Multicomponent Orbital Optimized Perturbation Theory with Density Fitting: Anharmonic Zero-Point Energies in Protonated Water Clusters

Jonathan H. Fetherolf,[†] Fabijan Pavošević,[‡] Zhen Tao,[†,¶] and Sharon Hammes-Schiffer[*,†]

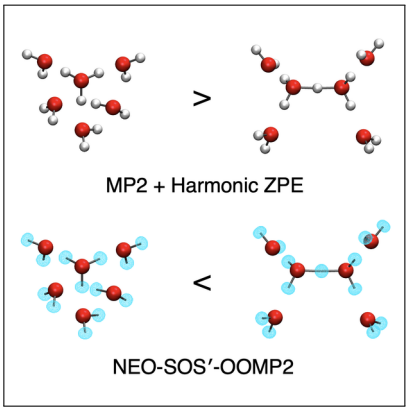
†*Department of Chemistry, Yale University, New Haven, Connecticut 06520, USA*
‡*Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, USA*
¶*Present address: Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

E-mail: sharon.hammes-schiffer@yale.edu

Abstract

Nuclear quantum effects such as zero-point energy are important in a wide range of chemical and biological processes. The nuclear-electronic orbital (NEO) framework intrinsically includes such effects by treating electrons and specified nuclei quantum mechanically on the same level. Herein, we implement the NEO scaled-opposite-spin orbital-optimized second-order Møller–Plesset perturbation theory with electron-proton correlation scaling (NEO-SOS′-OOMP2) using density fitting. This efficient implementation allows applications to larger systems with multiple quantum protons. Both the NEO-SOS′-OOMP2 method and its counterpart without orbital optimization predict proton affinities to within experimental precision and relative energies of protonated water tetramer isomers in agreement with previous NEO coupled cluster calculations. Applications to protonated water hexamers and heptamers illustrate that anharmonicity is critical for computing accurate relative energies. The NEO-SOS′-OOMP2 approach captures anharmonic zero-point energies at any geometry in a computationally efficient manner and hence will be useful for investigating reaction paths and dynamics in chemical systems.

TOC Graphic



The nuclear–electronic orbital (NEO) approach removes the Born-Oppenheimer separation between electrons and select nuclei, typically protons, treating them as a multicomponent quantum system.[1,2] By treating protons and electrons quantum mechanically at the same level of theory, NEO automatically captures important features such as vibrational zero-point energy (ZPE), proton delocalization, vibrational anharmonicity, and non-Born-Oppenheimer effects. NEO methods carry the same scaling as their conventional electronic structure counterparts, thus providing a computationally tractable framework for capturing the electronic structure and nuclear quantum effects in a single calculation. NEO density functional theory (NEO-DFT) provides a compromise between efficiency and accuracy when employing an appropriate electron-proton correlation functional.[3–5] As for any DFT-based theory, however, NEO-DFT performance is dependent on functionals that include parameters and is not strictly systematically improvable. NEO wavefunction methods, on the other hand, provide a systematically improvable set of theories, but at greater computational cost. The ideal method would be competitive with NEO-DFT in cost but competitive with high-level NEO wavefunction methods in accuracy and systematic improvability.

As in conventional electronic structure theory, NEO Hartree-Fock (NEO-HF) is the simplest multicomponent wavefunction theory and provides a reference state upon which higher-level wavefunction methods can be performed. Due to its mean-field treatment of the electron-proton interaction, however, NEO-HF is a poor starting point, predicting overly localized protonic densities and correspondingly inaccurate energies and geometries.[2] The failure of NEO-HF indicates that treating electron-proton correlation beyond the mean-field level is a requirement for a successful NEO method. NEO-HF is such a poor starting point that the lowest-order post-HF treatment of correlation, NEO Møller-Plesset second-order perturbation theory (NEO-MP2), does not significantly correct its deficiencies.[6,7] NEO coupled cluster with singles and doubles (NEO-CCSD), on the other hand, provides a robust and accurate *ab initio* benchmark, but at the cost of $\mathcal{O}(N^6)$ scaling, where N is a measure of system size.[6,8] NEO fourth-order Møller-Plesset methods have recently been shown to provide accuracy comparable to NEO-CCSD, but with $\mathcal{O}(N^7)$ formal scaling.[9]

The comparative failure of NEO-MP2 is due in part to a lack of orbital optimization, which occurs naturally in NEO-CCSD via orbital rotations arising from the single excitation terms. The necessity of orbital optimization is demonstrated by the observation that NEO coupled cluster doubles (NEO-CCD) is nearly as inaccurate as NEO-MP2. However, CCD methods with explicit orbital relaxation via Brueckner orbitals (NEO-BCCD) and variational orbital optimization (NEO-OOCCD) are much more competitive with NEO-CCSD.[10,11] The orbital-optimized NEO-MP2 (NEO-OOMP2) method shows improved protonic densities and energies compared to NEO-MP2.[11,12] The inclusion of scaling factors for opposite-spin and electron-proton correlation in NEO-SOS′-OOMP2[11] yields a method that is competitive in accuracy with NEO-OOCCD and NEO-CCSD but nominally scales as $\mathcal{O}(N^5)$ with the potential for $\mathcal{O}(N^4)$ scaling with the addition of density fitting and the Laplace transform treatment of energy denominators.[13]

