A Multireference Coupeld-Eletron Pair Approximation combined with Complete-Active Space Perturbation Theory in Local Pair-Natural Orbital Framework

Masaaki Saitow[1, a)] and Takeshi Yanai[2, 3]

[1)]*Department of Chemistry, Graduate School of Science, Nagoya University, 1-5 Chikusa-ku, Nagoya, Aichi 464-8602, Japan*

[2)]*Department of Chemistry, Graduate School of Science and Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, 1-5 Chikusa-ku, Nagoya, Aichi 464-8602, Japan*

[3)]*Japan Science and Technology Agency, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

(Dated: 9 December 2019)

The Complete Active Space Second-order Perturbation Theory (CASPT2) has been one of the most widely used methods for caluclating the electronic structure of the multireference systems with semi-quantitative accuracy. As a simple, yet higher-order approach, we have developed a hybrid theory of the CASPT2 and a multireference variant of the Coupled-Electron Pair Approximation (CEPA) in the fully-internally contracted (FIC) framework. In the newly developed theory (CEPT2), the external components of the wave functions, which usually give the most significant contribution to the dynamic correlation energy, are solved at the CEPA level while the rests are treated at the CASPT2 level. Moreover, using an automatic expression and code generation technique, we have implemented the CASPT2 and CEPT2 approaches in the pair-natural orbital (PNO) framework to reduce their computational costs with respect to the number of atomic orbital (AO) basis. To highlight the accuracy of the CEPT2 approach and to assess the errors caused by the PNO truncation, we have performed benchmark calculations on small- to medium-size molecules including dissociation of N_2 molecule, the singlet-triplet splitting of the free-base porphyrin molecule. We demonstrate that by tightening the truncation thresholds the PNO-CEPT2 energy converges toward the canonical counterpart and that the PNO-CEPT2 results are more accurate than that of PNO-CASPT2 as long as the same truncation thresholds are used.

[a)]Electronic mail: msaitow514@gmail.com

I. INTRODUCTION

Electronic structures that cannot be well approximated by a single Slater determinant are common in chemistry. Electronic wave function of the low-spin open-shell states where there are multiple anti-ferromagnetically coupled electrons is a very typical example of such cases.[1] The electron correlation effect that plays an important role in these problems is often referred to as static correlation.[2] However, to obtain quantitave accuracy, it is often necessary to consider dynamic correlation which is the remaining component of the correlation effect. The most stratghtforward way to take into account both types of correlation in a well-balanced manner is the multireference (MR) wave function theories[3,4] that are based on the multiconfigurational self-consistent field (CASSCF) reference function.[5–7] Note that there are tremendous attempts to utilize the single-reference (SR) framework[8–17] for capturing the static correlation effect.

Amongst the MR wave function theories, the Complete-Active-Space Second-order Perturbation Theory (CASPT2) pioneered by Roos and coworkers[18,19] in early 90s have been widely used as a fundamental tool for achieving quantitative accuracy for systems of complicated electronic structure including organic molecules of doubly-excited nature[20] and mono-[21,22] and multi-nuclear[23,24] transition metal complexes. The CASPT2 potential energy surface can often show singularity due to the near degeneracy especially if the active space is too small; this particular issue associated with CASPT2 model is known as the intruder state problem. There have been several regularization techniques proposed in order to remedy this issue including real[25] and imaginary[26] shifts. A newer variant of CASPT called N-Electron Valence-state Perturbvation Theory (NEVPT) was proposed by Angeli *et al.*[27–29] using Dyall's zeroth-order Hamiltonian[30] to overcome some of the shortcomings of CASPT approach. As both approaches converge to the conventional second-order Møller-Plesset (MP2) perturbation theory[31] in the closed-shell SR limit, floating-point operations (FPOs) needed to calculate CASPT2 and NEVPT2 energies scale at least as $O(N_c^2 N_v^2)$ for system with a fixed number of active molecular orbitals (MOs) where N_c and N_v represent the number of doubly-occupied and virtual MOs, respectively. Before calculating CASPT2 energy, one has to solve the first-order amplitude equations and hence the memory or storage requirement scales also as $O(N_c^2 N_v^2)$. For a system with more active MOs, the prefactors of both memory and FPO requiremtns steeply increase because their working equations involve the reduced-density matrices (RDMs) of up to four-body nature. Therefore, routine application of CASPT2 methods to systems with more than 1,000 atomic orbital (AO) functions often becomes prohibitively expensive.

To improve the applicability of CASPT2 approach, numerous efforts have been made including use of frozen-natural orbitals (FNOs)[32,33] in the virtual space in conjunction with the Cholesky decomposition technique for reducing the costs of MO integral transformation[34–37] and of the tensor-hyper-contraction-based (THC-) CASPT2.[38] The density-fitting (DF) or resolution-of-the-identity (RI) approximation has also been used for accelerating CASPT2 computation.[39] Recently, Katz and Werner developed a CASPT2 approach in the so-called pair-natural orbital (PNO) framework, which has been coined PNO-CASPT2, and demonstrated its applicability to large, real-life MR systems composed of up to approximately 3,000 AO functions.[40,41]

The original concept of PNOs was already realized in 60s in order to accelerate the concergence of the configuration interaction (CI) expansions.[42–48] Importantly, the PNO-based MRCI theory and the MR Coupled-Electron Pair Approximation (PNO-MCCEPA) using internally-contracted (IC) configuration functions were developed by Taylor[49] and Staemmler *et al.*,[50] respectively.

The use of PNOs in the context of modern wave function theories was first introduced by Neese and coworkers in their local PNO-based CEPA (LPNO-CEPA)[51] and Coupled-Cluster (LPNO-CCSD)[52–54] methods implemented in Orca package.[55] The LPNO-CEPA/CCSD is applicable to systems composed of up to 2,000 AO basis which is far beyond the reach of canonical counterpart while reconvering almost 99.9 % of the canonical correlation energy. Later they have developed a more sohisticated variant called domain-based LPNO-CC (DLPNO-CC) methods for large, real-life systems involving hundreds of atoms and thousands of AO functions.[56–61] The DLPNO scheme has been extended to MR wave function theories resulting in DLPNO-MR-CC based on the Mukherjee-Kutzelnigg formalism[62–64] and DLPNO-NEVPT2[65] method. In a similar vein, variants of PNO-based reduced-scaling wave function methods have been developed by many research groups including PNO-LCC methods by Werner *et al.*[66–72] and a variant of PNO-CC method and PNO-based theories for the excited states by Hättig *at al.*[73–81] to only name a few.

The IC basis functions spanning the first-order interacting space (FOIS) are classified into three types according to their excitation characters: Those configurations with single and double excitations into virtual space are called semi-internal and external, respectively. The remaining IC configurations which do not involve any excitations into virtual space are labeled as internal. In the closed-shell SR limit, the external configurations converge simply to the doubly-excited configurations while all the excitations vanish. In the general MR cases, actually, the semi-internal and internal IC configurations involve not only simple single-excitations but also the pseudo-double excitations because some of the active MOs may possess occupation number very close to zero.

Actually, in 80s, Roos *et al.* developed a primitive version of CASPT2 which involves only the external IC configurations[82] and, as summarized in Ref. 83, reported that the semi-internal subspaces in FOIS can have large impact on the dynamic correlation energy. Note that the first version of IC-MR-CC theory proposed by Banerjee and Simons also lacks the important semi-internal components in the wave function.[84]

In this paper, we report a hybrid of IC-MR-CEPA and CASPT2, which is conined CEPT2, as a higher-order, yet simple extention to the conventional CASPT2 method for capturing more dynamic correlation effects. The external configurations in the first-order CEPT wave function are treated at the IC-MR-CEPA level while the remaining semi-internal and the internal configurations are kept at the CASPT2 level. The coupling between two Ansätze has been accomplished in a similar vein to the CIPT2 model by Werner and coworkers.[85] In the closed-shell SR limit, the CEPT2 model converges to the Linearized Coupled-Cluster Doubles (L-CCD) or equivalently the CEPA/0 Doubles (CEPA/0(D)) that recovers more dynamic correlation energy than the MP2 model. Note that Fink and coworkers proposed a variant of MR perturbation theory[86–88] that also converges to the L-CCD or CEPA/0(D) model in the closed-shell SR limit at the second-order level and is known as the so-called Retantion of Excitaton degree Perturbation Theory (REPT).

In the original CIPT2 model,[89,90] the reference configurations, and all the excited configurations involving excitations solely from the active MO space[91] are treated at the MRCI level while leaving the remaining ones at the CASPT2 level in order to remedy the intruder state problem without employing any emprical shifts. Actually, treating those configurations in the IC framework at the MRCI level is known to cause several computational difficulties the most severe one of which is the appearance of 5-RDMs.[92] An active space 5-RDM is a tensor quantity associated with 10 active MO indices and hence for those systems with 16 active MOs generation and storage of the 5-RDM require more than 8.0 terabytes of disk space. Therefore, in the partially IC-MRCI models[90,93] implemented in Molpro package,[94] such configurations are uncontracted to remove them from the working equations.

Later, by rewriting the Hamiltonian matrix element into a commutator-based expression as done in the NEVPT2 formalism,[27] the MRCI theory in a fully IC framework (FIC-MRCI) without 5-RDM was formulated.[95–98] However, tensor contractions involving the 4-RDM and the t-amplitude in the FIC-MRCI σ-equation can still be a bottleneck even for systems with relatively small active space as their computational scaling is either of $O(N_c N_a^8)$ or $O(N_a^8 v)$ with N_a being the number of active MOs. In the CEPT2 model, on the other hand, such contractions do not appear in the

residua and all the 4-RDMs are contracted to the Fock operator thus appearing as an effective 3-body tensor with $O(N_a^8)$ costs just as in case of CASPT2. Such a Fock-contracted 4-RDM can be pre-computed only once a CEPT2 energy computation while in the FIC-MRCI the t-contracted 4-RDM should be constructed in each iteration for the diagonalization of the Hamiltonian matrix.

Even though the full IC-MR-CEPA/0 lacks size-consistency, all the CEPA terms introduced are fully-connected and thus the near-size-consistent nature of CASPT2 is directly inherited to the CEPT2 model. Use of MR-CEPA ansatz for the external subspaces introduces several terms involving 4-external MO integrals contracted to the t-amplitude, which may potentially make the whole computation quite demanding even for relatively small systems. This type of contraction terms are often referred to as the 4-External Exchange Operators (4-EEOs) and are known to be the most expensive terms in the SR-CEPA and CCSD computations with $O(N_c^2 N_v^4)$ computational cost. To overcome this issue, we introduce PNO expansion in the external subspace in a similar vein to Ref. 40. Derivation and implementation of our PNO-CASPT2/CEPT2 scheme have been achieved by an automatic code generation framework which was used in our previous works[95,96] and has been extended to handle PNOs.

This paper proceeds as follows: In Sec. II, we show how the PNOs are defined in the MR cases and used in the CASPT2 formalism. The PNO-CEPT2 Ansatz is given as a extension to the PNO-CASPT2 model and the strategy we have chosen to use for the automatic code generation in the PNO framework is addressed. In Sec. III, the accuracy of the canonical CEPT2 model and the PNO trucation errors are validiated for N_2 dissociation. The isomerization energy profile calculated at the PNO-CEPT2 level of theory is shown and compared to those by the CASPT2 and the previous theoretical results. The excitation energies (EEs) for the dipole-accessible and dipole-forbidden states of butadiene and longer polyene chains calculated by the PNO-CEPT2 theory are given and compared to the experimental and the theoretical values. Accuracy of the PNO-CEPT2 model are examined by using various PNO truncation thresholds in calculating the Singlet–Triplet energy (S–T) gap for the free-base porphyrin. In Sec. IV finally conclusions are drawn.

