Calculation of quasi-diabatic states within the DFT/MRCI(2) framework: the QD-DFT/MRCI(2) method

Simon P. Neville[1] and Michael S. Schuurman[1, 2]

[1)] *National Research Council Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada*

[2)] *Department of Chemistry and Biomolecular Sciences, University of Ottawa, 10 Marie Curie, Ottawa, Ontario, K1N 6N5, Canada*

(*Electronic mail: Simon.Neville@nrc-cnrc.gc.ca)

(Dated: 8 April 2024)

We describe a procedure for the calculation of quasi-diabatic states within the recently introduced DFT/MRCI(2) framework [J. Chem. Phys., **157**, 164103 (2022)]. Based on an effective Hamiltonian formalism, the proposed procedure, which we term QD-DFT/MRCI(2), has the advantageous characteristics of being simultaneously highly efficient and effectively black box in nature, while directly yielding both quasi-diabatic potentials and wave functions of high quality. The accuracy and efficiency of the QD-DFT/MRCI(2) formalism are demonstrated via the simulation of the vibronic absorption spectra of furan and chlorophyll a.

I. INTRODUCTION

The starting point for quantum dynamics simulations is the choice of representation. For studies of excited-state dynamics, the adoption of a diabatic representation is usually preferable as this ameliorates the problems arising due to singular non-adiabatic couplings in the adiabatic representation. Here, the coupling of the electronic and nuclear degrees of freedom is transformed into the off-diagonal elements of the potential matrix, which are smooth, well-behaved functions of the nuclear coordinates. For polyatomic molecules, strictly diabatic representations, for which the troublesome derivative couplings are rendered zero, are non-existent[1]. Instead, one seeks a quasi-diabatic representation, in which the singular components of the derivative couplings are removed and the, hopefully small, finite residual components can be safely neglected. Since such a transformation is not rigorously defined, there exist a plethora of proposed quasi-diabatisation procedures[2].

Beyond the choice of quasi-diabatisation procedure, one must choose an underlying level of electronic structure theory. Here, a level of theory that can account for the, in general, multi-reference nature of excited electronic states is highly desirable. Furthermore, in order to be reliably applicable to large molecular systems, computational efficiency (while ideally not sacrificing accuracy) is a necessity. One powerful class of methods that somewhat uniquely fulfils these three criteria are those based on the combined density functional theory and multi-reference configuration interaction (DFT/MRCI) formalism[3,4]. First formulated by Grimme and Waletzke[3], the DFT/MRCI method combines the flexibility of the MRCI method with semi-empirical DFT-specific corrections that allow for the recovery of the preponderance electron correlation whilst using short wave function expansions. Furthermore, the DFT/MRCI method can be implemented in a completely black box manner using an automated reference space selection and refinement algorithm[5], leaving the user to specify only the number of roots of interest.

Although already very efficient in its original formulation, recently introduced semi- and fully-perturbative approximations have resulted in new methods within the DFT/MRCI family, termed p-DFT/MRCI[5] and DFT/MRCI(2)[6], respectively, that achieve additional computational savings of 2-3 orders of magnitude but also retain the accuracy of the original method. The DFT/MRCI(2) approximation in particular will be the focus of this current work because, as shall be detailed in the following, its roots in effective Hamiltonian theory are found to lend itself to the solution of a number of difficulties often encountered in the calculation of quasi-diabatic states.

In Reference 7, we introduced a procedure for the quasi-diabatisation of DFT/MRCI electronic states based on the elegant and rigorous block diagonlisation diabatisation (BDD) framework of Pacher, Köppel and Cederbaum[8–10]. The method, which propagatively generates quasi-diabatic states (and was accordingly termed P-BDD), was shown to correctly capture the non-adiabatic coupling effects in a number of prototypical systems. However, the proposed methodology does have two significant limitations that need to be overcome for widespead, black box usage. First, the P-BDD working equations exhibit a singularity when a higher-lying state crosses into the subspace of interest, which in general cannot be avoided. Second, the P-BDD method requires the calculation of overlaps between wave functions at neighboring geometries expressed in terms of non-orthogonal molecular orbital (MO) bases, which can become prohibitively expensive for very large molecular systems.

The main result of this paper is to show that both of these limitations can be overcome by working within the DFT/MRCI(2) framework. In particular, as we shall demonstrate, the state-crossing problem may be eliminated by the adoption of a root-following procedure, which is made straightforward given that the eigenstates furnished by DFT/MRCI(2) are obtained from perturbation theory. At the same time, the effective Hamiltonian formulation of DFT/MRCI(2) may be exploited to arrive at a framework for the *direct* calculation of quasi-diabatic wave functions and potentials that may be considered as a perturbative approximation to those yielded by the P-BDD procedure. As shall be detailed, the proposed method requires the overlaps of only the reference space wave functions of a given geometry with the full quasi-diabatic wave functions of a neighboring one, leading to large computational savings compared to the original P-BDD procedure. As this new formalism makes direct use of the working components of a DFT/MRCI(2) calculation and directly yields quasi-diabatic potentials and wave functions, we term it QD-DFT/MRCI(2), the "QD" standing for "quasi-diabatic".

We note that this work builds on old ideas for constructing quasi-diabatic potentials and wave functions using quasi-degenerate perturbation theory (QDPT)[11–13]. However, previous QDPT-based quasi-diabatisation schemes have been severely hampered by the use of, in general, poor zeroth-order quasi-diabatic states expanded in terms of overly-small, hand-selected sets of electronic configurations. Moreover, such schemes were largely superseded by the BDD framework of Pacher, Köppel and Cederbaum[8–10], which provides a non-perturbative route to the same desired outcome. We show here, however, that an implementation within the DFT/MRCI(2) framework overcomes these problems. Namely, it is found possible to automatically construct zeroth-order quasi-diabatic states that lead to results of excellent quality.

The rest of this paper is as follows. In Section II A, we provide an overview of the DFT/MRCI and DFT/MRCI(2) methods, with a focus on aspects pertinent to the development of an improved diabatisation scheme. In Section II B, we discuss the P-BDD algorithm, focusing on its limitations. In Section II D, we introduce the QD-DFT/MRCI(2) method. In Section IV, we present some selected applications of QD-DFT/MRCI(2): (i) the simulation of absorption spectrum of furan, demonstrating the accuracy of the approach, and; (ii) calculation of the Q and B bands of the absorption spectrum of chlorophyll a, showing the applicabilty of the method to large systems. Finally, in Section V, we present our conclusions and outlook for potential applications of the QD-DFT/MRCI(2) method.

II. THEORY

A. The DFT/MRCI framework

We begin by providing a brief overview of the DFT/MRCI and DFT/MRCI(2) methods. In each case, the adiabatic electronic states $|\Psi_I\rangle$ are expanded in terms of a basis of spin-adapted configuration state functions (CSFs) $\{|\Omega\rangle\}$:

$$|\Psi_I\rangle = \sum_{\Omega\in\mathscr{R}} C_{\Omega I}|\Omega\rangle + \sum_{\Omega\in\mathscr{F}} C_{\Omega I}|\Omega\rangle. \tag{1}$$

Here, we separate the contributions from the reference space $\mathscr{R}$ and first-order interacting space (FOIS) $\mathscr{F}$. The CSFs belonging to the, in general, small reference space $\mathscr{R}$ are chosen to recover the static correlation and give a qualitatively correct description of the states of interest. The much larger FOIS, obtained via single and double excitations out of the reference space, accounts for the missing dynamic correlation to first-order in perturbation theory.

1. DFT/MRCI

In an *ab initio* MRCI calculation, the size of the FOIS quickly becomes prohibitively large with increasing reference space dimension and molecular size. In the DFT/MRCI method, only a small subset of the FOIS is used, which is made possible via the use of DFT-specific Hamiltonian corrections aimed at the recovery of the preponderance of the dynamic electron correlation in a manner removed from the interaction of the reference and FOIS CSFs. The starting point here is the recognition that the *ab initio* on-diagonal Hamiltonian matrix elements may be written in terms of the difference from the self-consistent field (SCF) energy as follows[14,15]:

$$\begin{aligned}\langle \Omega | \hat{H} - E_{\mathrm{SCF}} | \Omega \rangle &= \langle \mathbf{w}\omega | \hat{H} - E_{\mathrm{SCF}} | \mathbf{w}\omega \rangle \\ &= \sum_i F_{ii} \Delta \mathrm{w}_i + \frac{1}{2} \sum_{ij} \left(V_{iijj} - \frac{1}{2} V_{ijji} \right) \Delta \mathrm{w}_i \Delta \mathrm{w}_j \\ &+ \frac{1}{2} \sum_{i \in S_{\mathbf{w}}} \sum_{j \in S_{\mathbf{w}}} V_{ijji} \left[\eta_{ij}^{ji}(\mathbf{w}, \omega, \mathbf{w}, \omega)(1 - \delta_{ij}) - \frac{1}{2} \right].\end{aligned} \tag{2}$$

Here, each CSF $|\Omega\rangle$ is specified by a spatial occupation vector $\mathbf{w}$, giving the occupancies of the MOs, and a spin coupling patern ω: $|\Omega\rangle = |\mathbf{w}\omega\rangle$. $\Delta \mathrm{w}_p = \mathrm{w}_p - \overline{\mathrm{w}}_p$ denotes the difference of the occupation of the pth spatial orbital relative to a base, or anchor, occupation $\overline{\mathbf{w}}$, chosen as the Hartree-Fock (HF) occupation. $S_{\mathbf{w}}$ is the set of indices of singly-occupied orbitals in the configuration $\mathbf{w}$, $\boldsymbol{F}$ is the Fock matrix, V_{pqrs} denotes a two-electron integral in the Chemists' notation, and η_{ij}^{ji} is a spin-coupling coefficient,

$$\eta_{ij}^{ji}(\mathbf{w}, \omega, \mathbf{w}, \omega) = \langle \mathbf{w}\omega | \hat{E}_i^j \hat{E}_j^i | \mathbf{w}\omega \rangle, \tag{3}$$

where $\hat{E}_i^j$ denotes a singlet excitation operator. From Equation 2, we see that the on-diagonal matrix elements may be interpreted in a particle-hole picture, where the first term takes the form of the difference in on-diagonal Fock matrix elements between the orbitals involved in the excitation linking $\mathbf{w}$ and the base configuration $\overline{\mathbf{w}}$. The second term corresponds to Coulomb and exchange interactions between these particle and hole orbitals, and the third term accounts for the spin-coupling between the open shells in $\mathbf{w}$. Grimme's insight was to recognise that the differences between KS orbital energies, ε_p^{KS}, are typically closer to ground-to-excited-state excitation energies than the corresponding differences in Fock matrix elements. Substituting the Fock matrix elements for the corresponding KS orbital energies, accompanied by Coulomb and exchange corrections, leads to the general form of the on-diagonal DFT/MRCI Hamiltonian matrix elements:

$$\langle \mathbf{w}\omega | \hat{H} - E_{\mathrm{DFT}} | \mathbf{w}\omega \rangle = (E_{DFT} - E_{SCF}) + \sum_i \varepsilon_{ii}^{KS} \Delta \mathrm{w}_i + \Delta E_C + \Delta E_X, \tag{4}$$

where the Coulomb and exchange corrections, ΔE_C and ΔE_X, respectively, account for a down-scaling of the *ab initio* Coulomb and exchange integrals that enter into Equation 4, the exact forms of which vary between implementations[3,16,17].

