International Journal of QUANTUM CHEMISTRY

Pair Natural Orbitals in Explicitly Correlated Second-Order Møller–Plesset Theory

David P. Tew*[a] and Christof Hättig[b]

An improved formulation of pair natural orbital (PNO)–based explicitly correlated Møller–Plesset second-order perturbation theory is presented. Increased efficiency has been achieved by modifying the representation of the strong orthogonality projector used in the evaluation of the F12 intermediates. We demonstrate two simple relationships between the truncation errors arising from PNOs in the virtual space and those arising from the PNOs in the strong orthogonality operators. This permits robust error control through a single parameter. © 2012 Wiley Periodicals, Inc.

DOI: 10.1002/qua.24098

Introduction

In a recent article,[1] we reported a preliminary investigation of explicitly correlated[2–4] pair natural orbital[5–9] (PNO) methods for second-order Møller–Plesset perturbation theory (MP2). In this approach, the first-order pair functions are expanded using PNOs $\bar{a}_{ij}$, which are specific to every occupied pair ij, plus a geminal (F12) contribution that explicitly satisfies the first-order electron coalescence conditions.[10–12]

$$|\mu_{ij}\rangle = \sum_{\bar{a}_{ij}\bar{b}_{ij}} t^{ij}_{\bar{a}_{ij}\bar{b}_{ij}} |\bar{a}_{ij}\bar{b}_{ij}\rangle + \hat{Q}_{ij} f(r_{12}) |\widetilde{ij}\rangle \tag{1}$$

The amplitudes $t^{ij}_{\bar{a}_{ij}\bar{b}_{ij}}$ are optimized, but the F12 contribution is kept fixed. The correlation factor[13,14] is $f(r_{12}) = (1 - e^{-r_{12}})$ and $|\widetilde{ij}\rangle$ is the spin-flipped (or rational-generator) linear combination of spin-orbital determinants $|ij\rangle$.[15–18] Spin-orbital formulation is used throughout this article. The strong orthogonality operator for each pair $\hat{Q}_{ij}$ is central to this work and is discussed in detail below. The PNOs are chosen to diagonalize the pair difference density $\bar{D}^{ij}_{ab}$

$$\omega^{ij}_{\bar{a}_{ij}} \delta_{\bar{a}_{ij}\bar{b}_{ij}} = \sum_{ab} \bar{D}^{ij}_{ab} L^{ij}_{a\bar{a}_{ij}} L^{ij}_{b\bar{a}_{ij}} \tag{2}$$

$$\bar{D}^{ij}_{ab} = \sum_{c} d^{ij}_{ac} d^{ij}_{bc} \tag{3}$$

$$d^{ij}_{ab} = \left(\varepsilon_a + \varepsilon_b - f^i_i - f^j_j\right)^{-1} g^{ij}_{ab} - r^{ij}_{ab} \tag{4}$$

where $g^{ij}_{ab} = \langle ab|r_{12}^{-1}|ij\rangle$, $r^{ij}_{ab} = \langle ab|f(r_{12})|\widetilde{ij}\rangle$ and ε_a, f^i_i are the diagonal elements of the Fock matrix, where the virtual orbitals are canonical, but the occupied orbitals have been localized. This choice of PNOs is equivalent[19] to a single value decomposition of the difference amplitudes d^{ij}_{ab} for each pair and, therefore, leads to a very compact representation of the pair functions when PNOs with $\omega^{ij}_{\bar{a}_{ij}}$ under a selected threshold are discarded. Note that we have dropped the normalization factor from Eq. (3) to make this relationship explicit. We find that the normalization factor plays no useful role and is very close to unity, except for strongly multireference cases, where it serves to bias weak pairs over strong pairs, which does not seem physically justifiable.

With localized occupied orbitals, the number of significant pairs ij grows linearly with the number of atoms and the number of PNOs required to accurately represent each first-order pair function rapidly becomes independent of system size. In the PNO basis, therefore, storage of the significant doubles amplitudes and electron-repulsion integrals displays an asymptotic linear dependence on system size, rather than the apparent N^4 scaling of the four-index amplitudes and integrals. Because of the presence of the F12 geminal function, an average of less than 40 PNOs per significant pair are sufficient for an accuracy of better than 5 kJ/mol in energy differences. Therefore we expect the combination of PNO and F12 theory to form the basis of highly efficient approximate coupled-cluster methods, applicable to large molecules.

The incorporation of F12 functions into the wave function, vital for reducing basis set truncation errors, incurs significant overhead connected to the computation of Hamiltonian matrix elements over the geminal basis functions. The magnitude of the overhead depends critically on the number of complementary auxiliary (CA) functions used in the resolution of the identity approximation of the F12 integrals. In our preliminary study,[1] we extended PNO theory to encompass the CA orbitals, such that their number also tends to a constant with increasing system size. In this work, we refine our initial choice of CA-PNOs through a reconsideration of the representation of the strong orthogonality projector, demonstrating significantly reduced costs for building the F12 intermediates. Furthermore, we establish a relationship

[a] *D. P. Tew*
School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom
E-mail: david.tew@bristol.ac.uk

[b] *C. Hättig*
Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany

Contract grant sponsor: Deutsche Forschungsgemeinschaft; Royal Society (University Research Fellow scheme, D.P.T.)

© 2012 Wiley Periodicals, Inc.

between the virtual and CA-PNO truncation thresholds, which ensures that the CA truncation error is at least an order of magnitude less than the virtual basis truncation error. Both of these advances are important steps toward a robust and widely applicable PNO-F12 methodology.