II. THEORY

A. Notations and Convention

Throughout this paper, we use $ijkl\cdots$, $pqrs\cdots$ and $abcd\cdots$ to represent the doubly-occupied, active and virtual MOs, respectively, while the generic MOs are indicated by $wxyz\cdots$. Derivation and the implementation of the working equations of the PNO-CEPT approach are based on the reduced spin-free excitation operators[99–102]

$$E^w_x = \sum_{\sigma=\{\alpha,\beta\}} a^{w\sigma} a_{x\sigma} \tag{1}$$

$$E^{wx}_{yz} = \sum_{\sigma,\tau=\{\alpha,\beta\}} a^{w\sigma} a^{x\tau} a_{z\tau} a_{y\sigma} \tag{2}$$

where a^w (a_w) represents the Fermionic creation (annihilation) operator that acts on w-th spatial MO. By using Eqs. (1) and (2), the Born-Oppenheimer Hamiltonian is given as

$$H = \sum_{wx} h_{wx} + \frac{1}{2} \sum_{wxyz} (wy|xz) E^{wx}_{yz} \tag{3}$$

where h_{wx} and $(wy|xz)$ are one- and two-electron integrals. The redundant internally-contracted basis (r-ICBs) that span the first-order interacting space (FOIS) are generated by applying the following eight types of excitation operators to the reference function ($|\Psi_0\rangle$):

$$E^{ab}_{ij},\ E^{ab}_{pi},\ E^{ab}_{pq} \tag{4}$$

for the external subspaces,

$$\{E^{aq}_{ip},\ E^{qa}_{ip}\},\ E^{pa}_{ij},\ E^{ra}_{pq} \tag{5}$$

the semi-internal subspaces and

$$E^{pq}_{ij},\ E^{qr}_{ip} \tag{6}$$

for the internal subspaces. The energy and residual equations involve an expectation value of a string of E-operators with respect to the reference such as $\langle\Psi_0|E^{pq}_{ab}E^{yz}_{wx}\cdots|\Psi_0\rangle$. Such an expectation value is evaluated by invoking Wick's theorem for the E-operators and is reduced to a reduced-density matrix (RDM) composed only of active MO indices:

$$D^p_q = \langle\Psi_0|E^p_q|\Psi_0\rangle \tag{7}$$

$$D^{pq}_{rs} = \langle \Psi_0 | E^{pq}_{rs} | \Psi_0 \rangle \tag{8}$$

The r-ICBs are non-orthonornal and linearly-independent set of functions as a basis of FOIS. Therefore, the amplitude equations of MR wave function methods are solved in linearly-independent and orthonormalized IC representation spanned by the non-redundant (nr-) ICBs:

$$|\Psi^{ab}_{ij}\rangle,\ |\Psi^{ab}_{\rho i}\rangle,\ |\Psi^{ab}_{\rho}\rangle \tag{9}$$

for the external subspaces,

$$|\Psi^{a}_{\rho i}\rangle,\ |\Psi^{\rho a}_{ij}\rangle,\ |\Psi^{a}_{\rho}\rangle \tag{10}$$

the semi-internal subspaces and

$$|\Psi^{\rho}_{ij}\rangle,\ |\Psi_{i\rho}\rangle \tag{11}$$

for the internal subspaces where $\rho\tau\cdots$ are used to represent the eigenvectors of metric for each subspace (Eqs. (4) – (6)).

B. Brief Summary of PNO-CASPT2 Algorithm

As we follow a similar vein to the original formalism by Katz and Werner,[40,41] we only outline the PNO-CASPT2 theory in this section. In CASPT2 method, the generalized Fock operator is used as the zeroth-order Hamiltonian

$$F = \sum_{wx} f_{wx} E^{x}_{w} \tag{12}$$

where f represents the generalized Fock matrix

$$f_{wx} = (f_{\mathrm{core}})_{wx} + \sum_{pq} \left[(wx|pq) - \frac{1}{2}(wp|xq) \right] D^{p}_{q} \tag{13}$$

$$(f_{\mathrm{core}})_{wx} = h_{wx} + \sum_{i} [2(wx|ii) - (wi|xi)]. \tag{14}$$

The PNO-CASPT2 wave function is defined as

$$\begin{aligned} |\Psi\rangle = &\sum_{ij} \sum_{a_{ij} b_{ij}} t^{ij}_{a_{ij} b_{ij}} |\Psi^{a_{ij} b_{ij}}_{ij}\rangle + \sum_{\rho i} \sum_{a_{\rho i} b_{\rho i}} t^{\rho i}_{a_{\rho i} b_{\rho i}} |\Psi^{a_{\rho i} b_{\rho i}}_{\rho i}\rangle \\ &+ \sum_{\rho} \sum_{a_\rho b_\rho} t^{\rho}_{a_\rho b_\rho} |\Psi^{a_\rho b_\rho}_{\rho}\rangle + \sum_{\rho i j} \sum_{a_{ij}} t^{ij}_{a_{ij}\rho} |\Psi^{\rho a}_{ij}\rangle + \sum_{\rho i} \sum_{a} t^{\rho i}_{a} |\Psi^{a}_{\rho i}\rangle \end{aligned}$$

$$+ \sum_{\rho} \sum_{a} t_a^{\rho} |\Psi_{\rho}^{a}\rangle + \sum_{\rho ij} t_{\rho}^{ij} |\Psi_{ij}^{\rho}\rangle + \sum_{\rho i} t_{\rho}^{i} |\Psi_{i}^{\rho}\rangle \tag{15}$$

in the nr-ICB which is constructed by orthonormalizing and canonicalizing the redundant r-ICBs. For instance, an nr-ICB with double excitations to the active space, $\{\Psi_{ij}^{\rho}\}$, is related to its redundant counterpart as

$$|\Psi_{ij}^{\rho}\rangle = \sum_{pq} C_{pq}^{\rho} |\tilde{\Psi}_{ij}^{pq}\rangle = \sum_{pq} C_{pq}^{\rho} E_{ij}^{pq} |\Psi_0\rangle \tag{16}$$

where the orthonormalization matrix, $\mathbf{C}$, is obtained by solving the generalized eigenvalue problem:

$$\sum_{rs} \langle \Psi_0 | E_{pq}^{ij} F E_{rs}^{ij} | \Psi_0 \rangle C_{rs}^{\rho} = E_{\rho ij} \sum_{rs} \langle \Psi_0 | E_{pq}^{ij} E_{rs}^{ij} | \Psi_0 \rangle C_{rs}^{\rho}. \tag{17}$$

The PNO-CASPT2 amplitudes are determined by solving the residual equations

$$\begin{aligned} r^I &= \frac{\partial}{\partial t^I} \mathscr{F}_{\text{PNO-CASPT2}} \\ &= \sum_J \langle \Psi^I | F - \langle F \rangle | \Psi_J \rangle t^J - \langle \Psi^I | H | \Psi_0 \rangle \to 0 \end{aligned} \tag{18}$$

where IJ represent one of the eight non-redundant subspaces of FOIS. The Hylleraas-type energy functional for PNO-CASPT2 ansatz is given as

$$\mathscr{F}_{\text{PNO-CASPT2}} = \langle \Psi | H | \Psi_0 \rangle + \langle \Psi_0 | H | \Psi \rangle + \langle \Psi | F - \langle F \rangle | \Psi \rangle. \tag{19}$$

The amplitudes in the r-ICB representation ($\tilde{\mathbf{t}}$) can be obtained by back-transforming the nr-ICB counterpart ($\mathbf{t}$) as

$$\tilde{t}_{ij}^{pq} = \sum_{\rho} C_{pq}^{\rho} t_{ij}^{\rho}. \tag{20}$$

In typical canonical CASPT2 implementation, the residua that correspond to Eq. (18) are evaluated in r-ICB representation using the $\tilde{\mathbf{t}}$ amplitudes, and then transformed into the nr-ICB representation as

$$r_{ij}^{\rho} = \sum_{\rho} C_{pq}^{\rho} \tilde{r}_{ij}^{pq}. \tag{21}$$

for the $\{\Psi_{ij}^{\rho}\}$ subspace.

In Eq. (15), the PNOs are introduced only in the external subspaces and a semi-internal subspace spanned by $\{\Psi_{ij}^{\rho a_{ij}}\}$ function. Since the PNOs for $\{\Psi_{\rho}^{a_\rho b_\rho}\}$ and $\{\Psi_{\rho i}^{a_{\rho i} b_{\rho i}}\}$ subspaces are dependent

also on the nr-ICBs, in the PNO-CASPT2 implementation, one cannot back-transform the external amplitudes (**t**) and the residua (**r**) into the r-ICB representation while taking advantage of compactness of the PNO space. Therefore, in our PNO-CASPT2 implementation, we do not perform any back-transformation into the r-ICB representatation and the residua are constructed directly in the nr-ICB representation using the orthonormalized amplitudes. Thee explicit formuals for the orthonormalized amplitudes are given in Appendix I.

The PNOs for $ij-$, $\rho i-$ and $\rho-$pairs are generated by diagonalizing the pair-densities for the corresponding pairs:

$$(D^{ij})_{ab} = \langle \Psi_{ij} | E^a_b | \Psi_{ij} \rangle \quad \text{(for } \{\Psi^{ab}_{ij}\} \text{ subspsace)} \tag{22}$$

$$(D^{\rho i})_{ab} = \langle \Psi_{\rho i} | E^a_b | \Psi_{\rho i} \rangle \quad \text{(for } \{\Psi^{ab}_{\rho i}\} \text{ subspsace)} \tag{23}$$

$$(D^{\rho})_{ab} = \langle \Psi_{\rho} | E^a_b | \Psi_{\rho} \rangle \quad \text{(for } \{\Psi^{ab}_{\rho}\} \text{ subspsace)} \tag{24}$$

$$\tag{25}$$

For $\{\Psi^{\rho a_{ij}}_{ij}\}$ subspace, we use the $ij-$PNOs obtained as eigenfunctions of Eq. (22). Those PNOs with smaller occupation numbers than a user defined threshold are screened out. The PNO truncation error for a pair is estimated by taking difference between full and truncated semi-local (SL-) CASPT2-D pair-energies.

$$e_{ij} = \langle \Psi_0 | H | \Psi_{ij} \rangle \tag{26}$$

$$e_{\rho i} = \langle \Psi_0 | H | \Psi_{\rho i} \rangle \tag{27}$$

$$e_{\rho} = \langle \Psi_0 | H | \Psi_{\rho} \rangle \tag{28}$$

$$\tag{29}$$

Moreover, only those pairs with pair-energy larger than a given threshold proceeds to the iterative PNO-CASPT2/CEPT2 step. The SL-CASPT2-D pair-functions are used for defining the pair-densities

$$|\Psi_{ij}\rangle = \sum_{ab} \frac{(ia|jb)}{F_{ii} + F_{jj} - \epsilon_a - \epsilon_b} E^{ab}_{ij} |\Psi_0\rangle \tag{30}$$

$$|\Psi_{\rho i}\rangle = \sum_{ab} \frac{\langle \Psi^{ab}_{\rho i} | H | \Psi_0 \rangle}{F_{ii} - e_{\rho} - \epsilon_a - \epsilon_b} \sum_{p} C^{\rho}_{p} E^{ab}_{ip} |\Psi_0\rangle \tag{31}$$

$$|\Psi_{\rho}\rangle = \sum_{ab} \frac{\langle \Psi^{ab}_{\rho} | H | \Psi_0 \rangle}{e_{\rho} - \epsilon_a - \epsilon_b} \sum_{pq} C^{\rho}_{pq} E^{ab}_{pq} |\Psi_0\rangle. \tag{32}$$

The numerators in the SL-CASPT2-D amplitudes involve 2-external two-electron integrals and hence this part will be a computational bottleneck in a whole PNO-CASPT2 computation. In our current implementation, we generate a full set of canonical 2-external integrals using the resolution-of-the-identity (RI) approximation and store them on disk. This choice limits the applicability of our PNO-CASPT2 only up to systems composed of not more than 2,000 AO functions. To overcome such a difficulty, domain truncation should be introduced in the RI 3-index integral generation and the final 2-external integral construction steps as done in the original PNO-CASPT2 implementations,[40,41] in the DLPNO-CC and NEVPT2 algorithms by Neese and coworkers[56–61,65,103] and in the PNO-LCC programs by Werner *et al.*.[66–68,104]

In the canonical CASPT2 implementation, the nr-ICBs are constructed as eigenfunctions of diagonal blocks of the zeroth-order Hamiltonian by solving the generalized eigenvalue problems akin to Eq. (17) for all the eight subspaces. Therefore, programming the residua (Eq. 18) of diagonal blocks in the first term is quite simple. However, in the PNO-based counterpart, all the matrix blocks involving at least one DOMO index should be programmed, leading to a drastic increase in the number of terms that should be derived and programmed. To this end, we extended our code generator framework[95,96,105] such that codes for the PNO-based tensor contractions can also be generated. In our program, the PNO-CASPT2 residul equations are solved by the preconditioned conjugate gradient (PCG) algorithm.

We have also implemented real and imaginary shifts in our PNO-CASPT2 program in the same way as the canonical CASPT2 program. The conventional formulation of the imaginary shifts is based on the diagonality in the diagonal blocks of zeroth-order Hamiltonian matrix in FOIS which is not satisfied if the DOMOs are localized. Therefore, in our PNO-CASPT2 implementatin, the off-diagonal elements in the diagonal blocks are discarded, leading to a semi-canonical imaginary shift.