Building the KS orbital energies into the on-diagonal Hamiltonian matrix elements effectively incorporates a large amount of dynamic electron correlation that would otherwise be accounted for by the coupling of the reference and FOIS CSFs. To avoid a double counting of dynamic correlation, the off-diagonal Hamiltonian matrix elements must be appropriately adjusted. This is achieved through the introduction of a damping of the off-diagonal elements that is dependent on the energetic separation of the bra and ket CSFs:

$$\left\langle \Omega \middle| \hat{H} \middle| \Omega' \right\rangle \rightarrow \left\langle \Omega \middle| \hat{H} \middle| \Omega' \right\rangle \cdot D(\Delta E_{\Omega\Omega'}), \tag{5}$$

where

$$\Delta E_{\Omega\Omega'} = \frac{1}{n_\omega} \sum_{\omega}^{n_\omega} H_{\Omega,\Omega}^{DFT} - \frac{1}{n_{\omega'}} \sum_{\omega'}^{n_{\omega'}} H_{\Omega',\Omega'}^{DFT} \tag{6}$$

denotes the spin-coupling-averaged difference between the on-diagonal matrix elements corresponding to the bra and ket CSFs. The damping function $D(\Delta E)$ is chosen to decay rapidly with increasing ΔE: in practice chosen as either an exponential[3,18] or inverse arctangent[16,17] function. In this way, the coupling of energetically distant reference and FOIS CSFs is damped to near-zero, thereby avoiding to a large extent a double counting of dynamic correlation.

The decoupling of a large part of the FOIS from the reference space means that most of the FOIS CSFs are no longer required. These are identified *a priori* using a simple orbital energy-based selection criterion[3], which proceeds as follows. For each FOIS configuration **w**, the quantity

$$d_{\mathbf{w}} = \sum_{p} \Delta \mathrm{w}_p \varepsilon_p^{KS} - \delta E_{sel} \tag{7}$$

is computed, where δE_{sel} is a parameter with a value conventionally chosen as either 1.0 or 0.8 $\mathrm{E_h}$. If $d_{\mathbf{w}}$ is less than the highest reference space eigenvalue of interest, then all the CSFs generated from the configuration **w** are selected for inclusion, else they are discarded. This configuration selection step results in a massive reduction of the size of the CSF basis, typically by many orders of magnitude, and results in huge speedups relative to an *ab initio* MRCI calculation.

2. *The DFT/MRCI(2) approximation*

Although the above-described energy-based configuration selection procedure results in CSF bases that are exceptionally compact relative to an *ab initio* MRCI calculation, for large molecules the dimension of the DFT/MRCI Hamiltonian matrix may be of the order of $10^6 - 10^9$, and it's iterative diagonalisation can become prohibitively expensive. To obviate this, a new, perturbative approach to account for the interaction of the reference and FOIS spaces was recently introduced, termed DFT/MRCI(2)[6]. The main idea here is the replacement of the large, $N_{CSF} \times N_{CSF}$ DFT/MRCI Hamiltonian with a small, $N_{state} \times N_{state}$ effective Hamiltonian obtained via the application of the Epstein-Nesbet Hamiltonian partitioning[19,20] and 2nd-order generalised van Vleck perturbation theory[21]. The resulting effective Hamiltonian reads

$$\left[\mathscr{H}^{[2]}_{DFT}\right]_{IJ} = \delta_{IJ} E^{(0)}_I + \frac{1}{2} \sum_{\Omega \in \mathscr{F}'} B_{\Omega I} B_{\Omega J} \sum_{i=I,J} \left(E^{(0)}_i - E^{(0)}_\Omega \right)^{-1}, \tag{8}$$

where

$$B_{\Omega I} = \langle \Omega | \hat{H}^{DFT} | \Psi^{(0)}_I \rangle, \tag{9}$$

$\mathscr{F}'$ denotes the subset of FOIS CSFs that have survived the energy-based configuration selection, and $|\Psi^{(0)}_I\rangle$ and E^0_I are the eigenpairs of the DFT/MRCI Hamiltonian projected onto the reference space $\mathscr{R}$:

$$\left(\sum_{\Omega, \Omega' \in \mathscr{R}} |\Omega\rangle \langle \Omega | \hat{H}^{DFT} | \Omega' \rangle \langle \Omega' | \right) |\Psi^{(0)}_I\rangle = E^{(0)}_I |\Psi^{(0)}_I\rangle. \tag{10}$$

Here, and in the following, the notation $\langle \cdot | \hat{H}^{DFT} | \cdot \rangle$ is used to denote a Hamiltonian matrix element that has been subjected to the DFT/MRCI corrections.

Diagonalisation of the effective Hamiltonian $\mathscr{H}^{[2]}_{DFT}$ yields second-order approximations $E^{[2]}_I$ to the energies of the states of interest as well as first-order corrected wave functions

$$|\Psi^{[1]}_I\rangle = \sum_J X_{JI} \left[|\Psi^{(0)}_J\rangle + \sum_{\Omega \in \mathscr{F}'} \left(\frac{B_{\Omega J}}{E^{(0)}_J - E^{(0)}_\Omega} \right) |\Omega\rangle \right], \tag{11}$$

where $\boldsymbol{X}$ is the matrix of eigenvectors of $\mathscr{H}^{[2]}_{DFT}$.

Remarkably, the errors in excitation energies introduced by the DFT/MRCI(2) effective Hamiltonian approximation are found to be of the order of 10^{-2} eV[6], whilst speedups of between two and three orders of magnitude are realised for large molecular systems.

B. Propagative block diagonalisation diabatisation

We here give a brief overview of the P-BDD scheme, highlighting aspects that impact negatively on both stability and computational efficiency. For a more in depth description of the method, we direct the reader to Reference 7 as well as the original block diagonalisation diabatisation papers of Pacher, Köppel and Cederbaum[8–10]. In the following, where appropriate, the geometry dependence of the electronic states will be explicitly referenced, with $\boldsymbol{R}$ used to denote the vector of nuclear coordinates.

At the heart of the BDD formalism is the idea that, within the group Born-Oppenheimer approximation, it should be possible to identify a subset P of states of interest that are only weakly coupled to their orthogonal complement, termed the Q space. It then follows that the ADT matrix of interest can be approximated to only operate within the P space, thus leading to a block diagonal quasi-diabatic potential matrix. As such, within the BDD framework, we first represent the Hamiltonian in an, as of yet unspecified, 'initial state' basis $\{|\Phi_I(\boldsymbol{R})\rangle\}$,

$$H_{IJ}(\boldsymbol{R}) = \langle \Phi_I(\boldsymbol{R}) | \hat{H} | \Phi_J(\boldsymbol{R}) \rangle, \tag{12}$$

and seek a block-diagonalising transformation $\boldsymbol{T}(\boldsymbol{R})$ that: (i) operates only within the P space, and; (ii) results in a quasi-diabatic representation. As there are an infinite number of transformations that will yield the desired block diagonal structure, a constraint has to be imposed. In the BDD framework, this is taken as the following 'least action principle':

$$||\boldsymbol{T}(\boldsymbol{R}) - \mathbf{1}|| = \min, \tag{13}$$

where $||\cdot||$ denotes the Frobenius norm. This least action principle leads to a unique solution[22] but, in order to yield quasi-diabatic states, requires that the initial states $|\Phi_I(\boldsymbol{R})\rangle$ already behave somewhat quasi-diabatically themselves. In the P-BDD approach, the initial states $|\Phi_I(\boldsymbol{R}+\Delta\boldsymbol{R})\rangle$ are taken as the quasi-diabatic states $\{|\phi_I(\boldsymbol{R})\rangle\}$ taken from a neighboring geometry. Via the enforcement of the least action principle (Equation 13), the adiabatic-to-diabatic transformation (ADT) $\boldsymbol{U}(\boldsymbol{R}+\Delta\boldsymbol{R})$ is found to take the form[8]:

$$\boldsymbol{U}(\boldsymbol{R}+\Delta\boldsymbol{R}) = \boldsymbol{S}^{-1}(\boldsymbol{R}+\Delta\boldsymbol{R}) \left[\boldsymbol{S}(\boldsymbol{R}+\Delta\boldsymbol{R}) \boldsymbol{S}^{\dagger}(\boldsymbol{R}+\Delta\boldsymbol{R}) \right]^{\frac{1}{2}}, \tag{14}$$

with

$$S_{IJ}(\boldsymbol{R}+\Delta\boldsymbol{R}) = \langle \phi_I(\boldsymbol{R}) | \Psi_J(\boldsymbol{R}+\Delta\boldsymbol{R}) \rangle. \tag{15}$$

Here, the $|\Psi_I\rangle$ are the adiabatic electronic states to be used in the expansion of the quasi-diabatic states,

$$|\phi_I(\boldsymbol{R})\rangle = \sum_J U_{JI}(\boldsymbol{R}) |\Psi_I(\boldsymbol{R})\rangle. \tag{16}$$

To fix the global gauge of the ADT and to provide an initial set of quasi-diabatic states, the ADT is taken to be the unit matrix at a user-specified reference geometry $\boldsymbol{R}_0$. Quasi-diabatic states and potentials may then be computed along a chain of neighboring nuclear geometries $\boldsymbol{R}_0, \boldsymbol{R}_1, \dots, \boldsymbol{R}_n$, taking the quasi-diabatic states computed at geometry $\boldsymbol{R}_{i-1}$ as input for the calculation of the ADT matrix at geometry $\boldsymbol{R}_i$. We note that the P-BDD procedure may be trivially incorporated into a DFT/MRCI(2) calculation, with the first-order corrected wave functions $|\Psi_I^{[1]}\rangle$ (Equation 11) being substituted for the variationally determined adiabatic wave functions $|\Psi_I\rangle$.

