states,

$$\hat{H} = \begin{bmatrix} H_{00} & H_{01} \\ H_{10} & H_{11} \end{bmatrix}, \tag{6}$$

where $H_{IJ} = \langle \Psi_I | \hat{H} | \Psi_J \rangle$. An extensive discussion about the necessary conditions for this effective Hamiltonian to depict a conical intersection is provided in reference [70]. For simplicity, suppose that, in the absence of a perturbation, this low-energy effective Hamiltonian around a degeneracy at $\mathbf{R}_0$ takes the form

$$\hat{H} = h_0(\mathbf{R})I + h_X(\mathbf{R} - \mathbf{R}_0)\cdot\mathbf{R}_X X + h_Z(\mathbf{R} - \mathbf{R}_0)\cdot\mathbf{R}_Z Z, \tag{7}$$

where I is the 2×2 identity matrix, X and Z are the two real Pauli matrices, and $\mathbf{R}_X$ and $\mathbf{R}_Z$ are two arbitrary $3N_a$-dimensional vectors. This amounts to $H_{00} = h_0(\mathbf{R}) + h_Z(\mathbf{R} - \mathbf{R}_0)\cdot\mathbf{R}_Z$, $H_{01} = H_{10} = h_X(\mathbf{R} - \mathbf{R}_0)\cdot\mathbf{R}_X$ and $H_{11} = h_0(\mathbf{R}) - h_Z(\mathbf{R} - \mathbf{R}_0)\cdot\mathbf{R}_Z$. Such a degeneracy is known as a conical intersection, as the energy surfaces form a cone in the $\mathbf{R}_X$ and $\mathbf{R}_Z$ directions (modulo the background shift $h_0(\mathbf{R})$):

$$E_\pm(\mathbf{R}) = h_0(\mathbf{R}) \pm \sqrt{h_X|(\mathbf{R} - \mathbf{R}_0)\cdot\mathbf{R}_X|^2 + h_Z|(\mathbf{R} - \mathbf{R}_0)\cdot\mathbf{R}_Z|^2}. \tag{8}$$

This conical intersection necessarily connects the two surfaces $E_+(\mathbf{R})$ and $E_-(\mathbf{R})$; a continuous path along one surface passing through $\mathbf{R}_0$ will invert the bands. Note that a (real-valued) perturbation will shift the

degeneracy to some $\mathbf{R}_0'$, rather than breaking it (although a gap at $\mathbf{R} = \mathbf{R}_0$ is still induced) [70]. A proof is provided in appendix A.

In photochemistry, this degeneracy allows for non-radiative relaxation from an electronic excited state. Consider a photoinduced chemical reaction, triggered by the absorption of a photon that promotes the electrons of the molecule from the ground state to an excited state. As there is *a priori* no reason to expect that local minima of the ground state are also local minima of the excited state, the wavepacket will evolve in time, driven by atomic forces. Through its evolution, it may decay to a lower state thanks to different phenomena [21], finally reaching a new ground-state minima which may be in a drastically different nuclear geometry than the initial one. One such phenomena is called internal conversion, i.e. a non-radiative transition induced by the coupling between electronic and nuclear (vibrational) degrees of freedom (so-called vibronic coupling). Internal conversion is mediated by a degenerate point with double cone topography, so-called conical intersection [73–76], as described in the example above. Such an internal nuclear rearrangement is critical e.g. for sight, where the (*cis* to *trans*) isomerization of the retinal protonated Schiff base (RPSB) chromophore—in the visual receptor rhodopsin in the retina—is the key photochemical event following the absorption of a photon [77]. Conical intersections are also of fundamental interest, as the (fundamentally quantum) behaviour of the nuclear motion along the degrees of freedom ($\mathbf{R}_X$ and $\mathbf{R}_Z$ in the above example) is no longer well-described by the Born–Oppenheimer approximation. In other words, vibronic couplings diverge at the conical intersection. These quantum dynamics are critically important to consider when modelling e.g. photoinduced electron transfer processes relevant in solar energy conversion [78]. Also for the reverse process [79], emission rather than harvesting of light, (intersystem) crossings [80] of PES play a crucial role and require balanced treatments of all involved surfaces and their conical intersections.

2.3. Predicting conical intersections numerically

The accurate description of the PESs of a molecular system is essential to understand complex phenomena in excited-state dynamics [81, 82]. While vertical transitions and excited-state PES around equilibrium geometries are usually well described by standard chemistry methods such as linear response time-dependent density-functional theory (TDDFT) [83–86], this is much more challenging when states are degenerate or closely degenerate. Indeed, conical intersections between the ground and first-excited state are ill-defined in linear response TDDFT [70], in addition to other deficiencies regarding doubly excited [87] and charge transfer states [88]. In order to correctly describe a conical intersection, the crossing states have to be treated on the same footing, i.e. they have to result from a same high level theory. In this context, the full configuration interaction (FCI) approach, which consists in diagonalizing the entire electronic structure Hamiltonian in the full MO basis, would be the ideal method to describe the electronic spectrum. Unfortunately, FCI scales exponentially with the system size. Instead, it is common to employ the frozen core approximation, i.e. to assume that some core orbitals are always occupied by electrons, and some virtual orbitals are never occupied. These orbitals may be 'frozen', leaving us to solve the problem using only the remaining 'active' spin-orbitals. This method is typically known as complete active space configuration interaction (CASCI) approximation. In contrast to FCI, the active space selection is user-dependent and is not invariant with respect to orbital rotations. As a consequence, it is usual to variationally optimize the orbitals with respect to the ground-state CASCI energy, thus leading to the so-called state-specific complete active space self-consistent field method (CASSCF). These three methods are ordered both in terms of their complexity cost T and their accuracy in finding the ground-state energy E_0—we have $T_{\text{FCI}} \geqslant T_{\text{CASSCF}} \geqslant T_{\text{CASCI}}$, and $E_0 \leqslant E_0^{\text{FCI}} \leqslant E_0^{\text{CASSCF}} \leqslant E_0^{\text{CASCI}}$. Finding an appropriate level of theory for a given problem requires finding the best tradeoff in this cost-accuracy balance.

Regardless of the level of theory chosen, finding the exact solution of the electronic structure problem is a difficult task. Solving a single active space of m spatial orbitals containing n electrons, denoted by CASCI(m, n), requires finding the ground state eigenvalue of a $\binom{2m}{n} \times \binom{2m}{n}$ matrix. (The equivalent CASSCF (m, n) problem takes a similar amount of time, multiplied by the number of iterations needed to find the optimal single-particle basis rotation.) On a classical computer, this takes $O(\binom{2m}{n}^3) \sim O(e^n)$ time. This becomes impractical beyond $n \sim 20$ [89], though many approximations have been developed over the last century (see references [90–93] and references therein). Quantum computers promise a route beyond this boundary, as they can find eigenvalues of an (m, n) active space to error ϵ in time polynomial in n and ϵ [94, 95]. This has generated a huge flurry of experimental [1–8] and theoretical [9, 13, 14, 96] activity. As a promise, fault-tolerant quantum computer could open up areas of chemistry to accurate computational study that were previously inaccessible. However, a quantum computer only improves the solution of a chemistry problem within the active space; it targets E_0^{CASCI} (or E_0^{CASSCF} when orbital-optimization is considered), and not the true ground state energy E_0. Designing relevant active spaces is key to finding

useful applications of quantum devices within the field of chemistry, and is an active field of research [26, 29, 47, 97–105].
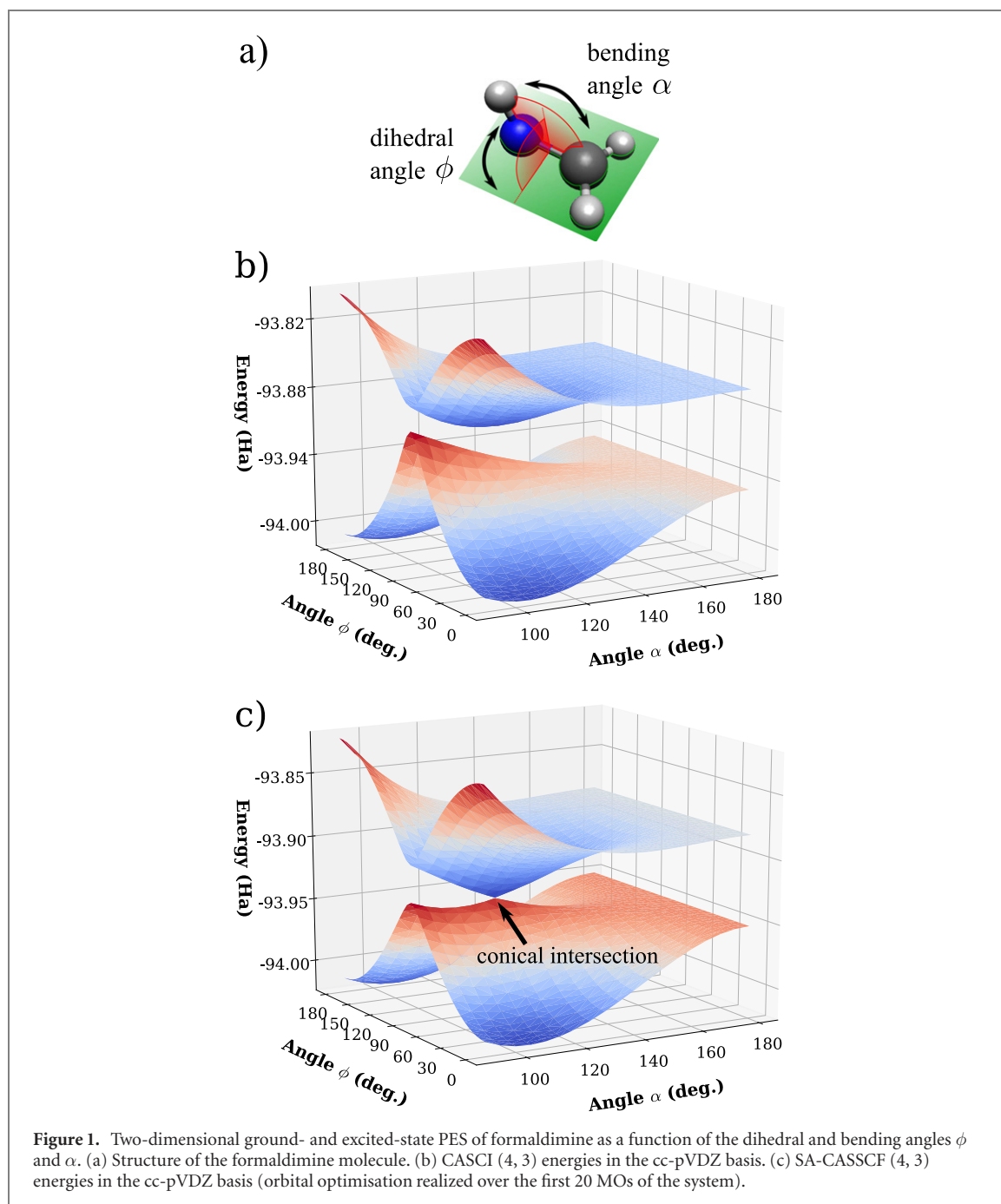
Conical intersections require careful consideration in active space approximations. Although the conical intersection might be qualitatively described by FCI, this is not guaranteed anymore when a truncated active space is considered (see appendix A). Switching to state-specific CASSCF, using the orbitals optimized with respect to the ground-state energy will by construction lower the ground state, but might have the opposite effect on the excited state. Thus, the gap between the two states will increase and the description of the conical intersection can become worse. One could think of optimizing the orbitals for the two states separately. In this case, two separate sets of state-specific orbitals are obtained. However, the two states expressed in those two different orbital sets are not guaranteed to be orthogonal anymore. Furthermore, when the states are closely degenerate (as around conical intersections or avoided crossings), optimizing the orbitals with respect to the excited state energy becomes impossible in practice. Indeed, the (optimized) excited state and the ground state will inevitably (and indefinitely) swap during the optimization process, thus preventing convergence of the orbitals. This practical problem is known as root flipping. Finally, active spaces can also be designed differently, for instance by downfolding electronic Hamiltonians into low-dimensional active spaces. However, those downfolded Hamiltonians are state-specific [97], and thus not appropriate for an equal footing treatment of ground and excited states.

In quantum chemistry, the standard method to describe conical intersections is given by the so-called state-averaged (SA)-CASSCF method. In this approach, the orbitals are variationally optimized with respect to the energy average of the states involved in the conical intersection [72]. In this sense, SA-CASSCF is a *democratic* method (i.e. it treats on an equal footing the optimization of all the states simultaneously) ensuring a coherent representation of an eventual degeneracy within a given active space, which strongly contrasts with the state-specific CASSCF approach. While the SA-CASSCF ground state might end up higher in energy than with state-specific CASSCF, results obtained from the former are most of the time more relevant as we are interested in energy difference in quantum chemistry. Note that to go beyond the SA-CASSCF calculation, one can use second order perturbation theory (CASPT2) [106], or similarly the N-electron valence second order perturbation theory (NEVPT2) [107] to recover the missing dynamical correlation (see also reference [70] and references therein).
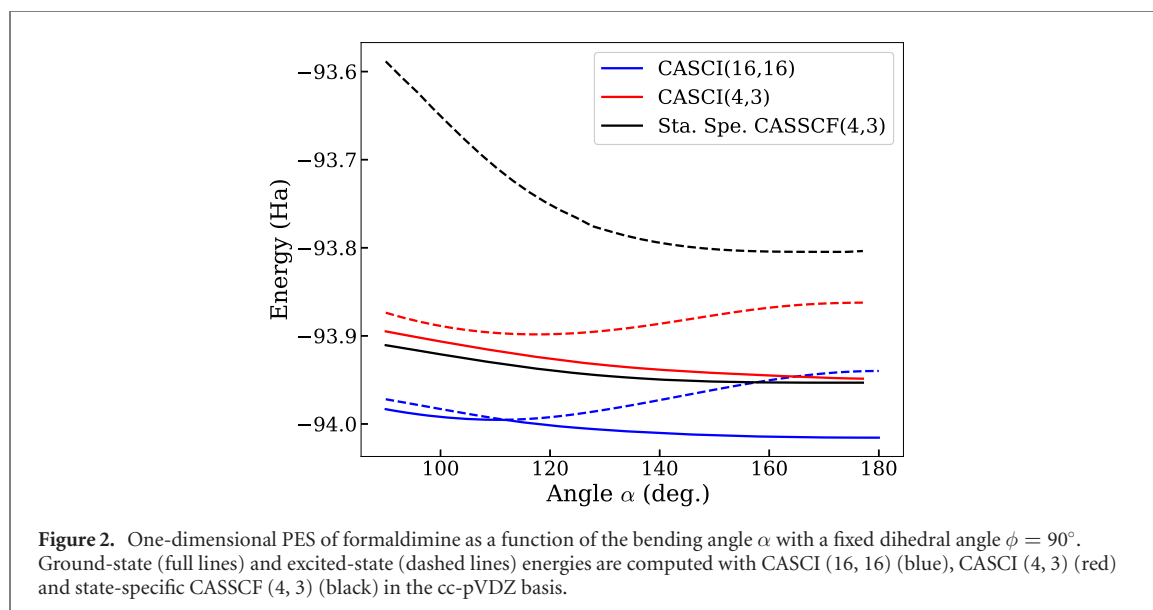
2.4. Example: photoisomerization of formaldimine

