Multicomponent Coupled Cluster Singles and Doubles with Density Fitting: Protonated Water Tetramers with Quantized Protons

Fabijan Pavošević, Zhen Tao, and Sharon Hammes-Schiffer*

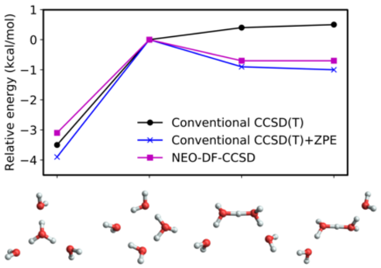
Department of Chemistry, Yale University
225 Prospect Street, New Haven, Connecticut 06520 USA

*Corresponding author email: sharon.hammes-schiffer@yale.edu

Abstract

Nuclear quantum effects such as zero-point energy are important for describing a wide range of chemical properties. The nuclear-electronic orbital (NEO) approach incorporates such effects into quantum chemistry calculations by treating specified nuclei, typically protons, quantum mechanically on the same level as electrons. Herein both the traditional and t_1-transformed NEO coupled cluster with singles and doubles (NEO-CCSD) methods are implemented with a density fitting (DF) scheme for approximating the four-center two-particle integrals. The enhanced computational efficiency enables calculations on larger molecules with multiple quantum protons. The NEO-DF-CCSD method predicts proton affinities within chemical accuracy. Its application to protonated water tetramers with all nine protons treated quantum mechanically produces the qualitatively correct ordering of the isomer energies, which are strongly influenced by the zero-point energy contributions inherently included in NEO energy calculations. This work showcases the capabilities of the NEO-DF-CCSD method and provides the foundation for future developments and applications.

TOC Graphic



Nuclear quantum effects play an important role in many chemical processes, such as hydrogen-bonding interactions, proton transfer reactions, and hydrogen tunneling. The nuclear-electronic orbital (NEO) approach[1-2] provides a computationally efficient framework for inclusion of the most significant nuclear quantum effects. In this approach, select nuclei, typically protons, are treated quantum mechanically at the same level as electrons with molecular orbital techniques, thereby avoiding the Born-Oppenheimer approximation between the different types of quantum particles (i.e., the electrons and protons). An advantage of the NEO methods over their purely electronic counterparts is that nuclear quantum effects such as zero-point energy and nuclear delocalization are included during the energy calculations, geometry optimizations, and reaction path calculations.[2]

The NEO Hartree-Fock (NEO-HF) method, in which the nuclear-electronic wave function is represented as a product of a nuclear and an electronic Slater determinant, is the simplest method within the NEO framework.[1] However, due to lack of correlation effects between the quantum particles, the calculated proton densities and resulting properties are highly inaccurate.[3-4] In order to incorporate the missing correlation effects, both wave function[1, 3-5] and density functional theory (DFT)[6-8] methods within the NEO framework have been developed. Among the various NEO wave function methods developed, the multicomponent coupled cluster methods,[9-10] which are parameter free and systematically improvable, have been the most successful.[3-5] Previously, we showed that the NEO coupled cluster with singles and doubles (NEO-CCSD) method predicts accurate proton densities, energies, and optimized geometries.[3-4] Its success is ascribed to the ability to relax orbitals in the presence of correlation effects through the exponential ansatz of the singles excitation operator.[3] For this reason, the NEO-CCSD method outperforms the NEO-CCD method, in which such orbital relaxation effects are missing.[3-4] This realization spawned different research directions in which orbitals are explicitly optimized either by the projective technique or variationally, leading to the NEO-BCCD method,[4] which uses Brueckner orbitals, and the orbital-optimized NEO-OOCCD method,[5] respectively. The results obtained with these methods confirmed our hypothesis about the importance of orbital optimization,[3-5] as also subsequently confirmed by other work.[11-12]

Our initial NEO-CCSD implementation was limited to systems composed of no more than a few atoms with only one proton treated quantum mechanically.[3-4] Because that implementation utilized a four-center integral code to store the two-particle integrals, the high memory requirement was the major bottleneck for studying larger systems. In this Letter, we present a computationally efficient implementation that relies on density fitting (DF) for approximating the four-center two-particle integrals,[13-14] thereby significantly reducing the memory requirements. Previous work[15] implemented DF in conjunction with NEO-DFT, but it is even more critical for coupled cluster methods. We used this DF implementation of the NEO-CCSD method to compute the proton affinities for a set of 23 molecules and to calculate the relative stabilities of four different protonated water tetramer isomers, treating all nine protons quantum mechanically. These applications highlight the capabilities of the NEO-CCSD method for computing accurate energies that inherently incorporate the zero-point energies of the quantum protons.

In the NEO-CC approach,[2-3] the wave function is represented by $\left|\Psi_{\text{NEO-CC}}\right\rangle = e^{\hat{T}}\left|0^{\text{e}}0^{\text{p}}\right\rangle$, where $\left|0^{\text{e}}0^{\text{p}}\right\rangle$ is the NEO-HF reference wave function and $\hat{T}$ is the excitation cluster operator that incorporates correlation effects between quantum particles through single, double, and higher excitations from the reference determinant. In the NEO-CCSD method, the cluster operator is

$$\hat{T} = \hat{T}_1 + \hat{T}_2 = t_a^i a_i^a + t_A^I a_I^A + \frac{1}{4} t_{ab}^{ij} a_{ij}^{ab} + \frac{1}{4} t_{AB}^{IJ} a_{IJ}^{AB} + t_{aA}^{iI} a_{iI}^{aA} \equiv \sum_{\alpha} t_{\alpha} a^{\alpha} \tag{1}$$