The P-BDD method is optimal in the sense that, in the limit of infinitesimal displacements $\Delta\boldsymbol{R}$, the integral of the squared L^2 norm of the resulting quasi-diabatic derivative coupling tensor along the path $\boldsymbol{R}_0, \boldsymbol{R}_1, \dots, \boldsymbol{R}_n$ is minimised[10]. However, in practice, two significant limitations appear. The first problem concerns the inversion of the overlap matrix $\boldsymbol{S}(\boldsymbol{R})$. If a higher-lying adiabatic state crosses into the manifold of interest at some geometry $\boldsymbol{R}_i$, then, taking the basis $\{|\Psi_I(\boldsymbol{R}_i)\rangle\}$ to comprise the N lowest-lying adiabats, the matrix $\boldsymbol{S}(\boldsymbol{R})$ can become near-singular, or even truly singular in high-symmetry cases. We shall refer to this as the 'state-crossing problem'. The second problem pertains to the calculation of the overlap matrix

$$\begin{aligned} S_{IJ}(\boldsymbol{R}_i) &= \langle \phi_I(\boldsymbol{R}_{i-1}) | \Psi_J(\boldsymbol{R}_i) \rangle \\ &= \sum_K U_{IK}^{\dagger}(\boldsymbol{R}_{i-1}) \langle \Psi_K(\boldsymbol{R}_{i-1}) | \Psi_J(\boldsymbol{R}_i) \rangle . \end{aligned} \tag{17}$$

This requires the evaluation of overlaps between bra and ket wave functions expressed in terms of non-orthogonal MOs. Even though the original DFT/MRCI P-BDD implementation employed the efficient intermediate-reuse algorithm of Plasser *et al.*[23], this step still becomes prohibitively expensive for very large molecules. We shall be refer to this as the 'overlap problem'.

C. Solution of the state-crossing problem: root following

In order to avoid (near) singular overlap matrices $\boldsymbol{S}$ (Equation 15), we employ a simple root-following procedure that allows the states that correlate with those computed at the reference geometry $\boldsymbol{R}_0$ to be calculated at all displaced geometries. At geometry $\boldsymbol{R}_i$, the reference space Hamiltonian is explicitly built and all its eigenpairs computed. As DFT/MRCI(2) employs highly tailored and compact individually selected reference spaces, the reference space is usually spanned by somewhere between a few tens and a few hundreds of CSFs. Thus, this full diagonalisation step is computationally trivial to perform. Next, the overlaps between all reference space wave functions $|\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle$ and the first-order corrected wave functions $|\Psi_I^{[1]}(\boldsymbol{R}_{i-1})\rangle$ of the previous geometry are computed. Here, highly-truncated Slater determinant expansions are used in order to maintain computational tracability. The reference space wave functions with greatest absolute overlap values are then selected for inclusion in the subsequent DFT/MRCI(2) effective Hamiltionian calculation.

For this root-following procedure to work, the reference space $\mathscr{R}$ at $\boldsymbol{R}_i$ must have good support for the DFT/MRCI(2) wave functions computed at $\boldsymbol{R}_{i-1}$. To ensure that this requirement holds, $\mathscr{R}$ is constructed from the reference space of the previous geometry as follows. Let $\{\mathbf{w}^{(0)}(\boldsymbol{R})\}$ denote the set of reference space configuration vectors at geometry $\boldsymbol{R}$: $\mathrm{w}_p^{(0)}(\boldsymbol{R})$ has values of 0, 1, or 2 depending on whether the pth MO is unoccupied, single-occupied, or doubly-occupied. The reference space configurations $\mathbf{w}^{(0)}(\boldsymbol{R}_i)$ at the ith geometry are constructed from those $\mathbf{w}^{(0)}(\boldsymbol{R}_{i-1})$ of the previous geometry via an approximate MO transformation:

$$\mathbf{w}^{(0)}(\boldsymbol{R}_i) = \boldsymbol{K}\mathbf{w}^{(0)}(\boldsymbol{R}_{i-1}), \tag{18}$$

$$K_{pq} = \delta_{p\mathscr{I}(q)}, \tag{19}$$

$$\mathscr{I}(q) = \overset{max}{p'} \left\{ \left| \langle \varphi_{p'}(\boldsymbol{R}_i) | \varphi_q(\boldsymbol{R}_{i-1}) \rangle \right| \right\}, \tag{20}$$

where $\{\varphi_p(\boldsymbol{R})\}$ is the MO basis at geometry $\boldsymbol{R}$. This reference space propagation algorithm is predicated upon a near-one-to-one mapping between the MOs at neighboring geometries, which will only hold for small nuclear displacements. However, as the P-BDD procedure also requires small displacements to give reliable results, this is not problematic.

Using the above described root-following algorithm, we have found that the states of interest may be reliably and selectively computed across a wide range of nuclear geometries, even when embedded in dense manifolds of other states. This essentially solves the state-crossing problem.

D. Solution of the overlap problem: the QD-DFT/MRCI(2) method

To arrive at approximations of the P-BDD quasi-diabatic poentials and states that do not require the calculation of full wave function overlaps between neighboring geometries, we consider the field of effective Hamiltonian theories[24–28]. Common to all effective Hamiltonian formalisms is the determination of a block diagonalisation of the Hamiltonian $\hat{H}$ initially represented in the basis of eigenfunctions of a zeroth-order Hamiltonian $\hat{H}_0$. Let the eigenfunctions of $\hat{H}$ be partitioned into a subspace P of interest and its orthogonal complement Q. The eigenfunctions of $\hat{H}_0$ are partitioned into the corresponding subspaces P_0, and Q_0. A transformation $\hat{T}$ is sought such that the similarity transformed Hamiltonian

$$\hat{\mathscr{H}} = \hat{T}^{-1}\hat{H}\hat{T} \tag{21}$$

is block diagonal between the P_0 and Q_0 subspaces. That is,

$$\hat{\mathscr{H}} = \hat{P}_0\hat{\mathscr{H}}\hat{P}_0 + \hat{Q}_0\hat{\mathscr{H}}\hat{Q}_0, \tag{22}$$

where $\hat{P}_0$ and $\hat{Q}_0$ are the projectors onto the P_0 and Q_0 subspaces, respectively. An effective Hamiltonian may then be defined as

$$\hat{\mathscr{H}}_{eff} = \hat{P}_0\hat{\mathscr{H}}\hat{P}_0, \tag{23}$$

which has the same eigenvalues as $\hat{H}$ within the P subspace.

This block diagonalisation of $\hat{H}$ provides a clear link to the BDD formalism. As the block diagonalisation transformation $\hat{T}$ is not uniquely defined, there exist a multitude of different effective Hamiltonian formalisms. As noted by Cederbaum *et al.*[22], however, the use of des Cloizeux's transformation[29]

$$\hat{T}_{dC} = \hat{P}\left(\hat{P}_0\hat{P}\hat{P}_0\right)^{-\frac{1}{2}} + \hat{Q}\left(\hat{Q}_0\hat{Q}\hat{Q}_0\right)^{-\frac{1}{2}} \tag{24}$$

is equivalent to the imposition of the least action principle of Equation 13. Thus, des Cloizeaux's effective Hamiltonian

$$\begin{aligned}\hat{\mathscr{H}}_{eff,dC} &= \hat{P}_0 \hat{T}_{dC}^{-1} \hat{H} \hat{T}_{dC} \hat{P}_0 \\ &= \left(\hat{P}_0 \hat{P} \hat{P}_0\right)^{-\frac{1}{2}} \hat{P} \hat{H} \hat{P} \left(\hat{P}_0 \hat{P} \hat{P}_0\right)^{-\frac{1}{2}},\end{aligned} \tag{25}$$

when constructed using a suitably chosen subspace partitioning, can be equated with a quasi-diabatic potential matrix that is equivalent to the transformation of the adiabatic potential using the BDD ADT of Equation 14. The main idea persued here is to directly compute the quasi-diabatic potential matrix using a perturbative approximation of des Cloizeaux's Hamiltonian in place of the calculation of the BDD transformation matrix. As shall be discussed below, although this is now an approximate approach, the associated errors incurred can be expected to be small for an implementation within the DFT/MRCI(2) framework. Moreover, this approximation leads to a resolution the overlap problem encountered in the P-BDD approach.

Through second-order in perturbation theory, des Cloizeaux's effective Hamiltonian reads

$$\begin{aligned}\left[\mathscr{H}_{eff,dC}^{[2]}\right]_{IJ} &= \langle \psi_I^{(0)} | \hat{H} | \psi_J^{(0)} \rangle \\ &+ \frac{1}{2} \sum_{\psi_K^{(0)} \in Q_0} \langle \psi_I^{(0)} | \hat{H} | \psi_K^{(0)} \rangle \langle \psi_K^{(0)} | \hat{H} | \psi_J^{(0)} \rangle \sum_{i=I,J} \left(E_i^{(0)} - E_K^{(0)}\right)^{-1}, \quad \psi_I^{(0)}, \psi_J^{(0)} \in P_0,\end{aligned} \tag{26}$$

where $|\psi_I^{(0)}\rangle$ and $E_I^{(0)}$ denote the eigenpairs of $\hat{H}_0$. Important to note here is that if the Epstein-Nesbet Hamiltonian partitioning is used to construct $\hat{H}_0$, and the DFT/MRCI Hamiltonian matrix elements are substituted for their *ab initio* counterparts, then the DFT/MRCI(2) working equations are obtained. If, however, a different choice of zeroth-order Hamiltonian is made, then the second-order effective Hamiltonian $\mathscr{H}_{eff,dC}^{[2]}$ may be used as an approximation of the quasi-diabatic potential matrix of interest. To arrive at such an effective Hamiltonian, we require a zeroth-order Hamiltonian $\hat{H}_0$ possessing a subset of eigenfunctions of qualitative similarity to the quasi-diabatic states of interest. Our choice of such a zeroth-order Hamiltonian reads as follows:

$$\hat{H}_0 = \sum_n |\tilde{\phi}_n\rangle \langle \tilde{\phi}_n | \hat{H} | \tilde{\phi}_n \rangle \langle \tilde{\phi}_n| + \sum_{IJ} |\overline{\Psi}_I^{(0)}\rangle \langle \overline{\Psi}_I^{(0)} | \hat{H} | \overline{\Psi}_J^{(0)} \rangle \langle \overline{\Psi}_J^{(0)}| + \sum_{\Omega \in \mathscr{F}'} |\Omega\rangle \langle \Omega | \hat{H} | \Omega \rangle \langle \Omega|. \tag{27}$$

Here, the $|\tilde{\phi}_n\rangle$ denote what we shall term 'prototype diabatic states', the $|\overline{\Psi}_I^{(0)}\rangle$ 'complement states', both of which shall be defined below. The $|\Omega\rangle$ are the FOIS CSFs that have survived the energy-based configuration selection. The prototype diabatic states are constructed as zeroth-order approximations to the quasi-diabatic states of interest and form our choice of P_0. They are constructed using the following procedure.