Our previous implementation of NEO-SOS′-OOMP2 was useful to demonstrate accuracy approaching that of NEO-CCSD, but we were limited to proof-of-concept calculations on small molecules with only a single quantum proton. The main limitation of the previous code was that it relied on explicitly computing and storing four-center two-electron repulsion integrals, four-center two-proton repulsion integrals, and four-center electron-proton attraction integrals. In this case, the MP2 energy and gradient calculations necessary for OOMP2 face a CPU timing bottleneck of $\mathcal{O}(N_{\text{occ}}N_{\text{AO}}^4)$ floating point operations, where N_{occ} is the number of occupied orbitals and N_{AO} is the total number of atomic orbitals. In contrast, the density fitting approximation of four-center integrals[14,15] will reduce this bottleneck to $\mathcal{O}(N_{\text{occ}}^2N_{\text{vir}}^2N_{\text{aux}})$, where N_{vir} is the number

of virtual orbitals and N_{aux} is the number of auxiliary basis functions.[16,17] Although N_{aux} is typically around $3N_{\text{AO}}$, the lower exponent leads to an overall smaller prefactor in front of the $\mathcal{O}(N^5)$ cost, and a practical savings of a factor of 10 or more in many cases.

To enable the study of larger systems with multiple quantum protons, we implemented a production-level NEO-SOS′-OOMP2 code that exploits the savings of density fitting. Previously we created a similar implementation for NEO-CCSD, which enabled us to generate benchmark calculations for the proton affinities of 23 molecules, as well as to study the effect of vibrational zero-point energy on the relative stabilities of protonated water tetramers, $H_3O(H_2O)_3^+$, with all nine protons treated quantum mechanically.[18] In this Letter, we replicate all of the NEO-CCSD results at comparatively low computational cost using NEO-SOS′-OOMP2 with density fitting. Taking advantage of the lower scaling and optimized implementation, we then apply NEO-SOS′-OOMP2 to systems that are nearly twice as large as any system studied with NEO-CCSD, computing the relative stabilities of protonated water hexamers and heptamers, $H_3O(H_2O)_5^+$ and $H_3O(H_2O)_6^+$, with all protons treated quantum mechanically. We also test the non-orbital-optimized method NEO-SOS′-MP2 as an even more affordable alternative to NEO-SOS′-OOMP2. Although NEO-SOS′-MP2 carries the same formal computational scaling as NEO-SOS′-OOMP2, in practice it is typically faster by a factor of between 10 and 100 because it requires only a single-point perturbative correction, rather than an iterative one.

The NEO Hamiltonian in second quantization is

$$\begin{aligned}\hat{H}_{\text{NEO}} = h_q^p a_p^q + \frac{1}{4}\overline{g}_{rs}^{pq} a_{pq}^{rs} \\ + h_Q^P a_P^Q + \frac{1}{4}\overline{g}_{RS}^{PQ} a_{PQ}^{RS} - g_{qQ}^{pP} a_{pP}^{qQ},\end{aligned} \tag{1}$$

where $a_{p_1p_2\ldots p_n}^{q_1q_2\ldots q_n} = a_{q_1}^\dagger a_{q_2}^\dagger \ldots a_{q_n}^\dagger a_{p_n} \ldots a_{p_2} a_{p_1}$ is the general electronic excitation operator composed of creation and annihilation operators, $a_p^\dagger$ and a_p, respectively, for electronic spin orbital p. Moreover, $h_q^p = \langle p|\hat{h}^{\text{e}}|q\rangle$ are electronic core Hamiltonian matrix elements and $\overline{g}_{rs}^{pq} = g_{rs}^{pq} - g_{rs}^{qp} = \langle rs|pq\rangle - \langle rs|qp\rangle$ are antisymmetrized electron repulsion tensor elements. Lowercase indices $i, j, k, l\ldots$ denote occupied electron spin orbitals, $a, b, c, d\ldots$ denote virtual (unoccupied) electronic spin orbitals, and $p, q, r, s\ldots$ denote general electronic spin orbitals. Uppercase indices are defined analogously to denote proton orbitals. Throughout this manuscript, summation over repeated indices is assumed.

The NEO-MP2 energy Lagrangian is

$$\begin{aligned}E_{\text{NEO-MP2}} &= E_{\text{NEO-HF}} + E_{\text{ee}}^{(2)} + E_{\text{ep}}^{(2)} + E_{\text{pp}}^{(2)} \\ &= \langle 0^{\text{e}}0^{\text{p}}|\hat{H}_{\text{NEO}}|0^{\text{e}}0^{\text{p}}\rangle \\ &+ \langle 0^{\text{e}}0^{\text{p}}|\hat{W}_{\text{N}}\hat{T}_2^{(1)}|0^{\text{e}}0^{\text{p}}\rangle \\ &+ \langle 0^{\text{e}}0^{\text{p}}|\hat{\Lambda}_2^{(1)}(\hat{F}_{\text{N}}\hat{T}_2^{(1)} + \hat{W}_{\text{N}})|0^{\text{e}}0^{\text{p}}\rangle,\end{aligned} \tag{2}$$

