\textbf{\textit{$\Delta$NO} and the complexities of electron correlation in simple hydrogen clusters}

Ismael A. Elayan,\textsuperscript{1} Rishabh Gupta,\textsuperscript{1} and Joshua W. Hollett\textsuperscript{1, 2, a)}

\textit{1) Department of Chemistry, University of Winnipeg, Winnipeg, Manitoba, R3B 2E9, Canada}
\textit{2) Department of Chemistry, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada}

\text{(Dated: 26 January 2022)}

The $\Delta$ natural orbital ($\Delta$NO) two-electron density matrix (2-RDM) and energy expression are derived from a multideterminantal wave function. The approximate $\Delta$NO 2-RDM is combined with an on-top density functional and a double-counting correction to capture electron correlation. A trust-region Newton’s method optimization algorithm for the simultaneous optimization of $\Delta$NO orbitals and occupancies is introduced and compared to the previous iterative diagonalization algorithm. The combination of $\Delta$NO and two different on-top density functionals, Colle-Salvetti (CS) and Opposite-spin exponential cusp and Fermi-hole correction (OF), is assessed on small hydrogen clusters and compared to density functional, single-reference coupled cluster, and multireference perturbation theory (MRMP2) methods. The $\Delta$NO-CS and $\Delta$NO-OF methods outperform the single-reference methods, and are comparable to MRMP2. However, there is a distinct qualitative error in the $\Delta$NO potential energy surface for H$_4$ compared to the exact. This discrepancy is explained through analysis of the $\Delta$NO orbitals, occupancies and the two-electron density.

\begin{multicols}{2}
\section{I. INTRODUCTION}

In an effort to create efficient electron structure methods, it has become commonplace to combine contrasting models of electron correlation (\textit{e.g.} wave function, density functional, density matrix functional) in order to exploit their most effective attributes.\textsuperscript{1–41} Current methods that employ such a strategy, vary in computational cost and accuracy, and in most cases their remains room for improvement in both departments. The most common strategy is to decompose electron correlation into static (strong) and dynamic (weak) correlation, and approach each differently. The development of these methods involves the design of models that precisely capture each component, and also blend the approaches together in a seamless fashion. The former aspect can be thoroughly evaluated by applying the method to relatively simple systems, such as small hydrogen clusters.

Although they are not chemically complex, the electronic structure of H atom clusters, particularly linear chains and regular lattices of H atoms, can serve as a systematic testing ground for new treatments of electronic structure of H atom clusters, particularly linear systems, such as small hydrogen clusters. 

The $\Delta$NO method combines a cumulant functional approach for the treatment of static correlation with an on-top density functional (or post-HF) to treat dynamic correlation.\textsuperscript{12,34} Amongst other approaches, it most resembles multiconfigurational-pair density functional theory (MC-PDFT).\textsuperscript{28–30,40,68} In which, a multireference one-electron density matrix (1-RDM), and subsequent one-electron density, is used to determine the kinetic, nuclear-potential, and classical Coulomb energy. The on-top pair density, obtained from the multireference two-electron density matrix (2-RDM), is used to create a modified spin-density\textsuperscript{28} which is inserted into the usual exchange-correlation density functionals (\textit{e.g.} PBE)\textsuperscript{69} to determine the exchange and correlation energy. The 2-RDM can be obtained from a traditional multireference self-consistent-field\textsuperscript{28–30,68}, a variational 2-RDM-driven complete active-space self-consistent field\textsuperscript{40}, or another economical alternative such as a pair-coupled cluster doubles calculation.\textsuperscript{70} The $\Delta$NO method uses an on-top density functional for only the dynamic correlation energy, where the functional depends directly on the on-top density. The remaining components of the energy are determined using the approximate $\Delta$NO 2-RDM.

This work begins by introducing a $\Delta$NO wave function (Subsection II A) from which the $\Delta$NO two-electron density matrix and energy expression can be derived. The $\Delta$NO 2-RDM is then supplemented with an \textit{ad hoc} term to correct for missing static correlation between electron pairs (Subsection II B). This is followed by a description of the determination of the dynamic correlation energy from the $\Delta$NO 2-RDM (Subsection II C), and then an approach to reduce correlation double-counting (Subsection II D). Then a new optimization algorithm for the $\Delta$NO 2-RDM is described (Subsection II E), which involves the simultaneous optimization of the orbitals and the electron transfer variables. The performance of the optimization method is then evaluated relative to the previous

\begin{flushleft}
\textsuperscript{a) Corresponding author: j.hollett@uwinnipeg.ca}
\end{flushleft}
The performance of ∆NO, in describing potential energy surfaces of small H clusters, is then compared to established single- and multireference methods (Section IV). The two-electron density is then used to analyze the static correlation provided by the ∆NO model in the rectangular H₄ cluster (Subsection IV G). Finally, the results of the ∆NO evaluation and analysis are summarized (Section V).

II. THEORY

A. The ∆NO wave function

The multideterminantal ∆NO wave function,

$$\Psi_{\Delta NO} = \sum_{I} c_{I} \Psi_{I}, \quad (1)$$

is defined with respect to a reference configuration in which all open-shell (singly occupied) orbitals are considered frozen (i.e., singly occupied in all determinants). Only doubly occupied orbitals are active and all excitations are 2n-tuples (doubles, quadruples, ...) where no electrons are unpaired. In other words, with respect to the active orbitals, $$\Psi_{\Delta NO}$$ is seniority-zero.\(^{53,63,71–77}\) Also, each active pair of electrons has its own active space, consisting of the doubly occupied orbital in the reference wave function and some number of vacant (virtual) orbitals. Each determinant of the expansion is then defined by 2n-tuple excitations of the reference configuration,

$$\Psi_{I} = \Psi_{\alpha i \beta j \ldots \gamma k \ldots}, \quad (2)$$

where $$i, j$$ and $$k$$ refer to reference occupied orbitals, and $$\alpha, \beta$$ and $$\gamma$$ refer to all active orbitals. The absence of a bar, $$i$$, indicates an $$\alpha$$-spin orbital while the presence of a bar, $$\bar{i}$$, indicates a $$\beta$$-spin orbital. Considering this, it is useful to define the determinantal expansion coefficients, $$c_{I}$$, as a product of coefficients corresponding to specific double excitations of the active electron pairs,

$$c_{I} = d_{i}^{a} d_{j}^{b} d_{k}^{c} \ldots, \quad (3)$$

In order to arrive at the ∆NO formalism, the double excitation coefficients, $$d_{i}^{a}$$, are defined in terms of orbital occupancies, $$n_{i}$$, and electron transfer variables, $$\Delta_{ia}$$,

$$d_{i}^{a} = \delta_{ia} \sqrt{n_{i}} - \sqrt{\Delta_{ia}}, \quad (4)$$

where $$\delta_{ia}$$ is the Kroenecker delta and $$\Delta_{i} = 0$$. Orbital occupancies, occupied and vacant (with respect to the reference configuration), are defined as

$$n_{i} = 1 - \sum_{a} \Delta_{ia}, \quad (5)$$

$$n_{a} = \sum_{i} \Delta_{ia}. \quad (6)$$

The N-representable components of the ∆NO energy expression, $$E^{\Delta NO}$$, can then be obtained by expanding the expectation value of the N-electron hamiltonian in terms of $$n_{i}, \Delta_{ia},$$ and one and two-electron integrals over orbitals. Initially, the energy may be divided into diagonal and off-diagonal contributions,

$$E^{\Delta NO'} = \langle \Psi_{\Delta NO}' | \hat{H} | \Psi_{\Delta NO}' \rangle = \sum_{I} c_{I}^{2} H_{II} + \sum_{I \neq J} c_{I} c_{J} H_{IJ}, \quad (7)$$

where

$$H_{IJ} = \langle \Psi_{I} | \hat{H} | \Psi_{J} \rangle, \quad (8)$$

and the electronic hamiltonian (in atomic units) is given by

$$\hat{H} = \sum_{i=1}^{N} -\frac{1}{2} \nabla_{i}^{2} - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_{A}}{r_{iA}} + \sum_{i=2}^{N} \sum_{j<i}^{N} \frac{1}{r_{ij}}, \quad (9)$$

where $$M$$ is the number of nuclei, $$Z_{A}$$ is the nuclear charge, $$r_{iA}$$ is the electron-nucleus distance, and $$r_{ij}$$ is the electron-electron distance. The diagonal contributions can be expressed as,

$$\sum_{I} c_{I}^{2} H_{II} = E^{\Delta NO}_{0-1RDM} + E^{\Delta NO}_{pair}, \quad (10)$$

where $$E^{\Delta NO}_{0-1RDM}$$ is the energy associated with the zeroth-order term in the expansion of the two-electron density matrix (2-RDM) in terms of the one-electron density matrix (1-RDM),\(^{34}\) which is the basis of cumulant functional theory.\(^{78}\) The 0-1RDM energy is expressed as,

$$E^{\Delta NO}_{0-1RDM} = \sum_{p}^{cl} 2n_{p} h_{pp} + \sum_{p}^{op} n_{p} h_{pp} + \sum_{p}^{cl} \sum_{q=1}^{cl} n_{p} n_{q} (2J_{pq} - K_{pq})$$

$$+ \sum_{p}^{cl} \sum_{q=1}^{op} n_{p} n_{q} (2J_{pq} - K_{pq})$$

$$+ \sum_{p}^{op} \sum_{q=1}^{op} n_{p} n_{q} (J_{pq} - K_{pq}), \quad (11)$$

where cl and op denote the set of closed and open-shell orbitals, respectively. The one-electron integrals over spatial orbitals, $$\phi_{p}$$, are defined as,

$$h_{pp} = \int \phi_{p}^{*}(r) \left[ -\frac{1}{2} \nabla^{2} - \sum_{A} \frac{Z_{A}}{r_{A}} \right] \phi_{p}(r) dr \quad (12)$$

where $$r = (x, y, z)$$ and the two-electron Coulomb, $$J_{pq}$$, and exchange, $$K_{pq}$$, integrals are defined as,

$$J_{pq} = \int \frac{\phi_{p}^{*}(r_{1}) \phi_{q}^{*}(r_{2}) \phi_{p}(r_{1}) \phi_{q}(r_{2})}{r_{12}} dr_{1} dr_{2} \quad (13)$$
and
\[ K_{pq} = \int \frac{\phi^*(p_1)\phi^*(p_2)\phi_q(p_1)\phi(p_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2. \] (14)

After expanding the expectation values for the diagonal contributions to \( \DeltaNO \) [first term of Equation (7)], simplifying, and removing the contribution of \( \Delta_{1RDM} \), the pair correction energy, \( \DeltaNO_{\text{pair}} \), remains. The pair correction energy is the first term associated with the cumulant within the strong-correlation limit. Nevertheless, it proved effective previously\(^{34} \) and is added to the \( \DeltaNO \) energy, derived in Subsection II A, to give the total \( \DeltaNO \) energy,
\[ \DeltaNO = \DeltaNO_{\text{stat}} + \DeltaNO_{\text{pair}}. \]

B. The high-spin correction

As defined above, the \( \DeltaNO \) wave function and the subsequent 2-RDM do not account for the correlation between active electron pairs (interpair correlation). The high-spin correction is added to \( \DeltaNO \) to account for this correlation exactly in the strong correlation limit (e.g. multiple bond dissociation). The term is \textit{ad hoc}, and therefore does not guarantee \( N \)-representability below the strong correlation limit. Nevertheless, it proved effective previously\(^{34} \) and is added to the \( N \)-representable \( \DeltaNO \) energy, derived in Subsection II A, to give the total \( \DeltaNO \) energy,
\[ \DeltaNO = \DeltaNO_{\text{stat}} + \DeltaNO_{\text{pair}} + \DeltaNO_{\text{HSC}}. \]