As an illustrative example of the previous discussion, we consider the formaldimine molecule (CH_2NH) which exhibits a conical intersection between its ground and first-excited PESs [108]. The photoisomerization of formaldimine can be considered as an illustrative minimal model of the more complex RPSB molecule whose photoisomerisation (*cis* to *trans*) plays a key role in the visual cycle process [77, 109]. The description of the conical intersection can be reduced to a two-dimensional problem by varying the bending angle $\alpha = H\overset{\frown}{-N-}C$ and the dihedral angle $\phi \equiv H - \overset{\frown}{C-N} - H$, while all other coordinates are kept frozen (see (a) in figure 1). It involves two states. The first one is dominated by a closed-shell configuration composed of the π-bonding orbital between the nitrogen and carbon $2p_x$ orbitals, and a lone pair in the nitrogen $2p_y$ orbital denoted by n. This corresponds to the HF determinant. The second state is dominated by an open-shell configuration called a $n\pi^*$ state, composed of one electron in the nitrogen lone pair orbital n, two electrons in the π-bonding orbital and one electron in the corresponding π^*-antibonding orbital [108]. This corresponds to the HOMO–LUMO excited configuration (HOMO and LUMO stand for highest occupied molecular orbital and lowest unoccupied molecular orbital, respectively). An active space comprising those orbitals should be sufficient to qualitatively capture the conical intersection, i.e. four electrons in three spatial-orbitals (4, 3). Note that one could even get rid of the π-bonding orbital, thus reducing the active space to two electrons in two spatial-orbitals (2, 2), and still capture a conical intersection. This could be used for an experiment on a real quantum device, as only four qubits are required. However, we decided to keep the π-bonding orbital in the active space due to its non-negligible quantitative contribution, thus leading to a (noiseless) classical simulation of a six qubits device.

As readily seen in figure 1(b), realizing a CASCI (4, 3) calculation in the canonical (HF) MO basis is not sufficient to capture the conical intersection. The latter is however well described within SA-CASSCF (4, 3) (figure 1(c)), which demonstrates the importance of working in the state-averaged orbital-optimized basis. As a proof of the existence of the conical intersection (see also references [108]), we considered a very large active space comprising all the electrons of the system (i.e. $n = 16$) in 16 orbitals, in the canonical MO basis. Results are shown in figure 2. For clarity, in this figure we only show the intersection plan of the 2D PES for $\phi = 90°$ which contains the conical intersection. By construction, the ground- and excited-state

Figure 1. Two-dimensional ground- and excited-state PES of formaldimine as a function of the dihedral and bending angles ϕ and α. (a) Structure of the formaldimine molecule. (b) CASCI (4, 3) energies in the cc-pVDZ basis. (c) SA-CASSCF (4, 3) energies in the cc-pVDZ basis (orbital optimisation realized over the first 20 MOs of the system).

CASCI (16, 16) energies are lower than the CASCI (4, 3) ones, which we recall does not capture the conical intersection, as also shown in figure 1. Interestingly, in contrast to CASCI (4, 3), the CASCI (16, 16) results feature the conical intersection around $\alpha = 112°$. This is already a numerical proof that the frozen core approximation does not guarantee the protection of the conical intersection (see appendix A for a mathematical proof). To recover the conical intersection, one has to either increase the size of the active space, like in CASCI (16, 16), or to switch to a more appropriate basis. By optimizing the orbitals with respect to the ground-state energy, we end up with the state-specific CASSCF (4, 3) result in figure 2. As expected, the ground-state energy is lower than the CASCI (4, 3) energy by construction. However, it remains higher than the CASCI (16, 16) one. What should be highlighted here is that while the ground-state energy is indeed getting better through the use of state-specific orbitals, this is at the expense of the description of the excited state. As a consequence, the energy difference between the ground and first excited state deteriorates by using state-specific CASSCF, thus impacting directly the description of the conical intersection.

Figure 2. One-dimensional PES of formaldimine as a function of the bending angle α with a fixed dihedral angle $\phi = 90°$. Ground-state (full lines) and excited-state (dashed lines) energies are computed with CASCI (16, 16) (blue), CASCI (4, 3) (red) and state-specific CASSCF (4, 3) (black) in the cc-pVDZ basis.

3. Methods

### 3.1. Variational quantum algorithms

The electronic structure problem may be solved on a quantum computer by mapping the fermionic modes onto qubits. This may be achieved by a number of non-local transformations, most of which mapping N spin-orbitals (not spatial orbitals) of an active space onto N qubits (see references [11, 12] for reference). The transformed qubit Hamiltonian $\hat{\mathcal{H}}^{q}$ has the same eigenvalue structure as the effective frozen core Hamiltonian $\hat{\mathcal{H}}^{FC}$, constructed with respect to a given active space (see appendix B for an introduction of the frozen core approximation). By virtue of the variational principle, all states within the N-qubit Hilbert space have energy $E \geqslant E_0^{FC}$, where E_0^{FC} is the ground-state energy of $\hat{\mathcal{H}}^{FC}$. A VQE [19, 20] exploits this fact by generating a set of states $|\Psi(\vec{\theta})\rangle$ dependent on some conditional control parameters $\vec{\theta}$. These parameters are used to tune quantum gates within a quantum circuit encoding an 'ansatz'. For a fixed choice of parameters $\vec{\theta}$, the energy of the generated state $|\Psi(\vec{\theta})\rangle$ under the Hamiltonian $\hat{\mathcal{H}}^{q}$ may be estimated by partial state tomography [33, 110], which may be achieved by as simple methods as single qubit rotation and readout. Traditionally, the energy $E(\vec{\theta}) = \langle \Psi(\vec{\theta})|\hat{\mathcal{H}}^{q}|\Psi(\vec{\theta})\rangle$ is used as a cost function in a (classical) optimization routine. When $E(\vec{\theta})$ is minimized, one hopes that the resulting state $|\Psi(\vec{\theta})\rangle$ is a good approximation for the ground state of $\hat{\mathcal{H}}^{q}$, and correspondingly that $E(\vec{\theta}) \sim E_0^{FC}$. Various methods exist to extend this scheme to find higher-energy eigenstates [38–49]. However, the accuracy of the eigenenergies of $\hat{\mathcal{H}}^{q}$ to the original problem is strictly limited by the active space, as we target $\hat{\mathcal{H}}^{FC}$ and not the original Hamiltonian $\hat{\mathcal{H}}$. Note that in the following, we will intentionally drop the notation $\hat{\mathcal{H}}^{q}$ when speaking about VQE to only use $\hat{\mathcal{H}}^{FC}$ for simplification.

3.2. State-averaged orbital-optimized VQE

With the current resources restriction affecting NISQ devices, VQE can only be employed to solve the electronic structure problem within very small-sized active spaces. In this context, trying to reproduce complex chemical features such as formaldimine's conical intersection becomes a real challenge. Indeed, as explained previously, VQE is energetically bounded by its classical analog CASCI which is known to drastically fail when applied to a small active space (see CASCI (4, 3) in figure 2). If one desires to solve such a problem on near term quantum devices, it is of primary importance to develop new approaches which give coherent results while working within small-sized active spaces.

To this end, we introduce a method specially dedicated to capture (near-)degeneracies on a NISQ device: the so-called SA-OO-VQE. Following the fundamental ideas that make the strength of the classical SA-CASSCF method, SA-OO-VQE provides a *democratic* treatment of the eigenstates of a system inside and outside the active-space. To this end, the SA-OO-VQE method is composed of two interdependent sub-algorithms working in a 'state-averaged' manner. The latter are respectively called state-averaged VQE (SA-VQE) and state-averaged orbital-optimizer (SA-OO).

In the following, we introduce how these two sub-algorithms work together in practice (as a visual support, figure 3 represents the general structure of the SA-OO-VQE method and provides a summary of the information exchanged between the two sub-algorithms). Motivated by the formaldimine problem, a
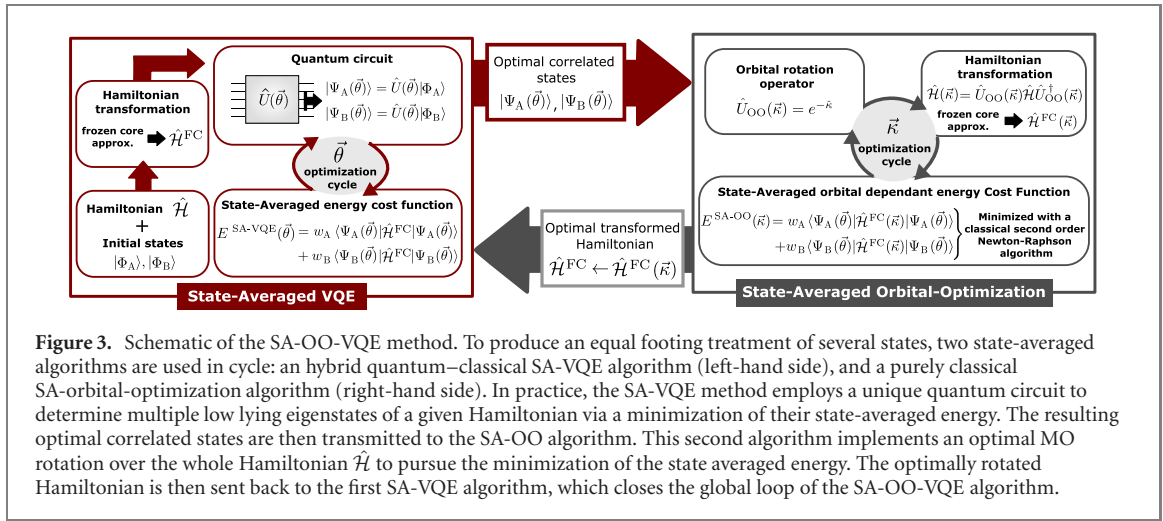
Figure 3. Schematic of the SA-OO-VQE method. To produce an equal footing treatment of several states, two state-averaged algorithms are used in cycle: an hybrid quantum–classical SA-VQE algorithm (left-hand side), and a purely classical SA-orbital-optimization algorithm (right-hand side). In practice, the SA-VQE method employs a unique quantum circuit to determine multiple low lying eigenstates of a given Hamiltonian via a minimization of their state-averaged energy. The resulting optimal correlated states are then transmitted to the SA-OO algorithm. This second algorithm implements an optimal MO rotation over the whole Hamiltonian $\hat{\mathcal{H}}$ to pursue the minimization of the state averaged energy. The optimally rotated Hamiltonian is then sent back to the first SA-VQE algorithm, which closes the global loop of the SA-OO-VQE algorithm.

two-states example will be considered in the following to illustrate the different steps of the method. SA-OO-VQE can however be straightforwardly generalized to any number of states, as well as to the particular case of a single state (thus leading to 'state-specific'-OO-VQE [26, 29]).

3.2.1. 1st algorithm: state-averaged VQE

The first algorithm acting in the SA-OO-VQE method is the SA-VQE approach, also called subspace-search VQE (SS-VQE) in reference [40], which consists in finding two (or more) low-lying eigenstates of a given Hamiltonian. To proceed, one starts by building the frozen core Hamiltonian $\hat{H}^{\mathrm{FC}}$. Then, we define a set of two (or more) orthonormal initial states $|\Phi_{\mathrm{A}}\rangle$ and $|\Phi_{\mathrm{B}}\rangle$ (with $\langle\Phi_{\mathrm{A}}|\Phi_{\mathrm{B}}\rangle = 0$) to be transformed into two correlated states via the implementation of a same unitary operator $\hat{U}(\vec{\theta})$,

$$|\Psi_{\mathrm{A}}(\vec{\theta})\rangle = \hat{U}(\vec{\theta})|\Phi_{\mathrm{A}}\rangle \quad \text{and} \quad |\Psi_{\mathrm{B}}(\vec{\theta})\rangle = \hat{U}(\vec{\theta})|\Phi_{\mathrm{B}}\rangle. \quad (9)$$