where $E_{\text{NEO-HF}}$, $E_{\text{ee}}^{(2)}$, $E_{\text{ep}}^{(2)}$ and $E_{\text{pp}}^{(2)}$ are the NEO-HF energy and the electron-electron, electron-proton, and proton-proton NEO-MP2 correlation energies, respectively. The Fock and fluctuation operators $\hat{F}_{\text{N}}$ and $\hat{W}_{\text{N}}$ are the components of the normal-ordered NEO Hamiltonian, $\hat{H}_{\text{N}} = \hat{F}_{\text{N}} + \hat{W}_{\text{N}} - \langle 0^{\text{e}}0^{\text{p}}|\hat{H}_{\text{NEO}}|0^{\text{e}}0^{\text{p}}\rangle$, where $|0^{\text{e}}0^{\text{p}}\rangle$ is the NEO-HF ground state. Moreover, $\hat{T}_2^{(1)} = \hat{\Lambda}_2^{(1)\dagger}$ is the first-order cluster operator, given as $\hat{T}_2^{(1)} = \frac{1}{4}t_{ab}^{ij(1)}a_{ij}^{ab} + \frac{1}{4}t_{AB}^{IJ(1)}a_{IJ}^{AB} + t_{aA}^{iI(1)}a_{iI}^{aA}$, where $t_{rs}^{pq(1)} = \lambda_{pq}^{rs(1)*}$ are the unknown amplitudes.

We use density fitting[14,15] to approximate the four-center integrals (using chemist notation) as

$$\begin{aligned}g_{ij}^{ab} &= (ia|jb) \\ &\approx \sum_{MN}(ia|M)(M|N)^{-1}(N|jb),\end{aligned} \tag{3a}$$

$$\begin{aligned}g_{IJ}^{AB} &= (IA|JB) \\ &\approx \sum_{M'N'}(IA|M')(M'|N')^{-1}(N'|JB)\end{aligned} \tag{3b}$$

where M and N are electronic auxiliary basis functions, M' and N' are protonic auxiliary basis functions, and we have a choice of using either the electronic or protonic auxiliary basis for the electron-proton attraction integrals:

$$\begin{aligned} g_{iI}^{aA} &= (ia|IA) \\ &\approx \sum_{M'N'} (ia|M')(M'|N')^{-1}(N'|IA) \\ &\approx \sum_{MN} (ia|M)(M|N)^{-1}(N|IA). \end{aligned} \tag{3c}$$

To perform orbital optimization, we rewrite the expression for the NEO-MP2 energy Lagrangian as a function of orbital rotation parameters $\mathbf{x}$:

$$\begin{aligned} &E_{\text{NEO-OOMP2}}(\mathbf{x}) = \langle 0^{\text{e}}0^{\text{p}}|\{\hat{H}_{\text{NEO}}\}_{\mathbf{x}}|0^{\text{e}}0^{\text{p}}\rangle \\ &+ \langle 0^{\text{e}}0^{\text{p}}|\{\hat{W}_{\text{N}}\}_{\mathbf{x}}\hat{T}_2^{(1)}|0^{\text{e}}0^{\text{p}}\rangle \\ &+ \langle 0^{\text{e}}0^{\text{p}}|\hat{\Lambda}_2^{(1)}(\{\hat{F}_{\text{N}}\}_{\mathbf{x}}\hat{T}_2^{(1)} + \{\hat{W}_{\text{N}}\}_{\mathbf{x}})|0^{\text{e}}0^{\text{p}}\rangle, \end{aligned} \tag{4}$$

where $\{\hat{O}\}_{\mathbf{x}} = e^{\hat{X}^\dagger - \hat{X}}\hat{O}e^{\hat{X}-\hat{X}^\dagger}$ is a unitary transformation with rotation operator $\hat{X} = \hat{X}^{\text{e}} + \hat{X}^{\text{p}} = x_a^i a_i^a + x_A^I a_I^A$ and $\mathbf{x} = \{x_a^i, x_A^I\}$ is the set of unknown electronic and protonic orbital rotation parameters. We minimize the energy as a function of $\mathbf{x}$ by finding the stationary point of the orbital gradient $\mathbf{w}$, which can be separated into electronic and protonic parts $\mathbf{w}_{\text{e}}$ and $\mathbf{w}_{\text{p}}$ with elements defined as

$$(\mathbf{w}_{\text{e}})_a^i = \frac{\partial E_{\text{NEO-OOMP2}}(\mathbf{x})}{\partial x_i^a}\Big|_{\mathbf{x}_{\text{e}}=0} \tag{5a}$$

and

$$(\mathbf{w}_{\text{p}})_A^I = \frac{\partial E_{\text{NEO-OOMP2}}(\mathbf{x})}{\partial x_I^A}\Big|_{\mathbf{x}_{\text{p}}=0}. \tag{5b}$$

We alternate optimization of the electronic and protonic orbital rotation parameters, holding the other constant, until self-consistency is achieved. We perform optimization using the direct inversion of the iterative subspace (DIIS) solver,[19] as suggested by Neese and coworkers.[20] Programmable equations for the NEO orbital gradients can be found in the supporting information of Ref. 11.