The high-spin correction energy is defined as
\[ \DeltaNO_{\text{HSC}} = - \sum_{pq} \kappa_{pq} K_{pq}, \]
where
\[ \kappa_{pq} = (1 - \delta_{pq}) \left[ \sum_{r \neq s} \xi_{pr}\xi_{ps} + W_p n_p n_q \frac{\sqrt{\Delta r_p \Delta r_q}}{2\sqrt{2}} \right] , \]

where \( \Delta_{1RDM} \) is an upper bound to the \( \Delta_{1RDM} \), the off-diagonal terms of the expectation value expansion are responsible for electron correlation. In the case of \( \DeltaNO \), excitations are limited to low-lying near-degenerate orbitals which is aimed at capturing the so-called “static” correlation energy,
\[ \sum_{IJ} c_{IJ} c_{IJ} H_{IJ} = \DeltaNO_{\text{stat}} \]
and
\[ \xi_{pq} = O_p V_q \sqrt{n_p \Delta_{pq}} + O_q V_p \sqrt{n_q \Delta_{qp}}. \] (22)

The time-inversion exchange integral is defined as
\[ L_{pq} = \int \frac{\phi_p^*(r_1)\phi_p^*(r_2)\phi_q(r_1)\phi_q(r_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2. \] (23)

The \( \DeltaNO \) wave function contains only 2\( \times \)2-tuple excitations, in which the electrons remained paired, and each electron pair is excited to its own (restricted) active space. Therefore, \( \DeltaNO \) is an upper bound to the more general doubly-occupied configuration interaction (DOCI) energy\(^{53,80} \).
C. Dynamic correlation

As mentioned above, the ∆NO wave function is designed to capture static correlation, which means dynamic correlation must be accounted for by some other means. In this study, the dynamic correlation energy is captured using an on-top density functional (ODF). The dynamic correlation energy, $E^\Delta_{\text{NO-dyn}}$, is added to the total ∆NO energy to give the total energy including static and dynamic correlation,

$$E^\Delta_{\text{NO-dyn}} = E^\Delta_{\text{NO}} + E^\Delta_{\text{dy}}. \quad (28)$$

Within the current formalism, it is necessary to remove extra correlation that is erroneously included, or “double counted”. Therefore, the total dynamic correlation energy is the sum of the dynamic correlation energy supplied by the ODF and a double-counting correction,

$$E^\Delta_{\text{dy}} = E_{\text{ODF}}[\Gamma^\Delta_{\text{NO}}(\mathbf{r}, \mathbf{r})] + E_{\text{DC}}[\Gamma^\Delta_{\text{NO}}(\mathbf{r}, \mathbf{r})], \quad (29)$$

where both are a function of the two-electron on-top density, $\Gamma(\mathbf{r}, \mathbf{r})$. The on-top density is the probability that two-electrons may be found at the same coordinate, $\mathbf{r}$. It may be calculated from the 2-RDM, which is given by

$$\Gamma(x_1, x_2, x'_1, x'_2) = \frac{N(N - 1)}{2} \int \Psi^*(x_1', x_2', x_3, \ldots, x_N) \times \Psi(x_1, x_2, x_3, \ldots, x_N) dx_3 \ldots dx_N, \quad (30)$$

where $\Psi$ is the $N$-electron wave function and $x = (\mathbf{r}, \omega)$ is a combination of spatial $\mathbf{r}$ and spin $\omega$ coordinates. The spinless 2-RDM may be obtained through integration over the spin-coordinates,

$$\Gamma(r_1, r_2, r'_1, r'_2) = \int \Gamma(x_1, x_2, x'_1, x'_2) \omega'_1 = \omega_1 \omega'_2 = \omega_2 \, d\omega_1 \, d\omega_2,$$

and the aforementioned two-electron density is the diagonal, $\Gamma(r_1, r_2) = \Gamma(r_1, r_2, r_1, r_2)$. The spinless 2-RDM can then be resolved into different spin-components corresponding to the relative spin of electrons 1 and 2,

$$\Gamma(r_1, r_2, r'_1, r'_2) = \Gamma^{\alpha\alpha}(r_1, r_2, r'_1, r'_2) + \Gamma^{\beta\beta}(r_1, r_2, r'_1, r'_2) + \Gamma^{\alpha\beta}(r_1, r_2, r'_1, r'_2) + \Gamma^{\beta\alpha}(r_1, r_2, r'_1, r'_2), \quad (32)$$

where $\alpha$ and $\beta$ denote spin-up and spin-down electrons, respectively.

The 2-RDM, and its components, can also be expanded in an orbital basis,

$$\Gamma(r_1, r_2, r'_1, r'_2) = \sum_{pqrs} \Gamma_{pqrs} \phi_p^*(r_1') \phi_q^*(r_2') \phi_p(r_1) \phi_q(r_2). \quad (33)$$

where, in this case, the $\phi_p$ are the ∆NO orbitals. Following a procedure similar to the derivation of $E^\Delta_{\text{NO}}$, but using second quantization, one can arrive at an expression for the elements of the spin-resolved ∆NO 2-RDM over the ∆NO basis,

$$\Gamma^\Delta_{pqrs} = \frac{f_{pq}^\sigma n_p n_q + \eta_{pq} s_{qs}}{2} \delta_{pr} \delta_{qs} \quad (34)$$

and

$$\Gamma^\Delta_{pqrs} = \frac{f_{pq}^\sigma' n_p n_q + \eta_{pq} s_{qs}}{2} \delta_{pr} \delta_{qs} \quad (35)$$

where $\sigma = \alpha$ or $\beta$ (with $\sigma \neq \sigma'$), $\delta_{pq} = \delta_{pr} \delta_{qs} - \delta_{ps} \delta_{qr}$ and the $f_{pq}^\sigma$ ensures there are only interactions with $\alpha$ electrons in singly-occupied orbitals,

$$f_{pq}^\alpha = 1, \quad (36a)$$
$$f_{pq}^\beta = 1 - W_q, \quad (36b)$$
$$f_{pq}^\alpha = 1 - W_p, \quad (36c)$$
$$f_{pq}^\beta = (1 - W_p)(1 - W_q). \quad (36d)$$

Two ODFs are used in this study. The first is a combination of the Opposite-spin exponential cusp (OSEC) and Fermi-hole correction (FHC) functionals, referred to as the OF functional.$^5$ The second is the Colle-Salvetti (CS) functional.$^{81,82}$ The OSEC functional depends on the total $\alpha\beta$ on-top density, whereas the FHC functional depends on the $\alpha\alpha$ and $\beta\beta$ on-top densities (more specifically, their Laplacians),

$$E_{\text{OF}}[\Gamma(\mathbf{r}, \mathbf{r})] = E_{\text{OSSEC}}[\Gamma^{\alpha\alpha}(\mathbf{r}, \mathbf{r}) + \Gamma^{\alpha\beta}(\mathbf{r}, \mathbf{r})] + E_{\text{FHC}}[\Gamma^{\beta\beta}(\mathbf{r}, \mathbf{r})] + E_{\text{FHC}}[\Gamma^{\beta\alpha}(\mathbf{r}, \mathbf{r})] \quad (37)$$

The FHC functional effectively widens the Fermi-hole, which exists in any 2-RDM that obeys antisymmetry, using an exponential-cusp ansatz for the second-order cusp. The functional is given by
where \( L^{\sigma\sigma}(\mathbf{r}) \) is the Laplacian of the two-electron density with respect to the interelectronic coordinate, \( \mathbf{u} = \mathbf{r}_1 - \mathbf{r}_2 \), at the electron-electron coalescence point (\( u = 0 \)),

\[
L^{\sigma\sigma}(\mathbf{r}) = \nabla^2 \Gamma^{\sigma\sigma}(\mathbf{r}, \mathbf{u}/2, \mathbf{r} - \mathbf{u}/2)|_{u=0},
\]

where \( \mathbf{r} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \). The \( \delta \) determines the correlation length,

\[
\delta \equiv \delta(\mathbf{r}) = \kappa L^{\sigma\sigma}(\mathbf{r})^{1/8}
\]

The OSEC functional approximates the input \( \alpha\beta \) two-electron density as constant in the neighbourhood of the electron-electron coalescence point, and determines the energy associated with introducing an exponential cusp. The OSEC energy is given by the functional

\[
E_{\text{OSEC}}[\Gamma^{\alpha\beta}(\mathbf{r}, \mathbf{r})] = \int \frac{2\pi \left(2\lambda \left(\sqrt{\pi\lambda} - 1\right) e^{-\frac{1}{2\lambda}} + \sqrt{\pi} \left(\sqrt{\pi\lambda} - 2\lambda^2 - 1\right) \left[1 + \text{erf} \left(\frac{1}{2\sqrt{\lambda}}\right)\right]\right) \Gamma^{\alpha\beta}(\mathbf{r}, \mathbf{r})}{\lambda^2 \left(2\lambda e^{-\frac{1}{2\lambda}} + \sqrt{\pi} (1 + 2\lambda^2) \left[1 + \text{erf} \left(\frac{1}{2\sqrt{\lambda}}\right)\right]\right)} d\mathbf{r},
\]

where the correlation length is determined by \( \lambda \),

\[
\lambda \equiv \lambda(\mathbf{r}) = q_{\text{OSEC}} \rho(\mathbf{r})^{1/3}
\]

with \( q_{\text{OSEC}} = 2.54 \) and \( \rho(\mathbf{r}) \) is the one-electron density.\(^{81}\)

The Colle-Salvetti ODF depends on the total (sum of all spin components) on-top density,

\[
E_{\text{CS}}[\Gamma(\mathbf{r}, \mathbf{r})] = -4a \int \frac{\Gamma(\mathbf{r}, \mathbf{r})}{\rho(\mathbf{r})} \left(\frac{1 + b\rho(\mathbf{r})^{-8/3} e^{-c\rho(\mathbf{r})^{-1/3}} L(\mathbf{r})}{1 + d\rho(\mathbf{r})^{-1/3}}\right) d\mathbf{r}
\]

where \( L(\mathbf{r}) \) is the Laplacian of the total on-top density at \( u = 0 \), \( a = 0.049 \), \( b = 0.132 \), \( c = 0.2533 \) and \( d = 0.349.\(^{82}\)

### D. Double-counting correction

The ODFs, OF and CS, are both derived from correcting the two-electron density in the neighbourhood of the electron-electron coalescence point, or in the short-range of interelectronic distance. Multideterminantal treatments of electron correlation, such as \( \Delta \text{NO} \), capture the full range of electron correlation. Therefore, there is overlap, or double-counting, that can occur when wave function methods are combined with density functional methods in the present manner. The main motivation for specifically using ODFs with \( \Delta \text{NO} \), is that as electrons become strongly correlated this information is revealed to the ODF via \( \Gamma^{\Delta \text{NO}} \). However, the ODF only depends on the OD, \( \Gamma^{\Delta \text{NO}}(\mathbf{r}, \mathbf{r}) \), which does not possess information regarding \( \Delta \text{NO} \) short-range correlation beyond the electron-electron coalescence point (in the neighbourhood of \( u = 0 \)). Because the ODF is not sensitive to the effects of correlation at small \( u \neq 0 \), a correction must be applied.

The general scheme behind the derivation of the OF functional\(^{81}\) (similar to that of the CS functional\(^{82}\)), is the approximation of the exact two-electron density \( \Gamma(\mathbf{r}_1, \mathbf{r}_2) \) by the uncorrelated two-electron density \( \Gamma^0(\mathbf{r}_1, \mathbf{r}_2) \) with a cusp, \( c(u) \), inserted at \( u = 0 \),

\[
\Gamma(\mathbf{r}_1, \mathbf{r}_2) = \Gamma^0(\mathbf{r}_1, \mathbf{r}_2) (1 - g(u) [1 - \Phi(\mathbf{r}) c(u)])
\]

where \( \Phi(\mathbf{r}) \) is a function that maintains the normalization of \( \Gamma(\mathbf{r}_1, \mathbf{r}_2) \). The correction is applied in the short-range (i.e. small \( u \)) and the range is controlled by the function \( g(u) \),

\[
g(u) = e^{-\lambda(\mathbf{r}) u^2},
\]

which depends on the correlation length \( \lambda(\mathbf{r}) \). The derivation is completed by Taylor-expanding \( \Gamma^0(\mathbf{r}_1, \mathbf{r}_2) \)
about \( u = 0 \). The expansion is zeroth-order for the OSEC functional,
\[
\Gamma^0 (\mathbf{r}_1, \mathbf{r}_2) \approx \Gamma^0 (\mathbf{r}, \mathbf{r}) ,
\]
and second-order for the FHC functional. The CS functional also uses a second-order expansion, but multiple approximations are made to ensure the integrand is well-behaved.

