A Density Functional Extension to Excited State Mean-Field Theory

Luning Zhao[†] and Eric Neuscamman[*,†,‡]

†*Department of Chemistry, University of California, Berkeley, California, 94720, USA*

‡*Chemical Science Division, Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA*

E-mail: eneuscamman@berkeley.edu

Abstract

We investigate an extension of excited state mean-field theory in which the energy expression is augmented with density functional components in an effort to include the effects of weak electron correlations. The approach remains variational and entirely time-independent, allowing it to avoid some of the difficulties associated with linear response and the adiabatic approximation. In particular, all of the electrons' orbitals are relaxed state specifically and there is no reliance on Kohn-Sham orbital energy differences, both of which are important features in the context of charge transfer. Preliminary testing shows clear advantages for single-component charge transfer states, but the method, at least in its current form, is less reliable for states in which multiple particle-hole transitions contribute significantly.

1 Introduction

The recently developed excited state mean-field theory (ESMF)[1] is intended to act as a mean-field platform for excited states in much the same way as Hartree-Fock (HF) theory[2]

does for ground states. As one might expect, these two theories share many properties: they rely on minimally correlated wave function forms, produce energy stationary points, relax orbital shapes variationally, and have the same cost-scaling. Also like HF theory, ESMF lacks a proper description of correlation effects, and so from a practical standpoint is expected to be more useful as a starting point for correlation methods than as a way of making energy predictions on its own. While there are many ways one could go about capturing correlation effects, it is hard to avoid thinking about density functional theory (DFT)[3] in this context given how useful it is for this purpose in ground states.

The Kohn-Sham (KS) formulation of DFT[4] is the most widely used electronic structure method in chemistry, physics, and materials science. Due to its favorable scaling with system size and reasonable accuracy in a variety of different circumstances, DFT is often regarded as one of the most powerful tools for studying the electronic and dynamic properties of materials and medium to large molecules. The KS-DFT method can also be considered as an extension to the HF method, by replacing the exchange energy in HF with the exchange-correlation (xc) energy in KS-DFT. With the exact xc functional, KS-DFT is able to capture correlation effects exactly. Comparing to other post-HF methods that account for weak correlation effects, such as configuration interaction, Moller-Plesset 2nd order perturbation theory, and coupled cluster, the most appealing feature of KS-DFT is its low cost-scaling, which allows it to be applied to systems with thousands of electrons or more.

Inspired by the success of KS-DFT in ground states, one may wonder whether similar extensions using DFT can also be achieved for ESMF. Intuitively, combining ESMF with DFT would allow one to go beyond the mean-field form of the ESMF wave function and be able to recover weak correlation effects while maintaining the mean-field cost scaling of ESMF. More importantly, such an approach need not rely on linear response (LR) theory or the adiabatic approximation (AA), both of which are central to the practical application of time-dependent density functional theory (TDDFT).[5–8] As the combination of LR and the AA can produce significant errors in some excited states, it would be very interesting to

instead augment ESMF theory by incorporating components from density functional theory while keeping the formulation entirely time-independent.

While the LR formulation of TDDFT is formally exact, approximations are needed to make the approach computationally tractable. The central quantity in TDDFT is the xc-kernel $f_{xc}(\mathbf{r}, t; \mathbf{r}', t')$, defined as the functional derivative of the xc-potential,[9]

$$f_{xc}(\mathbf{r}, t; \mathbf{r}'t') = \frac{\delta \upsilon_{xc}[n](\mathbf{r}, t)}{\delta n(\mathbf{r}', t')} \tag{1}$$

in which the $\upsilon_{xc}(\mathbf{r}, t)$ is the time-dependent analogy of the ground state xc-potential and $n(\mathbf{r}, t)$ is the electron density. The AA replaces the time-dependent xc-potential with the ground state xc-potential,[9]

$$\upsilon_{xc}^{adia}[n](\mathbf{r}, t) = \upsilon_{xc}^{GS}[n(t)](\mathbf{r}) \tag{2}$$

at which point the xc-kernel becomes

$$f_{xc}^{adia}(\mathbf{r}, t; \mathbf{r}'t') = \frac{\delta \upsilon_{xc}^{GS}[n(t)](\mathbf{r})}{\delta n(\mathbf{r}', t')} = \delta(t - t') \frac{\delta^2 E_{xc}[n]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')}. \tag{3}$$

Most notably, this approximation leads the xc-kernel to be local both in time and space if the ground state xc functional is local as in LDA, or local in time but nonlocal in space in the case of hybrid functionals.

While the AA is enormously convenient in that it makes TDDFT efficient and allows it to use existing ground state functionals, it does create important limitations when modeling charge transfer (CT), Rydberg, and double excitations. For example, TDDFT often drastically underestimates excitation energies for long-range CT states[10–12] and Rydberg states,[13–15] and it is completely incapable of describing doubly excited states.[16,17] Besides the underestimation of excitation energies, it is also well known that for long-range CT excited states, standard pure and hybrid functionals also fail to exhibit the correct $1/R$ dependence along the charge separation coordination.[10,18] Given the technological and bi-

ological importance of CT, the limitations of practical TDDFT in this area are especially frustrating.

To be more precise, these difficulties stem from two approximations: first, the usage of approximate xc functionals, and second, the AA. The former is responsible for the problems in Rydberg excited states and the missing $1/R$ behavior in long-range CT. For Rydberg states, the problem lies primarily in the fact that the ground state xc-potential of local and semi-local functionals decays exponentially with r, much faster than the $1/r$ decay of the exact xc-potential. In order to deal with this problem, the asymptotic correction approach[14] has been developed and results in dramatically improved Rydberg energetics. For CT excited states, range-separated hybrid functionals (RSH)[19–24] successfully yield the correct $1/R$ behavior of long range CT excited states. This approach eliminates the CT self-interaction error in which derivatives of the approximated exchange term fail to deliver the long range Coulomb term that should be present in the linear response equations.[16] Even with the $1/R$ issue repaired, though, long range CT still poses challenges. This is mainly due to fact that the excitation energy of long-range CT states should be determined by the ionization potential (IP) of the donor and the electron affinity (EA) of the acceptor. While KS-DFT's highest occupied molecular orbital (HOMO) energy corresponds to the IP, the lowest unoccupied molecular orbital (LUMO) energy does not, and is not supposed to, correspond to the EA, even when the xc functional is exact. The result is that the difference between the DFT LUMO and HOMO energies severely underestimates the excitation energy, which leads to a situation in which the kernel contribution to the TDDFT energy is asked to make up the difference. Such a difference between DFT LUMO energy and EA, is usually referred to as the "derivative discontinunity" in TDDFT literature.[25] However, the kernel contribution for most commonly used functionals is typically much too small to make an appreciable difference on this scale, and so CT energies get underestimated, sometimes quite severely. This difficulty, which we will refer to as the EA/IP imbalance, has been extensively studied[10,16,25,26] and can be seen clearly in the examples we investigate below.

Unlike the issues discussed above, TDDFT's failure to describe doubly excited states can be laid squarely at the feet of the AA, which converts the memory-dependent time-non-local xc-kernel $f_{xc}(\mathbf{r}, t; \mathbf{r}'t')$ into a time-local affair with no memory. One consequence of this simplification is that, when expressed in Fourier space, the AA xc-kernel is frequency independent. Given that it has been shown that the exact xc-kernel carries a strong frequency dependence near a double excitation,[26,27] adiabatic xc-kernels are thus not appropriate or accurate for doubly excited states. In practice, the failure of the AA in describing doubly excited states also creates difficulties for other excitations, especially in the context of CT. As pointed out by Ziegler and coworkers,[28,29] another consequence of the AA is that it fails to account for relaxations in the occupied orbitals that are not involved in the excitation. The orbital shapes for the particle and hole are relaxed by TDDFT, but the other orbital shapes are not, at least not when the AA is being used. An intuitive way to see this in light of the double excitation limitation is to consider that, after the single excitation itself, the leading order term in the Taylor expansion of a fully orbital-relaxed singly excited state is a linear combination of doubly excited determinants. Since CT excited states undergo substantial charge deformations and changes in dipole when compared to the ground state, allowing all of the orbitals to relax during the excitation is crucial.[30] Without full relaxation, errors in CT excitation energies often reach multiple eVs,[10] even when RSH functionals are employed. In sum, it would be highly desirable to have an excited state methodology that benefits from DFT's highly efficient incorporation of correlation effects but that is free from the difficulties created by the AA and EA/IP imbalances.

It is worth noting that within generalized KS (e.g. when using global hybrid or RSH functionals) the IP/EA imbalance can be partially resolved without going beyond the AA. For example, as one increases the degree of Hartree-Fock exchange in the xc functional, the LUMO orbital energy becomes more firmly related to an EA. In the limit of using 100% HF exchange and 0% correlation, TDDFT becomes equivalent to time-dependent HF (TDHF) and the LUMO energy becomes the HF EA (the correct 1/R dependence is also achieved,

although that already happens with RSH approaches). In practice, however, functionals do not reach this extreme as HF energetics are not desirable, and so even in the generalized KS context removing the EA/IP imbalance remains challenging. Further, it is important to remember that even if this issue is resolved within the AA, the lack of orbital relaxation remains an issue.

In this paper, we present a density functional extension of the ESMF method (DFE-ESMF). Instead of relying on the linear response formalism and AA of TDDFT, we directly modify the energy expression of ESMF theory by borrowing key ingredients from KS-DFT. While this does not lead to a formal density functional theory as it lacks some of the key properties of ground state DFT, the idea is to exploit density functionals' proven ability to add weak correlation effects to an uncorrelated reference wave function. As such correlations tend to be local, and a local region of a molecule should not be capable of knowing whether it is formally part of an excited state or a ground state, the hope is that the same ingredients that allow KS-DFT to capture weak correlation effects will remain effective in the excited state context. As in the original formulation of ESMF, the energy expression (including the newly incorporated DFT ingredients) is combined with an excited state variational principle to achieve excited-state-specific optimization of the orbitals. As we discuss below, this approach seeks to bypass both the orbital relaxation and EA/IP imbalance issues that show up in the practical application of TDDFT. In a variety of exploratory calculations, we find that, when paired with an xc functional with a high degree of exact exchange (necessary to help alleviate a self-interaction bias stemming from excited states' more open-shell character), this DFE-ESMF approach provides an accuracy comparable to TDDFT for simple single-configuration-state-funciton (single-CSF) valence excitations while far outperforming it in CT states, even when comparing against a RSH functional. The performance for multi-CSF single excitations is more mixed, which appears to be caused by double-counting issues as we discuss in some detail below.

This paper is organized as follows. We begin with a brief review of ground state KS-DFT

so as to make clear its parallels with our excited state formalism. We then develop the working equations for the DFE-ESMF method in the context of both single-configurational and multi-configurational wave functions. We then briefly review the ground state xc functionals that we employ and discuss concerns about possible double counting problems. At the end of the theory section, we compare DFE-ESMF with other excited state and multi-reference DFT methods and also with constrained DFT. Results and discussions are then presented for a variety of different valence, CT, and Rydberg excitations. We conclude our discussion by pointing out the merits and drawbacks of the current method, along with possible directions for future development.

