A DFT/MRCI Hamiltonian Parameterized Using Only *Ab Initio* Data: I. Valence Excited States

Teagan Shane Costain,[1] Victoria Ogden,[1] Simon P. Neville,[2, a)] and Michael S. Schuurman[1, 2, b)]
[1)]*Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Canada*
[2)]*National Research Council Canada, 100 Sussex Dr., Ottawa, Canada, K1A 0R6*

(Dated: 27 March 2024)

A new combined density functional theory and multi-reference configuration interaction (DFT/MRCI) Hamiltonian parameterized solely using benchmark *ab initio* electronic structure obtained from the QUEST databases is presented. This new formulation differs from all previous versions of the method in that the choice of the underlying exchange correlation (XC) functional employed to construct the one-particle (orbital) basis is considered, and a new XC functional, QTP17, is chosen for its ability to generate a balanced description of core and valence vertical excitation energies. The ability of the new DFT/MRCI Hamiltonian, termed QE8, to furnish accurate excitation energies is confirmed using benchmark quantum chemistry computations, and a mean absolute error of 0.16 eV is determined for the wide range of electronic excitations included the validation data set. In particular, the QE8 Hamiltonian dramatically improves the performance of DFT/MRCI for doubly-excited states. The performance of fast approximate DFT/MRCI methods, p-DFT/MRCI and DFT/MRCI(2), are also evaluated using the QE8 Hamiltonian and are found to yield excitation energies in quantitative agreement with the parent DFT/MRCI method, with the two methods exhibiting a mean difference of 0.01 eV with respect to DFT/MRCI over the entire benchmark set.

I. INTRODUCTION

The combined density functional theory and multi-reference configuration interaction (DFT/MRCI) electronic structure method,[1–5] in which compact MRCI expansions are supplemented with DFT-specific corrections, has a number of desirable properties. First, it is amenable to "black box" implementations, as the reference space is generated automatically and iteratively, rather than via a rigid active orbital subspace specified *a priori*. Secondly, it is able to furnish accurate vertical excitation energies at low computational cost through the use of DFT-specific Hamiltonian corrections to account for the preponderance of the dynamical electron correlation. Thirdly, it can describe a large range of electronic states, including those of multi-reference, Rydberg, and intramolecular charge-transfer character.

In part driven by this set of desirable, and comparatively unique, set of characteristics, the DFT/MRCI method has seen a recent increase in the pace of development over the past number of years. In particular, the original multiplicity-dependent formulation has been expanded to a spin-state agnostic parameterization,[3] extended to transition metal complexes,[4] and a new formulation was recently presented that focused on redressing previously observed errors in the description of doubly-excited electronic states.[6] Additionally, new methods based on the original DFT/MRCI approach, but with even more favorable computational scaling, have recently been described. This includes a "pruned" variant, p-DFT/MRCI,[7] and a perturbative approximation, DFT/MRCI(2),[8] in which the interaction of the reference and first-order interacting space (FOIS) is treated semi- and fully-perturbatively, respectively. Lastly, a core-valence separated DFT/MRCI method (CVS-DFT/MRCI) has been introduced[9] for the description of core-excited states and has been subsequently used to simulate both static and time-resolved X-ray spectra.[10]

The origin of the computational efficiency of DFT/MRCI lies in the use of short, truncated MRCI expansions to capture the static electronic correlation, in conjunction with DFT-specific Hamiltonian corrections aimed at the recovery of the remaining dynamic correlation. The DFT/MRCI Hamiltonian directly incorporates KS orbital energy differences, long-recognized to yield a good zeroth-order description of electronic transition energies, into the diagonal matrix elements, and thus effectively incorporates a significant amount dynamic correlation in a highly efficient manner. Accordingly, the off-diagonal elements of the Hamiltonian are damped to avoid a double-counting of dynamic correlation. Furthermore, the reference space is generated iteratively via a selected configuration interaction algorithm, thereby ensuring that the description of static correlation, captured by via the interaction of the reference configurations, is also highly computationally efficient, with only those configurations contributing significantly to the wave functions of interest being retained.

The semi-empirical parameters that arise in the working equations for DFT/MRCI Hamiltonians serve two primary purposes: *i*) to scale the Coulomb and exchange integral contributions to the diagonal Hamiltonian matrix elements, and *ii*) to damp the off-diagonal Hamiltonian matrix elements using a configuration energy-based criterion, as discussed above. The values of these parameters have historically been chosen so that DFT/MRCI vertical excitation energies reproduce the peak maxima in UV/VIS absorption spectra, although more recent parameterizations have employed mixed *ab initio* and experimental fitting data.[6] However, previous work has demonstrated that, in many cases, the positions of peak maxima in UV/VIS absorption spectra can exhibit significant deviations from electronic vertical excitation energies.[11–15] In par-

[a)]Electronic mail: Simon.Neville@nrc-cnrc.gc.ca
[b)]Electronic mail: Michael.Schuurman@uottawa.ca

ticular, structural relaxation, excited state anharmonicity, and vibronic coupling can all give rise to absorption band profiles for which the maximum of the band is shifted from the corresponding electronic vertical excitation energy. In short, the underlying transitions are *vibronic* in nature, not electronic. For organic choromophores, these effects generally conspire to give rise to positions of peak maxima that are red-shifted by around 0.1-0.3 eV in comparison to the corresponding electronic vertical excitation energies. As such, one can anticipate a significant underestimation of electronic vertical excitation energies by DFT/MRCI Hamiltonians whose fitting data sets contain such experimentally-derived excitation energy estimates.

In contrast to the DFT/MRCI Hamiltonian *ansatz* and parameterization, the choice of the underlying XC functional has not been the subject of significant study in the literature. In fact, all current DFT/MRCI implementations employ the same functional: BHLYP.[16] The large quantity of exact HF exchange (i.e. 50%) has been previously noted empirically to yield more accurate excitation energies relative to, for example, B3LYP.[17] However, to our knowledge, no systematic exploration of the choice of XC functionals has yet been performed.

Here we report a new DFT/MRCI formulation that differs from previous iterations in two fundamental characteristics: *i*) the values of the Hamiltonian parameters are optimized to reproduce benchmark *ab initio* estimates of electronic excitation energies only, and *ii*) the choice of XC functional is made to give a balanced description of both core and valence excitation and ionization energies. The motivation for the first property is to develop a method that reproduces as well as possible the eigenvalues of the *electronic* Hamiltonian; by incorporating experimental UV/VIS peak maxima into fitting sets, previous parameterizations can be viewed as having a bias to instead reproducing *vibronic* eigenvalues. As for the second property, the previously implemented CVS-DFT/MRCI method,[9] based on the R2017 Hamiltonian, has been observed to yield reasonably accurate core-excitation energies, albeit after applying large (multi-eV), edge-dependent shifts. Given that the KS orbital energy differences make up the leading contribution to the diagonal matrix elements of the DFT/MRCI Hamiltonian, we here explore whether the utilization of a different KS orbital basis will be able to furnish improved accuracy for vertical core-excitation energies, while retaining an accurate description of valence-excited states. This shall lead the way for the development of a DFT/MRCI Hamiltonian tailored to the accurate description of core-excited states, which will be the topic of a forthcoming publication.

We begin in Section II by providing a brief overview of the DFT/MRCI method, with an emphasis on the semi-empirical parameters that characterize the approach. This is followed by an articulation of the parameterization strategy, including the choice of *ab initio* data included in the fitting and validation sets. Section III describes technical details of the electronic structure computations and Section IV presents the results of the parameter optimization, including the choice of a new XC functional employed to generate the underlying KS orbital basis. We denote this new parameterization "QE8" and show incredibly favorable error metrics for the excitation energies of organic molecules. Section V showcases a set of representative applications of the QE8 parameterization including UV/VIS absorption spectra and vertical excitation energies for large molecular systems, which are reported alongside the results of a comparative benchmark. We summarize this work and include a prospectus for future refinements in Section VI.

II. THE DFT/MRCI METHOD

Let $\{|\Psi_I\rangle\}$ denote the set of eigenfunctions of the electronic Hamiltonian $\hat{H}$. The MRCI wave function ansatz may be written in the following general form:

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

Here, $|\Omega\rangle$ denote spin-adapted configuration state functions (CSFs). The total CSF space is partitioned into a reference space $\mathscr{R}$ and the first-order interacting space (FOIS) $\mathscr{F}$, where the latter is obtained by the application of one- and two-electron excitation operators to $\mathscr{R}$.

A. *Ab Initio* Diagonal Matrix Elements

Let each CSF $|\Omega\rangle$ be specified by a spatial occupation vector $\mathbf{w}$ and a spin-coupling pattern ω: $|\Omega\rangle = |\mathbf{w}\omega\rangle$. Where appropriate we shall switch between the $|\Omega\rangle$ and $|\mathbf{w}\omega\rangle$ notation, using the latter when the spatial occupations and/or spin-couplings enter explicitly into a given expression. The vector $\mathbf{w}$ encodes the occupations of each spatial orbital, while ω describes a particular spin-coupling amongst the singly-occupied orbitals, and can be thought of as specifying a valid walk through a genealogical spin coupling diagram. Following the work of Segal and Wetmore,[18,19] the diagonal matrix elements can be written in the following form in terms of their difference from a reference self-consistent field (SCF) energy E_{SCF}:

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

and

$$E_{SCF} = \sum_i F_{ii}\overline{\mathrm{w}}_i - \frac{1}{2}\sum_{ij}\left(V_{iijj} - \frac{1}{2}V_{ijji}\right)\overline{\mathrm{w}}_i\,\overline{\mathrm{w}}_j. \tag{3}$$

Here, $\Delta \mathrm{w}_i = \mathrm{w}_i - \overline{\mathrm{w}}_i$ denotes the difference of the occupation of the ith spatial orbital relative to a base, or anchor, occupation $\overline{\mathbf{w}}$, chosen here and in the following as the ground state

Hartree-Fock (HF) occupation. V_{ijkl} are standard two-electron integrals in the chemists' notation, S_{w} denotes the set of indices of singly-occupied orbitals in $\mathbf{w}$. The quantity F_{ii} is a diagonal Fock matrix element expressed in the chosen one-electron basis. Finally, η_{ij}^{ji} denotes a singlet spin-coupling coefficient in a notation that is explained in Appendix A.

B. *Ab Initio* Off-Diagonal Matrix Elements

The off-diagonal Hamiltonian matrix elements can be subdivided into three different classes: (i) those between bra and ket CSFs with the same spatial part but different spin-couplings, (ii) those with bra and ket spatial configurations linked by a single excitation, and (iii) those with bra and ket spatial configurations linked by a double excitation.