C. The Coupled-Electron Pair Perturbation Theory

In the PNO-CEPT2 ansatz, we minimize the following energy functonal

$$\mathscr{F}_{\text{PNO-CEPT2}} = \mathscr{F}^{\text{ext}}_{\text{PNO-CEPA/0}} + \mathscr{F}_{\text{PNO-CASPT2}} - \mathscr{F}^{\text{ext}}_{\text{PNO-CASPT2}} \tag{33}$$

$$\mathscr{F}^{\text{ext}}_{\text{PNO-CEPA/0}} = \langle \Psi^{\text{ext}} | H - E_0 | \Psi^{\text{ext}} \rangle \tag{34}$$

$$\mathscr{F}^{\text{ext}}_{\text{PNO-CASPT2}} = \langle \bar{\Psi}^{\text{ext}} | H | \Psi_0 \rangle + \langle \Psi_0 | H | \bar{\Psi}^{\text{ext}} \rangle + \langle \bar{\Psi}^{\text{ext}} | F - \langle F \rangle | \bar{\Psi}^{\text{ext}} \rangle \tag{35}$$

where $\bar{\Psi}^{\text{ext}}$ represents the external part of the wave function

$$\begin{aligned}|\Psi^{\text{ext}}\rangle &= |\Psi_0\rangle + |\bar{\Psi}^{\text{ext}}\rangle \\ &= |\Psi_0\rangle + \Bigg[\sum_{ij}\sum_{a_{ij}b_{ij}} t^{ij}_{a_{ij}b_{ij}}|\Psi^{a_{ij}b_{ij}}_{ij}\rangle \\ &+ \sum_{\rho i}\sum_{a_{\rho i}b_{\rho i}} t^{\rho i}_{a_{\rho i}b_{\rho i}}|\Psi^{a_{\rho i}b_{\rho i}}_{\rho i}\rangle + \sum_{\rho}\sum_{a_\rho b_\rho} t^{\rho}_{a_\rho b_\rho}|\Psi^{a_\rho b_\rho}_{\rho}\rangle\Bigg]. \end{aligned} \tag{36}$$

By differentiating the above functional with respect to the amplitudes, one obtains the PNO-CEPT2 residual equations as

$$r^E = \langle\Psi^E|H|\Psi_0\rangle + \sum_{E'}\langle\Psi_E|H - E_0|\Psi_{E'}\rangle t^{E'} + \sum_A \langle\Psi_E|F|\Psi_A\rangle t^A \to 0 \tag{37}$$

$$r^A = \langle\Psi^A|H|\Psi_0\rangle + \sum_{E}\langle\Psi_A|H|\Psi_E\rangle t^E + \sum_{A'}\langle\Psi_A|F - \langle F\rangle|\Psi_{A'}\rangle t^{A'} \to 0 \tag{38}$$

where $\{E\}$ and $\{A\}$ represent the external and the remaining subspaces in FOIS, respectively. If one evaluates the diagonal block for $\{\Psi^{a_\rho b_\rho}_\rho\}$ straightforwardly using Wick's theorem, there are two terms involving 4-body RDM whose size is approximately 32 gigabyte for systems composed of 16 active MOs. To eliminate such expensive contractions, we employ a commutator-based transformation originally proposed by Angeli and coworkers[27–29]

$$\begin{aligned} r^\rho_{a_\rho b_\rho} &= \sum_\tau\sum_{c_\tau d_\tau}\langle\Psi^\rho_{a_\rho b_\rho}|H - E_0|\Psi^{c_\tau d_\tau}_\tau\rangle t^\tau_{c_\tau d_\tau} \\ &= \sum_\tau\sum_{c_\tau d_\tau}\sum_{pqrs} C^\rho_{pq}\langle\Psi_0|E^{pq}_{a_\rho b_\rho}(H - E_0)E^{c_\tau d_\tau}_{pq}|\Psi_0\rangle C^\tau_{rs}t^\tau_{c_\tau d_\tau} \\ &= \sum_\tau\sum_{c_\tau d_\tau}\sum_{pqrs} C^\rho_{pq}\langle\Psi_0|E^{pq}_{a_\rho b_\rho}[H, E^{c_\tau d_\tau}_{pq}]|\Psi_0\rangle C^\tau_{rs}t^\tau_{c_\tau d_\tau} \end{aligned} \tag{39}$$

as we have also utilized in our previous works for developing an approximated FIC-MRCI variants with a DMRG reference function.[95,96] Note that both the CEPT2 and the CIPT2[85] methods use the energy functional of the similar form. The only difference between them consists in the choice of the CEPA subspaces; in the CEPT2 ansatz the external configurations are treated at the CEPA level while in the CIPT2 method the CEPA functional is used for in subspaces with three active MO indices i.e, $\{\Psi^\rho_i\}$ and $\{\Psi^a_\rho\}$.

The second term in Eq. (37) involves contractions of the 4-external exchange-type integrals with the 2-external amplitudes which is known as the most expensice terms in the canonical CC or CEPA implementations. However, such terms are typically not a bottleneck any more if the

virtual space is spanned by the PNOs. Some of the PNO-CEPA terms involve the semi-joint pair-pair interactions and in the MR case these terms appear to be more invlved. Apparently, in the single-reference limit, the PNO-CEPT2 energy and wave function converge to those of PNO-CEPA/0 with Doubles (PNO-CEPA/0(D)) or equivalently PNO-Linearlized Coupled-Cluster Doubles (PNO-LCCD) wave function.

All the CEPA terms in Eq. (37) are *linked* in the sense that the amplitude is either directly or indirectly via 1-body density cumulant connected to the integral. Therefore, as addressed in Ref. 106, even for a super system ($A \cdots B$) composed of infinitely separated two subsystems (A and B), no *unlinked* contractions appear. This indicates that the CEPT2 extension does not violate the near-size-consistent character of original CASPT2 model.

Unlike CIPT2 which is designed to mitigate the intruder state problems typically stemming from the $\{\Psi_i^\rho\}$, $\{\Psi_\rho^a\}$ and the reference subspaces, in CEPT2, we use the CASPT2 ansatz for those subspaces. In the single-reference case, it is known that the CEPA/0 energy can be irregular thus showing discontinuity in the potential energy surface. To remedy this drawback, Taube introduced a regularization scheme which is closely related to the imaginary shift in the CASPT2 framework.[107] Therefore, we have implemented the real- and semi-canonical imaginary shifts and the Hylleraas corrections to minimize the errors caused by such shifts.

Our current implementation of the PNO-CEPT2 is still not of optimal form since only the 4-external terms and the 1-body terms such as those similar to the contractions from $\mathscr{F}^{\text{ext}}_{\text{PNO-CASPT2}}$ are fully evaluated in the PNO space. All the remaining terms from $\mathscr{F}^{\text{ext}}_{\text{PNO-CEPA/0}}$, which involve either 4-internal integral or 2-external integral of semi-joint pair-pair interaction type,[51,52] are constructed in the canonical space using the back-transformed $\mathbf{t}$ amplitudes into the canonical space. Therefore, we only demonstrate and benchmark the accuracy of the PNO-CEPT2 model for small- to medium-size molecules and the benchmarking the efficiency is beyond the scope of this manuscript.

D. PNO-based Code Generation

The automatic derivation and code generation technique has been a quite powerful strategy for developing and testing the novel wave function methods which often involve hundreds or thousands of tensor contraction terms to implement.[90,97,98,106,108–127] To generate an efficient contraction code, each of the tensor contraction terms should be divided into a stream of binary contractions.

This is also the case for the code generations employing the PNOs spanning the virtual space. Even though the global optimization scheme proposed by Hanrath and Engels-Putzka for developing the efficient higher-order CC codes[128,129] can also also used in such a context, in order to keep the generator tool relatively simple yet effective, we use a more heuristic strategy for forming the optimal binary contractions.

In canonical MO basis, typical contraction terms in the CEPT2 model read

$$\begin{aligned} r^{\rho}_{ab} \leftarrow & \frac{1}{4}\sum_{\tau}\sum_{pqrs}\sum_{c}(f_{\mathrm{core}})_{ca}D^{pq}_{rs}t^{\tau}_{cb}C^{\tau}_{sr}C^{\rho}_{pq} \\ & - \frac{1}{4}\sum_{\tau}\sum_{pqrst}(f_{\mathrm{core}})_{ts}D^{pq}_{rs}t^{\tau}_{ba}C^{\tau}_{pq}C^{\rho}_{tr} \\ & + \frac{1}{4}\sum_{\tau}\sum_{pqrstu}\sum_{c}(uc|qa)D^{pqr}_{stu}t^{\tau}_{cb}C^{\rho}_{st}C^{\tau}_{pr} \\ & + \frac{1}{4}\sum_{\tau}\sum_{pqrstuv}(ur|tv)D^{pqr}_{stu}t^{\tau}_{ba}C^{\rho}_{pq}C^{\tau}_{vs} \\ & - \frac{1}{4}\sum_{\tau}\sum_{pqrstu}(st|ru)D^{pq}_{rs}t^{\tau}_{ba}C^{\rho}_{qp}C^{\tau}_{ut} \\ & + \frac{1}{4}\sum_{\tau}\sum_{pqrs}\sum_{cd}(ac|bd)D^{pq}_{rs}t^{\tau}_{cd}C^{\rho}_{sr}C^{\tau}_{qp} \\ & + (18\text{ more terms obtained by permuting pairs of indices}) \end{aligned} \tag{40}$$

for the $\{\Psi^{ab}_{\rho}\}$ subspace. Even though each term of Eq. (40) is composed of five tensor quantities resulting in XXX ways to form binary contractions, we find that the RDM is always contracted with two orthonormalization matrices (**C**). Moreover, if the MO integral involves at least one active MO indices, it is also contracted either with an RDM or with one **C** matrix. Therefore, after forming such amplitude-independent intermediates, Eq (40) becomes

$$\begin{aligned} r^{\rho}_{ab} \leftarrow & \frac{1}{4}\sum_{\tau}\sum_{c}(f_{\mathrm{core}})_{ca}(W_0)_{\rho\tau}t^{\tau}_{cb} - \frac{1}{4}\sum_{\tau}(fW_1)_{\rho\tau}t^{\tau}_{ba} + \frac{1}{4}\sum_{\tau}\sum_{c}(vW_2)_{\rho\tau,ca}t^{\tau}_{cb} \\ & + \frac{1}{4}\sum_{\tau}(W_3)_{\rho\tau}t^{\tau}_{ba} - \frac{1}{4}\sum_{\tau}\sum_{c}(W_4)_{\rho\tau}t^{\tau}_{ba} + \frac{1}{4}\sum_{\tau}\sum_{cd}(ac|bd)(W_5)_{\rho\tau}t^{\tau}_{cd} \\ & + (18\text{ more terms obtained by permuting pairs of indices}). \end{aligned} \tag{41}$$

By comparing the indices of the amplitudes, the generator can form an additional intermediate

$$(xW_6)_{\rho\tau} = (fW_1)_{\rho\tau} + (W_4)_{\rho\tau} \tag{42}$$

and the number of amplitude-dependent contractions are further reduced. In our current PNO-CASPT2/CEPT2 implementation, all the amplitude independent-intermediates such as (W_n), (fW_n), (vW_n) and (xW_n) are constructed in canonical MO basis and stored in fast core-memory before proceeding to the PCG iterations for solving the amplitude equations in PNO space. The amplitude-independent intermediates are surther subdivided into a stream of binary contractions and constructed using DGEMM subroutine assuming all the tensor quantities are stored in fast core-memory.