First, the normalised projections of a subset of $N+N_{buf}$ reference space eigenstates onto the quasi-diabatic states $|\phi_n(\boldsymbol{R}_{i-1})\rangle$ of the previous geometry are computed. Let these projections be denoted by $|\tilde{\phi}_n'(\boldsymbol{R}_i)\rangle$:

$$|\tilde{\phi}_n'(\boldsymbol{R}_i)\rangle = \frac{\sum_{I\in\mathscr{B}} |\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle\langle\Psi_I^{(0)}(\boldsymbol{R}_i)|\phi_n(\boldsymbol{R}_{i-1})\rangle}{\left|\left|\sum_{I\in\mathscr{B}} |\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle\langle\Psi_I^{(0)}(\boldsymbol{R}_i)|\phi_n(\boldsymbol{R}_{i-1})\rangle\right|\right|}, \quad n = 1,\ldots,N. \tag{28}$$

Here, the set basis states $\{|\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle | I \in \mathscr{B}\}$ is chosen as the N with greatest overlap with the quasi-diabatic states of the previous geometry plus a small number, N_{buf}, of ‘buffer’ states. These buffer states are simply taken as the lowest-energy reference states orthogonal to the N of greatest overlap. The non-orthogonal projections $|\phi_n'(\boldsymbol{R}_i)\rangle$ are then subjected to Löwdin’s symmetric orthogonalisation to yield the orthogonal ‘precursor’ states $|\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle$:

$$|\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle = \sum_{m=1}^{N} Y_{mn}^{-\frac{1}{2}} |\tilde{\phi}_m'(\boldsymbol{R}_i)\rangle, \tag{29}$$

where

$$Y_{mn} = \langle\tilde{\phi}_m'|\tilde{\phi}_n'\rangle. \tag{30}$$

The use of Löwdin’s symmetric orthogonalisation here is important as, amongst all possible orthogonalisations, it yields orthogonalised functions that are closest in the least squares sense to the original ones[30,31]. The precursor states form good zeroth-order approximations to the diabatic states of interest. Hence, the union of the precursor state basis with both its orthogonal complement within the reference space and the FOIS CSFs can be used to form a linearly independent basis with which to represent the zeroth-order Hamiltonian. However, as more reference than diabatic states are used in their construction, an additional step must be taken to construct their orthogonal complement within the reference space. To do so, the projections of the basis states $|\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle, I \in \mathscr{B}$ onto the orthogonal complement of the precursor states are computed:

$$\hat{Q}^{\#}|\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle = \left[1 - \sum_{n=1}^{N} |\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle\langle\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)|\right] |\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle. \tag{31}$$

The set of projected basis states $\{\hat{Q}^{\#}|\Psi_I^{(0)}(\boldsymbol{R}_i)\rangle\}$ is then subjected to canonical orthogonalisation followed by the discarding of the N null space vectors to yield the N_{buf} linearly-independent states $|\overline{\Psi}_I^{(0)\#}(\boldsymbol{R}_i)\rangle$. To do so, first the matrix

$$Z_{\nu(I)\nu(J)} = \langle\Psi_I^{(0)}(\boldsymbol{R}_i)|\hat{Q}^{\#}|\Psi_J^{(0)}(\boldsymbol{R}_i)\rangle, \quad I,J \in \mathscr{B} \tag{32}$$

is computed, where $\nu(I)$ is a map $\nu : \mathscr{B} \rightarrow \{1,2,\ldots,N+N_{buf}\}$. The matrix $\boldsymbol{Z}$ is then diagonalised:

$$\boldsymbol{Z} = \boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}^T, \quad \Lambda_{IJ} = \delta_{IJ}\lambda_J. \tag{33}$$

Let the eigenvalues λ_I be arranged in descending order. Then, the last N columns of the eigenvector matrix $\boldsymbol{A}$ correspond to the null space that we wish to discard. Let $\overline{\boldsymbol{A}}$ denote the truncation of $\boldsymbol{A}$ to the first N_{buf} columns. Linearly independent states $|\overline{\Psi}_I^{(0)\#}(\boldsymbol{R}_i)\rangle$ that are orthogonal to the precursor states may then be obtained as

$$|\overline{\Psi}_I^{(0)\#}(\boldsymbol{R}_i)\rangle = \sum_{J=1}^{N+N_{buf}} \overline{A}_{JI}|\Psi_{\nu^{-1}(I)}(\boldsymbol{R}_i)\rangle, \quad I = 1,\ldots,N_{buf} \tag{34}$$

At this point, we have a linearly independent basis with which we could represent the zeroth-order Hamiltonian, namely,

$$\{|\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle\} \cup \{|\overline{\Psi}_I^{(0)\#}\rangle\} \cup \{|\Omega\rangle | \Omega \in \mathscr{F}'\}.$$

However, it is preferable to first partially block diagonalise the Hamiltonian, decoupling the $\{|\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle\}$ and $\{|\overline{\Psi}_I^{(0)\#}\rangle\}$ blocks. This could be handled approximately in the subsequent effective Hamiltonian calculation, but it somewhat trivial to perform exactly in an additional step beforehand. To do so, we construct the representation of the Hamiltonian in the basis $\{|\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle\} \cup \{|\overline{\Psi}_I^{(0)\#}\rangle\}$. Let the Hamiltonian matrix projected onto this subspace be denoted by $\boldsymbol{h}$. A block diagonalising transformation, $\boldsymbol{b}$, of $\boldsymbol{h}$ is sought that: (i) decouples the $\{|\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle\}$ and $\{|\overline{\Psi}_I^{(0)\#}\rangle\}$ blocks, and; (ii) satisfies the least action principle $||\boldsymbol{b}-\boldsymbol{1}|| = \min$. The transformation that satisfies this reads as follows[22]:

$$\boldsymbol{b} = \boldsymbol{v}\boldsymbol{v}_{BD}^T\left(\boldsymbol{v}_{BD}\boldsymbol{v}_{BD}^T\right)^{-\frac{1}{2}}, \tag{35}$$

where $\boldsymbol{v}$ denotes the matrix of eigenvectors of $\boldsymbol{h}$, and $\boldsymbol{v}_{BD}$ is the portion of $\boldsymbol{v}$ corresponding to the desired block diagonal structure. The thus constructed matrix $\boldsymbol{b}$ defines the transformation from the coupled basis $\{|\tilde{\phi}_n^{\#}(\boldsymbol{R}_i)\rangle\} \cup \{|\overline{\Psi}_I^{(0)\#}\rangle\}$ to the decoupled basis $\{|\tilde{\phi}_n(\boldsymbol{R}_i)\rangle\} \cup \{|\overline{\Psi}_I^{(0)}\rangle\}$ of prototype diabatic states $|\tilde{\phi}_n(\boldsymbol{R}_i)\rangle$ and complement states $|\overline{\Psi}_I^{(0)}\rangle$. Along with the FOIS CSFs $|\Omega\rangle$, these are used to construct our zeroth-order Hamiltonian $\hat{H}_0$ (Equation 27).

Taking the P_0 space to be spanned by the prototype diabatic states, and inserting the DFT/MRCI Hamiltonian into Equation 26, we obtain the QD-DFT/MRCI(2) quasi-diabatic potential matrix, $\mathscr{W}_{DFT}^{[2]}$,

$$[\mathscr{W}_{DFT}^{[2]}]_{mn} = \langle \tilde{\phi}_m | \hat{H}_{DFT} | \tilde{\phi}_n \rangle + \frac{1}{2} \sum_{\Omega \in \mathscr{F}'} \langle \tilde{\phi}_m | \hat{H}_{DFT} | \Omega \rangle \langle \Omega | \hat{H}_{DFT} | \tilde{\phi}_n \rangle \sum_{k=m,n} \left(E_k^{(0)} - E_\Omega^{(0)} \right)^{-1}, \quad (36)$$

which may be considered as a second-order perturbative approximation of the DFT/MRCI(2) P-BDD potential. We note that this quasi-diabatic potential matrix corresponds to a symmetrized verson of that of Cimiralia *et al.*[12], but with a significantly improved zeroth-order Hamiltonian.

In addition to quasi-diabatic potentials, first-order corrected quasi-diabatic wave functions $|\phi_n^{[1]}\rangle$ may also be computed as

$$|\phi_n^{[1]}\rangle = |\tilde{\phi}_n\rangle + \sum_{\Omega \in \mathscr{F}'} \left(\frac{\langle \tilde{\phi}_n | \hat{H}_{DFT} | \Omega \rangle}{E_n^{(0)} - E_\Omega^{(0)}} \right) |\Omega\rangle, \quad (37)$$

Using the $|\phi_n^{[1]}\rangle$, quasi-diabatic properties and transition matrix elements may be computed. Furthermore, the quasi-diabatic states $|\phi_n^{[1]}\rangle$ computed at geometry $\boldsymbol{R}_{i-1}$ can be used in the construction of the precursor state basis $\{|\tilde{\phi}_n^{\#}\rangle\}$ at the neighboring geometry $\boldsymbol{R}_i$.