where $a^{\alpha} = a_{\alpha}^{\dagger} = \left\{ a_i^a, a_I^A, a_{ij}^{ab}, a_{IJ}^{AB}, a_{iI}^{aA} \right\}$ are excitation operators, the general electronic excitation operator $a_{p_1 p_2 \cdots p_n}^{q_1 q_2 \cdots q_n} = a_{q_1}^{\dagger} a_{q_2}^{\dagger} \ldots a_{q_n}^{\dagger} a_{p_n} \ldots a_{p_2} a_{p_1}$ is defined as a string of fermionic creation and annihilation ($a_p^{\dagger}$ and a_p) operators, and α represents the excitation manifold. The remaining excitation operators are defined analogously. Throughout this work, we utilize the Einstein summation convention for repeated indices. Furthermore, the $p, q, r, s, \ldots$ indices denote general electronic spin orbitals, $i, j, k, l, \ldots$ denote occupied electronic spin orbitals, and $a, b, c, d, \ldots$ denote unoccupied electronic spin orbitals. The protonic spin orbitals are denoted analogously using the corresponding upper-case indices.

The unknown amplitudes t_{α} (i.e., the wave function parameters) in Eq. (1) are determined using the projective technique for each α as

$$\left\langle 0^{\mathrm{e}} 0^{\mathrm{p}} \middle| a_{\alpha} e^{-\hat{T}_1 - \hat{T}_2} \hat{H}_{\mathrm{NEO}} e^{\hat{T}_1 + \hat{T}_2} \middle| 0^{\mathrm{e}} 0^{\mathrm{p}} \right\rangle = \sigma_{\alpha} = 0 \tag{2}$$

Here $\hat{H}_{\mathrm{NEO}} = h_q^p a_p^q + \frac{1}{4} \bar{g}_{rs}^{pq} a_{pq}^{rs} + h_Q^P a_P^Q + \frac{1}{4} \bar{g}_{RS}^{PQ} a_{PQ}^{RS} - g_{qQ}^{pP} a_{pP}^{qQ}$ is the NEO Hamiltonian expressed using the second-quantization formalism, $h_q^p = \langle q | \hat{h}^{\mathrm{e}} | p \rangle$ is an electronic core Hamiltonian matrix element, and $\bar{g}_{rs}^{pq} = g_{rs}^{pq} - g_{rs}^{qp} = \langle rs | pq \rangle - \langle rs | qp \rangle$ is an antisymmetrized two-electron repulsion tensor element. The corresponding protonic, h_Q^P and $\bar{g}_{RS}^{PQ}$, and mixed electronic-protonic, g_{qQ}^{pP}, counterparts are defined analogously. The total NEO-CCSD energy is calculated from $E_{\text{NEO-CCSD}} = \left\langle 0^{\mathrm{e}} 0^{\mathrm{p}} \middle| e^{-\hat{T}_1 - \hat{T}_2} \hat{H}_{\mathrm{NEO}} e^{\hat{T}_1 + \hat{T}_2} \middle| 0^{\mathrm{e}} 0^{\mathrm{p}} \right\rangle$. In the remainder of this Letter, this approach is denoted the traditional NEO-CCSD method.

In addition to the traditional NEO-CCSD method, herein we also introduce the t_1-transformed NEO-CCSD method, which is an extension of the analogous implementation in conventional electronic structure theory.[16-20] The details of this method are provided in the SI. The t_1-transformed NEO-CCSD method is appealing because the t_α amplitude equations contain fewer terms (i.e., 336 for traditional NEO-CCSD and 110 for t_1-transformed NEO-CCSD), leading to a more compact and efficient implementation. These two NEO-CCSD methods are mathematically equivalent and thus produce identical results. The t_1-transformed NEO-CCSD method bears a close resemblance to the NEO-BCCD method,[4] as discussed further in the SI, and therefore this implementation also enables the facile implementation of the NEO-BCCD method.

We applied the DF approximation to both the traditional and t_1-transformed NEO-CCSD implementations. The DF approximation is a type of tensor factorization that decomposes the four-center two-particle integrals as a product of three-center and two-center two-particle integrals.[13-14] This factorization is achieved by approximating the electron density of the four-center two-electron integrals $(\mu\nu|\rho\sigma) = \langle\mu\rho|\nu\sigma\rangle$ as $|\rho\sigma) \approx |\widetilde{\rho\sigma}) = \sum_M C_{\rho\sigma}^N |N)$, where $|N)$ is an auxiliary function that is used to fit that density, and $C_{\rho\sigma}^N$ are the fitting coefficients. In this work, the $\mu, \nu, \ldots$ and $\mu', \nu', \ldots$ indices denote electronic and protonic basis functions, respectively, and the $M, N, \ldots$ and $M', N', \ldots$ indices denote auxiliary electronic and protonic basis functions, respectively. Note that for the discussion of the DF approximation, we use the chemist notation for the two-particle integrals for simplicity. By employing the fitted density, the four-center two-electron integrals are approximated as

$$(\mu\nu|\rho\sigma) \approx \left(\widetilde{\mu\nu}\middle|\widetilde{\rho\sigma}\right) = \sum_{MN} C_{\mu\nu}^M (M|N) C_{\rho\sigma}^N \tag{3}$$

Minimization of the error in the Coulomb energy $\left(\mu\nu - \widetilde{\mu\nu}\middle|\rho\sigma - \widetilde{\rho\sigma}\right)$ leads to the set of linear equations $\sum_N (M|N) C_{\rho\sigma}^N = (M|\rho\sigma)$, and solution of these equations provides the optimal fitting coefficients $C_{\rho\sigma}^{N}$. Insertion of these optimal fitting coefficients into Eq. (3) provides the final expression for approximating the four-center two-electron integrals:

$$\left(\mu\nu\middle|\rho\sigma\right)\approx\sum_{MN}\left(\mu\nu\middle|M\right)\left(M\middle|N\right)^{-1}\left(N\middle|\rho\sigma\right) \tag{4}$$

where $\left(\mu\nu\middle|M\right)$ and $\left(M\middle|N\right)$ are three-center and two-center two-electron repulsion integrals, respectively.