1. Same spatial configuration, different spin-coupling

The off-diagonal Hamiltonian matrix elements for bra and ket CSFs that differ by spin-coupling only are given by

$$\left\langle \mathbf{w}\omega' \middle| \hat{H} \middle| \mathbf{w}\omega \right\rangle = \frac{1}{2} \sum_{i \in S_{\mathrm{w}}} \sum_{j \in S_{\mathrm{w}}} V_{ijji} \eta_{ij}^{ji} \left(1 - \delta_{ij}\right) \tag{4}$$

2. Single excitations

For bra and ket CSFs linked by a single excitation from orbital i to orbital a, the off-diagonal Hamiltonian matrix elements read as follows:

$$\begin{aligned} \left\langle \mathbf{w}'\omega' \middle| \hat{H} \middle| \mathbf{w}\omega \right\rangle = & \left[F_{ia} + \sum_k \left(V_{iakk} - \frac{1}{2} V_{ikka} \right) \Delta \mathrm{w}_k \right] \eta_a^i \\ & + \sum_{\substack{k \in S_{\mathrm{w}} \\ k \neq i,a}} V_{ikka} \left[\eta_{ak}^{ki} + \left(\frac{1}{2} \mathrm{w}_k - 1 \right) \eta_a^i \right] \\ & + \left[V_{aaai} \mathrm{w}_a + V_{aiii} (\mathrm{w}_i - 2) \right] \eta_a^i \end{aligned} \tag{5}$$

3. Double excitations

For bra and ket CSFs linked by a double excitation from orbitals i and j to orbitals a and b, the off-diagonal Hamiltonian matrix elements may be written as

$$\left\langle \mathbf{w}'\omega' \middle| \hat{H} \middle| \mathbf{w}\omega \right\rangle = \left[V_{aibj} \eta_{ab}^{ij} + V_{ajbi} \eta_{ab}^{ji} \right] \left[(1 + \delta_{ab})(1 + \delta_{ij}) \right]^{-1} \tag{6}$$

C. DFT/MRCI Parameterization

For MRCI wave functions, the comparatively small reference space $\mathscr{R}$ is associated primarily with a description of static correlation. In contrast, the FOIS $\mathscr{F}$ primarily captures dynamic correlation energy, and is in general large in size. Due to the size of the FOIS, the diagonalization of the electronic Hamiltonian projected onto the space spanned by the total CSF basis, $\mathscr{R} \cup \mathscr{F}$, quickly becomes an extremely demanding computational task.

With reference to Eq. 2, the on-diagonal Hamiltonian matrix elements may be interpreted in a particle-hole picture with respect to excitations relative to the base configuration $\overline{\mathbf{w}}$. The leading contribution to each element is the sum of differences of on-diagonal Fock matrix elements between the excited (particle) and annihilated (hole) orbitals. Supplementing these zeroth-order terms are exchange and Coulomb interactions between the particle and hole orbitals as well as spin-couplings between the open shell orbitals. The key idea in DFT/MRCI is to replace the zeroth-order Fock-matrix element difference term with a quantity that more closely correlates with ground-to-excited-state excitation energies. To achieve this, a KS orbital basis is adopted and these terms are replaced with the corresponding differences between KS orbital *energies*, which more closely approximate vertical excitation energies. Then, as the zeroth-order description is now improved, the Coulomb and exchange contributions must be down-scaled. In this way, one may effectively incorporate a large amount of dynamic electron correlation into the on-diagonal Hamiltonian matrix elements that would otherwise be accounted for by the coupling of the reference and FOIS CSFs.

A number of different schemes for down-scaling the exchange and Coulomb contributions have been proposed since the original formulation by Grimme and Waletzke.[1] In general, however, the DFT/MRCI on-diagonal Hamiltonian matrix elements may be written as a correction to the *ab initio* values as follows:

$$\begin{aligned} \left\langle \mathbf{w}\omega \middle| \hat{H}^{DFT} - E_{DFT} \middle| \mathbf{w}\omega \right\rangle = & \left\langle \mathbf{w}\omega \middle| \hat{H} - E_{SCF} \middle| \mathbf{w}\omega \right\rangle \\ & + \sum_i (\varepsilon_{ii}^{KS} - F_{ii}) \Delta \mathrm{w}_i + \Delta E_x + \Delta E_c . \end{aligned} \tag{7}$$

where the ε_i^{KS} denote the Kohn-Sham (KS) orbital energies, and ΔE_c and ΔE_x are Coulomb and exchange corrections, respectively. The exact form of ΔE_c and ΔE_x has varied with the different DFT/MRCI parameterizations.[1–4] In this work, we adopt the correction terms of the recent parameterizations of Marian and co-workers,[3] in which ΔE_c and ΔE_x are expressed in the following multiplicity-agnostic form:

$$\Delta E_c = -p_C \left[\sum_{i<j} V_{iijj} \Delta \mathrm{w}_i \Delta \mathrm{w}_j + \frac{1}{2} \sum_{i \in S_{\overline{\mathbf{w}}}} V_{iiii} |\Delta \mathrm{w}_i| + \sum_i V_{iiii} \delta_{\Delta \mathrm{w}_i, 2} \right] + \frac{1}{4} \sum_{i \in S_{\overline{\mathbf{w}}}} V_{iiii}$$

$$\Delta E_x = p_X \left[- \sum_{\substack{i,j \\ j \in S_{\mathrm{w}}}} V_{ijij} |\Delta \mathrm{w}_i| + \frac{1}{2} \sum_{\substack{i \in C_{\mathbf{w}} \\ j \in A_{\mathbf{w}}}} V_{ijij} \Delta \mathrm{w}_i \Delta \mathrm{w}_j + \sum_{\substack{i<j \\ i,j \in S_{\mathrm{w}}}} V_{ijij} \eta_{ij}^{ji} \right] \tag{8}$$

where $S_{\overline{\mathbf{w}}}$ is the set of singly occupied orbital indices in the base configuration, $C_{\mathbf{w}}$ and $A_{\mathbf{w}}$ are the sets of indices of orbitals created and annihilated in $\mathbf{w}$ relative to the base configuration $\overline{\mathbf{w}}$, and $p_X, p_C \in [0,1]$ are scalar parameters.[3,4]

In an *ab inito* MRCI calculation, the coupling of the reference and FOIS CSFs accounts for the preponderance of the dynamical correlation. As this is now captured by the modified on-diagonal Hamiltonian matrix elements, as it stands, there exists a double counting of the dynamic correlation. To ameliorate this, an energy-dependent damping of the off-diagonal Hamiltonian matrix elements is applied. To achieve this, the spin-coupling-averaged energy difference between configurations $\mathbf{w}$ and $\mathbf{w}'$, $\Delta E_{\mathbf{ww}'}$, is introduced,

$$\Delta E_{\mathbf{ww}'} = \frac{1}{n_\omega} \sum_{\omega}^{n_\omega} H^{DFT}_{\mathbf{w}\omega,\mathbf{w}\omega} - \frac{1}{n_{\omega'}} \sum_{\omega'}^{n_{\omega'}} H^{DFT}_{\mathbf{w}'\omega',\mathbf{w}'\omega'} \tag{9}$$

where n_ω is the number of spin-couplings generated by the configuration $\mathbf{w}$. All off-diagonal matrix elements are subsequently scaled as a function of $\Delta E_{\mathbf{ww}'}$. For matrix elements between CSFs that have the same spatial occupation and differ only in the spin coupling, the corresponding matrix element is given by:

$$\left\langle \mathbf{w}\omega \middle| \hat{H}^{DFT} - E_{DFT} \middle| \mathbf{w}\omega' \right\rangle = (1 - p_X) \left\langle \mathbf{w}\omega \middle| \hat{H} - E_{SCF} \middle| \mathbf{w}\omega' \right\rangle, \tag{10}$$

while CSFs linked by a single excitation from orbital i to orbital a, the corrections read:

$$\left\langle \Omega \middle| \hat{H}^{DFT} - E_{DFT} \middle| \Omega' \right\rangle = D(\Delta E_{\mathbf{ww}'}) \left(\left\langle \Omega \middle| \hat{H} - E_{SCF} \middle| \Omega' \right\rangle - F_{ia} \right), \tag{11}$$

where, prior to the application of the damping function, the Fock matrix element F_{ia} is subtracted. The matrix elements between CSFs that differ by two electrons are given by:

$$\left\langle \Omega \middle| \hat{H}^{DFT} - E_{DFT} \middle| \Omega' \right\rangle = D(\Delta E_{\mathbf{ww}'}) \left\langle \Omega \middle| \hat{H} - E_{SCF} \middle| \Omega' \right\rangle, \tag{12}$$

In both Eqs. 11 and 12, $D(\Delta E)$ denotes a damping function that is chosen to decay rapidly with increasing ΔE: in practice either an exponential[1,4] or inverse arctangent[2,3] function. The damping of these matrix elements decouples a large part of the FOIS from the reference space. This, now-redundant, subset of FOIS configurations may then be removed from the calculation entirely. To do so, a simple orbital energy-based selection criterion is used,[1] which proceeds as follows. For each FOIS configuration $\mathbf{w}$, the quantity

$$d_{\mathbf{w}} = \sum_i \Delta \mathrm{w}_i \varepsilon_i^{KS} - \delta E_{sel} \tag{13}$$

is computed, where δE_{sel} is a free parameter. If $d_{\mathbf{w}}$ is less than the highest reference space eigenvalue of interest, then all the CSFs generated from the configuration $\mathbf{w}$ are selected for inclusion, else they are discarded. The value of δE_{sel} has previously been chosen to be either 1.0 or 0.8 E_h, with the latter corresponding to a more aggressive truncation of the FOIS, thereby leading to increased computational savings at the cost of a modestly reduced accuracy. This configuration selection step results in a massive reduction of the size of the CSF basis - typically by many orders of magnitude - and yields huge speed-ups relative to an *ab initio* MRCI calculation.

III. COMPUTATIONAL DETAILS

There are two distinct aspects to the optimization of the parameters that arise in the DFT/MRCI Hamiltonian: *i)* the calculation of vertical excitation energies for multiple classes of molecules and electronic states using a given set of parameters, and *ii)* the iterative variation of those parameters so as to minimize the error between the computed excitation energies and those of a pre-determined fitting data set, calculated using benchmark-level *ab initio* quantum chemistry methods. In a subsequent step, optimized parameter values are validated *via* the computation of the errors that result for a separate sub-set (the validation set) of benchmark excitation energies not included in the fitting set. We begin by describing the molecules and electronic states included in the fitting and validation sets, as well as describing the technical elements of the computations. This is followed by a brief discussion of the considerations involved in the optimization of the semi-empirical parameters that arise in the DFT/MRCI Hamiltonian.