The most computationally expensive aspect of the evalution of the QD-DFT/MRCI(2) quasi-diabatc potential matrix $\mathscr{W}_{DFT}^{[2]}$ would the calculation of the matrix elements $\langle \tilde{\phi}_m | \hat{H}_{DFT} | \Omega \rangle$. However, these may be constructed via a recycling of the B-vectors $B_{\Omega I}$ used in the calculation of the DFT/MRCI(2) adiabatic states and potentials (see Equation 8), essentially eliminating the cost of this step altogether if both adiabatic and quasi-diabatic states are to be computed. The computational bottleneck then shifts to the calculation of the projections of the reference space eigenstates onto the quasi-diabatic states of the previous geometry (Equation 28). As in the P-BDD procedure, this requires the calculation of wave function overlaps involving non-orthogonal MOs. Taking N_{CSF} to be the total number of CSFs (assumed here to be equal for both bra and ket wave functions), the P-BDD procedure requires the consideration of N_{CSF}^2 non-orthogonal CSF pairs. This is reduced to $N_{CSF} \times N_{CSF}^{(0)}$ in the QD-DFT/MRCI(2) case, where $N_{CSF}^{(0)}$ is the number of reference space CSFs. In a DFT/MRCI(2) calculation, $N_{CSF}^{(0)} << N_{CSF}$ holds, with the ratio $N_{CSF}/N_{CSF}^{(0)}$

taking a value of up to 10^5-10^6 for very large systems. Thus, the overlap problem of the P-BDD procedure is ameliorated in the QD-DFT/MRCI(2) formalism.

Finally, we note that it can be expected that the QD-DFT/MRCI(2) approach will, in general, yield excellent approximations to DFT/MRCI(2) P-BDD quasi-diabatic potentials. The reasons for this are two-fold. Firstly, we note that, in the limit of $N_{buf} = 0$, the above-described procedure for the construction of the prototype diabatic states $|\tilde{\phi}_n\rangle$ reduces to a P-BDD calculation, but with the Hamiltonian replaced by its projection onto the reference space (see Appendix A of Reference 5). Secondly, the selected CI nature of a DFT/MRCI(2) calculation, combined with an iterative refinement of the reference space, results in reference space wave functions of very high quality. That is, the projections of these onto the DFT/MRCI(2) first-order corrected wave functions typically have large (near-unit) norms. Thus, the QD-DFT/MRCI(2) procedure can be seen as corresponding to the perturbative correction of the results of a P-BDD calculation performed with high-quality zeroth-order wave functions.

E. Enforcement of consistent phases

We end this section by noting that, in order to obtain continuous quasi-diabatic couplings and transition matrix elements, a consistent phase convention must be imposed on the QD-DFT/MRCI(2) first-order corrected quasi-diabatic wave functions. This is a consequence of the arbitrary phases of the reference space wave functions that enter into the QD-DFT/MRCI(2) working equations. Having wave function overlaps available, however, it is trivial to enforce a consistent phase. Namely, following the calculation of the first-order corrected quasi-diabatic wave functions $|\phi_n^{[1]}(\boldsymbol{R}_i)\rangle$ at geometry $\boldsymbol{R}_i$, one may compute overlaps with the quasi-diabatic wave functions $|\phi_n^{[1]}(\boldsymbol{R}_{i-1})\rangle$ of the previous geometry. With these available, the phase factors

$$\Gamma_n = \text{sgn}\left(\overset{max}{m}\{\langle \phi_n^{[1]}(\boldsymbol{R}_i) | \phi_m^{[1]}(\boldsymbol{R}_{i-1})\rangle\}\right) \tag{38}$$

are computed and $|\phi_n^{[1]}(\boldsymbol{R}_i)\rangle$ is multiplied by Γ_n. Here, to avoid ruinious computational costs, the required overlaps are computed using heavily truncated wave function expansions, using a norm-based truncation threshold of 0.95, which suffices for the calculation of the phase factors Γ_n but limits the wave function expansions to just the handful of dominant CSFs.

Furthermore, in order to obtain consistent wave function phases from calculation-to-calculation, the phases of the MOs entering into a QD-DFT/MRCI(2) calculation have to be fixed. To do so, we simply take the dominant atomic orbital (AO) coefficient for each MO to be positive. In the presence of non-Abelian point group symmetry and degenerate dominant AOs, the AO chosen can simply be taken as that generating the highest-indexed irrep in the highest Abelian sub-group according to some pre-defined list. This MO-phasing step is essential, e.g., for the calculation of multi-dimensional quasi-diabatic potential surfaces.

III. METHODOLOGY

A. Calculation of vibronic spectra

To assess the accuracy of the QD-DFT/MRCI(2) quasi-diabatic potentials, we considered their use in the construction of model Hamiltonians for the simulation of vibronic absorption spectra. We focused on two systems: (i) the first band in the absorption spectrum of furan, and; (ii) the first two bands in the absorption spectrum of chlorophyll a. For both molecules, it is known that strong vibronic coupling effects are manifest in their absorption spectra. The first system, furan, is small enough that both P-BDD and QD-DFT/MRCI(2) calculations are tractable, enabling a comparison of the quasi-diabatic potentials yielded by each. The size of the second system, chlorophyll a, is large enough to prohibit the application of the P-BDD methodology, while QD-DFT/MRCI(2) remains eminently tractable, highlighting the efficiency and scope of the approach.

Let $\sigma_I(E)$ denote the linear absorption spectrum corresponding to excitation from the ground state to the excited quasi-diabatic state $|\phi_I\rangle$. We adopt a time-depenedent formalism, in which $\sigma_I(E)$ may be computed from the Fourier transform of the wave packet autocorrelation function $a_I(t)$ corresponding to an initial wave packet $|\Psi(t=0)\rangle$ prepared *via* vertical excitation of the ground vibronic eigenstate to the electronic state $|\phi_I\rangle$:

$$\sigma_I(E) \propto E \int_{-\infty}^{\infty} a_I(t) \exp(iEt) dt, \tag{39}$$

with

$$a_I(t) = \langle \Psi(0) | \Psi(t) \rangle, \tag{40}$$

$$|\Psi(0)\rangle = \{ |\phi_I\rangle \langle \phi_1| + h.c. \} |\Psi_{GS}\rangle. \tag{41}$$

Here, $|\Psi_{GS}\rangle$ denotes the ground vibronic state, which was obtained *via* wave packet relaxation[32], and $|\phi_1\rangle$ is the ground quasi-diabatic electronic state.

All wave packet propagations were performed using the multi-layer multi-configurational time-dependent Hartree (ML-MCTDH) method[33–40]. For furan, all 21 vibrational degrees of freedom were included in the wave packet propagations. For chlorophyll a, 204 out of the total of 213 vibrational degrees of freedom where included, with those excluded corresponding to highly anharmonic, high-frequency modes that were found to be spectroscopically inactive. The full details of these calculation, including the ML-MCTDH wave function trees, are given in the Supplementary Information.

In order to perform the ML-MCTDH wave packet propagations, the total molecular Hamiltonian, $\hat{H}$, was represented in a basis of quasi-diabatic states $\{|\phi_I\rangle\}$,

$$\begin{aligned} \hat{H} &= \sum_{IJ} |\phi_I\rangle \langle\phi_I|\hat{H}|\phi_J\rangle \langle\phi_J| \\ &= \sum_{I} |\phi_I\rangle \hat{T}_n \langle\phi_I| + \sum_{IJ} |\phi_I\rangle W_{IJ}(\boldsymbol{R}) \langle\phi_J|, \end{aligned} \tag{42}$$

where $\hat{T}_n$ denotes the nuclear kinetic energy operator, and $\boldsymbol{W}(\boldsymbol{R})$ is the nuclear-coordinate-dependent quasi-diabatic potential matrix. To proceed, the latter must be cast into a form compatible with ML-MCTDH. Specifically, as a sum of products of monomodal operators[35]. This was achieved by approximating $\boldsymbol{W}(\boldsymbol{R})$ using the vibronic coupling Hamiltonian model of Köppel *et al.*[41]. In brief, $\boldsymbol{W}(\boldsymbol{R})$ is represented by a Taylor expansion in terms of mass- and frequency-scaled normal modes, Q_α, about the ground state minimum $\boldsymbol{Q}_0$. Our model potential is complete up to fourth-order in the one-mode terms, and reads as follows:

$$W_{IJ}(\boldsymbol{R}) \approx W_{IJ}^{mod}(\boldsymbol{Q}) = \tau_0^{(I,J)} + \sum_{p=1}^{4} \frac{1}{p!} \sum_{\alpha} \tau_{p\alpha}^{(I,J)} Q_\alpha^p. \tag{43}$$

For each molecule, expansion coefficients $\tau_{p\alpha}^{(I,J)}$ were computed using the normal equations approach detailed in Reference 7 using either DFT/MRCI(2) P-BDD or QD-DFT/MRCI(2) quasi-diabatic potential matrix elements as input. We note that the model potential of Equation 43 offers only approximations to the true DFT/MRCI(2) P-BDD or QD-DFT/MRCI(2) quasi-diabatic potentials. However, it does capture the leading order vibronic coupling effects and essential physics of the problem. Thus, if the these models furnish accurate vibronic absorption spectra then we may conclude that the quasi-diabatic potentials to which they are fitted are of good quality.

B. DFT/MRCI(2) and QD-DFT/MRCI(2) calculations

All DFT/MRCI(2) and QD-DFT/MRCI(2) calculations were performed using the General Reference Configuration Interaction (GRaCI) program[42]. In all calculations, the recently developed QE8 DFT/MRCI Hamiltonian[43] was used, which is parameterised for use with the QTP17 exchange-correlation functional[44]. In the furan and chlorophyll a calculations, the aug-cc-pVTZ and def2-SVP basis sets were used, respectively.