If the ODF is to exclusively account for the correlation energy in the short-range, then the short-range correlation already present in \( \Gamma \) must be removed,
\[
E_{\text{ODF}}^{\Delta \text{NO}} = E_{\text{ODF}}^{\Delta \text{NO}} [\Gamma] - \frac{1}{2} \int \frac{[\Gamma(\mathbf{r}_1, \mathbf{r}_2) - \Gamma^0(\mathbf{r}_1, \mathbf{r}_2)] g(u)}{u} d\mathbf{r}_1 d\mathbf{r}_2 .
\]

To arrive at a practical expression for the double-counting correction, \( \Gamma(\mathbf{r}_1, \mathbf{r}_2) \) and \( \Gamma^0(\mathbf{r}_1, \mathbf{r}_2) \), are treated with a zeroth-order Taylor expansion about \( u = 0 \) (\textit{i.e.} constant), identical to the OSEC functional. Also, the double-counting correction is not required in the strong correlation limit, when correlation causes electrons to completely avoid each other and \( \Gamma(\mathbf{r}, \mathbf{r}) = 0 \), and must be turned off. Therefore, the correction is multiplied by an electron-pair specific factor, \( \Gamma_{pp} \), and the resulting double-counting correction is given as
\[
E_{\text{DC}}^{\Delta \text{NO}} [\Gamma(\mathbf{r}, \mathbf{r})] = 2\pi c_{\text{DC}} \int \sum_p O_p \left( \frac{\Gamma_{pp}}{\Gamma_{pp}^0} \right) \frac{\Gamma_{pp}^0(\mathbf{r}, \mathbf{r}) - \Gamma_{pp}(\mathbf{r}, \mathbf{r})}{\lambda(\mathbf{r})^2} d\mathbf{r}
\]

where \( \Gamma^0 \) is the uncorrelated two-electron density, which for \( \Delta \text{NO} \) is given by
\[
\Gamma^{\Delta \text{NO},0} = \Gamma_{0-1\text{RDM}} + \Gamma_{\text{pair}}^{\Delta \text{NO}} .
\]

The two-electron densities are decomposed into electron-pair components, which for \( u = 0 \), are given by
\[
\Gamma_{pp}^0 (\mathbf{r}, \mathbf{r}) = n_p |\phi_p(\mathbf{r})|^4 + \sum_q \Delta_{pq} |\phi_q(\mathbf{r})|^4 ,
\]

and
\[
\Gamma_{pp} (\mathbf{r}, \mathbf{r}) = \Gamma_{pp}^0 (\mathbf{r}, \mathbf{r}) + \sum_{q,s} \Delta_{pq} \Delta_{rs} |\phi_s(\mathbf{r})|^2 |\phi_q(\mathbf{r})|^2
\]
\[
- 2 \sum_q \xi_{pq} |\phi_p(\mathbf{r})|^2 |\phi_q(\mathbf{r})|^2 .
\]

The above correction is added to the OF functional energy to give the total dynamic correlation energy. In the case of the CS functional,
\[
\lambda(\mathbf{r}) = q_{\text{CS}} \rho(\mathbf{r})^{1/3}
\]
where \( q_{\text{CS}} = 2.29 \) (original value from CS derivation).

### E. 2-RDM optimization

The \( \Delta \text{NO} \) 2-RDM depends upon the \( \Delta \text{NO} \) orbitals, \( \{\phi_p\} \), and the electron transfer variables, \( \{\Delta_{ia}\} \). The optimal \( \Delta \text{NO} \) 2-RDM is found by minimizing \( E^{\Delta \text{NO}} \) with respect to both \( \{\phi_p\} \) and \( \{\Delta_{ia}\} \). The optimization can be performed for both \( \{\phi_p\} \) and \( \{\Delta_{ia}\} \) simultaneously, using the trust-region method.

In order to obtain the optimal \( \Delta \text{NO} \) orbitals, \( \{\phi_p\} \), the initial set of orthonormalized orbitals, \( \{\phi_r\} \), are rotated amongst themselves via a unitary transformation,
\[
\phi_p = \sum_r U_{rp} \phi_r .
\]

To determine such a transformation, the unitary matrix is parameterized with the exponential of a skew matrix,
\[
Y = e^Y ,
\]

where
\[
Y_{pq} = y_{pq}, 
Y_{qp} = -y_{pq} .
\]

Any energy, \( E[y] \), that is a functional of a set of orbitals, may be expanded about the optimal orbitals, in a Taylor series, in terms of the skew matrix parameters, \( y \),
\[
E[y] \approx E[0] + y \cdot g^\phi + \frac{1}{2} y^T H^{\phi\phi} y ,
\]

and approximated to second-order. The gradient, \( g^\phi \), and hessian, \( H^{\phi\phi} \), consist of the first and second derivatives of the energy with respect to the skew-matrix parameters,
\[
g_{pq}^\phi = \frac{\partial E^{\Delta \text{NO}}}{\partial y_{pq}} \bigg|_{y=0} , \quad H_{pq,rs}^{\phi\phi} = \frac{\partial^2 E^{\Delta \text{NO}}}{\partial y_{pq} \partial y_{rs}} \bigg|_{y=0} .
\]

General expressions for the above derivatives may be found elsewhere, but for \( \Delta \text{NO} \) specifically see Appendix A. In the case of \( E^{\Delta \text{NO}} \), the energy is invariant (\( g_{pq}^\phi = 0 \)) with respect to rotations between orbital pairs where both belong to the inactive (frozen) closed-shell, open-shell, or inactive vacant orbital subspaces. Consequently, optimization must be performed with respect to all other possible rotations.

In the case of \( \Delta \text{NO} \), the energy is also a function of the electron transfer variables, \( \{\Delta_{ia}\} \). To simultaneously optimize the energy with respect to both \( \{\phi_p\} \) and \( \{\Delta_{ia}\} \), the orbital gradient and hessian are combined with those for \( \{\Delta_{ia}\} \). However, in order to avoid violation of N-representability of the 1-RDM, the \( \{\Delta_{ia}\} \) are first parameterized such that they remain between 0 and 1/2,
\[
\Delta_{ia} = \frac{\cos^2 \theta_{ia}}{2} .
\]

The total gradient and hessian are then given by
\[
g = \begin{bmatrix} g^\phi & g^\theta \end{bmatrix} , \quad H = \begin{bmatrix} H^{\phi\phi} & H^{\phi\theta} \\ H^{\theta\phi} & H^{\theta\theta} \end{bmatrix} ,
\]

where
\[
\Delta_{ia} = \frac{\cos^2 \theta_{ia}}{2} .
\]
where
\[ g_{pq}^\theta = \frac{\partial E^{\Delta\text{NO}}}{\partial \theta_{pq}}, \quad H_{pq,rs}^{\phi\theta} = \frac{\partial^2 E^{\Delta\text{NO}}}{\partial \theta_{pq} \partial \theta_{rs}}, \] (60)
and
\[ H_{pq,rs}^{\phi\theta} = \frac{\partial^2 E^{\Delta\text{NO}}}{\partial \theta_{pq} \partial \theta_{rs}}. \] (61)

Minimization can be performed using the trust-region Newton method, in which an augmented hessian is prepared and a step is taken within a dynamic trust-region radius. A standard trust-region approach is described in Nocedal and Wright (Algorithm 4.1). \(^90\) The procedure begins with the calculation of an augmented hessian, \( A \),
\[ A = H + \epsilon I \] (62)
where \( I \) is the identity matrix and \( \epsilon \) is chosen such that \( A \) is positive definite. Then a Newton step, \( p \), is calculated by solving
\[ Ap = -g. \] (63)

If the step is within the trust-region radius, \( p = \sqrt{p \cdot p} < t \), then the step is taken, otherwise a subproblem must be solved (Equation 4.5 in reference 90) in which a step within the trust-region radius is found iteratively. Each step, whether pure Newton-Raphson or a solution to the subproblem, is evaluated by comparing the actually change in the function (\( E \)) to the change predicted by a second-order Taylor expansion. If the reduction is less than a minimum threshold (\( e.g. \ 25\% \) of predicted) then no step is taken and \( t \) is reduced. If the reduction is more than the minimum, a step is taken, and if it is more than a maximum threshold (\( e.g. \ 75\% \) of predicted) then \( t \) is increased.

In order to avoid optimizations to saddle points in \{\( \phi_p \)\} and \{\( \Delta a \)\} space, a modification can be made before commencing the trust-region algorithm. The non-augmented hessian is diagonalized to evaluate the eigenvalues,
\[ Z^T H Z = H' \] (64)
where \( H' \) is a diagonal matrix of hessian eigenvalues and \( Z \) is the matrix that transforms to normal coordinates. For any normal coordinate (mode) with a negative eigenvalue, \( H'_{ii} < 0 \), and a gradient component with a magnitude less than a threshold, \( |g'_i| < g_0 \), the gradient is amplified to promote escape from the saddle point,
\[ g_{\text{mod}} = Z g'_{\text{mod}} \] (65)
where
\[ g'_{\text{mod},i} = \begin{cases} \text{sign}(g'_i)g_0, & \text{If } H'_{ii} < 0 \text{ and } |g'_i| < g_0 \\ g'_i, & \text{otherwise} \end{cases} \] (66)
The trust-region method then continues as usual.

When an optimization step is calculated, the unitary matrix for the orbital rotation is prepared by first approximating \( e^Y \) to second order,
\[ U \approx I + Y + \frac{1}{2} YY^\dagger, \] (67)
and then ensuring it is unity by applying Gram-Schmidt orthonormalization.

F. Conditions of \( N \)-representability

Due to the addition of the high-spin correction, the \( \Delta\text{NO} \) 2-RDM is not strictly \( N \)-representable. Therefore, it is useful to have some means of evaluating the extent to which \( N \)-representability might be violated. The necessary, but not sufficient, \( N \)-representability conditions that depend only on the \( PQG \) (or \( DQG \)) conditions. \(^91\)-\(^93\) They require that the following matrices, in a spin-orbital basis and defined in terms of the second quantization, be positive semidefinite (\( i.e. \) all non-negative eigenvalues)
\[ P_{ijkl} = \langle \Psi|a_j^{\dagger}a_i^{\dagger}a_k a_l|\Psi\rangle, \] (68a)
\[ Q_{ijkl} = \langle \Psi|a_j a_i a_k^{\dagger}a_l^{\dagger}|\Psi\rangle, \] (68b)
\[ G_{ijkl} = \langle \Psi|a_j^{\dagger}a_i^{\dagger}a_k a_l|\Psi\rangle, \] (68c)
where \( \Psi \) is the \( N \)-electron wave function, \( a_i^{\dagger} \) and \( a_i \) are creation and annihilation operators, respectively. Sufficient conditions for \( N \)-representability require the calculation of higher-order matrices and are not considered here. \(^93\)-\(^94\) For a spin-resolved 2-RDM over spatial orbitals \([?\,?] \) and \( (35)\), and a diagonal 1-RDM, the elements of \( P, Q, \) and \( G \) can be expressed as
\[ P_{pqrs}^{\sigma\sigma'} = \Gamma_{pqrs}^{\sigma\sigma'}, \] (69a)
\[ Q_{pqrs}^{\sigma\sigma'} = \Gamma_{pqrs}^{\sigma\sigma'} + \frac{\delta_{pq}\delta_{rs}}{2} (1 - n_p - n_q) - \frac{\delta_{\sigma\sigma'}\delta_{ps}\delta_{qr}}{2} (1 - n_p - n_q), \] (69b)
\[ G_{pqrs}^{\sigma\sigma'} = \frac{\delta_{\sigma\sigma'}\delta_{ps}n_q}{2} \Gamma_{pqrs}^{\sigma\sigma'}, \] (69c)
\[ G_{pqrs}^{\sigma\sigma',\sigma''} = \Gamma_{pqrs}^{\sigma\sigma'} (\sigma \neq \sigma'), \] (69d)
where \( \sigma = \alpha \) or \( \beta \), and \( \sigma \) can be equivalent to \( \sigma' \) (except for \( G_{pqrs}^{\sigma\sigma',\sigma''} \)). The expression of \( P, Q, \) and \( G \) in terms of spatial-orbital indices (subscript) requires specifying the spin-function (superscript) used to form the corresponding spin orbital. The matrices can be decomposed into spin-blocks for construction and analysis, which is
evident from their structure.\textsuperscript{95} 

\[
P = \begin{bmatrix}
P_{\alpha\alpha\alpha\alpha} & 0 & 0 & 0 \\
0 & P_{\beta\beta\beta\beta} & 0 & 0 \\
0 & 0 & P_{\alpha\beta\alpha\beta} & 0 \\
0 & 0 & 0 & P_{\beta\alpha\beta\alpha}
\end{bmatrix}, \quad (70a)
\]

\[
Q = \begin{bmatrix}
Q_{\alpha\alpha\alpha\alpha} & 0 & 0 & 0 \\
0 & Q_{\beta\beta\beta\beta} & 0 & 0 \\
0 & 0 & Q_{\alpha\beta\alpha\beta} & 0 \\
0 & 0 & 0 & Q_{\beta\alpha\beta\alpha}
\end{bmatrix}, \quad (70b)
\]

\[
G = \begin{bmatrix}
G_{\alpha\alpha\alpha\alpha} & G_{\alpha\beta\alpha\beta} & 0 & 0 \\
G_{\beta\beta\alpha\beta} & G_{\beta\beta\beta\beta} & 0 & 0 \\
0 & 0 & G_{\alpha\beta\alpha\beta} & 0 \\
0 & 0 & 0 & G_{\beta\alpha\beta\alpha}
\end{bmatrix}. \quad (70c)
\]