The nuclear geometries and benchmark vertical excitation energies comprising the fitting and validation sets were obtained from the QUEST databases.[20–27] In all cases, geometries optimized at the aug-cc-pVTZ / CC3 level of theory were used. At the time of writing, the QUEST database is comprised of 7 datasets, grouped by the type of electronic excitation and/or molecule size. The fitting set employed here included vertical excitation energies from six of the data sets (DSs): 51 from small molecules (DS1), 9 from double excitations (DS2), 50 from medium molecules (DS3), 27 from exotic molecules and radicals (DS4), 24 from larger molecules (DS5), and 4 from charge transfer excitations (DS6), for a total of 93, 19, and 60 singlet, doublet, and triplet vertical excitation energies, respectively. A complete list of molecules and excitation energies used in the fitting and validation sets are given as Supplementary Information (SI).

Where available, near Full-CI (FCI) benchmark results were used, computed using the configuration interaction using a perturbative selection made iteratively (CIPSI) method.[28–31] For the remaining molecules, the QUEST Theoretical Best Estimate (TBE), computed using higher-order coupled-cluster theory methods, was employed as the benchmark value. The computation of the DFT/MRCI excitation energies was always performed using the same basis set as that employed in the benchmark computation; typically aug-cc-pVTZ[32,33] when comparing to the TBE results and aug-cc-pVDZ for the near-FCI values. All DFT/MRCI computations were performed using the GRaCI software package.[34] Supplementary calculations using the equation of motion coupled cluster singles and doubles (EOM-CCSD) method were performed using the using QChem5.4[35].

In all DFT/MRCI calculations, the difference dedicated CI (DDCI)[36] configuration reduction procedure was used. Here, the size of the FOIS is reduced via the exclusion of configurations that do not contribute through second-order in perturbation theory to transition energies. Consistent with previous implementations, the number of configurations in the CSF basis was further reduced by constraining the number of open shells to be at most 10, as well as limiting the maximum excitation degree relative to the base configuration to 8. The initial guess reference spaces were generated using a previously described automatic selection procedure[7] that we term AutoRAS. In short, a preliminary DFT/CIS[37] computation is used to identify the subset of orbitals necessary to construct the leading configurations for the states of interest. The initial guess reference space is then taken as a restricted active space CI (RASCI) one,[38] with a maximum excitation level of two and only the RAS1 and RAS3 spaces occupied. Finally, we note that all p-DFT/MRCI calculations were performed using a pruning threshold α_p of 0.90 (see Reference 7 for a definition of this quantity).

The optimization of the semi-empirical parameters in the DFT/MRCI Hamiltonian was performed using a dedicated module in GRaCI[34] that employs a Nelder-Mead non-linear optimization algorithm[39,40] as implemented in SciPy.[41] At each iteration of the optimization procedure, those electronic transitions corresponding to the fitting data were identified via the computation of wave functions overlaps between a large number of roots of the DFT/MRCI Hamiltonian and previously determined reference wave functions of the correct electronic character.

IV. RESULTS

A. Selection of XC Functional

The primary motivation of the present work is the determination of a DFT/MRCI parameterization that reproduces benchmark *ab initio* vertical excitation energies. However, a secondary consideration is to develop a class of methods that can provide a consistent description of both valence and core excitation processes. While a previous implementation of the CVS approximation applied to the valence-parameterized R2017 DFT/MRCI Hamiltonian[9] generally produced good *relative* core-excitation energies, the absolute errors were found to be quite large: around 5 eV for K-edge transitions. The main source of error is found to be attributable to the previous choice of XC functional, BHLYP[16]: while the BHLYP KS orbital-energy differences yield good zeroth-order approximations of valence excitation energies, there is a significant degradation when considering core excitations.

This motivates investigating the use of other XC functionals to generate the KS orbitals and orbital energies that enter into the DFT-specific corrections of the DFT/MRCI Hamiltonian. In lieu of performing numerous computationally expensive parameter optimizations for each XC functional, the expected performance of each XC functional was evaluated via a comparison of the errors for valence and core excitation energies using a related method. Specifically, at the Tamm-Dancoff approximation time-dependent DFT (TDA-TDDFT)[42] level of theory. The use of this surrogate method was motivated by the fact that the leading contribution to the diagonal matrix elements of the corresponding matrix eigenvalue problem within TDA-TDDFT is that of the KS orbital energy differences, analogous to the form of the diagonal matrix elements of the DFT/MRCI Hamiltonian (Eq. 7). While the similarities here are more evocative than quantitative, the following results are sufficiently useful to identify candidate XC functionals.

To evaluate various XC functionals' simultaneous ability to describe core and valence excitation energies, we consider the mean absolute error (MAE) of core and valence excitation energies for a multitude of functionals. The reference valence excitation data were taken from the QUEST databases and the core excitation energies were computed at the aug-cc-pCVTZ / CCSDT level of theory. These MAEs are shown in Figure 1a), and are also tabulated in Table S6 of the SI.

As Figure 1a) shows, the errors in the core-excitation energies can be as large as 20 eV (for the functionals shown), while all of the XC functionals reproduce the lowest valence excitations to better than 1 eV. Note that the TDA time-dependent Hartree Fock (i.e., CIS) results are included (labeled as "HF" in Figure 1a)) for reference and yield MAE errors in the valence and core-excitation energies of 0.7 eV and 12.6 eV, respectively. XC functionals that will be potentially useful in the development of new DFT/MRCI parameterizations, namely those that provide a consistent (and accurate) description of core and valence excitation energies, are in the lower left quadrant Figure 1a). This region of the plot is populated by hybrid functionals: hybrid generalized gradient approximation (GGA) and meta-GGA functionals. Shown plotted in Figure 1b) are the mean errors (MEs) for the subset of XC functionals with the smallest MAEs. In general, the same XC functionals perform best using either error metric.

Firstly, we note that BHLYP, the up-to-now de facto standard XC functional for the DFT/MRCI approach, exhibits MAEs of 0.28 and 2.6 eV for valence and core excitations, respectively. These values correspond well with the previously observed accuracies for BHLYP-based DFT/MRCI Hamiltonians for valence and core excitation energies.[3,9] This provides a post-hoc sanity check of the indirect XC functional

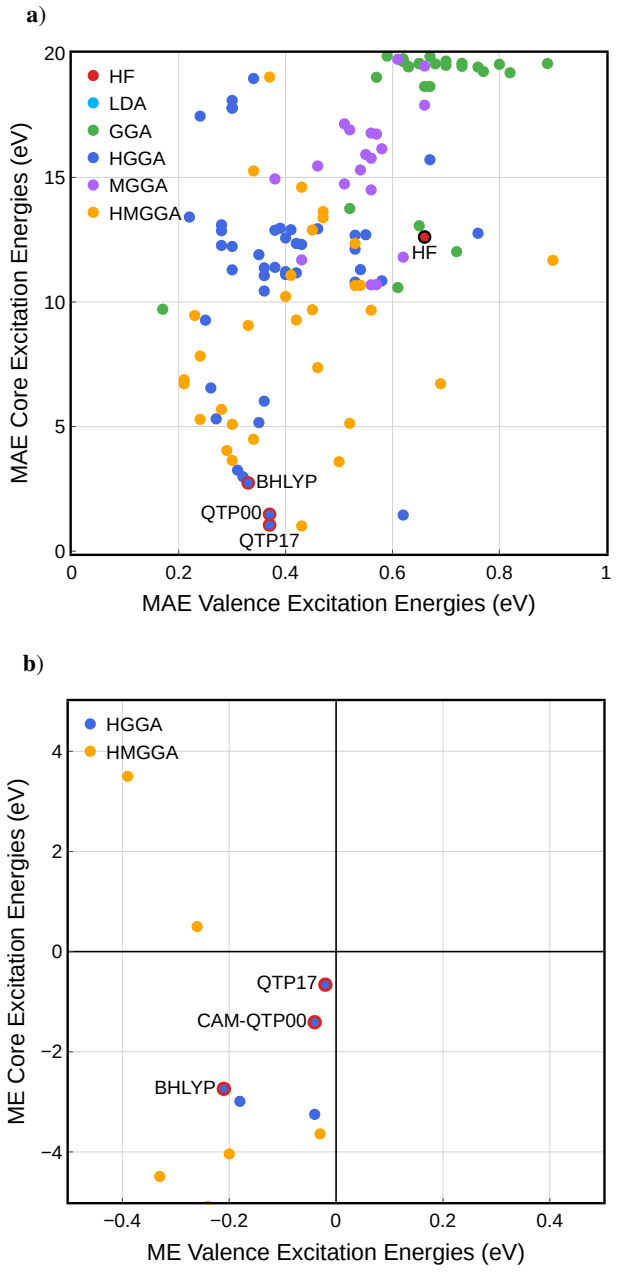

FIG. 1. Mean absolute errors (MAEs) for core and valence excitation energies for a variety of functionals computed using TDA/TDDFT are given in panel **a**. Panel **b** presents a subset of functionals exhibiting the smallest mean errors (MEs) for valence and core excitation energies. The numerical values for the data in this figure is given in Table S6 of the supplementary material.

assessment procedure used here. While the authors are not aware of any explicit analysis of this type in the previous literature, the performance of BHLYP with respect to these metrics is impressive and is validated by the accuracy of the results furnished by previous parameterizations that have been based on this XC functional. However, it is clear that a more balanced description of core and valence excited states is provided by other functionals.

Curiously, the XC functionals dlDF[43] and WC04[44] also exhibit relatively balanced errors with respect to valence and core-excitation. The former is a functional optimized to determine intermolecular interaction energies via the calculation of a separate dispersion correction added to the (dispersionless) dlDF result. The latter is optimized to reproduce ^{13}C NMR chemical shifts. Our desire to employ a general XC functional means that we will not consider these functionals further for the present purposes, but the role of "non-standard" functionals for the generation of orbital bases may warrant subsequent study.