The use of a same unitary $\hat{U}(\vec{\theta})$ where the parameters $\vec{\theta}$ are optimized for both states simultaneously has multiple advantages. First, a single circuit is used to determine eigenstates which is a great asset for the NISQ era. Second, all target states are treated on an equal footing, thus avoiding biases that would naturally appear when using different levels of theory for each state. In appendix F, we discuss on the ability of several methods to treat different states on an equal footing. Finally, the orthonormality property of the initial states is transferred to the final correlated states, so that $\langle\Psi_{\mathrm{A}}(\vec{\theta})|\Psi_{\mathrm{B}}(\vec{\theta})\rangle = \langle\Phi_{\mathrm{A}}|\Phi_{\mathrm{B}}\rangle = 0$. The two correlated states are then used as ansatze to approximate the two low-lying eigenstates of $\hat{\mathcal{H}}^{\mathrm{FC}}$. To this end, the parameters $\vec{\theta}$ of the VQE circuit are optimized to minimize the state-averaged energy

$$E^{\mathrm{SA-VQE}}(\vec{\theta}) = w_{\mathrm{A}}\langle\Psi_{\mathrm{A}}(\vec{\theta})|\hat{\mathcal{H}}^{\mathrm{FC}}|\Psi_{\mathrm{A}}(\vec{\theta})\rangle + w_{\mathrm{B}}\langle\Psi_{\mathrm{B}}(\vec{\theta})|\hat{\mathcal{H}}^{\mathrm{FC}}|\Psi_{\mathrm{B}}(\vec{\theta})\rangle, \quad (10)$$

where the weights w_{A} and w_{B} (with $w_{\mathrm{A}} + w_{\mathrm{B}} = 1$ and $w_{\mathrm{A}} \geqslant w_{\mathrm{B}}$) define the weighted contribution of each correlated state to the cost function. This ensemble energy is governed by the Rayleigh–Ritz variational principle [111] ensuring that the lower energy bound reachable is given by the ensemble energy of the exact two low lying eigenstates of $\hat{H}^{\mathrm{FC}}$ (denoted by $|\Psi_0\rangle$ and $|\Psi_1\rangle$),

$$E^{\mathrm{SA}} = w_{\mathrm{A}}\langle\Psi_0|\hat{\mathcal{H}}^{\mathrm{FC}}|\Psi_0\rangle + w_{\mathrm{B}}\langle\Psi_1|\hat{\mathcal{H}}^{\mathrm{FC}}|\Psi_1\rangle$$

$$\leqslant E^{\mathrm{SA-VQE}}(\vec{\theta}). \quad (11)$$

In the particular case of an equi-ensemble (i.e. $w_{\mathrm{A}} = w_{\mathrm{B}}$), it is easy to show that the state-averaged energy is invariant with respect to any rotation of the states $|\Psi_0\rangle$ and $|\Psi_1\rangle$. Indeed, by replacing $|\Psi_0\rangle$ and $|\Psi_1\rangle$ by $|\Psi_-(\varphi)\rangle = \cos\varphi|\Psi_0\rangle - \sin\varphi|\Psi_1\rangle$ and $|\Psi_+(\varphi)\rangle = \sin\varphi|\Psi_0\rangle + \cos\varphi|\Psi_1\rangle$, one gets an ensemble energy $E^{\mathrm{SA}}(\varphi) \geqslant E^{\mathrm{SA}}(0)$ where the equality is always fulfilled (for any rotation angle φ) if and only if $w_{\mathrm{A}} = w_{\mathrm{B}}$ [111]. Hence, taking an equi-ensemble only ensures that the subspace spanned by $|\Psi_{\mathrm{A}}(\vec{\theta})\rangle$ and $|\Psi_{\mathrm{B}}(\vec{\theta})\rangle$ best approximates the subspace spanned by $|\Psi_0\rangle$ and $|\Psi_1\rangle$, instead of approximating each individual states directly. To extract the expected individual states and energies from the equi-ensemble, additional treatments are required such as additional maximizations [40] or classical diagonalization [45]. Alternatively, one can just take $w_{\mathrm{A}} > w_{\mathrm{B}}$ to impose a constraint in the ordering of the states [40], which may complicate the optimization considerably [49]. We refer to this alternative as the fully-weighted SS-VQE, described in section 2.3 of reference [40].

In this work, we considered the equi-ensemble formalism. While there is no guarantee to capture the correct individual states after a single minimization of the state-average energy [equation (10)], this remains possible by choosing appropriate initial states and wavefunction ansatz. Indeed, in section 5 we show that the correct individual energies are obtained with errors even below chemical accuracy. Furthermore, as no ordering is specified, the PES of the final states ($|\Psi_A\rangle$ and $|\Psi_B\rangle$) can cross: $|\Psi_A\rangle$ and $|\Psi_B\rangle$ can respectively be the ground and excited state for a given set of nuclear coordinates **R**, and reverse order for another set **R′**. This interesting feature is not without significance and will be discussed further in section 5.

Nevertheless, in general one has to ensure the correct capture of the eigenstates. For this, one possibility is to complement the cost function in equation (10) by the state-averaged variance [49]

$$\Delta^{SA}(\vec{\theta}) = w_A \Delta_A(\vec{\theta}) + w_B \Delta_B(\vec{\theta}), \tag{12}$$

where

$$\Delta_k(\vec{\theta}) = \langle \Psi_k(\vec{\theta})|(\hat{\mathcal{H}}^{FC})^2|\Psi_k(\vec{\theta})\rangle - \langle \Psi_k(\vec{\theta})|\hat{\mathcal{H}}^{FC}|\Psi_k(\vec{\theta})\rangle^2 \tag{13}$$

is the energy variance of Hamiltonian $\hat{\mathcal{H}}^{FC}$ with wavefunction $|\Psi_k(\vec{\theta})\rangle$. In practice, the measurement of the energy variance on a quantum computer can be realized via the averaging of a quantum covariance matrix depending on the output states and the elements of the Hamiltonian $\hat{H}$ (expressed in qubit form). For more information about this method, we refer the interested reader to reference [49] (and references within).

3.2.2. 2nd algorithm: state-averaged orbital-optimization

Once the SA-VQE sub-algorithm reaches convergence, the next step consists in using a classical SA-OO procedure. Directly inspired by the SA-CASSCF method, the goal of the SA-OO algorithm is to pursue the minimization of the state-averaged energy of $|\Psi_A(\vec{\theta})\rangle$ and $|\Psi_B(\vec{\theta})\rangle$ via the implementation of an optimal rotation of the system's MOs. To do so, throughout the whole SA-OO process, one considers that the shape of the correlated states never changes (i.e. the $\vec{\theta}$ parameters are frozen). Then, one introduces an unitary operator $U_{OO}(\vec{\kappa})$ such as

$$\hat{U}_{OO}(\vec{\kappa}) = e^{-\hat{\kappa}} \quad \text{with} \quad \hat{\kappa} = \sum_{p>q}^{MOs} \kappa_{pq}(\hat{E}_{pq} - \hat{E}_{qp}), \tag{14}$$

where p, q denote any arbitrary spatial orbitals. This operator is used to transform the system's MOs. Such a transformation acts simultaneously inside and outside the active space [72]. In practice, a restricted set of MOs is considered, including all doubly occupied MOs, all active space MOs and a few low virtuals. Using the $\hat{U}_{OO}(\vec{\kappa})$ operator, the MO basis transformation is applied to the full second quantized Hamiltonian of the system,

$$\hat{\mathcal{H}}(\vec{\kappa}) = \hat{U}_{OO}^\dagger(\vec{\kappa}) \, \hat{\mathcal{H}} \, \hat{U}_{OO}(\vec{\kappa}). \tag{15}$$

This transformation parametrizes the system's Hamiltonian which becomes explicitly dependent on the MO basis, through rotation $\vec{\kappa}$ (see appendix D for details about this transformation). The resulting transformed Hamiltonian $\hat{\mathcal{H}}(\vec{\kappa})$ still shares the same spectrum as the original operator $\hat{\mathcal{H}}$ since we use a unitary transformation. However, this is not the case for the associated frozen core Hamiltonian $\hat{\mathcal{H}}^{FC}(\vec{\kappa})$ which has a different spectrum compared to the original frozen core Hamiltonian $\hat{\mathcal{H}}^{FC}$ (in the non-optimized MO basis). The energy of the correlated states computed in the active space will therefore change with the parameters $\vec{\kappa}$. Using this new definition of the frozen core Hamiltonian, the state-averaged energy of the two correlated states now reads (for fixed parameters $\vec{\theta}$)

$$E^{SA-OO}(\vec{\kappa}) = w_A \langle \Psi_A(\vec{\theta})|\hat{\mathcal{H}}^{FC}(\vec{\kappa})|\Psi_A(\vec{\theta})\rangle + w_B \langle \Psi_B(\vec{\theta})|\hat{\mathcal{H}}^{FC}(\vec{\kappa})|\Psi_B(\vec{\theta})\rangle. \tag{16}$$

By minimizing equation (16) with respect to the parameters $\vec{\kappa}$, one obtains a variationally optimized MO basis designed to treat both correlated states on an equal footing.

Throughout this work, this optimization is carried out by using a state-averaged version of the classical Newton–Raphson method [112] (see appendix C for technical details). In practice, this algorithm requires the knowledge of the two converged correlated states $|\Psi_A(\vec{\theta})\rangle$ and $|\Psi_B(\vec{\theta})\rangle$ obtained from the SA-VQE algorithm. More precisely, it requires their respective one- and two-electron reduced density matrices, which are measured from the VQE circuit anyhow to compute the energy. Hence, no additional measurements are required. These matrices are used in the analytical formula for the so-called orbital gradient and Hessian which represent fundamental ingredients of the orbital-optimisation process [71, 72, 113]. Note that this approach based on reduced density matrices contrasts with regular VQE gradients which require additional measurements [114, 115].

To close the loop of the SA-OO-VQE algorithm, the (now orbital-optimized) frozen core Hamiltonian $\hat{\mathcal{H}}^{FC}(\vec{\kappa})$ is sent back to the SA-VQE algorithm which will determine the two low-lying states again. Such a

loop between the SA-VQE and the SA-OO algorithms is carried out until reaching a global convergence of the state-averaged energy (in addition to the state-averaged variance if used).

3.2.3. Generalized unitary coupled cluster with double-excitation operators
In VQE-type algorithms, like SA-VQE, one has to choose a specific (and usually approximate) wavefunction ansatz. One promising (and now traditional) ansatz considered for quantum chemistry applications within VQE is the unitary coupled cluster ansatz with single- and double-excitation operators (UCCSD) [20]:

$$U(\vec{\theta}) = e^{\hat{T}(\vec{\theta}) - \hat{T}^\dagger(\vec{\theta})}, \tag{17}$$

$$\hat{T} \equiv \hat{T}_{\mathrm{SD}} = \sum_a^{\mathrm{virt.}} \sum_i^{\mathrm{occ.}} \theta_{ai} \, \hat{a}_a^\dagger \hat{a}_i + \sum_{a>b}^{\mathrm{virt.}} \sum_{i>j}^{\mathrm{occ.}} \theta_{abij} \, \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j, \tag{18}$$

where i, j and a, b denote occupied and virtual spin-orbitals (in reference to the HF slater determinant) *within the active space*, respectively. However, it might be of interest to consider generalizations of equation (18) [23, 42, 52]. In this work, we use a UCC ansatz based on a generalized spin-free form of double-excitation operator

$$\hat{T} = \sum_{t,v,w,u}^{\mathrm{active}} \theta_{tuvw} \sum_{\sigma,\tau=\uparrow,\downarrow} \hat{a}_{t\sigma}^\dagger \hat{a}_{v\tau}^\dagger \hat{a}_{w\tau} \hat{a}_{u\sigma}, \tag{19}$$
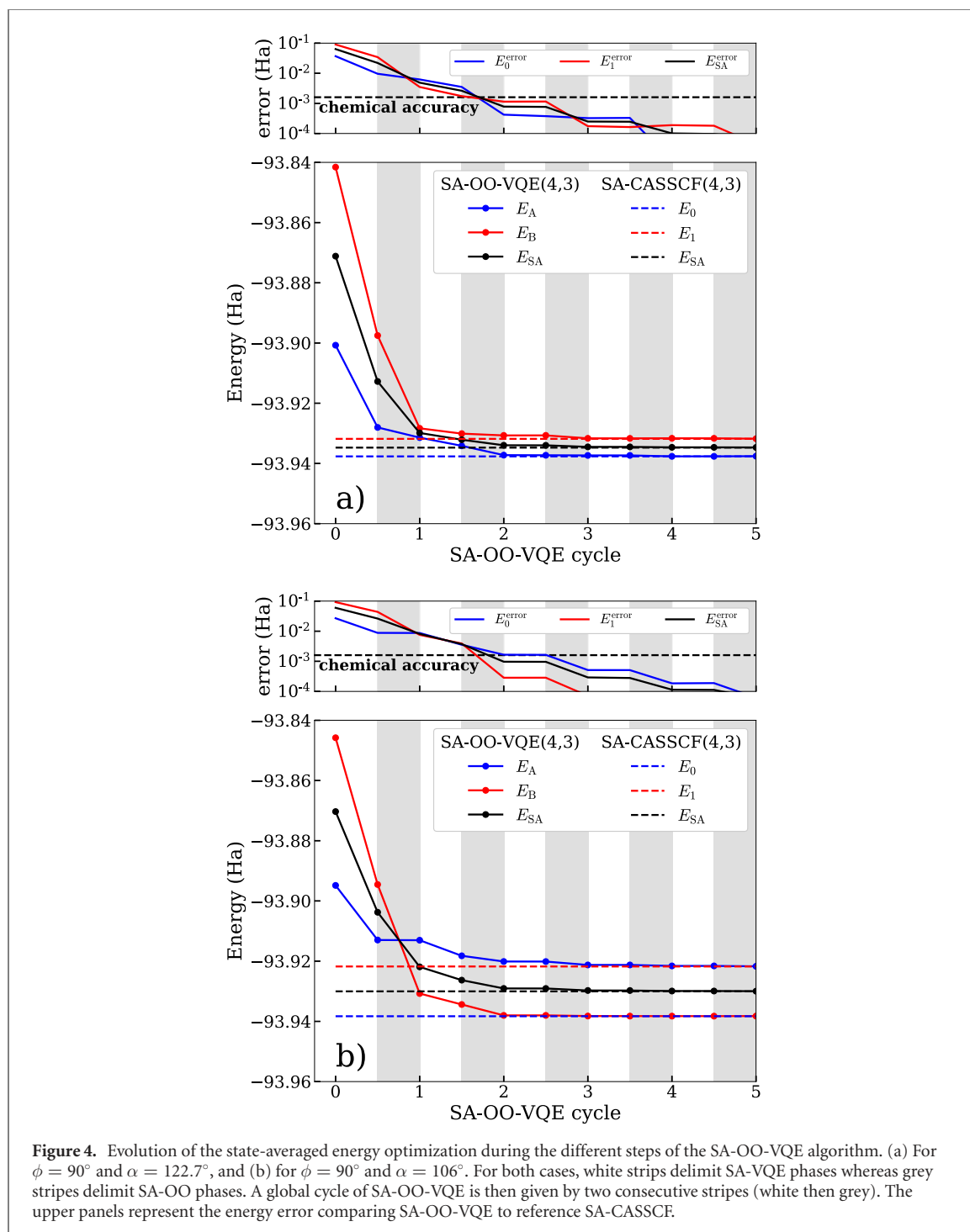
where t, u, v, w denote any active spatial orbital. Comparing equation (19) to equation (18), there are three key alterations using a generalized spin-free double-excitation operator. First, all single excitations can be removed in equation (19) as these are already included in the unitary operator $\hat{U}_{\mathrm{OO}}(\vec{\kappa})$ [equation (14)] as part of the SA-OO algorithm, which effectively encodes all single excitations in the active space. This allows savings of quantum resources by performing the single-excitations classically instead of within the quantum circuit. Second, the traditional UCCSD ansatz only considers excitations from occupied to unoccupied spin-orbitals, and are usually defined according to one specific slater determinant (the HF determinant). However, different initial states are considered in the SA-OO-VQE method. As a direct consequence, operators such as in equation (18) are not flexible enough to simultaneously generate multiple correlated states. For this task, generalized excitation operators are more suitable as the latter remain intrinsically state-agnostic. Finally, equation (19) is chosen to match the spin-free form of the two-body operator present in the electronic Hamiltonian in equation (3). This further reduces the number of independent cluster parameters θ to optimize (four chains of creation-annihilation operators share the same parameters and the symmetry $\theta_{tuvw} = \theta_{vwtu}$ can be exploited).