This DF procedure for electrons can also be used for approximating the four-center two-proton integrals as[15]

$$\left(\mu'\nu'\middle|\rho'\sigma'\right)\approx\sum_{M'N'}\left(\mu'\nu'\middle|M'\right)\left(M'\middle|N'\right)^{-1}\left(N'\middle|\rho'\sigma'\right) \tag{5}$$

Moreover, the four-center electron-proton attraction integrals, $\left(\mu\nu\middle|\mu'\nu'\right)=\left\langle\mu\mu'\middle|\nu\nu'\right\rangle$, can be approximated using three possible approaches as $\left(\widetilde{\mu\nu}\middle|\widetilde{\mu'\nu'}\right)$, $\left(\mu\nu\middle|\mu'\nu'\right)$, or $\left(\mu\nu\middle|\widetilde{\mu'\nu'}\right)$. Previous multicomponent work[15] used only the second approach, but in this Letter, we investigate both the second and third approaches. The first approach requires the calculation of the pseudo inverse of the two-center electron-proton attraction integrals, $\left(M\middle|N'\right)$, and therefore will not be considered further. The second and third approaches are implemented as

$$\left(\mu\nu\middle|\mu'\nu'\right)\approx\left(\mu\nu\middle|\mu'\nu'\right)=\sum_{MN}\left(\mu\nu\middle|M\right)\left(M\middle|N\right)^{-1}\left(N\middle|\mu'\nu'\right) \tag{6}$$

$$\left(\mu\nu\middle|\mu'\nu'\right)\approx\left(\mu\nu\middle|\widetilde{\mu'\nu'}\right)=\sum_{M'N'}\left(\mu\nu\middle|M'\right)\left(M'\middle|N'\right)^{-1}\left(N'\middle|\mu'\nu'\right) \tag{7}$$

The DF approximation reduces the memory requirements for storing four-center two-particle integrals from N_{bf}^{4} to $N_{\mathrm{bf}}^{2}N_{\mathrm{aux}}$, where N_{bf} and N_{aux} are the number of basis and auxiliary functions, respectively (electronic or protonic). Below we analyze the errors introduced by approximating each of the four-center two-particle integrals with the DF procedure.

We implemented both the traditional NEO-CCSD and the t_1-transformed NEO-CCSD methods in a developmental version of the Q-Chem quantum chemistry software.[21] We confirmed that these two methods provide the same NEO-CCSD correlation energies, thereby validating both implementations. Furthermore, we implemented both the traditional NEO-CCSD and the t_1-transformed NEO-CCSD methods using the four-center (without DF) and the three-center (with DF) approaches. We will refer to these methods as the NEO-CCSD and NEO-DF-CCSD methods, respectively. Note that the error in the Coulomb energy associated with the DF approximation is always positive, and therefore the DF coefficients are determined variationally. However, the CCSD method is not variational because of the projective technique for determining the unknown coupled cluster amplitudes. All of the numerical results given in this Letter were obtained with the t_1-transformed NEO-CCSD implementation because it is slightly more computationally efficient.

Prior to chemical applications of this approach, we analyzed the numerical errors introduced by the DF approximation for the four-center two-particle integrals, namely $\left(\mu\nu|\rho\sigma\right)$, $\left(\mu'\nu'|\rho'\sigma'\right)$, and $\left(\mu\nu|\mu'\nu'\right)$. These tests were performed using the NEO-(DF)-CCSD method on the protonated water dimer optimized at the CCSD/aug-cc-pVDZ level, where all five protons were treated quantum mechanically. The NEO calculations employed the aug-cc-pVDZ[22-23] and PB4-F2 (4s3p2d2f)[24] electronic and protonic basis sets, respectively. The aug-cc-pVDZ-RI (aDZ-RI), aug-cc-pVTZ-RI (aTZ-RI), and aug-cc-pVQZ-RI (aQZ-RI) auxiliary electronic basis sets,[25] as well as the even-tempered 8s8p8d, 8s8p8d8f, and 8s8p8d8f8g auxiliary protonic basis sets with exponents spanning the range from $2\sqrt{2}$ to 32,[26] were used for these tests.

Table 1. Absolute Errors (kcal/mol) Introduced by the Density Fitting Approximation.