The the best performing XC functionals based on this analysis are QTP17[45] and CAM-QTP00.[46] The former is found to perform the best, yielding MAEs of 0.36 and 0.62 eV for valence- and core-excitations, respectively. These functionals are designed to obey the "IP-eigenvalue theorem",[47] which states that each of the KS-orbital energies (i.e. Koopmans' ionization potentials) should make an accurate approximation to the corresponding vertical IP. Within this ansatz, the IP Theorem[48] interpretation of KS DFT is extended to a "complete" one. To this end, both functionals were optimized to reproduce each of the vertical ionization potentials of the water molecule.[45] The excellent performance of this functional for core-excitation has been previously noted.[49]

Given that QTP17 performs best under the current analysis, as well as our desire to keep the method as computationally efficient as is possible by avoiding the use of range-separated functionals, we select this XC functional to form the basis of our DFT/MRCI Hamiltonian parameterization. We note that even though BHLYP actually exhibits a smaller MAE for valence excitation energies than QTP17, the QTP17 description of valence and core-excitation energies is more *balanced*. Additionally, the ME plot in Figure 1(b) indicates smaller systematic errors in the QTP17 results than the BHLYP ones.

B. New Parameterization: The QE8 Hamiltonian

1. *Form of the Damping Function*

As discussed in Sec. II, the DFT/MRCI Hamiltonian employs a semi-empirical damping of the off-diagonal Hamiltonian matrix elements (see Eqs. 11 and 12). The form of this function has yet to be defined. In the current work, a flexible three-parameter exponential damping function was adopted, defined as follows:

$$D(\Delta E_{\mathbf{ww}'}) = d_1 \exp\left(-d_2 \Delta E_{\mathbf{ww}'}^{d_3}\right) \tag{14}$$

We found that the quality of the Hamiltonian parameterization was not strongly dependent on the functional form of the damping function, beyond the requirement that it decay rapidly as a function of $\Delta E_{\mathbf{ww}'}$, and the current form was chosen for its generality. Specifically, the parameter d_1 determines the value of the damping function in the limit of zero spin-coupling-averaged energy differences between

CSFs, while the values of d_2 and d_3 (and the exponential functional form) ensures that the both the value of $\Delta E_{\mathbf{ww}'}$ at which the damping function begins to significantly decrease, and the rapidity with which the damping function tends to zero, can be independently varied. For formal simplicity, and to ensure well-behaved non-linear optimizations, we have chosen to constrain the parameter d_3 to positive integer values, i.e., $d_3 \in \mathbb{N}$.

Lastly, we have enforced an additional constraint in the present parameterization as a consequence of the iterative generation of the reference space. Specifically, a selected CI algorithm is used that tailors the reference space $\mathscr{R}$ to provide good support for all the reference space states of interest. A consequence of this is that the dimension of the reference space Hamiltonian will depend on the number of roots, N_{root}. As N_{root} increases, the span of the reference space eigenspectrum will correspondingly broaden. Thus, the damping function, which is a function of $\Delta E_{\mathbf{ww}'}$, will exhibit a different effective scaling for different roots of the DFT/MRCI Hamiltonian. We have this observed this to manifest as a moderate dependence of the DFT/MRCI state energies on N_{root}, leading to changes in excitation energies of up to 0.2 eV depending on the number of roots computed. This is particularly a problem if a large number ($\gtrsim 20$) of roots are sought. To ameliorate this effect, we have chosen to constrain d_2 such that $D(\Delta E_{\mathbf{ww}'}) \leq 0.01$ for $\Delta E_{\mathbf{ww}'} = 1.0$ $\mathrm{E_h}$ which corresponds to $d_2 \geq -\ln(0.01) \approx 4.605$. We found this choice to strike a balance between maintaining the accuracy of the VEEs, while mitigating the N_{root} dependence of the excitation energies.

TABLE I. Optimized QE8 Hamiltonian parameter values and MAEs (in eV) for both the fitting and validation sets. Values of the configuration selection parameter δE_{sel} are given in units of E_h. For comparison, the parameters of, and MAEs yielded by the R2017 Hamiltonian are shown alongside. The MAEs are split into subsets for the different spin multiplicities present in the fitting and validation sets, as well as the total across all multiplicities.

Hamiltonian	δE_{sel}	p_C	p_X	d_1	d_2	d_n
QE8	1.0	0.426	0.255	0.690	4.61	8
	0.8	0.419	0.258	0.712	4.69	8
	δE_{sel}	p_C	p_X	p_1 [a]	p_2	
R2017	1.0	0.503	0.359	0.564	1.857	
	0.8	0.501	0.357	0.574	1.927	

Fitting Data MAEs

Hamiltonian	δE_{sel}	Singlet	Triplet	Doublet	Total
QE8	1.0	0.18	0.14	0.17	0.16
	0.8	0.22	0.15	0.20	0.20
R2017	1.0	0.37	0.29	0.28	0.33
	0.8	0.37	0.30	0.28	0.34

Validation Data MAEs

Hamiltonian	δE_{sel}	Singlet	Triplet	Doublet	Total
QE8	1.0	0.15	0.15	0.22	0.16
	0.8	0.21	0.18	0.23	0.20
R2017	1.0	0.27	0.25	0.38	0.27
	0.8	0.28	0.26	0.40	0.28

[b] A different form of the damping function is employed in R2017 such that the parameters (d_1,d_2,d_3) and (p_1,p_2) are directly comparable.

2. *Analysis of Fitted Parameters*

In the following analysis we will use the R2017 Hamiltonian[3] as the primary DFT/MRCI method of comparison. The parameters in that Hamiltonian were optimized employing experimental ultraviolet/visible (UV/VIS) absorption spectra peak maxima, in clear contrast to the current approach. Furthermore, the fit set used in that work was comprised exclusively of organic chromophores, which is also the emphasis of the present work. Subsequent DFT/MRCI parameterizations focused on an improvement of the description of transition metal complexes[4] and doubly-excited states[6]. More specific comparisons for these classes of excitations will be the subject of future work.

In the following, we denote the DFT/MRCI Hamiltonian with the form given in Section II C, and the damping function given in Eq. 14 as QE*n*. This choice is meant to signify a parameterization obtained by fitting to the QUEST databases and a Hamiltonian employing an exponential decay function in which the exponential argument is raised to the power $d_3 = n$. For reasons detailed below, the final choice of the value of d_3 was 8. As such, our final DFT/MRCI Hamiltonian is denoted as QE8.

Table I presents a summary of the optimized QE8 DFT/MRCI Hamiltonian parameter values, as well as the corresponding MAEs relative to the benchmark *ab initio* excitation energies of the fitting set. We include parameter sets for configuration selection parameter values δE_{sel} of 1.0 and 0.8 E_h. For comparison, we also show the same results for the R2017 DFT/MRCI Hamiltonian.[3] In the subsequent discussion, we will refer exclusively to the $\delta E_{sel} = 1.0$ $\mathrm{E_h}$ results, but the slightly less accurate $\delta E_{sel} = 0.8$ $\mathrm{E_h}$ results are shown for consistency with previous parameterizations.[1,3,4,6]

We begin with a comment on the optimized value of the Coulomb integral scaling term, p_C. As previously noted, this parameter primarily scales the integrals corresponding to the hole-particle Coulomb interaction. In the TDDFT formalism, integrals of this form are explicitly scaled by $1 - x_{HF}$, where x_{HF} denotes the proportion of exact exchange present in the XC functional. A broadly similar relation can be expected to be followed by the parameter p_c, in line with previous studies.[1,3] Indeed, we find an optimized value of $p_C = 0.426$, which is to be compared with the value of $1 - x_{HF} = 0.38$ for the QTP17 XC functional.

In the Hamiltonian optimization procedure, the damping function parameters d_1 and d_2 were optimized simultaneously using the Nelder-Mead optimization algorithm, while the exponent d_3 was scanned over integer values between 4 and 12. The resulting optimized QE8 damping functions, as well as the damping function of the R2017 Hamiltonian, are shown plotted in Figure 2. We also include for comparison the damping functions that result from setting $d_3 = 6$ and 10, which corresponds to the Hamiltonians QE6 and QE10, respectively. It is immediately apparent that the new QE*n* parameterizations incorporate contributions from the off-diagonal matrix

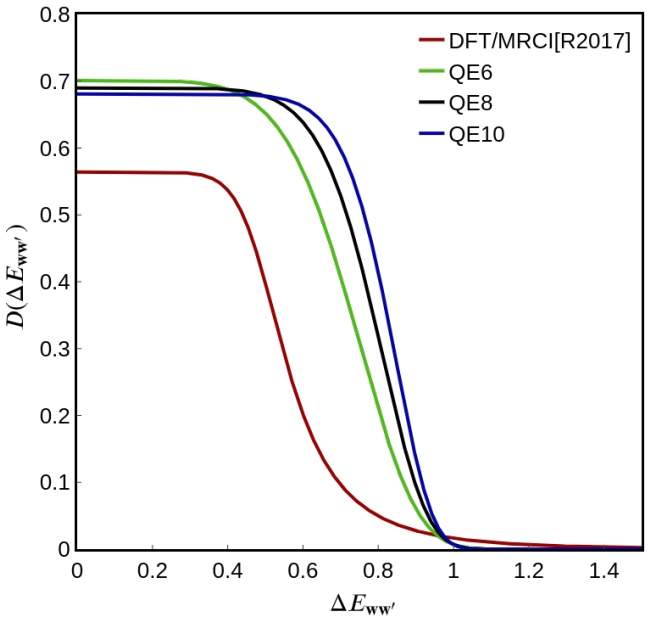

FIG. 2. The damping function applied to the off-diagonal matrix elememnts of the DFT/MRCI Hamiltonian for current the (QE8) and previous (R2017) parameterizations.

elements of the DFT/MRCI Hamiltonian at much higher level than the previous R2017 formulation. The limiting value of the damping function as $\Delta E_{ww'} \to 0$, as well as the value of $\Delta E_{ww'}$ at which the damping function begins to decrease significantly are markedly larger than that for the R2017 damping function. Additionally, as d_3 is increased, the onset of the damping occurs correspondingly later. While the QE8 ($d_3 = 8$) error metrics were significantly better than QE4 ($d_3 = 4$) and QE6 ($d_3 = 6$), the QE10 ($d_3 = 10$) parameterization resulted in no appreciable improvement (ΔMAE ≤ 0.01 eV) in the error metrics. Hence, the value of $d_3 = 8$ was adopted for the final Hamiltonian parameterization.

As shown in Table I, the QE8 Hamiltonian is able to achieve highly impressive agreement with the theoretical best estimates of VEEs. The fitting set is comprised primarily of near-FCI VEEs for small molecules and CC3 values for larger molecules. For these values, not only was the MAE 0.16 eV, but it was largely independent of multiplicity. This value is approximately half the error of 0.33 eV yielded by the R2017 Hamiltonian.