The CEPT2 residua (Eq. 41) with an additional intermediate (Eq. 42) are then transformed into PNO-based computer code assuming only the amplitudes and the residua are the native-PNO quantities. At each step of the PCG iteration, the amplitude-dependent intermediates involving a Fock matrix are constructed such as

$$(fvvt)^{\rho}_{a_\rho b_\rho} = \sum_{c_\rho} (f_{\rm core})_{a_\rho c_\rho} t^{\rho}_{c_\rho b_\rho} \tag{43}$$

$$(tfcc)^{\rho}_{a_{\rho i} b_{\rho i}} = \sum_{k} \sum_{a_{\rho k} b_{\rho k}} (f_{\rm core})_{ik} S_{a_{\rho i} a_{\rho k}} t^{\rho k}_{c_{\rho k} b_{\rho k}} S_{b_{\rho k} b_{\rho i}} \tag{44}$$

where **S** represent the pair-pair overlap matrices which are all stored on disk. As in case of Eq. (41), if the pair indices of the ampliudes and the residuum, which are τ and ρ, respectively, are different, the generator call a subroutine for projecting the amplitude from the τ-PNOs space to the ρ-PNO space. By this procedure, all the PNO-based tensor contractions which do not involve any two-electron integrals in Eq. (41) are transformed into

$$\begin{aligned} r^{\rho}_{a_\rho b_\rho} \leftarrow & \frac{1}{4} \sum_{\tau} (W_0)_{\rho\tau} \left[\sum_{a_\tau b_\tau} S_{a_\rho a_\tau} (fvvt)^{\tau}_{a_\tau b_\tau} S_{b_\tau b_\rho} \right] - \frac{1}{4} \sum_{\tau} (xW_6)_{\rho\tau} \left[\sum_{a_\tau b_\tau} S_{b_\rho b_\tau} t^{\tau}_{b_\tau a_\tau} S_{a_\tau a_\rho} \right] \\ & + \frac{1}{4} \sum_{\tau} (W_3)_{\rho\tau} \left[\sum_{a_\tau b_\tau} S_{b_\rho b_\tau} t^{\tau}_{ba} S_{a_\tau a_\rho} \right] \end{aligned} \tag{45}$$

and are constructed efficiently using DGEMM for projecting the PNOs from τ-pair to ρ-pair and DAXPY for summing up each contribution into the residua.

All the 4-external terms that appear in PNO-CEPT2 such as the sixth term in Eq. (41) have the similar structure

$$r^{\rho}_{a_\rho b_\rho} \leftarrow \frac{1}{4} \sum_{c_\rho d_\rho} (a_\rho c_\rho | b_\rho d_\rho) \left[\sum_{\tau} (W_5)_{\rho\tau} \left[\sum_{c_\tau d_\tau} S_{a_\rho a_\tau} t^{\tau}_{c_\tau d_\tau} S_{b_\tau b_\rho} \right] \right] \tag{46}$$

$$r^{\rho i}_{a_{\rho i} b_{\rho i}} \leftarrow 2 \sum_{c_{\rho i} d_{\rho i}} (a_{\rho i} c_{\rho i} | b_{\rho i} d_{\rho i}) \left[\sum_{\tau} (W_{124})_{\rho\tau} \left[\sum_{c_{\tau i} d_{\tau i}} S_{a_{\rho i} a_{\tau i}} t^{\tau i}_{c_{\tau i} d_{\tau i}} S_{b_{\tau i} b_{\rho i}} \right] \right] \tag{47}$$

and so we have written an optimized subroutine for efficiently calculate those terms in the PNO basis which are called by the generator.

For those subspaces which do not involve any PNOs and the coupling matrix elements of the canonical and the PNO subspaces, the PNO-based amplitudes are back-transformed into the canonical space and used to form the similar contractions.

III. RESULTS AND DISCUSSION

A. Non-Parallelity Errors for N_2 Dissociation

To assess the performance of canonical CEPT2 model and the validity of the PNO truncation, we calculated the dissociation curve of N_2 molecule in the singlet state using CEPT2 model with and without the PNO truncations in comparison with the FIC-NEVPT2 and CASPT2 results. The CEPT2 results are also compared with those obtained by higher-order approaches such as Davidson-corrected FIC-MRCI (MRCI+Q) and the Averaged-Coupled-Pair Functional (FIC-ACPF), and FIC-MR-CEPA/0. In all the CASPT2 and CEPT2 computations, no IPEA and level shifts were used. Errors from the FIC-MRCI+Q values are shown in Fig. 1. The non-parallelity errors (NPEs) are given in Table I. Orca 4.2 program package[55] was used for FIC-NEVPT2, FIC-MRCI,[98] FIC-CEPA/0 and FIC-ACPF[130] computations. In Table I, NPE by the CEPT2 model is shown to be smaller than that of the CASPT2 by about 3.8 kcal/mol. Even though the NEVPT2 shows the largest deviations from the reference FIC-MRCI+Q results and the dynamical correlation energies recovered are smaller than those by CASPT2 by more than 15 kcal/mol, the NPE in the NEVPT2 energies is smaller than those by CASPT2 and CEPT2 models. The FIC-MRCI-based and its approximately size-consistent variants produce quite similar results showing the NPE by up to about 2.0 kcal/mol with respect to the FIC-MRCI+Q results even though their computational costs are higher than any of CASPT2 and CEPT2 models.

In Table II, the dynamical correlation energies recovered by PNO-CEPT2 model are shown for N_2 molecule with various internuclear distances. To obtain approximately 99.9 % of the correlation energy, the PNO truncation threshold, which is coined `TCutPNO`, should be tightened by up to 1.0×10^{-8}. Surprisingly, the correlation energy recovered appears to be smallest at 1.0 Å which is close to the equilibrium internuclear distance where the static correlation effect is very minor in comparison with the dynamic counterpart. On the other hand, in the strongly-correlated region

where the triple bond is being broken, the PNO approximation performs relatively well recovering approximately 99.9 % of the canonical correlation energy even with TCutPNO= 5.0×10^{-8}.

B. Isomerization of $[Cu_2O_2]^{2+}(NH_3)_6$

We applied the PNO-CEPT2 model to isomerization of $[Cu_2O_2]^{2+}(NH_3)_6$ complex from the bis(μ-oxo) to the $\mu-\eta^2:\eta^2$-peroxo structures, which is a typical benchmark system for MR methods.[106,131,132] The O_2 activation mechanism in enzymes such as hemocyanin is known th be mediated by interconversion of $[Cu_2(\mu-\eta^2:\eta^2\text{-peroxo})]^{+2}$ and $[Cu_2(\mu\text{-oxo})]^{+2}$ cores. To accurately determine the isomerization pathways, numerous attempts have been made employing $[Cu_2O_2]^{2+}(NH_3)_6$ complex as an active-site model for O_2 activation reactions.

A particular difficulty of the isomerization of this complex consists in the fact that the CASPT2 and MRCI+Q results strongly disagree with each other: With CAS(8e, 10o), the CASPT2 predicts that the bis(μ-oxo) structure is lower in the energy while using the MRCI+Q and MR-ACPF models $\mu-\eta^2:\eta^2$-peroxo structure appears to be more stable. Recently, Köhn and coworkers[106] have performed a set of benchmark calculatons on this system using IC-MR Coupled-Cluster (IC-MRCC)[116,133] and the higher-order MRPT methods and found that all the higher-order methods than the second-order of perturbation expansion predicts that the $\mu-\eta^2:\eta^2$-peroxo structure is lower in the energy. This is actually consistent with the IC-MRCI+Q calculations published by Rode and Werner in 2005.[131] To produce the consistent energy ordering to the IC-MRCI+Q and IC-MRCC calculations using CASPT2-type wave function, one has to employ a larger active space involving all the 3d and the double-shell 4d orbitals of copper atoms thus leading to active space icomposed of 24 electrons in 28 orbitals as reported in RAS(24e, 28o)PT2 study by Gagliardi and coworkers[10] and cumulant-based DMRG-CASPT2 study by Yanai and coworkers[132] Therefore, it is interesting to see how much improvement over the PNO-CASPT2 is obtained by the PNO-CEPT2 model while keeping the active space CAS(8e, 10o).

Isomerization energy profiles calculated by PNO-CASPT2/CEPT2 methods are shown in Fig. 2 as a function of Cu-Cu distance. The bis(μ-oxo) and $\mu-\eta^2:\eta^2$-peroxo structures were taken from Ref. 90 and all the intermediates geometries were generated as a superposition of two isomeric geometries. The isomerization energies (E($\mu-\eta^2:\eta^2$-peroxo)-E(bis(μ-oxo))) calculated at CASSCF, PNO-CASPT2 and PNO-CEPT2 are -27.5 kcal/mol, +11.4 kcal/mol and +9.6 kcal/mol, respectively as can be seen in Table III. Difference between the previous CASPT2(8e, 10o) results

by Rode *et al.* and our PNO-CASPT2(8e, 10o) calculations with larger quadruple ζ basis sets is stemming from the difference in the basis sets, and PNO and weak-pair truncation errors. Still, the difference reaches only up to about 1.2 kcal/mol and is quite small in comparison with descrepancy between the MRCI+Q(8e, 10o) and the CASPT2(8e, 10o) results. The PNO-CEPA extension in the PNO-CEPT2 model corrects the PNO-CASPT2 results in the direction of MRCI+Q(8e, 10o) result by up to about 1.8 kcal/mol. In Table IV, the magnitude of the PNO-CEPT2 corrections to each subspace is analyzed. As can be seen, the CEPA extension to (c,c)→(v,v) subspace corrects the dynamical correlation energy by about up to -21.0 kcal/mol for both bis(μ-oxo) and $\mu-\eta^2:\eta^2$-peroxo states thus resulting in almost perfect cancellation in the isomerization energy. The 2-external subspaces with at least one active MO indices give rise to larger change in the isomerization energy by up to 1.52 kcal/mol. Even though the zeroth Hamiltonian is left unmodified, (c,c)→(a,v) semi-internal subspace shows a relatively large correction of about 0.45 kcal/mol in comparison with those from the other semi-internal and internal subspaces. However, the isomerization energy profile calculated by PNO-CEPT2 model is still similar to that by PNO-CASPT2 and the isomerization energy still appears to be far off from the IC-MRCI+Q and the IC-MRCC values. Therefore, at least in this case, the CEPA extension in the 2-external subspaces gives only a minor correction to the CASPT2 isomerization energy.

C. Ground and Excited States of Butadiene and Linear Polyenes

Predicting low-lying dark ($2^1A_g^-$) and bright ($1^1B_u^+$) excited states of butadiene has been known as notoriously challenging despite as reported in Ref. 134. To achieve qualitative accuracy in calculating the EE for the accessible $1^1B_u^+$ state, inclusion of the dynamic correlation plays a key role. On the other hand, the dipole-forbidden $2^1A_g^-$ state is dominated by doubly-excited character and is of strong MR nature even though there is no reliable experimental EEs reported for this state. Therefore, the typical single-reference excited state methods such as Equation-Of-Motion Coupled-Cluster with Singles and Doubles (EOM-CCSD) are not the appropriate choices for the $2^1A_g^-$ state.

In Table V, EEs for the $1^1B_u^+$ and $2^1A_g^-$ states calculated by canonical CASPT2 and CEPT2 models for butadiene are shown in comparison with various theoretical and experimental values. The reference CASSCF wave functions were calculated in a state specific (SS) manner to each of $1^1A_g^-$, $2^1A_g^-$ and $1^1B_u^+$. Geometries for butadiene and longer polyene chains were taken from

Ref. 132. In Ref. 134, Watson *et al.* have proposed the best theoretical EEs by using EOM-CC method including up to quadruple excitations (EOM-CCSDTQ). It is seen that the CC2, EOM-CCSD and EOM-CCSD(T) values largely deviate from the EOM-CCSDTQ values for the $2^1A_g^-$ state that possesses doubly-excited character while giving reasonably well agreement for $1^1B_u^+$. To obtain good agreement with the EOM-CCSDTQ results for both states in a single-reference framework, at least, CC3 level of theory should be used. The CASSCF values which lacks dynamic correlation effects show large deviation by up to 2.3 eV from the EOM-CCSDTQ result for $1^1B_u^+$. Addition of the dynamic correlation correction on top of the CASSCF energies significantly remedy the CASSCF results and the CASPT2 gives reasonably well agreement with the EOM-CCSDTQ values while in the CASPT2 results in Ref. 135 $1^1B_u^+$ and $2^1A_g^-$ are predicted to be almost degenerated. The EEs calculated by CEPT2 model are quite close to those by MR-AQCC results and in good agreement with the EOM-CCSDTQ results.

The $1^1B_u^+$ and $2^1A_g^-$ states of longer polyene chains were used as a benchmark set for the cumulant-approximated CASPT2[132] and NEVPT2[136] studies to assess the errors caused by the cumulant approximations in the 3- and 4-RDMs. In this study, we do not utilize the cumulant approximations in the RDMs. To assess the impact of CEPA/0 extensions in the PNO-CEPT2 model, we calculated the vertical EEs for $1^1B_u^+$ and $2^1A_g^-$ states of conjugated linear polyene molecules ($C_{2n}H_{2n+2}$) using CAS(2n*e*, 2n*o*). As SS-CASSCF calculations did not converge to C_6H_8 and longer polyenes, we performed state-average (SA-) CASSCF calculations involving the three lowest A_g and B_u states for obtaining the MOs for the $^1A_g^-$ and $1^1B_u^+$ states, respectively. The following CASPT2 and CEPT2 computations were performed in a SS manner to each state.