To apply spin-component scaling, we further separate the electron-electron NEO-MP2 correlation energy in Eq. 2 into the same-spin contribution ($E_{\text{ee}}^{\text{ss}(2)}$) and opposite-spin contribution ($E_{\text{ee}}^{\text{os}(2)}$):

$$\begin{aligned} E_{\text{NEO-SOS}'\text{-MP2}} = E_{\text{NEO-HF}} &+ c_{\text{ss}}E_{\text{ee}}^{\text{ss}(2)} \\ &+ c_{\text{os}}E_{\text{ee}}^{\text{os}(2)} + c_{\text{ep}}E_{\text{ep}}^{(2)} + E_{\text{pp}}^{(2)}. \end{aligned} \tag{6}$$

We use the well-established scaled-opposite-spin (SOS) prefactors of $c_{\text{ss}} = 0$ and $c_{\text{os}} = 1.3$ for unoptimized orbitals and $c_{\text{os}} = 1.2$ when orbital optimization has been applied.[13,21,22] For NEO-SOS′-OOMP2, we scale the electron-proton correlation energy, $E_{\text{ep}}^{(2)}$, by a prefactor that was determined previously to be $c_{\text{ep}} = 1.3$ or 1.2,[11] depending on whether orbital optimization has been applied. Although this is no longer a fully parameter-free theory, the scaling parameters are correcting for a well-known deficiency of MP2 theory, the systematic underestimation of opposite-spin electron-electron correlation energy.[21] A similar underestimation of electron-proton correlation is found for the extension to multicomponent systems. Applying this scaling adds no computational cost, and by neglecting the same-spin electron-electron correlation, we could in principle create a fully $\mathcal{O}(N^4)$ implementation. In practice, our $\mathcal{O}(N^5)$ implementation with density fitting is quite efficient, and it is unclear whether the benefits of lower scaling would be apparent except in very large systems.

We implemented NEO-(SOS′)-OOMP2 with density fitting in a development version of the Q-Chem 5.4 quantum chemistry software package.[23] Our implementation was based on the conventional OOMP2 code described in Ref. 24, which uses the mixed Lagrangian method[17] to efficiently compute orbital gradients. Throughout this study, we treat all quantum protons with the PB4-F2 (4s3p2d2f) protonic basis set,[25] with an even-tempered 8s8p8d8f auxiliary protonic basis set with exponents spanning $2\sqrt{2}$ to 32. Our testing shows that for a given aug-cc-pVXZ (X=D,T,Q) electronic basis set[26,27] and the PB4-F2 protonic basis, density fitting with the aug-cc-pVXZ-RI electronic auxiliary basis set[28,29] and the even-tempered 8s8p8d8f pro-

tonic auxiliary basis set only introduces an error on the order of 0.01 kcal/mol in the total energy compared to results using four-center integrals. This level of density fitting error is consistent with previous results for density fitting with NEO-CCSD.[18] The error in density fitting does not systematically increase for calculations with multiple quantum protons. We provide further analysis of convergence with respect to the auxiliary basis set in the Supporting Information. Throughout this paper, all MP2-based calculations, including both conventional and NEO, are all-electron in that we have not employed any frozen-core approximation.

We first tested our implementation of NEO-SOS′-OOMP2 by calculating proton affinities for the 23 molecule test set previously used to benchmark NEO-DFT and NEO-CCSD.[5,18,30] The proton affinity PA(A) of molecule A is given by $\mathrm{PA(A)} = E_\mathrm{A} - E_{\mathrm{HA}^+} + \frac{5}{2}RT$, where E_A is the energy of molecule A, E_{HA^+} is the energy after adding a proton, R is the universal gas constant, and T is temperature. The quantity $\frac{5}{2}RT$ accounts for conversion from energy to enthalpy and for the change in translational energy. We calculated E_A with conventional electronic structure theory and E_{HA^+} with the analogous NEO method, with only the added proton treated quantum mechanically.

NEO methods are well-suited for calculating proton affinities because they automatically include the anharmonic vibrational ZPE of the proton without requiring the calculation of a nuclear coordinate Hessian or anharmonic force constants. When calculating proton affinities with the NEO approach, typically we assume that the vibrational frequencies of the other nuclei are unaffected by the addition of a proton, and thus the associated ZPEs cancel in $E_\mathrm{A} - E_{\mathrm{HA}^+}$, as justified by our previous calculations.[8] We use the same geometries as used in our previous study with NEO-CCSD. These geometries were optimized at the level of conventional electronic MP2/aug-cc-pVTZ. For the quantum proton, we use the aug-cc-pVQZ electronic basis set with the aug-cc-pVQZ-RI auxiliary basis set. We use the aug-cc-pVTZ electronic basis set with the aug-cc-pVQZ-RI auxiliary basis set for the other nuclei.