	$(\mu\nu\vert\rho\sigma)$					
aux. basis	aDZ-RI		aTZ-RI		aQZ-RI	
abs. error[a]	0.135		0.001		0.002	
abs. error (conv)[b]	0.130		0.001		0.002	
	$(\mu'\nu'\vert\rho'\sigma')$					
aux. basis	8s8p8d		8s8p8d8f		8s8p8d8f8g	
abs. error[a]	0.035		0.007		0.005	
	$(\mu\nu\vert\mu'\nu')$					
	$(\widetilde{\mu\nu}\vert\mu'\nu')^c$			$(\mu\nu\vert\widetilde{\mu'\nu'})^d$		
aux. basis	aDZ-RI	aTZ-RI	aQZ-RI	8s8p8d	8s8p8d8f	8s8p8d8f8g
abs. error[a]	0.825	0.720	0.229	0.142	0.025	0.025

[a] Absolute errors correspond to energy difference between NEO-CCSD and NEO-DF-CCSD when only the specified type of four-center two-particle integral is approximated with DF.

[b] Absolute errors correspond to energy difference between conventional electronic CCSD and DF-CCSD.

[c] Approximation given in Eq. (6).

[d] Approximation given in Eq. (7).

The absolute errors introduced by the DF approximation with respect to our four-center integral code are provided in Table 1. The DF approximation of the four-center two-electron integrals in the NEO-DF-CCSD method produces nearly identical absolute errors as those produced by the conventional electronic DF-CCSD method. The DF approximation of the four-center two-proton integrals introduces errors of similar magnitude (i.e., less than 0.01 kcal/mol) for the largest auxiliary basis sets studied. Approximating the four-center electron-proton integrals, $(\mu\nu|\mu'\nu')$, via Eq. (6) is controlled by the size of the auxiliary electronic basis set and leads to significantly larger absolute errors of ~0.2 kcal/mol even for the largest basis set employed. Approximating the four-center electron-proton integrals via Eq. (7) is controlled by the size of the auxiliary protonic basis set and leads to much smaller absolute errors of ~0.02 kcal/mol for the largest two basis sets employed. Therefore, we will use this approximation as a default. Furthermore, the results from Table 1 indicate that the 8s8p8d8f basis set is an adequate auxiliary protonic basis set for approximating the four-center proton-proton and electron-proton integrals.

Finally, the total absolute error introduced by the DF approximation with all types of four-center two-particle integrals approximated is 0.167 kcal/mol using the aDZ-RI and 8s8p8d8f auxiliary basis sets and 0.031 kcal/mol using the aTZ-RI and 8s8p8d8f auxiliary basis sets. Given this level of accuracy for the energies, the DF approximation presumably also produces accurate proton densities, as shown previously for the NEO-CCSD method.[4]

We used the NEO variants of HF theory, second-order Møller-Plesset perturbation theory (MP2), and CCSD to compute the proton affinities for a set of 23 molecules[8, 27] and compared the predicted proton affinities to the experimentally determined values.[28-31] The proton affinity of a molecule A within the NEO framework was computed via the expression $\mathrm{PA(A)} = E_{\mathrm{A}} - E_{\mathrm{HA^+}} + 5/2\, RT$,[8, 27] where E_{A} is the energy of molecule A calculated using the conventional electronic approach and $E_{\mathrm{AH^+}}$ is the energy of molecule $\mathrm{AH^+}$ calculated with the NEO method treating the relevant hydrogen nucleus quantum mechanically. In this expression, 5/2 RT accounts for the change in translational energy and the conversion from energy to enthalpy, where R is the universal gas constant and T is the temperature. Note that this approach does not require the calculation of a Hessian or zero-point energies because the NEO method inherently includes the zero-point energy contributions from the quantum proton. This procedure assumes that the zero-point energies associated with the other nuclei are approximately unchanged upon protonation, as validated by us previously.[3]

All of the proton affinity calculations were performed on the geometries optimized at the conventional electronic MP2/aug-cc-pVTZ level of theory. Tests on several of the smaller systems confirmed that the changes in geometry upon optimization at the conventional CCSD level do not significantly impact the calculated proton affinities (see Table S1). In our NEO calculations, the aug-cc-pVQZ electronic basis set[22-23, 32] and the PB4-F2 protonic basis set[24] were used for the quantum mechanical hydrogen, and these basis functions were centered at the hydrogen position in the MP2/aug-cc-pVTZ optimized geometries. We used the aug-cc-pVQZ-RI auxiliary basis set to approximate the four-center two-electron integrals.[25] Because only one proton was treated quantum mechanically, there was no need to approximate the other types of four-center two-particle integrals.

The mean unsigned errors (MUEs) and absolute deviations of each of these NEO methods relative to the experimental data are given in Table 2. The NEO-HF method has the largest overall MUE of 0.51 eV, followed by the NEO-DF-MP2 method, which has an overall MUE of 0.32 eV. The NEO-DF-CCSD has an overall MUE of 0.05 eV, which is within both the experimental (~0.09 eV)[31] and chemical (~0.05 eV)[20] accuracy. For comparison, the NEO-DFT method with the epc17-2 or epc19 electron-proton correlation functional produced an overall MUE of 0.06 eV for the same set of molecules.[8, 33] However, the advantages of the wave function methods studied herein over the NEO-DFT methods are the avoidance of parameterization and the potential for systematic improvement.