IV. RESULTS

A. Furan

We consider the absorption spectrum of furan over the interval 5.6 to 6.7 eV, which is shown for reference in Figure 1. In this region, the spectrum consists, broadly speaking, of: (i) a broad, relatively featureless background, and; (ii) several sharp, well defined peaks superimposed towards higher energy. The former is a hallmark of strong vibronic coupling effects, making furan an excellent test case for the application of QD-DFT/MRCI(2). The electronic states underlying this regions are the $A_2(\pi 3s)$, $B_2(\pi\pi^*)$, $A_1(\pi\pi^*)$, $B_1(\pi 3p_y)$, and $A_2(\pi 3p_z)$ states. From the work of Gromov *et al.*[45,46], it is known that this manifold of states forms a complex system of vibronic coupling, with strong direct coupling of the $B_2(\pi\pi^*)$ state to both the $A_2(\pi 3s)$ and $A_1(\pi\pi^*)$ states. Additionally, there exists an important indirect coupling of the $A_2(\pi 3s)$ and $B_2(\pi\pi^*)$ states via coupling to both the $A_1(\pi\pi^*)$ abd $B_1(\pi 3p_y)$ states. This non-trivial vibronic coupling system can thus be seen to constitute a suitable test of the ability of the QD-DFT/MRCI(2) to correctly describe non-adiabatic coupling.

We begin out analysis with a brief discussion of the vertical excitation energies yielded by the QD-DFT/MRCI(2) method at the ground state minumum geometry, $\boldsymbol{Q}_0$, at which the quasi-diabatic and adiabatic representations are taken to be equal. These values, as computed using the aug-cc-pVTZ basis, are shown in Table I. Shown alongside are the current theoretical best estimates, computed at the CCSDT/aug-cc-pVTZ level of theory and taken from the QUESTDB database[47,48]. Overall, the agreement with the theoretical best estimates is very good, with a mean absolute error of 0.06 eV and a maximum deviation of 0.11 eV.

We next consider the ability of the DFT/MRCI(2), P-BDD- and QD-DFT/MRCI(2)-parameterised model Hamiltonians to reproduce the vibronic absorption spectrum. The spectra obtained using

State	QD-DFT/MRCI(2)	CCSDT
$A_2(\pi 3s)$	6.14 (6.09)	6.09
$B_2(\pi\pi^*)$	6.48 (6.28)	6.37
$A_1(\pi\pi^*)$	6.48 (6.56)	6.56
$B_1(\pi 3p_y)$	6.66 (6.71)	6.64
$A_2(\pi 3p_z)$	6.87 (6.81)	6.81

TABLE I. Ground state minimum energy geometry vertical excitation energies computed using QD-DFT/MRCI(2) and CCSDT in conjunction with the aug-cc-pVTZ basis set. All values are given in units of eV. The CCSDT values are taken from Reference 48. The values in parentheses are the shifted vertical excitation energies found to best reproduce the experimental absorption spectrum

these models are shown in Figure 1. Shown alongside is the experimental spectrum of Holland *et al.*[49], obtained from the Mainz spectral database[50]. The simulated spectra were computed using the slightly-shifted vertical excitation energies given in parentheses in Table I, chosen so as to best reproduce the experimental spectrum. We note that all applied shifts are within the expected accuracy of the DFT/MRCI(2) method. Overall, the computed spectra are in excellent agreement with their experimental counterpart, with the broad background arising from vibronic coupling effects correctly reproduced by both Hamiltonians. In addition, the sharp peaks between around 6.5 and 6.7 eV, attributable to the $B_1(\pi 3p_y)$ state, are also correctly described. Importantly, the spectra computed using the two model Hamiltonians are barely distinguishable, implying that QD-DFT/MRCI(2) provides an excellent approximation to the DFT/MRCI(2) P-BDD diabatic potentials and couplings.

B. Chlorophyll a

Having established the accuracy of the QD-DFT/MRCI(2) diabatic potentials and couplings relative to the more formally rigorous, but more expensive, DFT/MRCI(2) P-BDD approach, we now consider the application of the QD-DFT/MRCI(2) method to a system whose size prohibits the application of the P-BDD approach. For this purpose, we consider the simulation of the first two bands in the absorption spectrum of cholorophyll a: the so-called Q and B bands. The relevant region of the spectrum is shown in Figure 2, adapted from the action spectroscopy measurements of Gruber *et al.*[51] Inset is the ground state minimum energy geometry of cholorophyll a. Across

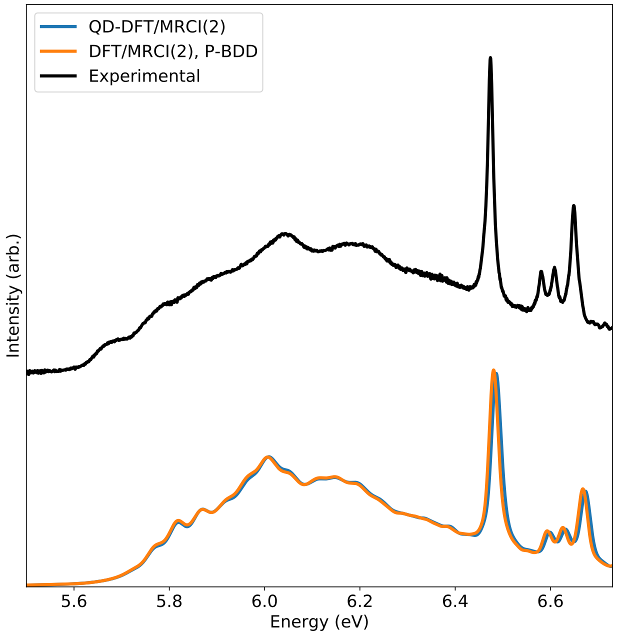

FIG. 1. Vibronic absorption spectra computed using vibronic coupling Hamiltonians parameterised via fitting to quasi-diabatic potentials computed at the DFT/MRCI(2), P-BDD and QD-DFT/MRCI(2) levels of theory using the aug-cc-pVTZ basis. Also shown is the experimental spectrum of Holland *et al.*[49], obtained from the Mainz spectral database[50].

this energy range (1.9 to 3.8 eV), there exist four ‘Gouterman-type’ states[52,53], corresponding to excitation amongst the HOMO-1, HOMO, LUMO and LUMO+1 π and π^* orbitals localised on the chlorin ring. In order of increasing vertical excitation energy, these are the Q_y, Q_x, B_x and B_y states. It is believed that the Q band contains contributions from vibronic states of mixed Q_x/Q_y character as a result of strong non-adiabatic coupling between the two[54], whereas the current interpretation of the B band assumes transitions to vibronic states of independent B_x and B_y character. We note, however, that there are some caveats to these interpretations. First, the current interpretation of the role of intensity borrowing in the Q band relied on the construction of model Hamiltonians parameterised by fitting to spectroscopic data and did not account for the potential interaction of the Q and B states[54]. Second, as far as we are aware, the role of intensity borrowing in the B band has not yet been assessed using high-level calculations. As a first step towards an assessment of these factors, we consider the a reduced dimensionality model obtained via the replacement of the phytyl chain with a methyl group. This model structure is shown in Figure 3. We note that this structual modification is known to have little effect on the energies of the Q and B states.

We begin with a discussion of the QD-DFT/MRCI(2) vertical excitation energies computed at the ground state minimum energy geometry, $\boldsymbol{Q}_0$. These are shown in Table II for the def2-SVP basis. Shown alongside are the near-benchmark quality DLPNO-STEOM-CCSD/def2-TZVP values of Sirohiwal *et al.*[55]. Also shown are the previous best estimates obtained by Sirohiwal *et al.* via the adjustment of the the DLPNO-STEOM-CCSD values to best match the experimental spectrum when used in a vibronic spectrum simulation. Overall, the QD-DFT/MRCI(2)/def2-SVP vertical excitation energies are in excellent agreement with best estimates, with a maximum deviation of just 0.08 eV. An important caveat here, however, is that the best estimate values here were extracted from spectrum simulations performed within the Born-Oppenheimer approximation, a point that we shall return to later.

Shown Figure 4 is the total simulated vibronic absorption spectrum alongside the experimental spectrum. The QD-DFT/MRCI(2)-parameterised model Hamiltonian is found to reproduce all main features of both the Q and B bands, providing validation of both the form of the model and the underlying QD-DFT/MRCI(2) calculations. We note that, in order to maximise agreement with the experimental spectrum, the vertical excitation energies entering into the model Hamiltonian were adjusted slightly. The values yielding maximal agreement are given in the final column of Table II and, we believe, constitute the current best estimates of the veritcal excitation energies of the Q and B states of chlorophyll a. Overall, our current best estimates are in good agreement with those of Sirohiwal *et al.*[55], with the exception of the vertical excitation energy of the B_x state, which we predict to lie 0.24 eV higher in energy. However, given the increased level of agreement with the experimental spectrum, we are lead to favour our vertical excitation energy estimates.

Finally, we briefly consider the role of vibronic coupling effects in both the Q and B bands. This is most clearly seen by re-computing the spectrum using a modified Hamiltonian in which all interstate coupling coefficients $\tau_{p\alpha}^{(I,J)}, I \neq J$ are set to zero. The result is shown in Figure 5. As is expected from the conclusions of previous work[54], the Q band is described incorrectly in the absence of non-adiabatic coupling, both in terms of the band origin and structure. Additionally, we find that the B band is also significantly affected when these coupling terms are ignored. In particular, the peak at around 3.5 eV is found to shift down in energy by around 0.2 eV and increase appreciably in intensity. We thus conclude that there exist hitherto unreported vibronic coupling effects that are operative within the B band of chlorophyll a. A full analysis of this is beyond the scope of the current work, but will be the focus of future work within our group.

State	QD-DFT/MRCI(2)	DLPNO-STEOM-CCSD	Previous best estimate	Current best estimate
Q_y	2.07	1.75	1.99	1.94
Q_x	2.37	2.24	2.30	2.21
B_x	3.18	3.17	3.12	3.36
B_y	3.36	3.40	3.38	3.40
CT	3.28	3.51	-	-

TABLE II. QD-DFT/MRCI(2)/def2-SVP vertical excitation energies (in units of eV) for the five excited states included in the ChlA model Hamiltonian. DLPNO-STEOM-CCSD and previous best estimate values taken from Reference 55. Experimental estimates are taken from Reference 55.