2 Theory

2.1 Ground State DFT

In ground state DFT, the Levy constrained search formulation provides a formally exact energy functional,[3]

$$E[n] = \min_{\Psi \to n} \left\langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \right\rangle + V_{ext}[n] \tag{4}$$

in which $\hat{T}$ and $\hat{V}_{ee}$ are the kinetic and electron-electron repulsion operators, and $V_{ext}[n]$ is the external potential. In practice, KS-DFT re-writes this functional as[3,4]

$$E[n] = T_s[n] + V_{ext}[n] + J[n] + E_{xc}[n], \tag{5}$$

in which $T_s[n]$ is the kinetic energy of a fictitious Slater determinant that shares the same density as the actual interacting system, $J[n]$ is the Hartree part of the electron-electron repulsion energy, and $E_{xc}[n]$ is the exchange-correlation functional. Considering the common case of a closed-shell, spin-restricted KS determinant for N electrons, one can re-express the energy in terms of the orbitals $\phi_i(\mathbf{r})$ for $i \in [1, 2, \ldots, N/2]$. The external potential and

Hartree pieces,

$$V_{ext}[n] = \int v_{ext}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \tag{6}$$

$$J[n] = \frac{1}{2} \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{7}$$

are dependent only on the density, which is in turn now determined by the orbitals,[4]

$$n(\mathbf{r}) = 2 \sum_{i}^{N/2} |\phi_i(\mathbf{r})|^2. \tag{8}$$

Note that we follow the convention that i, j, k refer to occupied orbitals, a, b, c to virtual orbitals, and p, q, r, s to all orbitals. In the original KS formulation,[4] the xc functional depends only on density. However, due to the development of generalized KS schemes[31,32] and hybrid functionals, it becomes more appropriate to write the xc function as a direct function of the orbitals,

$$E_{xc}[n] \rightarrow E_{xc}\big(\phi_1, \phi_2, \ldots, \phi_{N/2}\big). \tag{9}$$

Of course, the KS kinetic energy is also an orbital functional,

$$T_s[n] = -\sum_{i}^{N/2} \int \phi_i(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) d\mathbf{r}. \tag{10}$$

With this orbital-based formulation, one then minimizes Eq. (5) under the constraint that the orbitals remain orthonormal in order to arrive at the KS orbital eigenvector equation,

$$\hat{F}_{KS}\, \phi_i = \epsilon_i\, \phi_i, \tag{11}$$

in which $\hat{F}_{KS}$ is the KS Fock operator.

Crucially, we note that the same energies, orbitals, and densities are arrived at if one

performs the minimization

$$E = \min_{\boldsymbol{X}} \Big\{ T_s[n] + V_{ext}[n] + J[n] + E_{xc}\big(\phi_1, \phi_2, \ldots, \phi_{N/2}\big) \Big\} \tag{12}$$

in terms of the elements of the anti-Hermitian matrix $\boldsymbol{X}$ that transforms some initial orthonormal set of orbitals (such as those that diagonalize the one-electron parts of $\hat{H}$) into the final KS orbitals.

$$\phi_p(\mathbf{r}) = \sum_q \left[e^{\boldsymbol{X}}\right]_{pq} \phi_q^{(0)}(\mathbf{r}) \tag{13}$$

Given ESMF's similarities to HF, it is worthwhile to write the HF energy in this same form.

$$E_{HF} = \min_{\boldsymbol{X}} \Big\{ T_s + V_{ext} + J + E_x \Big\} \tag{14}$$

Here E_x is the HF exchange energy

$$E_x = -\sum_{ij} (ij|ji) \tag{15}$$

which we have expressed in terms of the two-electron integrals in the relaxed orbital basis.

$$(pq|rs) = \int\int \frac{\phi_p(\mathbf{r})\phi_q(\mathbf{r})\phi_r(\mathbf{r}')\phi_s(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{16}$$

By comparing Equation 12 and 14, KS-DFT can be seen as the pairing of a minimally-correlated ansatz (the Slater determinant) and a variational principle (the total energy) in which the energy expression within the latter has been augmented by modifying the exchange term. To formulate DFE-ESMF, we will follow a similar route, but with the ESMF wave function as the minimally-correlated ansatz and using an excited state variational principle instead of simple energy minimization.

2.2 DFE-ESMF: Single-CSF Formalism

In ESMF, excited states are targeted by applying the following Lagrangian form of an excited state variational principle,[1]

$$L = \Big\langle \Psi | (\omega - \hat{H})^2 | \Psi \Big\rangle - \boldsymbol{\mu} \cdot \frac{\partial E}{\partial \boldsymbol{\nu}} \tag{17}$$

Here ω is an energy used to select which excited state is being targeted, $\boldsymbol{\nu}$ is the vector of variational parameters within the ansatz, and $\boldsymbol{\mu}$ is a vector of Lagrange multipliers by which we constrain the the minimization of L so that it must converge to an energy stationary point. In essence, the first term in L is a rigorous excited state variational principle with the energy eigenstate closest to ω as its global minimum, but because approximate ansatzes will prevent us from reaching this minimum, we add the energy stationarity constraint to ensure that at least this important property of exact excited states is maintained. In other words, the idea is for the first term to drive the optimization to the energy stationary point that best corresponds to the desired excited state. In preliminary work on ESMF,[1] it has been found that computationally tractable approximations to this Lagrangian

$$\widetilde{L} = (\omega - E)^2 - \boldsymbol{\mu} \cdot \frac{\partial E}{\partial \boldsymbol{\nu}} \tag{18}$$

are in practice effective at achieving the same goal, and so for expediency's sake we will adopt $\widetilde{L}$ as our working variational principle for DFE-ESMF.

Before augmenting the ESMF energy expression with density functional components, we should consider the choice of the wave function ansatz used in ESMF. To start, consider a singly excited configuration state function (CSF), which is perhaps the simplest spin-pure excited state ansatz.

$$|\Psi_i^a\rangle = \frac{1}{\sqrt{2}} \Big(a_a^+ a_i \, |\Psi_0\rangle \pm a_{\bar{a}}^+ a_{\bar{i}} \, |\Psi_0\rangle \Big) \tag{19}$$

This is a superposition between alpha ($i \to a$) and beta ($\bar{i} \to \bar{a}$) excitations from the closed

shell Slater determinant Ψ_0 in which the excitations both occur within the same pair of spatial orbitals $\{i, a\}$. The sign determines whether the excitation is a singlet or triplet, and although we will develop the mathematics for the singlet case below, the triplet is equally straightforward. As in the ground state presentation above, the (spin-restricted) molecular orbitals will be defined via an anti-Hermitian matrix $\boldsymbol{X}$ as in Eq. (13), but with the corresponding ground state KS-DFT orbitals now acting as the initial orbitals $\phi^{(0)}$ and $\boldsymbol{X}$ encoding excited-state-specific relaxations. This single-CSF ansatz leads to the electron density

$$n_{ia}(\mathbf{r}) = |\phi_a(\mathbf{r})|^2 - |\phi_i(\mathbf{r})|^2 + 2\sum_k^{N/2} |\phi_k(\mathbf{r})|^2 \tag{20}$$

and a one-body reduced density matrix (1RDM) $\boldsymbol{P}$ that is diagonal in the basis of the relaxed molecular orbitals.

$$\begin{aligned} P_{kj} &= 2\delta_{kj} - \delta_{ki}\delta_{ji} \\ P_{bc} &= \delta_{ba}\delta_{ca} \\ P_{jb} &= P_{bj} = 0 \end{aligned} \tag{21}$$

The ESMF kinetic energy can be computed as

$$T_{ia} = \left\langle \Psi_i^a | \hat{T} | \Psi_i^a \right\rangle = \mathrm{Tr}\left[\boldsymbol{P}_{AO}\boldsymbol{T}_{AO}\right] \tag{22}$$

where $\boldsymbol{P}_{AO}$ is the 1RDM rotated into the atomic orbital basis, and $\boldsymbol{T}_{AO}$ are the kinetic energy integrals in that basis. Likewise, the external potential contribution may be evaluated as

$$V_{ia}^{ext} = \mathrm{Tr}\left[\boldsymbol{P}_{AO}\boldsymbol{h}_{AO}\right] \tag{23}$$

where $\boldsymbol{h}_{AO}$ are the corresponding one-electron integrals.

Turning our attention now to the electron-electron repulsion energy, we start with the

Hartree term, which for this singlet CSF's density is

$$J[n_{ia}] = \frac{1}{2}(ii|ii) + \frac{1}{2}(aa|aa) - (aa|ii) + 2\sum_{kj}(kk|jj) + 2\sum_{j}\Big((aa|jj) - (ii|jj)\Big), \tag{24}$$

For hybrid functionals, we will need a definition for the wave-function-based exchange energy (i.e. an excited state analogue of HF exchange), which we choose to arrive at by making the usual index exchanges in the two-electron integrals of the corresponding Hartree term.

$$E_{ia}^{x(\mathrm{wfn})} = -\frac{1}{2}(ii|ii) - \frac{1}{2}(aa|aa) + (ai|ia) - \sum_{kj}(kj|jk) - \sum_{j}\Big((aj|ja) - (ij|ji)\Big) \tag{25}$$

Now, for the closed shell Slater determinant used in KS-DFT, the Hartree term (Equation 7) and the wave-function exchange term (Equation 15) sum to the electron-electron repulsion energy of the Slater determinant. However, things are not so simple in the excited state, and even for this single-CSF singlet wave function, the full wave-function-based electron-electron repulsion energy contains one additional term:

$$E_{ia}^{ee} \equiv \langle \Psi_i^a | \, \hat{V}_{ee} \, | \Psi_i^a \rangle = J[n_{ia}] + E_{ia}^{x(\mathrm{wfn})} + (ai|ia). \tag{26}$$

For the triplet CSF, we have a similar situation, but the additional term takes on the opposite sign:

$$E_{ia}^{ee(\mathrm{triplet})} = J[n_{ia}] + E_{ia}^{x(\mathrm{wfn})} - (ai|ia). \tag{27}$$

This extra term, which we will denote as the wave function correlation energy (WFCE)

$$E_{ia}^{\mathrm{WFCE}} = \pm(ai|ia), \tag{28}$$

determines the singlet-triplet splitting and arises from the fact that $\hat{V}_{ee}$ connects the two different terms in our CSF.

The energy of the wave function in Equation 19 is then the sum of the kinetic energy, the external potential, Hartree and exchange energy, and WFCE:

$$E_{ia}^{\mathrm{wfn}} = T_{ia} + V_{ia}^{ext} + J[n_{ia}] + E_{ia}^{x(\mathrm{wfn})} + E_{ia}^{\mathrm{WFCE}}. \tag{29}$$

In the same way that one can arrive at KS-DFT starting from HF by replacing the exchange term with an exchange correlation functional (which converts Equation 14 into Equation 12), we now replace the exchange term in our excited state wave function energy in order to arrive at the energy expression for single-CSF DFE-ESMF.

$$E_{ia} = T_{ia} + V_{ia}^{ext} + J[n_{ia}] + E_{ia}^{xc} + E_{ia}^{\mathrm{WFCE}} \tag{30}$$

Note especially that the WFCE term is retained. As this term originates from a strong correlation effect (the two electrons involved in the excitation are taking care to never be in the same orbital at the same time), we assume that it will not create significant double counting issues when used in conjunction with standard formulations of ground state exchange-correlation functionals, as these are geared towards weak correlation and are not designed to capture open-shell spin recoupling correlations. In ground state KS-DFT, $E^{\mathrm{WFCE}} = 0$ and HF is recovered by using a functional with no correlation and 100% HF exchange. The analogous property is maintained by DFE-ESMF: when using a functional consisting soley of 100% wave function exchange as defined in Equation 25, the DFE-ESMF energy reverts back to the ESMF expression for a single CSF's energy.