Table 2. Mean Unsigned Errors (MUEs) and Absolute Deviations of the Proton Affinities Relative to Experimentally Determined Proton Affinities.[a]

Molecule	Experiment	NEO-HF	NEO-DF-MP2	NEO-DF-CCSD
Amines				
NH_3	8.85	0.52	0.28	0.03
CH_3NH_2	9.32	0.49	0.26	<0.01
$CH_3CH_2NH_2$	9.45	0.48	0.26	0.01
$CH_3CH_2CH_2NH_2$	9.51	0.47	0.26	0.01
$(CH_3)_2NH$	9.63	0.47	0.24	0.03
$(CH_3)_3N$	9.84	0.46	0.25	0.05
MUE		0.48	0.26	0.02
Inorganics				
CN^-	15.31	0.93	0.37	0.10
HS^-	15.31	0.84	0.41	0.08
NO_2^-	14.75	0.54	0.38	0.03
MUE		0.77	0.39	0.07
Carboxylates				
$HCOO^-$	14.97	0.48	0.35	0.01
CH_3COO^-	15.11	0.44	0.32	0.04
$CH_3CH_2COO^-$	15.07	0.40	0.30	0.06

$CH_3CH_2CH_2COO^-$	15.03	0.37	0.28	0.08
$CH_3CH_2CH_2CH_2COO^-$	15.01	0.35	0.27	0.10
CH_3COCOO^-	14.46	0.48	0.30	0.05
CH_2FCOO^-	14.71	0.48	0.37	0.01
CHF_2COO^-	14.32	0.40	0.28	0.08
CF_3COO^-	13.99	0.38	0.22	0.14
CH_2ClCOO^-	14.58	0.48	0.32	0.04
$CH_2ClCH_2COO^-$	14.78	0.56	0.43	0.07
MUE		0.44	0.31	0.06
Aromatics				
$C_6H_5O^-$	15.24	0.45	0.39	<0.01
$C_6H_5COO^-$	14.75	0.73	0.56	0.19
$C_6H_5NH_2$	9.15	0.54	0.31	0.01
MUE		0.57	0.42	0.07
overall MUE		0.51	0.32	0.05

[a] All values are given in units of eV. Experimental data obtained from Ref. (28-31).

To further showcase the power of the NEO-DF-CCSD method, we calculated the relative stabilities of four protonated water tetramer isomers,[34-37] treating all nine protons quantum mechanically. This application is particularly challenging because inclusion of zero-point energy contributions is known to change the relative stabilities of these isomers at the coupled cluster level.[35-36] According to the electronic energies without zero-point energy contributions at the CCSD(T) level of theory, the relative stabilities of these isomers are eigen < ring < *cis*-Zundel < *trans*-Zundel (black line in Figure 1A).[36] When zero-point energy contributions are included within the harmonic approximation by diagonalizing the Hessian at this same level of theory, the ordering changes to be eigen < *trans*-Zundel < *cis*-Zundel < ring (red dashed line in Figure 1A). The inclusion of anharmonic corrections to the zero-point energy contributions does not alter this ordering (blue dashed line in Figure 1). The anharmonic corrections were calculated as the differences between anharmonic zero-point energies obtained with the vibrational second-order perturbative (VPT2) approach[38] and harmonic zero-point energies at the MP2/aug-cc-pVTZ level of theory using Q-Chem.[21] Moreover, when the oxygen atoms are assigned infinite mass so that only the zero-point energy contributions from the hydrogen nuclei are included, these trends do not change for either the harmonic or anharmonic zero-point energy calculations (red and blue dotted lines in Figure 1A). These results indicate that the zero-point energy contributions from the hydrogen nuclei alter the ordering of the relative stabilities of these isomers. Thus, this system serves as an excellent test of the NEO-DF-CCSD approach, where the nine hydrogen nuclei are treated quantum mechanically, and the associated zero-point energies are inherently included in a single-point energy calculation.

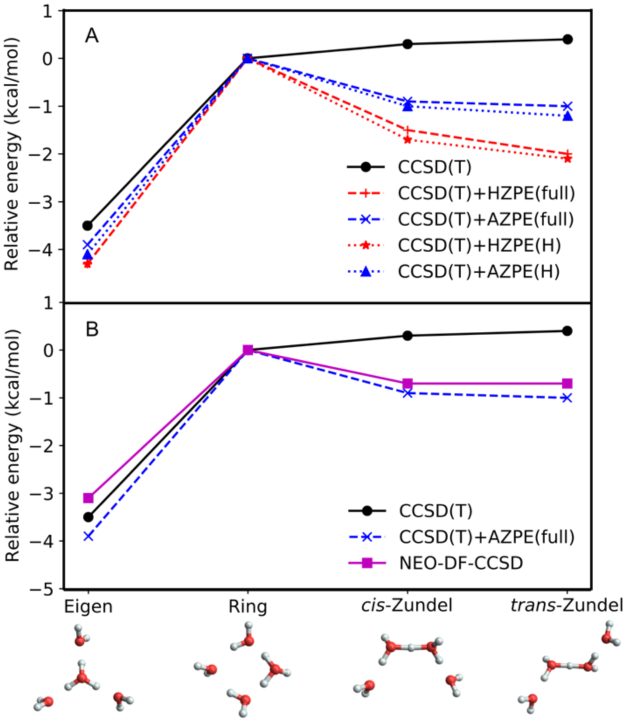


Figure 1. Relative energies of the protonated water tetramer isomers calculated with the conventional electronic CCSD(T) method without any zero-point energy (ZPE) contributions (solid black curve), the conventional CCSD(T) method with harmonic ZPE contributions (CCSD(T)+HZPE, dashed and dotted red curves), the conventional CCSD(T) method with anharmonic ZPE contributions (CCSD(T)+AZPE, dashed and dotted blue curves), and the NEO-DF-CCSD method (solid purple curve). The notations (full) and (H) indicate ZPE contributions from all nuclei and from only the hydrogen nuclei (i.e., with infinitely heavy oxygen nuclei), respectively. For visualization purposes, all energies are plotted relative to the ring isomer energy at each level of theory, and lines are drawn between data points to depict the trends.