Strong Orthogonality Operators

The strong orthogonality operator $\hat{Q}_{ij}$ projects onto the space of doubly excited pair functions that are complementary to the basis of PNO pairs $|\bar{a}_{ij}\bar{b}_{ij}\rangle$. Two representations of $\hat{Q}_{ij}$ are required in F12 theory, ${}^{A}\hat{Q}_{ij}$ and ${}^{B}\hat{Q}_{ij}$. The latter is less accurate and is only adopted when the former leads to three-electron integrals. In our preliminary study, the choice of ${}^{A}\hat{Q}_{ij}$ was derived from the form that gives compact expressions for canonical F12 theory,[20] which was readily available from our MP2-F12 implementation in the Turbomole package.[21] Here, we introduce three further choices, which are better suited to PNO-F12 theory. A summary of the indices used to denote the relevant molecular orbital (MO) and PNO spaces for the Hartree–Fock (HF) and CA basis sets is presented in Figure 1. We define the following one-electron projection operators:

$$\begin{aligned} O &= \sum_{k} |k\rangle\langle k| & \bar{O} &= \sum_{\bar{k}_{ij}} |\bar{k}_{ij}\rangle\langle\bar{k}_{ij}| \\ \bar{V} &= \sum_{\bar{a}_{ij}} |\bar{a}_{ij}\rangle\langle\bar{a}_{ij}| & \bar{P} &= \sum_{\bar{p}_{ij}} |\bar{p}_{ij}\rangle\langle\bar{p}_{ij}| \\ V' &= \sum_{a'_{ij}} |a'_{ij}\rangle\langle a'_{ij}| & \bar{V}' &= \sum_{\bar{a}'_{ij}} |\bar{a}'_{ij}\rangle\langle\bar{a}'_{ij}| \\ V'' &= \sum_{a''} |a''\rangle\langle a''| & \bar{V}'' &= \sum_{\bar{a}''_{ij}} |\bar{a}''_{ij}\rangle\langle\bar{a}''_{ij}| \end{aligned} \tag{5}$$

The bar indicates that the orbitals have been transformed to a PNO basis. We will refer to virtual PNOs as V-PNOs and we will use CA-PNOs and O-PNOs to refer to PNOs in the CA and occupied spaces, respectively. X-PNOs will be used to denote any of the auxiliary PNOs (CA, occupied, but not virtual) used to approximate the strong orthogonality projector. We will also use one-electron projectors with a tilde in place of the bar to differentiate between two different choices of PNOs (e.g., $\bar{a}'_{ij}$ and $\tilde{a}'_{ij}$) for the same orbital space (a'_{ij}). Each operator is pair specific (except O and V''), but the indices ij have been dropped to improve readability.

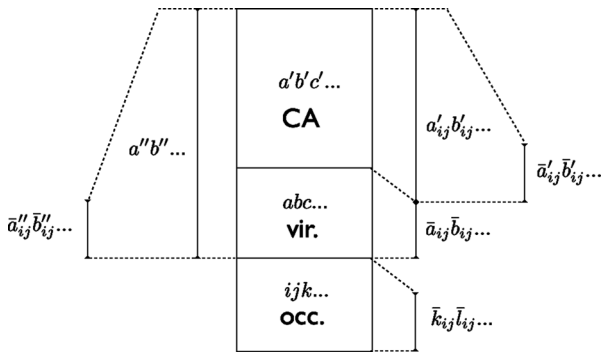


Figure 1. Orbital spaces. Overbars indicate a PNO truncated space. Dotted lines indicate the contraction, expansion, or retention of an orbital space, starting from the parent occupied (occ.) virtual (vir.), and complementary virtual (CA) spaces.

For ${}^{B}\hat{Q}_{ij}$, our recommendation is unchanged from our previous work.

$${}^{B}\hat{Q}_{ij} = \bar{V}_1\bar{V}'_2 + \bar{V}'_1\bar{V}_2 + \bar{V}'_1\bar{V}'_2 \tag{6}$$

where $\bar{a}'_{ij}$ are chosen to diagonalize the pair density

$$\bar{D}^{ij}_{a'_{ij}b'_{ij}} = \sum_{c_{ij}} r^{ij}_{a'_{ij}c_{ij}} r^{ij}_{b'_{ij}c_{ij}} \tag{7}$$

This is equivalent to a single value decomposition of the $\bar{V}_1V'_2$ (or equivalently $V'_1\bar{V}_2$) components of ${}^{B}\hat{Q}_{ij}f(r_{12})|\widetilde{ij}\rangle$ and is, therefore, not optimally suited to represent the $V'_1V'_2$ component. However, as this term occurs in only one small energy contribution that is usually negligible for energy differences,[22] it is our opinion that it is not worth the computational expense required to construct CA-PNOs suited to $V'_1V'_2$.

In this work, we consider four choices for ${}^{A}\hat{Q}_{ij}$

$${}^{A_0}\hat{Q}_{ij} = 1 - \bar{P}_1\bar{P}_2 - O_1\tilde{V}'_2 - \tilde{V}'_1O_2 \tag{8}$$
$${}^{A_1}\hat{Q}_{ij} = 1 - \tilde{O}_1\tilde{O}_2 - \bar{V}_1\bar{V}_2 - O_1\tilde{V}''_2 - \tilde{V}''_1O_2 \tag{9}$$
$${}^{A_2}\hat{Q}_{ij} = 1 - \tilde{O}_1\tilde{O}_2 - \bar{V}_1\bar{V}_2 - \bar{O}_1V''_2 - V''_1\bar{O}_2 \tag{10}$$
$${}^{A_3}\hat{Q}_{ij} = 1 - \tilde{O}_1\tilde{O}_2 - \bar{V}_1\bar{V}_2 - \bar{O}_1\bar{V}''_2 - \bar{V}''_1\bar{O}_2 \tag{11}$$