In Fig. 3, the EEs calculated at the canonical CASPT2 and PNO-CEPT2 levels are shown. As can be seen from Fig. 3, the EEs calculated by the PNO-CEPT2 model are always higher than those by the CASPT2 theory. This means that the PNO-CEPA/0 correction on the ground $1^1A_g^-$ state is even larger than those on the $1^1B_u^+$ and $2^1A_g^-$ states.

The doubly-excited $2^1A_g^-$ state is higher in energy than the bright $1^1B_u^+$ state by 0.06 eV for $C_{14}H_{16}$ and the crossover is not observed at PNO-CEPT2 level in Fig. 3. On the other hand, at the CASPT2 level, the $2^1A_g^-$ is calculated to be very slightly lower in the energy for $C_{14}H_{16}$ by up to 0.002 eV. By treating the external IC configurations at the CEPA/0 level in PNO-CEPT2 model for recovering more dynamic correlation effect, the valence $1^1B_u^+$ state is more lowered in the energy than the doubly-excited $2^1A_g^-$ state. This behavior of CASPT2 model is actually in contrast to previous studies such as Ref. 132 where the $2^1A_g^-$ state becomes lower in the energy

for C_8H_{10} with cc-pVDZ basis set. We found that use of larger def2-TZVPP basis set drastically stabilizes the $1^1B_u^+$ state in comparison with the $2^1A_g^-$ state of strong MR character thus pushing the crossover point to even longer polyenes. Moreover, the $1^1B_u^+$ and the higher $1^1B_u^-$ states are closely-lying in the energy at the CASSCF level for C_6H_8 thus showing a discontinuity. For a more thorough computation of the quasi-degenerated electronic states at the PNO-CEPT level, multi-state (MS)[137] or the extended multi-state (XMS)[138] extension is necessary.

D. Singlet-Triplet Gap of Free-base Porphyrin

We calculated the non-vertical singlet-triplet (S–T) energy gaps for free-base porphyrin molecule The geometries for S_0 and T_0 states of free-base porphyrin were taken from Ref. 95. In Table VI, the S–T gaps and the dynamic correlation energies recovered by various PNO cutoff thresholds are shown. It is seen that the PNO-CASPT2 model shows good agreement with the experimental value for all the PNO truncation thresholds even for quite loose threshold such as 1.0×10^{-6}, thus indicating that the error cancellation works surprisingly fine for this systems. This is most probably due to the fact that the S_0 and T_0 geometries are similar enough to each other. The PNO-CEPT2 model gives even closer values to the experimental one. The compression ratios of the first-order amplitudes, which are calculated as a ratio between total number of amplitudes in Eq. (15) and the canonical counterpart, are also shown. Even with the tightest threshold used, 1.0×10^{-9}, the PNO-CASPT2/CEPT2 models require only less than 6 % of the amplitudes while giving quite good agreement with the untruncated value. The PNO-CEPT2 values agree well also to the cumulant-approximated MRCI+Q and ACPF results that use the DMRG-CASSCF reference function with full-valence π active space. Note that none of computational values in Table VI include the solvent-effects.

In Table VII dynamic correlation contribution from each subspace is shown for both PNO-CASPT2 and PNO-CEPT2 models. Interestingly, unlike the isomerization of $[Cu_2O_2]^{2+}(NH_3)_6$ complex, the $(c,c)\rightarrow(v,v)$ subspace, which does not involve any active MO indices, gives dominant correction to the S–T gap by up to -2.98 kcal/mol (-0.15 eV). On the other hand, the other two types of the external subspaces give quite minor corrections that reach only to -0.15 kcal/mol in total.

IV. CONCLUSION

To obtain qualitative accuracy in predicting the reaction energies and various chemical energetics, a balanced treatment between static and dynamic electron correlation effects plays a key role. One of the most commonly used methods is the MR second-order perturbation theory such as CASPT2 and NEVPT2, both of which converge to the MP2 method in the closed-shell SR limit. As a simple yet higher order extension to the CASPT2 method, we have developed a hybrid of MR-CEPA/0 and CASPT2 Ansätze that converges to the CEPA/0(D) in the closed-shell SR limit, which is coined CEPT2 Ansatz. One of the largest bottleneck of the CEPT2 computation that is stemming from the MR-CEPA/0 terms in the residual equations is the presence of the so-called 4-EEOs involving a contraction of the 4-external two-electron integrals and the amplitude. To accelerate the CEPT2 computation, the CEPT2 equations are expanded in the local PNO basis which a compact set of virtual orbitals specifically optimized to each pair of localized orbitals and the nr-ICBs. Development of the PNO-CEPT2 computer programs was conducted using the automatic code generation technique and interfaced to our in-house quantum chemistry program package.

For N_2 dissociation, the CEPT2 model was shown to give a smaller NPE than that of CASPT2 model in comparison with the reference FIC-MRCI+Q results. The errors caused by the PNO truncation were examined and shown to be the largest at around the equilibrium internuclear distance.

The PNO-CEPT2 model was applied to calculating the isomerization energy profiles of $[Cu_2O_2]^{2+}(NH_3)_6$ complex from the bis(μ-oxo) to the $\mu - \eta^2 : \eta^2$-peroxo structures. The PNO-CEPT2 results show a similar trend to the CASPT2 as long as CAS($8e$, $10o$) is used and strongly disagree with the previous IC-MRCI+Q and IC-MRCC results. To remedy the accuracy of the PNO-CEPT2 model, the higher-order treatment of the subspaces with three active MO indices (E^{ip}_{qr} and E^{ra}_{pq}) will be necessary. However, simple use of the IC-MR-CEPA ansatz into those subspaces introduce a complicated tensor contraction involving the 4-RDMs and the amplitude and the PNOs cannot accelerate construction of such terms. Another possibility to improve the accuracy of the PNO-CEPT2 model is to use unconventionally large active space using the DMRG reference function. As all the terms involving the 4-RDMs are not more complicated than those in the CASPT2 mode, combination of the PNO-CEPT2 model and the DMRG reference function via use of the cumulant-approximated 4-RDM is quite straightforward. Nevertheless, previous studies have shown that simple use of cumulant-approximated RDM often introduces unavoidable

instability in the residual equation. Therefore, as a future development, we prefere to pursue other possibilities such as uncontracting the E^{ip}_{qr} and E^{ra}_{pq} subspaces using the Matrix-Product State (MPS) basis.

The CEPT2 was shown to give good agreement with the EOM-CCSDTQ results on calculating the excitation energies of butadiene achieving a good balance between dynamic and static electron correlations. For longer polyene systems, the PNO-CEPT2 produces the larger EEs than those by the CASPT2 model by up to 0.29 eV.

For calculating the non-vertical S–T gap of free-base porphyrin where the relaxation of the geometry in the triplet state is relatively small, the PNO and weak-pair truncations work surprisingly fine thus producing a S–T gap in a good agreemtns with the experimental value. Size of the PNO-CEPT2 amplitude is only up to less than 6 % of the canonical counterpart even with the PNO and the weak-pair truncation thresholds are set to 1.0×10^{-9} and 1.0×10^{-5} Eh, respectively.

We find that the IC-MR-CEPA extension to the CASPT2 model in te 2-external subspaces leads to the improvement in the accuracy. However, the PNO-CEPT2 model is not alwaysaccurate enough especially if the active space is too small. To improve the accuracy while keeping the active space as small as possible, higher order extension to E^{ip}_{qr} and E^{ra}_{pq} subspaces will be important. A production level implementation of the PNO-CEPT2 model is currently in progress.

SUPPLEMENTARY MATERIAL

ACKNOWLEDGMENTS

REFERENCES

[1]F. Neese, Coordination Chemistry Reviews **253**, 526 (2009), theory and Computing in Contemporary Coordination Chemistry.

[2]D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, Chemical Reviews **112**, 182 (2012), pMID: 22220988, https://doi.org/10.1021/cr2001417.

[3]P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, Chem. Rev. **112**, 108 (2012).

[4]J. P. Malrieu, R. Caballol, C. J. Calzado, C. de Graaf, and N. Guihry, Chemical Reviews **114**, 429 (2014), pMID: 24102410, https://doi.org/10.1021/cr300500z.

[5]B. O. Roos, P. R. Taylor, and P. E. Siegbahn, Chem. Phys. **48**, 157 (1980).

[6]B. O. Roos, Adv. Chem. Phys. **69**, 399 (1987).

[7]K. Ruedenberg, L. M. Cheung, and S. T. Elbert, Int. J. Quantum Chem. **16**, 1069 (1979).

[8]K. Kowalski and P. Piecuch, The Journal of Chemical Physics **113**, 18 (2000), https://doi.org/10.1063/1.481769.

[9]K. Kowalski and P. Piecuch, The Journal of Chemical Physics **113**, 5644 (2000), https://doi.org/10.1063/1.1290609.

[10]C. J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, and L. Gagliardi, The Journal of Physical Chemistry A **110**, 1991 (2006), pMID: 16451035, https://doi.org/10.1021/jp056791e.

[11]J. J. Eriksen and J. Gauss, Journal of Chemical Theory and Computation **0**, null (0), pMID: 31381327, https://doi.org/10.1021/acs.jctc.9b00456.

[12]Y. Qiu, T. M. Henderson, J. Zhao, and G. E. Scuseria, The Journal of Chemical Physics **147**, 064111 (2017), https://doi.org/10.1063/1.4991020.

[13]Y. Qiu, T. M. Henderson, J. Zhao, and G. E. Scuseria, The Journal of Chemical Physics **149**, 164108 (2018), https://doi.org/10.1063/1.5053605.

[14]J. A. Gomez, T. M. Henderson, and G. E. Scuseria, The Journal of Chemical Physics **150**, 144108 (2019), https://doi.org/10.1063/1.5085314.

[15]T. Tsuchimochi and S. Ten-no, Journal of Chemical Theory and Computation **12**, 1741 (2016), pMID: 26950651, https://doi.org/10.1021/acs.jctc.6b00137.

[16]T. Tsuchimochi and S. L. Ten-no, The Journal of Chemical Physics **149**, 044109 (2018), https://doi.org/10.1063/1.5036542.

[17]E. Xu and S. L. Ten-no, Journal of Computational Chemistry **39**, 875 (2018), https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.25163.

[18]K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, The Journal of Physical Chemistry **94**, 5483 (1990), https://doi.org/10.1021/j100377a012.

[19]K. Andersson, P. Malmqvist, and B. O. Roos, The Journal of Chemical Physics **96**, 1218 (1992), https://doi.org/10.1063/1.462209.

[20]M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, The Journal of Chemical Physics **128**, 134110 (2008), https://doi.org/10.1063/1.2889385.

[21]S. Vancoillie, H. Zhao, M. Radoń, and K. Pierloot, Journal of Chemical Theory and Computation **6**, 576 (2010), pMID: 26617311, https://doi.org/10.1021/ct900567c.

[22]Q. M. Phung and K. Pierloot, Journal of Chemical Theory and Computation **15**, 3033 (2019), pMID: 30995039, https://doi.org/10.1021/acs.jctc.9b00166.

[23]P.-Å. Malmqvist, K. Pierloot, A. Rehaman, M. Shahi, C. J. Cramer, and L. Gagliardi, J. Chem.

Phys. **128**, 204109 (2008).

[24]Q. M. Phung and K. Pierloot, Chemistry A European Journal **0**, 10.1002/chem.201902766, https://onlinelibrary.wiley.com/doi/pdf/10.1002/chem.201902766.

[25]B. O. Roos and K. Andersson, Chemical Physics Letters **245**, 215 (1995).

[26]N. Forsberg and P.-Å. Malmqvist, Chemical Physics Letters **274**, 196 (1997).

[27]C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J. P. Malrieu, J. Chem. Phys. **114**, 10252 (2001).

[28]C. Angeli, R. Cimiraglia, and J. Malrieu, J. Chem. Phys. **117**, 9138 (2002).

[29]C. Angeli, M. Pastore, and R. Cimiraglia, Theor. Chem. Acc. **117**, 743 (2006).

[30]K. G. Dyall, J. Chem. Phys. **102**, 4909 (1995).

[31]C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[32]F. Aquilante, T. K. Todorova, L. Gagliardi, T. B. Pedersen, and B. O. Roos, The Journal of Chemical Physics **131**, 034113 (2009), https://doi.org/10.1063/1.3157463.

