Electronic correlation without double counting via a combination of spin projected Hartree-Fock and density functional theories

Alejandro J. Garza,1 Carlos A. Jiménez-Hoyos,1 and Gustavo E. Scuseria2

1Department of Chemistry, Rice University, Houston, Texas 77251-1892, USA
2Department of Chemistry and Department of Physics and Astronomy, Rice University, Houston, Texas 77251-1892, USA and Chemistry Department, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

(Received 29 April 2014; accepted 3 June 2014; published online 23 June 2014)

Several schemes to avoid the double counting of correlations in methods that merge multireference wavefunctions with density functional theory (DFT) are studied and here adapted to a combination of spin-projected Hartree-Fock (SUHF) and DFT. The advantages and limitations of the new method, denoted SUHF+\(_f\), DFT, are explored through calculations on benchmark sets in which the accounting of correlations is challenging for pure SUHF or DFT. It is shown that SUHF+\(_f\), DFT can greatly improve the description of certain molecular properties (e.g., singlet-triplet energy gaps) which are not improved by simple addition of DFT dynamical correlation to SUHF. However, SUHF+\(_f\), DFT is also shown to have difficulties dissociating certain types of bonds and describing highly charged ions with static correlation. Possible improvements to the current SUHF+\(_f\), DFT scheme are discussed in light of these results. © 2014 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4883491]

I. INTRODUCTION

In quantum chemistry, correlation energy is usually defined as the difference between the Hartree-Fock and exact (electronic) energies. 1 It is thus by definition that we are interested in calculating correlations. This task, however, remains a significant challenge and to simplify the problem one often distinguishes between two types of correlations: static (or strong) and dynamic (or weak). Static correlations arise due to degeneracies or near degeneracies in frontier orbitals and need to be dealt with multireference (MR) methods. Dynamic correlations occur as a result of instantaneous electron repulsions; MR techniques are broadly inefficient for capturing these since very large determinant expansions are required to describe the electron correlation cusp. Instead, one uses other methods to treat dynamic correlations, e.g., coupled cluster methods of which CCSD(T) is considered the “gold standard.” Still, the \(O(M^3)\) computational scaling of CCSD(T) remains daunting and hinders the method’s applicability to large systems.

A hugely successful, 2 low cost, alternative to wavefunction methods is Kohn-Sham density functional theory (DFT). 3 This approach owes its popularity to its mean-field scaling (typically \(O(M^3)\)), as well as the good accuracy that modern functionals can achieve. Kohn-Sham DFT is, nonetheless, a single-reference theory and is therefore unreliable for strongly correlated systems. Considering the above discussion, and the fact that DFT correlation functionals capture exclusively dynamic correlations, it would appear like a very attractive idea to combine MR and DFT in an attempt to simultaneously (and cheaply) describe static and dynamic correlations. That is, to approximate the exact energy as

\[
E \approx E^{\text{MR}} + E^{\text{DFT}},
\]

where the total MR energy \(E^{\text{MR}}\) takes into account the static correlations and \(E^{\text{DFT}}\) is pure dynamical correlation. However, the combination of MR and DFT is not straightforward, mainly because of two problems: symmetry inconsistencies and double counting. The first issue refers to modern DFT functionals being designed to work with densities that break spin \((S^2)\) symmetry, whereas MR methods yield symmetry adapted densities. The second one originates because static and dynamic correlations cannot be fully separated from a MR wavefunction, resulting in double counting (i.e., overcorrelating) when adding \(E^{\text{DFT}}\) to \(E^{\text{MR}}\).

Because of the potential of MR+DFT to describe both types of correlation at moderate cost, much effort has been devoted to develop useful MR+DFT techniques that avoid the above mentioned problems. Lie and Clementi 4,5 combined a multiconfiguration self-consistent field (MCSCF) wavefunction of minimal active space with a reparametrized Gombás’ functional, 6 which depended on natural orbitals and their occupation numbers. This is conceptually similar to a more recent MCSCF+DFT implementation 7 that uses modern functionals with alternative densities defined by a mapping of the occupation numbers. 8 To describe molecular dissociations, Kraka 9 added LDA correlation to generalized valence bond (GVB) energies using the correction by Stoll et al. 10

\[
E^{\text{DFT}} = E^{\text{LDA}}[\gamma_0, \gamma_\rho] - E^{\text{LDA}}[\gamma_\rho, 0] - E^{\text{LDA}}[0, \gamma_\rho],
\]

which effectively suppresses the self-interaction error (at the cost of neglecting equal spin correlations). A different way to couple MR and DFT is that of Colle and Salvetti, 11,12 who designed a functional depending on the charge density and the pair density, and combined it with two-configuration wavefunctions. Along these lines, alternative densities may be defined by the charge density \(\gamma(r)\) and the on-top pair density...
A. SUHF

The SUHF method has, in fact, a long history in quantum chemistry.\cite{1, 36, 37} Nevertheless, it never became widely utilized mainly because its initial formulation\cite{1} was inefficient, involving the evaluation of products of two-body operators. Recently, however, a mean-field SUHF methodology has been proposed and implemented by our research group.\cite{34} This implementation draws from the form of the projection operators used in nuclear physics,\cite{38} and can capture strong correlations arising from spin fluctuations while retaining the low computational scaling of regular Hartree-Fock.\cite{39} For a clearer exposition of our SUHF+DFT scheme, a brief description of SUHF is germane here. The interested reader may find further details in Ref. 34 (see also Refs. 38, 40, and 41).

In SUHF, one has a Slater determinant |Φ\rangle that is an eigenfunction of \(\hat{S}_z\) but is allowed to break \(\hat{S}^2\) symmetry (i.e., |Φ\rangle is of UHF type). A projection operator \(\hat{P}_{\sigma\sigma}^\pm\) restores the symmetry of |Φ\rangle, yielding an eigenfunction of \(\hat{S}^2\) and \(\hat{S}_z\) with eigenvalues \(s(s + 1)\) and \(m\), respectively. More precisely, the form of \(\hat{P}_{\sigma\sigma}^\pm\) that we utilize is mathematically defined as

\[
\hat{P}_{\sigma\sigma}^\pm = \frac{2s + 1}{8\pi^2} \int d\Omega \left[D_{\sigma\sigma}^{\pm}(\Omega)\right]^2 \hat{R}(\Omega),
\]

where \(\Omega = (\alpha, \beta, \gamma)\) is the set of Euler angles, \(D_{\sigma\sigma}^{\pm}(\Omega) = (s, m|R(\Omega)|s, m)\) is Wigner’s D-matrix, and \(\hat{R}(\Omega) = \exp[-i\alpha\hat{S}_x] \exp[-i\beta\hat{S}_y] \exp[-i\gamma\hat{S}_z]\) is the spin rotation operator. With these definitions, the description of SUHF is straightforward: SUHF minimizes the energy of the wavefunction |Ψ\rangle = \(\hat{N}\hat{P}_{\sigma\sigma}^\pm|\Phi\rangle\), where \(\hat{N}\) is a normalization constant. That is, we use a variation after projection approach, which is more physically sound than the simpler, more common, projection after variation. We also note that |Ψ\rangle is multireference because |Φ\rangle can be written as a linear combination of non-orthogonal determinants that, when written in the traditional particle-hole excitation picture, involve replacements to all levels (collective excitations).\cite{38} The projection recovers the components having the desired quantum numbers or, in other words, eliminates all the contaminants from the expansion irrespectively of their excitation level.