A notable by-product of employing a different XC functional to generate the KS orbital basis is the significant decrease in the size of the FOIS for the QTP17-based QE8 parameterization in comparison to the BHLYP-based R2017 Hamiltonian. The QE8 Hamiltonian resulted in CSF bases that were roughly half as large as the corresponding R2017 expansion, with an average ratio of 0.53 being found for the molecules contained in the fitting set. Furthermore, this ratio is found to decrease with increasing molecular size; the QE8 CSF expansions become *more* compact relative to the R2017

TABLE II. Errors of the QE8 and R2017 DFT/MRCI Hamiltonians for doubly-excited states relative to theoretical best estimates (TBEs). All values are given in units of eV.

			Error	
Molecule	Transition	TBE	R2017	QE8
butadiene	$1^1A_g \to 2^1A_g$	6.51	-0.26	0.02
ethylene	$1^1A_g \to 2^1A_g$	13.07	-0.12	0.39
glyoxal	$1^1A_g \to 2^1A_g$	5.48	-0.53	0.30
formaldehyde	$1^1A_1 \to 3^1A_1$	10.45	-1.24	-0.69
nitrosomethane	$1^1A' \to 2^1A'$	4.84	-0.66	-0.07
nitroxyl	$1^1A' \to 2^1A'$	4.40	-0.86	-0.04

equivalent as the dimension of the orbital basis increases. This can be attributed to the lower occupied valence orbital energies furnished by the QTP17 functional, in concert with the orbital energy-based selection criterion given in Eq. 13, yielding a smaller effective orbital basis from which to generate FOIS configurations.

While the fitting set was comprised of electronic excitations involving states of varied electronic character (e.g., optically allowed single excitations, Rydberg, intramolecular charge-transfer, etc.), we emphasize here that VEEs involving doubly-excited electronic character were also included. A recent work[6] has recently attempted to improve the performance of DFT/MRCI with respect to this historically problematic class of transitions via a new parameterization that introduces configuration-specific hole-particle and two-hole-two-particle Coulomb and exchange scaling parameters. Interestingly, the present results show that a comparable improvement in performance can be achieved without the introduction of additional semi-empirical parameters. Table IV B 2 summarizes the performance of the QE8 and R2017 Hamiltonians for the doubly-excited states included in the fitting set. The MAE furnished by QE8 for these excitations is only 0.26 eV, compared to 0.61 eV for R2017. The largest error by a significant margin is for the 3^1A_1 state of formaldehyde (0.69 eV) for which the TBE of 10.34 eV differs dramatically from the CCSDT (10.79 eV) and CC3 (11.20 eV) results; a challenging excitation to converge. Conversely, the $\pi^2 \to \pi^{*2}$ transition in butadiene is reproduced to 0.02 eV.

3. *Performance on Validation Data*

The validation set, again culled from transitions across all 6 QUEST databases (Table IV B 2), is comprised of 165 singlet, 102 triplet, and 14 doublet vertical excitation energies. Impressively, the MAE across all excitation types is identical to that for the fitting set: 0.16 eV. Figure 3 breaks down the MAEs for the singlet, triplet and doublet excitation classes for both the fitting and validation sets. As the Figure evinces, the errors for the QE8 Hamiltonian are consistent across all excitation classes (< 0.2 eV) and roughly half the magnitude of the errors for the R2017 parameterization.

Further insight is gained from the histograms of errors for the training and validation sets for both the QE8 and R2017

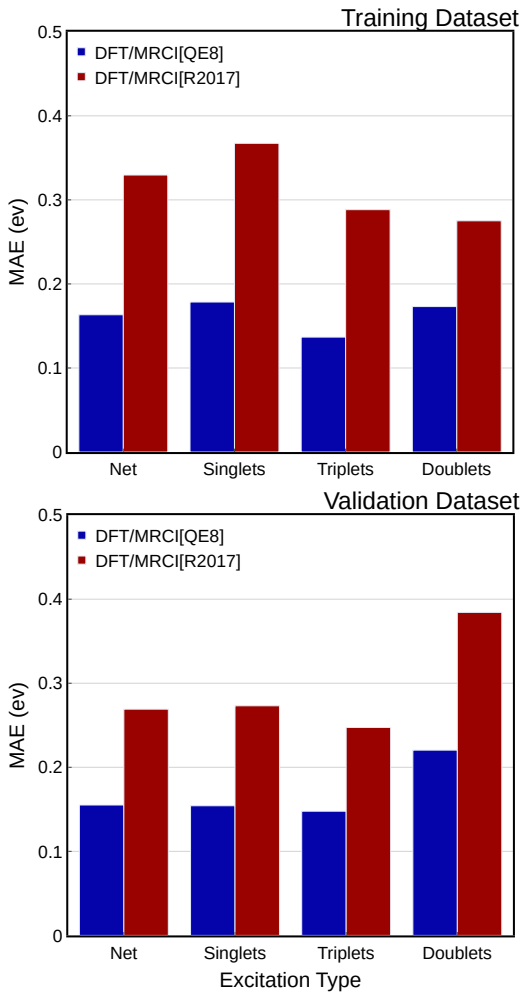

FIG. 3. Mean absolute errors for singlets, triplet, and doublets in the fitting and validation sets for the new (QE8) and a previous (R2017) DFT/MRCI parameterization.

Hamiltonians, as shown in Figure 4. Firstly, the QE8 error distribution is approximately centered around zero, exhibiting mean errors of −0.001 and 0.06 eV for the fitting and validation sets, respectively. In contrast, the R2017 Hamiltonian yields excitation energies that are evidently systematically red-shifted from the *ab initio* benchmark values, and yields mean errors of −0.32 and −0.26 eV, for the fitting and validation sets, respectively. The most likely origin of this is the previously noted empirical observation that for most organic chromophores, UV/VIS band maxima will generally be red-shifted relative to the corresponding electronic vertical excitation energies.[12,13] The inclusion of such data in the fitting set for the R2017 Hamiltonian thus manifests as a systematic underestimation of VEEs.

4. *Performance of Approximate DFT/MRCI methods*

Recent years have seen the development of approximate DFT/MRCI methods that exhibit an even higher level of computational efficiency, termed p-DFT/MRCI[7] and DFT/MRCI(2).[8] The p-DFT/MRCI method prunes the FOIS by removing those configurations that do not contribute appreciably to a second-order perturbation theory estimate of the energies of the states of interest. The surviving configurations are used to construct the DFT/MRCI Hamiltonian, while the contributions of the pruned ones are treated approximately at the level of second-order Rayleigh-Schrödinger perturbation theory. In the DFT/MRCI(2) approach, the construction of Hamiltonian matrix elements between FOIS configurations is entirely obviated. Instead, a small effective Hamiltonian is constructed using second-order generalized van Vleck perturbation theory (GVVPT2)[50], the eigenvalues of which provide approximations of those of the DFT/MRCI Hamiltonian. It has previously been demonstrated that both the p-DFT/MRCI and DFT/MRCI(2) approaches can yield computational savings of upwards of two orders of magnitude for large molecules while retaining the accuracy of the parent DFT/MRCI method.[7,8]

TABLE III. Comparison of errors for DFT/MRCI, p-DFT/MRCI and DFT/MRCI(2) QE8 Hamiltonian vertical excitation energies relative to both the fitting and validation datasets. All values are given in units of eV. All values were computed using a configuration selection parameter value of $\delta E_{sel} = 1.0$ E$_h$

Fitting Data MAEs				
Method[QE8]	Singlet	Triplet	Doublet	Total
DFT/MRCI	0.18	0.14	0.17	0.16
p-DFT/MRCI	0.18	0.14	0.17	0.16
DFT/MRCI(2)	0.17	0.14	0.17	0.16
Validation Data MAEs				
Method[QE8]	Singlet	Triplet	Doublet	Total
DFT/MRCI	0.15	0.15	0.22	0.16
p-DFT/MRCI	0.16	0.15	0.22	0.16
DFT/MRCI(2)	0.15	0.14	0.22	0.15

Table III shows the MAEs furnished by the p-DFT/MRCI and DFT/MRCI(2) approaches in combination with the QE8 Hamiltonian with respect to fitting and validation data sets discussed above. The DFT/MRCI results are shown alongside for ease of comparison. For both the fitting and validation sets, the MAEs are reproduced to within 0.01 eV for both p-DFT/MRCI and DFT/MRCI(2) relative to the parent DFT/MRCI method. Additionally, the error histograms in Figure 4 are particularly illustrative in demonstrating how closely p-DFT/MRCI and DFT/MRCI(2) reproduce the DFT/MRCI results. The mean errors for the fitting (validation) sets are 0.004 (0.06) eV for p-DFT/MRCI and 0.006 (0.07) eV for DFT/MRCI(2), which is in excellent agreement the DFT/MRCI value of −0.001 (0.06) eV.

In summary, the QE8 Hamiltonian does not negatively affect the previously observed agreement between parent DFT/MRCI and related p-DFT/MRCI and DFT/MRCI(2)