where $\bar{a}_{ij}$ are those that diagonalize the pair difference density Eq. (3) and the remaining PNOs diagonalize the following pair densities

$$\tilde{a}'_{ij}: \quad \tilde{D}^{ij}_{a'_{ij}b'_{ij}} = \sum_{k} r^{ij}_{a'_{ij}k} r^{ij}_{b'_{ij}k} \tag{12}$$
$$\tilde{a}''_{ij}: \quad \tilde{D}^{ij}_{a''b''} = \sum_{k} r^{ij}_{a''k} r^{ij}_{b''k} \tag{13}$$
$$\bar{a}''_{ij}: \quad \bar{D}^{ij}_{a''b''} = \sum_{\bar{k}_{ij}} r^{ij}_{a''\bar{k}_{ij}} r^{ij}_{b''\bar{k}_{ij}} \tag{14}$$
$$\tilde{k}_{ij}: \quad \tilde{D}^{ij}_{kl} = \sum_{m} r^{ij}_{km} r^{ij}_{lm} \tag{15}$$
$$\bar{k}_{ij}: \quad \bar{D}^{ij}_{kl} = \sum_{c''} r^{ij}_{kc''} r^{ij}_{lc''} \tag{16}$$

${}^{A_0}\hat{Q}_{ij}$ is the form used in our earlier work. Expanding to

$${}^{A_0}\hat{Q}_{ij} = 1 - O_1O_2 - \bar{V}_1\bar{V}_2 - O_1\bar{V}_2 - \bar{V}_1O_2 - O_1\tilde{V}'_2 - \tilde{V}'_1O_2 \tag{17}$$

it is clear that this is suboptimal, as the O_1O_2 contribution depends quadratically on system size and the $O_1\bar{V}_2$, $\bar{V}_1O_2$, $O_1\tilde{V}'_2$, and $\tilde{V}'_1O_2$ contributions depend linearly on the system size. In both ${}^{A_1}\hat{Q}_{ij}$ and ${}^{A_2}\hat{Q}_{ij}$, the quadratic dependence has been removed and the four terms linear with system size have been combined into two. Finally, ${}^{A_3}\hat{Q}_{ij}$ is independent of system size. The number of CA-PNOs with nonzero contributions to ${}^{A_0}\hat{Q}_{ij}$ is bounded by the number of occupied orbitals. The number that have non-negligible contribution is further restricted by the number of

occupied orbitals spatially close to the orbitals *i* or *j*. Exactly the same considerations hold for the CA orbitals of $^{A_1}\hat{Q}_{ij}$ and our investigations show that for a given truncation threshold, the average number of CA-PNOs in $^{A_1}\hat{Q}_{ij}$ is typically only two or three more than in $^{A_0}\hat{Q}_{ij}$. Typical values for molecules containing H, C, N, O, and F range between 10 and 40. The number of O-PNOs is also obviously bounded by the number of occupied orbitals, and in practice, we find that the number of O-PNOs retained with a given threshold is two or three less than the number of CA-PNOs.

In Tables 1–3, we have collected the scaling of the operations necessary to form the various PNOs, and to use them to construct the key F12 intermediate $V_{ij}^{ij} = \langle \widetilde{ij}|f(r_{12})\hat{Q}_{ij}r_{12}^{-1}|ij\rangle$, required in explicitly correlated MP2 and coupled cluster theory. The relative computational merits of the three new choices for $\hat{Q}_{ij}$ can be read from Tables 2 and 3. Table 2 refers to the steps required to compute the PNO vectors. In our density fitting implementation, these are

Table 1. Notation for scaling behavior.

N	Number of atomic orbitals in the HF basis
N''	Number of atomic orbitals in the HF plus CA basis
A	Number of orbitals in the density fitting basis
N_{ij}	Number of significant occupied pairs
O	Number of occupied-MOs
V	Number of virtual-MOs
V''	Number of virtual plus CA molecular orbitals
k	A constant in place of the number of PNO orbitals

Table 2. Scaling behavior for the evaluation of intermediates in the determination of various PNO vectors.

PNO		B_{xi}^{Q}	B_{yi}^{Q}	r_{xy}^{ij}, g_{xy}^{ij}	D_{xy}^{ij}	$L_{x\bar{x}}^{ij}$
$\bar{a}_{ij}$	$\bar{V}\bar{V}$	ON^2A	ON^2A	O^2V^2A	$N_{ij}V^3$	$N_{ij}V^3$
$\tilde{k}_{ij}$	$\tilde{O}\tilde{O}$	ON^2A	ON^2A	$N_{ij}O^2A$	$N_{ij}O^3$	$N_{ij}O^3$
$\tilde{a}''_{ij}$	$O\tilde{V}''$	$ONN''A$	ON^2A	$N_{ij}OV''A$	$N_{ij}OV''^2$	$N_{ij}V''^3$
$\bar{k}_{ij}$	$\bar{O}V''$	ON^2A	$ONN''A$	$N_{ij}OV''A$	$N_{ij}O^2V''$	$N_{ij}O^3$
$\bar{a}''_{ij}$	$\bar{O}\bar{V}''$	$ONN''A$	ON^2A	$kN_{ij}V''A$	$kN_{ij}V''^2$	$N_{ij}V''^3$

Table 3. Scaling behavior for the evaluation of intermediates in the determination of various contributions to V_{ij}^{ij}.