We are now poised to define the key ingredients required by SUHF+\(f_\sigma\)DFT: the densities generated by |Φ\rangle and |Ψ\rangle. We follow Harriman’s notation\cite{42} and denote the true (projected) SUHF one-particle density matrix (1PDM) of spin \(\sigma\) as \((\gamma_{\sigma})_{ij} = \langle\Psi|a_i^{\dagger}\sigma a_j\sigma|\Psi\rangle\), whereas the deformed (unprojected) 1PDM is \((\varrho_{\sigma})_{ij} = \langle\Phi|a_i^{\dagger}\sigma a_j\sigma|\Phi\rangle\). The matrix \(\varrho_{\sigma}\) is evaluated straightforwardly from the MO coefficients of |Φ\rangle, \(C_{\sigma}\), as \(C_{\sigma} C_{\sigma}^\dagger\). It was noted in Ref. 40 that computing \(\varrho_{\sigma}\) would require a double integral over \(\Omega\) because, in general, \(\hat{P}_{\sigma\sigma}^\pm\) and \(a_i^{\dagger}\sigma a_j\sigma\) do not commute. However, an equivalent but more efficient approach is employed here: using the Wigner-Eckart theorem,\cite{43} one can demonstrate that, for any rank-\(k\) tensor operator \(T_k\), the following relationship holds:\cite{44, 45}

\[
\hat{P}_{\sigma\sigma}^\pm T_k^{\pm\pm} = \delta_{\varrho,0}(s, m, k, 0|s, m) \\
\times \sum_{\mu, \mu'}(-1)^{\mu'}(s, m, k, -\mu'|s, m, -\mu')T_k^{\pm\pm} \hat{P}_{\sigma\sigma}^\pm|_{\mu, \mu', m}.\tag{5}
\]

where \((j_1, m_1, j_2, m_2|j, m)\) are Clebsch-Gordan coefficients. Equation (5) effectively allows to evaluate \(\hat{P}_{\sigma\sigma}^\pm a_i^{\dagger}\alpha a_k\beta \hat{P}_{\sigma\sigma}^\pm\) in terms of linear combinations of \(a_i^{\dagger}\alpha a_k\beta\hat{P}_{\sigma\sigma}^\pm\) (with \(\sigma^\prime\) running over spins \(\alpha\) and \(\beta\)), circumventing the more costly double integration. This is helpful in SUHF+\(f_\sigma\)DFT because both \(\varrho_{\sigma}\) and \(\gamma_{\sigma}\) are needed to couple SUHF and DFT information in this scheme.

B. Coupling of SUHF and DFT

As mentioned in the Introduction, modern DFT functionals are designed to work with densities that break \(\hat{S}^2\)
symmetry. This means that the true SUHF densities
\[
\gamma_{\sigma}(r, r') = \sum_{ij} (\gamma_{\sigma})_{ij} \varphi_\sigma^*(r) \varphi_\sigma(r'),
\]
obtained by setting \(r' = r\) (here, \(\varphi_\sigma^*(r)\) is a spatial orbital of \(\sigma\) spin), cannot be used if we are to take advantage of the vast resources available for current DFT methods. Although the alternative densities of Eq. (3) could in principle be used for SUHF+DFT, certain complications arise that hamper their direct applicability. In particular, Eq. (3) was derived from the expression of \(P_{\sigma}(r)\) for a single-determinant state, \(P_{\sigma}(r) = \gamma_{\sigma}(r) \gamma_{\sigma}(r)\). This has the consequence that, for a \(P_{\sigma}(r)\) derived from a multireference wavefunction, it is possible to have \(4P_{\sigma}(r) > \gamma(r)^2\) leading to unphysical complex densities in Eq. (3) and a formally imaginary spin density\(^{48}\) (typical in, e.g., open-shell singlets). Thus, one would need to extrapolate DFT correlation functionals to imaginary spin density values. We opt here for using a different set of densities to avoid this complication.

Suitable densities for DFT can be obtained straightforwardly from SUHF if, rather than forming \(\gamma_{\sigma}(r)\) using \((\gamma_{\sigma})_{ij}\), we form a deformed density \(\varphi_{\sigma}(r)\) from the unprojected matrix \((\varphi_{\sigma})_{ij}\). It follows from the above discussions that the resulting \(\varphi_{\sigma}(r)\) and \(\varphi_{\beta}(r)\) components break spin symmetry and are thus compatible with DFT correlation functionals. This is one of the approaches that we used in our previous SUHF+DFT implementation in Ref. 35, where it was shown to work reasonably well. Here, \(\varphi_{\sigma}(r)\) and \(\varphi_{\beta}(r)\) are also used as input densities for \(E_{c}^{\text{DFT}}\). However, this time we shall scale the DFT correlation energy density locally by a factor \(f(\gamma)\), which uses the projected \(\gamma(r)\) to estimate the fraction of dynamic correlation already captured by SUHF. That is, we now use an extended functional of the form
\[
E_{c}^{\text{DFT}}[\varphi_{\alpha}, \varphi_{\beta}, \gamma] = \int d^3r \varrho(r) f(\gamma) E_{c}^{\text{DFT}}[\varphi_{\alpha}, \varphi_{\beta}],
\]
where \(\varrho(r) = \varrho_{\alpha}(r) + \varrho_{\beta}(r)\), and \(E_{c}^{\text{DFT}}\) is a standard DFT correlation energy density. It is important to note that, in our implementation, \(E_{c}^{\text{DFT}}\) is evaluated in a single-point calculation using converged SUHF densities, rather than in a self-consistent manner. The effect of self-consistency, however, has been shown to be small when adding \(E_{c}^{\text{DFT}}\) to wavefunction methods.\(^{16,49}\) We thus proceed to describe how to calculate \(f(\gamma)\).

C. Scaling of the correlation energy density

To compute the scaling factor \(f(\gamma)\), we draw from the methods designed for CAS+DFT in Refs. 16–18 and adapt them to SUHF+DFT. Following Miehlich et al.,\(^{16}\) we begin by defining a reference density \(\varrho_{\text{ref}}\) that describes the spatial regions spanned by the occupied (integer or fractional) orbitals:
\[
\varrho_{\text{ref}}(r) = \sum_{\gamma \neq \gamma_{\text{ref}}} n_{\gamma} \psi_{\gamma}(r)^2,
\]
where \(\psi_{\gamma}(r)\) are natural orbitals of \(\gamma(r, r') = \gamma_{\alpha}(r, r') + \gamma_{\beta}(r, r')\), and the sum runs over all orbitals with nonzero occupation numbers \(n_{\gamma}\). Thus, in CAS+DFT the sum is over all core and active orbitals. In SUHF, there are not strictly defined core and active orbitals. However, for reasons that will in short be evident, the results by Harriman\(^{12}\) (see also Refs. 50 and 51) on the eigendecomposition of \(\gamma(r, r')\) in SUHF are worth mentioning here:

- The spectrum of \(\gamma(r, r')\) has the structure \(\sigma(\gamma) = \{[2 - x, x], 1, 0\}\), with \(0 \leq x \leq 1\).
- The natural orbitals \(\psi_{\gamma}(r)\) that diagonalize \(\gamma(r, r')\) also diagonalize the unprojected \(\rho_{\gamma}(r, r')\).
- The spectrum of the unprojected \(\rho_{\gamma}(r, r')\) also has the structure \(\sigma(\rho) = \{[2 - y, y], 1, 0\}\), with \(0 \leq y \leq 1\), but, in general, \(y \neq x\).

Hence, in SUHF, the sum in Eq. (8) runs over at most \(N\) orbitals, where \(N\) is the number of electrons. This is so because the structure of \(\sigma(\gamma)\) and the conservation of particle number trace(\(\gamma\)) = \(N\) lead to \(\max\{\text{Rank}(\gamma)\} = N\). Furthermore, due to the two last points noted above, \(\rho_{\text{ref}}(r, r')\) can be evaluated from \(\rho_{\gamma}(r, r')\) solely.

With \(\varrho_{\text{ref}}(r)\), a local measure of the size of the active space can be defined as\(^{16–18}\)
\[
\eta_{\text{ref}}(r) = \left(\frac{\varrho_{\text{ref}}(r)}{\gamma(r)}\right)^{1/3}.
\]
It is easy to see from Eqs. (8) and (9) that, for an RHF density where all orbitals are doubly occupied, \(\eta_{\text{ref}}(r) = 1\). This value is bound to increase with the number of fractionally occupied orbitals.