In appendix E, we discuss how we implemented this ansatz in practice to produce the results presented in our study. Notably, we introduce step-by-step every manipulations realized to make the ansatz tractable in our simulations (i.e. using restrictions to reduce the number of parameters to optimize and invoking symmetries). We also provide additional information about the resulting quantum circuit (such as the ordering of the unitaries and the total number of quantum gates).

4. Computational details

The simulation of the SA-OO-VQE algorithm is realized using the python quantum computing packages **OpenFermion** [116] and **Cirq**[117] []. In practice, SA-VQE parameters $\vec{\theta}$ are always initialized to zero and optimized using the gradient-free sequential least squares programming method (SLSQP) from the python **Scipy** package. For each call of SA-VQE, the SLSQP method is run with a maximum number of 400 iterations and a precision threshold of 10^{-4} Ha (i.e. we set the optimizer parameter 'ftol' to 10^{-4} as implemented in the scipy package). This low expectation on the performance of SA-VQE is rendered possible in SA-OO-VQE thanks to alternating repetitions of SA-VQE and SA-OO algorithms. The threshold for the global convergence of SA-OO-VQE is also set to 10^{-4} Ha. The two initial orthonormal states considered are the HF slater determinant $|\Phi_A\rangle = |HF\rangle$, and its HOMO–LUMO excited version $|\Phi_B\rangle = \hat{E}_{\mathrm{LH}}|HF\rangle = (1/\sqrt{2})\sum_\sigma \hat{a}_{L,\sigma}^\dagger \hat{a}_{H,\sigma}|HF\rangle$ (with 'L' and 'H' referring to the LUMO and HOMO spatial orbitals, respectively). Concerning the SA-OO algorithm, a homemade state-averaged Netwon–Raphson code has been developed based on references [71, 72, 112].

Turning to the geometry of formaldimine, the N–CH$_2$ part of the molecule is frozen and constrained in a same plane (see green plane in figure 1(a)). The interatomic distances are $d_{\mathrm{N-C}} = 1.498$ Å and $d_{\mathrm{C-H}} = 1.067$ Å and the internuclear angles are $N \overset{\frown}{-C-} H = 118.36^\bullet$ The second H atom is symmetric to the first one with respect to the N–C axis. The two remaining degrees of freedom characterize the out-of-plane bending angle $\alpha \equiv H \overset{\frown}{-N-} C$ and the dihedral angle $\phi \equiv H - \overset{\frown}{N-C} - H$.

Figure 4. Evolution of the state-averaged energy optimization during the different steps of the SA-OO-VQE algorithm. (a) For $\phi = 90°$ and $\alpha = 122.7°$, and (b) for $\phi = 90°$ and $\alpha = 106°$. For both cases, white strips delimit SA-VQE phases whereas grey stripes delimit SA-OO phases. A global cycle of SA-OO-VQE is then given by two consecutive stripes (white then grey). The upper panels represent the energy error comparing SA-OO-VQE to reference SA-CASSCF.

In the following, the cc-pVDZ basis is used and an active space of four electrons in three orbitals (4, 3) is considered, unless stated otherwise. The SA-OO procedure is always realized over the first 20 spatial-orbitals of the system (for SA-OO-VQE and SA-CASSCF) composed of the six frozen occupied, the three active and the first eleven virtual orbitals. CASCI and SA-CASSCF calculations are realized with the quantum chemistry code Psi4 [118].

5. Results

We now present the results obtained with SA-OO-VQE for formaldimine's ground- and excited-state spectra around the conical intersection.

5.1. Convergence of the SA-OO-VQE method

We start by illustrating how SA-OO-VQE converges towards the minimal state-averaged energy for a given geometry of the molecule ($\phi = 90°$ and $\alpha = 122.7°$). The different energy values produced by the algorithm after each step of optimization are given in figure 4(a) (where E_0 and E_1 stand for the ground- and first excited-state energies, E_A and E_B are the associated energies of $|\Psi_A\rangle$ and $|\Psi_B\rangle$, and E_{SA} is the state-averaged energy). SA-VQE phases are represented by white stripes while SA-OO phases are represented with grey stripes. One SA-OO-VQE cycle is then given by two consecutive white and grey stripes, repeated until global convergence of the state-averaged energy is reached. We compare the convergence of SA-OO-VQE to the SA-CASSCF reference energies (horizontal dashed lines), which form a natural lower bound on the performance of the method. As readily seen in figure 4(a), alternating between the SA-VQE and the SA-OO algorithms progressively lowers the energies, requiring five SA-OO-VQE cycles to reach global convergence, and only two cycles to reach chemical accuracy compared to the SA-CASSCF reference. Similar convergence was also observed across all simulated formaldimine geometries, with all SA-OO-VQE simulations converging within a maximum of 10 cycles to within chemical accuracy (1.6 mHa) of the SA-CASSCF reference. This fast convergence on the whole one-dimension PES is a consequence of the equi-ensemble theory. Indeed, each state can naturally evolve towards the final state which they have the strongest overlap with, without having to obey any *a priori* ordering (see section 3.2.1). This is depicted in figure 4(b) where the energies E_A and E_B cross each other in the first cycle of SA-OO-VQE, which is by construction not favoured by setting $w_A > w_B$. The importance of such a feature will be discussed in the following.
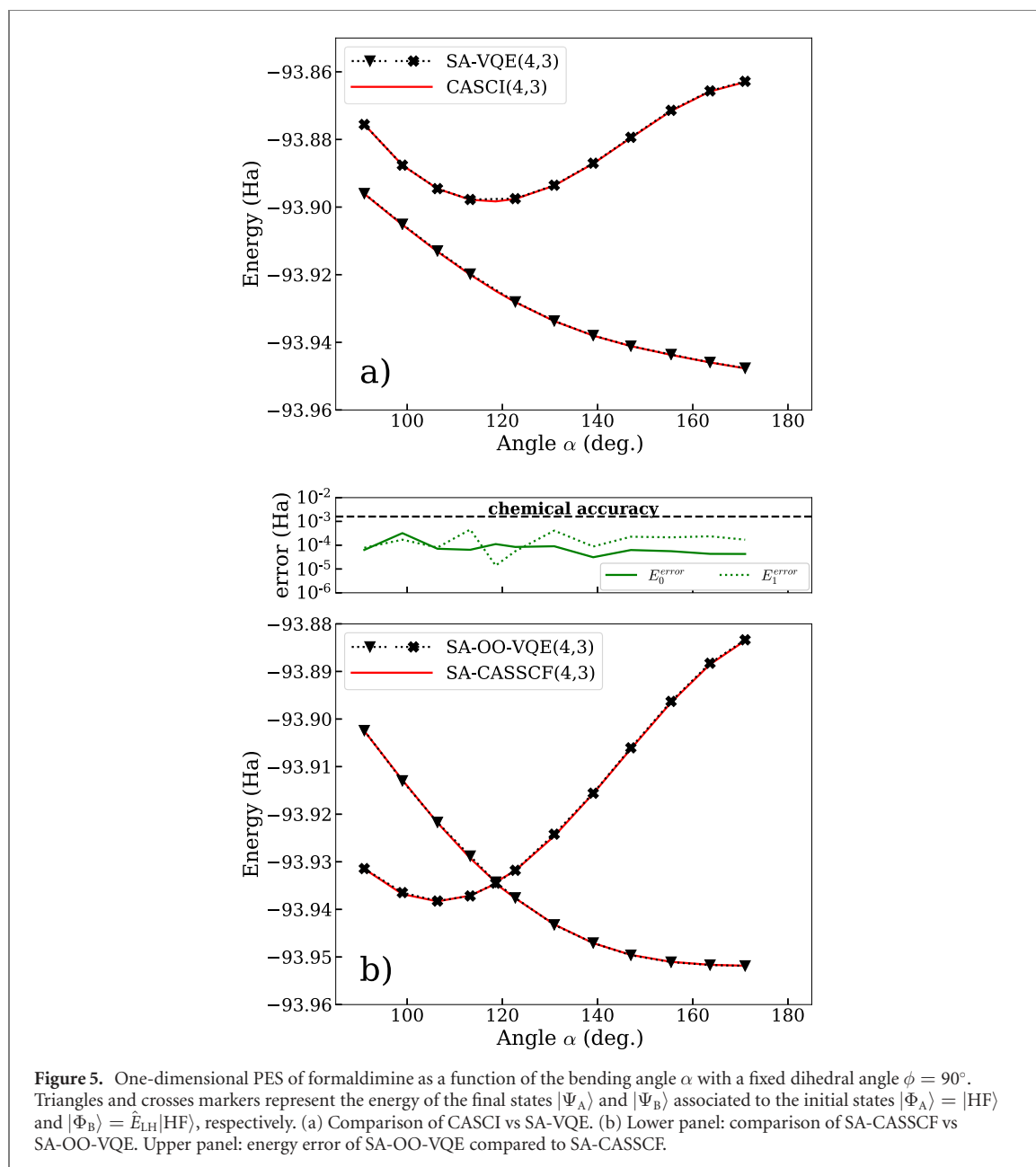
5.2. Description of the conical intersection

Let us now investigate the ability of SA-OO-VQE to reproduce the conical intersection of the formaldimine molecule. We focus on the calculation of the one-dimensional PES of the system as function of the bending angle α with a fixed dihedral angle $\phi = 90°$ (figure 5). In the absence of orbital optimization (i.e. when the SA-VQE algorithm acts alone), the energies are in very good agreement with the classical CASCI calculation [figure 5(a)]. However, no conical intersection is present in the energy profiles thus demonstrating a qualitative failure of the frozen core approximation in the canonical MO basis (see sections 2.3 and 2.4). To reproduce the conical intersection, the SA-OO algorithm is required. In figure 5(b), we plot the energies obtained following global convergence of the full SA-OO-VQE method. We see excellent qualitative and quantitative agreements between SA-OO-VQE and its classical analog SA-CASSCF. Crucially, both methods are able to capture the presence of a conical intersection around $\alpha = 118.5°$. Furthermore, we measure all over the PES an individual energy error between SA-OO-VQE and SA-CASSCF that is always below chemical accuracy (see upper panel of figure 5(b)).

Another important quantity in photochemistry is the excitation energy—the energy difference between the ground and the first excited states. We compare the estimated excitation energies from SA-CASSCF and SA-OO-VQE in figure 6. Here again, SA-OO-VQE matches SA-CASSCF perfectly. Nevertheless, even though the SA-CASSCF reference is reproduced with errors always below chemical accuracy, we recall that the latter will still deviate from the exact energy as defined by the FCI value. As FCI is intractable for the formaldimine in the cc-pVDZ basis, we instead consider CASCI (16, 16), expected to be very close to the FCI value. We see that SA-CASSCF (or SA-OO-VQE) and CASCI (16, 16) agree qualitatively on the existence of a conical intersection, but differ quantitatively: the location of the conical intersection is shifted by 6.5°. This is not a failure of the SA-OO-VQE algorithm itself, but rather that the small active space is insufficient to capture the full dynamic correlation present in the FCI space. This can be solved by enlarging the active space, or potentially by perturbative approaches such as considered in reference [98]. However, in many cases the qualitative agreement is sufficient, and the correct reproduction of the conical intersection at all is vindication of the SA-OO-VQE method.
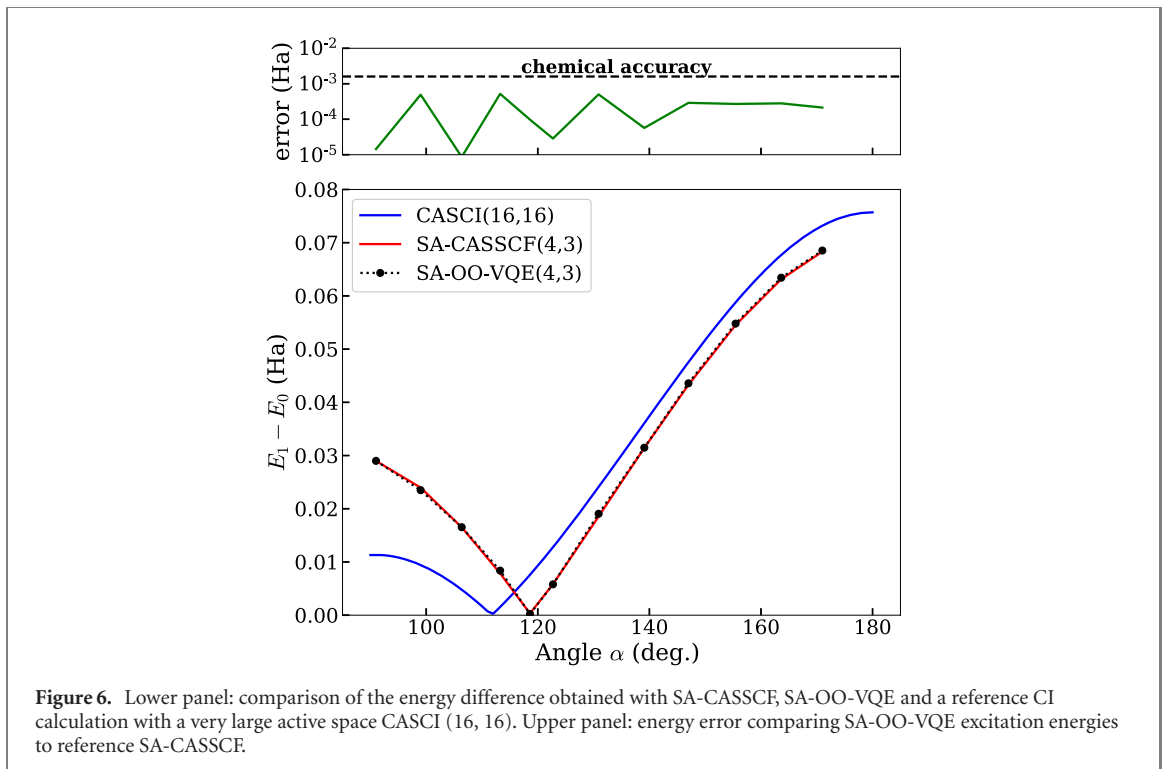
5.3. Diabatic versus adiabatic representation

One interesting feature of our SA-OO-VQE implementation is that the converged states $|\Psi_A\rangle$ and $|\Psi_B\rangle$ cross each other at the conical intersection point, as shown by the triangles and crosses markers in figure 5(b). As discussed in section 3.2.1 (see also reference [49]), this is only possible when considering the equi-weighted version of SA-VQE, as the fully-weighted version of SS-VQE with condition $w_A > w_B$ will force an *a-priori* (energy based) ordering of the states $|\Psi_A(\vec{\theta})\rangle$ and $|\Psi_B(\vec{\theta})\rangle$ [40].