PNO		$B_{x_{ij}i}^{Q}$	$B_{y_{ij}i}^{Q}$	r_{xy}^{ij}, g_{xy}^{ij}	V_{ij}^{ij}
$\bar{a}_{ij}$	$\bar{V}\bar{V}$	$kN_{ij}VA$	$kN_{ij}VA$	$k^2N_{ij}A$	k^2N_{ij}
$\tilde{k}_{ij}$	$\tilde{O}\tilde{O}$	$kN_{ij}OA$	$kN_{ij}OA$	$k^2N_{ij}A$	k^2N_{ij}
$\tilde{a}''_{ij}$	$O\tilde{V}''$	—	$kN_{ij}V''A$	$kN_{ij}OA$	$kN_{ij}O$
$\bar{k}_{ij}$	$\bar{O}V''$	$kN_{ij}OA$	—	$kN_{ij}V''A$	$kN_{ij}V''$
$\bar{a}''_{ij}$	$\bar{O}\bar{V}''$	$kN_{ij}OA$	$kN_{ij}V''A$	$k^2N_{ij}A$	k^2N_{ij}

- Computation of the density fitting intermediates B_{xi}^{Q} and B_{yi}^{Q}, the cost of which depends on the size of the required orbital spaces *x* and *y* (for simplicity, we have assumed that the time-dependent step is the first partial transformation to the occupied space and that this is done without screening).
- Computation of the two-electron integrals r_{xy}^{ij} and g_{xy}^{ij}, where, for example, $g_{xy}^{ij} = \sum_Q B_{xi}^{Q}B_{yj}^{Q} - \sum_Q B_{xj}^{Q}B_{yi}^{Q}$.
- Computation of the relevant pair densities from Eqs. (3), (7), (12–16).
- Diagonalization of the pair densities to obtain the PNO vectors $L_{x\bar{x}_{ij}}^{ij}$.

Table 3 refers to the cost of computing the various contributions to V_{ij}^{ij}.

- Transformation of the density-fitting intermediates from *x* to $\bar{x}_{ij}$.
- Computation of r_{xy}^{ij} and g_{xy}^{ij}, where, for example, $g_{xy}^{ij} = \sum_Q B_{x_{ij}i}^{Q}B_{y_{ij}j}^{Q} - \sum_Q B_{x_{ij}j}^{Q}B_{y_{ij}i}^{Q}$.
- Computing the contribution to V_{ij}^{ij} from, for example, $\sum_{x_{ij}y_{ij}} r_{xy}^{ij}g_{xy}^{ij}$

Although $^{A_3}\hat{Q}_{ij}$ is attractive in that it has the correct scaling behavior, this choice is more expensive than both $^{A_1}\hat{Q}_{ij}$ and $^{A_2}\hat{Q}_{ij}$ in the present PNO-F12 scheme. The additional cost arises because the PNO vectors for both $\bar{k}_{ij}$ and $\bar{a}''_{ij}$ must be computed. Similarly, transformation of the three-index intermediates must be transformed for both $\bar{k}_{ij}$ and $\bar{a}''_{ij}$. In our work, no *a priori* localization assumptions are applied, that is, the PNOs for each pair are constructed using the entire set of virtual MOs. Although this avoids problems associated with domain choices, it precludes the possibility of linear scaling algorithms, as the PNO vectors are themselves n^2, where n is a measure of system size. The linear scaling property of $^{A_3}\hat{Q}_{ij}$ cannot currently be exploited.

Table 3 shows that the cost of using $^{A_1}\hat{Q}_{ij}$ and $^{A_2}\hat{Q}_{ij}$ differs only in the final n^2 contraction and is, therefore, essentially the same. This also holds true for the construction of $V_{\bar{a}\bar{b}}^{ij}$ required in explicitly correlated coupled-cluster theory, where the transformation of $B_{a''b}^{Q}$ into the PNO basis is the most expensive step. The cost of forming $\bar{k}_{ij}$, however, is significantly cheaper than forming $\tilde{a}''_{ij}$ principally because the diagonalization of $\tilde{D}_{a''b''}^{ij}$ is avoided (Table 2). Furthermore, the cost of storage for the transformation matrices is much less for the O-PNOs than for the CA-PNOs. $^{A_2}\hat{Q}_{ij}$ is, therefore, the most efficient choice.

Some general remarks are pertinent concerning the PNO-MP2-F12 method. The only n^5 step is the initial build of the first-order amplitudes. Thereafter, all steps are n^4 or less if a strong pair approximation is applied. In canonical density fitted MP2-F12 theory, the dominant n^5 step is the contraction of the intermediates B_{xi}^{Q} to form the two-electron integrals r_{xy}^{ij}. In PNO-MP2-F12, this step becomes n^3 at worse. However, in the regime where few pairs are neglected, the savings gained in the construction of the MP2-F12 intermediates through transforming to PNOs are largely (but not entirely) cancelled by the cost of computing the PNO vectors. For an efficient program, it is imperative to reduce the overhead of forming and transforming to the PNOs. Work is in progress in this direction in our laboratories.

PNO Truncation Thresholds

One of the attractive features of PNO methods is that the truncation error can be smoothly reduced by adjusting a single parameter, the occupation number threshold, which determines the number of V-PNOs used in the wave function expansion. In PNO-F12 methods, several additional auxiliary PNOs are required for the evaluation of the F12 Hamiltonian matrix elements. For

reliable error control, the error in the energy due to truncation of these X-PNOs should be at least an order of magnitude less than the V-PNO truncation error. Here, we report calculations that reveal two simple relationships between the V-PNO truncation threshold and the corresponding X-PNO truncation thresholds.

We have chosen to use two test sets for our numerical investigations: Bakowies'[23] set of 105 small molecules and the S66 set of noncovalent complexes.[24] For the 105 set, the H_2 molecule was added and all structures were taken from Ref. [25]. The basis set limit MP2 energies are available from Ref. [26]. For the S66 set, the structures and basis set limit estimates were taken from Ref. [27]. Liakos et al. [28] have recently assessed the PNO-CEPA method using the S22 set and the interaction energies of noncovalent complexes form a relatively stringent test for PNO methods.