To actually estimate the fraction of dynamic correlation already contained in \(\gamma(r)\) from \(\eta_{\text{ref}}(r)\), one can lean on a suitable model system and introduce approximations that satisfy known limits. These limits can be \(f(\gamma) = 1\) and \(f(\gamma) = 0\) for a density with no dynamic correlation (e.g., RHF) and an exact density, respectively. A possibility for the model system is then provided by the calculations by Savin,\(^{52}\) which describe the correlation density of a homogeneous electron gas, \(\epsilon_{\text{hom}}(\gamma, \infty)\), where the excitations are restricted to levels with energies lower or equal than those of the orbitals used to construct \(\eta_{\text{ref}}\) (i.e., orbitals with \(n_{\gamma} \neq 0\), see Figure 1). If all excitations are considered, this energy becomes \(\epsilon_{\text{hom}}(\gamma, \infty)\); the types of excitations taken into

![FIG. 1. Types of excitations taken into account by \(\epsilon_{\text{hom}}(\gamma, \infty)\) and \(\epsilon_{\text{hom}}(\gamma, \eta_{\text{ref}})\). The latter suppresses contributions to the correlation from excitations that go to levels higher in energy than those with occupations \(n_{\gamma} \neq 0\).](image-url)
account by \(\varepsilon\) and \(\varepsilon\) are both illustrated in Figure 1. Thus, the fraction of dynamic correlation in \(\gamma(r)\) is estimated to be \(\varepsilon\) and the scaling factor should therefore be

\[
f(\gamma, \eta_{\text{ref}}) = 1 - \frac{\varepsilon_{\text{hom}}(\gamma, \eta_{\text{ref}})}{\varepsilon_{\text{hom}}(\gamma, \infty)}. \tag{10}
\]

The actual evaluation of \(f(\gamma, \eta_{\text{ref}})\) is conveniently carried out via the parametrization given in Ref. 16,

\[
f(\gamma, \eta_{\text{ref}}) = \left( \sum_{mn} b_{mn} x^{m-1} y^{n-1} \right)^{-1}, \tag{11}
\]

where \(b_{mn}\) is a \(6 \times 5\) matrix, \(y = \eta_{\text{ref}}\), and \(x = \ln r_s\), with \(r_s = [3/(4\pi \gamma)]^{1/3}\). The values of \(b_{mn}\) can be consulted in the original paper by Miehlich et al.16

It is easy to verify that \(f(\gamma, \eta_{\text{ref}}) = f(\gamma)\), as given by Eq. (10), satisfies \(f(\gamma) = 1\) and \(f(\gamma) = 0\) for the limiting cases of RHF and an exact wavefunction, respectively. For RHF, \(\gamma(r) \approx \eta_{\text{ref}}(r)\) so that there are no occupied-virtual excitations in \(\varepsilon_{\text{hom}}(\gamma, \eta_{\text{ref}})\), which leads to \(\varepsilon_{\text{hom}}(\gamma, \eta_{\text{ref}}) = 0\) and \(f(\gamma) = 1\). In contrast, if all excitations are taken into account, \(\eta_{\text{ref}} \rightarrow \infty\) in the infinite basis set limit and therefore \(f(\gamma)\) becomes zero trivially from Eq. (10). These trends can be appreciated in Figure 2, which shows a contour plot of the values of \(f(\gamma)\) as a function of \(x\) and \(y = \eta_{\text{ref}}\). It is seen that \(f(\gamma) = 1\) for \(\eta_{\text{ref}} = 1\), and that this value decays rather rapidly to zero with increasing \(\eta_{\text{ref}}\), specially in regions of greater density (lower \(x\)).

D. Scaling factor’s core correction

If there are orbitals with \(n_k \notin 0\) that do not (but should) contribute to the MR dynamic correlation, then the above described scaling factor will need further corrections. The reason for this is that \(\varepsilon_{\text{hom}}(\gamma, \eta_{\text{ref}})\) considers all dynamic correlations from excitations between levels with \(n_k \neq 0\). Thus, if the MR method employed to get \(\gamma(r)\) is not capturing correlations between some of these levels, they will be excluded by the \(f(\gamma)\) given in Eq. (10). This problem was noted by Gräfenstein and Cremer17,18 in CAS+DFT, where it affects the excitations from core (inactive) into weakly populated active orbitals. An analogous problem, illustrated in Figure 3, occurs in SUHF+DFT because SUHF captures mostly correlations from low-lying excitations, and transitions from low energy orbitals with \(n_k \geq 2\) do not really contribute to the correlation energy. The authors in Refs. 17 and 18 proposed a correction for this issue that we adjust here to SUHF+DFT.

The missing correlations from excitations of core \(n_k \geq 2\) into weakly populated \((n_k = x)\) orbitals, \(\varepsilon_{\text{core\rightarrow weak}}\), can be approximated as

\[
\varepsilon_{\text{core\rightarrow weak}} = \varepsilon_{\text{core\rightarrow act}} - \varepsilon_{\text{core\rightarrow strong}}, \tag{12}
\]

where the active occupied orbitals are those with \(n_k = 2 - x\), whereas for strongly occupied \(n_k = 2 - x\). The terms in Eq. (12) are estimated as

\[
\varepsilon_{\text{core\rightarrow act}} = \frac{\varepsilon_{\text{hom}}(\gamma_{\text{core}}, \eta_{\text{ref}})}{\varepsilon_{\text{hom}}(\gamma_{\text{core}}, \infty)} \varepsilon_{\text{DFT}}(\gamma_{\text{core}}), \tag{13}
\]

\[
\varepsilon_{\text{core\rightarrow strong}} = \frac{\varepsilon_{\text{hom}}(\gamma_{\text{core}}, \eta_{\text{ref}})}{\varepsilon_{\text{hom}}(\gamma_{\text{core}}, \infty)} \varepsilon_{\text{DFT}}(\gamma_{\text{core}}), \tag{14}
\]

where \(\gamma_{\text{core}}\) is a density constructed from the orbitals classified as inactive or core, and \(\eta_{\text{ref}}\) is analogous to \(\eta_{\text{ref}}\) but defined instead in terms of \(\gamma_{\text{core}}\) and \(\gamma\), i.e., \(\eta_{\text{ref}} = (\gamma/\gamma_{\text{core}})^{1/3}\). Using Eq. (10), this immediately leads to

\[
\varepsilon_{\text{core\rightarrow weak}} = [f(\gamma_{\text{core}}, \eta_{\text{ref}}) - f(\gamma_{\text{core}}, \eta_{\text{ref}})] \varepsilon_{\text{core}}. \tag{15}
\]

Hence, the correction that should be added to \(E_{\text{DFT}}\) in Eq. (7) is

\[
E_{\text{DFT}}(\gamma, \gamma_{\text{core}}) = \int d^3r \gamma_{\text{core}} f(\gamma_{\text{core}}, \eta_{\text{ref}}) \varepsilon_{\text{core}}(\gamma_{\text{core}})
- \int d^3r \gamma_{\text{core}} f(\gamma_{\text{core}}, \eta_{\text{ref}}) \varepsilon_{\text{core}}(\gamma_{\text{core}}), \tag{16}
\]

The above expression always yields nonpositive energies because \(f(\gamma_{\text{core}}, \eta_{\text{ref}}) \geq f(\gamma_{\text{core}}, \eta_{\text{ref}})\). Also, \(E_{\text{DFT}}\) will depend on the choice of core orbitals. This dependence can introduce problems in certain cases and we shall discuss this issue in Secs. III and IV.
E. Alternative densities for special cases

Having fully described the SUHF+DFT methods used here, it is worth noting that alternative densities other than $\varrho_\sigma$ can be used to couple SUHF and DFT. In particular, Gräfenstein and Cremer\cite{17,18} have shown that using the densities defined by Staroverov and Davidson\cite{53}:
\begin{align}
 u_A(r) &= \frac{1}{2} \sum_k n_k (1 + n_k (2 - n_k)) |\psi_k(r)|^2, \\
 u_B(r) &= \frac{1}{2} \sum_k n_k (1 - n_k (2 - n_k)) |\psi_k(r)|^2,
\end{align}
yields improved results in the prediction of singlet-triplet splittings involving low-spin open-shell states. Thus, in Sec. III, we present also SUHF+DFT results for this kind of calculations using the $u_A$ and $u_B$ densities instead of $\varrho_\sigma$ and $\varrho_\beta$. We also note that, for low-spin open-shell singlets, SUHF typically yields occupation numbers close to two, one, and zero. Hence, the results given using $u_A$ and $u_B$ would be nearly identical to those obtained from the densities by Pérez-Jiménez and Pérez-Jordá.\cite{7,8}

\begin{align}
 \rho_\beta(r) &= \sum_{n_\beta \geq 1} (n_\beta - 1) |\psi_k(r)|^2, \\
 \rho_\alpha(r) &= \sum_{n_\alpha \geq 1} |\psi_k(r)|^2 + \sum_{n_\alpha < 1} n_\alpha |\psi_k(r)|^2,
\end{align}

because $u_A \simeq \rho_\beta$ and $u_B \simeq \rho_\alpha$ for $n_k \simeq 2, 1, 0$. 