As in ESMF theory, the minimization of Eq. (18) requires the evaluation of certain sums over the second derivatives of our energy expression. Although the density functional energy expression of Eq. (30) differs from that of ESMF theory, we can exploit the same automatic differentiation (AD) approach in order to perform the optimization at a cost whose scaling with system size is the same as a ground state KS Fock build. For an explanation of how this is achieved, we refer the reader to the original ESMF paper.[1] As in that case, we have formulated our pilot code using the convenient AD capabilities of the TensorFlow framework[33] and have carried out the minimization via a quasi-Newton approach.[34] In addition to what is necessary for ESMF, this requires AD through the grid integration involved in density functional components such as the LDA exchange and correlation terms, which we have now achieved with the correct scaling.

2.3 DFE-ESMF: Multiple-CSF Formalism

In cases where a state contains major contributions from multiple different single excitations, we may generalize the approach into a multi-CSF form with a wave function similar to configuration interaction singles (CIS),[35]

$$|\Psi_{\mathrm{MCSF}}\rangle = \sum_{ia} c_{ia} \, |\Psi_i^a\rangle, \tag{31}$$

in which we still relax the orbitals as above. In this case, the density becomes

$$\begin{aligned} n_{\mathrm{MCSF}}(\mathbf{r}) = 4 \sum_{ia} |c_{ia}|^2 \sum_{k} |\phi_k(\mathbf{r})|^2 + 2 \sum_{iab} c_{ia} c_{ib} \phi_a(\mathbf{r}) \phi_b(\mathbf{r}) \\ - 2 \sum_{ija} c_{ia} c_{ja} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) \end{aligned} \tag{32}$$

and the 1RDM in the relaxed MO basis is no longer diagonal.

$$\begin{aligned} P_{ij} &= \delta_{ij} - \sum_a c_{ia} c_{ja} \\ P_{ia} &= P_{ai} = 0 \\ P_{ab} &= \sum_i c_{ia} c_{ib} \end{aligned} \tag{33}$$

Nonetheless, we can still take the KS approach and evaluate both the kinetic energy and external potential via the wave function's 1RDM using Eqs. (22) and (23).

Although the one-electron components are quite similar to the single-CSF approach, the electron-electron repulsion energy is less straightforward. In order to define the Hartree term, one possibility is to use the density from Eq. (32) in the standard $J[n]$ form of Eq. (7). However, doing so introduces unphysical virtual-virtual Coulomb repulsion terms in the form of $(aa|bb)$, similar to the ghost interactions encountered in ensemble DFT. In order to avoid these in the multi-CSF case, we generalize the Hartree term as the weighted statistical average of the Hartree terms from each separate CSF as given in Eq. (24).

$$J_{\mathrm{MCSF}} \equiv \sum_{ia} |c_{ia}|^2 J[n_{ia}] \tag{34}$$

If we now apply the index-exchange approach, we simply arrive at an "exact" wave function exchange that is the weighted average of the single-CSF pieces from Eq. (25).

$$E_{MCSF}^{x(\mathrm{wfn})} \equiv \sum_{ia} |c_{ia}|^2 E_{ia}^{x(\mathrm{wfn})} \tag{35}$$

As before, the Hartree and exchange pieces do not add up to the full wave function electron-electron repulsion energy,

$$\langle \Psi | \, \hat{V}_{ee} \, | \Psi \rangle = J_{\mathrm{MCSF}} + E_{\mathrm{MCSF}}^{x(\mathrm{wfn})} + E_{\mathrm{MCSF}}^{\mathrm{WFCE}}, \tag{36}$$

and the additional correlation effects are now more involved.

$$
\begin{aligned}
E^{\mathrm{WFCE}}_{\mathrm{MCSF}} &= 2\sum_{iajb} c_{ia}c_{jb}[2(ai|jb) - (ab|ji)] \\
&+ \sum_{abi} c_{ia}c_{jb}[\sum_{k} 4(ab|kk) - 2(ak|kb)] \\
&- \sum_{ija} c_{ia}c_{ja}[\sum_{k} 4(ij|kk) - 2(jk|ki)] \\
&+ \sum_{ia} |c_{ia}|^2 \sum_{k} [-4(aa|kk) + 2(ak|ka) + 4(ii|kk) - 2(ik|ki)] \\
&+ \sum_{ia} |c_{ia}|^2 [2(aa|ii) - 2(ai|ia)]
\end{aligned}
\tag{37}
$$

Using the same logic as before (although see Section 2.4 regarding double counting concerns), we define the multi-CSF density functional form for the energy in Eq. (18) to be

$$
E_{\mathrm{MCSF}} = T + V_{ext} + J_{\mathrm{MCSF}} + E_{xc} + E^{\mathrm{WFCE}}_{\mathrm{MCSF}} \tag{38}
$$

in which the T and V_{ext} are as in Eqs. (22) and (23) but with the multi-CSF 1RDM, and E_{xc} is as in the ground state functional but with the density taken from Eq. (32) and the wave function exchange component set to $E^{x\,(\mathrm{wfn})}_{\mathrm{MCSF}}$.

2.4 Double Counting Problems

The DFE-ESMF energy, in both the single and multiple CSF formalisms, contains correlation terms that do not exist in the energy expression of ground state DFT. However, one potential problem of adding these correlation terms into the energy formula as we have done is that, in principle, they could be accounted for again in the xc functional, leading to a double counting problem. In the single-CSF formalism, the WFCE term $(ai|ia)$ arises completely due to the fact that the wave function contains two determinants with equal weights. Such a strong correlation effect is (typically) not built in to practical forms for E_{xc} which instead aim to include weak correlation effects.[36] Therefore, we do not expect to have significant

double counting problems in the single-CSF case.

However, if one employs the full CIS-style multi-CSF formalism, the wave function definitely includes both some kinetic energy correlation effects and some electron-electron interaction correlation effects. In order to illustrate this, consider the case where the multi-CSF expansion is dominated by one CSF with an excitation between the ith and ath orbitals. We can treat this dominant piece as the zeroth-order reference in a perturbative expansion. As some other singly-excited CSFs are coupled to this reference by $\hat{T}$ and even more by $\hat{V}_{ee}$, such couplings would be part of any 2nd-order perturbation correction starting from this reference. Thus, a simple Moller-Plesset-style argument suggests that many and perhaps most of the contributions within $E_{\mathrm{MCSF}}^{\mathrm{WFCE}}$ would be part of the system's weak correlation physics and so at significant risk of double counting within our multi-CSF formalism. Indeed, in early testing, we found that excitation energies with the full multi-CSF formalism were worse than those from the single-CSF formalism, which we now understand was primarily a double counting issue.

In order to avoid this problem, one might try to separate contributions from the weak and strong correlations within the multi-CSF expansion. Although there is no unique way to do this, we have for now taken the expedient approach of including in our multi-CSF expansion only those CSFs whose TDDFT coefficients are above a threshold. While it may become clear once more data is available what the least-bad threshold choice would be, we have for now set this threshold at a relatively large value of 0.2 to help ensure that retained CSFs are playing a larger-than-perturbative role in the excitation and are therefore more likely to contribute energetic correlation effects of the type that are not built in to common density functionals. For simplicity, and in contrast to ESMF theory, we do not optimize these coefficients in our minimization of $\widetilde{L}$ and instead hold them fixed at their TDDFT values. Admittedly, such a 0.2 threshold will become troublesome in cases where the excited states are composed of a large collection of excitations, such as plasmons. Therefore, developing alternative approaches to avoid the double counting problem will be highly desired in future

developments of DFE-ESMF.

2.5 Discussion of DFE-ESMF

It is important to note that the DFE-ESMF method in its current form is not an excited state generalization of the ground state KS-DFT. Based on the Hohenberg-Kohn theorem,[3] which establishes a one-to-one mapping between external potential and density, the ground state energy depends solely on density. However, it has been shown[37] that such a one-to-one mapping between external potential and density does not exist for excited states. Therefore, the excited state density alone can not uniquely determine its energy. The simplest example would the singlet and triplet excited state of a given configuration. These two states have the same density, but different energies. In previous developments that try to generalize the ground state KS-DFT to excited states, Levy and Nagy use bi-functionals[38] that depends on both excited state and ground state density, and Görling uses totally symmetric part of the density[39] and a generalized adiabatic connection scheme,[40,41] in order to enforce the correct symmetry of excited states. Thus, we suggest that it is more useful to view DFE-ESMF as a practical extension to ESMF rather than as a formal density functional theory. That said, DFE-ESMF does share some similarities with the exact generalized adiabatic connection (GAC) approach.[41] Both methods use a symmetry-determined linear combination of Slater determinant to compute kinetic energy, external potential, and exchange energy. In addition, both methods try to enforce the correct excited state symmetry. In DFE-ESMF, the excited state symmetry is taken care by the WFCE term, while GAC uses the symmetrized density.[41]

In the context of spin symmetry, it is important to note that the WFCE term is essential for our optimization approach. Without this term, singlet and triplet excited states formulated by the same excitation have the same density and energy. Consequently, our energy-based excited state variational approach would not be able to distinguish these two states and its results would be arbitrary. As discussed before, the difference between singlet and triplet states is encoded in the WFCE term, and adding this term to the energy

expression greatly helps the optimization procedure to pick the desired state.

At present, DFE-ESMF uses functionals developed for ground state to treat excited states. While there is no reason to think that this approach is optimal, it is a very common procedure to treat excited states using ground state functionals. For example, the aforementioned GAC approach, the Δ self-consistent field (ΔSCF) method,[42] constricted variational density functional theory,[43] and spin-restricted ensemble-referenced Kohn-Sham method (REKS)[44] all use ground state functionals to describe excited states. While it may in future be worthwhile to develop functionals specifically for use with DFE-ESMF, we do not explore this direction here.

Although we do use ground state functionals, it is important to distinguish the present approach from the use of such functionals in TDDFT via the AA. First, the AA is a statement about the time dependence of the exchange correlation kernel, which has no direct analogue in DFE-ESMF, as it is a time-independent theory. Second, the AA, when combined with LR theory to produce practical versions of TDDFT, creates issues that are not present in DFE-ESMF, regardless of whether ground state functionals are employed. Most importantly, the AA prevents TDDFT from incorporating the effects of orbital relaxations for electrons not involved in the excitation. Due to its many-electron variational nature, DFE-ESMF explicitly includes these relaxations, in direct analogy to how ground state KS-DFT variationally relaxes all the electrons' orbitals. Thus, although both the AA and the current formulation of DFE-ESMF lead in practice to the use of ground state functionals for treating excited states, the approximations being made in these two approaches are distinct.