[33]J. Segarra-Mart, M. Garavelli, and F. Aquilante, Journal of Chemical Theory and Computation **11**, 3772 (2015), pMID: 26574459, https://doi.org/10.1021/acs.jctc.5b00479.

[34]F. Aquilante, T. B. Pedersen, R. Lindh, B. O. Roos, A. Snchez de Mers, and H. Koch, The Journal of Chemical Physics **129**, 024113 (2008), https://doi.org/10.1063/1.2953696.

[35]J. Boström, F. Aquilante, T. B. Pedersen, and R. Lindh, Journal of Chemical Theory and Computation **5**, 1545 (2009), pMID: 26609847, https://doi.org/10.1021/ct9000284.

[36]F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh, and B. O. Roos, Journal of Chemical Theory and Computation **4**, 694 (2008), pMID: 26621084, https://doi.org/10.1021/ct700263h.

[37]J. Boström, M. G. Delcey, F. Aquilante, L. Serrano-Andrs, T. B. Pedersen, and R. Lindh, Journal of Chemical Theory and Computation **6**, 747 (2010), pMID: 26613305, https://doi.org/10.1021/ct900612k.

[38]C. Song and T. J. Martnez, The Journal of Chemical Physics **149**, 044108 (2018), https://doi.org/10.1063/1.5037283.

[39]W. Győffy, T. Shiozaki, G. Knizia, and H.-J. Werner, J. Chem. Phys. **138**, 104104 (2013).

[40]F. Menezes, D. Kats, and H.-J. Werner, J. Chem. Phys. **145**, 124115 (2016), http://dx.doi.org/10.1063/1.4963019.

[41]D. Kats and H.-J. Werner, The Journal of Chemical Physics **150**, 214107 (2019), https://doi.org/10.1063/1.5097644.

[42]W. Kutzelnigg, J. Chem. Phys. **40**, 3640 (1964).

[43]R. Ahlrichs and W. Kutzelnigg, J. Chem. Phys. **48**, 1819 (1968).

[44]C. Edmiston and M. Krauss, J. Chem. Phys. **42**, 1119 (1965).

[45]C. Edmiston and M. Krauss, J. Chem. Phys. **45**, 1833 (1966).

[46]W. Meyer, Int. J. Quantum Chem. **5**, 341 (1971).

[47]W. Meyer, J. Chem. Phys. **58**, 1017 (1973).

[48]R. Ahlrichs and W. Kutzelnigg, Theor. Chim. Acta. **10**, 377 (1968).

[49]P. R. Taylor, J. Chem. Phys. **74**, 1256 (1981).

[50]R. Fink and V. Staemmler, Theor. Chim. Acta. **87**, 129 (1993).

[51]F. Neese, F. Wennmohs, and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).

[52]F. Neese, A. Hansen, and D. G. Liakos, J. Chem. Phys. **131**, 064103 (2009).

[53]A. Hansen, D. G. Liakos, and F. Neese, J. Chem. Phys. **135**, 214102 (2011).

[54]D. G. Liakos, A. Hansen, and F. Neese, Journal of Chemical Theory and Computation **7**, 76 (2011), pMID: 26606220, https://doi.org/10.1021/ct100445s.

[55]F. Neese, Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1327 (2018), e1327.

[56]C. Riplinger and F. Neese, J. Chem. Phys. **138**, 034106 (2013).

[57]C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, J. Chem. Phys. **139**, 134101 (2013).

[58]P. Pinski, C. Riplinger, E. F. Valeev, and F. Neese, J. Chem. Phys. **143**, 034108 (2015).

[59]C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, and F. Neese, J. Chem. Phys. **144**, 024109 (2016).

[60]M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, and F. Neese, The Journal of Chemical Physics **146**, 164105 (2017), https://doi.org/10.1063/1.4981521.

[61]D. Datta, S. Kossmann, and F. Neese, J. Chem. Phys. **145**, 114101 (2016).

[62]O. Demel, J. Pittner, and F. Neese, J. Chem. Theory Comput. **11**, 3104 (2015), pMID: 26575747, http://dx.doi.org/10.1021/acs.jctc.5b00334.

[63]J. Brabec, J. Lang, M. Saitow, J. Pittner, F. Neese, and O. Demel, Journal of Chemical Theory and Computation **14**, 1370 (2018), pMID: 29345924, https://doi.org/10.1021/acs.jctc.7b01184.

[64]J. Lang, J. Brabec, M. Saitow, J. Pittner, F. Neese, and O. Demel, Phys. Chem. Chem. Phys. **21**, 5022 (2019).

[65]Y. Guo, K. Sivalingam, E. F. Valeev, and F. Neese, J. Chem. Phys. **144**, 094111 (2016), http://dx.doi.org/10.1063/1.4942769.

[66]H.-J. Werner, G. Knizia, C. Krause, M. Schwilk, and M. Dornbach, J. Chem. Theory Comput.

11, 484 (2015), pMID: 26580908, http://dx.doi.org/10.1021/ct500725e.

[67]M. Schwilk, D. Usvyat, and H.-J. Werner, J. Chem. Phys. **142**, 121102 (2015).

[68]M. Schwilk, Q. Ma, C. Köppl, and H.-J. Werner, Journal of Chemical Theory and Computation **13**, 3650 (2017), pMID: 28661673, http://dx.doi.org/10.1021/acs.jctc.7b00554.

[69]H. J. Werner, J. Chem. Phys. **145** (2016), 10.1063/1.4968595.

[70]Q. Ma and H. J. Werner, J. Chem. Theory Comput. **14**, 198215 (2018).

[71]C. Krause and H.-J. Werner, J. Chem. Theory Comput. (2018), 10.1021/acs.jctc.8b01012.

[72]Q. Ma and H.-J. Werner, J. Chem. Theory Comput. (2019), 10.1021/acs.jctc.8b01098.

[73]G. Schmitz, B. Helmich, and C. Hättig, Molecular Physics **111**, 2463 (2013), https://doi.org/10.1080/00268976.2013.794314.

[74]M. S. Frank, G. Schmitz, and C. Hättig, Molecular Physics **115**, 343 (2017), https://doi.org/10.1080/00268976.2016.1263762.

[75]G. Schmitz and C. Hättig, The Journal of Chemical Physics **145**, 234107 (2016), http://aip.scitation.org/doi/pdf/10.1063/1.4972001.

[76]D. P. Tew and C. Hättig, International Journal of Quantum Chemistry **113**, 224 (2013).

[77]G. Schmitz, C. Hättig, and D. P. Tew, Phys. Chem. Chem. Phys. **16**, 22167 (2014).

[78]D. P. Tew, B. Helmich, and C. Hättig, J. Chem. Phys. **135**, 074107 (2011), http://dx.doi.org/10.1063/1.3624370.

[79]B. Helmich and C. Hättig, J. Chem. Phys. **135**, 214106 (2011), http://dx.doi.org/10.1063/1.3664902.

[80]C. Hättig, D. P. Tew, and B. Helmich, J. Chem. Phys. **136**, 204105 (2012), http://dx.doi.org/10.1063/1.4719981.

[81]B. Helmich and C. Hättig, J. Chem. Phys. **139**, 084114 (2013), http://dx.doi.org/10.1063/1.4819071.

[82]B. O. Roos, P. Linse, P. E. Siegbahn, and M. R. Blomberg, Chemical Physics **66**, 197 (1982).

[83]P. Pulay, International Journal of Quantum Chemistry **111**, 3273 (2011), https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.23052.

[84]A. Banerjee and J. Simons, Int. J. Quantum Chem. **19**, 207 (1981).

[85]P. Celani, H. Stoll, H. Werner, and P. Knowles, Mol. Phys. **102**, 2369 (2004).

[86]R. F. Fink, Chemical Physics Letters **428**, 461 (2006).

[87]R. F. Fink, Chemical Physics **356**, 39 (2009), moving Frontiers in Quantum Chemistry:.

[88]S. Sharma, G. Knizia, S. Guo, and A. Alavi, Journal of Chemical Theory and Computation **13**,

488 (2017), pMID: 28060507, https://doi.org/10.1021/acs.jctc.6b00898.

[89]P. Celani and H.-J. Werner, J. Chem. Phys. **112**, 5546 (2000).

[90]K. R. Shamasundar, G. Knizia, and H.-J. Werner, J. Chem. Phys. **135**, 054101 (2011).

[91]This indicate the $(a, a)\rightarrow(a, v)$ and $(a, a)\rightarrow(v, v)$ configurations where a and v represent the active and virtual MOs.

[92]H. Werner and E. Reinsch, The Journal of Chemical Physics **76**, 3144 (1982), https://doi.org/10.1063/1.443357.

[93]H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).

[94]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, WIREs. Comput. Mol. Sci. **2**, 242 (2012).

[95]M. Saitow, Y. Kurashige, and T. Yanai, J. Chem. Phys. **139**, 044118 (2013).

[96]M. Saitow, Y. Kurashige, and T. Yanai, Journal of Chemical Theory and Computation **11**, 5120 (2015), pMID: 26574310, https://doi.org/10.1021/acs.jctc.5b00270.

[97]M. K. MacLeod and T. Shiozaki, The Journal of Chemical Physics **142**, 051103 (2015), https://doi.org/10.1063/1.4907717.

[98]K. Sivalingam, M. Krupička, A. A. Auer, and F. Neese, The Journal of Chemical Physics **145**, 054104 (2016), https://doi.org/10.1063/1.4959029.

[99]J. Planelles, C. Valdemoro, and J. Karwowski, Phys. Rev. A **43**, 3392 (1991).

[100]J. Planelles, C. Valdemoro, and J. Karwowski, Phys. Rev. A **41**, 2391 (1990).

[101]W. Kutzelnigg, The Journal of Chemical Physics **82**, 4166 (1985), https://doi.org/10.1063/1.448859.

[102]W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **107**, 432 (1997).

[103]M. Saitow and F. Neese, The Journal of Chemical Physics **149**, 034104 (2018), https://doi.org/10.1063/1.5027114.

[104]Q. Ma and H.-J. Werner, J. Chem. Theory Comput. **11**, 5291 (2015), pMID: 26574323, http://dx.doi.org/10.1021/acs.jctc.5b00843.

[105]FEMTO :: An Integrated Toolset For The Automated Tensor Generation, Version 0.1.0, M. Saitow (https://github.com/msaitow), Accessed September 17, 2015.

[106]Y. A. Aoto, A. Bargholz, D. Kats, H.-J. Werner, and A. Köhn, J. Chem. Theory Comput. **15**, 2291 (2019).

[107]A. G. Taube and R. J. Bartlett, The Journal of Chemical Physics **130**, 144112 (2009), https://doi.org/10.1063/1.3115467.

[108]C. L. Janssen and H. F. Schaefer, Theor. Chim. Acta. **79**, 1 (1991).

[109]S. Hirata, J. Phys. Chem. A **107**, 9887 (2003), http://dx.doi.org/10.1021/jp034596z.

[110]S. Hirata, J. Chem. Phys. **121**, 51 (2004).

[111]S. Hirata, Theor. Chem. Acc. **116**, 2 (2006).

[112]A. A. Auer, G. Baumgartner, D. E. Bernholdt, A. Bibireata, V. Choppella, D. Cociorva, X. Gao, R. Harrison, S. Krishnamoorthy, S. Krishnan, C.-C. Lam, Q. Lu, M. Nooijen, R. Pitzer, J. Ramanujam, P. Sadayappan, and A. Sibiryakov, Molecular Physics **104**, 211 (2006), http://dx.doi.org/10.1080/00268970500275780.

[113]A. Köhn, J. Chem. Phys. **130**, 104104 (2009).

[114]A. Köhn, J. Chem. Phys. **130**, 131101 (2009).

[115]M. Hanauer and A. Köhn, J. Chem. Phys. **131**, 124118 (2009).

[116]M. Hanauer and A. Köhn, J. Chem. Phys. **134**, 20411 (2011).

[117]J. A. Black and A. Köhn, J. Chem. Phys. **150**, 194107 (2019).

[118]T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, J. Chem. Phys. **129**, 071101 (2008).

[119]T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, Phys. Chem. Chem. Phys. **10**, 3358 (2008).

[120]T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, J. Chem. Phys. **130**, 054101 (2009).

[121]T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, J. Chem. Phys. **131**, 044118 (2009).

[122]M. Krupička, K. Sivalingam, L. Huntington, A. A. Auer, and F. Neese, Journal of Computational Chemistry **38**, 1853 (2017), https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.24833.