In contrast to the fully-weighted SS-VQE (or SA-CASSCF) method which generate adiabatic states, our states $|\Psi_A(\vec{\theta})\rangle$ and $|\Psi_B(\vec{\theta})\rangle$ keep their nature along the whole PES. This can be rationalized by looking at the decomposition of the wavefunction in the basis of slater determinants (i.e. the configuration interaction coefficients). Consider the converged SA-OO-VQE wavefunctions $|\Psi_A(\vec{\tilde{\kappa}})\rangle$ and $|\Psi_B(\vec{\tilde{\kappa}})\rangle$, where $\vec{\tilde{\kappa}}$ denote the converged SA-OO basis. The dominant configuration in $|\Psi_A(\vec{\tilde{\kappa}})\rangle$ is $|\Phi_A(\vec{\tilde{\kappa}})\rangle$—the HF slater determinant

Figure 5. One-dimensional PES of formaldimine as a function of the bending angle α with a fixed dihedral angle $\phi = 90°$. Triangles and crosses markers represent the energy of the final states $|\Psi_A\rangle$ and $|\Psi_B\rangle$ associated to the initial states $|\Phi_A\rangle = |HF\rangle$ and $|\Phi_B\rangle = \hat{E}_{LH}|HF\rangle$, respectively. (a) Comparison of CASCI vs SA-VQE. (b) Lower panel: comparison of SA-CASSCF vs SA-OO-VQE. Upper panel: energy error of SA-OO-VQE compared to SA-CASSCF.

expressed in the SA-OO basis. The dominant configuration in $|\Psi_B(\vec{\kappa})\rangle$ is $|\Phi_B(\vec{\kappa})\rangle)$—the singly-excited HF slater determinant expressed in the SA-OO basis. These respective dominating configurations contribute to more than 93% for $|\Psi_A(\vec{\kappa})\rangle$ ($|\langle\Phi_A(\vec{\kappa})|\Psi_A(\vec{\kappa})\rangle|^2 > 0.93$) and 88% for $|\Psi_B(\vec{\kappa})\rangle$ ($|\langle\Phi_B(\vec{\kappa})|\Psi_B(\vec{\kappa})\rangle|^2 > 0.88$) along the whole PES, while their contribution on the other state is negligible ($\langle\Phi_A(\vec{\kappa})|\Psi_B(\vec{\kappa})\rangle$ and $\langle\Phi_B(\vec{\kappa})|\Psi_A(\vec{\kappa})\rangle \sim 0$). Hence, it is expected that the SA-OO-VQE states $|\Psi_A(\vec{\kappa})\rangle$ and $|\Psi_B(\vec{\kappa})\rangle$ will always be associated to $|\Phi_A(\vec{\kappa})\rangle$ and $|\Phi_B(\vec{\kappa})\rangle$, respectively, which are the initial states of the SA-VQE algorithm. As the ground and first-excited state of the frozen-core Hamiltonian in the SA-OO basis exhibit a conical intersection (in contrast to the canonical MO basis), $|\Psi_A(\vec{\kappa})\rangle$ and $|\Psi_B(\vec{\kappa})\rangle$ correspond to the ground and first-excited states for $\alpha > 118.5°$ and conversely for $\alpha < 118.5°$, respectively. This change of ordering in the energies of $|\Psi_A(\vec{\kappa})\rangle$ and $|\Psi_B(\vec{\kappa})\rangle$ in figure 5(b) is reflected by the crossing in the energy curves in figure 4(b) and is made possible by using an equi-ensemble.

These observations are not without significance, and suggest that SA-OO-VQE can lead to a diabatic representation of the states, in contrast to SA-CASSCF or fully-weighted SS-VQE. In multistate quantum dynamics simulations, considering diabatic states is crucial to enforce smoothness of physical properties and to reduce the magnitude of non-adiabatic couplings, which explode around conical intersections or avoided crossings, signalling a break down of the Born–Oppenheimer approximation [81, 119]. However, switching to a diabatic representation is not straightforward [120]. Hence, representing diabatic states like in figure 5(b) with SA-OO-VQE will allow the calculation of, for instance, smooth ground- and excited-state

Figure 6. Lower panel: comparison of the energy difference obtained with SA-CASSCF, SA-OO-VQE and a reference CI calculation with a very large active space CASCI (16, 16). Upper panel: energy error comparing SA-OO-VQE excitation energies to reference SA-CASSCF.

energy gradient to perform quantum dynamics around challenging PES featuring conical intersections or avoided crossings.

This diabatic representation in SA-OO-VQE needs to be assessed for more challenging systems, where more states could be required together with more complicated initial states. A rigorous way of capturing the expected individual states in an equi-ensemble formalism is provided by adding the state-averaged variance to the cost function in equation (12). We implemented and tested this augmented cost-function on the formaldimine, but did not see any noticeable improvement compared to using the state-averaged energy alone, as in our case the latter was sufficient to capture the correct eigenstates. As an alternative approach (not employed in this work), one can also use the fully-weighted SS-VQE method (starting with $w_A > w_B$) together with the equi-ensemble SA-OO. As already mentioned in this work and in reference [49], taking $w_A > w_B$ imposes an ordering of the states based on their energy (i.e. $E_A < E_B$), which might not be ideal and may complicate the optimization process. When combining this method with the equi-ensemble SA-OO, the ordering constraint during the SA-OO cycle is relaxed and the energy ordering is allowed to swap freely (just as in figure 4(b)). If this swapping leads to $E_A > E_B$, one can simply change the ordering of the weights for the next fully-weighted SS-VQE cycle (i.e. with $w_A < w_B$), potentially leading to a much easier optimization.
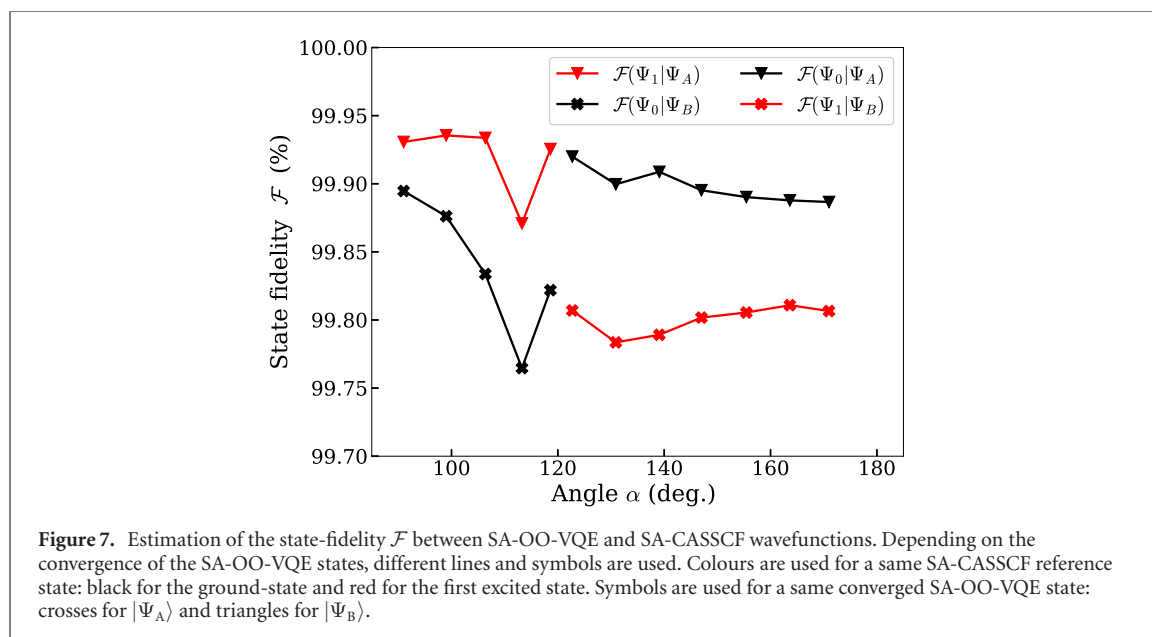
5.4. State fidelity

As a final study, we estimate the similarity of the converged correlated states generated with SA-OO-VQE ($|\Psi_A\rangle$ or $|\Psi_B\rangle$) compared to SA-CASSCF reference wavefunctions ($|\Psi_0\rangle$ and $|\Psi_1\rangle$) by measuring the fidelity

$$\mathcal{F}(\Psi_{\text{ref}}|\Psi) = |\langle\Psi_{\text{ref}}|\Psi\rangle|^2. \qquad (20)$$

The estimation of the state fidelity involves the evaluation of the overlap between two wavefunctions expressed in two different optimized MO basis (the one from SA-OO-VQE and the other one from SA-CASSCF). In practice, this calculation is non-trivial and requires manipulating projections of the MOs from the two different basis [121–123] (generally non-orthogonal to each other, as shown in appendix G).

The state-fidelity over the one dimensional PES of formaldimine is shown in figure 7. The correlated states obtained with the SA-OO-VQE method exhibit a very high fidelity with an average of $\sim 99.85\%$ and a lower bound of $\sim 99.75\%$, thus demonstrating an excellent correspondence to SA-CASSCF wavefunctions. Interestingly, $|\Psi_B\rangle$ shows a slightly lower fidelity compared to $|\Psi_A\rangle$. This could be attributed to the difference in the overlap between the initial states and the final one ($|\langle\Phi_A(\vec{\kappa})|\Psi_A(\vec{\kappa})\rangle|^2 > 0.93$ and $|\langle\Phi_B(\vec{\kappa})|\Psi_B(\vec{\kappa})\rangle|^2 > 0.88$), suggesting that the exact eigenstate $|\Psi_B\rangle$ is more difficult to prepare. This high fidelity together with the accurate energy representation and the diabatic representation of the states clearly

Figure 7. Estimation of the state-fidelity $\mathcal{F}$ between SA-OO-VQE and SA-CASSCF wavefunctions. Depending on the convergence of the SA-OO-VQE states, different lines and symbols are used. Colours are used for a same SA-CASSCF reference state: black for the ground-state and red for the first excited state. Symbols are used for a same converged SA-OO-VQE state: crosses for $|\Psi_A\rangle$ and triangles for $|\Psi_B\rangle$.

encourages the use of SA-OO-VQE wavefunctions as a good starting point for realizing more complicated studies such as, for example, excited state dynamics.

6. Conclusions

In this work, we implemented a variational hybrid quantum/classical algorithm able to capture non-trivial spectral features such as conical intersections. We showed the importance of treating several states on an equal footing, as this is a necessary condition to correctly describe closely degenerate states such as in avoided crossings or conical intersections, which are prominent in photochemistry. Our method remains appropriate in the NISQ era, where quantum resources are limited, as it does not increase the circuit depth of each call of VQE. Instead, it only requires a careful choice of orbital-optimization, i.e. the state-averaged orbital-optimization, together with a quantum solver able to treat all states on an equal footing, such as SA-VQE [40]. We called this method the SA-OO-VQE. As this orbital-optimization is performed classically, the only additional cost on a quantum computer is given by the number of updates of the frozen core Hamiltonian to be solved by SA-VQE, which happens to be smaller than 8 in our simulation (for a convergence criteria of 10^{-4} hartree on the state-averaged energy). As a proof of concept, the performance of SA-OO-VQE was simulated classically and successfully applied to the description of the conical intersection in the formaldimine molecule, a minimal model for a photoisomerization process similar to that occurring in the rhodopsine chromophore. The latter is known to participate in the process of vision, through a photoisomerization reaction mediated by the presence of a conical intersection. Therefore, we expect our algorithm to provide accurate PES for both the ground and excited states, especially for challenging cases in photochemistry applications such as presented in this work. It should be noted that the orbital-optimization procedure (requiring repeated two-electron integral transformations to obtain an updated Hamiltonian) can be costly when applied to large active space and basis sets. As quantum computers promise to increase in size and in accuracy, this could be a potential bottleneck of the algorithm. However, state-averaged orbital-optimization can also be performed differently, as demonstrated for instance in reference [124] on large active spaces. The procedure described in reference [124] alleviates the need of updating the frozen core Hamiltonian at each step, such that a single SA-VQE step would be performed on a quantum computer.

This work highlights important perspectives for near-term quantum computers. Combined with our approach, extending the calculation of atomic forces [46, 125–128] to excited states will pave the way towards excited-state quantum dynamics beyond what is currently tractable on a classical device. Having a control on the excited state dynamics can have important implications in energy conversion and catalysis [129]. This is a clear direction for future research, as it can be used to unravel photochemical processes that are hard to describe with classical methods for which the size of the active spaces becomes prohibitive. This encompasses industrially relevant applications such as the development as OLED devices with improved phosphorescent emitters [130] as well as fundamental questions on how nature has evolved protection mechanisms to prevent damages caused by absorption of too many [131] or too energetic [132] photons.