Finally, it is important to emphasize that state-specific formulations do come with limitations alongside their advantages. As for some other excited state specific DFT methods discussed in the next section, it is not obvious how to arrive at rigorous transition moments for DFE-ESMF. Although one could simply define these in terms of the underlying wave function and the ab initio Hamiltonian, this approach would miss the fact that the states have been optimized based on a DFT-modified energy expression, creating a disconnect be-

tween the evaluations of energy differences and transition strengths. Another issue with state specific methods is that they do not in general satisfy known sum rules and sum-over-state expressions.[3] Thus, although the approach pursued here possesses some formal advantages when compared to TDDFT, it also suffers from some formal disadvantages.

2.6 Comparisons to Other State-Specific DFT Methods

DFE-ESMF is not the first attempt to combine wave function based methods with density functionals. For example, multi-reference (MR) DFT methods such as multiconfiguration Pair-Density Functional Theory (MC-PDFT)[45,46] and density matrix renormalization group pair-density functional theory (DMRG-PDFT)[47] also modify a wave function's energy expression by using an xc energy functional to capture correlation effects. A major difference between these methods and DFE-ESMF is that they target strong correlation in ground states, whereas DFE-ESMF targets weak correlation in singly excited states. Another difference is that, because CASSCF and DMRG-SCF wave functions already incorporate state-specific orbital relaxations, the DFT part of their methodology need not address the orbital shapes. The central feature of DFE-ESMF, on the other hand, is its excited-state-specific orbital relaxation. Finally, these MR-DFT approaches use the on-top pair density functional,[48] which is more capable of addressing strong correlation issues.

With regards to variational DFT methods for excited states, many approaches distinct from DFE-ESMF already exist. The Δ self-consistent field (ΔSCF) approach[42] relaxes excited state orbitals by using the SCF cycle in an attempt to converge onto open-shell solutions to KS equations, employing the maximum overlap method (MOM) to help avoid collapsing back to the ground state or to lower-lying excited states.[49,50] The related restricted open-shell Kohn-Sham (ROKS) method[51] may also collapse to lower excited states, but its enforced open-shell nature prevents collapse to the ground state and it has shown advantages relative to ΔSCF for CT excitations' singlet-triplet splittings.[52] Finally, ensemble DFT in the form of REKS and SA-REKS optimizes excited state orbitals in a state-averaged manner,

trading some state-specificity in return for a further reduction in the risk of variational collapse. In contrast to these approaches, the DFE-ESMF approach makes direct use of an excited state variational principle. Although this does not rigorously guarantee that the correct stationary point will be found (all of these variational methods are nonlinear minimizations with at least some starting point dependence, after all) the global minimum of the variational principle it employs is the desired excited state, offering a strong formal advantage in the effort to avoid collapse to lower states. In our preliminary explorations, we have yet to encounter a case where the optimization does not converge to the stationary point corresponding to the targeted excited state, even in cases where ΔSCF encounters variational collapse. It is also worth noting that, although the multi-CSF version of DFE-ESMF comes with double counting concerns, it can at least be applied to states that strongly mix two or more excitation components, while ΔSCF, ROKS, and REKS all assume that excitations are single-component in nature.

The constrained DFT (CDFT) method[53] represents another route towards excited state orbital relaxation that is especially relevant for long range CT, where it is straightforward to impose physically motivated density constraints in cases where the donor and acceptor can be clearly identified. As shown by numerous applications, CDFT can provide accurate estimates of excitation energies,[54] coupling elements,[55] forces,[56] and diabatic surfaces.[57] A particularly strong parallel with DFE-EMSF can be seen in long range CT, where single-CSF DFE-ESMF is expected to be equivalent to CDFT in the limit of complete donor-acceptor separation (see for example Figure S1). To understand this equivalence, consider that both methods will move an electron from the donor's HOMO to the acceptor's LUMO and then make their energy expression stationary with respect to orbital rotations. As the WFCE term in DFE-ESMF vanishes in the limit of long range CT, the two methods will have the same energy expression in this case and so will produce the same results. In shorter-ranged CT where donor and acceptor are less well defined, DFE-ESMF has the formal advantage of not having to impose a user-specified charge constraint, and so can in principle predict

the distribution of the particle and hole rather than having it imposed from some external source. DFE-ESMF also avoids having to worry about the ambiguities inherent to assigning formal atomic charges and the difficulties these create.[53]

3 Results

3.1 Computational Details

To assess the performance of DFE-ESMF and to compare to existing excited state DFT methodologies, we have carried out tests in the following atomic and molecular excitations:

1) singlet and triplet $n \rightarrow \sigma^*$ in H_2O.

2) singlet and triplet $n \rightarrow \pi^*$ and singlet $\pi \rightarrow \pi^*$ in CH_2O.

3) singlet and triplet $\sigma \rightarrow \sigma^*$ in LiH.

4) singlet $\pi \rightarrow \pi^*$ in CO.

5) singlet He 1s$\rightarrow$Be 2p in He-Be dimer.

6) singlet NH_3 2pz$\rightarrow$$F_2$ 2pz in NH_3-F_2 dimer.

7) singlet 2s$\rightarrow$3s and 2p$\rightarrow$3p in the Ne atom.

All of the DFE-ESMF results are obtained via our own pilot code, which extracts one- and two-electron integrals from PySCF.[58] The Lebdev-Laikov grid[59] is used to perform the numerical integration to compute the xc energy. The TDDFT, CIS, ROKS, and ΔSCF DFT results were obtained from QChem.[60] In ROKS the spin purification techniques[61] is used, and in ΔSCF DFT we report the broken-spin open-shell energy. Equation-of-Motion Coupled Cluster with singles and doubles (EOM-CCSD) results were computed by MOLPRO.[62] It is also worth noting that the implementation of ROKS in QChem is limited to only HOMO$\rightarrow$LUMO excitations. In the current study, the CSF expansions in both the single-CSF formalism and multiple-CSF formalism are selected by the CI vector of TDDFT using the same xc functional. We choose a large threshold of $\epsilon = 0.2$ for CSF truncation, and switch to the multi-CSF formalism of DFE-ESMF when there are more than one CSF

left after truncation.

For DFE-ESMF, we employ three xc functionals: LDA, the Becke3-Lee-Yang-Parr functional (B3LYP),[63–65] and the Becke-Half-Half functional (BHHLYP),[66] which have 0%, 20%, and 50% wave function exchange fractions, respectively. We expect results to be somewhat sensitive to this fraction, as ground and excited states have different amounts of open shell character and thus are likely to suffer from differing degrees of self-interaction error. As existing functionals have mostly been optimized for closed shell ground states, it would not be surprising if a higher than usual wave function exchange fraction was necessary to create a fair playing field for the open-shell state. Note that our excitation energies come from energy differences between DFE-ESMF (for the excited state) and KS-DFT (for the ground state) in which both have used the same xc functional.

For basis sets, we employed the cc-pVDZ basis[67] for H_2O, CH_2O, LiH, and CO, the cc-pVTZ basis[68] for the He-Be dimer, the aug-cc-pVTZ basis[68] for Ne, and the 6-31G basis[69] for the NH_3-F_2 dimer. Results will be presented in terms of excitation energy errors relative to EOM-CCSD. The molecular geometries and absolute values of excitation energies can be found in the Supplemental Information.

3.2 Excited State Dipole Shifts

Before presenting the results of DFE-ESMF, we first categorize the excited states into CT and non-CT types by computing the difference between the ground and excited state dipole moment. In atomic units, the dipole moment is

$$\vec{\mu} = \sum_A Z_A \mathbf{R}_A - \int \mathbf{r}\, n(\mathbf{r}) d\mathbf{r} \tag{39}$$

in which Z_A and $\mathbf{R}_A$ are the charge and position of the Ath nuclei. For simplicity, the electron density for excited states is estimated using Equation 20 using ground state KS orbitals without any relaxation. The dipole moment difference ($—\Delta\vec{\mu}—$) between the ground

and excited state yields information about the electron charge distribution between these two states.

The computed —$\Delta\vec{\mu}$—s are shown in Figure 1. For Ne, H_2O, and the $n \rightarrow \pi^*$ state in CH_2O, the —$\Delta\vec{\mu}$— are fairly small, indicating that the charge distributions are similar for the ground and excited states. These states can thus be viewed as purely Rydberg (Ne) and valence (H_2O, CH_2O) excited states with little charge deformation. As expected, the long-range CT excited states[18,70] of He-Be and NH_3-F_2 have a significantly larger —$\Delta\vec{\mu}$—. Note that LiH also sees significant charge deformations, which are a consequence of its partially ionic nature in which the bonding and anti-bonding orbitals are shifted towards opposite ends of the molecule.

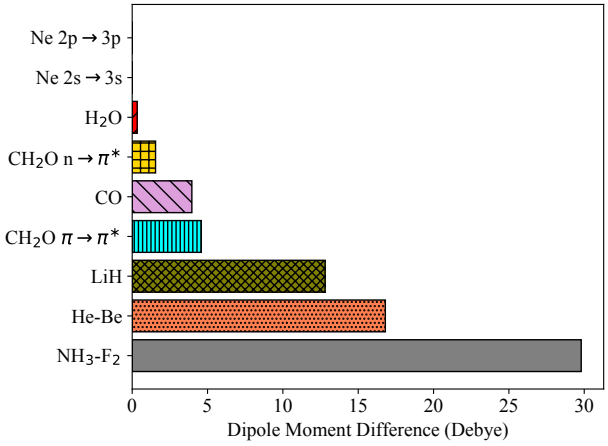

Figure 1: The norms of the dipole moment differences between ground and excited states. The BHHLYP functional is used in all dipole calculations.

3.3 Single-CSF Excited States

Let us begin with an analysis of how the theory performs for excited states dominated by a single CSF, looking at different types of excitations within this category. We will look first at valence excitations, followed by CT states and finally Rydberg excitations. These cases covered, we will then consider states that contain a superposition of multiple CSFs.

3.3.1 Valence Excitations

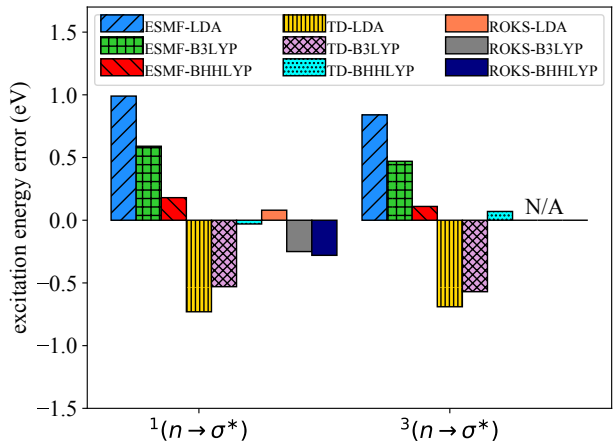

Figure 2: The excitation energy error of singlet (left) and triplet (right) $n \to \sigma^*$ excited states in H_2O compared to EOM-CCSD results.