The sum of the negative eigenvalues of these matrices can be used to quantify the extent to which they are not positive semi-definite and subsequently violate the PQG-conditions.\textsuperscript{96} These sums are denoted by 

\[
X_{\text{neg}} = \sum_{x_i < 0} x_i \quad (71)
\]

where \( K \) is the size of the orbital basis, \( X = P, Q \) or \( G \) and \( x_i \) is an eigenvalue of \( X \).

III. METHOD

A. Reference energies and properties

The energies from single-reference methods: completely renormalized coupled-cluster (CR-CCL)\textsuperscript{97,98} and spin-unrestricted density functional approximation (B3LYP),\textsuperscript{99–101} and multireference energies: multireference self-consistent field (MRSCF)\textsuperscript{102} and multireference Möller-Plesset second-order perturbation theory (MRMP2),\textsuperscript{103–106} were calculated using GAMESS.\textsuperscript{107} The active space for the MRSCF and MRMP2 calculations was chosen to be all occupied orbitals of the reference configuration, and an equivalent number of the lowest lying vacant (virtual) orbitals. Full configuration interaction (FCI) energies, and 2-RDMs, were determined using Quantum Package 2.\textsuperscript{108} All calculations were performed using the cc-pVTZ basis set.\textsuperscript{109,110}

B. ∆NO energies and properties

The ∆NO, ∆NO-OF, and ∆NO-CS calculations were performed with MUNgauss.\textsuperscript{111} The active space for each electron pair consisted of the orbital they occupy in the reference configuration and one vacant (virtual) orbital. The ∆NO 2-RDMs were optimized using the the trust-region method outlined in Subsection II E. The iterative subproblem was solved using the dgqt routine from MINPACK-2,\textsuperscript{112} translated to Fortran 90. Restricted Hartree-Fock orbitals are used as the initial guess, and the initial trust-region radius was \( t = 0.05 \) with a maximum radius value of 2.00. The gradient threshold, required for saddle-point escape, was \( g_0 = 10^{-4} \). The convergence criteria was \( \bar{g} < 10^{-6} \) where \( \bar{g} = \frac{\|g\|}{n} \) is the normalized total gradient length, where \( n \) is the total number of orbital optimization parameters, \( \{y_{pq}\} \) and \( \{\Delta_{ia}\} \).

The OF and CS energies are calculated post-∆NO using an SG-1 grid\textsuperscript{113} and Becke atomic weights.\textsuperscript{114} The value of the double-counting correction parameter, \( c_{\text{DC}} \), was chosen to minimize the error in the ∆NO-OF, and ∆NO-CS, \( \text{H}_2 \) potential energy curves, relative to FCI. The values were 0.40 and 0.35 for ∆NO-OF and ∆NO-CS, respectively.

IV. RESULTS

A. Cluster geometries

The geometries of the H clusters studied are presented in Figure 1. The potential energy surfaces were calculated with established single-reference (FCI, CR-CCL and B3LYP) and multireference methods (MRSCF and MRMP2) as well as with ∆NO, ∆NO-OF and ∆NO-CS. The potential energy surface is defined as,

\[
U(X) = E(X) + V_{\text{nn}}(X) \quad (72)
\]

where \( E(X) \) is the electronic energy provided by the electronic structure method, \( V_{\text{nn}}(X) \), is the nuclear-repulsion energy, and both are a function of the nuclear coordinates, \( X \).

B. 2-RDM optimization

The most time consuming part of the ∆NO 2-RDM optimization is the transformation of the two-electron integrals from the atomic orbital basis to the ∆NO orbital basis, which scales as \( N_{\text{act}}K^4 \), where \( N_{\text{act}} \) is the number of active orbitals and \( K \) is the number of atomic basis
TABLE I. Comparison of the number of integral transformations required to optimize the ∆NO 2-RDM

| Cluster | R (Å) | ID | TR |
|---------|------|----|----|
| H₂      | 1.2  | 170 | 10 |
| H₃      | 1.2  | 924 | 13 |
| H₄ linear | 1.2 | 252 | 47 |
| H₄ square | 1.2 | 320 | 20 |

a ID = iterative diagonalization of pseudo-Fock matrix
b TR = trust-region method

functions. Recently, this cost has been reduced for similar calculations in natural orbital functional methods by introducing the resolution-of-the-identity (DoNOF-RI).\textsuperscript{115}

For the systems studied, the use of the trust-region method, on average, reduced the number of optimization steps (and therefore the number of integral transformations) by an order of magnitude compared to the previous iterative diagonalization algorithm\textsuperscript{12,34,116} (Table I). Furthermore, the introduction of an analytical hessian, and a saddle-point escape procedure, enabled the trust-region method to find global minima that the previous iterative-diagonalization algorithm could not locate.

In some cases, in order to determine whether optimized 2-RDMs were indeed global minima, multiple different initial guesses were used. This includes using orbitals different from RHF, including MRSCF orbitals and FCI natural orbitals, choosing different vacant orbitals for the active space, and also using 2-RDMs (both the orbitals and the electron-transfer variables) from other cluster geometries. The minima found by the 2-RDM optimization algorithm were invariant with respect to the choice of initial orbitals or which vacant orbitals to include in the active space. Local minima were only located when an alternative 2-RDM guess was supplied from another cluster geometry (e.g. 2-RDM from H₄ at θ = 90° for H₄ at θ = 85°).

C. H₂

The error in the H₂ potential energy curves relative to the FCI result is presented in Figure 2. To illustrate the importance of the double-counting correction, the ∆NO-OF and ∆NO-CS result without including the double-counting correction (c_DC = 0) is shown. Without the double-counting correction, both ∆NO-OF and ∆NO-CS predict ∼ 30 kJ mol⁻¹ too much correlation energy near the equilibrium bond length (R = 0.743 Å). However, as the amount a static correlation grows with increasing R, the error approaches zero. The static correlation included in the ∆NO 2-RDM causes the on-top density to vanish, and therefore there is no superfluous correlation energy at dissociation. If the double-counting correction is included, the error in ∆NO-OF and ∆NO-CS is dramatically reduced with a maximum error of ∼ 3 kJ mol⁻¹ (a positive deviation near equilibrium, and negative at stretched bond lengths). The error in the B3LYP potential energy curve is also included, where the most notable deviation is the energy at the dissociation limit. The dissociation limit predicted by B3LYP is 5.6 kJ mol⁻¹ higher than FCI.

The equilibrium bond lengths and dissociation energies predicted by each method for the various H clusters are presented in Tables II and III. In the case of H₂, ∆NO-OF and ∆NO-CS predict equilibrium bond lengths within 0.004 Å of the FCI value, 0.743 Å, while the other methods are slightly more accurate. Also for H₂, ∆NO and MRSCF are identical and therefore predict that same bond length (0.756 Å) and dissociation energy (399 kJ mol⁻¹).

Other than ∆NO and MRSCF, all methods predict fairly accurate bond dissociation energies for H₂. The high accuracy of ∆NO-OF (0 kJ mol⁻¹ error) and ∆NO-CS (1 kJ mol⁻¹ error) is due to the calibration of the double-counting correction coefficient, c_DC. The strategy is to use the bond of H₂ as a prototype for the static and dynamic correlation energy balance between paired electrons.
TABLE III. Dissociation energies (in kJ mol\(^{-1}\)) of H clusters

| cluster | method   | \(\Delta\text{NO}\) | \(\Delta\text{NO}-\text{OF}\) | \(\Delta\text{NO}-\text{CS}\) | B3LYP     | CR-CCL    | MRSCF     | MRMP2     | FCI       |
|---------|----------|----------------------|-----------------------------|-----------------------------|-----------|-----------|-----------|-----------|-----------|
|         | \(\text{H}_2\) | 399                  | 454                         | 455                         | 461       | 454       | 444       | 444       | 454       |
|         | \(\text{H}_3\) | 317                  | 411                         | 413                         | 444       | 412       | 400       | 400       | 412       |
|         | \(\text{H}_4\) linear | 588                  | 722                         | 718                         | 765       | 726       | 713       | 713       | 733       |
|         | \(\text{H}_4\) square | 168                  | 288                         | 288                         | 349       | 292       | 202       | 202       | 288       |

FIG. 3. Potential energy of a linear \(\text{H}_3\) cluster with neighbouring H atoms separated by a distance \(R\).

D. \(\text{H}_3\)

The potential energy curves (PECs) for \(\text{H}_3\), in a linear arrangement, are presented in Figure 3. The \(\Delta\text{NO}\) and MRSCF PECs are above the PECs of the other methods designed to also capture dynamic correlation. Unlike \(\text{H}_2\), the \(\Delta\text{NO}\) and MRSCF PECs differ and it is apparent that, even with the high-spin correction applied to the unpaired electron (Equation 25), \(\Delta\text{NO}\) is missing correlation at moderately stretched bond lengths. Nevertheless, the correct dissociation limit is reached by \(\Delta\text{NO}\), \(\Delta\text{NO}-\text{OF}\) and \(\Delta\text{NO}-\text{CS}\).

The error in the PECs, for the methods that include both static and dynamic correlation (\(\Delta\text{NO}-\text{OF}\), \(\Delta\text{NO}-\text{CS}\), MRMP2 and B3LYP), compared to the FCI PEC is shown in Figure 4. At small \(R\), \(\Delta\text{NO}-\text{OF}\) and \(\Delta\text{NO}-\text{CS}\) overestimate the correlation energy by a maximum of 4 and 11 kJ mol\(^{-1}\), respectively. At moderately stretched bond lengths, both \(\Delta\text{NO}-\text{OF}\) and \(\Delta\text{NO}-\text{CS}\) underestimate the correlation energy (23 and 29 kJ mol\(^{-1}\), respectively), which is due to the missing static correlation in \(\Delta\text{NO}\) between the electron pair and the unpaired electron.

The largest deviation of the MRMP2 PEC from the FCI PEC is at small \(R\) (16 kJ mol\(^{-1}\)), and the error decreases monotonically as \(R\) increases. In the case of B3LYP, the correlation energy is significantly overestimated around equilibrium \(R\) (24 kJ mol\(^{-1}\)) and like \(\text{H}_2\), the energy at the dissociation limit is incorrect. These errors are reflected in the predicted equilibrium \(R\) values and dissociation energies (Tables II and III).