In Table 1, for each NEO method we present the absolute deviation per molecule and the mean unsigned error (MUE) for each group of molecules and for the entire set, relative to experiment.[31,32,34] For reference, NEO-HF (not shown here) results in a very large overall MUE of 0.51 eV (12 kcal/mol).[18] NEO-MP2 results in only a modest improvement over NEO-HF, with an overall MUE of 0.32 eV (7.4 kcal/mol). Orbital optimization leads to a small improvement in overall accuracy, with NEO-OOMP2 giving an MUE of 0.29 eV (6.7 kcal/mol). NEO-SOS′-OOMP2 and NEO-SOS′-MP2 produce a similar MUE of 0.07 eV (1.6 kcal/mol) and 0.06 eV (1.4 kcal/mol), respectively. NEO-CCSD achieves the best agreement with experiment among all of these NEO methods, with a MUE of only 0.05 eV (1.2 kcal/mol). Both NEO-SOS′-OOMP2 and NEO-SOS′-MP2 can be considered to achieve accuracy that is within the experimental error bars, which is estimated to be around 2 kcal/mol.[34]

From these results, it is clear that the SOS′ electron-proton correlation scaling is crucial for predicting accurate proton affinities. Orbital optimization leads to only a marginal improvement over NEO-MP2. The 0.01 eV difference in error between NEO-SOS′-MP2 and NEO-SOS′-OOMP2 is too small to be meaningful, as both methods produce proton affinities within the range of experimental accuracy. NEO-SOS′-OOMP2 and NEO-SOS′-MP2 show similar accuracy for the smaller set of molecules used to empirically derive the c_ep scaling parameter in Ref. 11. The full results for the smaller test set using density fitting are given in the Table S2.

The NEO-SOS-OOMP2 method, without scaling of the electron-proton correlation, was previously shown to result in much higher overall errors than the SOS′ variant.[11] According to our current calculations, it gives errors well in excess of 2 kcal/mol and therefore does not achieve chemical or experimental accuracy. On the basis of this analysis, we conclude that the major source of error in NEO-MP2 energies is the systematic underestimation of electron-proton correlation. Although orbital optimization marginally alleviates this systematic error, the SOS′ electron-proton correlation scaling

Table 1: Absolute Deviation and Mean Unsigned Error (MUE) of Proton Affinities (in electron volts) Relative to Experiment[a]

Molecule	Experiment[a]	NEO-CCSD[b]	NEO-MP2[b]	NEO-SOS′-MP2	NEO-OOMP2	NEO-SOS′-OOMP2
Amines						
NH_3	8.85	0.03	0.28	0.05	0.20	0.05
CH_3NH_2	9.32	<0.01	0.26	0.01	0.19	0.09
$CH_3CH_2NH_2$	9.45	0.01	0.26	0.01	0.18	0.10
$CH_3CH_2CH_2NH_2$	9.51	0.01	0.26	0.01	0.19	0.10
$(CH_3)_2NH$	9.63	0.03	0.24	0.03	0.17	0.13
$(CH_3)_3N$	9.84	0.05	0.25	0.05	0.19	0.16
MUE	–	0.02	0.26	0.03	0.19	0.10
Inorganics						
CN^-	15.31	0.10	0.37	0.10	0.29	0.01
HS^-	15.31	0.08	0.41	0.15	0.31	0.03
NO_2^-	14.75	0.03	0.38	0.03	0.36	0.02
MUE	–	0.07	0.39	0.09	0.32	0.02
Carboxylates						
$HCOO^-$	14.97	0.01	0.35	0.05	0.34	0.02
CH_3COO^-	15.11	0.04	0.32	0.03	0.31	0.05
$CH_3CH_2COO^-$	15.07	0.06	0.30	0.01	0.29	0.07
$CH_3CH_2CH_2COO^-$	15.03	0.08	0.28	0.02	0.27	0.10
$CH_3CH_2CH_2CH_2COO^-$	15.01	0.10	0.27	0.03	0.26	0.11
CH_3COCOO^-	14.46	0.05	0.30	<0.01	0.36	0.04
CH_2FCOO^-	14.71	0.01	0.37	0.07	0.38	0.01
CHF_2COO^-	14.32	0.08	0.28	0.01	0.27	0.08
CF_3COO^-	13.99	0.14	0.22	0.07	0.20	0.14
CH_2ClCOO^-	14.58	0.04	0.32	0.03	0.34	0.03
$CH_2ClCH_2COO^-$	14.78	0.07	0.43	0.14	0.44	0.07
MUE	–	0.06	0.31	0.04	0.31	0.06
Aromatics						
$C_6H_5O^-$	15.24	<0.01	0.39	0.10	0.44	0.05
$C_6H_5COO^-$	14.75	0.19	0.56	0.25	0.55	0.18
$C_6H_5NH_2$	9.15	0.01	0.31	0.02	0.25	0.07
MUE	–	0.07	0.42	0.12	0.41	0.10
Overall MUE	–	0.05	0.32	0.06	0.29	0.07

[a] Experimental values obtained from refs. 31–34.
[b] NEO-CCSD and NEO-MP2 values obtained from ref. 18.

is necessary to achieve good agreement with experiment and competitiveness with NEO-CCSD. We point out, however, that properties depending more directly on the wavefunction benefit more tangibly from orbital relaxation, as is the case for protonic densities and nuclear geometries.