We performed NEO-DF-CCSD calculations for the geometries optimized at the conventional CCSD(T)/aug-cc-pVTZ level of theory, as obtained from Ref. ([36]). For our NEO-

DF-CCSD calculations at these geometries, we used the aug-cc-pVTZ electronic basis set[22-23, 32] and the PB4-F2 protonic basis set,[24] in conjunction with the aug-cc-pVTZ-RI[25] and 8s8p8d8f[26] auxiliary electronic and protonic basis sets, respectively. As discussed above, the NEO energy includes the zero-point energy contributions from the quantum protons, and the contributions from the oxygen nuclei to the relative zero-point energies among the four isomers is negligible (Figure 1A). Thus, although the NEO energy can be viewed as the sum of the electronic energy and the anharmonic zero-point energy contributions from the quantum hydrogen nuclei with infinitely heavy oxygen nuclei, the target reference can be viewed as the sum of the electronic energies and the anharmonic zero-point energy contributions from all nuclei. A comparison of the energies computed with the NEO-DF-CCSD approach to these reference energies, as well as the electronic energies computed at the conventional CCSD(T) level without any zero-point energy contributions, is provided in Figure 1B. Moreover, Table 3 provides all of the energies relative to the energy of the Eigen isomer at each level of theory discussed.

Table 3. Energies of the Protonated Water Tetramer Isomers Relative to the Eigen Isomer Energy Computed with Different Methods.[a]

Method	Ring	*cis*-Zundel	*trans*-Zundel
CCSD(T)	3.5	3.8	3.9
CCSD(T)+HZPE(full)	4.3	2.8	2.3
CCSD(T)+HZPE(H)	4.3	2.6	2.2
CCSD(T)+AZPE(full)	3.9	3.0	2.9
CCSD(T)+AZPE(H)	4.1	3.1	2.9
NEO-DF-CCSD	3.1	2.4	2.4

[a] Relative energies are given in kcal/mol. HZPE and AZPE indicate harmonic and anharmonic zero-point energy contributions, respectively; (full) and (H) indicate zero-point energy contributions from all nuclei and from only the hydrogen nuclei (i.e., with infinitely heavy oxygen nuclei), respectively.

The NEO-DF-CCSD approach predicts the correct qualitative trends among the isomer energies. In particular, the NEO-DF-CCSD results are in agreement with the reference energies in that the ring isomer is the least stable among the four isomers. However, the NEO-DF-CCSD method underestimates the energy of the ring isomer relative to the Eigen isomer by ~1 kcal/mol.

Moreover, the energies of the *cis*-Zundel and *trans*-Zundel isomers are predicted to be the same with the NEO-DF-CCSD method, whereas the *trans*-Zundel isomer is predicted to be ~0.1 kcal/mol lower in energy with the reference method. These relatively small discrepancies are ascribed to several factors: (1) the use of geometries optimized with the conventional electronic structure method, (2) the size of the electronic and nuclear basis sets, and (3) the truncation of the NEO coupled cluster ansatz.

We were unable to optimize the geometries at the NEO-DF-CCSD level because of the computational expense of numerical gradients and the unavailability of analytic gradients. However, we were able to perform full geometry optimizations of the oxygen nuclei and the nuclear basis function centers at the NEO-DFT level with the B3LYP electronic exchange-correlation functional[39-41] and the epc17-2 electron-proton correlation functional[7-8] using analytic gradients available in Q-Chem.[21] As shown in Figure 2, the geometry optimizations do not impact the relative energies significantly. Further analysis of the NEO-DFT data is provided in Table S2. In addition, we investigated the impact of using smaller basis sets and found that the results are qualitatively similar (see Table S3), suggesting that the basis sets may be sufficient. Based on these analyses, we hypothesize that the minor quantitative discrepancies in the NEO-DF-CCSD results arise from the omission of triple excitations, particularly those corresponding to a double electronic excitation and a single protonic excitation. Another possibility is that the VPT2 anharmonic corrections at the MP2 level are problematic. In any case, the NEO-DF-CCSD relative energies are in reasonable agreement with the reference results.

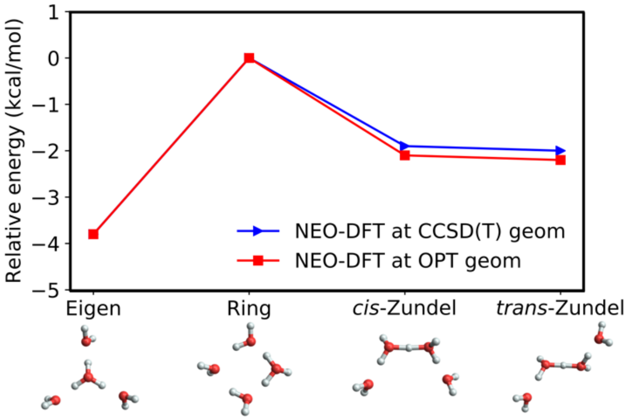


Figure 2. Relative energies of the protonated water tetramer isomers calculated with the NEO-DFT/B3LYP/epc17-2 method at the geometries optimized at the conventional CCSD(T)/aug-cc-pVTZ level of theory (solid blue curve), and with full geometry optimizations of the oxygen nuclei and the nuclear basis function centers (solid red curve). For visualization purposes, all energies are plotted relative to the ring isomer energy at each set of geometries, and lines are drawn between data points to depict the trends.