Similar to \(\text{H}_2\), both \(\Delta\text{NO}\) and MRSCF overestimate the equilibrium separation. Whereas, \(\Delta\text{NO}-\text{OF}\) and \(\Delta\text{NO}-\text{CS}\) slightly underestimate \(R_e\), by 0.006 Å and 0.012 Å, respectively. The values of \(R_e\) from both MRMP2 (0.934 Å) and B3LYP (0.931 Å) differ only slightly from the FCI value. Also, like \(\text{H}_2\), the CR-CCL PEC is equivalent to FCI.

For those methods designed to capture all the correlation energy, it is only the dissociation energy predicted by B3LYP (444 kJ mol\(^{-1}\)) that differs substantially from the FCI value (412 kJ mol\(^{-1}\)).

E. Linear \(\text{H}_4\)

The PECs for the linear \(\text{H}_4\) cluster, with interatomic spacing \(R\), are presented in Figure 5. Similar to \(\text{H}_3\), the \(\Delta\text{NO}\) and MRSCF PECs appear above the others (as expected), and the B3LYP PEC lies below the others at equilibrium and above at dissociation. Interestingly, while the CR-CCL PEC is essentially exact close to \(R_e\), and eventually approaches the correct dissociation limit, there is a significant negative deviation from the FCI PEC at large \(R\).

The error in the linear \(\text{H}_4\) PECs, compared to FCI, are explicitly shown in Figure 6. The behaviour of the error in the \(\Delta\text{NO}-\text{OF}\) and \(\Delta\text{NO}-\text{CS}\) PECs for linear \(\text{H}_4\) is the same as that seen for \(\text{H}_2\), although substantially larger. Both \(\Delta\text{NO}-\text{OF}\) and \(\Delta\text{NO}-\text{CS}\) underestimate the correlation energy at small \(R\), and overestimate at large \(R\), while eventually reaching the correct dissociation limit. Like \(\text{H}_3\), the error in the MRMP2 PEC is largest at small \(R\) and decreases as \(R\) is increased. In the case of CR-CCL, the PEC is essentially exact for small \(R\) and then decreases.
begins overestimating the correlation energy at $\sim 1.4 \, \text{Å}$. The error in the B3LYP PEC is similar that seen for both $\text{H}_2$ and $\text{H}_3$.

With the exception of $\Delta \text{NO}$ and MRSCF, all methods provide good estimates of $R_c$ compared to FCI, with $\Delta \text{NO}$ and $\Delta \text{NO-CS}$ showing the most deviation (0.005 and 0.008 Å, respectively). Both $\Delta \text{NO}$ and MRSCF predict larger $R_c$ values, and are within close agreement with each other (differing by 0.002 Å).

The dissociation energies predicted by each method are more varied. The dissociation energies predicted by $\Delta \text{NO}$ and $\Delta \text{NO-CS}$, 722 kJ mol$^{-1}$ and 718 kJ mol$^{-1}$, respectively, are in decent agreement with the FCI value, 733 kJ mol$^{-1}$. Only CR-CCL provides a better estimate, 726 kJ mol$^{-1}$.

F. Square $\text{H}_4$

The PECs for the dissociation of $\text{H}_4$ in a square arrangement ($\theta = 90^\circ$, see Figure 1) are shown in Figure 7. The square arrangement of H atoms results in much more varied performance amongst the methods assessed compared to the linear arrangement. The B3LYP PEC is far below the others near $R_c$, but above again at the dissociation limit. There is substantial separation between the $\Delta \text{NO}$ and MRSCF PECs near $R_c$. Furthermore, this system is a known issue for single-reference coupled-cluster methods, and while CR-CCL performs better here than CCSD(T), the error is significant. The explicit errors are shown in Figure 8.

The error in the $\Delta \text{NO-CS}$ PEC is relatively small, and provides, arguably, the best agreement with the FCI PEC. The usual underestimation of the correlation energy at small $R$ by MRMP2 is relatively small for this system, and therefore MRMP2 also provides a good estimate of the FCI PES. The negative deviation of $\Delta \text{NO}$ at
moderate $R$ values reaches a maximum of $\sim 14$ kJ mol$^{-1}$ before going to the correct dissociation limit. Whereas, the CR-CCL PES overestimates the correlation energy by more than 30 kJ mol$^{-1}$ before approaching the dissociation limit.

Despite the errors in the PECs, accurate predictions of the FCI $R_e$, 0.864 Å, are provided by ∆NO-OF, ∆NO-CS, MRMP2 and CR-CCL, with values of 0.863, 0.857, 0.868 and 0.878, respectively. As usual, the dissociation energies vary more widely. Both ∆NO-OF and ∆NO-CS provide very accurate estimates of the dissociation energy, and MRMP2 and CR-CCL differ by only 10 and 4 kJ mol$^{-1}$, respectively. While B3LYP overestimates the dissociation energy by 61 kJ mol$^{-1}$.

### G. Rectangle to square H$_4$

At fixed $R$, if $\theta$ of the, $D_{4h}$ symmetry, square arrangement of H$_4$ (Figure 1) is changed from 90°, the result is a rectangular, $D_{2h}$ symmetry, arrangement. The PECs for $\theta = 70^\circ$ to $\theta = 110^\circ$ ($D_{2h} \rightarrow D_{4h} \rightarrow D_{2h}$) for fixed $R$ values of 0.8, 1.2 and 1.7 Å are presented in Figure 9.

Like the PECs for the other H clusters, the ∆NO and MRSCF PECs lie above those of the methods designed to capture all of the correlation energy. As $R$ increases, and static correlation increases while dynamic correlation decreases, the ∆NO and MRSCF PECs approach the others (note the scale of the $y$-axis in Figure 9).

In the case of CR-CCL, coupled-cluster methods are known to have a concave (downward pointing) derivative discontinuity, 54,58,67 or cusp, at $\theta = 90$, which becomes more prominent as $R$ increases. This is in contrast to the convex (upward pointing) cusp seen for RHF and spin-restricted density functional methods. The cusp in the CR-CCL PEC is barely noticeable at $R = 0.8$ Å, but is evident at the larger $R$ values. The MRMP2 PEC closely resembles the FCI PEC at all $R$ values, and the absolute error decreases with increasing $R$. The height of the PEC is significantly underestimated by B3LYP at all $R$ values.

The ∆NO-OF and ∆NO-CS methods accurately model the PEC near $\theta = 70^\circ$ and 90°, however there is an obvious cusp at $\theta \sim 85^\circ$ and $\theta \sim 95^\circ$. This is due to an inadequate description of the 2-RDM upon transitioning from a $D_{4h}$ to $D_{2h}$ arrangement, which is analysed, and elaborated upon, later.

The barrier height, defined as the difference in energy at $\theta = 90^\circ$ and $70^\circ$, predicted by each method at each $R$ value is presented in Table IV. At $R = 0.8$ Å, ∆NO-OF and ∆NO-CS underestimate the barrier by 6 and 9 kJ mol$^{-1}$, respectively. Compare that to ∆NO, which overestimates the barrier by 12 kJ mol$^{-1}$. MRMP2 provides accuracy similar to ∆NO-OF and ∆NO-CS, while CR-CCL only differs from the FCI barrier by 1 kJ mol$^{-1}$. As $R$ is increased, the agreement of ∆NO with FCI improves, which highlights the importance of static correlation to these PECs. At $R = 1.2$, ∆NO-OF and ∆NO-CS both differ from FCI by 8 kJ mol$^{-1}$, and differ by 5 and 3 kJ mol$^{-1}$, respectively, at $R = 1.7$ Å. At the same time, the agreement of MRMP2 with FCI also improves, to within 4 kJ mol$^{-1}$ at $R = 1.2$ Å and within 1 kJ mol$^{-1}$ at $R = 1.7$ Å. In contrast, due to the cusp, the CR-CCL values worsen.

Although the barrier heights predicted by ∆NO-OF and ∆NO-CS are in good agreement with FCI values, the deviation of the PECs from FCI approximately 5° from $\theta = 90^\circ$ is concerning. An initial analysis of the correlation captured by ∆NO is provided by the electron transfer variables, $\Delta_{\alpha}$. The values of $\Delta_{14}$ and $\Delta_{23}$ from $\theta = 70^\circ$ to $\theta = 110^\circ$, at $R = 1.7$ Å, are presented in Figure 10. From $\theta = 70^\circ$ to $\theta = 77^\circ$, $\Delta_{14}$ and $\Delta_{23}$ are equivalent, indicating that the motion of both electron pairs is correlated in the same manner. This points to the rectangular cluster behaving as two H$_2$ molecules, which is not surprising at particularly large separations like $R = 1.7$ Å. Beyond $\theta = 77^\circ$, the values of $\Delta_{14}$ and $\Delta_{23}$ diverge, and there is an obvious transition that occurs from $\theta = 84^\circ$ to $\theta = 87^\circ$, and from $\theta = 93^\circ$ to $\theta = 96^\circ$ (grey rectangles). At the $\theta = 90^\circ$ side of the transition, the electron pairs have significantly different amounts of static correlation. The $\Delta_{23}$ electron pair is very near the strong correlation limit, $\Delta \rightarrow 1/2$, while $\Delta_{14} \approx 0.28$. Essentially, there are two regimes of correlation. For $\theta$ near 70° (and 110°) there is the 2H$_2$ regime, and for $\theta$ near 90° there is the H$_4$ regime.

The presence of different two regimes, or models, for the correlation in H$_4$ by ∆NO, and their transition, is supported further by an analysis of the ∆NO orbitals (Figure 11). From Figure 11, it is seen that from $\theta = 84^\circ$ to 85° the ∆NO orbitals resemble those of two separate H$_2$ molecules, albeit with small tails on the opposite pair of protons. At $\theta = 86^\circ$, the orbitals begin to extend significantly to the neighbouring pair of protons, and by $\theta = 87^\circ$ they are almost evenly distributed. This transition occurs quickly with respect to the PEC, in just 3°.

| method      | $R$ (Å) | 0.8  | 1.2  | 1.7  |
|-------------|---------|--     |     |     |
| ∆NO        |         | 305  | 212  | 76   |
| ∆NO-OF     |         | 287  | 200  | 76   |
| ∆NO-CS     |         | 284  | 200  | 78   |
| B3LYP      |         | 238  | 157  | 50   |
| CR-CCL     |         | 292  | 186  | 51   |
| MRSCF      |         | 281  | 190  | 72   |
| MRMP2      |         | 286  | 204  | 80   |
| FCI        |         | 293  | 208  | 81   |

*a Barrier height is defined as $E[90^\circ] - E[70^\circ]$*

### Table IV. Barrier height$^*$ (in kJ mol$^{-1}$) for $D_{2h} \rightarrow D_{4h} \rightarrow D_{2h}$ transition of H$_4$ cluster at varying $R$. 

Analysis of the ∆ values and the orbitals of ∆NO partially explains the behaviour of the model for the this challenging H$_4$ PEC, but it is does not explain why it differs from the FCI result. A comparison of the ∆NO and
FCI two-electron densities, $\Gamma^{\Delta NO}(r_1, r_2)$ and $\Gamma(r_1, r_2)$, can show conclusively how the $\Delta NO$ model of electron correlation differs from how the electrons actually behave. The $\Delta NO$ and FCI opposite-spin ($\alpha\beta$) and parallel-spin ($\alpha\alpha$) two-electron densities for $H_4$ at $\theta = 70^\circ$, $85^\circ$ and $90^\circ$, with $R = 1.7$ Å, are shown in Figure 12. The two-electron densities are plotted for a test-electron placed at the upper right $H$ atom (indicated by an orange dot). Therefore, the plots show the probability of finding the other electron(s), opposite or parallel-spin, within the $H_4$ cluster.

At $\theta = 70^\circ$ (left column of Figure 12), the $\Delta NO$ and FCI two-electron densities closely resemble each other. For opposite-spin electrons (top half of Figure 12), if an $\alpha$ electron is found at the upper-right $H$ atom, then it is most likely to find a $\beta$ electron on the closest $H$ atom, to which it is paired. It would then be equally likely to find the other $\beta$ electron on either of the two more distant $H$ atoms. In the case of parallel-spin electrons (bottom half of Figure 12), if an $\alpha$ electron is found at the upper-right $H$ atom then the other $\alpha$ electron must be found in the “other” $H_2$ molecule, with an equal probability of being found at either proton.