The value of the electron-proton correlation scaling prefactor c_{ep} used in these calculations was determined previously by minimizing the MUE for the proton affinities of a smaller set of molecules calculated using a mixed electronic basis set with aug-cc-pVTZ on the classical nuclei and aug-cc-pVQZ on the quantum proton.[11] The use of a larger electronic basis set on the quantum proton is common practice for NEO methods, as the higher angular momentum basis functions assist in capturing electron-proton correlation. We find, however, that the proton affinities are not as accurate if the larger aug-cc-pVQZ electronic basis set is used for *all* the atoms. The physical basis for this discrepancy is that the mixed electronic basis set has extra electronic basis functions from aug-cc-pVQZ only for HA^+, thereby providing more variational flexibility for the electrons and enhancing the electron-proton and electron-electron correlation in HA^+ in compar-

ison to A. When the aug-cc-pVQZ electronic basis set is used for all atoms, the optimal c_{ep} is 1.4 for NEO-SOS′-OOMP2 and 2.0 for NEO-SOS′-MP2. These scaling factors also perform well when the aug-cc-pVTZ electronic basis set is used for all atoms (Table S2). The larger scaling factors are needed to quantitatively predict proton affinities, but for applications that compare relative energies with the same number of quantum protons, the exact value of c_{ep} has less impact as long as it remains within a reasonable range. The remainder of our results are of this latter type, and therefore we continue using the c_{ep} of 1.2 for NEO-SOS′-OOMP2 and 1.3 for NEO-SOS′-MP2. More details on the electronic basis set dependence of the electron-proton correlation scaling factor is provided in the Supporting Information.

After demonstrating excellent agreement in proton affinities, we continued to benchmark NEO-SOS′-OOMP2 against NEO-CCSD by calculating the relative energies of isomers for protonated water tetramers with all nine protons treated quantum mechanically. At the conventional CCSD(T)/aug-cc-pVTZ level of theory, the relative electronic energies of the four most stable isomers are ordered Eigen < ring < *cis*-Zundel < *trans*-Zundel (red data in Fig. 1). However, inclusion of harmonic vibrational ZPE via the conventional CCSD(T) Hessian changes the ordering to Eigen < *trans*-Zundel < *cis*-Zundel < ring. This ordering is not changed any further by the inclusion of an anharmonic correction to the ZPE computed with vibrational second-order perturbation theory (VPT2)[35] at the conventional MP2/aug-cc-pVTZ level. Therefore, we choose to show only CCSD(T) with anharmonic ZPE included (CCSD(T)+AZPE, blue data), as it represents the most rigorous estimate of vibrational ZPE available from conventional electronic structure. The disagreement in the relative energies with and without ZPE renders this an ideal test case for the NEO approach, which captures the anharmonic ZPE from the quantum protons without the need to compute any Hessian or potential energy surface.

We computed the energies of the four $H_3O(H_2O)_3^+$ isomers with NEO-SOS′-OOMP2 and NEO-SOS′-MP2 using geometries optimized at the conventional CCSD(T)/aug-cc-pVTZ level. For direct comparison to the NEO-CCSD results, we used the same aug-cc-pVTZ electronic basis set with the aug-cc-pVTZ-RI auxiliary basis set. We previously showed agreement between the relative energies from NEO-CCSD and those of conventional CCSD(T)+AZPE (the green and blue curves in Fig. 1, respectively). Fig. 1 shows that both NEO-SOS′-OOMP2 and NEO-SOS′-MP2 (dashed purple and orange data, respectively) produce excellent agreement with NEO-CCSD, with all relative energies matching the NEO-CCSD result to within 0.1 kcal/mol.

This system was quite computationally demanding to treat with NEO-CCSD, even with

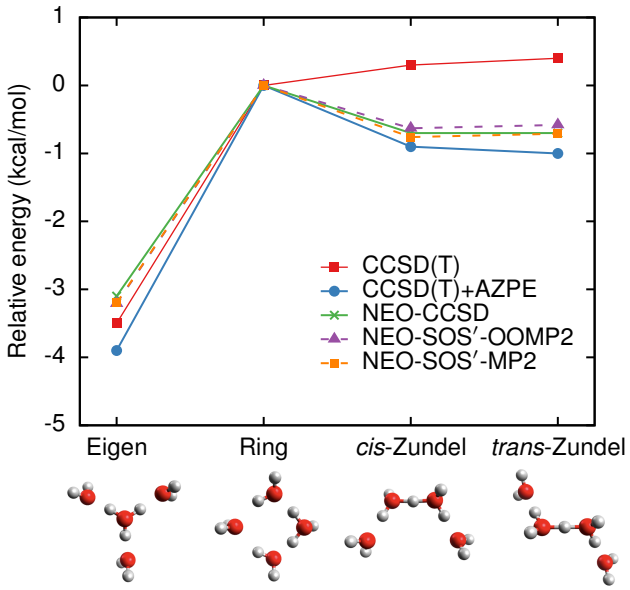


Figure 1: Relative ground-state energies of the four most stable isomers of the protonated water tetramer, $H_3O(H_2O)_3^+$. The NEO-CCSD (solid green data), NEO-SOS′-OOMP2 (dashed purple data) and NEO-SOS′-MP2 (dashed orange curve) methods are compared to conventional CCSD(T) with and without an anharmonic ZPE correction (solid red and blue data, respectively). The CCSD(T), CCSD(T)+AZPE and NEO-CCSD results are from Ref. 18. All three NEO methods qualitatively agree with CCSD(T)+AZPE, whereas CCSD(T) is qualitatively incorrect without the ZPE correction.