This Letter presents an efficient implementation of the NEO-CCSD method with density fitting that allows calculations of molecular systems significantly larger than those previously studied, as well as the quantum treatment of multiple protons. In addition to the traditional NEO-CCSD method, we also implemented the t_1-transformed NEO-CCSD method, which is related to the NEO orbital-optimized coupled cluster methods. Our calculations show that the NEO-DF-CCSD method predicts proton affinities within chemical and experimental accuracy. We also show that the NEO-DF-CCSD method can predict the relative stabilities of four protonated water tetramer isomers, where zero-point energy contributions are known to change the energetic ordering at the coupled cluster level. Because the zero-point energies associated with the hydrogen nuclei are inherently included in NEO calculations, only single-point NEO-DF-CCSD energy calculations are required to reproduce the qualitative trends. The accuracy of the NEO-CCSD method can be improved systematically by including higher excitations in the coupled cluster ansatz, such as triples, and its computational efficiency can be improved by low-rank tensor factorizations for approximating the wave function parameter.[42] Moreover, the NEO-CCSD method can serve as a reference in the development of more computationally efficient methods, such as NEO-DFT[7-8, 33] and NEO scaled-opposite-spin orbital-optimized MP2 (NEO-SOS′-OOMP2),[5] which require parameterization. Thus, this work opens up many research directions in terms of methodological development and applications to chemically interesting systems.

Acknowledgement

The authors thank Dr. Qi Yu, Dr. Saswata Roy, Benjamin Rousseau, Patrick Schneider, and Prof. John Tully for useful discussions. F.P. was supported by the U.S. Department of Energy, Office of Science, Offices of Basic Energy Sciences and Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program. Z.T. was supported by the National Science Foundation Grant No. CHE-1954348.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website: t_1-transformed NEO-CCSD formulation, geometry effect analysis on calculated proton affinities, DFT results for relative energies of protonated water tetramers, and basis set analysis for relative energies of protonated water tetramers.

References

(1) Webb, S. P.; Iordanov, T.; Hammes-Schiffer, S., Multiconfigurational Nuclear-Electronic Orbital Approach: Incorporation of Nuclear Quantum Effects in Electronic Structure Calculations. *J. Chem. Phys.* **2002,** *117*, 4106-4118.
(2) Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S., Multicomponent Quantum Chemistry: Integrating Electronic and Nuclear Quantum Effects via the Nuclear–Electronic Orbital Method. *Chem. Rev.* **2020,** *120*, 4222-4253.
(3) Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S., Multicomponent Coupled Cluster Singles and Doubles Theory within the Nuclear-Electronic Orbital Framework. *J. Chem. Theory Comput.* **2018,** *15*, 338-347.

(4) Pavošević, F.; Hammes-Schiffer, S., Multicomponent Coupled Cluster Singles and Doubles and Brueckner Doubles Methods: Proton Densities and Energies. *J. Chem. Phys.* **2019,** *151*, 074104.

(5) Pavošević, F.; Rousseau, B. J. G.; Hammes-Schiffer, S., Multicomponent Orbital-Optimized Perturbation Theory Methods: Approaching Coupled Cluster Accuracy at Lower Cost. *J. Phys. Chem. Lett.* **2020,** *11*, 1578-1583.

(6) Pak, M. V.; Chakraborty, A.; Hammes-Schiffer, S., Density Functional Theory Treatment of Electron Correlation in the Nuclear−Electronic Orbital Approach. *J. Phys. Chem. A* **2007,** *111*, 4522-4526.

(7) Yang, Y.; Brorsen, K. R.; Culpitt, T.; Pak, M. V.; Hammes-Schiffer, S., Development of a Practical Multicomponent Density Functional for Electron-Proton Correlation to Produce Accurate Proton Densities. *J. Chem. Phys.* **2017,** *147*, 114113.

(8) Brorsen, K. R.; Yang, Y.; Hammes-Schiffer, S., Multicomponent Density Functional Theory: Impact of Nuclear Quantum Effects on Proton Affinities and Geometries. *J. Phys. Chem. Lett.* **2017,** *8*, 3488-3493.

(9) Nakai, H.; Sodeyama, K., Many-Body Effects in Nonadiabatic Molecular Theory for Simultaneous Determination of Nuclear and Electronic Wave Functions: Ab Initio NOM/MBPT and CC Methods. *J. Chem. Phys.* **2003,** *118*, 1119-1127.

(10) Ellis, B. H.; Aggarwal, S.; Chakraborty, A., Development of the Multicomponent Coupled-Cluster Theory for Investigation of Multiexcitonic Interactions. *J. Chem. Theory Comput.* **2015,** *12*, 188-200.

(11) Brorsen, K., Quantifying Multireference Character in Multicomponent Systems with Heat-Bath Configuration Interaction. *J. Chem. Theory Comput.* **2020,** *16*, 2379-2388.

(12) Fajen, O. J.; Brorsen, K. R., Separation of Electron–Electron and Electron–Proton Correlation in Multicomponent Orbital-Optimized Perturbation Theory. *J. Chem. Phys.* **2020,** *152*, 194107.

(13) Whitten, J. L., Coulombic Potential Energy Integrals and Approximations. *J. Chem. Phys.* **1973,** *58*, 4496.

(14) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R., On Some Approximations in Applications of Xα Theory. *J. Chem. Phys.* **1979,** *71*, 3396-3402.

(15) Mejía-Rodríguez, D.; de la Lande, A., Multicomponent Density Functional Theory with Density Fitting. *J. Chem. Phys.* **2019,** *150*, 174115.