At $\theta = 90^\circ$ (right column of Figure 12), the behaviour is much different. For electrons of opposite-spin, when an $\alpha$ electron is found at the upper-right $H$ atom, it is equally likely to find the $\beta$ electron at either of the nearest protons, with a reduced probability of finding a $\beta$ electron at the furthest proton. Comparing the $\Delta NO$ and FCI two-electron densities, the same qualitative behaviour is seen but the $\Delta NO$ 2-RDM results in too high a probability of finding the $\beta$ electron on the furthest proton compared to the probability at the nearest protons. In the case of parallel-spin electrons, if an $\alpha$ electron is found at the upper-right $H$ atom then it is most likely that another $\alpha$ electron is on the furthest proton, with only a slight chance of an $\alpha$ electron on the nearest protons. Again, $\Delta NO$ models this behaviour qualitatively, but compared to FCI, $\Delta NO$ predicts the probability of finding an $\alpha$ electron on the nearest protons to be too high relative to the furthest proton. This suggests that the amount of exchange, or Fermi correlation, which occurs between parallel-spin electrons, in the $\Delta NO$ 2-RDM is slightly inadequate at $\theta = 90^\circ$.

Finally, at $\theta = 85^\circ$ (middle column of Figure 12), the

---

**FIG. 9.** Potential energy of $H_4$ cluster as a function of $\theta$ and $R$ (see Figure 1).

**FIG. 10.** $\Delta$ values for rectangle-to-square $H_4$ cluster as a function of $\theta$ at $R = 1.7$ Å.

**FIG. 11.** Change in $\Delta NO$ active orbitals of $H_4$ with $\theta$. 
FIG. 12. FCI and \( \Delta \text{NO} \) two-electron densities of \( \text{H}_4 \) cluster when a test electron is placed at a H nucleus, \( r_0 \) (orange dot). Densities are shown for \( \text{H}_4 \) at \( R = 1.7 \) Å and \( \theta = 70^\circ \) (left), \( 85^\circ \) (middle), and \( 90^\circ \) (right).
difference between ∆NO and FCI two-electron densities is the most pronounced. It is evident that the electrons of the ∆NO 2-RDM at θ = 85° behave more like the electrons at θ = 70° (2 H₂ regime) than the actual behaviour given by the FCI two-electron density. In the case of the ∆NO two-electron densities, both opposite and parallel-spin, the probabilities of finding electrons on the left-side protons (other H₂) are near equivalent. Whereas, the FCI two-electron density shows uneven probabilities for the left-side protons, which more resembles the behaviour at θ = 90° (H₄ regime).

The inability of ∆NO to accurately model the PEC from θ = 70° to 90°, is due to the energetic preference of ∆NO to model H₄ as 2 H₂ at θ values close to 90°. The error observed in the parallel-spin two-electron density suggests that ∆NO is incorrectly describing the correlation between different electron pairs (interpair correlation), which is similar to the conclusion reached in a study with PNOF6.⁵⁴ For ∆NO, this could mean an error in the exchange component (Fermi correlation), or possibly something else, but it implies that the high-spin correction (HSC) is not adequate. The HSC provides the correct description in the strong correlation limit (both pairs), but it has deficiencies for moderate amounts of static correlation and must be improved upon.

The inadequacy of the HSC is highlighted further by examining the N-representability of the ∆NO 2-RDM; in particular, the violation of the PQG-conditions (Subsection II F). The sums of the negative eigenvalues of the P, Q, and G matrices corresponding to the ∆NO 2-RDM for rectangle-to-square H₄ are shown in Figure 13. The sums are separated into contributions from different spin-blocks of the P, Q, and G matrices. The P and Q matrices are separated into spin-blocks for each specific spin-pair, whereas the G matrix is separated into only three blocks, Gαβαβ, Gαβαβ, and Gσσσσ, due to the coupling of the parallel-spin blocks [Equation (70c)]. For all values of θ, the parallel-spin blocks of the P and Q matrices have no negative eigenvalues. This is completely expected considering that all parallel-spin terms of the ∆NO 2-RDM are derived from the ∆NO wave function. The only terms not originating from the ∆NO wave function are those corresponding to the HSC, which are applied to the opposite-spin block of the 2-RDM only [Equation (35)]. The off-diagonal, Gαβαβ and Gβααα, spin-blocks of Gσσσσ contain 2-RDM elements corresponding to the HSC and they are responsible for the contribution to Gneg. The most severe violation of the PQG-condition is seen for Gαβαβ, also known as the particle-hole density matrix. The sum of the negative eigenvalues of Gαβαβ reach a minimum at θ = 90°, where Gαβαβ = -0.229. Due to the fact that the HSC is applied to the exchange-like terms (i.e. Γpqq), of the ∆NO 2-RDM, the sum of the negative eigenvalues of the Pαβαβ and Qαβαβ matrices are equivalent for all values of θ. In the case of Fneg and Qneg, the minimum value observed is -0.0818, which occurs at θ = 86° and θ = 94°.

As seen for the energy and orbital occupancies, the violation of the PQG-conditions exhibits different behaviour, with respect to θ, in the neighbourhood of θ = 90° (between the grey rectangles) compared to θ closer to 70° or 110° (outside the grey rectangles).

V. CONCLUSIONS

The ∆NO method is one of many that combines different models for different aspects of electron correlation in an effort to efficiently capture the total. The ∆NO-ODF approach is unique in the manner that it combines an approximate 2-RDM for static correlation with an ODF for dynamic correlation. With the exception of the high-spin correction, the ∆NO 2-RDM and energy can be derived from a multideterminantal wave function with restricted 2n-tuple excitations. Combining the ∆NO 2-RDM for static correlation with an ODF for dynamic correlation requires the inclusion of a double-counting correction, which can be done by replacing the short-range correlation counted by the ∆NO 2-RDM.

Previously, ∆NO, and other NOF methods, optimized the 2-RDM through alternating optimization of the occupation numbers and the orbitals until convergence was reached. The orbitals themselves were optimized by iterative diagonalization of a pseudo-Fock matrix. The trust-region Newton method introduced here, which allows for simultaneous optimization of the orbitals and occupancies, drastically reduces the number of optimization steps required and has the ability to identify and avoid saddle points.

Even simple hydrogen clusters can prove challenging for electronic structure methods, particularly single-reference methods. Overall, the ∆NO-ODF approach performs well when describing the potential energy surfaces of H₂, H₃, and H₄, relative to the more accurate, but costly, multireference wave function method, MRMP2.

Analysis of the two-electron density of H₄ reveals that
ΔNO incorrectly describes interorbital correlation. Despite being accurate in the strong-correlation limit, the ad hoc HSC does have deficiencies, particularly near the $D_{2h} \rightarrow D_{1h}$ transition. A more complete description of interorbital correlation within ΔNO is a target of future development.

**ACKNOWLEDGMENTS**

JWH thanks the Natural Sciences and Engineering Research Council of Canada (NSERC) for a Discovery Grant, Compute/Calcul Canada for computing resources and the Discovery Institute for Computation and Synthesis for useful consultations.

**DATA AVAILABILITY**

The data that support the findings of this study are available from the corresponding author upon request.

**Appendix A: ΔNO energy derivatives**

The optimization of the ΔNO 2-RDM is achieved by simultaneously finding the optimal ΔNO orbitals, $\{\phi_p\}$, and electron transfer variables, $\{\Delta_p\}$. The optimization is performed using a trust-region Newton algorithm, which requires the analytical calculation of the energy gradient, $g$, and hessian, $H$, with respect to the skew-matrix parameters, $y$, and the variables used to parameterize the electron-transfer variables, $\theta$.

The gradient, with respect to $y$, is evaluated at $y = 0$. The first and second derivatives of the individual orbitals, 

\[ \frac{\partial \phi_p}{\partial y_p} \bigg|_{y=0} = \delta_{aq} \phi_p - \delta_{ap} \phi_q \]  

(A1)

and

\[ \frac{\partial^2 \phi_p}{\partial y_p \partial y_{rs}} \bigg|_{y=0} = (\delta_{qr} \delta_{as} - \delta_{qs} \delta_{ar}) \phi_p - (\delta_{pr} \delta_{as} - \delta_{ps} \delta_{ar}) \phi_q + (\delta_{ps} \delta_{aq} - \delta_{qs} \delta_{ap}) \phi_r - (\delta_{pr} \delta_{aq} - \delta_{qr} \delta_{ap}) \phi_s, \]  

(A2)

can be used to derive expressions for derivatives of the total energy, $E^{\Delta\text{NO}}$. The total gradient with respect to the orbital-mixing parameters, is given by

\[ g^1_{pq} = 2 (\lambda_{qp} - \lambda_{pq}), \]  

(A3)

where

\[ \lambda_{pq} = (2 - W_p) n_p (h_{pq} + G^c_{pq}) + 2 (1 - W_p) n_p C^{oc}_{pq} + W_p n_p C^{op}_{pq} + 2 \sum_r n_{pr} (2 (pr | qr) - (pr | rq)) + 2 \sum_r (\xi_{pr} - \xi_{rp}) (pq | rr) - 2 \sum_r \kappa_{pr} (pr | rq) \]  

(A4)

and the two-electron integrals are defined as follows

\[ \langle pq | rs \rangle = \int \phi^*_p (r_1) \phi^*_q (r_2) \phi_r (r_1) \phi_s (r_2) \frac{1}{r_{12}} \, dr_1 \, dr_2, \]  

(A5)

and

\[ G^{cl}_{pq} = \sum_{\mu \nu} P^{cl}_{\mu \nu} \left( \langle pq | \mu \nu \rangle - \frac{1}{2} \langle \mu \mu | \nu \nu \rangle \right), \]  

(A6)

\[ G^{op}_{pq} = \sum_{\mu \nu} P^{op}_{\mu \nu} \left( \langle pq | \mu \nu \rangle - \langle \mu \mu | \nu \nu \rangle \right), \]  

(A7)

\[ G^{oc}_{pq} = \sum_{\mu \nu} P^{op}_{\mu \nu} \left( \langle \mu \mu | \nu \nu \rangle - \frac{1}{2} \langle \mu \mu | \nu \nu \rangle \right), \]  

(A8)

where $\mu$ and $\nu$ are indices of atomic basis functions. The density matrices are defined as

\[ P^{cl}_{\mu \nu} = \sum_p 2 n_p C_{\mu \nu} C^*_{\nu \mu}, \]  

(A9)

and

\[ P^{op}_{\mu \nu} = \sum_p n_p C_{\mu \nu} C^*_{\nu \mu}, \]  

(A10)

where $C_{\mu \nu}$ is a ΔNO orbital coefficient.