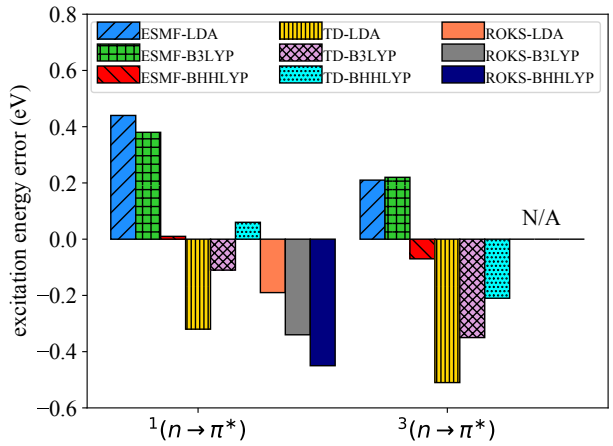

Figure 3: The excitation energy error of singlet (left) and triplet (right) $n \to \pi^*$ excited states in CH_2O compared to EOM-CCSD results.

Excitation energy errors relative to EOM-CCSD for H_2O and CH_2O are shown in Figures 2 and 3, respectively. Compared to the CT examples in the next section, TDDFT performs relatively well for these valence excitations, with B3LYP and especially BHHLYP providing excitation energies within about half an eV of the reference and LDA performing only a little worse. These relatively good TDDFT results are not particularly surprising in light of the

fact that the charge density deformations are small, and so the lack of orbital relaxation due to the AA is not especially concerning. In fact, we have explicitly analyzed the importance of orbital relaxation by evaluating the Frobenius norm of the DFE-ESMF orbital rotation matrix $\boldsymbol{X}$. Averaging over the three values from the three different functionals tested, we find $||\boldsymbol{X}||$ to be 0.16 and 0.15 for H_2O and CH_2O, respectively, which is smaller than in the CT examples we will see below.

The basic trend in DFE-ESMF accuracies for different functionals follows that of TDDFT in these excitations, with BHHLYP giving the most accurate predictions, followed by B3LYP and then LDA. As we will see, the accuracy ordering for different functionals in DFE-ESMF plays out is the same way in most of our test systems, with BHHLYP's high fraction of wave function exchange outperforming the other two functionals in valence, CT, and Rydberg states. Our understanding of this trend is that a larger fraction of wave function exchange is most likely helping to balance self-interaction errors in the ground and excited states. We expect that these errors are larger in the excited states due to their open-shell nature, and so a higher fraction of wave function exchange than is typically used in ground state models appears to be helpful for balancing these errors between the ground and excited states. We should emphasize that in all of our DFE-ESMF results, energy differences were evaluated based on the same functional for both ground and excited states, but using the density and wave function exchange definition for the state in question (see discussion surrounding Eq. (30)). Although this preliminary test of four single-CSF valence excitations is far from exhaustive or systematic, the efficacy of DFE-ESMF/BHHLYP in these cases provides an encouraging proof of principle.

For singlet excited states, we also compare to the predictions of ROKS, which, like DFE-ESMF, provides excited state orbital relaxation. The most striking difference between the ROKS results and those of TDDFT and DFE-ESMF is that they do not follow the same trend with respect to the fraction of wave function exchange. Indeed, the accuracy ordering differs in H_2O and CH_2O, with LDA seeming to offer the best average ROKS performance.

Note that we have not carried out ROKS comparisons in the triplet states because QChem currently only implements ROKS for singlet states.

3.3.2 Charge-Transfer Excitations

Although DFE-ESMF and TDDFT provide similar accuracies in the simple valence excitations discussed above, the story is very different for CT excitations. Before looking in detail at the numerical CT examples, it is worth considering two important potential sources of error TDDFT faces in CT contexts, which we will refer to as the EA/IP imbalance and the orbital relaxation error. To see these clearly, consider a simple CT excitation consisting of a single $i \to a$ transition, in which case the TDDFT excitation energy is given by[71]

$$\Delta E(i \to a) = \epsilon_a^{KS} - \epsilon_i^{KS} + \langle ia|f_{xc}|ia\rangle \tag{40}$$

in which ϵ_i^{KS} and ϵ_a^{KS} are the ground state KS orbital energies of the donor and acceptor orbitals, respectively. This equation has been used extensively[10,16,25,26] to analyze the TDDFT's failure in CT excited states and we refer readers to those references for details. In a nutshell, because ϵ_a^{KS} does not correspond to the EA of the acceptor (it undercounts the new repulsions created by the extra electron), the orbital energy difference in this equation tends to severely underestimate CT excitation energies. In principle this should be repaired by the xc term, and it has been shown that a fully non-adiabatic frequency-dependent xc kernel, such as the time-dependent exact exchange (TDEXX),[72] is able to produce the necessary corrections.[25] However, once the AA is applied, the loss of frequency dependence in the kernel leads to corrections that are too small with existing functionals, even RSHs.

While the EA/IP imbalance is a significant concern, it is typically offset in practice by the fact that AA TDDFT works with unrelaxed orbitals. It is well known in electronic structure theory that the omission of orbital relaxation effects tends to raise the energy of the excited state. For example, configuration interaction singles (CIS), which is free from the IP/EA

imbalance problem because of Koopmann's theorem, usually overestimates the CT excitation energy by more than 1eV due to the lack of orbital relaxation.[73] Ideally, the third term in Equation 40 would eliminate both orbital relaxation issues and the EA/IP imbalance, but even when the third term is zero, these two errors do at least work to cancel each other because they push in opposite directions. However, the EA/IP imbalance in long-range CT is often much too large for orbital relaxation errors to counteract, resulting in TDDFT excitation energies that are much too low as in the $NH_3 \rightarrow F_2$ and He→Be transitions shown below. However, as one moves to increasingly shorter range CT with correspondingly larger overlaps between the donor and acceptor, the EA/IP imbalance becomes less and less of an error and more and more a positive feature of the TDDFT formalism. Indeed, in the valence excitation limit, the difference in how DFT accounts for electron-electron repulsion energy in the occupied and virtual orbitals increases the accuracy of using orbital energy differences as excitation energy estimates, because in this limit the "donor" and "acceptor" are one and the same. One can imagine that for very short-ranged CT, any small remaining errors from the EA/IP imbalance could cancel with orbital relaxation errors precisely enough for the exchange correlation term to clean up the details. Such an effect seems to be at work in LiH, to which we now turn our attention, which despite having a substantial dipole change and thus CT is nonetheless treated well by TDDFT.

LiH Figure 4 shows excitation energy errors for the lowest singlet and triplet excitations in LiH. For both of these states, the balancing process between EA/IP issues and missing orbital relaxations appears to work in TDDFT's favor, especially in the case of the BHHLYP functional. To check that such a trade off really does appear to be at work here, we again evaluated $||\boldsymbol{X}||$ as a measure of orbital relaxation importance and found it to be 0.41, significantly higher than for H_2O or CH_2O. As in the valence states of those molecules, BHHLYP is also very effective in DFE-ESMF's single-CSF formalism for LiH's singlet excitation. For the triplet, however, the single-CSF formalism shows relatively poor accuracy regardless of

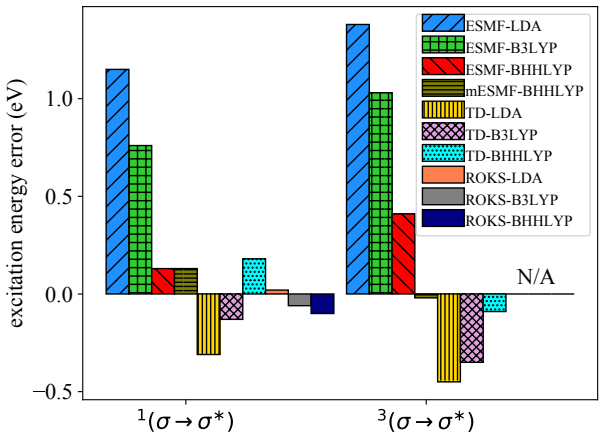

Figure 4: The excitation energy error of singlet (left) and triplet (right) $\sigma \rightarrow \sigma^*$ excited states in LiH compared to EOM-CCSD results.

functional, and indeed this state has more than one CSF above our $\epsilon = 0.2$ threshold in TDDFT. When we include both of the CSFs whose coefficients breach this threshold via the multi-CSF approach, the DFE-ESMF/BHHLYP result improves from an error above 0.4 eV to an error of just -0.02 eV relative to EOM-CCSD. Note that the multi-CSF approach has no effect on the singlet state, as in that case only the primary CSF was above the threshold.

NH_3 to F_2 We now turn our attention to the first of two long-range CT excitations: the NH_3-F_2 dimer shown in Figure S1. In Figure 5, we see that, after excited-state orbital relaxation via the minimization of Eq. (18), DFE-ESMF is substantially more accurate than TDDFT regardless of the functionals chosen. Even when comparing ESMF-LDA against TDDFT with the ωB97X RSH, the variational approach makes an excitation energy error roughly half as large. Using BHHLYP, which continues to outperform the others for DFE-ESMF, the variational approach achieves an excitation energy error of just 0.26 eV, compared to multi-eV errors for TDDFT when using either ωB97X or BHHLYP.

Although this excited state is predominantly a HOMO→LUMO excitation, we find that ROKS consistently collapses to a lower excited state (see the Supplemental Information for absolute excitation energies). Hence, we compare our results to MOM ΔSCF instead.

We found that ΔSCF DFT yields similarly accurate predictions as DFE-ESMF. This is as expected since ΔSCF DFT are known to perform well in long-range CT excited states, if the states are dominated by one CSF.

It would appear that TDDFT's error cancellation between its EA/IP imbalance and its lack of orbital relaxations breaks down here, with the magnitude of the former overwhelming that of the latter. We can verify this picture in two ways: first, with the DFE-ESMF approach, and second, by looking at ground state KS-DFT IP-EA estimates at very long range. Start with DFE-ESMF. At the top of Figure 5, we show the excitation energies (i.e. the energy differences between Eq. (30) and the ground state KS energy) *before* the DFE-ESMF orbital optimization has been carried out, meaning that the excited state DFE-ESMF energy is being evaluated using the ground state KS orbitals. This excitation energy is thus the difference between two many-electron DFT energies (one DFE-ESMF and one KS-DFT) and so does not suffer from the EA/IP imbalance. While the EA/IP issue has thus been removed, orbital relaxation effects have yet to be included, and as expected the excitation energies are now too large. When we then relax the orbitals ($||\boldsymbol{X}|| = 0.21$), we see in the middle of Figure 5 that the predicted excitation energies decrease to more accurate values. Thus, by looking step-wise at how DFE-ESMF changes the energy from TDDFT, we can watch the staged removal of first the EA/IP imbalance and then the fixed-orbital error. This process appears to confirm the idea that TDDFT suffers from both, and that in long-range CT the EA/IP part dominates, leading TDDFT to underestimate the excitation energy.