We performed a series of calculations to determine the individual V- and X-PNO truncation errors for counterpoise corrected[29,30] interaction energies of the 24 smallest noncovalent complexes in the S66 set. The aug-cc-pVDZ atomic orbital basis set[31,32] was used with the specially optimized CA basis set.[33] The occupied orbitals were localized using Boys'[34] localization. The V- and X-PNO truncation errors were computed as the difference between the untruncated MP2-F12 interaction energy and the interaction energy from a PNO-MP2-F12 calculation where only the PNO type under consideration was truncated using a threshold of V- or X-PNOTOL. From these data, we determined the value of X-PNOTOL where the absolute X-PNO truncation error is 0.1 times the V-PNO truncation error at a given V-PNOTOL. This was repeated for a series of V-PNOTOL $= 10^{-6}$, $10^{-7}, \ldots 10^{-10}$ and for each X-PNO and averaged over the 24 interaction energies. The results are plotted in Figure 2.

The X-PNOTOL for all the X-PNOs of the ${}^{A}\hat{Q}_{ij}$ projectors show similar relationships to V-PNOTOL but the relationship for the X-PNOs of ${}^{B}\hat{Q}_{ij}$ is strikingly different. For ${}^{A}\hat{Q}_{ij}$, the relationships between the logarithms of V-PNOTOL and the X-PNOTOLs are remarkably linear, with slopes varying between 0.5 and 0.6. The relationship between X-PNOTOL and V-PNOTOL for the X-PNOs of the ${}^{B}\hat{Q}_{ij}$ projector is less well defined. In Figure 2, two curves are plotted, one where the hybrid approximation has been applied,[35,36]

$$ {}^{B}\hat{Q}_{ij}^{\mathrm{Hyb}} = \bar{V}_1\bar{V}'_2 + \bar{V}'_1\bar{V}_2 \tag{18} $$

and one where it has not. The X-PNO truncation error for $\bar{a}'_{ij}$ is clearly dominated by the $\bar{V}'_1\bar{V}'_2$ contributions. Both curves deviate from linearity on the log–log plot, but for loose truncation thresholds the gradient is close to 1.0. These relationships between the X-PNO and V-PNO truncation errors must originate from the spectral properties of the integrals r^{ij}_{kl}, $r^{ij}_{ka''}$ and $r^{ij}_{aa'}$ in relation to that of t^{ij}_{ab}. The fact that all the X-PNO truncation errors from the ${}^{A}\hat{Q}_{ij}$ contributions are similar results from the fact that for all of the contributing integrals, the number of nonzero eigenvalues is limited by the number of occupied orbitals close in space to orbitals i or j.

To ensure that the maximum X-PNO truncation errors are less than 0.1 times the V-PNO truncation error, our calculations suggest to use

$$ -\log_{10}\ \text{X-PNOTOL} = 5.5 - 0.5\log_{10}\ \text{V-PNOTOL} \tag{19} $$

for all X-PNOs of the ${}^{A}\hat{Q}_{ij}$ projector and to use

$$ -\log_{10}\text{X-PNOTOL} = 0.5 - \log_{10}\text{V-PNOTOL} \tag{20} $$

for the CA-PNOs of the ${}^{B}\hat{Q}_{ij}$ projector. This is sufficient to ensure that the maximum ratio between the X-PNO and V-PNO truncation errors is less than 0.1 when the hybrid approximation is applied. This choice is also safe when the full ${}^{B}\hat{Q}_{ij}$ is used, except for the limit of tight V-PNOTOL thresholds, where the truncation error for $\bar{a}'_{ij}$ can limit the accuracy of the calculation.

In Table 4, we present the average (and maximum) ratio between the V-PNO and total X-PNO truncation errors in noncovalent interaction energies, using X-PNOTOLs according to Eqs. (19) and (20). We present values for the aug-cc-pVXZ orbital (and CA) sets with X = D, T, and Q. The 24 smallest complexes of the S66 were used as our test set and only the smallest 12 complexes were computed using the T and Q sets. In all calculations, ${}^{A2}\hat{Q}_{ij}$ was used and values for both the MP2-F12/A and MP2-F12/B methods are presented (the $\bar{V}'_1\bar{V}'_2$ component of ${}^{B}\hat{Q}_{ij}$ is not used in MP2-F12/A[37]).

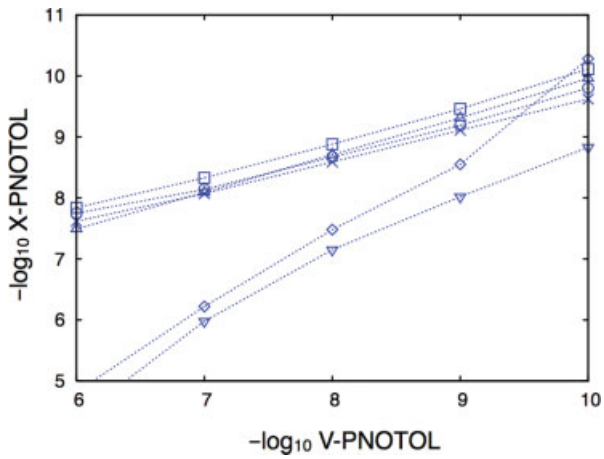


Figure 2. Relationship between X-PNOTOL and V-PNOTOL such that the X-PNO truncation error is 0.1 times the V-PNO truncation error on average. ○ $\bar{k}_{ij}$, × $\bar{k}_{ij}$, □ $\bar{a}''_{ij}$, △ $\bar{a}'_{ij}$, ◇ $\bar{a}'_{ij}$ and ▽ $\bar{a}'_{ij}$ using the hybrid approximation. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 4. Mean (and maximum) ratios between V-PNO and total X-PNO truncation errors in noncovalent MP2-F12/A and MP2-F12/B interaction energies for various values of $T = -\log_{10}$ V-PNOTOL, using Eqs. (19) and (20) for X-PNOTOL.