For the second derivatives with respect to the orbital-mixing parameters, it is useful to decompose $E^{\Delta\text{NO}}$. The one-electron energy is given by

\[ E^{\Delta\text{NO}}_{1e} = \sum_p 2 n_p h_{pp} + \sum_p n_p h_{pp}, \]  

(A11)

and the contribution of $E^{\Delta\text{NO}}_{1e}$ to the orbital hessian is given by

\[ H^1_{pq,rs} = \delta_{pr} \left( n^\alpha_p + n^\beta_p + n^\alpha_r + n^\beta_r - n^\alpha_q - n^\beta_q - n^\alpha_s - n^\beta_s \right) h_{qs} + \delta_{qs} \left( n^\alpha_q + n^\beta_q + n^\alpha_s + n^\beta_s - n^\alpha_p - n^\beta_p - n^\alpha_r - n^\beta_r \right) h_{pr} + \delta_{ps} \left( n^\alpha_p + n^\beta_p + n^\alpha_s + n^\beta_s - n^\alpha_q - n^\beta_q - n^\alpha_r - n^\beta_r \right) h_{qr} + \delta_{qr} \left( n^\alpha_r + n^\beta_r + n^\alpha_s + n^\beta_s - n^\alpha_q - n^\beta_q - n^\alpha_p - n^\beta_p \right) h_{ps}, \]  

(A12)
The two-electron energy associated with the 0-1RDM component of the energy is given by

\[ E_{\text{0-1RDM-2e}}^{\Delta \text{NO}} = E_{\text{0-1RDM}}^{\Delta \text{NO}} - E_{1e}^{\Delta \text{NO}}, \quad (A13) \]

and the corresponding contribution to the orbital hessian is

\[ H_{pq,rs}^{\phi \phi, 0-1\text{RDM-2e}} = \]

\[ \delta_{pq} \left( \begin{bmatrix} n_p^a + n_p^a + n_r^a + n_r^a - n_q^a - n_q^a - n_s^a \end{bmatrix} C_{qs}^{cl} + \left[ n_p^a + n_r^a - n_q^a - n_s^a \right] C_{qs}^{op} + \left[ n_p^b + n_r^b - n_q^b - n_s^b \right] \left[ 2G_{qs}^{op} - C_{qs}^{op} \right] \right) + \delta_{qs} \left( \begin{bmatrix} n_q^a + n_q^a + n_r^a + n_r^a - n_p^a - n_p^a - n_s^a \end{bmatrix} C_{pr}^{cl} + \left[ n_q^a + n_s^a - n_p^a - n_r^a \right] C_{pr}^{op} + \left[ n_q^b + n_r^b - n_p^b - n_s^b \right] \left[ 2G_{pr}^{op} - C_{pr}^{op} \right] \right) + \delta_{ps} \left( \begin{bmatrix} n_p^a + n_r^a + n_r^a + n_r^a - n_p^a - n_p^a - n_s^a \end{bmatrix} C_{qr}^{cl} + \left[ n_p^a + n_s^a - n_p^a - n_r^a \right] C_{qr}^{op} + \left[ n_p^b + n_r^b - n_p^b - n_s^b \right] \left[ 2G_{qr}^{op} - C_{qr}^{op} \right] \right) \]

and can be found elsewhere. In the case of \( \Delta \text{NO} \), for which the 2-RDM is a JKL-functional, the above expression can be simplified due to the fact that each term of the 2-RDM is described by only two indices, rather than four. The simplification can be applied specifically to terms involving \( J, K \) or \( L \) integrals. In the case of Coulomb integrals, \( J \), the 2-RDM simplifies as \( \Gamma_{abcd} = \delta_{ac} \delta_{bd} \gamma_{ab} \), giving

\[ E_{XJ}^{\Delta \text{NO}} = \sum_{ab} \gamma_{ab} J_{ab}, \quad (A18) \]

and leading to the following orbital hessian contribution,

\[ H_{pq,rs}^{\phi \phi, XJ} = 4 \left( \gamma_{pr} + \gamma_{qs} - \gamma_{ps} - \gamma_{qr} \right) \left( \langle pr|qs \rangle + \langle ps|rq \rangle \right) + \sum_a \left[ \delta_{pa} \left( \gamma_{qa} + \gamma_{ra} - \gamma_{pa} - \gamma_{sa} \right) \langle qa|ra \rangle + \delta_{pr} \left( \gamma_{pa} + \gamma_{ra} - \gamma_{qa} - \gamma_{sa} \right) \langle qa|sa \rangle + \delta_{qr} \left( \gamma_{pa} + \gamma_{sa} - \gamma_{qa} - \gamma_{ra} \right) \langle pa|a \rangle + \delta_{qs} \left( \gamma_{qa} + \gamma_{sa} - \gamma_{qa} - \gamma_{ra} \right) \langle pa|ra \rangle \right) \quad (A19) \]

In the case of exchange integrals, \( K \), the 2-RDM simplifies as \( \Gamma_{abcd} = \delta_{ac} \delta_{bd} \gamma_{ab} \), giving

\[ E_{XK}^{\Delta \text{NO}} = \sum_{ab} \gamma_{ab} K_{ab} \quad (A20) \]

and leading to the following orbital hessian contribution,

\[ H_{pq,rs}^{\phi \phi, XK} = 4 \left( \gamma_{pr} + \gamma_{qs} - \gamma_{ps} - \gamma_{qr} \right) \left( \langle pr|sq \rangle + \langle ps|rq \rangle \right) + \sum_a \left[ \delta_{pa} \left( \gamma_{qa} + \gamma_{ra} - \gamma_{pa} - \gamma_{sa} \right) \langle qa|ar \rangle + \delta_{pr} \left( \gamma_{pa} + \gamma_{ra} - \gamma_{qa} - \gamma_{sa} \right) \langle qa|as \rangle + \delta_{qr} \left( \gamma_{pa} + \gamma_{sa} - \gamma_{qa} - \gamma_{ra} \right) \langle pa|a \rangle + \delta_{qs} \left( \gamma_{qa} + \gamma_{sa} - \gamma_{qa} - \gamma_{ra} \right) \langle pa|ar \rangle \right) \quad (A21) \]
Similarly, for time-inversion exchange integrals, \( L \), the 2-RDM simplifies as \( \Gamma_{abcd} = \delta_{ab}\delta_{cd} \gamma_{ac} \) and the orbital hessian contribution simplifies in the same manner.

Simultaneous optimization of the electron-transfer variables, \( \{ \Delta_{pq} \} \), and consequently the orbital occupancies, requires including the gradient and hessian contributions from \( \theta \), where \( \theta_{pq} \) is related to \( \Delta_{pq} \) via Equation (58). The gradient contributions are given by,

\[
g_{pq}^{\theta} = \frac{\sin(2\theta_{pq})}{2} \frac{\partial E^{\Delta NO}}{\partial \Delta_{pq}}
\]  
(A22)
The derivative of \( E^{\Delta NO}_{0-1}\text{RDM} \) with respect to \( \Delta_{pq} \) is derived using the following identity,

\[
\frac{\partial n_t}{\partial \Delta_{pq}} = \delta_{qt} - \delta_{pt},
\]  
(A23)
which leads to

\[
\frac{\partial E^{\Delta NO}_{0-1}\text{RDM}}{\partial \Delta_{pq}} = 2 \left( h_{qp} - h_{pp} + G_{qq}^{cl} - G_{pp}^{cl} + G_{qq}^{oc} - G_{pp}^{oc} \right).
\]  
(A24)
The contribution to the gradient from the remaining components of the energy can be calculated by replacing the coefficients, \( \eta_{tu}, \zeta_{tu}, \xi_{tu} \) and \( \kappa_{tu} \) [Equations (15), (20) and (25)] with their corresponding derivatives. The derivatives of the coefficients are given by

\[
\frac{\partial \eta_{tu}}{\partial \Delta_{pq}} = \delta_{tu} \left[ (1 - 2n_q) - \delta_{pt} (1 - 2n_p) \right]
+ (1 - \delta_{tu}) \left[ \Omega_{t}V_{tt} \left( \delta_{pt} \delta_{qu} (n_u - n_t - \Delta_{tu}) + \Delta_{tu} (\delta_{qu} + \delta_{pt} - \delta_{pt} \delta_{qu}) \right) \right]
+ V_{t}O_{u} \left( \delta_{pt} \delta_{qu} (n_u - n_t - \Delta_{ut}) + \Delta_{ut} (\delta_{qt} + \delta_{pu} - \delta_{pt} \delta_{qt}) \right)
+ V_{t}V_{u} (\delta_{qt} \Delta_{pu} + \delta_{qu} \Delta_{pt}),
\]  
(A25)

\[
\frac{\partial \xi_{tu}}{\partial \Delta_{pq}} = \frac{V_{t}V_{u}}{2} \left( \delta_{qt} \sqrt{\frac{\Delta_{pu}}{\Delta_{pt}}} + \delta_{pu} \sqrt{\frac{\Delta_{pt}}{\Delta_{pu}}} \right),
\]  
(A26)

\[
\frac{\partial \xi_{tu}}{\partial \Delta_{pq}} = \frac{O_{t}V_{u}}{2} \left( \delta_{pt} \delta_{qu} \sqrt{\frac{n_t}{\Delta_{tu}}} - \delta_{pt} \sqrt{\frac{\Delta_{tu}}{n_t}} \right)
+ \frac{O_{u}V_{t}}{2} \left( \delta_{pu} \delta_{qt} \sqrt{\frac{n_u}{\Delta_{ut}}} - \delta_{pu} \sqrt{\frac{\Delta_{ut}}{n_u}} \right),
\]  
(A27)

\[
\frac{\partial \kappa_{tu}}{\partial \Delta_{pq}} = \frac{\partial \kappa_{tu}}{\partial \Delta_{pq}} = \frac{O_{t}V_{u}}{4} \left( \frac{-\delta_{pt} \delta_{qu} \sqrt{\frac{n_t}{\Delta_{tu}}} - \delta_{pt} \delta_{qu} \sqrt{\frac{\Delta_{tu}}{n_t}}}{n_t \Delta_{tu} \sqrt{\Delta_{tu}}} \right)
+ \frac{O_{u}V_{t}}{4} \left( \frac{-\delta_{pu} \delta_{qt} \sqrt{\frac{n_u}{\Delta_{ut}}} - \delta_{pu} \delta_{qt} \sqrt{\frac{\Delta_{ut}}{n_u}}}{n_u \Delta_{ut} \sqrt{\Delta_{ut}}} \right).
\]  
(A28)

Using the chain rule, the \( \theta \) contribution to the hessian can be expressed in terms of derivatives with respect to the \( \{ \Delta_{pq} \} \),

\[
H_{pq,rs}^{\theta \theta} = \frac{\sin(2\theta_{pq}) \sin(2\theta_{rs})}{4} \frac{\partial^2 E^{\Delta NO}_{0-1}\text{RDM}}{\partial \Delta_{pq} \partial \Delta_{rs}}
- \delta_{pr} \delta_{qs} \cos(2\theta_{pq}) \frac{\partial E^{\Delta NO}_{0-1}\text{RDM}}{\partial \Delta_{pq}}
\]  
(A29)

In the case of the second derivative, it is useful to express the 0-1RDM and pair contributions together, where

\[
E^{\Delta NO}_{0-1}\text{RDM-pair} = E^{\Delta NO}_{0-1}\text{RDM} + E^{\Delta NO}_{\text{pair}},
\]  
(A30)

and

\[
\frac{\partial^2 E^{\Delta NO}_{0-1}\text{RDM-pair}}{\partial \Delta_{pq} \partial \Delta_{rs}} = 4 \left( J_{pr} + J_{qs} - J_{ps} - J_{qr} \right)
- 2 \left( K_{pr} + K_{qs} - K_{ps} - K_{qr} \right)
\]  
(A31)

For the remainder of the contributions to the hessian, as with the gradient, the second derivatives can be calculated by replacing the coefficients, \( \zeta_{tu}, \xi_{tu} \) and \( \kappa_{tu} \), with their corresponding second derivative,

\[
\frac{\partial^2 \xi_{tu}}{\partial \Delta_{pq} \partial \Delta_{rs}} = \frac{O_{t}V_{u}}{4} \left( \frac{-\delta_{pt} \delta_{qu} \sqrt{\frac{n_t}{\Delta_{tu}}} - \delta_{pt} \delta_{qu} \sqrt{\frac{\Delta_{tu}}{n_t}}}{n_t \Delta_{tu} \sqrt{\Delta_{tu}}} \right)
+ \frac{O_{u}V_{t}}{4} \left( \frac{-\delta_{pu} \delta_{qt} \sqrt{\frac{n_u}{\Delta_{ut}}} - \delta_{pu} \delta_{qt} \sqrt{\frac{\Delta_{ut}}{n_u}}}{n_u \Delta_{ut} \sqrt{\Delta_{ut}}} \right).
\]  
(A32)
\[ \frac{\partial^2 k_{tu}}{\partial \Delta_{pq} \partial \Delta_{rs}} = (1 - \delta_{tu}) \left[ \sum_{v \neq w} \left( \frac{\partial^2 \xi_{tv}}{\partial \Delta_{pq} \partial \Delta_{rs}} + \xi_{tv} \frac{\partial^2 \xi_{uw}}{\partial \Delta_{pq} \partial \Delta_{rs}} + \frac{\partial^2 \xi_{uw}}{\partial \Delta_{pq} \partial \Delta_{rs}} + \frac{\partial^2 \xi_{tv}}{\partial \Delta_{pq} \partial \Delta_{rs}} \right) + W_t \frac{n_t}{2\sqrt{2}} \sum_v \frac{\partial^2 \xi_{tu}}{\partial \Delta_{pq} \partial \Delta_{rs}} + W_u \frac{n_u}{2\sqrt{2}} \sum_v \frac{\partial^2 \xi_{tv}}{\partial \Delta_{pq} \partial \Delta_{rs}} \right]. \] (A34)

The hessian terms involving derivatives with respect to both \( y \) and \( \theta \), can be expressed in terms of derivatives of \( \lambda \),

\[ H_{pq,rs}^{\phi} = -\sin(2\theta_{rs}) \left( \frac{\partial \lambda_{qp}}{\partial \Delta_{rs}} - \frac{\partial \lambda_{pq}}{\partial \Delta_{rs}} \right). \] (A35)

The derivative of the \( \lambda_{pq} \) component corresponding to \( E^0_{0-1RDM}^{\text{NO}} \) [first three terms of Equation (A4)] is given by

\[ \frac{\partial \lambda_{pq}^{0-1RDM}}{\partial \Delta_{rs}} = (2 - W_p) \left( \delta_{ps} - \delta_{pr} \right) \left( h_{pq} + G_{pq}^{1c} + \left( 1 - W_p \right) G_{pq}^{sc} + W_p G_{pq}^{sp} \right) \right] + 2 \left( (ps)q - (pr)q \right) - \left( (ps)q - (pr)q \right) \] (A36)

The remaining components of \( \frac{\partial \lambda_{pq}}{\partial \Delta_{rs}} \) are derived by substituting the coefficients, \( \xi_{tu} \), \( \xi_{tu} \), and \( \kappa_{tu} \), with their corresponding derivative.