F. Nomenclature

It is convenient at this point to introduce notation to refer to the techniques so far described. We denote the direct addition of the unscaled $E_c^{DFT}$ to $E_{SUHF}^{SUHF}$ as SUHF+DFT (this is the approach used in our previous paper\cite{35}); if $E_c^{DFT}$ is scaled by $f(\gamma)$, the method is referred to as SUHF+fDFT; if the core correction $E_c^{DFT}$ is added to SUHF+fDFT, the resulting method is denoted SUHF+f+DFT. We shall also adopt an analogous notation to refer to earlier CAS+DFT methods, in order to facilitate comparison with SUHF+DFT results given here. As in Ref. 35, we indicate within square brackets the particular set of alternative densities used in the DFT correlation evaluation. Thus, DFT[\$] refers to the deformed $\varrho_\alpha(r)$ and $\varrho_\beta(r)$ densities of SUHF; DFT[\$] corresponds with $\rho_\beta$ and $\rho_\alpha$, and DFT[\$] uses $u_A$ and $u_B$. For brevity, DFT[\$] = DFT, since most calculations presented here use the $\varrho_\sigma$ densities. We also remind the reader that DFT[\$] and DFT[\$] are constructed using the density matrix of the projected state.

G. Computational details

Our SUHF+DFT schemes were implemented in a development version of the GAUSSIAN suite of programs.\cite{54} Calculations were carried out using Dunning’s correlation consistent basis sets,\cite{36} cc-pVnZ. Unrestricted Hartree-Fock wavefunctions were used to generate the initial guesses for SUHF. Frozen-core unrestricted coupled cluster singles and doubles with perturbative triplets, UCCSD(T),\cite{56} served as comparison reference in certain benchmark calculations. Extrapolation techniques\cite{57,58} (direct inversion of the iterative subspace (DIIS)) were also applied when feasible to accelerate self-consistent field convergence. The DFT correlation energies were evaluated using the Tao-Perdew-Staroverov-Scuseria (TPSS)\cite{59} functional without any reparametrization. This functional was selected based on the facts that (1) it is nonempirical; (2) it is a meta-GGA free of one-electron self-interaction error, meaning we do not need to use Eq. (2) (which ignores equal spin correlations) to avoid this issue; and (3) SUHF+TPSS provided good results in Ref. 35. However, for comparison purposes, we also show some results with PBE\cite{60,61} and VWN5\cite{62} which are, respectively, GGA and LDA type functionals.

III. RESULTS AND DISCUSSION

A. Helium isoelectronic series

Pure DFT methods are known to have difficulties in describing the correlation energies for the He isoelectronic series\cite{53} (i.e., two-electron $X(Z-2)^{2+}$ ions). For example, as the nuclear charge $Z$ increases, one gets $E_{\text{LDA}}^{\text{TPSS}} \to \infty$, whereas the experimental trend is $E_{\text{exp}}^{\text{TPSS}} \to \text{constant}$.\cite{63,64} Because of this, the He isoelectronic series was used in Ref. 16 to benchmark CAS+DFT. Thus, we use this same benchmark set here as a first test for the analogous SUHF+DFT. The core correction is not needed here because there are only two electrons and all SUHF orbitals can be considered to be active.

Figure 4 shows the correlation energies (with respect to RHF) for the He isoelectronic calculated by SUHF, SUHF+f/TPSS, SUHF+f/PBE, and SUHF+f/LDA with the cc-pV5Z basis. These data are compared with results from variational wavefunctions reported in Ref. 65, which are accurate to about one part in 10$^{14}$. The correlation in the $X(Z-2)^{2+}$ ions is purely dynamical and its magnitude increases slightly with $Z$ before reaching a constant value. It is seen that SUHF, better suited to calculate static correlation, yields a qualitatively incorrect trend by predicting a decrease in the absolute value of $E_{\text{SUHF}}^{\text{exp}}$ as $Z$ increases, before reaching a constant. Adding DFT dynamic correlation to SUHF corrects this trend; however, only SUHF+f/TPSS and SUHF+f/PBE yield results that are really close (~1–3 kcal/mol) to the accurate energies.

![Correlation Energy (Hartrees) vs. Nuclear Charge Z](image-url)  
**FIG. 4.** Correlation energies (with respect to RHF) for the He isoelectronic series (i.e., two-electron $X(Z-2)^{2+}$ ions). Accurate values were taken from Ref. 65.
TABLE I. Mean errors (ME) and nonparallelity errors (NPE) for the correlation energies of the He isoelectronic series. All values are in kcal/mol.

<table>
<thead>
<tr>
<th>Method</th>
<th>ME</th>
<th>NPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUHF</td>
<td>19.3</td>
<td>4.3</td>
</tr>
<tr>
<td>SUHF+fLDA</td>
<td>−45.0</td>
<td>27.3</td>
</tr>
<tr>
<td>SUHF+fPBE</td>
<td>2.8</td>
<td>2.5</td>
</tr>
<tr>
<td>SUHF+fTPSS</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>SUHF+TPSS</td>
<td>−8.6</td>
<td>3.6</td>
</tr>
</tbody>
</table>

In contrast, raw SUHF+TPSS, as well as SUHF+fLDA, both overestimate correlations significantly.

Table I complements Figure 4 by listing the mean errors (ME) and nonparallelity errors (NPE)—the difference between minimum and maximum errors—of the data in said figure. It is clear from this table that the best results are given by SUHF+fTPSS. The overcorrelation in SUHF+TPSS, reflected in a ME of −8.6 kcal/mol, comes from double counting, and SUHF+fTPSS (ME = 2.6 kcal/mol) corrects this problem. In the case of SUHF+fLDA (ME = −45.0 kcal/mol), the overestimation of $E_\text{DFT}$ comes from the functional itself; LDA is known to yield correlation energies that are too large by a factor of about 2.66 While in certain applications of LDA this error is partially compensated by an underestimation (by ∼10%) of the exchange energy,10,66 this is clearly not the case for MR+DFT where the exchange is exact. Thus, the use of GGAs and meta-GGAs over LDA functionals is highly recommended for MR+DFT applications.

B. Molecular dissociations

The He isoelectronic series serves to illustrate the effect of the scaling factor when only dynamic correlation is present. However, actual applications of MR+DFT are more likely to be concerned with situations where both dynamic and static correlations are important. Molecular dissociations provide perhaps the most paradigmatic examples of such situations. In a typical dissociation, dynamic correlations are the most prominent at equilibrium bond length, whereas static correlations dominate at dissociation (electrons become entangled). Thus, the quantitative description of said phenomenon demands the inclusion of both types of correlation, making it a fitting test for MR+DFT methods.