We can corroborate this view by moving the molecules to a very large distance and comparing DFE-ESMF, TDDFT, and a simple difference of ground state KS energies between the cation, anion, and neutral species that provides a many-electron evaluation of the IP and EA. At very long distance, Figure 6 shows that the performance of DFE-ESMF and TDDFT is quite similar to what we saw at the shorter separation. At the bottom of the figure, we see that if we simply perform four single-molecule ground state KS calculations for the donor cation, acceptor anion, and the two neutral species, the resulting difference

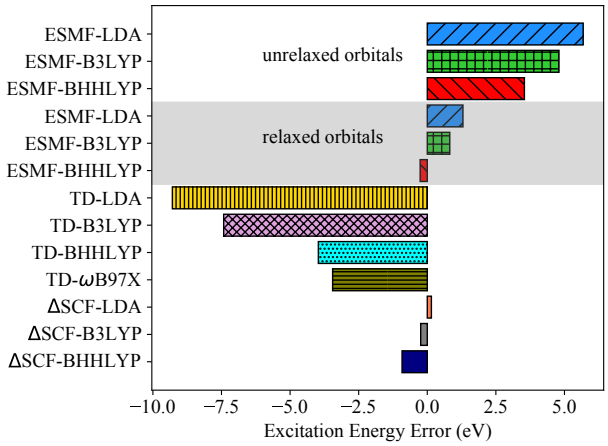

Figure 5: Comparison of excitation energy errors relative to EOM-CCSD for the NH_3 2pz $\rightarrow F_2$ 2pz CT excitation at an intermolecular separation of 6 $\mathring{A}$. For DFE-ESMF, we show the results both before and after the orbital relaxation is performed.

between the IP and EA is a very accurate predictor of the charge transfer energy, as we would expect from previous work on CDFT. Thus, if both the IP/EA imbalance born of single-particle orbital energy differences and the orbital relaxation errors are removed via this ground state KS approach or by CDFT, accuracy is restored. The key point is that, unlike these approaches, the DFE-ESMF formalism should allow both of these issues to be addressed even in systems where clear-cut foreknowledge distinguishing between the donor and acceptor is not available, and without having to worry about the spin-symmetry breaking inherent to ΔSCF.

He to Be In our second long-range CT example, we investigate the excitation from the He 1s orbital to the Be 2pz orbital as a function of the distance between the atoms. In Figure 7, we see the familiar failure of local functionals and simple hybrids to predict the correct $1/R$ trend in the excitation energy. While this problem is repaired by the use of a RSH, we see that absolute accuracy is still poor, at least for the specific ωB97X functional we tested here. As in the previous example, DFE-ESMF out-performs the accuracy of the RSH regardless of which functional it is paired with while also correctly capturing the $1/R$ behavior. The advantage of DFE-ESMF becomes especially clear when looking at the non-

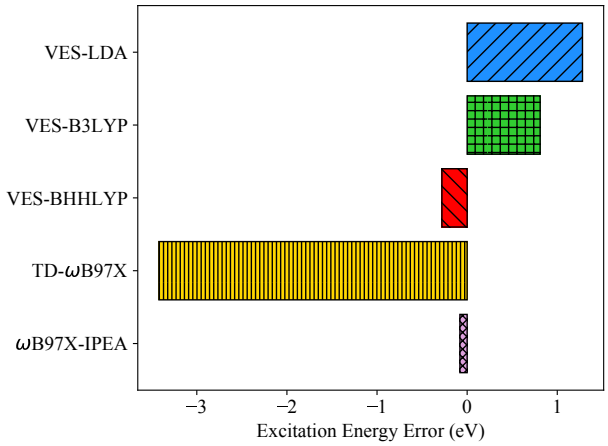

Figure 6: Excitation energy errors relative to EOM-CCSD for the NH_3 2pz $\rightarrow F_2$ 2pz CT at a 120 $\AA$ intermolecular separation. ωB97X-IPEA refers to the difference between (a) the sum of ground state KS energies for the donor cation and acceptor anion and (b) the ground state KS energy of the neutral ground states.

parallelity error (NPE) plotted in Figure 8. NEP is defined as the difference between the largest and smallest errors relative to EOM-CCSD across the distance coordinate. DFE-ESMF with LDA, B3LYP, and BHHLYP all produce NPEs below 1eV, as compared to an NPE of almost 2 eV for TDDFT with the ωB97X functional. In Figure 8 we also plot the shifted root-mean-square (RMS) error. Shifted RMS error is calculated by first shifting the PES in Figure 7 so that the excitation energy at 8 $\AA$ matches the prediction of EOM-CCSD, and then computing the RMS error along the shifted PES. Unlike NPE, shifted RMS error is less affected by the repulsive regime and we find it to be significantly smaller than NPE for all methods. Even so the shifted RMS error of TDDFT with the ωB97X functional (0.29 eV) is still much larger than that of DF-ESMF, which is around 0.1 eV.

While ΔSCF produces a visibly not-smooth curve when when paired with the LDA functional (possibly due to variational collapse issues), its performance with B3LYP and BHHLYP is quite good. In the long-range limit these potential curves overlap that of EOM-CCSD and their shifted RMS error is comparable to DF-ESMF, although accuracy is a bit lower at shorter ranges where the broken spin symmetry is expected to matter more. Consequently, the NPEs of ΔSCF with the hybrid functionals are a bit larger than those of

DFE-ESMF, but a major improvement over TDDFT.

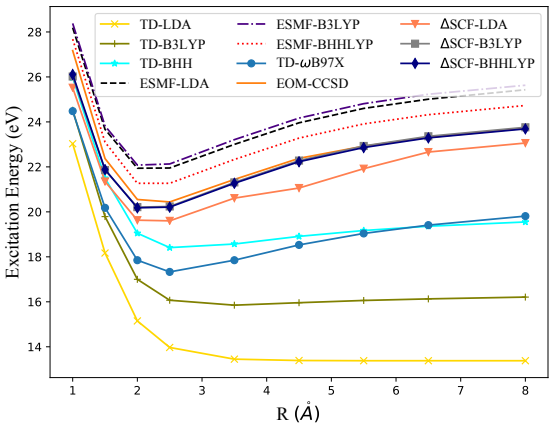

Figure 7: Comparison of He 1s→Be 2p CT excitation energy as a function of R(He-Be) between DFE-ESMF, TDDFT, ΔSCF, and EOM-CCSD.

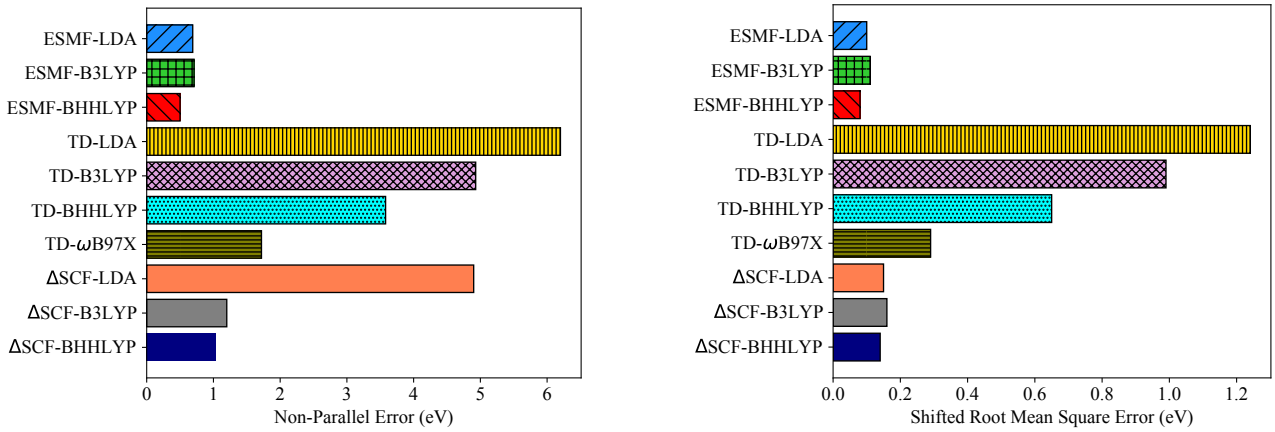

Figure 8: Left: Comparison of non-parallelity errors in the He 1s→Be 2p CT excitation energy. Right: Comparison of shifted root-mean-square errors in the He 1s→Be 2p CT excitation energy.

3.3.3 Rydberg Excitations

Unlike in CT excitations, it is not obvious that the ability of DFE-ESMF to address the EA/IP imbalance and orbital relaxations will be of great benefit in the context of Rydberg excitations. In these states, the challenge faced by TDDFT arises primarily from the failure

of practical xc functionals to produce a potential that decays as $1/r$ at large distances, which is not the same as the failure of error cancellation that causes trouble in the CT case. However, work by Van Voorhis[74] has shown that although the ground state xc potential has little resemblance to the exact potential, the xc potential associated with an excited state density can behave much more sensibly at long distance. Thus, it is interesting to ask whether DFE-ESMF's inherently excited state nature and its ability to relax the orbitals in an excited-state-specific manner may in practice lead to improvements for Rydberg states.

As an initial probe of this question, we have studied the 2s→3s excitation in the Ne atom. The excitation energy error relative to EOM-CCSD is plotted in Figure 9. As expected, TDDFT drastically underestimates the excitation energy by as large as 8.21eV using LDA and 2.94eV using BHHLYP. Notably, although our DFE-ESMF method is able to reduce the error of TDDFT by some amount, it is still very far from being quantitatively accurate. The most accurate functional in DFE-ESMF, the BHHLYP functional, still underestimates the excitation energy by 2.74eV. Although there is no dipole change in this excitation, the charge deformation in Rydberg states is still large since the virtual orbitals are much more diffuse than and share little overlap with the occupied orbitals. Consequently, the averaged $||\boldsymbol{X}||$ is as large as 0.25, comparable to that of CT excitations, indicating that orbital relaxation is also important in Rydberg excitations.

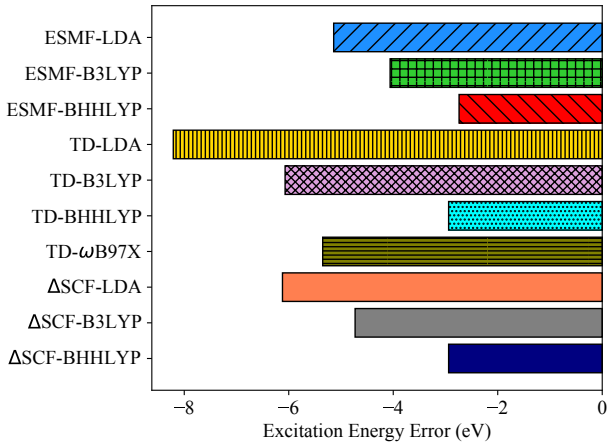

Figure 9: Comparison of excitation energy error of the Ne 2s → 3s transition.