[123]L. M. J. Huntington, M. Krupička, F. Neese, and R. Izsák, The Journal of Chemical Physics **147**, 174104 (2017), https://doi.org/10.1063/1.5001320.

[124]A. Sen, B. de Souza, L. M. J. Huntington, M. Krupička, F. Neese, and R. Izsák, The Journal of Chemical Physics **149**, 114108 (2018), https://doi.org/10.1063/1.5048688.

[125]D. Datta, L. Kong, and M. Nooijen, J. Chem. Phys. **134**, 214116 (2011).

[126]L. Kong, K. R. Shamasundar, O. Demel, and M. Nooijen, The Journal of Chemical Physics **130**, 114101 (2009), https://doi.org/10.1063/1.3089302.

[127]L. Kong and E. F. Valeev, The Journal of Chemical Physics **135**, 214105 (2011), https://doi.org/10.1063/1.3664729.

[128]M. Hanrath and A. Engels-Putzka, J. Chem. Phys. **133**, 064108 (2010).

[129]A. Engels-Putzka and M. Hanrath, J. Chem. Phys. **134**, 124106 (2011).

[130]R. Gdanitz and R. Ahlrichs, Chem. Phys. Lett. **143**, 413 (1988).

[131]M. F. Rode and H.-J. Werner, Theoretical Chemistry Accounts **114**, 309 (2005).

[132]Y. Kurashige, J. Chalupský, T. N. Lan, and T. Yanai, The Journal of Chemical Physics **141**, 174111 (2014), https://doi.org/10.1063/1.4900878.

[133]M. Hanauer and A. Köhn, J. Chem. Phys. **137**, 131103 (2012).

[134]M. A. Watson and G. K.-L. Chan, Journal of Chemical Theory and Computation **8**, 4013 (2012), pMID: 26605568, https://doi.org/10.1021/ct300591z.

[135]L. SerranoAndrés, M. Merchán, I. NebotGil, R. Lindh, and B. O. Roos, The Journal of Chemical Physics **98**, 3151 (1993), https://doi.org/10.1063/1.465071.

[136]D. Zgid, D. Ghosh, E. Neuscamman, and G. K.-L. Chan, J. Chem. Phys. **130**, 194107 (2009).

[137]J. Finley, P. ke Malmqvist, B. O. Roos, and L. Serrano-Andrs, Chemical Physics Letters **288**, 299 (1998).

[138]T. Shiozaki, W. Győffy, P. Celani, and H.-J. Werner, J. Chem. Phys. **135**, 081106 (2011).

[139]F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).

[140]F. Weigend, Phys. Chem. Chem. Phys. **8**, 1057 (2006).

[141]J. D. Watts, S. R. Gwaltney, and R. J. Bartlett, The Journal of Chemical Physics **105**, 6979 (1996), https://doi.org/10.1063/1.471988.

[142]O. Lehtonen, D. Sundholm, R. Send, and M. P. Johansson, The Journal of Chemical Physics **131**, 024301 (2009), https://doi.org/10.1063/1.3158990.

[143]P. Szalay, A. Karpfen, and H. Lischka, Chemical Physics **130**, 219 (1989).

[144]M. Dallos and H. Lischka, Theoretical Chemistry Accounts **112**, 16 (2004).

[145]J. P. Doering and R. McDiarmid, The Journal of Chemical Physics **73**, 3617 (1980), https://doi.org/10.1063/1.440587.

[146]L. Serrano-Andrés, M. Merchán, M. Rubio, and B. O. Roos, Chem. Phys. Lett. **295**, 195 (1998).

[147]A. Aspuru-Guzik, O. E. Akramine, J. C. Grossman, and W. A. Lester, J. Chem. Phys. **120**, 3049 (2004).

[148]M. Gouterman and G.-E. Khalil, J. Mol. Spectrosc. **53**, 88 (1974).

TABLE I: The NPEs in kcal/mol for N_2 dissociation by various FIC methods with respect to the FIC-MRCI+Q energies.

	CASPT2	CEPT2	FIC-NEVPT2	FIC-MRCI	FIC-CEPA/0	FIC-ACPF
NPEs / kcal/mol	9.98	6.21	2.99	2.03	0.98	0.68

TABLE II: Dynamic correlation energies recovered (%) by PNO-CEPT2 model for N_2 with various PNO truncation thresholds with def2-QZVPP[139,140] basis set. No pair-energy-based screening of weak-pairs were performed. 1*s* orbitals of nitrogen atoms were kept frozen while the 2*p* orbitals were included in the active space (CAS(6*e*, 6*o*)).

	TCutPNO											
R (N-N) / Å	1.0×10^{-11}	1.0×10^{-10}	1.0×10^{-9}	5.0×10^{-9}	1.0×10^{-8}	5.0×10^{-8}	1.0×10^{-7}	5.0×10^{-7}	1.0×10^{-6}	5.0×10^{-6}	1.0×10^{-5}	1.0×10^{-4}
1.0	99.95	99.95	99.93	99.93	99.88	99.83	99.79	99.68	99.64	99.37	99.15	96.58
1.1	99.95	99.95	99.94	99.92	99.90	99.84	99.80	99.72	99.67	99.52	99.30	96.37
1.2	99.95	99.95	99.94	99.91	99.90	99.86	99.81	99.73	99.71	99.62	99.50	96.13
1.3	99.95	99.95	99.94	99.92	99.91	99.87	99.83	99.75	99.73	99.65	99.60	96.13
1.4	99.95	99.95	99.94	99.93	99.92	99.88	99.85	99.80	99.75	99.69	99.71	96.79
1.6	99.96	99.96	99.95	99.94	99.93	99.90	99.89	99.86	99.81	99.86	99.85	97.74
1.8	99.97	99.97	99.96	99.95	99.95	99.93	99.92	99.91	99.86	100.02	100.06	98.52
2.0	99.98	99.98	99.97	99.96	99.96	99.96	99.95	99.95	99.87	100.09	100.18	99.18
2.2	99.98	99.98	99.98	99.98	99.97	99.97	99.96	99.96	99.86	100.12	100.16	99.76
2.4	99.99	99.98	99.98	99.98	99.98	99.97	99.97	99.96	99.85	100.14	100.10	99.94
2.6	99.99	99.99	99.98	99.98	99.98	99.97	99.97	99.96	99.85	100.15	100.15	100.19
2.8	99.99	99.99	99.98	99.98	99.98	99.97	99.97	99.96	99.85	100.13	100.16	100.34
3.0	99.99	99.99	99.99	99.98	99.98	99.97	99.96	99.96	99.85	100.13	100.18	100.43

TABLE III: Isomerization energies calculated by size-extensive version of the completely-renormalized CCSD with perturbative triples (CR-CCSD(T)$_L$) and various MR methods. For the PNO-CASPT2/CEPT2 calculations, the PNO and weak-pair truncation thresholds were set to 5.0×10^{-8} and 1.0×10^{-5} Eh, respectively and the relativistic effects were not included.

Method	E($\mu - \eta^2 : \eta^2$-peroxo)-E(bis(μ-oxo)) / kcal/mol
CR-CCSD(T)$_{\mathrm{L}}$ [a,b]	-10.1
FIC-MRCCSD(8e,10o)[a,b]	-9.4
IC-MRCI+Q(8e,10o)[a,b]	-12.4
DMRG-cu(4)-CASPT2(24e, 28o)[c]	-4.9
DMRG-CASSCF(24e, 28o)[c]	-21.3
RASPT2(24e,28o)[b,d]	-7.6
CASSCF(16e,14o)[a,b]	-29.8
CASPT2(16e,14o)[a,b]	+17.2
CASSCF(8e,10o)[b,e]	-23.4
CASPT2(8e,10o)[b,e]	+10.2
IC-MRCI+Q(8e,10o)[b,e]	-11.6
PNO-CASPT2(8e,10o) (This work)	+11.4
PNO-CEPT2(8e,10o) (This work)	+9.6

[a] Taken from Ref. 106

[b] Relativistic effects were included by the effective core potentials (ECPs) on copper atoms

[c] Taken from Ref. 132

[d] Taken from Ref. 10

[d] Taken from Ref. 131

TABLE IV: Contributions from each subspace in the PNO-CASPT2 and PNO-CEPT2 dynamical correlation energies using `TCutPairs`$=1.0\times10^{-5}$ Eh and `TCutPNO`$=5.0\times10^{-8}$ for $[Cu_2O_2]^{2+}(NH_3)_6$ complex. Magnitude of PNO-CEPT2 correction (δE_{CEPA}) in each subspace is compared with that of PNO-CASPT2. In addition, difference between PNO-CEPT2 corrections for two geometries, $\delta E_{CEPA}(R_{Cu\text{-}Cu}=1.6$ Å$)$-$\delta E_{CEPA}(R_{Cu\text{-}Cu}=2.2$ Å$)$, are also shown as Δ-Correction.

	$R_{Cu\text{-}Cu}$=1.6 Å			$R_{Cu\text{-}Cu}$=2.2 Å			
Subspace	PNO-CASPT2 / Eh	PNO-CEPT2 / Eh	δE_{CEPA} / kcal/mol	CASPT2 / Eh	CEPT2 / Eh	δE_{CEPA} / kcal/mol	Δ-Correction / kcal/mol
			External				
$(c,c)\rightarrow(v,v)$	-2.641 691	-2.675 058	(-20.94)	-2.632 592	-2.666 005	(-20.97)	(0.03)
$(c,a)\rightarrow(v,v)$	-0.320 316	-0.303 014	(10.86)	-0.331 461	-0.312 859	(11.67)	(-0.82)
$(a,a)\rightarrow(v,v)$	-0.051 163	-0.052 895	(-1.09)	-0.053 697	-0.054 310	(-0.39)	(-0.70)
Sum							(-1.49)
			Semi-internal				
$(c,c)\rightarrow(a,v)$	-0.132 298	-0.130 746	(0.97)	-0.154 270	-0.152 005	(1.42)	(-0.45)
$(c,a)\rightarrow(a,v)$, $(c,a)\rightarrow(v,a)$	-0.206 663	-0.207 394	(-0.46)	-0.238 979	-0.239 850	(-0.55)	(0.09)
$(a,a)\rightarrow(a,v)$	-0.015 565	-0.015 542	(0.01)	-0.013 961	-0.013 962	(0.00)	(0.02)
Sum							(-0.34)
			Internal				
$(c,c)\rightarrow(a,a)$	-0.006 854	-0.006 793	(0.04)	-0.009 053	-0.008 963	(0.06)	(-0.02)
$(c,a)\rightarrow(a,a)$	-0.006 606	-0.006 607	(0.00)	-0.009 019	-0.009 018	(0.00)	(0.00)
Sum							(-0.02)
Total sum							(-1.85)

TABLE V: Computed EEs for $1^1B_u^+$ and $2^1A_g^-$ states at canonical CASPT2 and CEPT2 model in comparison with various theoretical and experimental data. The def2-TZVPP basis and def2-TZVPP/JK auxiliary basis sets were used. The C($1s$) orbitals were frozen during CASPT2 and CEPT2 computations while all the C($2p_z$) orbitals were included in the active space leading to CAS(4e, 4o) treaement. No shifts were used in the zeroth-order Hamiltonian.

Method	$1^1B_u^+$	$2^1A_g^-$
EOM-CCSD[141]	6.42	7.23
EOM-CCSD(T)[141]	6.13–6.36	6.76–6.92
CC2[142]	6.15	7.04
CC3[142]	6.19	6.58
CASSCF[135]	8.54	6.64
CASPT2[135]	6.23	6.27
MRCI[143]	6.70	6.78
MR-AQCC[144]	6.18	6.55
SS-CASSCF (This work)	8.18	6.72
CASPT2 (This work)	6.00	6.29
CEPT2 (This work)	6.12	6.50
EOM-CCSDTQ[134]	6.21	6.41
Exp.[145]	5.96–6.05	–

TABLE VI: Dynamic electron correlation energies and S–T gaps for free-base porphyrin calculated at PNO-CASPT2/CEPT2 level with various PNO cutoff thresholds in comparison with experimental S–T gap. The def2-SVP basis set was used. The threshold for weak-pair screening was set to 1×10^{-5} Eh. All the $1s$ orbitals of carbon atoms were kept frozen. 8 π MOs composed of $2p_z$ orbitals of carbon atoms were included in the active space thus leading to CAS(8e, 8o) settings which is consistent with CAS treatment of Ref. 146. Ratios between total number of PNO-CASPT2/CEPT2 amplitudes to optimize and that of canonical CASPT2, both of which include internal and semi-internal amplitudes, are also shown. Total number of AO basis functions is 406.