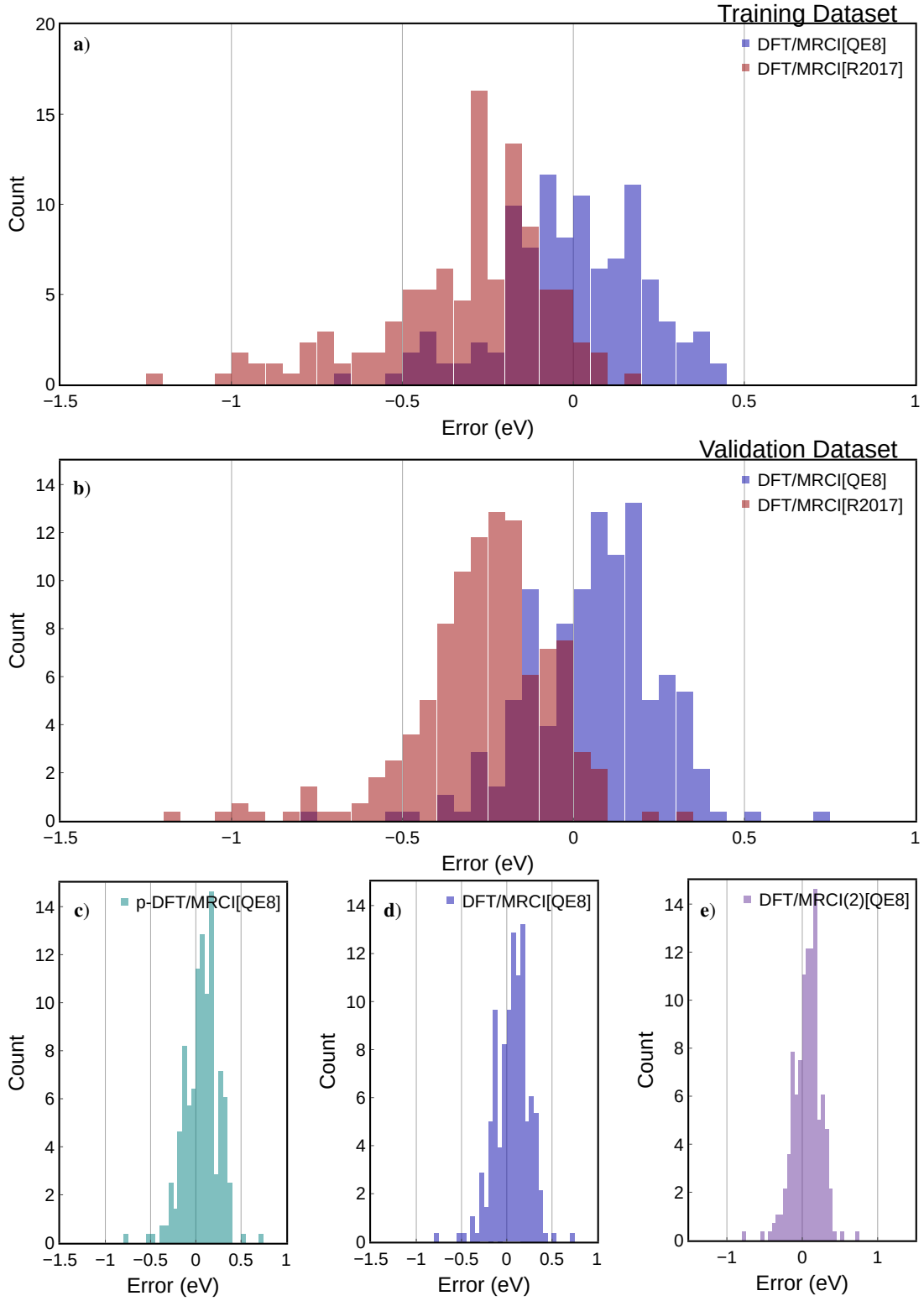

FIG. 4. Signed vertical excitation energy errors for the QE8 and R2017 Hamiltonians relative to both the fitting and validation data sets. The performance of the p-DFT/MRCI and DFT/MRCI(2) methods (panels (c) and (e)) are also compared to the parent DFT/MRCI method (panel (d). In all calculations a value of the configuration selection parameter $\delta E_{sel} = 1.0$ E$_h$ was used.

methods. The consistency in the computed excitation energies is impressive, with p-DFT/MRCI and DFT/MRCI both

exhibiting a mean difference of ≤ 0.01 eV with DFT/MRCI overall excitation classes in the validation set.

V. REPRESENTATIVE APPLICATIONS

We conclude this work with two representative applications that illustrate the range of applicability of the DFT/MRCI family of methods in conjunction with the QE8 Hamiltonian: (i) the calculation of electronic spectra of the five nucleobases, and (ii) the calculation of the low-lying states of chlorophyll a. For each example, the results are compared to established high-level *ab initio* electronic structure methods. The purpose of the former application is to show the ability of the QE8 Hamiltonian to furnish oscillator strengths of good quality in addition to VEEs. Through the second application, we seek to highlight the ability to treat large molecules at low cost while retaining high levels of accuracy.

A. Electronic Spectra of Nucleobases

One important application area for DFT/MRCI is the simulation of optical absorption spectra of molecules. The ability to reliably compute vertical excitation energies and oscillator strengths for a potentially large number of valence electronic states is crucial for these simulations. Accurate simulations of absorption spectra that explicitly describe non-adiabatic vibronic coupling effects have been carried out previously using DFT/MRCI potentials and non-adiabatic couplings,[51,52] and the present QE8 parameterization has been conceived with applications such as these in mind.

The QE8 parameterization was performed using benchmark vertical excitation energies only, typically including ≤ 5 states per molecule. Although there has been recent work to construct equivalent benchmark QUEST databases for oscillator strengths[53], such information was not included in the parameter optimization process. In order to be useful in the simulation of absorption spectra, the QE8 Hamiltonian needs to be capable of furnishing both accurate excitation energies and oscillator strengths. To assess this capability, we consider the calculation of purely electronic absorption spectra for the five nucleobases (adenine, cytosine, guanine, thymine, and uracil). The results are compared to a well-established *ab initio* method: EOM-CCSD. This method, in combination with the aug-cc-pVTZ basis, has a similar accuracy to the QE8 Hamiltonian for valence excitation energies of organic molecules, exhibiting an MAE of 0.13 eV relative to the theoretical best estimates for the 442 excitation energies that span all of the QUEST datasets. Additionally, this method is capable of furnishing accurate oscillator strengths.[53]

For each molecule, VEEs and oscillator strengths were computed at both the EOM-CCSD and DFT/MRCI[QE8] levels of theory using the aug-cc-pVDZ basis. The resulting electronic spectra are shown in ig. V A. Here, in order to aid a visual comparison, the spectra are also shown convoluted with a Gaussian lineshape with a full width at half maximum (FWHM) of 0.2 eV. Broadly speaking, the agreement between

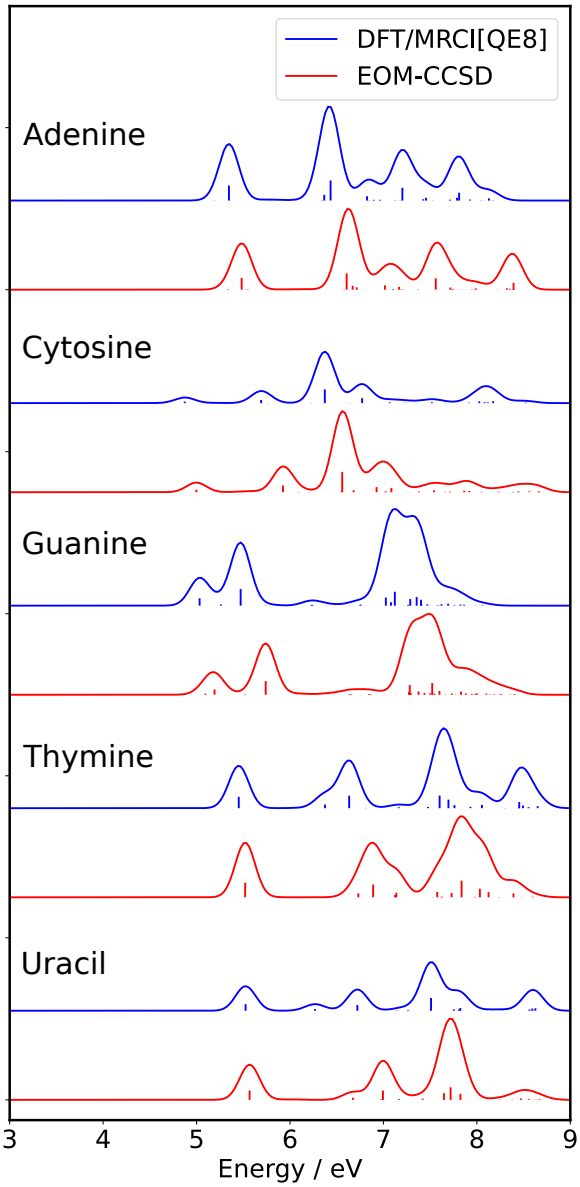

FIG. 5. Vertical excitation energies and oscillator strengths for nucleobases adenine, cytosine, guanine, thymine, and uracil. DFT/MRCI[QE8] is compared to EOM-CCSD results.

the DFT/MRCI and EOM-CCSD vertical excitation energies is excellent, with differences between excitations involving equivalent states observed to be ≤ 0.3eV; none of the spectra shown in Fig. V A have had any shifts applied Moreover, the relative oscillator strengths are found to be in good agreement. Notably, however, the percent difference for some of the bright transition oscillator strengths was observed to be up to 40% and may warrant subsequent study. For present purposes, we wish only to demonstrate that the vertical excitation energies of the optically bright electronic states and relative oscillator strengths agree well with the corresponding EOM-CCSD results over a range of multiple eV.

B. Application to Large Molecules

The development of perturbative DFT/MRCI approaches enables the application of these methods to very large molecular system at low computational cost. Here, we consider the calculation of the VEEs of the low-lying singlet excited states of chlorophyll a using DFT/MRCI(2) and the QE8 Hamiltonian. In particular, we consider the four "Gouterman-type" states, corresponding to excitation between the HOMO-1, HOMO, LUMO and LUMO+1 π and π^* orbitals localized on the chlorin ring.[54,55] In order of increasing energy, these are labeled the Q_y, Q_x, B_x, and B_y states. The ground state minimum energy geometry is shown in Figure 6 b). It is commonly accepted that the phytyl chain is essentially electronically decoupled from the chlorin ring on which the Q and B state excitations are localized. As such, when computing these states, it is common practice to replace the phytyl chain with a methyl group in order to reduce computational costs. The geometry of this reduced system is shown in Figure 6 a). The computational expediency of the DFT/MRCI(2) method, however, allows for the treatment of the full, 137-atom system.

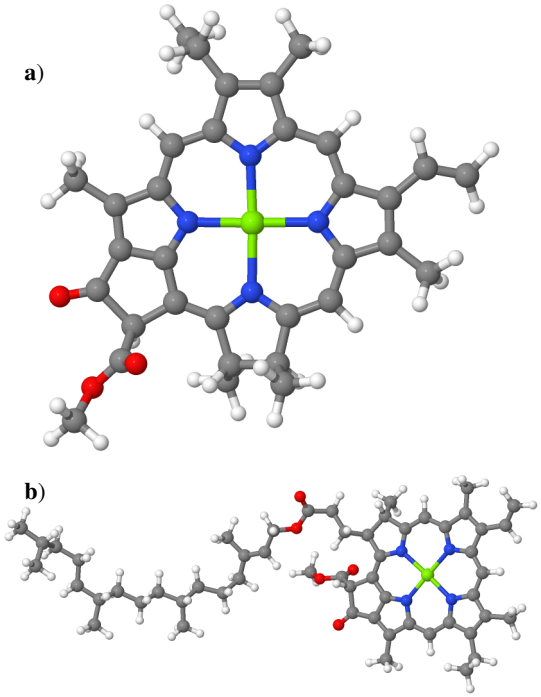

FIG. 6. Molecular geometry for a model of chlorophyll *a* (top, a)) in which the hydrocarbon tail is terminated with a methyl group, as well as the complete structure of chlorophyll *a* (bottom, b))

In a recent study, Sirohiwal and co-workers[56] published benchmark domain local pair natural orbital similarity transformed CCSD (DLPNO-STEOM-CCSD) calculations of the VEEs of the Q and B states for the model system in which a methyl group is substituted for the phytyl chain. In that work, the UV/VIS absorption spectrum of the Q and B bands was simulated in the Franck-Condon approximation, in which the vibrational structure of the excited states was incorporated into the simulation while neglecting the non-adiabatic coupling between relevant electronic states. On the basis of these calculations, and engendered by good agreement with experiment, the authors could report the best estimates of the VEEs that resulted in the best agreement between the simulated and experimental spectra.