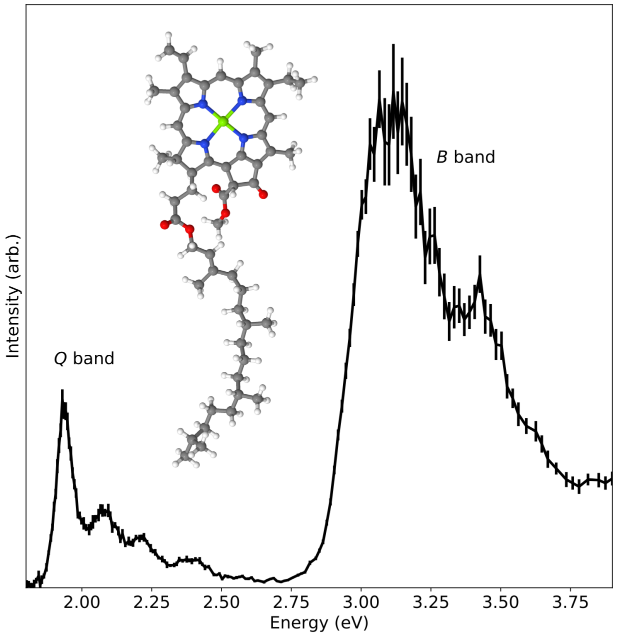

FIG. 2. Experimental action spectrum of cholorophyll a of Gruber *et al.*[51] across the energy range encompassing the *Q* and *B* bands.

V. CONCLUSIONS

To conclude, we have presented a new procedure for the calculation of diabatic potentials and couplings within the DFT/MRCI(2) framework. The method, termed QD-DFT/MRCI(2), exploits the intimate links between block diagonalisation, diabatisation and effective Hamiltonian theories to directly yield quasi-diabatic potential matrices and wave functions of high quality at low computational cost. Furthermore, the procedure is entirely black box in nature, and rendered

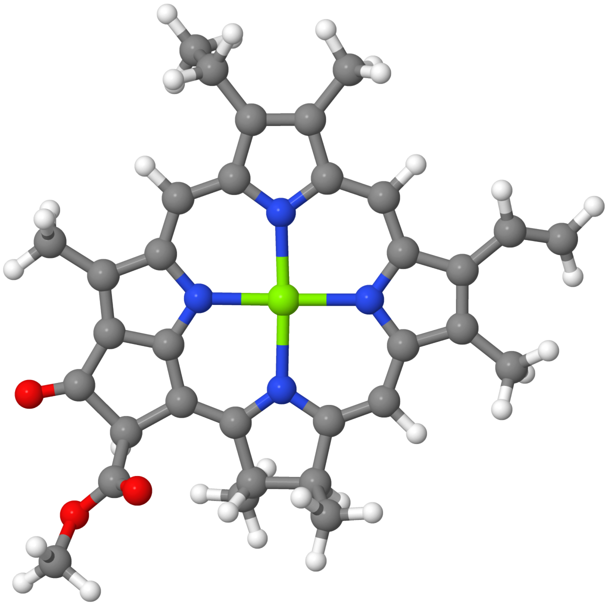

FIG. 3. Geometry of the chlorophyl a model used, corresponding to the replacement of the phytyl chain with a methyl group.

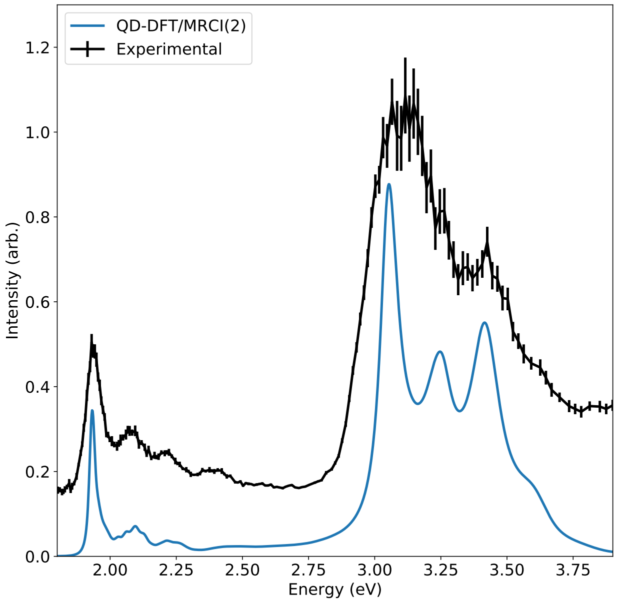

FIG. 4. Chlorophyll a absorption spectrum simulated using ML-MCTDH wave packet propagations employing a model Hamiltonian parameterised via fitting the QD-DFT/MRCI(2)/def2-SVP diabatic potentials and couplings.

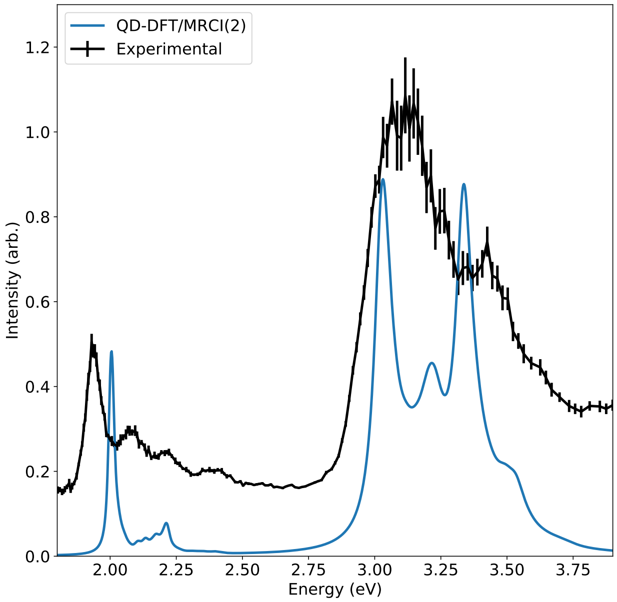

FIG. 5. Chlorophyll a absorption spectrum simulated using ML-MCTDH wave packet propagations employing a modified model Hamiltonian with all interstate coupling coefficients set to zero.

problem-free via the application of a simple yet robust root-following procedure; all that a user need input is: (i) the number of roots of interest, and; (ii) paths along which the quasi-diabatic potentials are to be computed.

To test the proposed method, absorption spectra for two prototypical vibronic coupling systems were computed: the first band in the absorption spectrum of furan, and the Q and B bands in chlorophyll a. For furan, the small system size permitted the application of both QD-DFT/MRCI(2) and the formally more rigorous P-BDD approach, to which QD-DFT/MRCI(2) can be viewed as an approximation. The vibronic spectra furnished by these two methods are essentially indistinguishable, and in excellent agreement with experimental spectrum, providing validation of the QD-DFT/MRCI(2) method. For chlorophyll a, the P-BDD approach becomes prohibitively expensive, whilst QD-DFT/MRCI(2) remains eminently tractable, demonstrating the reach of the method to large molecular systems. Moreover, the simulated spectrum is found to be in excellent agreement with experiment, allowing for the extraction of revised best estimates of the vertical excitation energies of the Q and B states. In addition, the results presented here suggest that there exist significant, previously unidentified vibronic coupling effects that are operative within the B band, challenging previous interpretations of this region of the spectrum.

ACKNOWLEDGMENTS

The authors thank Prof. Lars H. Andersen for providing experimental chlorophyll a absorption spectrum from Reference 51.

REFERENCES

[1] C. A. Mead and D. G. Truhlar, "Conditions for the definition of a strictly diabatic electronic basis for molecular systems," The Journal of Chemical Physics **77**, 6090–6098 (1982), https://pubs.aip.org/aip/jcp/article-pdf/77/12/6090/11149561/6090_1_online.pdf.

[2] Y. Shu, Z. Varga, S. Kanchanakungwankul, L. Zhang, and D. G. Truhlar, "Diabatic states of molecules," The Journal of Physical Chemistry A **126**, 992–1018 (2022), pMID: 35138102, https://doi.org/10.1021/acs.jpca.1c10583.

[3] S. Grimme and M. Waletzke, "A combination of kohn–sham density functional theory and multi-reference configuration interaction methods," J. Chem. Phys. **111**, 5645–5655 (1999), https://doi.org/10.1063/1.479866.

[4] C. M. Marian, A. Heil, and M. Kleinschmidt, "The dft/mrci method," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **9**, e1394 (2019), https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1394.

[5] S. P. Neville and M. S. Schuurman, "Removing the deadwood from dft/mrci wave functions: The p-dft/mrci method," Journal of Chemical Theory and Computation **17**, 7657–7665 (2021), pMID: 34861111, https://doi.org/10.1021/acs.jctc.1c00959.

[6] S. P. Neville and M. S. Schuurman, "A perturbative approximation to dft/mrci: Dft/mrci(2)," The Journal of Chemical Physics **157**, 164103 (2022), https://doi.org/10.1063/5.0118285.

[7] S. P. Neville, I. Seidu, and M. S. Schuurman, "Propagative block diagonalization diabatization of dft/mrci electronic states," J. Chem. Phys. **152**, 114110 (2020), https://doi.org/10.1063/1.5143126.

[8] T. Pacher, L. S. Cederbaum, and H. Köppel, "Approximately diabatic states from block diagonalization of the electronic hamiltonian," The Journal of Chemical Physics **89**, 7367–7381 (1988), https://doi.org/10.1063/1.455268.

[9] T. Pacher, C. A. Mead, L. S. Cederbaum, and H. Köppel, "Gauge theory and quasidiabatic states in molecular physics," The Journal of Chemical Physics **91**, 7057–7062 (1989), https://doi.org/10.1063/1.457323.

[10] T. Pacher, L. S. Cederbaum, and H. Köppel, "Adiabatic and quasidiabatic states in a gauge theoretical framework," in *Advances in Chemical Physics* (John Wiley & Sons, Ltd, 1993) pp. 293–391, https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470141427.ch4.

[11] F. Spiegelmann and J. P. Malrieu, "The use of effective hamiltonians for the treatment of avoided

crossings. ii. nearly diabatic potential curves," Journal of Physics B: Atomic and Molecular Physics **17**, 1259 (1984).

[12] R. Cimiraglia, J. P. Malrieu, M. Persico, and F. Spiegelmann, "Quasi-diabatic states and dynamical couplings from ab initio ci calculations: a new proposal," Journal of Physics B: Atomic and Molecular Physics **18**, 3073 (1985).

[13] R. Cimiraglia, "Adiabatic and diabatic basis sets in molecular calculations," in *Time-Dependent Quantum Molecular Dynamics*, edited by J. Broeckhove and L. Lathouwers (Springer US, Boston, MA, 1992) pp. 11–26.