Figure 5 shows the potential energy curves for several of the molecules that were also studied in our previous SUHF+DFT paper.35 The mandatory curve for the H$_2$ molecule is shown in panel (a) of this figure; it is seen that SUHF+fTPSS matches very closely the full-CI (FCI) curve with a NPE of 2 kcal/mol and a mean absolute error (MAE) of only 1 kcal/mol. We emphasize that this good agreement...
is achieved even though no reoptimization of the TPSS parameters has been carried out here, in contrast with what was done in Ref. 35. Also interesting to look at are the plots of the scaling factor \( f(\gamma, \eta_{\text{ref}}) \) in real space for \( \text{H}_2 \) that appear in Figure 6. The scaling factor is largest in the regions between the atoms, where it is more likely to find the pair of electrons close together. This mirrors the fact that dynamic correlations arise due to instantaneous electron repulsions, and that \( f(\gamma, \eta_{\text{ref}}) \) should be largest where dynamic correlations are important.

Figure 5(b) depicts the dissociation of the BH molecule; in lieu of FCI, we give as reference UCCSD(T) data from Ref. 35 since this method has been shown to give results close to FCI (NPE \( \approx 3 \) kcal/mol) for this molecule.\(^{67}\) Here, SUHF+TPSS also provides a good improvement over SUHF and agrees very well with the UCCSD(T) energy profile. The symmetric dissociation of a triangular, \( D_3h \), \( \text{H}_3 \) molecule (Figure 5(c)) represents a simple three-electron system which is already challenging to describe for SUHF, which misses a large amount of dynamic correlation. However, SUHF+TPSS sharply improves the results over SUHF in this case, again matching well with UCCSD(T). Furthermore, this improvement is gained with virtually no increase in computational cost, since evaluating the \( f_{\text{TPSS}} \) correlation is inexpensive and does not contribute significantly to the total CPU time of the calculation.

For the purposes of the present paper, the most enlightening case is, however, provided by the dissociation of the \( \text{N}_2 \) molecule shown in Figure 5(d). This is the first case we have analyzed so far in which the correlations \( \epsilon_{\text{core-weak}}^{\text{core-weak}} \) from core into weakly occupied orbitals (Eq. (12)), which are suppressed in SUHF+TPSS (Figure 3), appear to be important. Despite improving upon SUHF, SUHF+TPSS underestimate the correlation energy near the equilibrium bond length, most likely due to the neglect of \( \epsilon_{\text{core-weak}}^{\text{core-weak}} \) contributions. Near equilibrium, however, the \( \sigma_{1s}, \sigma_{\sigma}, \sigma_{2s}, \) and \( \sigma_{2s}^\ast \) orbitals have occupation numbers which are very close to two. Thus, we can define \( (\sigma_{1s})^2(\sigma_{\sigma})^2(\sigma_{2s})^2(\sigma_{2s}^\ast)^2 \) to be the core for SUHF+\( f_i \)TPSS (i.e., four core orbitals). With this choice, SUHF+\( f_i \)TPSS yields results that compare very well with UCCSD(T) near equilibrium. Nonetheless, the correlation is now overestimated at dissociation. In fact, SUHF+\( f_i \)TPSS and SUHF+\( f_i \)TPSS are nearly parallel to each other. Hence, it would appear that we need a variable core to describe the whole dissociation curve. A seemingly simple fix for this issue would be to select a threshold on the occupation numbers of orbitals defining the core. However, such an approach would in turn introduce the possibility of discontinuities along the potential energy curve. One is thus faced with a dilemma on how to choose core orbitals in the current SUHF+\( f_i \)TPSS scheme. Feasible solutions to this problem are discussed in Sec. IV.

C. Singlet-triplet splittings

While SUHF+TPSS and SUHF+\( f_i \)TPSS do improve the description of molecular dissociations over SUHF, so does the less involved SUHF+TPSS.\(^{35}\) However, the latter methodology was shown to provide only marginal improvements upon SUHF for the prediction of singlet-triplet splittings,\(^{35}\) i.e., the energy gap between the lowest singlet and triplet states \( \Delta E_{ST} = E_S - E_T \). Moreover, singlet-triplet splittings are highly sensitive to both static and dynamic correlation effects. Therefore, calculating \( \Delta E_{ST} \) provides another good benchmark test for SUHF+DFT methods.

In the following discussion regarding singlet-triplet gaps, we shall borrow the benchmark sets and (when applicable) zero-point energy corrections from Slipchenko and Krylov.\(^{68}\) These sets include (1) methylene and isovalent molecules, (2) atoms, (3) diatomics, (4) trimethyleneimethane, and (5) benzene isomers. Subsets of these have also been studied in, e.g., Refs. 69–74. The benchmarks all consist of diradicals as defined by Salem and Rowland,\(^{75}\) i.e., species with two electrons occupying two degenerate or near degenerate orbitals. The accurate prediction of \( \Delta E_{ST} \) in diradicals is particularly difficult because, while the triplets are typically well described by a single Slater determinant, even the simplest possible singlet diradical wavefunction\(^{68}\)

\[
\Psi_i = \frac{1}{2} \left[ \lambda (\phi_1)^2 - \sqrt{1 - \lambda^2} (\phi_2)^2 \right] (\alpha \beta - \beta \alpha), \tag{21}
\]
The above mentioned deficiency of $f(y, \eta_{\text{ref}})$ is diminished by including the core correction to SUHF+TPSS. The SUHF+TPSS splittings in Table II show very good agreement with the experimental data (MAE = 0.8 kcal/mol), outclassing all of the other mean-field approaches in this table. These calculations used a core equivalent to that of a minimal active space CAS(2,2) calculation; this choice was based on the occupation numbers of the projected density matrix. In general, our tests indicate that including the core correction is necessary in order to reduce the systematic error of SUHF+TPSS and obtain a better description of $\Delta E_{ST}$. Further improvement could be achieved by taking into consideration the spin density to recompute the scaling factor,18 for this could permit to differentiate between strongly correlated singlets and weakly correlated open-shells.

1. Methylene and other isovalent molecules

For methylene and the isovalent NH$_2$, SiH$_2$, and PH$_2$, the lowest energy singlets are of type $\Psi^+_1$,68 Table II shows the singlet-triplet splittings calculated by different mean-field methods for these molecules. The errors for these methods, with respect to experiment,16,76–79 are also listed in this table. SUHF yields a good improvement over UHF, roughly halving the 9.0 kcal/mol MAE of the latter. In this case, adding TPSS correlation to SUHF worsens the results; $E_{\text{TPSS}}$ does not contribute to the singlet and triplet energies in a balanced way. This error is further exaggerated by SUHF+TPSS, which consistently underestimates the stability of the triplet states (ME = 11.5 kcal/mol and MAE = 11.5 kcal/mol). This is, in fact, not surprising because very little correlation is actually being captured by SUHF for the triplets, but, for their densities, it is easy to see that—because of the unpaired electrons—$\eta_{\text{ref}} > 1$ and thus $f(y, \eta_{\text{ref}}) < 1$. That is, the scaling factor does not properly distinguish that the half-filled orbitals ($\eta_t = 1$) in triplets are not really active. The result is a substantial, erroneous, reduction of the triplets’ DFT correlation. The exact same problem occurs in CAS+DFT.17,18

The splittings are given by $E_S - E_T$, ME is the mean error (theory − experiment), MAE is the mean absolute error.

### Table III. Singlet-triplet splittings (in kcal/mol) for carbon, oxygen, and silicon atoms using the cc-pVTZ basis. Experimental data were taken from Ref. 80.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>29.0</td>
<td>5.5</td>
<td>12.7</td>
<td>23.0</td>
<td>24.0</td>
<td>24.1</td>
<td>24.2</td>
<td>31.1</td>
<td>29.0</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>22.6</td>
<td>17.6</td>
<td>20.2</td>
<td>39.2</td>
<td>38.1</td>
<td>39.8</td>
<td>41.0</td>
<td>41.0</td>
<td>45.3</td>
<td></td>
</tr>
<tr>
<td>Si</td>
<td>6.6</td>
<td>4.4</td>
<td>6.4</td>
<td>9.2</td>
<td>12.4</td>
<td>12.6</td>
<td>12.1</td>
<td>21.9</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>ME</td>
<td>−11.1</td>
<td>−21.4</td>
<td>−17.4</td>
<td>−6.7</td>
<td>−5.7</td>
<td>−5.0</td>
<td>−4.8</td>
<td>−0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>11.1</td>
<td>21.4</td>
<td>17.4</td>
<td>6.7</td>
<td>5.7</td>
<td>5.0</td>
<td>4.8</td>
<td>3.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[a]Core = all occupied orbitals except valence $p$ orbitals.
[b]Core = all orbitals with $n_t > 1.99$. 