TABLE IV. Vertical excitation energies of four Gouterman-type excited states of chlorophyll a for both the full and model structures. In all cases, the def2-TZVP basis set was used. All values are given in units of eV.

State	STEOM-CCSD	DFT/MRCI(2)		Best
	model[a]	model	full	estimate
Q_y	1.75	2.10	2.06	1.99
Q_x	2.24	2.40	2.36	2.30
B_x	3.17	3.21	3.11	3.12
B_y	3.40	3.40	3.33	3.38

[a] Ref. 56.

In Table IV we report VEEs of the Q and B states for both the full chlorophyl a structure, including the phtyl chains, as well as the reduced model geometry, with the phytyl chain removed, as computed at the def2-TZVP/DFT/MRCI(2) level of theory using the QE8 Hamiltonian. Shown alongside are the DLPNO-STEOM-CCSD/def2-TZVP results of Sirohiwal *et al.*[56] as well as the current best estimates. It is found that DFT/MRCI(2) furnishes results in very close agreement with the DLPNO-STEOM-CCSD results as well as the best estimate values. The computations on the full molecule, in particular, are generally in excellent agreement with the best estimate values, with a maximum deviation of only 0.07 eV observed for the Q_y state. We thus conclude that the QE8 Hamiltonian is capable of producing excellent quality results for this prototypical system, albeit at very modest computational costs.

C. Computational Efficiency: Timings

We close by commenting on the computational efficiency of the DFT/MRCI family of methods: DFT/MRCI, p-DFT/MRCI, and DFT/MRCI(2). Table V shows the computation CPU times for each of the applications presented in Sec. V for each of these methods. We note that these timings include only the post-DFT components of the calculations. The results shown here were generated using the GRaCI package[34], which uses the PySCF program package to solve the KS DFT equations. However, this aspect of the computation is highly modular and other packages may be employed in this respect, the computational efficiency of which will vary widely.

We first consider the nucleobase calculations. Table V clearly demonstrates how computationally tractable even extremely large computations are using the DFT/MRCI family of methods. No computation of the UV/VIS spectra shown in Fig. V A, which involved the determination of the 30 lowest-energy roots, took longer than 45 minutes using the DFT/MRCI method running in serial. Using the

TABLE V. CPU times for representative applications (in units of seconds).[a]

Molecule	DFT/MRCI	p-DFT/MRCI	DFT/MRCI(2)	N_{roots}
Adenine	1851	194	50	30
Cytosine	735	47	27	30
Guanine	2604	273	61	30
Thymine	880	101	35	30
Uracil	276	38	22	30
Chlorophyll a (model)	–	–	265	6
Chlorophyll a (full)	–	–	2044	6

[a] Values in the table are CPU times (excluding KS-DFT computation) and correspond to calculations run using a single thread of a Intel Xeon Gold 6130 processor. The aug-cc-pVDZ basis set was employed for the nucleobases, and the def2-TZVP basis was used for both chlorophyll a systems.

DFT/MRCI(2) method results in large computational savings, with an average CPU time of just 39 seconds. Moreover, the resulting spectra are found to be essentially identical to the more expensive DFT/MRCI results, consistent with the agreement in vertical excitation energies presented in Table III. Likewise, the p-DFT/MRCI method also yields significant savings relative to the parent DFT/MRCI method, whilst furnishing VEEs within 0.01 eV of the DFT/MRCI values. However, the p-DFT/MRCI CPU times are found to be significantly greater than those for DFT/MRCI(2). We note here that, due to the increased amount of correlation captured by the QE8 Hamiltonian relative to previous DFT/MRCI Hamiltonians, the configuration pruning algorithm used in p-DFT/MRCI is rendered less effective than previously reported[7], and DFT/MRCI(2) now emerges as the clear method of choice for reduced cost DFT/MRCI calculations.

Finally, we consider the timings for the cholorophyll a calculations, as performed at the DFT/MRCI(2)/def2-TZVP level of theory. For the model structure, the required CPU time is less than 5 minutes to compute six roots. For the full structure calculation, including the phytyl chain, this increases to 34 minutes, still a remarkably low computational cost for a system of this size, especially given the impressive accuracy of the results (see Table IV).

VI. CONCLUSION

We have presented a newly-developed DFT/MRCI Hamiltonian, termed QE8, that is parameterized using only benchmark-level *ab inito* data in the form of VEEs taken from the QUEST database. With an eye to the development of a DFT/MRCI Hamiltonian that gives a balanced description of core-excited states, a new XC functional was adopted, namely, the QTP17 functional. The performance of a Hamiltonian designed to yield highly accurate core-excitation energies, CVS-QE8, will be discussed in a subsequent publication. The present Hamiltonian yields VEEs with an MAE of 0.16 eV with respect to theoretical best estimates for valence-excited states of organic molecules. Importantly, by fitting the Hamiltonian parameters to only *ab initio* data, the systematic underestimation of VEEs exhibited by previous DFT/MRCI Hamiltonians is ameliorated. Furthermore, doubly-excited states, a previously problematic class of states for DFT/MRCI methods, can now be computed with high accuracy. In combination with the newly-introduced DFT/MRCI(2) variant, the QE8 Hamiltonian opens the door to the accurate simulation of excited states of large organic molecules at very low cost, as exemplified by the current application to the chlorophyll a molecule.

The Hamiltonian design principle employed in the present work involved only the reproduction of benchmark VEEs. Future efforts will examine, and potentially improve, the performance of DFT/MRCI methods for the computation of other quantities relevant for simulations in spectroscopy and dynamics. Specifically, benchmark electronic properties, such as oscillator strengths and electronic moments, are becoming increasingly available[53,57]. The degree to which DFT/MRCI wave functions furnish accurate electronic properties will be investigated in a subsequent publication.

The ability of DFT/MRCI to produce accurate potential energy surfaces has been studied only indirectly via, for example, the parameterization of vibronic Hamiltonians. That the DFT/MRCI reference space is iteratively generated anew at each nuclear configuration means that the underlying potential is not entirely smooth and complicates the evaluation of energy gradients and nonadiabatic couplings. However, recent work in our group employing gaussian process and kernel ridge regression approaches to learn potential energy surfaces[58] will likewise be applicable to DFT/MRCI potentials, providing a route to structure optimization and dynamics simulations. The accuracy of excited structural minima derived from DFT/MRCI potential surfaces, including minimum energy conical intersections and the corresponding branching spaces, will likewise be evaluated in an upcoming work.

VII. ACKNOWLEDGEMENT

M.S.S. thanks the Natural Sciences and Engineering Research Council (NSERC) Discovery grant program for financial support.

VIII. DATA AVAILABILITY STATEMENT

Full fitting and validation sets, as well as the nuclear geometries of the molecules employed in representative applications are provided as Supplementary Material.

Appendix A: Notation: spin-coupling coefficients

Following Segal and Whetmore[18], we adopt the following notation for the one- and two-electron singlet spin-coupling coefficients:

$$\eta_i^j(\mathbf{w},\omega,\mathbf{w}',\omega') \equiv \langle \mathbf{w}\omega|\hat{E}_i^j|\mathbf{w}'\omega'\rangle, \tag{A1}$$

and

$$\eta_{ik}^{jl}(\mathbf{w},\omega,\mathbf{w}',\omega') \equiv \langle \mathbf{w}\omega|\hat{E}_i^j\hat{E}_k^l|\mathbf{w}'\omega'\rangle, \tag{A2}$$

respectively, where $\hat{E}_i^j$ denotes a singlet excitation operator,

$$\hat{E}_i^j = \sum_{\sigma=\alpha,\beta} \hat{a}_{i\sigma}^\dagger \hat{a}_{j\sigma}, \tag{A3}$$

where $\hat{a}_{i\sigma}^\dagger$ ($\hat{a}_{i\sigma}$) denotes the elementary Fermionic creation (annihilation) operator associated with the spin-orbital with spatial index i and spin index σ.

In the following, we shall, where appropriate, drop the explicit dependence of the spin-coupling coefficients on the bra and ket spatial occupations and spin-couplings, i.e.,

$$\eta_i^j(\mathbf{w},\omega,\mathbf{w}',\omega') \to \eta_i^j, \tag{A4}$$

$$\eta_{ik}^{jl}(\mathbf{w},\omega,\mathbf{w}',\omega') \to \eta_{ik}^{jl}, \tag{A5}$$

with the now implicit dependence always being clear from the context in which the simplified notation is used.

[1] S. Grimme and M. Waletzke, "A combination of kohn–sham density functional theory and multi-reference configuration interaction methods," The Journal of chemical physics **111**, 5645–5655 (1999).

[2] I. Lyskov, M. Kleinschmidt, and C. M. Marian, "Redesign of the dft/mrci hamiltonian," The Journal of chemical physics **144**, 034104 (2016).

[3] A. Heil and C. M. Marian, "Dft/mrci hamiltonian for odd and even numbers of electrons," The Journal of Chemical Physics **147**, 194104 (2017).

[4] 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," The Journal of Chemical Physics **149**, 164106 (2018).

[5] C. M. Marian, A. Heil, and M. Kleinschmidt, "The dft/mrci method," WIREs Computational Molecular Science **9**, e1394 (2019), https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1394.

[6] D. R. Dombrowski, T. Schulz, M. Kleinschmidt, and C. M. Marian, "R2022: A dft/mrci ansatz with improved performance for double excitations," The Journal of Physical Chemistry A **127**, 2011–2025 (2023), pMID: 36799533, https://doi.org/10.1021/acs.jpca.2c07951.

[7] 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.

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

[9] I. Seidu, S. P. Neville, M. Kleinschmidt, A. Heil, C. M. Marian, and M. S. Schuurman, "The simulation of x-ray absorption spectra from ground and excited electronic states using core-valence separated dft/mrci," The Journal of Chemical Physics **151**, 144104 (2019).