(16) Koch, H.; Christiansen, O.; Kobayashi, R.; Jørgensen, P.; Helgaker, T., A Direct Atomic Orbital Driven Implementation of the Coupled Cluster Singles and Doubles (CCSD) Model. *Chem. Phys. Lett.* **1994,** *228*, 233-238.

(17) Crawford, T. D.; Schaefer, H. F., An Introduction to Coupled Cluster Theory for Computational Chemists. *Rev. Comp. Chem.* **2000,** *14*, 33-136.

(18) Bartlett, R. J.; Musiał, M., Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007,** *79*, 291.

(19) Shavitt, I.; Bartlett, R. J., *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*. Cambridge university press: 2009.

(20) Helgaker, T.; Jørgensen, P.; Olsen, J., *Molecular Electronic-Structure Theory*. John Wiley & Sons: 2014.

(21) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; et. al., Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015,** *113*, 184-215.

(22) Dunning, T. H., Jr., Gaussian Basis Sets for Use in Correlated Molecular Calculations: 1. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989,** *90*, 1007-1023.

(23) Kendall, R. A.; Dunning Jr, T. H.; Harrison, R. J., Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992,** *96*, 6796-6806.

(24) Yu, Q.; Pavošević, F.; Hammes-Schiffer, S., Development of Nuclear Basis Sets for Multicomponent Quantum Chemistry Methods. *J. Chem. Phys.* **2020,** *152*, 244123.

(25) Weigend, F.; Köhn, A.; Hättig, C., Efficient Use of the Correlation Consistent Basis Sets in Resolution of the Identity MP2 Calculations. *J. Chem. Phys.* **2002,** *116*, 3175-3183.

(26) Culpitt, T.; Yang, Y.; Pavošević, F.; Tao, Z.; Hammes-Schiffer, S., Enhancing the Applicability of Multicomponent Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2019,** *150*, 201101.

(27) Díaz-Tinoco, M.; Romero, J.; Ortiz, J.; Reyes, A.; Flores-Moreno, R., A Generalized Any-Particle Propagator Theory: Prediction of Proton Affinities and Acidity Properties with the Proton Propagator. *J. Chem. Phys.* **2013,** *138*, 194108.

(28) Cumming, J. B.; Kebarle, P., Summary of Gas Phase Acidity Measurements Involving Acids Ah. Entropy Changes in Proton Transfer Reactions Involving Negative Ions. Bond Dissociation Energies D (A—H) and Electron Affinities EA (A). *Can. J. Chem.* **1978,** *56*, 1-9.

(29) Jolly, W. L., *Modern Inorganic Chemistry*. McGraw-Hill College: 1984.

(30) Graul, S. T.; Schnute, M. E.; Squires, R. R., Gas-Phase Acidities of Carboxylic Acids and Alcohols from Collision-Induced Dissociation of Dimer Cluster Ions. *Int. J. Mass Spectrom. Ion. Proc.* **1990,** *96*, 181-198.

(31) Hunter, E. P. L.; Lias, S. G., Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update. *J. Phys. Chem. Ref. Data* **1998,** *27*, 413-656.

(32) Woon, D. E.; Dunning Jr, T. H., Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993,** *98*, 1358-1371.

(33) Tao, Z.; Yang, Y.; Hammes-Schiffer, S., Multicomponent Density Functional Theory: Including the Density Gradient in the Electron-Proton Correlation Functional for Hydrogen and Deuterium. *J. Chem. Phys.* **2019,** *151*, 124102.

(34) Fournier, J. A.; Wolke, C. T.; Johnson, M. A.; Odbadrakh, T. T.; Jordan, K. D.; Kathmann, S. M.; Xantheas, S. S., Snapshots of Proton Accommodation at a Microscopic Water Surface: Understanding the Vibrational Spectral Signatures of the Charge Defect in Cryogenically Cooled $H^+(H_2O)_{N= 2–28}$ Clusters. *J. Phys. Chem. A* **2015,** *119*, 9425-9440.

(35) Yu, Q.; Bowman, J. M., High-Level Quantum Calculations of the IR Spectra of the Eigen, Zundel, and Ring Isomers of $H^+(H_2O)_4$ Find a Single Match to Experiment. *J. Am. Chem. Soc* **2017,** *139*, 10984-10987.

(36) Heindel, J. P.; Yu, Q.; Bowman, J. M.; Xantheas, S. S., Benchmark Electronic Structure Calculations for $H_3O^+(H_2O)_N$, N= 0–5, Clusters and Tests of an Existing 1, 2, 3-Body Potential Energy Surface with a New 4-Body Correction. *J. Chem. Theory Comput.* **2018,** *14*, 4553-4566.

(37) Finney, J. M.; DiRisio, R. J.; McCoy, A. B., Guided Diffusion Monte Carlo: A Method for Studying Molecules and Ions That Display Large Amplitude Vibrational Motions. *J. Phys. Chem. A* **2020,** *124*, 9567-9577.

(38) Barone, V., Anharmonic Vibrational Properties by a Fully Automated Second-Order Perturbative Approach. *J. Chem. Phys.* **2005,** *122*, 014108.

(39) Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988,** *37*, 785.

(40) Becke, A. D., Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988,** *38*, 3098.
(41) Becke, A. D., Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993,** *98*, 5648-5652.
(42) Riplinger, C.; Neese, F., An Efficient and near Linear Scaling Pair Natural Orbital Based Local Coupled Cluster Method. *J. Chem. Phys.* **2013,** *138*, 034106.