The splittings are given by $E_S - E_T$, ME is the mean error (theory − experiment), MAE is the mean absolute error.

### Table II. Singlet-triplet splittings (in kcal/mol) for methylene and other isovalent molecules predicted by different mean-field methods using the cc-pVTZ basis. Experimental data were taken from Refs. 68 and 76–79 geometries for CH$_2$ are from Ref. 71; geometries for NH$_2$, SiH$_2$, and PH$_2$ are from Ref. 68. The splittings are given by $E_S - E_T$, ME is the mean error (theory − experiment), MAE is the mean absolute error.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CH$_2$</td>
<td>16.9</td>
<td>10.3</td>
<td>13.1</td>
<td>15.6</td>
<td>12.8</td>
<td>−1.3</td>
<td>10.0</td>
<td>9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH$_2$</td>
<td>30.5</td>
<td>23.6</td>
<td>27.3</td>
<td>36.2</td>
<td>33.1</td>
<td>18.5</td>
<td>30.2</td>
<td>30.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SiH$_2$</td>
<td>−6.9</td>
<td>−15.4</td>
<td>−12.9</td>
<td>−17.7</td>
<td>−19.8</td>
<td>−28.5</td>
<td>−20.5</td>
<td>−21.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PH$_2$</td>
<td>−3.7</td>
<td>−11.1</td>
<td>−8.5</td>
<td>−15.3</td>
<td>−35.1</td>
<td>−33.9</td>
<td>−19.0</td>
<td>−17.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ME</td>
<td>9.0</td>
<td>1.6</td>
<td>4.5</td>
<td>4.5</td>
<td>−2.5</td>
<td>−11.5</td>
<td>−0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>9.0</td>
<td>4.8</td>
<td>5.9</td>
<td>4.5</td>
<td>6.4</td>
<td>11.5</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The projection in SUHF eliminates spin contamination and the results are significantly better (MAE = 6.7 kcal/mol) compared to UHF, UTPSS, or UHF+TPSS. In turn, SUHF+DFT methods all have lower MAEs than SUHF (3.7–5.7 kcal/mol). For reasons outlined in Sec. II E, we calculated here also TPSS[\(u\)] correlations, which should be very close to TPSS[\(\rho\)]. The choice of alternative densities, however, does not appear to be critical in SUHF+DFT methods, in agreement with our previous observations in Ref. 35. SUHF+\(f\), TPSS and SUHF+\(f\), TPSS[\(u\)] provide similar results with MAEs of 5.0 and 4.8 kcal/mol, respectively. The cores for these two calculations consisted of all occupied orbitals except for the valence \(p\) orbitals. Another plausible choice for the core (all orbitals with \(n_k > 1.99\)) shows further improvement (MAE = 3.7 kcal/mol) utilizing the Staroverov-Davidson densities. Although it is certainly a problem that we do not have a clear definition for inactive orbitals in SUHF, these two reasonable choices for the core have both led to results that outperform SUHF and SUHF+TPSS.

3. Diatomic molecules

The results for the diatomics NH, OH\(^+\), O\(_2\), and NF are contrasted with experimental data\(^{81}\) in Table IV. The lowest energy singlets for these species are, like in the case of the above discussed atoms, open-shells of the type \(\Psi_1^+\). The splittings for UHF, UTPSS, and UHF+TPSS are therefore highly underestimated (by ~15–20 kcal/mol) for this benchmark set too due to the large spin contamination. Again, the spin projection operator gets rid of this issue and thus SUHF sharply improves the splittings (MAE = 2.6 kcal/mol). SUHF+TPSS and SUHF+\(f\), TPSS (with a core analogous to a minimal active space calculation) also provide good results, but are slightly worse than SUHF, with errors around 3.5 kcal/mol. This could be attributed to introduction of spin contamination in the DFT correlation, or perhaps some sort of error cancellation occurring in SUHF. We also note that SUHF yields particularly good results for this specific benchmark set.\(^{54}\) The predictions that are closest to experiment are, however, given by SUHF+\(f\), TPSS[\(u\)], which has a MAE of 2.6 kcal/mol. The use of the Staroverov-Davidson densities in calculations involving open-shell singlets may thus be recommendable for SUHF+DFT methods.

4. Trimethylenemethane

In the case of trimethylenemethane (TMM), the lowest energy singlet is a strongly-correlated closed shell of type \(\Psi_1^+\). The experimental\(^{82}\) and calculated data are presented in Table V. The singlets here have less spin contamination than those of the atoms and diatomics discussed above, leading to more moderate errors for the unrestricted methods (from ~10–20 to ~5 kcal/mol). SUHF yields an excellent result that is within 1 kcal/mol of the experimental value. This approximation is not improved by SUHF+TPSS, which overestimates \(\Delta E_{ST}\) by 4.3 kcal/mol. Even though the error is still larger than in the case of SUHF, SUHF+TPSS and SUHF+\(f\), TPSS provide very good estimates of the splitting, having errors of 1.3 and 2.0 kcal/mol, respectively. Since SUHF normally misses large amounts of dynamic correlation, the better result given by this method in this case could be due to error cancellation and/or a balanced (albeit incomplete) description of correlations. It is worth noting that sophisticated CASPT2N(10, 10)/cc-pVTZ calculations predict a \(\Delta E_{ST}\) of 19.1 kcal/mol for TMM,\(^{80}\) which is very close to the values given here by SUHF and different SUHF+DFT methods.

5. Benzyne isomers

Our last example concerning singlet-triplet energy gaps consists of the \(o\)-, \(m\)-, and \(p\)-isomers of benzyne. Before continuing, a brief discussion on the geometry of \(m\)-benzyne is pertinent. According to theoretical calculations,\(^{84–87}\) there are two competing structures (see Figure 7) for the singlet of \(m\)-benzyne that are feasible minima in the potential energy surface: (1) a monocyclic diradical species, and (2) a bicyclic structure (bicyclo[3.1.0]hexatriene) with no diradical character that is well described by a single Slater determinant. Based on the comparison of high-level CCSD(T) and CASPT2 data with experimental spectra, the most plausible structure for

<table>
<thead>
<tr>
<th>Method</th>
<th>(\Delta E_{ST})</th>
<th>ME</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHF</td>
<td>23.7</td>
<td>5.6</td>
</tr>
<tr>
<td>UTPSS</td>
<td>12.0</td>
<td>6.1</td>
</tr>
<tr>
<td>UHF+TPSS</td>
<td>23.7</td>
<td>5.6</td>
</tr>
<tr>
<td>SUHF+TPSS</td>
<td>19.1</td>
<td>1.0</td>
</tr>
<tr>
<td>SUHF</td>
<td>22.4</td>
<td>4.3</td>
</tr>
<tr>
<td>SUHF+(f), TPSS</td>
<td>19.4</td>
<td>1.3</td>
</tr>
<tr>
<td>SUHF+(f), TPSS(^a)</td>
<td>20.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Expt.</td>
<td>18.1</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Core = all orbitals with \(n_k > 1.999\).
m-benzene is $^1$87,88 However, in Ref. 35 we used the benzylene UB3LYP/6-31G(d) geometries from Ref. 71, which predicted 2 to be more stable. Here, we correct that oversight by reoptimizing the geometry of m-benzene singlet at the UBLYP/6-31G(d) level, since this functional is known to converge to structure 1.89 We also remark that the SUHF and SUHF+DFT energies of the UBLYP 1 geometry are lower than those of structure 2 predicted by UB3LYP.