[10] K. S. Zinchenko, F. Ardana-Lamas, I. Seidu, S. P. Neville, J. van der Veen, V. U. Lanfaloni, M. S. Schuurman, and H. J. Wörner, "Sub-7-femtosecond conical-intersection dynamics probed at the carbon k-edge," Science **371**, 489–494 (2021), https://www.science.org/doi/pdf/10.1126/science.abf1656.

[11] F. J. Avila Ferrer, J. Cerezo, E. Stendardo, R. Improta, and F. Santoro, "Insights for an accurate comparison of computational data to experimental absorption and emission spectra: Beyond the vertical transition approximation," Journal of Chemical Theory and Computation **9**, 2072–2082 (2013), pMID: 26583553, https://doi.org/10.1021/ct301107m.

[12] B. Lasorne, J. Jornet-Somoza, H.-D. Meyer, D. Lauvergnat, M. A. Robb, and F. Gatti, "Vertical transition energies vs. absorption maxima: Illustration with the uv absorption spectrum of ethylene," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **119**, 52–58 (2014), frontiers in molecular vibrational calculations and computational spectroscopy.

[13] S. Bai, R. Mansour, L. Stojanović, J. M. Toldo, and M. Barbatti, "On the origin of the shift between vertical excitation and band maximum in molecular photoabsorption," Journal of Molecular Modeling **26**, 107 (2020).

[14] S. P. Neville and G. A. Worth, "A reinterpretation of the electronic spectrum of pyrrole: A quantum dynamics study," The Journal of Chemical Physics **140**, 034317 (2014), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4861223/15472151/034317_1_online.pdf.

[15] S. P. Neville, A. Stolow, and M. S. Schuurman, "Vacuum ultraviolet excited state dynamics of the smallest ring, cyclopropane. I. A reinterpretation of the electronic spectrum and the effect of intensity borrowing," The Journal of Chemical Physics **149**, 144310 (2018), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.5044392/15549648/144310_1_online.pdf.

[16] A. D. Becke, "A new mixing of Hartree–Fock and local density-functional theories," The Journal of Chemical Physics **98**, 1372–1377 (1993), https://pubs.aip.org/aip/jcp/article-pdf/98/2/1372/19125465/1372_1_online.pdf.

[17] I. Lyskov, *Redesign and Reparameterization of the DFT/MRCI Hamiltonian and its Application to Electronically Excited Linear Polyenes*, Ph.D. thesis (2016).

[18] R. W. Wetmore and G. A. Segal, "?" Chem. Phys. Lett. **36**, 478 (1975).

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

[20] P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin, "A mountaineering strategy to excited states: Highly accurate reference energies and benchmarks," Journal of Chemical Theory and Computation **14**, 4360–4379 (2018), pMID: 29966098, https://doi.org/10.1021/acs.jctc.8b00406.

[21] P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, and D. Jacquemin, "Reference energies for double excitations," Journal of Chemical Theory and Computation **15**, 1939–1956 (2019), https://doi.org/10.1021/acs.jctc.8b01205.

[22] 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.

[23] P.-F. Loos, A. Scemama, M. Boggio-Pasqua, and D. Jacquemin, "Mountaineering strategy to excited states: Highly accurate energies and benchmarks for exotic molecules and radicals," Journal of Chemical Theory and Computation **16**, 3720–3736 (2020), pMID: 32379442, https://doi.org/10.1021/acs.jctc.0c00227.

[24] 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.

[25] P.-F. Loos, M. Comin, X. Blase, and D. Jacquemin, "Reference energies for intramolecular charge-transfer excitations," Journal of Chemical Theory and Computation **17**, 3666–3686 (2021), pMID: 33955742, https://doi.org/10.1021/acs.jctc.1c00226.

[26] P.-F. Loos and D. Jacquemin, "A mountaineering strategy to excited states: Highly accurate energies and benchmarks for bicyclic systems," The Journal of Physical Chemistry A **125**, 10174–10188 (2021), pMID: 34792354, https://doi.org/10.1021/acs.jpca.1c08524.

[27] P.-F. Loos, A. Scemama, and D. Jacquemin, "The quest for highly accurate excitation energies: A computational perspective," The Journal of Physical Chemistry Letters **11**, 2374–2383 (2020), pMID: 32125872, https://doi.org/10.1021/acs.jpclett.0c00014.

[28] B. Huron, J. P. Malrieu, and P. Rancurel, "Iterative perturba-

tion calculations of ground and excited state energies from multi-configurational zeroth-order wavefunctions," The Journal of Chemical Physics **58**, 5745–5759 (1973), https://pubs.aip.org/aip/jcp/article-pdf/58/12/5745/18885418/5745_1_online.pdf.

[29] E. Giner, A. Scemama, and M. Caffarel, "Using perturbatively selected configuration interaction in quantum monte carlo calculations," Canadian Journal of Chemistry **91**, 879–885 (2013), https://doi.org/10.1139/cjc-2013-0017.

[30] E. Giner, A. Scemama, and M. Caffarel, "Fixed-node diffusion Monte Carlo potential energy curve of the fluorine molecule F2 using selected configuration interaction trial wavefunctions," The Journal of Chemical Physics **142**, 044115 (2015), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4905528/15490864/044115_1_online.pdf.

[31] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner, and A. Scemama, "Quantum package 2.0: An open-source determinant-driven suite of programs," Journal of Chemical Theory and Computation **15**, 3591–3609 (2019), pMID: 31082265, https://doi.org/10.1021/acs.jctc.9b00176.

[32] R. A. Kendall, J. Dunning, Thom H., and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," The Journal of Chemical Physics **96**, 6796–6806 (1992), https://pubs.aip.org/aip/jcp/article-pdf/96/9/6796/18998924/6796_1_online.pdf.

[33] D. E. Woon and J. Dunning, Thom H., "Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties," The Journal of Chemical Physics **100**, 2975–2988 (1994), https://pubs.aip.org/aip/jcp/article-pdf/100/4/2975/19303178/2975_1_online.pdf.

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

[35] Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. W. III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. D. Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. G. III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. S. III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill, and M. Head-Gordon, "Advances in molecular quantum chemistry contained in the q-chem 4 program package," Molecular Physics **113**, 184–215 (2015), https://doi.org/10.1080/00268976.2014.952696.

[36] J. Miralles, O. Castell, R. Caballol, and J.-P. Malrieu, "Specific ci calculation of energy differences: Transition energies and bond energies," Chemical Physics **172**, 33–43 (1993).

[37] S. Grimme, "Density functional calculations with configuration interaction for the excited states of molecules," Chemical Physics Letters **259**, 128–137 (1996).

[38] J. Olsen, B. O. Roos, P. Jo/rgensen, and H. J. A. Jensen, "Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces," The Journal of Chemical Physics **89**, 2185–2192 (1988), https://pubs.aip.org/aip/jcp/article-pdf/89/4/2185/18972558/2185_1_online.pdf.

[39] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," The Computer Journal **7**, 308–313 (1965), https://academic.oup.com/comjnl/article-pdf/7/4/308/1013182/7-4-308.pdf.

[40] F. Gao and L. Han, "Implementing the nelder-mead simplex algorithm with adaptive parameters," Computational Optimization and Applications **51**, 259–277 (2012).

[41] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," Nature Methods **17**, 261–272 (2020).

[42] S. Hirata and M. Head-Gordon, "Time-dependent density functional theory within the tamm–dancoff approximation," Chemical Physics Letters **314**, 291–299 (1999).

[43] K. Pernal, R. Podeszwa, K. Patkowski, and K. Szalewicz, "Dispersionless density functional theory," Phys. Rev. Lett. **103**, 263201 (2009).

[44] K. W. Wiitala, T. R. Hoye, and C. J. Cramer, "Hybrid density functional methods empirically optimized for the computation of 13c and 1h chemical shifts in chloroform solution," Journal of Chemical Theory and Computation **2**, 1085–1092 (2006), pMID: 26633067, https://doi.org/10.1021/ct6001016.

[45] 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.

[46] P. Verma and R. J. Bartlett, "Increasing the applicability of density functional theory. IV. Consequences of ionization-potential improved exchange-correlation potentials," The Journal of Chemical Physics **140**, 18A534 (2014), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4871409/15482116/18a534_1_online.pdf.

[47] R. J. Bartlett, V. F. Lotrich, and I. V. Schweigert, "Ab initio density functional theory: The best of both worlds?" The Journal of chemical physics **123** (2005).

[48] M. K. Harbola, "Relationship between the highest occupied kohn-sham orbital eigenvalue and ionization energy," Physical Review B **60**, 4545 (1999).

[49] 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.

[50] 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://pubs.aip.org/aip/jcp/article-pdf/75/2/798/18929991/798_1_online.pdf.

[51] S. P. Neville, I. Seidu, and M. S. Schuurman, "Propagative block diagonalization diabatization of DFT/MRCI electronic states," The Journal of Chemical Physics **152**, 114110 (2020), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.5143126/15573990/114110_1_online.pdf.

[52] R. Forbes, S. P. Neville, M. A. B. Larsen, A. Röder, A. E. Boguslavskiy, R. Lausten, M. S. Schuurman, and A. Stolow, "Vacuum ultraviolet excited state dynamics of the smallest ketone: Acetone," The Journal of Physical Chemistry Letters **12**, 8541–8547 (2021), pMID: 34464141, https://doi.org/10.1021/acs.jpclett.1c02612.

[53] A. Chrayteh, A. Blondel, P.-F. Loos, and D. Jacquemin, "Mountaineering strategy to excited states: Highly accurate oscillator strengths and dipole moments of small molecules," Journal of Chemical Theory and Computation **17**, 416–438 (2021), pMID: 33256412, https://doi.org/10.1021/acs.jctc.0c01111.

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

[55] 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.

[56] 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.

[57] R. Sarkar, M. Boggio-Pasqua, P.-F. Loos, and D. Jacquemin, “Benchmarking td-dft and wave function methods for oscillator strengths and excited-state dipole moments,” Journal of Chemical Theory and Computation **17**, 1117–1132 (2021), pMID: 33492950, https://doi.org/10.1021/acs.jctc.0c01228.

[58] T. Y. Wang, S. P. Neville, and M. S. Schuurman, “Machine learning seams of conical intersection: A characteristic polynomial approach,” The Journal of Physical Chemistry Letters **14**, 7780–7786 (2023), pMID: 37615964, https://doi.org/10.1021/acs.jpclett.3c01649.