T	aug-cc-pVDZ	aug-cc-pVTZ	aug-cc-pVQZ
6 A	0.08 (0.17)	0.10 (0.23)	0.16 (0.83)
B	0.09 (0.22)	0.10 (0.18)	0.09 (0.17)
7 A	0.05 (0.13)	0.04 (0.10)	0.04 (0.08)
B	0.05 (0.11)	0.05 (0.06)	0.04 (0.06)
8 A	0.04 (0.23)	0.02 (0.04)	0.02 (0.06)
B	0.05 (0.16)	0.03 (0.05)	0.02 (0.03)
9 A	0.04 (0.10)	0.03 (0.10)	0.02 (0.05)
B	0.10 (0.83)	0.01 (0.02)	0.01 (0.02)
10 A	0.06(0.28)	0.02 (0.06)	0.02 (0.08)
B	0.27 (1.89)	0.01 (0.06)	0.01 (0.03)

On average, the sum of the X-PNO truncation errors resulting from the ${}^{A}\hat{Q}_{ij}$ operator are less than 0.1 times the V-PNO truncation error for the full range of V-PNOTOL thresholds, as desired. The maximum value of the total X-PNO truncation error over our set is never greater than 0.3 times the V-PNO truncation error, for the MP2-F12/A method. For MP2-F12/B, the truncation errors are similar for loose V-PNOTOL thresholds, but for tight thresholds of 10^{-9} and 10^{-10}, the error in the $\bar{V}_1'\bar{V}_2'$ term dominates. For larger orbital basis sets, the situation is reversed and the $\bar{V}_1'\bar{V}_2'$ error is insignificant for tight thresholds, but dominates for small thresholds. We do not envisage that PNO methods will be used with tight V-PNOTOL thresholds and small orbital basis sets, or with loose thresholds and large orbital sets, so these issues are inconsequential. In Table 5, we present example data from PNO-MP2-F12/aug-cc-pVDZ calculations on the interaction energy of Uracil with AcOH. An average of only 16 V-PNOs per pair is sufficient for a truncation error of only 1.1 kJ/mol and a deviation from the basis set limit of 2.8 kJ/mol. For this accuracy, less than 20 X-PNOs per significant pair were required on average.

Table 5. Data from PNO-MP2-F12/aug-cc-pVDZ calculations on the correlation energy contribution to the interaction of Uracil with AcOH: mean (and maximum) number of V- and X-PNOs per significant pair in the complex, signed V-PNO, X-PNO truncation errors, and total deviation from the basis set limit of −16.75. $T = -\log_{10}$ V-PNOTOL, and energies are in kJ/mol.

	Mean No. PNOs				Truncation error		
T	$\bar{a}_{ij}$	$\tilde{k}_{ij}$	$\bar{k}_{ij}$	$\bar{a}'_{ij}$	V	X	BSL
6	9 (28)	11 (22)	15 (25)	9 (36)	3.15	0.21	5.01
7	16 (46)	12 (23)	16 (28)	18 (68)	1.07	0.04	2.76
8	28 (74)	14 (26)	18 (30)	35 (103)	0.40	0.01	2.07
9	46 (115)	17 (27)	22 (32)	62 (164)	0.13	0.00	1.79
10	73 (156)	19 (29)	25 (34)	103 (228)	0.04	0.00	1.69

In Table 6, we report V-PNO and X-PNO truncation errors for a set of 105 reactions involving the 106 small molecules (taken from the Supporting Information of Ref. [1]). We tabulate mean absolute and maximum V-PNO truncation errors in the reaction energies for a series of V-PNOTOL thresholds, where no X-PNO truncation is applied. We additionally tabulate the mean absolute and maximum change in the reaction energy due to X-PNO truncation using X-PNOTOLs of Eqs. (19) and (20). Ratios of truncation errors are not particularly informative for reaction energies. We report MP2-F12/A and MP2-F12/B values computed using the cc-pVDZ-F12, cc-pVTZ-F12, and cc-pVQZ-F12 orbital and CA basis sets,[33,38] and as before, ${}^{A_2}\hat{Q}_{ij}$ was used.

Table 6. Mean (and maximum) V-PNO and total X-PNO truncation errors in MP2-F12/A and MP2-F12/B reaction energies.

	cc-pVDZ-F12		cc-pVTZ-F12		cc-pVQZ-F12	
T	A	B	A	B	A	B
6 V	52.2 (198)	79.6 (260)	41.0 (172)	81.6 (287)	55.9 (273)	83.9 (312)
X	0.3 (2)	5.8 (21)	0.3 (1)	5.5 (19)	0.3 (1)	5.7 (22)
7 V	19.2 (79)	31.1 (95)	17.1 (61)	32.7 (110)	17.4 (61)	31.9 (108)
X	0.1 (1)	0.9 (3)	0.1 (1)	1.1 (4)	0.1 (1)	1.1 (4)
8 V	9.5 (31)	11.2 (31)	5.6 (35)	12.4 (36)	6.6 (35)	12.1 (36)
X	0.0 (0)	0.4 (2)	0.0 (0)	0.2 (0)	0.0 (0)	0.2 (1)
9 V	3.0 (9)	2.9 (9)	4.4 (18)	3.9 (11)	1.9 (8)	4.1 (13)
X	0.0 (0)	0.5 (2)	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)
10 V	0.9 (9)	0.7 (2)	2.1 (9)	1.0 (3)	0.8 (4)	1.1 (3)
X	0.0 (0)	0.6 (3)	0.0 (0)	0.0 (0)	0.0 (0)	0.0 (0)

Values are in J/mol per valence electron. $T = -\log_{10}$ V-PNOTOL and X-PNOTOL is set using Eqs. (19) and (20).