1. J. Cizek, Adv. Chem. Phys. 14, 35 (1969).
2. B. Zejzorski and H. J. Monkhorst, Phys. Rev. A 24, 1668 (1981).
3. K. Andersson, P. Mainqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. 94, 5483 (1990).
4. D. K. W. Mok, R. Neumann, and N. C. Handy, J. Phys. Chem. 100, 6225 (1996).
5. N. C. Handy and A. J. Cohen, Mol. Phys. 99, 403 (2001).
6. D. Cremer, Mol. Phys. 99, 1899 (2001).
7. P. Pulay, Int. J. Quantum Chem. 111, 3273 (2011).
8. A. D. Becker, J. Phys. Chem. 138, 074109 (2013).
9. D. L. Crittenden, J. Phys. Chem. A 117, 3852 (2013).
10. T. Tschimochi and T. V. Voorhis, J. Chem. Phys. 141, 164117 (2014).
11. A. J. Wallace and D. L. Crittenden, J. Phys. Chem. A 118, 2138 (2014).
12. J. W. Hollett, H. Hoseini, and C. Menzies, J. Chem. Phys. 145, 084106 (2016).
13. E. Ramos-Cordoba, P. Salvador, and E. Matito, Phys. Chem. Chem. Phys. 18, 24015 (2016).
14. C. L. Benavides-Riveros, N. N. Lathiotakis, and M. A. L. Marques, Phys. Chem. Chem. Phys. 19, 12655 (2017).
15. F. A. Evangelista, J. Chem. Phys. 149, 039901 (2018).
16. M. Via-Nadal, M. Rodriguez-Mayorga, E. Ramos-Cordoba, and E. Matito, J. Phys. Chem. Lett. 10, 4032 (2019).
17. E. Fromager, J. Toulouse, and H. Jørgen Aa. Jensen, J. Chem. Phys. 126, 074111 (2007).
18. M. Weimer, F. D. Sala, and A. Görling, J. Chem. Phys. 128, 144109 (2008).
19. J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, and J. G. Angyán, Phys. Rev. Lett. 102, 096404 (2009).
20. B. G. Janesko, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. 131, 034110 (2009).
21. J.-D. Chai and M. Head-Gordon, J. Chem. Phys. 131, 174105 (2009).
22. A. Stoyanova, A. M. Teale, Toulouse, T. Helgaker, and E. Fromager, J. Chem. Phys. 139, 134113 (2013).
23. Y. Yamatani, K. Nakata, T. Uta, T. Takada, and K. Yamaguchi, Int. J. Quantum Chem. 106, 3312 (2006).
24. D. R. Rohr, J. Toulouse, and K. Pernal, Phys. Rev. A 82, 052502 (2010).
25. M. Piris, Phys. Rev. Lett. 119, 063002 (2017).
26. M. Piris, Phys. Rev. A 98, 022504 (2018).
27. A. J. Pérez-Jiménez, J. M. Pérez-Jordá, and J. C. Sancho-García, J. Chem. Phys. 127, 104102 (2007).
28. G. L. Manti, R. K. Carlson, S. Luo, D. Ma, J. Olsen, D. G. Truhlar, and L. Gagliardi, J. Chem. Theory Comput. 10, 3669 (2014).
29. R. K. Carlson, D. G. Truhlar, and L. Gagliardi, J. Chem. Theory Comput. 11, 4077 (2015).
30. L. Gagliardi, D. G. Truhlar, G. L. Manti, R. K. Carlson, C. E. Hoyer, and J. L. Bao, Acc. Chem. Res. 50, 66 (2017).
31. N. O. J. Malcolm and J. J. W. McDouall, Chem. Phys. Lett. 282, 121 (1998).
32. A. Kohn, M. Hanauer, L. A. Mück, T.-C. Jagau, and J. Gauss, WIREs Comput. Mol. Sci. 3, 176 (2013).
33. M. Hubert, E. D. Hedegård, and A. J. H. J. J. Chem. Theory Comput. 12, 2203 (2016).
34. J. W. Hollett and P.-F. Loos, J. Chem. Phys. in press (2019).
35. K. Hira, Chem. Phys. Lett. 190, 374 (1992).
36. S. Grimm and M. Waletzke, J. Chem. Phys. 111, 5645 (1999).
37. S. J. Grafenstein and D. Cremer, Chem. Phys. Lett. 316, 569 (2000).
38. C. Angeli, R. Cimring-Clark, and J. Malrieu, Chem. Phys. Lett. 350, 297 (2001).
39. C. Angeli, R. Cimring-Clark, and J. Malrieu, J. Chem. Phys. (2002).
P. A. Limacher, P. W. Ayers, P. A. Johnson, J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review A. M. Sand, D. G. Truhlar, and L. Gagliardi, J. Chem. Phys. T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria, The Journal of chemical physics 152, 174104 (2020).

A. Marie, F. Kosselki, and P.-F. Loos, The Journal of Chemical Physics 155, 104105 (2021).

A. M. Sand, D. G. Truhlar, and L. Gagliardi, J. Chem. Phys. 146, 034101 (2017).

J. P. Perdew, K. Burke, and M. Ernzerhof, Physical review letters 77, 3865 (1996).

A. J. Garza, I. W. Bulik, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. 142, 044109 (2015).

T. Stein, T. M. Henderson, and G. E. Scuseria, The Journal of chemical physics 140, 214113 (2014).

W. Poelmans, M. Van Raemdonck, B. Verstichel, S. De Baerdemacker, A. Torre, L. Lain, G. E. Massacesi, D. R. Alcoba, P. Bultinck, and D. Van Neck, Journal of chemical theory and computation 11, 4064 (2015).

Z. Chen, C. Zhou, and W. Wu, Journal of chemical theory and computation 11, 4102 (2015).

P. A. Limacher, Journal of chemical theory and computation 11, 3029 (2015).

R. Van Meer, O. Gritsenko, and E. Baerends, The Journal of chemical physics 148, 104102 (2018).

F. Brzdek, K. Boguslawski, P. Tecmer, and P. S. Zuchowski, Journal of chemical theory and computation 15, 4021 (2019).

N. Vu and A. E. DePrince III, The Journal of Chemical Physics 152, 244103 (2020).

W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. 110, 2800 (1999).

A. Szabo and N. S. Ostlund, Modern quantum chemistry (Dover, New York, 1996).

L. Bytautas, G. E. Scuseria, and K. Ruedenberg, The Journal of chemical physics 143, 094105 (2015).

J. W. Hollett and N. Pegoretti, J. Chem. Phys. 148, 164111 (2018).

R. Colle and O. Salvetti, Theor. Chim. Acta 37, 329 (1975).

R. Colle and O. Salvetti, Theor. Chim. Acta 53, 55 (1979).

This value differs from the value determined in reference 81, \( x + 1.73 \), because it was found to overestimate the correlation energy between electrons of parallel-spin in these H clusters.

T. Helgaker, P. Jorgensen, and J. Olsen, Molecular electronic-structure theory (John Wiley & Sons, 2014).

U. Bozkaya, J. M. Turney, Y. Yamaguchi, H. F. Schaefer III, and C. D. Sherrill, The Journal of chemical physics 135, 104103 (2011).

U. Bozkaya and C. D. Sherrill, The Journal of chemical physics 139, 054104 (2013).

U. Bozkaya and C. D. Sherrill, The Journal of chemical physics 141, 204105 (2014).

U. Bozkaya, Journal of chemical theory and computation 12, 1179 (2016).

J. Nocedal and S. J. Wright, Numerical Optimization (Springer, New York, New York, 2006).

C. Garrod and J. K. Percus, Journal of Mathematical Physics 5, 1756 (1964).

D. A. Mazzotti, Accounts of chemical research 39, 207 (2006).

D. A. Mazzotti, Phys. Rev. Lett. 108, 263002 (2012).

R. R. Li, M. Liebenthal, and E. DePrince, The Journal of Chemical Physics (2021).

D. A. Mazzotti, Physical Review A 65, 062511 (2002).

M. Rodríguez-Mayorga, E. Ramos-Cordoba, M. Via-Nadal, M. Piris, and E. Matito, Physical Chemistry Chemical Physics 19, 24029 (2017).

P. Piecuch, S. A. Kucharski, K. Kowalski, and M. Musial, Computer Physics Communications 149, 71 (2002).

P. Piecuch and M. Włoch, The Journal of chemical physics 123, 224105 (2005).

A. D. Becke, J. Chem. Phys. 98, 5648 (1993).

P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, The Journal of physical chemistry 98, 11623 (1994).

R. H. Hertwig and W. Koch, Chemical Physics Letters 268, 345 (1997).

G. Chaban, M. W. Schmidt, and M. S. Gordon, Theoretical Chemistry Accounts 97, 88 (1997).

K. Hirao, Chemical physics letters 190, 374 (1992).

K. Hirao, Chemical physics letters 196, 397 (1992).

K. Hirao, International Journal of Quantum Chemistry 44, 517 (1992).

K. Hirao, Chemical physics letters 201, 59 (1993).

G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Guna, M. E. Guido, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruit, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasatuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Wloch, P. Xu, F. Zahariev, and M. S. Gordon, The Journal of Chemical Physics 152, 154102 (2020).

Y. Garniron, K. Gasperich, T. Appelcourt, A. Benali, A. Ferte, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, Y. Garniron, K. Sundrival, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Wloch, P. Xu, F. Zahariev, and M. S. Gordon, The Journal of Chemical Physics 152, 154102 (2020).

M. Rodríguez-Mayorga, E. Ramos-Cordoba, M. Via-Nadal, M. Piris, and E. Matito, Theoretical Chemistry Accounts 97, 88 (1997).

K. Hirao, Chemical physics letters 190, 374 (1992).

K. Hirao, Chemical physics letters 196, 397 (1992).

K. Hirao, International Journal of Quantum Chemistry 44, 517 (1992).

K. Hirao, Chemical physics letters 201, 59 (1993).

G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Guana, M. E. Guido, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruit, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasatuchana, M. W. Schmidt, J. Shen, L. Slpenchenko, M. Sosonkina, V. Sundrival, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Wloch, P. Xu, F. Zahariev, and M. S. Gordon, The Journal of Chemical Physics 152, 154102 (2020).

Y. Garniron, K. Gasperich, T. Appelcourt, A. Benali, A. Ferte, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, Y. Garniron, K. Sundrival, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Wloch, P. Xu, F. Zahariev, and M. S. Gordon, The Journal of Chemical Physics 152, 154102 (2020).

M. Piris and J. M. Ugalde, J. Comput. Chem. 30, 2078 (2009).