As discussed in the previous subsection, the main source of error in DFE-ESMF is the self-interaction error with approximate xc functionals: The functional parametrized for the ground state is incapable of correcting the self-interaction error of the excited electron residing in virtual orbitals. This problem is not too concerning in valence excitations since the occupied and virtual orbitals have similar characters, and hence similar amount of self-interaction. Therefore thanks to error cancellation, balanced description between the ground and the excited states can still be achieved. However, in Rydberg states, the virtual orbitals are much more diffuse than the occupied orbitals. Consequently, the amount of SIEs left in $J[n]$ after adding the UEG exchange energy $E_x^{UEG}[n]$ to it, as in LDA, becomes clearly different in the occupied and virtual orbitals. This results in an unbalanced treatment between the ground and the excited state if one uses the same xc functionals for both states, leading to a large error in this Rydberg excitation energy. In future, testing asymptotically corrected functionals in DFE-ESMF would appear to be warranted.

The unbalanced treatment of SIEs should not be exclusive to DFE-ESMF and could affect ΔSCF DFT as well, since it also uses ground state functionals to describe excited states. In fact, in Figure 9 we found that the size and trend of errors of ΔSCF DFT with different xc functionals is quite comparable to DFE-ESMF, corroborating our speculations. In this case, the recovery of spin-symmetry in DFE-ESMF does not appear to have had much effect on the excitation energy.

3.4 Multi-CSF Excited States

We now turn our attention to states in which multiple CSFs are important. While DFE-ESMF can be formulated to treat such states and thus may be expected to offer a significant advantage over single-determinant methods such as ΔSCF, such optimism must be tempered by the increased risk of making double counting errors and the complication of needing a way to choose the values of the CI coefficients. As we see in the following subsections, these issues result in the present formulation of DFE-ESMF being less effective in the multi-CSF

case than it was in the single-CSF cases discussed above.

CH_2O The $\pi \rightarrow \pi^*$ excited state in CH_2O contains significant contributions from two excited CSFs: HOMO-1→LUMO and HOMO→LUMO+2. In contrast to LiH, where TDDFT only predicts multiple important CSFs for the BHHLYP functional, this CH_2O excitation is strongly multi-CSF in TDDFT regardless of the choice of funcitonal. Excitation energy errors relative to EOM-CCSD are plotted in Figure 10, where the ΔSCF results have been obtained by optimizing the orbitals for the open-shell determinant that has the largest weight in TDDFT's CI vector. Curiously, both TDDFT and DFE-ESMF give the most accurate prediction with LDA in this case, with DFE-ESMF showing especially poor accuracy with BHHLYP. This sharp contrast to the trends we saw in the single-CSF cases is explained, at least in part, by the strong dependence of the CI coefficients on the choice of functional. For LDA and B3LYP, the HOMO→LUMO+2 CSF has a larger weight that the HOMO-1→LUMO CSF, whereas the relative importance of these CSFs is reversed when using BHHLYP. Noting that ΔSCF places the energy of the orbital-optimized HOMO-1→LUMO CSF 1.86eV below the energy of the orbital-optimized HOMO→LUMO+2 CSF, we can understand much of DFE-ESMF/BHHLYP's lowering of the excitation energy simply in terms of the change in the CI coefficients coming from TDDFT. It therefore appears that, in future, re-optimizing the CI coefficients within the DFE-ESMF framework may be important.

CO In the $\pi \rightarrow \pi^*$ excitation in CO, spatial symmetry ensures that the p_x π orbital is degenerate with its p_y counterpart, and likewise for the two π^* orbitals, so that the excited state is an equal mixture of the $\pi_x \rightarrow \pi_x^*$ and $\pi_y \rightarrow \pi_y^*$ CSFs. The excitation energy errors for this state are shown in Figure 11, where we have used our multi-CSF formalism to treat this two-CSF state. In this case, the DFE-ESMF trend is back in line with what we saw in most other states, with BHHLYP providing the most accurate result, although TDDFT is notably more accurate than DFE-ESMF. The trend for ΔSCF is quite different, and it faces multiple difficulties here, including the breaking of spin and spatial symmetry as well

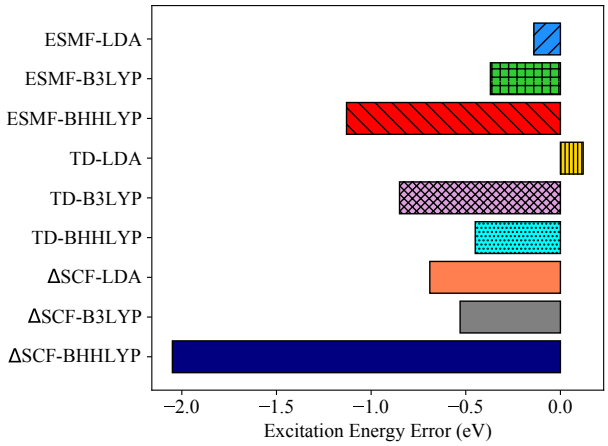

Figure 10: The excitation energy error of singlet $\pi \rightarrow \pi^*$ excited states in CH_2O compared to EOM-CCSD results.

as variational collapse to the ground state with BHHLYP, which explains the unusually large error in that case.

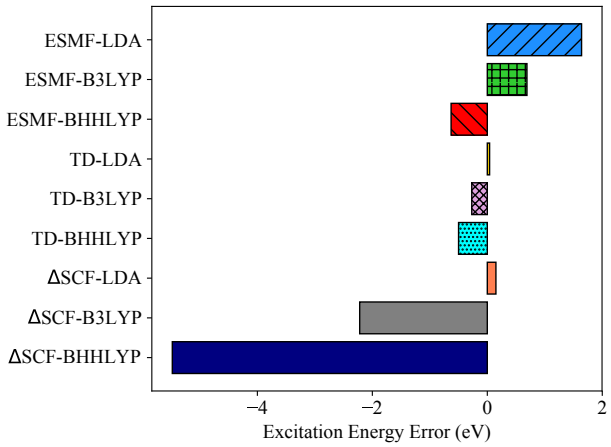

Figure 11: The excitation energy errors relative to EOM-CCSD for the CO $\pi \rightarrow \pi^*$ singlet excitation. Energy minimization of Δ SCF with BHHLYP functional ends up collapsing back to ground state.

Ne Unlike its 2s→3s excitation, the Ne atom's 2p→3p excited state is a equal mixture of two CSFs. In the plot of excitation energy errors (Figure 12) we see that DFE-ESMF and TDDFT have similar accuracies in this case, although the former errors high while the latter errors low. Again, we see the tendency of DFE-ESMF to benefit from a relatively

high fraction of wave function exchange. Unlike the other two multi-CSF cases discussed above, ΔSCF proves considerably more accurate for this excitation than either DFE-ESMF or TDDFT.

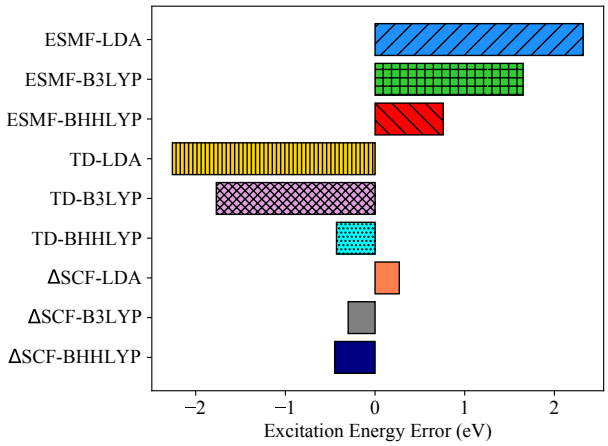

Figure 12: Comparison of excitation energy errors relative to EOM-CCSD for the Ne 2p → 3p transition.

4 Conclusions

We have presented a density functional extension to the recently developed excited state mean-field theory in an attempt to capture the effects of weak electron correlation. By augmenting the ESMF wave function's energy expression with components from density functional theory and inserting the resulting expression into an approximate excited variational principle, the approach provides excited-state-specific orbital optimization in the presence of a correlation treatment. In the same way that KS-DFT loosely parallels many aspects of HF theory, this DFE-ESMF approach parallels ESMF theory. Being a variational, time-independent approach, the method differs from TDDFT in some important aspects, most notably in its ability to fully relax orbitals and in the fact that it does not depend on the ground state KS orbital eigenvalues and so avoids concerns related to the EA/IP imbalance. In preliminary testing, these advantages result in significantly improved excitation

energies for simple charge transfer examples, while accuracy for other excitations is more comparable to TDDFT. Compared to ROKS and ΔSCF, which can also deliver full orbital relaxation, DFE-ESMF is less prone to variational collapse thanks to its use of an excited state variational principle.

Although these advantages and strengths are promising, the current formulation of DFE-ESMF also has a number of shortcomings. For excited states in which multiple CSFs make major contributions, double counting becomes a serious concern and accuracy, although still close to that of TDDFT, is reduced. This challenge is especially concerning in the context of larger systems, where multi-CSF states become increasingly common. Rydberg states are also challenging, although this may have more to do with the choice of density functional than with the DFE-ESMF approach itself: both it and TDDFT make substantial errors in these states. Looking forward, we therefore see a number of directions for possible improvement. As we saw in all cases (except for the $\pi \rightarrow \pi^*$ transition in formaldehyde where the sensitivity of TDDFT's CI coefficients to the fraction of wave function exchange appears to be the culprit) DFE-ESMF was most accurate when using a much higher fraction of wave function exchange (50%) than is typically found in ground state functionals. This is not surprising given the open shell nature of excited states, and suggests that a straightforward route to improved energetics may come from retraining functionals in this direction. Of course, this is just one question in a much broader array of functional design issues, such as whether it would be advantageous to include range-separation or asymptotic corrections, the latter of which may help with the Rydberg problem. Functional design questions aside, our observation that accuracy tends to be lower in multi-CSF states raises the question of whether this is due to relying on TDDFT CI coefficients or comes from some other source. If the former, then variational re-optimization of the CI coefficients alongside the orbitals may be advantageous. Finally, as a practical issue, our pilot implementation achieves the desired cost scaling (equivalent to ground state KS-DFT) but could be sped up considerably if implemented in a production level quantum chemistry package.

Acknowledgement

This work was supported by the National Science Foundation's CAREER program under Award Number 1848012. Calculations were performed using the Berkeley Research Computing Savio cluster. We also thank Diptarka Hait for helpful discussions.

References

(1) Shea, J. A.; Neuscamman, E. Communication: A mean field platform for excited state quantum chemistry. *J. Chem. Phys.* **2018**, *149*.

(2) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications: Mineola, N.Y., 1996.

(3) Parr, R. G.; Yang, W. *Density-functional theory of atoms and molecules*; Oxford University Press: New York, 1989.

(4) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(5) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997.

(6) Casida, M.; Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Annu. Rev. Phys. Chem.* **2012**, *63*, 287–323.

(7) Ullrich, C. A.; Yang, Z.-h. A brief compendium of time-dependent density-functional theory. *Braz. J. Phys.* **2014**, *44*, 154–188.

(8) Burke, K.; Werschnik, J.; Gross, E. K. U. Time-dependent density functional theory: Past, present, and future. *J. Chem. Phys.* **2005**, *123*, 062206.

(9) Petersilka, M.; Grossmann, U. J.; Gross, E. K. U. Excitation Energies from Time-Dependent Density-Functional Theory. *Phys. Rev. Lett.* **1995**, *76*, 1212.

(10) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *J. Chem. Phys.* **2003**, *119*, 2943–2946.