Reaction energies are much less sensitive to X-PNO truncation errors than noncovalent interaction energies. For the reaction energies, the X-PNO truncation errors are completely negligible compared to the V-PNO truncation errors. The only exception is for the MP2-F12/B method with the cc-pVDZ-F12 basis and tight V-PNOTOL thresholds, where the $\bar{V}_1'\bar{V}_2'$ term dominates, but since using tight thresholds with a cc-pVDZ-F12 basis is nonsensical, this situation should not arise in practical applications. In Figure 3, we plot mean absolute deviations from the basis set limit reaction energies as a function of the number of V-PNOs. The curves using untruncated and truncated X-PNOs, with our recommended thresholds, are almost indistinguishable on this plot. It is somewhat curious that the cc-pVQZ-F12 curve lies above the cc-pVTZ-F12 curve, except for V-PNOTOL $< 10^{-9}$. This indicates that when the cc-pVQZ-F12 basis is used, functions are retained in the wave function expansion that are unimportant for energy differences. The exponents of the cc-pVQZ-F12 basis functions are those of the cc-pV5Z basis and have not been optimized for F12 calculations. Consequently, p, d, and f functions with tight exponents are present, which have significant contribution to r_{ab}^{ij} but not to t_{ab}^{ij}. As our choice of density Eq. (3) for the PNO construction is based on fixed F12 amplitudes, these tight virtuals are erroneously retained in the PNO expansion. This issue is not easy to overcome through modified-PNO definitions, but if the tightest p, d, and f functions are removed from each cc-pVQZ-F12 atomic basis, a much improved truncation behavior is obtained.

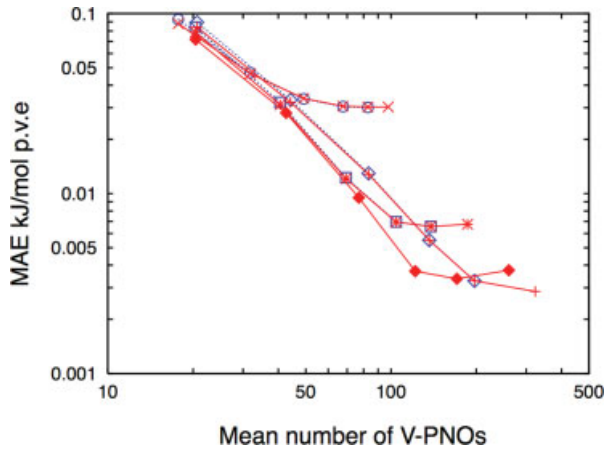


Figure 3. Mean absolute deviation of reaction energies from the basis set limit in kJ/mol per valence electron of the reactants with $-\log_{10}$ V-PNOTOL = 6, 7, 8, 9, 10, ∞. Solid lines are without X-PNO truncation, × cc-pVDZ-F12 and ∗ cc-pVTZ-F12, + cc-pVQZ-F12, ♦ cc-pVQZ-F12′ (no tight p, d, or f), dashed lines with default X-PNO truncation, ○ cc-pVDZ-F12, □ cc-pVTZ-F12, ◇ cc-pVQZ-F12. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Conclusions

We have presented an improved formulation of local explicitly correlated Møller–Plesset theory with PNOs. The efficiency of our method has been enhanced by choosing an alternative set of auxiliary PNOs for the F12 intermediates, which reduces the computational overhead required to form the PNO vectors. Furthermore, we have linked the truncation thresholds for virtual and auxiliary PNOs, so that the accuracy of a calculation can be controlled through a single parameter. We find that the truncation energy errors arising from each of the PNO types used in our method can be broadly related through two simple formulas that originate from the spectral properties of the integrals involved.

Using a cc-pVTZ-F12 basis set, only 40 virtuals on average per pair are required to compute reaction energies within 2 kJ/mol of the basis set limit in our calculations. The number of PNOs used to evaluate F12 intermediates is also less than 40 per pair on average. The energy error arising from the approximations in the F12 intermediates is at least 1 order of magnitude less than the truncation error arising from the virtual PNOs. This was achieved by linking the V-PNO and X-PNO truncation thresholds based on errors for noncovalent interaction energies, which form a stringent test for PNO methods.

The main advantage of PNO approaches to low-scaling methods for computing electronic energies is the compactness of the representation of the pair functions. The space of PNOs adapts naturally to the complexity of each pair function, which means that weak pairs are automatically treated with few PNOs or none at all. This is not the case for domain-based local explicitly correlated methods,[39–44] or for the recent orbital specific virtual approach.[45] However, in contrast to domain-based local methods,[46] where the correlating orbitals are chosen *a priori*, PNO methods have large overheads related to the formation of and transformation between the PNO vectors for each pair. It is essential to reduce these overheads, if the PNO-F12 method is to be a cost-effective approach to low-scaling coupled-cluster energies.