Acknowledgments

Appendix A. Protection of conical intersections

Adding a real-valued perturbation $V = V(\mathbf{R})$ to a system with a conical intersection will not break the degeneracy, but will instead shift it to a new point $\mathbf{R}$ (as long as equation (7) remains valid). We may write $V(\mathbf{R})$ in a general form

$$V(\mathbf{R}) = V_0(\mathbf{R})I + V_X(\mathbf{R})X + V_Z(\mathbf{R})Z, \tag{A1}$$

and then the conical intersection will re-appear around any points $\mathbf{R}$ where the following two conditions are satisfied

$$V_X(\mathbf{R}) + h_X(\mathbf{R} - \mathbf{R}_0) \cdot \mathbf{R}_X = 0 \tag{A2}$$

$$V_Z(\mathbf{R}) + h_Z(\mathbf{R} - \mathbf{R}_0) \cdot \mathbf{R}_Z = 0. \tag{A3}$$

A solution to these equations exists as long as $\mathbf{R}_X$ and $\mathbf{R}_Z$ are linearly independent and V_X and V_Z are continuous. To lowest order in $\mathbf{R}$, we may write

$$V_X(\mathbf{R}) \sim V_X(\mathbf{R}_0) + \nabla_{\mathbf{R}} V_X(\mathbf{R}_0) \cdot (\mathbf{R} - \mathbf{R}_0) \tag{A4}$$

$$V_Z(\mathbf{R}) \sim V_Z(\mathbf{R}_0) + \nabla_{\mathbf{R}} V_Z(\mathbf{R}_0) \cdot (\mathbf{R} - \mathbf{R}_0), \tag{A5}$$

which may be substituted into equations (A2) and (A3) to obtain

$$(\mathbf{R} - \mathbf{R}_0) \cdot (h_X \mathbf{R}_X + \nabla_{\mathbf{R}} V_X(\mathbf{R}_0)) = -V_X(\mathbf{R}_0) \tag{A6}$$

$$(\mathbf{R} - \mathbf{R}_0) \cdot (h_Z \mathbf{R}_Z + \nabla_{\mathbf{R}} V_Z(\mathbf{R}_0)) = -V_Z(\mathbf{R}_0). \tag{A7}$$

Then, if we define

$$h'_X \mathbf{R}'_X = h_X \mathbf{R}_X + \nabla_{\mathbf{R}} V_X(\mathbf{R}_0) \tag{A8}$$

$$h'_Z \mathbf{R}'_Z = h_Z \mathbf{R}_Z + \nabla_{\mathbf{R}} V_Z(\mathbf{R}_0), \tag{A9}$$

we recover the form of equation (7) with $h_X \to h'_X$, $h_Z \to h'_Z$, $\mathbf{R}_X \to \mathbf{R}'_X$, $\mathbf{R}_Z \to \mathbf{R}'_Z$, and

$$\mathbf{R}_0 \to \mathbf{R}_0 - \frac{V_X(\mathbf{R}_0)}{h'_X |\mathbf{R}'_X|} \mathbf{R}'_X - \left(\frac{V_Z(\mathbf{R}_0)}{h'_Z |\mathbf{R}'_Z|} + \frac{V_X(\mathbf{R}_0)}{h'_X |\mathbf{R}'_X|} \mathbf{R}'_X \cdot \mathbf{R}'_Z \right)$$
$$\times \frac{|\mathbf{R}'_Z|^2 |\mathbf{R}'_X|^2}{|\mathbf{R}'_Z|^2 |\mathbf{R}'_X|^2 - |\mathbf{R}_Z \cdot \mathbf{R}_X|^2} \left(\mathbf{R}'_Z - \frac{\mathbf{R}'_Z \cdot \mathbf{R}'_X}{|\mathbf{R}'_X|^2} \mathbf{R}'_X \right). \tag{A10}$$

Although conical intersections cannot be lifted by a perturbation, they may be lifted by projection. Consider a three-level system spanned by eigenstates $|\Psi_0(\mathbf{R})\rangle$, $|\Psi_1(\mathbf{R})\rangle$, and $|\Psi_2(\mathbf{R})\rangle$, with eigenenergies $E_0(\mathbf{R})$, $E_1(\mathbf{R})$, and $E_2(\mathbf{R})$. Suppose now that there exists a conical intersection between $E_0(\mathbf{R})$ and $E_1(\mathbf{R})$ at $\mathbf{R} = \mathbf{R}_0$, but $E_2(\mathbf{R}) > E_1(\mathbf{R}) \geqslant E_0(\mathbf{R})$ everywhere. The following linear combination of eigenstates

$$|\Phi(\mathbf{R})\rangle = a_1 |\Psi_1(\mathbf{R})\rangle + a_2 |\Psi_2(\mathbf{R})\rangle \tag{A11}$$

will have energy

$$E_\Phi(\mathbf{R}) = |a_1|^2 E_1(\mathbf{R}) + |a_2|^2 E_2(\mathbf{R}). \tag{A12}$$

If we project our three-level system into the subspace spanned only by $|\Phi(\mathbf{R})\rangle$ and $|\Psi_0\rangle$, the resulting Hamiltonian has eigenenergies $E_0(\mathbf{R})$ and $E_\Phi(\mathbf{R})$, which are by our assumption never equal for any $\mathbf{R}$. As the active space Hamiltonian $\hat{\mathcal{H}}^{\text{FC}}$ is generated by projecting the full Hamiltonian $\hat{H}$ onto the active space, this gives a mechanism for such a process breaking the conical intersection. Unless carefully treated (e.g. by the methods in this text), this represents a qualitative failure of the frozen core approximation.

Appendix B. Frozen core Hamiltonian

We describe how to build the frozen core Hamiltonian in practice. We start from the original shape of the second quantized electronic Hamiltonian

$$\hat{H} = \sum_{pq}^{N_o} h_{pq}\hat{E}_{pq} + \frac{1}{2}\sum_{pqrs}^{N_o} g_{pqrs}\hat{e}_{pqrs}. \tag{B1}$$

Applying the frozen core approximation to this Hamiltonian consists in assuming the existence of a set of frozen orbitals (always occupied), another set of active orbitals (belonging to an active space), and a set of virtual orbitals (always unoccupied). Based on this partitioning, every slater determinant $|\Phi\rangle$ used to describe properties of the system will always take the form

$$|\Phi\rangle = |\Phi_{\text{frozen}}\Phi_{\text{active}}\rangle, \tag{B2}$$

where the left contribution Φ_{frozen} represents a part of the determinant encoding the frozen orbitals of the system (always occupied) whereas Φ_{AS} is a part encoding the occupancy of the remaining electrons in the active orbitals of the system. In this context, if one considers that every correlated electronic wavefunction is always expanding in terms of slater determinants following equation (B2), one can demonstrate by projections onto equation (B1) that the system Hamiltonian takes an effective form

$$\langle\Phi|\hat{H}|\Phi\rangle \equiv \langle\Phi_{\text{active}}|\hat{H}^{\text{FC}}|\Phi_{\text{active}}\rangle, \tag{B3}$$

with $\hat{H}^{\text{FC}}$ the so-called 'frozen core Hamiltonian' defined as follows,

$$\hat{H}^{\text{FC}} = \hat{H}_{\text{active}} + E_{\text{frozen}}^{\text{MF}} + \hat{\mathcal{V}}. \tag{B4}$$

Here, $\hat{H}_{\text{AS}}$ is the Hamiltonian encoding the one- and two-body terms only acting in the active space,

$$\hat{H}_{\text{active}} = \sum_{tu}^{\text{active}} h_{tu}\hat{E}_{tu} + \sum_{tuvw}^{\text{active}} g_{tuvw}\hat{e}_{tuvw}, \tag{B5}$$

where t, u, v, w denote active space orbitals. The second term $E_{\text{frozen}}^{\text{MF}}$ is a scalar representing the mean-field-like energy obtained from the frozen orbitals,

$$E_{\text{frozen}}^{\text{MF}} = 2\sum_{i}^{\text{frozen}} h_{ii} + \sum_{ij}^{\text{frozen}} (2g_{iijj} - g_{ijji}), \tag{B6}$$

and the third term

$$\hat{\mathcal{V}} = \sum_{tu}^{\text{active}} \mathcal{V}_{tu}\hat{E}_{tu}, \quad \text{with } \mathcal{V}_{tu} = \sum_{i}^{\text{frozen}} (2g_{tuii} - g_{tiiu}) \tag{B7}$$

represents an effective one body potential which encodes the interaction of the frozen electrons with the active space electrons. To summarize, the main effect of the frozen core approximation [equation (B4)] is first to introduce an energetic shift [equation (B6)], and second to augment the one body term of the Hamiltonian operator [equation (B5)] (that only lives in the active space) with an additional effective one body operator [equation (B7)].

Appendix C. Orbital optimization with Newton–Raphson

We give here some additional details about the SA-OO algorithm in section 3.2.2, based on the classical Netwon–Raphson approach. For simplicity, let us first focus on the case of a single correlated state noted $|\Psi\rangle$ for which we want to optimize the orbitals. The parametrized energy of the state thus reads

$$E_{\Psi}(\vec{\kappa}) = \langle\Psi|\,e^{\hat{\kappa}}\hat{\mathcal{H}}\,e^{-\hat{\kappa}}|\Psi\rangle. \tag{C1}$$

Using a second-order Baker–Campbell–Hausdorff development of $e^{-\hat{\kappa}}\hat{\mathcal{H}}\,e^{\hat{\kappa}}$, equation (C1) becomes

$$E_{\Psi}(\vec{\kappa}) \simeq \langle\Psi|\hat{\mathcal{H}}|\Psi\rangle + \langle\Psi|[\hat{\kappa},\hat{\mathcal{H}}]|\Psi\rangle + \frac{1}{2}\langle\Psi|\left[\hat{\kappa},[\hat{\kappa},\hat{\mathcal{H}}]\right]|\Psi\rangle. \tag{C2}$$

In parallel, we also develop $E_\Psi(\vec{\kappa})$ using a second-order Taylor expansion with respect to the $\vec{\kappa}$ parameters. In matrix form, this expansion reads

$$E_\Psi(\vec{\kappa}) \simeq E_\Psi(0) + \mathbf{G}^\dagger \boldsymbol{\kappa} + \frac{1}{2}\boldsymbol{\kappa}^\dagger \mathbf{H} \boldsymbol{\kappa}, \tag{C3}$$

where $\mathbf{G}$ and $\mathbf{H}$ are the column MO-gradient vector and the MO-hessian matrix of the energy, and $\boldsymbol{\kappa}$ is a vector encoding the parameters $\vec{\kappa}$. From equations (C2) and (C3), we obtain the MO-gradient and MO-Hessian elements

$$\mathbf{G}_{pq} = \langle \Psi | [\hat{E}_{pq}^-, \hat{\mathcal{H}}] | \Psi \rangle$$

$$\mathbf{H}_{pq,rs} = \frac{1}{2}(1 + \mathcal{S}_{(pq)}^{(rs)}) \langle \Psi | \left[\hat{E}_{pq}^-, [\hat{E}_{rs}^-, \hat{\mathcal{H}}] \right] | \Psi \rangle, \tag{C4}$$

where $\hat{E}_{pq}^- = \hat{E}_{pq} - \hat{E}_{qp}$ and $\mathcal{S}_{(pq)}^{(rs)}$ is an operator that permutes the two couples of indices (pq) and (rs). In practice, $\mathbf{G}$ and $\mathbf{H}$ can be expressed analytically in terms of the one- and two-electron reduced density matrices, and one- and two-electron integrals (see references [71, 72, 113]). Once $\mathbf{G}$ and $\mathbf{H}$ are built, a Newton–Raphson step is computed as follows:

$$\Delta \boldsymbol{\kappa} = -\mathbf{H}^{-1}\mathbf{G}, \tag{C5}$$

and used to update the value of the rotation parameters

$$\boldsymbol{\kappa} \leftarrow \boldsymbol{\kappa} + \Delta\boldsymbol{\kappa}. \tag{C6}$$

The new parameters are used to transform the MO coefficient matrix $\mathbf{C}$ [equation (D4)] and by extension the system Hamiltonian in equation (D7) (via a redefinition of the one- and two-electron integrals of the system, see equations (D5) and (D6)]. In practice, the initial conditions considered in the optimisation process are usually located far away from the minimum targeted. Hence, the convergence of the algorithm is not necessarily guaranteed, mainly because the Hessian matrix used to drive the step in equation (C5) can be either non-positive or singular. To solve the issue, a more complicated definition of the Newton–Raphson step can be employed, using for example the so-called 'augmented Hessian' approach:

$$\mathbf{H} \leftarrow \mathbf{H} + \nu\mathbf{1}, \tag{C7}$$

with ν a positive number used to ensure the positive-definiteness of the matrix, and $\mathbf{1}$ the identity matrix. Naturally, the value of ν has to be wisely chosen and many different techniques can be employed to this end (see for instance references [72, 112]). Using this definition to compute the step in equation (C5) ensures convergence of the Newton–Raphson algorithm to a local minimum.

The above description was for a single state. Let us discuss how to optimize MO for a group of N_S correlated states $\{\Psi_k\}_{k=1,\ldots,N_S}$. In this case, a same number of individual Hessians $\mathbf{H}_{\Psi_k}$ and gradients $\mathbf{G}_{\Psi_k}$ have to be computed. One then builds a state-averaged version of these matrices as follows:

$$\mathbf{H}_{SA} = \sum_k^{N_S} w_k \mathbf{H}_{\Psi_k}, \quad \text{and} \quad \mathbf{G}_{SA} = \sum_k^{N_S} w_k \mathbf{G}_{\Psi_k}, \tag{C8}$$

where $\mathbf{H}_{SA}$ and $\mathbf{G}_{SA}$ are respectively the state-averaged Hessian and gradient of the ensemble $\{\Psi_k\}_{k=1,\ldots,N_S}$. The weights w_k (with $\sum_k^{N_S} w_k = 1$) encode the contribution of each state into the orbital-optimization process. Using these state-averaged matrices, a Newton–Raphson step like in equation (C5) is built and the system's MOs can be optimized in *a democratic way* with respect to all states considered in the ensemble.

To close this discussion, we mention two potential numerical bottlenecks that can arise when using orbital optimization. The first one is the repetitive transformations used to build two-electron integrals in new optimised MO basis. Such a transformation is known to scale as N^5 (see reference [133]), when not employing approximations such as density fitting (see reference [134]). The second bottleneck is related to the explicit construction of the orbital Hessian during each Newton–Raphson step. This is a very active research problem in quantum chemistry for which many new approaches and methods are being developed currently (see for example references [89, 135, 136]).