The lowest energy singlets for the benzenes are, like in the case of TMM, closed shell strongly-correlated systems of type $\Psi_f^s$. A comparison of the calculated and experimental $\Delta E_{ST}$ is given in Table VI. UHF has a large average error of 19.5 kcal/mol, which is slightly reduced by adding TPSS correlation. UTPSS does surprisingly well, in comparison with the rest of the methods, having the lowest MAE of 4.1 kcal/mol. In this case, SUHF has much more difficulty in accurately describing the splittings (MAE = 18.9 kcal/mol) than in the previous benchmarks. SUHF+TPSS provides small improvement over SUHF (MAE = 15.5 kcal/mol), but SUHF+TPSS and SUHF+fc/TPSS yield much better results with MAEs of 8.1 and 6.6 kcal/mol, respectively. For SUHF+fc/TPSS calculations, the core was defined in the same way as for TMM (all orbitals with $n_k > 1.999$), in the absence of a clear better option.

We close this discussion with a compilation of the singlet-triplet splitting MAEs for all the benchmark sets considered so far, which is given graphically in Figure 8. In the left panel of this figure, it is seen that SUHF+fc/TPSS is usually the best or second best option among the various mean-field methods. The right panel is, however, most interesting: it shows that UHF, UTPSS, and SUHF+TPSS have nearly the same aggregate error (~13 kcal/mol); SUHF improves upon these spin-contaminated methods (MAE = 7.1 kcal/mol), but SUHF+TPSS has virtually the same MAE as pure SUHF. Nonetheless, SUHF+fc/TPSS is capable of providing better results than SUHF, with a global MAE of 3.1 kcal/mol. Overall, it appears that SUHF+TPSS adds similar amounts of correlation to the singlet and triplet states, often resulting in a $\Delta E_{ST}$ values close to SUHF. The reason for this could be that SUHF+TPSS does not make use of information from the true, correlated, SUHF charge density $\gamma(r)$, and therefore is unable to compensate for any lack of balance in the correlations captured by SUHF. It follows from this argument that, since SUHF+fc/TPSS does utilize information from $\gamma(r)$, this method has the possibility of correcting errors in the correlations accounted for by SUHF.

### D. Ring opening of dioxirane

Static and dynamic correlation effects are both important in the ring opening process converting dioxirane 3 into methylenebis(oxy) 4 via transition state TS (see Figure 9): as the O–O bond is elongated, static correlation increases whereas dynamic correlation decreases. Accurate data for this reaction is available from MR-AQCC calculations,91 which were also used to benchmark CAS+fc/DFT in Ref. 17. Additionally, the ring opening of 3 and similar processes are important in atmospheric chemistry.91–93 All of this makes the conversion of 3 into 4 a nice example to be studied with the present SUHF+DFT methods.

Table VII shows the energies for the ring opening of dioxirane computed by a variety of schemes. Since there is more dynamic correlation in 3 than in TS or 4, methods not suited for capturing this type of correlations, such as UHF, SUHF, and CAS, predict a reaction barrier $\Delta E($TS–3) and a total change in energy $\Delta E($4–3) that are too low compared to MR-AQCC. UHF+TPSS reduces this error by about 10 kcal/mol with respect to UHF, but the average error of 31.9 kcal/mol is unacceptably large and much worse than errors in SUHF or CAS. All of the MR+DFT techniques improve upon these methods as well as UTPSS. The raw SUHF+TPSS provides a very good approximation with a MAE of 2.6 kcal/mol. This could be due to the fact that the process is dominated by the O–O bond breaking, and SUHF+TPSS provides accurate relative energies for non-metallic bond dissociations.35 The best results are given by CAS(6,4)+fc/CS (CS = Colle-Salvetti) and SUHF+fc/TPSS using five core orbitals, which have MAEs of 2.9 and 2.4 kcal/mol, respectively. The fact that MR+DFT methods provide the closest results to MR-AQCC reflects the potential of these methods to describe both types of correlation simultaneously. However, Table VII also reveals one of the problems of SUHF+fc/TPSS: the definition of the core. Predictions with different core sizes are shown in this table and it is seen that results can differ significantly (~5 kcal/mol in...

### Table VI. Singlet-triplet splittings (in kcal/mol) for the isomers of benzene using cc-pVTZ basis. Experimental data were taken from Refs. 68 and 90. For the geometries, see text.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>UHF</th>
<th>UTPSS</th>
<th>SUHF</th>
<th>SUHF+TPSS</th>
<th>SUHF+fc/TPSS</th>
<th>SUHF+fc/TPSS</th>
<th>Expt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>o-benzene</td>
<td>–15.8</td>
<td>–30.3</td>
<td>–20.6</td>
<td>–51.4</td>
<td>–51.6</td>
<td>–39.9</td>
<td>–36.3</td>
</tr>
<tr>
<td>m-benzene</td>
<td>9.4</td>
<td>–15.7</td>
<td>10.0</td>
<td>–1.6</td>
<td>–5.6</td>
<td>–8.2</td>
<td>–6.9</td>
</tr>
<tr>
<td>ME</td>
<td>15.0</td>
<td>3.9</td>
<td>14.9</td>
<td>–6.6</td>
<td>–5.9</td>
<td>–0.3</td>
<td>3.3</td>
</tr>
<tr>
<td>MAE</td>
<td>19.5</td>
<td>4.1</td>
<td>16.8</td>
<td>18.9</td>
<td>15.5</td>
<td>8.1</td>
<td>6.6</td>
</tr>
</tbody>
</table>

$^a$Core = all orbitals with $n_k > 1.999$. 

FIG. 7. Two possible isomers of $m$-benzene: the diradical (1) and bicyclo[3.1.0]hexatriene (2).
E. Beryllium isoelectronic series

We close this section by returning to the theme with which we opened it: isoelectronic series. In contrast with the He isoelectronic series in which there is only dynamic correlation, the Be isoelectronic series is characterized by an important static $s^2 \rightarrow p^2$ contribution to the energy. The trend of the series is radically changed by the presence of static correlation: the correlation energy scales linearly with the atomic charge $Z$. We notice too that the Be isoelectronic series has been used for benchmarking other MR+DFT methods, e.g., in Refs. 16 and 27.

Figure 10 compares the correlation energies given by SUHF and different SUHF+DFT methods with accurate values from Ref. 94. Because SUHF captures static correlations, it correctly reproduces the linear increase in the correlation energy as a function of $Z$. However, the overall SUHF energies are significantly higher than the accurate values and the slope in SUHF correlation is not steep enough. Addition of TPSS dynamic correlation to SUHF improves the total energies, but the quality of all of the SUHF+DFT methods decreases as $Z$ increases: the errors of about 1–3 kcal/mol for Li$^-$ become 30–40 kcal/mol for Ne$^+$. Extending the basis set from cc-pVTZ to cc-pV5Z does not eliminate this large error; the energy is reduced by only $\sim 5$ kcal/mol. Also, the trend in the SUHF+DFT correlation is no longer linear, but rather closer to the ln$Z$ behavior of CAS(2,2)$+_{DFT}$. In CAS+DFT, this erroneous tendency can be corrected by extending the size of the active space, something that cannot be done in SUHF. Thus, the Be isoelectronic series serves to illustrate some of the limitations of SUHF+DFT methods. A possible solution to this problem is discussed in Sec. IV.