density fitting. In contrast, the NEO-SOS′-OOMP2 calculation can be performed on a consumer-grade desktop PC in around two hours, with the NEO-SOS′-MP2 calculation finishing in less than an hour. In both cases, the NEO approach is more practical than conventional methods for this problem because the NEO approach provides the anharmonic ZPE without requiring a nuclear coordinate Hessian or anharmonic force constants. Of course, NEO does not capture contributions to the ZPE from the heavy nuclei. In the case of water clusters, however, the vibrational energy is dominated by the proton modes, which are by far the most numerous and highest in energy. We confirmed that the ZPE is dominated by proton motion by assigning the oxygen atoms infinite mass and recomputing the conventional Hessian, which decreased the total ZPE by less than 10%, and did not affect the relative energies of the isomers. To further demonstrate the value of NEO-SOS′-OOMP2, we calculated the relative energies of much larger protonated water clusters that are inaccessible via our current implementation of NEO-CCSD. We focus on hexamers and heptamers, $H_3O(H_2O)_5^+$ and $H_3O(H_2O)_6^+$, where all 13 and 15 protons, respectively, are treated quantum mechanically. Even single-point energy calculations on systems of this size are impractical for NEO-CCSD with density fitting, whereas they require only modest computational effort with our implementation of NEO-SOS′-OOMP2. Due to the system size, we cannot benchmark against NEO-CCSD, and therefore we must assess the accuracy by comparing to conventional methods for calculating electronic and vibrational ground state energies.

We obtained the structures of nine low-energy isomers of the protonated water hexamer optimized at the MP2/aug-cc-pVTZ level from Ref. 36. These structures represent a challenging application for the NEO approach because their relative energies differ by only approximately 1–2 kcal/mol, making them sensitive to subtle nuclear quantum effects.[36] We include data for all nine isomers in the Supporting Information (Fig. S1), but in this Letter we focus on five isomers that show a qualitative difference with the inclusion of ZPE. We use MP2 as the conventional electronic structure method for consistency with Ref. 36 and 37 and because implementations with normal mode analysis are widely available. Using SOS-MP2 or SOS-OOMP2 does not change the ordering of the electronic energies. Fig. 2(a) shows the energies relative to the lowest energy isomer, T1.

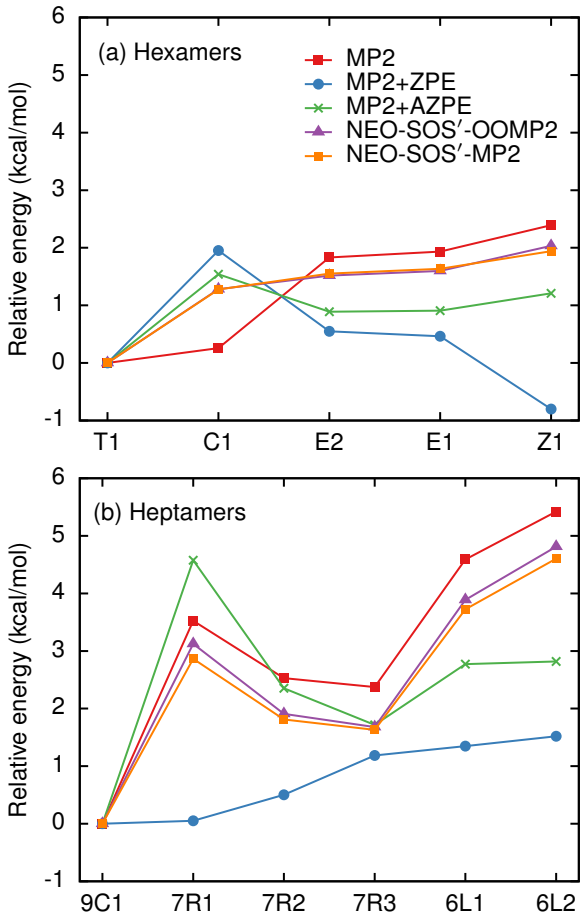


Figure 2: Relative ground-state energies for low-lying isomers of (a) protonated water hexamers, $H_3O(H_2O)_5^+$, and (b) protonated water heptamers, $H_3O(H_2O)_6^+$. The NEO-SOS′-OOMP2 (purple data) and NEO-SOS′-MP2 (orange data) methods are compared to conventional MP2 (red data), conventional MP2 with a harmonic ZPE correction (blue data), and conventional MP2 with an anharmonic ZPE correction (green data). Both NEO methods qualitatively disagree with the MP2+ZPE relative energies but agree much more closely with the MP2+AZPE energies, highlighting the importance of anharmonicity.