Keywords: explicit correlation • localization • pair natural orbitals

How to cite this article: D. P. Tew, C. Hättig, *Int. J. Quantum Chem.* **2013**, *113*, 224–229. DOI: 10.1002/qua.24098

[1] D. P. Tew, B. Helmich, C. Hättig, *J. Chem. Phys.* **2011**, *135*, 074107.
[2] W. Kutzelnigg, W. Klopper, *J. Chem. Phys.* **1991**, *94*, 1985.
[3] C. Hättig, W. Klopper, A. Köhn, D. P. Tew, *Chem. Rev.* **2012**, *112*, 4.
[4] L. Kong, F. A. Bischoff, E. F. Valeev, *Chem. Rev.* **2012**, *112*, 75.
[5] W. Meyer, *Int. J. Quantum Chem. Symp.* **1971**, *5*, 341.
[6] W. Meyer, *J. Chem. Phys.* **1975**, *58*, 1017.
[7] R. Ahlrichs, F. Driessler, H. Lischka, V. Staemmler, W. Kutzelnigg, *J. Chem. Phys.* **1975**, *62*, 1225.
[8] F. Neese, F. Wennmohs, A. Hansen, *J. Chem. Phys.* **2009**, *130*, 114108.
[9] F. Neese, A. Hansen, D. G. Liakos, *J. Chem. Phys.* **2009**, *131*, 064103.
[10] T. Kato, *Commun. Pure Appl. Math.* **1957**, *10*, 151.
[11] R. T. Pack, W. Byers Brown, *J. Chem. Phys.* **1966**, *45*, 556.
[12] D. P. Tew, *J. Chem. Phys.* **2008**, *129*, 014104.
[13] S. Ten-no, *Chem. Phys. Lett.* **2004**, *398*, 56.
[14] D. P. Tew, W. Klopper, *J. Chem. Phys.* **2005**, *123*, 074101.
[15] S. Ten-no, *J. Chem. Phys.* **2007**, **126**, 014108.
[16] G. Knizia, T. B. Adler, H.-J. Werner, *J. Chem. Phys.* **2009**, *130*, 054104.
[17] D. Bokhan, S. Bernadotte, S. Ten-no, *J. Chem. Phys.* **2009**, *131*, 084105.
[18] D. P. Tew, W. Klopper, *Mol. Phys.* **2010**, *108*, 315.
[19] F. Bell, D. Lambrecht, M. Head-Gordon, *Mol. Phys.* **2010**, *108*, 2759.
[20] E. F. Valeev, *Chem. Phys. Lett.* **2004**, *395*, 190.
[21] R. A. Bachorz, F. A. Bischoff, A. Glöß, C. Hättig, S. Höfener, W. Klopper, D. P. Tew, *J. Comput. Chem.* **2011**, *32*, 2492.
[22] O. Marchetti, H.-J. Werner, *Phys. Chem. Chem. Phys.* **2008**, *10*, 3400.
[23] D. Bakowies, *J. Chem. Phys.* **2007**, *127*, 084105.
[24] J. Řezáč, K. E. Riley, P. Hobza, *J. Chem. Theory Comput.* **2011**, *7*, 2427.
[25] W. Klopper, B. Ruscic, D. P. Tew, F. A. Bischoff, S. Wolfsegger, *Chem. Phys.* **2010**, *356*, 14.
[26] F. A. Bischoff, S. Wolfsegger, D. P. Tew, W. Klopper, *Mol. Phys.* **2009**, *107*, 963.
[27] J. Řezáč, P. Jurečka, K. E. Riley, J. Černý, H. Valdes, K. Pluháčková, K. Berka, T. Řezáč, M. Pitoňák, J. Vondrášek, P. Hobza, *Collect. Czech. Chem. Commun.* **2008**, *73*, 1261.
[28] D. G. Liakos, A. Hansen, F. Neese, *J. Chem. Theory Comput.* **2011**, *7*, 76.
[29] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553.
[30] D. P. Tew, W. Klopper, *J. Chem. Phys.* **2006**, *125*, 094302.
[31] T. H. Dunning, Jr., *J. Chem. Phys.* **1989**, *90*, 1007.
[32] D. E. Woon, T. D., Jr., *J. Chem. Phys.* **1994**, *100*, 2975.
[33] K. E. Yousaf, K. A. Peterson, *J. Chem. Phys.* **2008**, *129*, 184108.
[34] S. F. Boys, *In Quantum Theory of Atoms, Molecules and the Solid State*, P. O. Löwdin, Ed.; Academic: New York, 1966, p. 253.
[35] H.-J. Werner, T. B. Adler, F. Manby, *J. Chem. Phys.* **2007**, *126*, 164102.
[36] W. Klopper, *J. Chem. Phys.* **2004**, *120*, 10890.
[37] W. Klopper, C. C. M. Samson, *J. Chem. Phys.* **2002**, *116*, 6397.
[38] K. A. Peterson, T. B. Adler, H.-J. Werner, *J. Chem. Phys.* **2008**, *128*, 084102.
[39] F. R. Manby, H.-J. Werner, T. B. Adler, A. J. May, *J. Chem. Phys.* **2006**, *124*, 094103.
[40] H.-J. Werner, *J. Chem. Phys.* **2008**, *129*, 101103.
[41] T. B. Adler, H.-J. Werner, F. R. Manby, *J. Chem. Phys.* **2009**, *130*, 054106.
[42] T. B. Adler, H.-J. Werner, *J. Chem. Phys.* **2009**, *130*, 241101.
[43] T. B. Adler, H.-J. Werner, *J. Chem. Phys.* **2011**, *135*, 144117.
[44] J. Friedrich, D. P. Tew, W. Klopper, M. Dolg, *J. Chem. Phys.* **2010**, *132*, 164114.
[45] J. Yang, Y. Kurashige, F. R. Manby, G. K. L. Chan, *J. Chem. Phys.* **2011**, *134*, 044123.
[46] S. Saebo, P. Pulay, *Annu. Rev. Phys. Chem.* **1993**, *44*, 213.

Received: 15 January 2012
Revised: 26 February 2012
Accepted: 29 February 2012
Published online on 14 April 2012