(11) Nitta, H.; Kawata, I. A close inspection of the charge-transfer excitation by TDDFT with various functionals: An application of orbital- and density-based analyses. *Chem. Phys.* **2012**, *405*, 93–99.

(12) Eriksen, J. J.; Sauer, S. P.; Mikkelsen, K. V.; Christiansen, O.; Jensen, H. J. A.; Kongsted, J. Failures of TDDFT in describing the lowest intramolecular charge-transfer excitation in para-nitroaniline. *Mol. Phys.* **2013**, *111*, 1235–1248.

(13) Yang, Y.; Van Aggelen, H.; Yang, W. Double, Rydberg and charge transfer excitations from pairing matrix fluctuation and particle-particle random phase approximation. *J. Chem. Phys.* **2013**, *139*.

(14) Casida, M. E.; Salahub, D. R. Asymptotic correction approach to improving approximate exchange-correlation potentials: Time-dependent density-functional theory calculations of molecular excitation spectra. *J. Chem. Phys.* **2000**, *113*, 8918–8935.

(15) Van Meer, R.; Gritsenko, O. V.; Baerends, E. J. Physical meaning of virtual kohn-sham orbitals and orbital energies: An ideal basis for the description of molecular excitations. *J. Chem. Theory Comput.* **2014**, *10*, 4432–4441.

(16) Dreuw, A.; Head-gordon, M. Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules. *Sciences-New York* **2005**, 4009–4037.

(17) Elliott, P.; Goldson, S.; Canahui, C.; Maitra, N. T. Perspectives on double-excitations in TDDFT. *Chem. Phys.* **2011**, *391*, 110 – 119.

(18) Gritsenko, O.; Baerends, E. J. Asymptotic correction of the exchange-correlation kernel of time-dependent density functional theory for long-range charge-transfer excitations. *J. Chem. Phys.* **2004**, *121*, 655–660.

(19) Chai, J. D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(20) Chai, J. D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 0–15.

(21) Mardirossian, N.; Head-Gordon, M. ωb97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

(22) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*.

(23) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett* **2011**, *2*, 2810–2817.

(24) Lin, Y.-S.; Tsai, C.-W.; Li, G.-D.; Chai, J.-D. Long-range corrected hybrid meta-generalized-gradient approximations with dispersion corrections. *J. Chem. Phys.* **2012**, *136*, 154109.

(25) Maitra, N. T. Charge transfer in time-dependent density functional theory. *J. Phys.: Condens. Matter* **2017**, *49*, 423001.

(26) Maitra, N. T. Perspective: Fundamental aspects of time-dependent density functional theory. *J. Chem. Phys.* **2016**, *144*, 220901.

(27) Maitra, N. T.; Zhang, F.; Cave, R. J.; Burke, K. Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.* **2004**, *120*, 5923.

(28) Park, Y. C.; Krykunov, M.; Ziegler, T. On the relation between adiabatic time dependent density functional theory (TDDFT) and the Δ SCF-DFT method. Introducing a numerically stable Δ SCF-DFT scheme for local functionals based on constricted variational DFT. *Mol. Phys.* **2015**, *113*, 1636–1647.

(29) Ziegler, T.; Seth, M.; Krykunov, M.; Autschbach, J.; Wang, F. On the relation between time-dependent and variational density functional theory approaches for the determination of excitation energies and transition moments. *J. Chem. Phys.* **2009**, *130*, 154102.

(30) Ziegler, T.; Seth, M.; Krykunov, M.; Autschbach, J. A revised electronic Hessian for approximate time-dependent density functional theory. *J. Chem. Phys.* **2008**, *129*, 184114.

(31) Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372.

(32) Seidl, A.; Görling, A.; Vogl, P.; Majewski, J. A.; Levy, M. Generalized Kohn-Sham schemes and the band-gap problem. *Phys. Rev. B* **1996**, *53*, 3764.

(33) Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Goodfellow, I.; Harp, A.; Irving, G.; Isard, M.; Jia, Y.; Jozefowicz, R.; Kaiser, L.; Kudlur, M.; Levenberg, J.; Mané, D.; Monga, R.; Moore, S.; Murray, D.; Olah, C.; Schuster, M.; Shlens, J.; Steiner, B.; Sutskever, I.; Talwar, K.; Tucker, P.; Vanhoucke, V.; Vasudevan, V.; Viégas, F.; Vinyals, O.; Warden, P.; Wattenberg, M.; Wicke, M.; Yu, Y.; Zheng, X. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015; `http://tensorflow.org/`, Software available from tensorflow.org.

(34) Nocedal, J. Updating quasi-Newton matrices with limited storage. *Math. Comp.* **1980**, *35*, 773 – 782.

(35) Hirata, S.; Head-Gordon, M.; Bartlett, R. J. Configuration interaction singles, time-dependent Hartree-Fock, and time-dependent density functional theory for the electronic excited states of extended systems. *J. Chem. Phys.* **1999**, *111*, 10774–10786.

(36) Gräfenstein, J.; Cremer, D. The combination of density functional theory with multi-configuration methods - CAS-DFT. *Chem. Phys. Lett.* **2000**, *316*, 569–577.

(37) Gaudoin, R.; Burke, K. Lack of Hohenberg-Kohn Theorem for Excited States. *Phys. Rev. Lett.* **2004**, *93*, 173001.

(38) Levy, M.; Nagy, A. Variational Density-Functional Theory for an Individual Excited State. *Phys. Rev. Lett.* **1999**, *83*, 4361.

(39) Görling, A. Symmetry in density-functional theory. *Phys. Rev. A* **1993**, *47*, 2783.

(40) Görling, A. Density-functional theory beyond the Honhenberg-Kohn theorem. *Phys. Rev. A* **1999**, *59*, 3359.

(41) Görling, A. Proper Treatment of Symmetries and Excited States in a Computationally Tractable Kohn-Sham Method. *Phys. Rev. Lett.* **2000**, *85*, 4229.

(42) T.Ziegler,; Rauk, A.; Baerends, E. J. *Theor. Chim. Acta* **1977**, *43*, 261.

(43) Ferré, N.; Filatov, M.; Huix-Rotllant, M. *Density-Functional Methods for Excited States*; Springer, Cham, 2016.

(44) Filatov, M. Spin-restricted ensemble-referenced Kohn-Sham method: basic principles and application to strongly correlated ground and excited states of molecules. *WIREs Comput Mol Sci* **2015**, *5*, 146–167.

(45) Manni, G. L.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2014**, *10*, 3669–3680.

(46) Carlson, R. K.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory: A Fully Translated Gradient Approximation and Its Performance for Transition Metal Dimers and the Spectroscopy of $Re_2Cl_8^{-2}$. *J. Chem. Theory Comput.* **2015**, *11*, 4077–4085.

(47) Sharma, P.; Bernales, V.; Knecht, S.; Truhlar, D.; Gagliardi, L. Density matrix renormalization group pair-density functional theory (DMRG-PDFT): singlet-triplet gaps in polyacenes and polyacetylenes. *Chem. Sci.* **2019**, *10*, 1716.

(48) Staroverov, V. N.; Davidson, E. R. Charge Densities for Singlet and Triplet Electron Pairs. *Int J. Quantum Chem.* **1999**, *77*, 651–660.

(49) Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method. *J. Phys. Chem. A* **2008**, *112(50)*, 13164–13171.

(50) Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. *J. Chem. Phys.* **2009**, *130*, 124308.

(51) Kowalczyk, T.; Tsuchimochi, T.; Chen, P. T.; Top, L.; Van Voorhis, T. Excitation energies and Stokes shifts from a restricted open-shell Kohn-Sham approach. *J. Chem. Phys.* **2013**, *138*.

(52) Hait, D.; Zhu, T.; McMahon, D. P.; Van Voorhis, T. Prediction of excited-state energies and singlet-triplet gaps of charge-transfer states using a restricted open-shell Kohn-Sham approach. *J. Chem. Theory Comput.* **2016**, *12*, 3353–3359.

(53) Kaduk, B.; Kowalczyk, T.; Voorhis, T. V. Constrained Density Functional Theory. *Chem. Rev.* **2012**, *112*, 321–370.

(54) Wu, Q.; Voorhis, T. V. Constrained Density Functional Theory and Its Application in Long-Range Electron Transfer. *J. Chem. Theoy Comput.* **2006**, *2*, 765–774.

(55) Wu, Q.; Voorhis, T. V. Extracting electron transfer coupling elements from constrained density functional theory. *J. Chem. Phys.* **2006**, *125*, 164105.

(56) Wu, Q.; Voorhis, T. V. Direct Calculation of Electron Transfer Parameters through Constrained Density Functional Theory. *J. Phys. Chem. A* **2006**, *110*, 9212–9218.

(57) Kaduk, B.; Voorhis, T. V. Communication: Conical intersections using constrained density functional theoryconfiguration interaction. *J. Chem. Phys.* **2010**, *133*, 061102.

(58) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K. L. PySCF: the Python-based simulations of chemistry framework. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*.

(59) Lebedev, V. I.; Laikov, D. N. A quadrature formula for the sphere of the 131st algebraic order of accuracy. *Dokl. Math.* **1999**, *59*, 477–481.

(60) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Ku, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; III, H. L. W.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; Jr., R. A. D.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W.; Harbach, P. H.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Lu-

enser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; ONeill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stck, D.; Su, Y.-C.; Thom, A. J.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K.; Chipman, D. M.; Cramer, C. J.; III, W. A. G.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; III, H. F. S.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Voorhis, T. V.; Herbert, J. M.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

(61) Filatov, M.; Shaik, S. A spin-restricted ensemble-referenced Kohn-Sham method and its application to diradicaloid situations. *Chem. Phys. Lett.* **1999**, *304*, 429–437.

(62) Knowles, P. J.; Manby, F. R.; Werner, H.-J.; Knizia, G.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *2*, 242–253.

(63) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *162*, 165–169.

(64) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098.

(65) Becke, A. D. A new inhomogeneity parameter in density-functional theory. *J. Chem. Phys.* **1998**, *109*, 2092–2098.

(66) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(67) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1988**, *90*, 1007.

(68) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796.

(69) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257.

(70) Zhao, Y.; Truhlar, D. G. Density functional for spectroscopy: No long-range self-interaction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states. *J. Phys. Chem. A* **2006**, *110*, 13126–13130.

(71) Baerends, E. J.; Gritsenko, O. V.; van Meer, R. The Kohn-Sham gap, the fundamental gap and the optical gap: the physical meaning of occupied and virtual Kohn-Sham orbital energies. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16408.

(72) Gonze, X.; Scheffler, M. Exchange and Correlation Kernels at the Resonance Frequency: Implication for Excitations Energies in Density Functional Theory. *Phys. Rev. Lett.* **1999**, *82*, 4416.

(73) Subotnik, J. E. Communication: Configuration interaction singles has a large systematic bias against charge-transfer states. *J. Chem. Phys.* **2011**, *135*, 071104.

(74) Cheng, C.-L.; Wu, Q.; Voorhis, T. V. Rydberg energies using excited state density functional theory. *J. Chem. Phys.* **2008**, *129*, 124112.

Graphical TOC Entry