[14] R. W. Wetmore and G. A. Segal, "Efficient generation of configuration interaction matrix elements," Chem. Phys. Lett. **36**, 478–483 (1975).

[15] G. A. Segal, R. W. Wetmore, and K. Wolf, "Efficient methods for configuration interaction calculations," Chem. Phys. **30**, 269–297 (1978).

[16] I. Lyskov, M. Kleinschmidt, and C. M. Marian, "Redesign of the dft/mrci hamiltonian," J. Chem. Phys. **144**, 034104 (2016), https://doi.org/10.1063/1.4940036.

[17] A. Heil and C. M. Marian, "Dft/mrci hamiltonian for odd and even numbers of electrons," J. Chem. Phys. **147**, 194104 (2017), https://doi.org/10.1063/1.5003246.

[18] A. Heil, M. Kleinschmidt, and C. M. Marian, "On the performance of dft/mrci hamiltonians for electronic excitations in transition metal complexes: The role of the damping function," J. Chem. Phys. **149**, 164106 (2018), https://doi.org/10.1063/1.5050476.

[19] P. S. Epstein, "The stark effect from the point of view of schroedinger's quantum theory," Phys. Rev. **28**, 695–710 (1926).

[20] R. K. Nesbet and D. R. Hartree, "Configuration interaction in orbital theories," Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **230**, 312–321 (1955), https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1955.0134.

[21] B. Kirtman, "Simultaneous calculation of several interacting electronic states by generalized van vleck perturbation theory," The Journal of Chemical Physics **75**, 798–808 (1981), https://doi.org/10.1063/1.442123.

[22] L. S. Cederbaum, J. Schirmer, and H. D. Meyer, "Block diagonalisation of hermitian matrices," Journal of Physics A: Mathematical and General **22**, 2427 (1989).

[23] F. Plasser, M. Ruckenbauer, S. Mai, M. Oppel, P. Marquetand, and L. González, "Efficient and flexible computation of many-electron wave function overlaps," Journal of Chemical Theory and Computation **12**, 1207–1219 (2016), pMID: 26854874,

https://doi.org/10.1021/acs.jctc.5b01148.

[24] D. J. Klein, “Degenerate perturbation theory,” The Journal of Chemical Physics **61**, 786–798 (1974), https://pubs.aip.org/aip/jcp/article-pdf/61/3/786/11544368/786_1_online.pdf.

[25] I. Shavitt and L. T. Redmon, “Quasidegenerate perturbation theories. a canonical van vleck formalism and its relationship to other approaches,” The Journal of Chemical Physics **73**, 5711–5717 (1980), https://doi.org/10.1063/1.440050.

[26] C. E. Soliverez, “General theory of effective hamiltonians,” Phys. Rev. A **24**, 4–9 (1981).

[27] P. Durand, “Direct determination of effective hamiltonians by wave-operator methods. i. general formalism,” Phys. Rev. A **28**, 3184–3192 (1983).

[28] V. Hurtubise and K. F. Freed, “The algebra of effective hamiltonians and operators: Exact operators,” in *Advances in Chemical Physics* (John Wiley Sons, Ltd, 1992) pp. 465–541, https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470141410.ch6.

[29] J. des Cloizeaux, “Extension d’une formule de lagrange à des problèmes de valeurs propres,” Nuclear Physics **20**, 321–346 (1960).

[30] B. C. Carlson and J. M. Keller, “Orthogonalization procedures and the localization of wannier functions,” Phys. Rev. **105**, 102–103 (1957).

[31] I. Mayer, “On löwdin’s method of symmetric orthogonalization*,” International Journal of Quantum Chemistry **90**, 63–65 (2002), https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.981.

[32] R. Kosloff and H. Tal-Ezer, “A direct relaxation method for calculating eigenfunctions and eigenvalues of the schrödinger equation on a grid,” Chemical Physics Letters **127**, 223–230 (1986).

[33] H.-D. Meyer, U. Manthe, and L. Cederbaum, “The multi-configurational time-dependent hartree approach,” Chemical Physics Letters **165**, 73–78 (1990).

[34] U. Manthe, H. Meyer, and L. S. Cederbaum, “Wave-packet dynamics within the multiconfiguration Hartree framework: General aspects and application to NOCl,” The Journal of Chemical Physics **97**, 3199–3213 (1992), https://pubs.aip.org/aip/jcp/article-pdf/97/5/3199/11323495/3199_1_online.pdf.

[35] M. Beck, A. Jäckle, G. Worth, and H.-D. Meyer, “The multiconfiguration time-dependent hartree (mctdh) method: a highly efficient algorithm for propagating wavepackets,” Physics Reports **324**, 1–105 (2000).

[36] H.-D. Meyer and G. A. Worth, “Quantum molecular dynamics: propagating wavepackets and density operators using the multiconfiguration time-dependent hartree method,” Theoretical

Chemistry Accounts **109**, 251–267 (2003).

[37]H. Wang and M. Thoss, "Multilayer formulation of the multiconfiguration time-dependent Hartree theory," The Journal of Chemical Physics **119**, 1289–1299 (2003), https://pubs.aip.org/aip/jcp/article-pdf/119/3/1289/10852931/1289_1_online.pdf.

[38]U. Manthe, "A multilayer multiconfigurational time-dependent Hartree approach for quantum dynamics on general potential energy surfaces," The Journal of Chemical Physics **128**, 164116 (2008), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.2902982/15410377/164116_1_online.pdf.

[39]U. Manthe, "Layered discrete variable representations and their application within the multiconfigurational time-dependent Hartree approach," The Journal of Chemical Physics **130**, 054109 (2009), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.3069655/15422973/054109_1_online.pdf.

[40]O. Vendrell and H.-D. Meyer, "Multilayer multiconfiguration time-dependent Hartree method: Implementation and applications to a Henon–Heiles Hamiltonian and to pyrazine," The Journal of Chemical Physics **134**, 044135 (2011), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.3535541/13789371/044135_1_online.pdf.

[41]H. Köppel, W. Domcke, and L. S. Cederbaum, "Multimode molecular dynamics beyond the born-oppenheimer approximation," in *Advances in Chemical Physics* (John Wiley Sons, Ltd, 1984) pp. 59–246, https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470142813.ch2.

[42]S. Neville and M. Schuurman, "GRaCI: General Reference Configuration Interaction," (2021).

[43]T. Costain, V. Ogden, S. Neville, and M. Schuurman, "A dft/mrci hamiltonian parameterized using only ab initio data: I. valence excited states," ChemRxiv.; doi:10.26434/chemrxiv-2024-clzns (2024).

[44]Y. Jin and R. J. Bartlett, "Accurate computation of X-ray absorption spectra with ionization potential optimized global hybrid functional," The Journal of Chemical Physics **149**, 064111 (2018), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.5038434/13501453/064111_1_online.pdf.

[45]E. V. Gromov, A. B. Trofimov, N. M. Vitkovskaya, J. Schirmer, and H. Köppel, "Theoretical study of the low-lying excited singlet states of furan," The Journal of Chemical Physics **119**, 737–753 (2003), https://pubs.aip.org/aip/jcp/article-pdf/119/2/737/10852652/737_1_online.pdf.

[46]E. V. Gromov, A. B. Trofimov, N. M. Vitkovskaya, H. Köppel, J. Schirmer, H.-D. Meyer, and L. S. Cederbaum, "Theoretical study of excitations in furan: Spectra and molecular dynamics,"

The Journal of Chemical Physics **121**, 4585–4598 (2004), https://pubs.aip.org/aip/jcp/article-pdf/121/10/4585/10859183/4585_1_online.pdf.

[47] P.-F. Loos, F. Lipparini, M. Boggio-Pasqua, A. Scemama, and D. Jacquemin, "A mountaineering strategy to excited states: Highly accurate energies and benchmarks for medium sized molecules," Journal of Chemical Theory and Computation **16**, 1711–1741 (2020), pMID: 31986042, https://doi.org/10.1021/acs.jctc.9b01216.

[48] M. Véril, A. Scemama, M. Caffarel, F. Lipparini, M. Boggio-Pasqua, D. Jacquemin, and P.-F. Loos, "Questdb: A database of highly accurate excitation energies for the electronic structure community," WIREs Computational Molecular Science **11**, e1517 (2021), https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1517.

[49] D. Holland, E. Seddon, A. Trofimov, E. Gromov, M. Wormit, A. Dreuw, T. Korona, N. de Oliveira, L. Archer, and D. Joyeux, "A study of the excited electronic states of normal and fully deuterated furan by photoabsorption spectroscopy and high-level ab initio calculations," Journal of Molecular Spectroscopy **315**, 184–195 (2015), spectroscopy with Synchrotron Radiation.

[50] H. Keller-Rudek, G. K. Moortgat, R. Sander, and R. Sörensen, "?" Earth Syst. Sci. Data **5**, 365 (2013).

[51] E. Gruber, C. Kjær, S. B. Nielsen, and L. H. Andersen, "Intrinsic photophysics of light-harvesting charge-tagged chlorophyll a and b pigments," Chemistry – A European Journal **25**, 9153–9158 (2019), https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/chem.201901786.

[52] M. Gouterman, "Spectra of porphyrins," Journal of Molecular Spectroscopy **6**, 138–163 (1961).

[53] A. Ceulemans, W. Oldenhof, C. Gorller-Walrand, and L. G. Vanquickenborne, "Gouterman's "four-orbital" model and the mcd spectra of high-symmetry metalloporphyrins," Journal of the American Chemical Society **108**, 1155–1163 (1986), https://doi.org/10.1021/ja00266a007.

[54] J. R. Reimers, Z.-L. Cai, R. Kobayashi, M. Rätsep, A. Freiberg, and E. Krausz, "Assignment of the q-bands of the chlorophylls: Coherence loss via qx qy mixing," Scientific Reports **3**, 2761 (2013).

[55] A. Sirohiwal, R. Berraud-Pache, F. Neese, R. Izsák, and D. A. Pantazis, "Accurate computation of the absorption spectrum of chlorophyll a with pair natural orbital coupled cluster methods," The Journal of Physical Chemistry B **124**, 8761–8771 (2020), pMID: 32930590, https://doi.org/10.1021/acs.jpcb.0c05761.