Appendix D. Hamiltonian transformation

We describe how the full second-quantized Hamiltonian transformation is realized in practice during the orbital-optimization process. Let us start with the $\mathbf{C}$ matrix encoding the MO coefficients of the system,

$$\phi_p(\mathbf{r}) = \sum_\mu \mathbf{C}_{lp}\chi_\mu(\mathbf{r}), \tag{D1}$$

where $\phi_p(\mathbf{r})$ is a generic MO expressed as a linear combination of atomic orbitals $\chi_\nu(\mathbf{r})$. Considering the initial MO basis $\{\phi_p\}$, we build a unitary operator

$$\mathbf{U}_{OO} = e^{-\mathbf{K}} \quad \text{with } \mathbf{K} = \text{skew}(\boldsymbol{\kappa}), \tag{D2}$$

where $\mathbf{K}$ is a skew-symmetric matrix shaped by the column vector $\boldsymbol{\kappa}$ containing the rotational parameters. This operator is used to change the original MO basis into a new one $\{\phi'_q\}$ with

$$\phi'_q(\mathbf{r}) = \sum_p \phi_p(\mathbf{r})(\mathbf{U}_{OO})_{pq}. \tag{D3}$$

To realize this transformation and obtain a $\mathbf{C}$ matrix encoding the new MO basis, one applies $\mathbf{U}_{OO}$ on the right of $\mathbf{C}$ and then update the latter such that

$$\mathbf{C} \leftarrow \mathbf{C}\mathbf{U}_{OO}. \tag{D4}$$

The new MO coefficients matrix $\mathbf{C}$ is then used to transform the one- and two-electron integrals in the new MO basis. To do so, one starts from the one- and two-electron integrals expressed in the atomic orbital basis, respectively $h_{\mu\nu}$ and $g_{\mu\nu\gamma\delta}$, and implements the following two- and four-indexes transformations:

$$h_{pq}^{\text{new}} = \sum_{\mu,\nu} h_{\mu\nu}\mathbf{C}_{\mu p}\mathbf{C}_{\nu q}, \tag{D5}$$

and

$$g_{pqrs}^{\text{new}} = \sum_{\mu,\nu,\gamma,\delta} g_{\mu\nu\gamma\delta}\mathbf{C}_{\mu p}\mathbf{C}_{\nu q}\mathbf{C}_{\gamma r}\mathbf{C}_{\delta s}. \tag{D6}$$

The Hamiltonian of the system expressed in the new MO basis then reads

$$\hat{H} = \sum_{pq}^{N_{MO}} h_{pq}^{\text{new}}\hat{E}_{pq}^{\text{new}} + \frac{1}{2}\sum_{pqrs}^{N_{MO}} g_{pqrs}^{\text{new}}\hat{e}_{pqrs}^{\text{new}}. \tag{D7}$$

Appendix E. Numerical implementation of the generalized spin-free double-excitation ansatz

We detail here our practical implementation of the generalized spin-free double excitation ansatz. Starting from the original definition of the excitation operator

$$\hat{T}(\vec{\theta}) = \sum_{t,v,w,u}^{\text{active}} \theta_{tuvw} \sum_{\sigma,\tau=\uparrow,\downarrow} \hat{a}_{t\sigma}^\dagger \hat{a}_{v\tau}^\dagger \hat{a}_{w\tau} \hat{a}_{u\sigma}, \tag{E1}$$

the scaling of the total number of independent variational parameters θ to optimize with SA-VQE is N^4 (with N the number of active space MOs). In practice, this scaling could represent a bottleneck for the energy optimization realized with the SA-VQE sub-algorithm (especially when multiple calls of SA-VQE are required to reach global convergence of the SA-OO-VQE method). To reduce the number of parameters and favour a faster convergence, we simplify equation (E1) by imposing two conditions. First, we considered the indices restriction $t \geqslant v \geqslant w \geqslant u$ over the active space MOs (in addition to discarding terms where $t = v = w = u$). In practice, this reduces the total number of parameters from 81 to 12 for the three MO active-space considered in our simulations. Second, we invoke the symmetry property $\theta_{tuvw} = \theta_{vwtu}$ to recover part of the variational parameters that were discarded with the previous restriction on the indices. This is inspired by the properties $\hat{e}_{pqrs} = \hat{e}_{rspq}$ of the spin-free two-body term of the Hamiltonian [equation (3)] (see reference [72]). Within the above conditions, the excitation operator simplifies as

$$\hat{T}(\vec{\theta}) = \sum_{\substack{t \geqslant v \geqslant w \geqslant u \\ \text{not}(t=v=w=u)}}^{\text{active}} \theta_{tuvw}(1 + \mathcal{S}_{(tu)}^{(vw)}) \sum_{\sigma,\tau=\uparrow,\downarrow} \hat{a}_{t\sigma}^\dagger \hat{a}_{v\tau}^\dagger \hat{a}_{w\tau} \hat{a}_{u\sigma}, \tag{E2}$$

where $\mathcal{S}_{(tu)}^{(vw)}$ is an operator that permutes the couples of indices (tu) and (vw). Based on this simplified version of the excitation operator, we built up a quantum circuit implementing the unitary

Algorithm 1. Ansatz encoding.

```
1: N = 3                                    ▷Number of active MOs
2: for u = 1 : N do
3:    for t = 1 : N do
4:       for w = 1 : N do
5:          for v = 1 : N do
6:             if not(t = u = v = w) then
7:                if (t ⩾ v ⩾ w ⩾ u) then
8:                   Set σ =↑, τ =↑
9:                   Encode e^{θ_{tuvw}(â†_{tσ}â†_{vτ}â_{wτ}â_{uσ}−h.c.)}
10:                  Set σ =↓, τ =↑
11:                  Encode e^{θ_{tuvw}(â†_{tσ}â†_{vτ}â_{wτ}â_{uσ}−h.c.)}
12:                  Set σ =↑, τ =↓
13:                  Encode e^{θ_{tuvw}(â†_{tσ}â†_{vτ}â_{wτ}â_{uσ}−h.c.)}
14:                  Set σ =↓, τ =↓
15:                  Encode e^{θ_{tuvw}(â†_{tσ}â†_{vτ}â_{wτ}â_{uσ}−h.c.)}
16:                  Swap the indices: t ↔ v, u ↔ w
17:                  Encode the same exponentials
18:                  (with the same parameter θ_{tuvw})
```

$\hat{U}(\vec{\theta}) = e^{\hat{T}(\vec{\theta}) - \hat{T}(\vec{\theta})^\dagger}$. To proceed, we chose to apply a single Trotter step to split up the original unitary into products of exponentials as follows:

$$\hat{U}(\vec{\theta}) = e^{\hat{T}(\vec{\theta}) - \hat{T}(\vec{\theta})^\dagger}$$

$$\approx \prod_{\substack{t \geqslant v \geqslant w \geqslant u \\ \text{not}(t=v=w=u)}} \prod_{\sigma,\tau=\uparrow,\downarrow} e^{\theta_{tuvw}(\hat{a}^\dagger_{t\sigma}\hat{a}^\dagger_{v\tau}\hat{a}_{w\tau}\hat{a}_{u\sigma} - \text{h.c.})} \tag{E3}$$

$$\prod_{\sigma,\tau=\uparrow,\downarrow} e^{\theta_{tuvw}(\hat{a}^\dagger_{v\sigma}\hat{a}^\dagger_{t\tau}\hat{a}_{u\tau}\hat{a}_{w\sigma} - \text{h.c.})},$$

where 'h.c.' stands for 'hermitian conjugate'. Then, we encode each of the operators into a quantum circuit with an order described by the following pseudo-code:(algorithm 1)

In practice, every time we have to encode a fermionic exponential in the quantum circuit, we apply the Jordan–Wigner transformation to the associated difference of chain of fermionic operators. We then obtain a series of exponentiated Pauli-strings we translate into circuit form using pretty common implementation techniques (namely, chains of CNOT gates and single gate rotations, as described in reference [137]). Within this scheme, the total number of quantum gates in our six-qubits quantum circuit is estimated at 3904 gates (2080 single-qubit gates and 1824 two-qubit gates).

Appendix F. Variational quantum algorithms for excited states

While variational quantum algorithms were originally designed to extract the ground state of a given Hamiltonian, several extensions to excited states have been recently developed. We provide here a brief discussion on these methods and their capability to describe several states on an equal footing.

In section 2, we insisted on the importance of a democratic description of the states involved in the conical intersection. This democratisation can be achieved by performing a single minimization for all states sharing the same ansatz, as in multistate-contracted VQE (MC-VQE) [45, 46], fully-weighted subspace-search VQE (SS-VQE) [40] and variance-VQE [49]. The other versions described in the SS-VQE paper require additional minimizations or maximizations to extract the excited states one by one, thus potentially breaking the democratic treatment of the states as the optimization landscape might be more complex from one state to another.

Other methods to extract excited states can be grouped in three types.

First, methods which consists in expanding about the reference ground state to form a linear (or beyond) subspace. The optimal solutions within this subspace are approximations to the excited states and are obtained by solving a generalized eigenvalue equation on a classical computer. These methods include quantum subspace expansion [3, 38] and the quantum equation-of-motion coupled-cluster [39]. Although the excited states are treated on the same footing, the ground state is favoured by construction. Hence, an accurate description of a conical intersection between the ground and first excited state is unlikely, as in linear response TDDFT [70]. Similar conclusions also hold for the quantum Lanczos algorithm described in

reference [48]. Note that other methods imply to solve generalized eigenvalue equations on a classical computer but without favouring any state and are therefore democratic, for instance in the quantum filter diagonalization method [138, 139].

Second, variational algorithms based on penalization of the Hamiltonian. In order to get the *k*th excited state, these methods penalize the Hamiltonian by the (orthogonal) ground and $(k-1)$th excited states [41–44]. This procedure is sequential, which may lead to error accumulation [42]. As each state is determined by a separate minimization, these methods are by construction 'state-specific'. Hence, although in principle exact, the latter are not expected to coherently treat a conical intersection, except maybe for relatively simple cases.

Third, methods based on phase estimation (like for instance the witness-assisted variational eigenspectra solver [140]) are democratic, as phase estimation is the quantum algorithm equivalent to matrix diagonalization. However, for NISQ devices these approaches remain ill-adapted due to their very high computational cost.

Appendix G. Overlap of wavefunctions expressed in different MO basis

We describe how to compute the non-trivial overlap between two many-electron wavefunctions expressed in two different MO basis. Let us consider two distinct correlated wavefunctions $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$ which form a linear combination of slater determinants

$$|\Psi\rangle = \sum_I d_I |\Phi_I\rangle \quad \text{and} \quad |\tilde{\Psi}\rangle = \sum_J \tilde{d}_J |\tilde{\Phi}_J\rangle, \tag{G1}$$

where $\{|\Phi_I\rangle\}$ and $\{|\tilde{\Phi}_J\rangle\}$ represent two (sometimes different) sets of slater determinant expressed in two different MO basis $\{|\phi_p\rangle\}$ and $\{|\tilde{\phi}_q\rangle\}$. Our goal here is to compute

$$\langle\Psi|\tilde{\Psi}\rangle = \sum_I \sum_J d_I^* \tilde{d}_J \langle\Phi_I|\tilde{\Phi}_J\rangle. \tag{G2}$$

To proceed, the overlap $\langle\Phi_I|\tilde{\Phi}_J\rangle$ between two slater determinants needs to be estimated. This calculation is non-trivial as the two determinants considered are expressed in different (and usually non-orthogonal) MO basis which generally leads to $\langle\Phi_I|\tilde{\Phi}_J\rangle \neq \delta_{IJ}$. Based on the spin-orbital occupancy defining the slater determinants,

$$|\Phi_I\rangle = |\phi_1 \dots \phi_{N_e}| \quad \text{and} \quad |\tilde{\Phi}_J\rangle = |\tilde{\phi}_1 \dots \tilde{\phi}_{N_e}|, \tag{G3}$$

their overlap can be computed as follows (see references [121, 122] for a proof):

$$\langle\Phi_I|\tilde{\Phi}_J\rangle = \begin{vmatrix} \langle\phi_1|\tilde{\phi}_1\rangle & \dots & \langle\phi_1|\tilde{\phi}_{N_e}\rangle \\ \vdots & \ddots & \vdots \\ \langle\phi_{N_e}|\tilde{\phi}_1\rangle & \dots & \langle\phi_{N_e}|\tilde{\phi}_{N_e}\rangle \end{vmatrix}. \tag{G4}$$

In other words, the overlap of two slater determinants expressed in two different MO basis is given by the determinant of the matrix containing all mutual spin-orbital overlaps. Here, a generic spin-orbital overlap $\langle\phi_p|\tilde{\phi}_q\rangle$ is defined as

$$\langle\phi_p|\tilde{\phi}_q\rangle = \int \phi_p^*(\mathbf{x})\tilde{\phi}_q(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{G5}$$

where $\mathbf{x} \equiv (\mathbf{r}, \sigma)$ denotes both spatial and spin coordinates of the spin-orbitals. Note that when a same MO basis is used, many simplifications occur leading to $\langle\phi_p|\tilde{\phi}_q\rangle = \delta_{pq}$ and to the trivial result $\langle\Phi_I|\tilde{\Phi}_J\rangle = \delta_{IJ}$.

ORCID iDs

Saad Yalouz https://orcid.org/0000-0002-8818-3379

References

[1] O'Malley P J *et al* 2016 *Phys. Rev. X* **6** 031007
[2] Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow J M and Gambetta J M 2017 *Nature* **549** 242
[3] Colless J I, Ramasesh V V, Dahlen D, Blok M S, Kimchi-Schwartz M, McClean J, Carter J, De Jong W and Siddiqi I 2018 *Phys. Rev. X* **8** 011021

[4] Hempel C *et al* 2018 *Phys. Rev.* X **8** 031022

[5] Bruzewicz C D, Chiaverini J, McConnell R and Sage J M 2019 *Appl. Phys. Rev.* **6** 021314

[6] Arute F *et al* 2019 *Nature* **574** 505

[7] Arute F *et al* 2020 *Science* **369** 1084

[8] Nam Y *et al* 2020 *npj Quantum Inf.* **6** 1

[9] Reiher M, Wiebe N, Svore K M, Wecker D and Troyer M 2017 *Proc. Natl Acad. Sci. USA* **114** 7555

[10] Li Z, Li J, Dattani N S, Umrigar C J and Chan G K-L 2019 *J. Chem. Phys.* **150** 024302

[11] Cao Y *et al* 2019 *Chem. Rev.* **119** 10856

[12] McArdle S, Endo S, Aspuru-Guzik A, Benjamin S C and Yuan X 2020 *Rev. Mod. Phys.* **92** 015003

[13] Bauer B, Bravyi S, Motta M and Chan G K 2020 arXiv:2001.03685

[14] von Burg V, Low G H, Häner T, Steiger D S, Reiher M, Roetteler M and Troyer M 2020 arXiv:2007.14460

[15] Preskill J 2018 *Quantum* **2** 79

[16] Abrams D S and Lloyd S 1999 *Phys. Rev. Lett.* **83** 5162

[17] Aspuru-Guzik A, Dutoi A D, Love P J and Head-Gordon M 2005 *Science* **309** 1704

[18] O'Brien T E, Tarasinski B and Terhal B M 2019 *New J. Phys.* **21** 023022

[19] Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love P J, Aspuru-Guzik A and O'brien J L 2014 *Nat. Commun.* **5** 4213

[20] McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 *New J. Phys.* **18** 023023

[21] Mai S and González L 2020 *Angew. Chem., Int. Ed.* **59** 16832

[22] Romero J, Babbush R, McClean J R, Hempel C, Love P J and Aspuru-Guzik A 2018 *Quantum Sci. Technol.* **4** 014008

[23] Lee J, Huggins W J, Head-Gordon M and Whaley K B 2018 *J. Chem. Theory Comput.* **15** 311