TABLE VII. Energies (in kcal/mol) for the ring opening of dioxirane 3 computed by different methods. UHF, SUHF, UTPSS, and SUHF+DFT calculations use cc-pVTZ basis. CAS+DFT and MR-AQCC use the 6-311+G(2df,2pd) basis and were taken from Refs. 17 and 91. The number of core orbitals in SUHF+fcTPSS is indicated in parenthesis.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\Delta E$(TS-3)</th>
<th>$\Delta E$(4-3)</th>
<th>ME</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHF</td>
<td>−26.6</td>
<td>−37.8</td>
<td>−43.7</td>
<td>43.7</td>
</tr>
<tr>
<td>SUHF</td>
<td>9.9</td>
<td>−1.1</td>
<td>−8.0</td>
<td>8.0</td>
</tr>
<tr>
<td>UTPSS</td>
<td>23.2</td>
<td>14.5</td>
<td>6.4</td>
<td>6.4</td>
</tr>
<tr>
<td>SUHF+TPSS</td>
<td>−15.0</td>
<td>−23.9</td>
<td>−31.9</td>
<td>31.9</td>
</tr>
<tr>
<td>SUHF+fcTPSS</td>
<td>16.3</td>
<td>5.8</td>
<td>−1.4</td>
<td>2.6</td>
</tr>
<tr>
<td>SUHF+(f,3)TPSS</td>
<td>12.5</td>
<td>1.5</td>
<td>−5.4</td>
<td>5.4</td>
</tr>
<tr>
<td>SUHF+(f,5)TPSS</td>
<td>14.4</td>
<td>3.0</td>
<td>−3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>SUHF+(f,6)TPSS</td>
<td>18.9</td>
<td>8.0</td>
<td>1.0</td>
<td>2.4</td>
</tr>
<tr>
<td>CAS(6,4)</td>
<td>11.1</td>
<td>2.3</td>
<td>−5.8</td>
<td>5.8</td>
</tr>
<tr>
<td>CAS(6,4)+CS</td>
<td>10.5</td>
<td>1.9</td>
<td>−6.3</td>
<td>6.3</td>
</tr>
<tr>
<td>CAS(6,4)+CS</td>
<td>16.9</td>
<td>8.8</td>
<td>0.4</td>
<td>3.8</td>
</tr>
<tr>
<td>CAS(6,4)+fCS</td>
<td>13.1</td>
<td>4.4</td>
<td>−3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>CAS(6,4)+fCS</td>
<td>17.4</td>
<td>7.4</td>
<td>0.1</td>
<td>2.9</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>20.3</td>
<td>4.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
avoided by using active space based spin-projected Hartree-Fock (SUHF	extsuperscript{a}), rather than standard SUHF. In SUHF	extsuperscript{a}, there are core orbitals analogous to inactive orbitals in CAS, that is, symmetry breaking in the deformed determinant can only occur within an active space. SUHF	extsuperscript{a} calculations have been described, e.g., in Refs. 34, 95, and 96 and they yield energies comparable to, but higher than, CAS (provided that active spaces are equal, of course). This does not represent a problem for SUHF	extsuperscript{a}+f,DFT since the scaling factor would take care of the dynamic correlations missing from SUHF	extsuperscript{a} with respect to CAS.

- **Distinction between open- and closed-shells:** As noted in the discussion regarding the SUHF+f,DFT splittings for methane in Sec. III C 1, the scaling factor does not properly distinguish between singly occupied orbitals in triplets and strongly correlated singlets. More precisely, for the triplets, \( \eta_{\text{ref}} > 1 \) (always) and thus \( f(\gamma, \eta_{\text{ref}}) < 1 \), regardless of whether or not there is dynamic correlation in \( \gamma(r) \). Triplet states are therefore unevenly destabilized as compared to the singlet states. Gräfenstein and Cremer\textsuperscript{18} have proposed a fix for this shortcoming that consists in a reparametrization of the scaling factor using the spin density. An alternative possibility is to calculate spin-dependent reference densities

\[
\rho^\gamma_{\text{ref}}(r) = \sum_k \left[ \psi^\alpha_k(r) \right]^2, 
\]

and measures of the active space

\[
\eta^\gamma_{\text{ref}}(r) = \left( \frac{\rho^\alpha_{\text{ref}}(r)}{\gamma(r)} \right)^{1/3},
\]

to then compute the scaling factor as

\[
f(r) = \frac{f(\gamma, \eta_{\text{ref}}^\alpha) \rho^\alpha(r) + f(\gamma, \eta_{\text{ref}}^\beta) \rho^\beta(r)}{\rho^\alpha(r) + \rho^\beta(r)}.
\]

For singlets we have that \( \gamma^\alpha = \gamma^\beta = \gamma/2 \), and hence this new scaling factor leads results that are indistinguishable from those given by \( f(\gamma, \eta_{\text{ref}}) \) in Eq. (10). However, the \( f(r) \) in Eq. (26) will not artificially scale down the DFT correlation for triplet states dominated by a single Slater determinant, which is the problem that we wanted to solve.

- **Spin contamination:** Our current SUHF+DFT schemes all use spin-contaminated (symmetry-broken) densities as inputs for \( E^\text{DFT}_c \). This can be the source of error in the description of spin multiplets\textsuperscript{18,69,97}. To alleviate this issue, one may use functionals that take as input the total and pair densities. The most well known example of such a functional is Colle-Salvetti,\textsuperscript{11} which could be reparametrized for SUHF+DFT applications. Other possible two-body functionals have been derived by Moscardó and San-Fabían,\textsuperscript{98} as well as Moscardó and Pérez-Jiménez,\textsuperscript{99} an analysis of these functionals, along with Colle-Salvetti, has been carried out in Ref. 100. An alternative solution is to employ the densities defined by Eq. (3), and adapt modern functionals so that correlation energies can be extrapolated to imaginary values of the spin density. Also, instead of this extrapolation, a redefinition of \( P_2(r) \) as \( P_2(r) = 0 \) or \( P_2(r) = \gamma(r)^2/4 \) when the spin density turns imaginary is possible too.\textsuperscript{16}

- **Inadequacy of the SUHF reference:** This problem is illustrated in the case of the Be isoelectronic series (Sec. III E), where it was noted that SUHF+DFT errors were similar to those in CAS(2,2)+DFT, and that the latter can improve results by enlarging the active space. Even though SUHF will generally provide an improved wavefunction over HF, it is clear that not all strong correlations can be accurately described in terms of spin symmetry breaking. There are strong correlations that are due to spatial-symmetry degeneracies, and breaking and restoring such symmetries may provide a better reference for our projected Hartree-Fock (PHF)+DFT formalism.

Yet another way to improve the SUHF reference is to break and restore further symmetries of the Hamiltonian (see Refs. 34 and 40), e.g., \( \hat{S}_z \) and complex conjugation. The more symmetries that are broken and restored, the more correlations that are captured and the better the reference will become. An example where the SUHF reference is inadequate for MR+DFT, but KSUHF (which breaks and restores both \( \hat{S}_z \) and complex conjugation) improves results, occurs in the calculations on the chromium dimer detailed in Ref. 35. In that case, KSUHF+DFT provided a qualitatively correct description of the dissociation of \( \text{Cr}_2 \), whereas the SUHF+DFT curve was strongly repulsive in the region near the experimental bond length. It is worth noting that the inclusion of additional correlations by PHF methods more general than SUHF leads to severe double counting in raw PHF+DFT, but this issue is resolved by using the scaling factor of the correlation \( f(\gamma, \eta_{\text{ref}}) \). Finally, another strategy that we could pursue is to consider a short expansion in terms of symmetry-projected configurations (Ref. 101).
 Hence, all of the major problems leading to inaccuracies in the SUHF+DFT can be mitigated via relatively simple modifications to the current schemes. Also, the fine tuning of DFT correlation functional parameters for application in SUHF+DFT (or PHF+DFT) can be helpful to achieve further accuracy. All of this opens the possibility to improve upon the result reported here.

V. CONCLUSIONS

We have adapted available schemes to avoid double counting in CAS+DFT to SUHF+DFT. The derived SUHF+/DFT and SUHF+fc/DFT methods were shown to improve upon SUHF for the description benchmarks such as isoelectronic series, molecular dissociations, and the ring opening of dioxirane. Furthermore, we demonstrated that straightforward addition of DFT correlation to SUHF does not improve results in very sensitive tests such as singlet-triplet splittings. However, SUHF+fc/DFT can improve upon SUHF and other mean-field methods in these challenging tests. The SUHF+DFT methods presented in this paper are not without their own problems; nevertheless, we identified these problems and proposed concrete, feasible, solutions to them which are rather simple extensions of the current schemes. Improving upon the results reported here is therefore a very real possibility. Since our SUHF+DFT implementations have a mean-field computational scaling, we believe that these kind of methods should be viewed as low cost alternatives within the MR+DFT family. Finally, the results and ideas here discussed and reviewed should also be of utility for anyone trying to couple other MR techniques with DFT.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (CHE-1110884) and the Welch Foundation (C-0036).