TCutPNO	Total Energy (S_0) / Eh	Correlation Energy / Eh	Accuracy (%)	Compression Ratio (%)	Total Energy (T_0) / Eh	Correlation Energy / Eh	Accuracy (%)	Compression Ratio (%)	S-T gap / eV
				PNO-CASPT2					
Full	-985.795110	-3.233234	(100.00)	(100.00)	-985.733659	-3.219267	(100.00)	(100.00)	1.67
1.0×10^{-9}	-985.791357	-3.229481	(99.88)	(5.83)	-985.729366	-3.214973	(99.87)	(4.69)	1.69
1.0×10^{-8}	-985.790258	-3.228382	(99.85)	(3.02)	-985.728332	-3.213940	(99.83)	(2.56)	1.69
5.0×10^{-8}	-985.788902	-3.227025	(99.81)	(1.90)	-985.726937	-3.212544	(99.79)	(1.68)	1.69
1.0×10^{-7}	-985.787947	-3.226070	(99.78)	(1.57)	-985.725979	-3.211586	(99.76)	(1.41)	1.69
1.0×10^{-6}	-985.781456	-3.219580	(99.58)	(0.94)	-985.719703	-3.205310	(99.57)	(0.90)	1.68
				PNO-CEPT2					
1.0×10^{-9}	-985.982732	-3.420855	-	(5.83)	-985.925605	-3.411212	-	(4.69)	1.55
1.0×10^{-8}	-985.981451	-3.419574	-	(3.02)	-985.924399	-3.410006	-	(2.56)	1.55
5.0×10^{-8}	-985.979903	-3.418026	-	(1.90)	-985.922854	-3.408461	-	(1.68)	1.55
1.0×10^{-7}	-985.978852	-3.416975	-	(1.57)	-985.921792	-3.407399	-	(1.41)	1.55
1.0×10^{-6}	-985.973026	-3.411150	-	(0.94)	-985.916193	-3.401800	-	(0.90)	1.55
				Other Methods					
DMRG-CASSCF(26e,24o)[95]									1.26
DMRG-cu(4)-MRCI(26e,24o)[95]									1.41
DMRG-cu(4)-MRCI+Q(26e,24o)[95]									1.48
DMRG-cu(4)-ACPF(26e,24o)[95]									1.55
Diffusion Monte-Carlo (DMC)[147]									1.60
Exp.[148]									1.58

TABLE VII: Contribution from each subspace in the PNO-CASPT2 and PNO-CEPT2 dynamical correlation energies using `TCutPairs`=1.0×10^{-5} Eh and `TCutPNO`=5.0×10^{-8} for free-base porphyrin in singlet and triplet states. Magnitude of PNO-CEPT2 correction (δE_{CEPA}) in each subspace is compared with that of PNO-CASPT2. In addition, difference between PNO-CEPT2 corrections for two geometries, $\delta E_{CEPA}(T_0)$-$\delta E_{CEPA}(S_0)$), are also shown as Δ-Correction.

	S_0 State			T_0 State			
Subspace	PNO-CASPT2 / Eh	PNO-CEPT2 / Eh	δE_{CEPA} / kcal/mol	CASPT2 / Eh	CEPT2 / Eh	δE_{CEPA} / kcal/mol	Δ-Correction / kcal/mol
			External				
$(c,c)\rightarrow(v,v)$	-2.38 0046	-2.571 048	(-119.86)	-2.416 554	-2.612 158	(-122.74)	(-2.89)
$(c,a)\rightarrow(v,v)$	-0.401 660	-0.401 882	(-0.14)	-0.377 783	-0.377 975	(-0.12)	(0.02)
$(a,a)\rightarrow(v,v)$	-0.027 777	-0.027 339	(0.27)	-0.022 754	-0.022 585	(0.11)	(-0.17)
Sum							(-3.04)
			Semi-internal				
$(c,c)\rightarrow(a,v)$	-0.223 743	-0.223 854	(-0.07)	-0.208 821	-0.209 014	(-0.12)	(-0.05)
$(c,a)\rightarrow(a,v)$, $(c,a)\rightarrow(v,a)$	-0.166 721	-0.166 812	(-0.06)	-0.155 565	-0.155 614	(-0.03)	(0.03)
$(a,a)\rightarrow(a,v)$	-0.008 111	-0.008 122	(-0.01)	-0.012 407	-0.012 445	(-0.02)	(-0.02)
Sum							(-0.04)
			Internal				
$(c,c)\rightarrow(a,a)$	-0.014 448	-0.014 450	(0.00)	-0.009 582	-0.009 587	(0.00)	(0.00)
$(c,a)\rightarrow(a,a)$	-0.004 442	-0.004 442	(0.00)	-0.009 023	-0.009 026	(0.00)	(0.00)
Sum							(0.00)
Total sum							(-3.08)

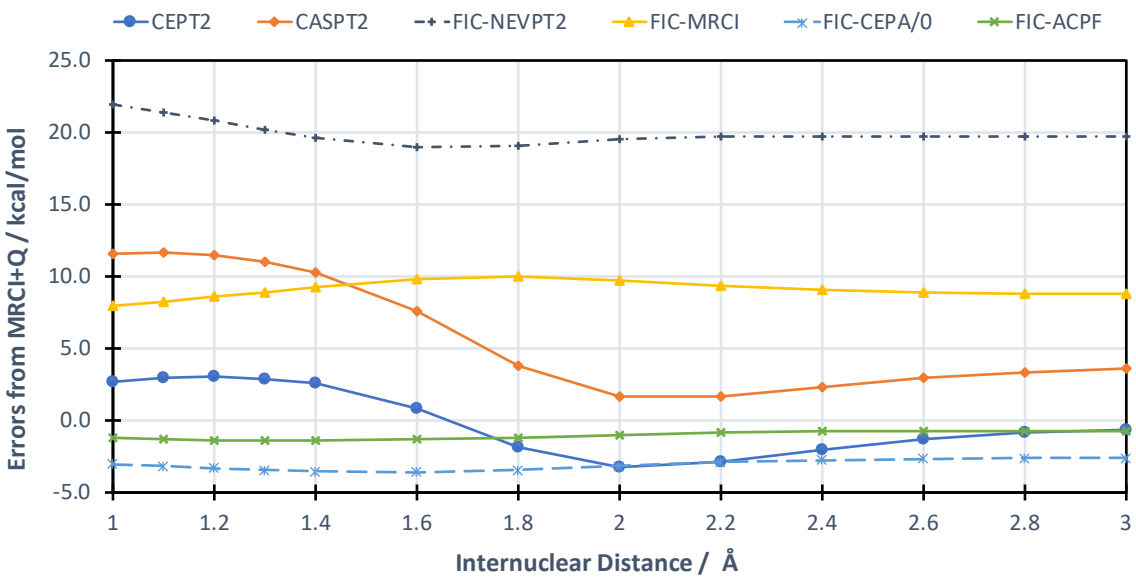

FIG. 1. Deviations of canonical CEPT2 energy from those of MRCI+Q for N_2 with various internuclear distances with def2-QZVPP basis set. 1*s* orbitals of nitrogen atoms were kept frozen while all the other orbitals were correlated. All the 2*p* orbitals were included in the active space leading to CAS(6*e*, 6*o*) treatment.

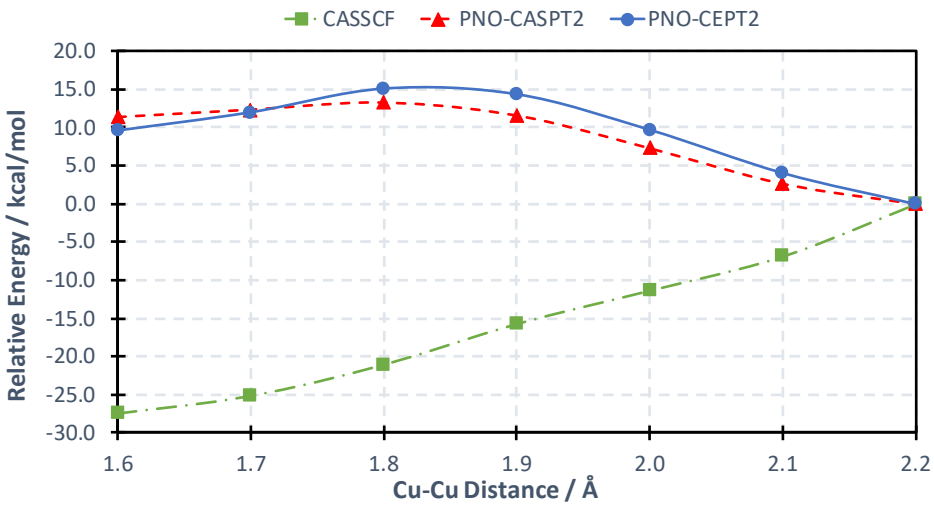

FIG. 2. Isomerization curves of $[Cu_2O_2]^{+2}(NH_3)_6$ calculated by PNO-CASPT2 and PNO-CEPT2 using def2-QZVPP (O, Mn) and def2-TZVPP (H, N) basis sets. For the RI treatment, def2-TZVPP/C auxiliary basis was used. All the O($1s$), N($1s$), Fe($1s$, $2s$, $2p$) orbitals were kept frozen at the PNO-CASPT2/CEPT2 steps. The active space is composed of O($2p_x$, $2p_y$, $3p_x$, $3p_y$) and Fe($3d_{xy}$) thus leading to CAS($8e$, $10o$) treatment. The PNO and weak-pair truncation thresholds were set to 5×10^{-1} and 1×10^{-5} Eh, respectively. Total number of AO basis functions is 752.

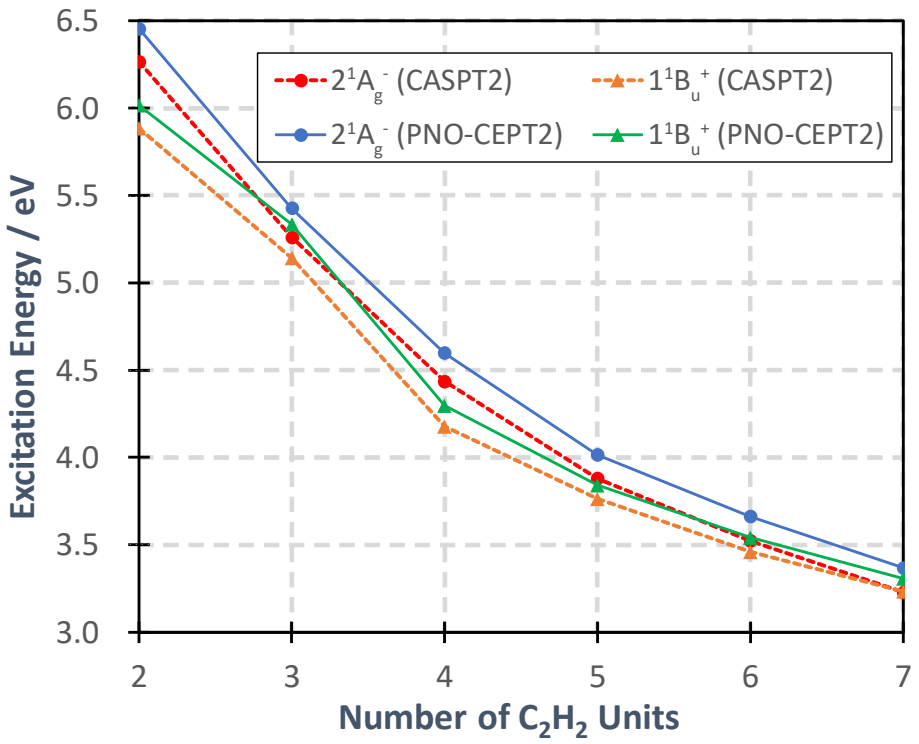

FIG. 3. Vertical EEs for linear polyene, $C_{2n}H_{2n+2}$ where n = 2–7, calculated by canonical CASPT2 and the PNO-CEPT2 methods. For the PNO-CEPT2 computation, PNO and weak-pair truncation thresholds were set to 5.0×10^{-8} and 1.0×10^{-5} Eh, respectively. The def2-TZVPP basis and def2-TZVPP/JK auxiliary basis sets were used. All the C($1s$) orbitals were kep frozen after the reference CASSCF calculation while C($2p_z$) orbitals were included in the active space, thus leading to CAS(2n e, 2n o) treatment. No shifts were used in the zeroth-order Hamiltonian.