[24] Ryabinkin I G, Yen T-C, Genin S N and Izmaylov A F 2018 *J. Chem. Theory Comput.* **14** 6317

[25] McArdle S, Jones T, Endo S, Li Y, Benjamin S C and Yuan X 2019 *npj Quantum Inf.* **5** 1

[26] Mizukami W, Mitarai K, Nakagawa Y O, Yamamoto T, Yan T and Ohnishi Y-y 2019 arXiv:1910.11526

[27] Kühn M, Zanker S, Deglmann P, Marthaler M and Weiß H 2019 *J. Chem. Theory Comput.* **15** 4764

[28] Rattew A G, Hu S, Pistoia M, Chen R and Wood S 2019 arXiv:1910.09694v4

[29] Sokolov I O, Barkoutsos P K, Ollitrault P J, Greenberg D, Rice J, Pistoia M and Tavernelli I 2020 *J. Chem. Phys.* **152** 124107

[30] Ryabinkin I G, Lang R A, Genin S N and Izmaylov A F 2020 *J. Chem. Theory Comput.* **16** 1055

[31] Lang R A, Ryabinkin I G and Izmaylov A F 2020 arXiv:2002.05701

[32] Matsuzawa Y and Kurashige Y 2020 *J. Chem. Theory Comput.* **16** 944

[33] Huggins W J, Lee J, Baek U, O'Gorman B and Whaley K B 2020 *New J. Phys.* **22** 073009

[34] Gomes N, Zhang F, Berthusen N F, Wang C-Z, Ho K-M, Orth P P and Yao Y 2020 arXiv:2006.15371

[35] Meitei O R, Gard B T, Barron G S, Pappas D P, Economou S E, Barnes E and Mayhall N J 2020 arXiv:2008.04302

[36] Wang Q, Li M, Monroe C and Nam Y 2020 arXiv:2004.04151

[37] Zhang Z-J, Kyaw T H, Kottmann J S, Degroote M and Aspuru-Guzik A 2020 arXiv:2008.07553

[38] McClean J R, Kimchi-Schwartz M E, Carter J and de Jong W A 2017 *Phys. Rev.* A **95** 042308

[39] Ollitrault P J *et al* 2019 arXiv:1910.12890

[40] Nakanishi K M, Mitarai K and Fujii K 2019 *Phys. Rev. Res.* **1** 033062

[41] Ibe Y, Nakagawa Y O, Yamamoto T, Mitarai K, Gao Q and Kobayashi T 2020 arXiv:2002.11724

[42] Higgott O, Wang D and Brierley S 2019 *Quantum* **3** 156

[43] Jones T, Endo S, McArdle S, Yuan X and Benjamin S C 2019 *Phys. Rev.* A **99** 062304

[44] Jouzdani P, Bringuier S and Kostuk M 2019 arXiv:1908.05238

[45] Parrish R M, Hohenstein E G, McMahon P L and Martínez T J 2019 *Phys. Rev. Lett.* **122** 230401

[46] Parrish R M, Hohenstein E G, McMahon P L and Martinez T J 2019 arXiv:1906.08728

[47] Bauman N P, Low G H and Kowalski K 2019 *J. Chem. Phys.* **151** 234114

[48] Motta M, Sun C, Tan A T K, O'Rourke M J, Ye E, Minnich A J, Brandão F G S L and Chan G K-L 2020 *Nat. Phys.* **16** 205

[49] Zhang D-B, Yuan Z-H and Yin T 2020 arXiv:2006.15781

[50] Ryabinkin I G, Genin S N and Izmaylov A F 2018 *J. Chem. Theory Comput.* **15** 249

[51] Ryabinkin I G and Genin S N 2018 arXiv:1812.09812

[52] Greene-Diniz G and Ramo D M 2019 arXiv:1910.05168

[53] Zhang F, Gomes N, Berthusen N F, Orth P P, Wang C-Z, Ho K-M and Yao Y-X 2020 arXiv:2006.11213

[54] Gard B T, Zhu L, Barron G S, Mayhall N J, Economou S E and Barnes E 2020 *npj Quantum Inf* **6** 1

[55] Seki K, Shirakawa T and Yunoki S 2020 *Phys. Rev.* A **101** 052340

[56] Domcke W, Yarkony D R and Köppel H 2011 *Conical Intersections: Theory, Computation and Experiment* vol 17 (Singapore: World Scientific)

[57] Yarkony D R 2012 *Chem. Rev.* **112** 481

[58] Garavelli M, Celani P, Bernardi F, Robb M A and Olivucci M 1997 *J. Am. Chem. Soc.* **119** 6891

[59] González-Luque R, Garavelli M, Bernardi F, Merchán M, Robb M A and Olivucci M 2000 *Proc. Natl Acad. Sci.* **97** 9379

[60] Polli D *et al* 2010 *Nature* **467** 440

[61] Valsson O, Campomanes P, Tavernelli I, Rothlisberger U and Filippi C 2013 *J. Chem. Theory Comput.* **9** 2441

[62] Manathunga M, Yang X, Luk H L, Gozem S, Frutos L M, Valentini A, Ferrè N and Olivucci M 2016 *J. Chem. Theory Comput.* **12** 839

[63] Olaso-González G, Merchán M and Serrano-Andrés L 2006 *J. Phys. Chem* B **110** 24734

[64] Kang H, Lee K T, Jung B, Ko Y J and Kim S K 2002 *J. Am. Chem. Soc.* **124** 12958

[65] Groenhof G, Schäfer L V, Boggio-Pasqua M, Goette M, Grubmüller H and Robb M A 2007 *J. Am. Chem. Soc.* **129** 6812

[66] Barbatti M, Aquino A J A, Szymczak J J, Nachtigallová D, Hobza P and Lischka H 2010 *Proc. Natl Acad. Sci.* **107** 21453

[67] Francés-Monerris A, Gattuso H, Roca-Sanjuán D, Tuñón I, Marazzi M, Dumont E and Monari A 2018 *Chem. Sci.* **9** 7902

[68] May V and Kühn O 2008 *Charge and Energy Transfer Dynamics in Molecular Systems* (New York: Wiley)

[69] Ho E K-L and Lasorne B 2019 *Comput. Theor. Chem.* **1156** 25

[70] Gozem S *et al* 2014 *J. Chem. Theory Comput.* **10** 3074

[71] Siegbahn P E M, Almlöf J, Heiberg A and Roos B O 1981 *J. Chem. Phys.* **74** 2384

[72] Helgaker T, Jorgensen P and Olsen J 2014 *Molecular Electronic-Structure Theory* (New York: Wiley)

[73] Klessinger M and Michl J 1995 *Excited States and Photochemistry of Organic Molecules* (New York: VCH Publishers)

[74] Robb M A, Bernardi F and Olivucci M 1995 *Pure Appl. Chem.* **67** 783

[75] Bernardi F, Olivucci M and Robb M A 1996 *Chem. Soc. Rev.* **25** 321

[76] Domcke W *et al* 2004 *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy* vol 15 (Singapore: World Scientific)

[77] Birge R R 1990 *Biochim. Biophys. Acta Bioenerg.* **1016** 293

[78] Menzel J P, de Groot H J M and Buda F 2019 *J. Phys. Chem. Lett.* **10** 6504

[79] Baldo M A, O'Brien D F, You Y, Shoustikov A, Sibley S, Thompson M E and Forrest S R 1998 *Nature* **395** 151

[80] Marian C M 2012 *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2** 187

[81] Gatti F, Lasorne B, Meyer H-D and Nauts A 2017 *Applications of Quantum Dynamics in Chemistry* vol 98 (Berlin: Springer)

[82] González L and Lindh R 2020 *Quantum Chemistry and Dynamics of Excited States: Methods and Applications* (New-York: Wiley)

[83] Casida M E 1995 *Recent Advances in Density Functional Methods: (Part I)* (Singapore: World Scientific) pp 155–92

[84] Marques M A L and Gross E K U 2004 *Annu. Rev. Phys. Chem.* **55** 427

[85] Ullrich C A 2011 *Time-dependent Density-Functional Theory: Concepts and Applications* (Oxford: Oxford University Press)

[86] Casida M E and Huix-Rotllant M 2012 *Annu. Rev. Phys. Chem.* **63** 287

[87] Maitra N T, Zhang F, Cave R J and Burke K 2004 *J. Chem. Phys.* **120** 5932

[88] Fuks J I and Maitra N T 2014 *Phys. Chem. Chem. Phys.* **16** 14504

[89] Vogiatzis K D, Ma D, Olsen J, Gagliardi L and De Jong W A 2017 *J. Chem. Phys.* **147** 184111

[90] Williams K T *et al* 2020 *Phys. Rev. X* **10** 011041

[91] Stair N H and Evangelista F A 2020 arXiv:2005.11349

[92] Eriksen J J *et al* 2020 arXiv:2008.02678

[93] Loos P-F, Damour Y and Scemama A 2020 arXiv:2008.11145

[94] Lloyd S 1996 *Science* **273** 1073

[95] Whitfield J D, Biamonte J and Aspuru-Guzik A 2011 *Mol. Phys.* **109** 735

[96] Berry D W, Gidney C, Motta M, McClean J R and Babbush R 2019 *Quantum* **3** 208

[97] Bauman N P, Bylaska E J, Krishnamoorthy S, Low G H, Wiebe N, Granade C E, Roetteler M, Troyer M and Kowalski K 2019 *J. Chem. Phys.* **151** 014107

[98] Takeshita T, Rubin N C, Jiang Z, Lee E, Babbush R and McClean J R 2020 *Phys. Rev. X* **10**

[99] Kowalski K and Bauman N P 2020 *J. Chem. Phys.* **152** 244127

[100] Motta M, Gujarati T P, Rice J E, Kumar A, Masteran C, Latone J A, Lee E, Valeev E F and Takeshita T Y 2020 arXiv:2006.02488

[101] Metcalf M, Bauman N P, Kowalski K and de Jong W A 2020 arXiv:2004.07721

[102] Urbanek M, Camps D, Beeumen R V and de Jong W A 2020 arXiv:2002.12902

[103] McArdle S and Tew D P 2020 arXiv:2006.11181

[104] Bylaska E J, Song D, Bauman N P, Kowalski K, Claudino D and Humble T S 2020 arXiv:2009.00080

[105] Rossmannek M, Barkoutsos P K, Ollitrault P J and Tavernelli I 2020 arXiv:2009.01872

[106] Andersson K, Malmqvist P Å and Roos B O 1992 *J. Chem. Phys.* **96** 1218

[107] Angeli C, Cimiraglia R, Evangelisti S, Leininger T and Malrieu J-P 2001 *J. Chem. Phys.* **114** 10252

[108] Bonačić-Koutecký V and Michl J 1985 *Theor. Chim. Acta* **68** 45

[109] Chahre M 1985 *Annu. Rev. Biophys. Biophys. Chem.* **14** 331

[110] Bonet-Monroig X, Babbush R and O'Brien T E 2019 arXiv:1908.05628

[111] Gross E K U, Oliveira L N and Kohn W 1988 *Phys. Rev.* A **37** 2805

[112] Fletcher R 2013 *Practical Methods of Optimization* (New York: Wiley)

[113] Yarkony D 1995 *Modern Electronic Structure Theory* vol 2 (Singapore: World Scientific)

[114] Grimsley H R, Claudino D, Economou S E, Barnes E and Mayhall N J 2019 *J. Chem. Theory Comput.* **16** 1

[115] Tang H L, Barnes E, Grimsley H R, Mayhall N J and Economou S E 2019 arXiv:1911.10205

[116] McClean J R *et al* 2020 *Quantum Sci. Technol.* **5** 034014

[117] Quantum AI team and collaborators Quantum AI team and collaborators Cirq Zenodo Oct 2020 https://doi.org/10.5281/zenodo.4062499

[118] Smith D G A *et al* 2020 *J. Chem. Phys.* **152** 184108

[119] Baer M 2006 *Beyond Born–Oppenheimer: Electronic Nonadiabatic Coupling Terms and Conical Intersections* (New York: Wiley)

[120] Zhang Y, Su P, Lasorne B, Braïda B and Wu W 2020 *J. Phys. Chem. Lett.* **11** 5295–301

[121] Plasser F, Ruckenbauer M, Mai S, Oppel M, Marquetand P and González L 2016 *J. Chem. Theory Comput.* **12** 1207

[122] Löwdin P-O 1955 *Phys. Rev.* **97** 1474

[123] Malmqvist P Å, Roos B O and Schimmelpfennig B 2002 *Chem. Phys. Lett.* **357** 230

[124] Shepard R and Brozell S R 2019 *Mol. Phys.* **117** 2374

[125] Kassal I and Aspuru-Guzik A 2009 *J. Chem. Phys.* **131** 224102

[126] O'Brien T E, Senjean B, Sagastizabal R, Bonet-Monroig X, Dutkiewicz A, Buda F, DiCarlo L and Visscher L 2019 *npj Quantum Inf.* **5** 1

[127] Mitarai K, Nakagawa Y O and Mizukami W 2020 *Phys. Rev. Res.* **2** 013129

[128] Sokolov I O, Barkoutsos P K, Moeller L, Suchsland P, Mazzola G and Tavernelli I 2020 arXiv:2008.08144

[129] Paulus B C, Adelman S L, Jamula L L and McCusker J K 2020 *Nature* **582** 214

[130] Ligthart A, de Vries X, Zhang L, Pols M C W M, Bobbert P A, van Eersel H and Coehoorn R 2018 *Adv. Funct. Mater.* **28** 1804618

[131] Ostroumov E E, Götze J P, Reus M, Lambrev P H and Holzwarth A R 2020 *Photosynth. Res.* **144** 171

[132] Markovitsi D 2016 *Photochem. Photobiol.* **92** 45

[133] Yoshimine M 1973 *J. Comput. Phys.* **11** 449

[134] Feyereisen M, Fitzgerald G and Komornicki A 1993 *Chem. Phys. Lett.* **208** 359

[135] Kreplin D A, Knowles P J and Werner H-J 2020 *J. Chem. Phys.* **152** 074102

[136] Kreplin D A, Knowles P J and Werner H-J 2019 *J. Chem. Phys.* **150** 194106

[137] Seeley J T, Richard M J and Love P J 2012 *J. Chem. Phys.* **137** 224109

[138] Parrish R M and McMahon P L 2019 arXiv:1909.08925

[139] Bespalova T A and Kyriienko O 2020 arXiv:2009.03351

[140] Santagati R *et al* 2018 *Sci. Adv.* **4** eaap9646