The relative energies calculated at the conventional MP2/aug-cc-pVTZ level (red data in Fig. 2(a)) are arranged in ascending order, T1 < C1 < E2 < E1 < Z1. When we include the harmonic ZPE calculated via the MP2 Hessian (MP2+ZPE, blue data) the ZPE has a dramatic effect on the ordering, reversing the trend of the latter four structures (Z1 < E1 < E2 < C1), and increasing the T1-C1 relative energy from 0.3 kcal/mol to 2 kcal/mol.

The NEO-SOS′-OOMP2 and NEO-SOS′-MP2 methods predict a much more subtle change due to nuclear quantum effects. We performed these computations using the aug-cc-pVTZ electronic basis set with the aug-cc-pVTZ-RI auxiliary basis set. These NEO methods predict the same ordering as conventional MP2 (purple and orange data in Fig. 2(a), with the most notable difference being an increase in the T1-C1 relative energy from 0.3 kcal/mol to 1.3 kcal/mol. To examine the discrepancy between these NEO methods and the MP2+ZPE approach, we tested two possible causes: the NEO methods lacking ZPE involving the heavy (oxygen) atoms and the MP2+ZPE approach lacking anharmonicity. The MP2+ZPE results computed with infinitely heavy oxygen atoms (shown in Fig. S1) are very similar to our original MP2+ZPE results, ruling out the treatment of the oxygen atoms as the source of the discrepancy.

Calculating the necessary anharmonic force constants to approximate the anharmonic ZPE for a system of this size at the conventional MP2 level is computationally challenging. To estimate the qualitative impact of anharmonicity, we calculated the VPT2 frequencies at the B3LYP/6-311G(d,p) level.[35,38–40] We then subtracted the harmonic ZPE from the anharmonic ZPE at this same level of theory to obtain an estimated correction due to anharmonicity. The MP2+AZPE results (green data in Fig. 2(a)) are the sum of the MP2 electronic energy and harmonic ZPE, as well as the B3LYP anharmonic correction. Including anharmonicity significantly moderates the effect of the ZPE and restores the E2 < E1 < Z1 ordering. The MP2+AZPE approach still predicts C1 to be higher in energy than E2, E1 or Z1, but the T1-C1 relative energy is now 1.5 kcal/mol compared to 1.3 kcal/mol from NEO-SOS′-OOMP2 and 2 kcal/mol from MP2+ZPE. Therefore, we conclude that the anharmonicity plays a major role in determining the relative energies of large water clusters, which exhibit many nearly-degenerate configurations. Excluding anharmonicity leads to an overestimation of the impact of ZPE, whereas the NEO approach may slightly underestimate the impact, although a definitive benchmark is not available, and the NEO methods may be more accurate than the MP2+AZPE approach. Regardless, the NEO approach provides a substantially more computationally efficient way to include effects of anharmonicity than calculating the anharmonic force constants.

Finally, we used NEO-SOS′-OOMP2 to calculate the relative energies of protonated water heptamers. We obtained the structures of six topologically distinct low-energy isomers from Ref. 37, in which the geometry optimizations and harmonic ZPEs were computed at the MP2/aug-cc-pVDZ level. Fig. 2(b) shows their MP2 electronic energies (red data) and MP2 energies with harmonic ZPE corrections (blue data). The addition of harmonic ZPE changes the relative energies significantly, most notably causing 9C1 and 7R1 to be nearly degenerate as the lowest-energy isomer, in contrast to the separation of 3.5 kcal/mol for the purely electronic energies.

We do not observe such a stark difference for the relative energies computed with NEO-SOS′-OOMP2 and NEO-SOS′-MP2 (purple and orange data in Fig. 2(b)). We used the same MP2-optimized geometries in conjunction with the aug-cc-pVDZ electronic basis set with the aug-cc-pVTZ-RI auxiliary basis set. The NEO methods only produce a subtle change compared to the conventional MP2 electronic energies. To investigate the difference between the NEO methods and the MP2+ZPE approach, we calculated the anharmonic correction to the conventional harmonic ZPE using VPT2 at the B3LYP/6-311G(d,p) level of theory. The MP2+AZPE relative energies (green data) again show that anharmonicity offsets the effects of the harmonic ZPE, leading to a much

more subtle change relative to the conventional electronic energies. Although the MP2+AZPE results are not in quantitative agreement with the NEO-SOS′-(OO)MP2 results, the trends are qualitatively similar. Moreover, as mentioned above, the NEO approach may be more accurate than the MP2+AZPE approach and is definitely more computationally tractable.

In this Letter, we present an efficient implementation of NEO-SOS′-OOMP2 with density fitting, enabling the study of larger systems with many nuclei treated quantum mechanically. Both NEO-SOS′-OOMP2 and NEO-SOS′-MP2 predict proton affinities to within experimental accuracy. We also calculated the relative energies of four protonated water tetramer isomers with all protons treated quantum mechanically. We found that both NEO-SOS′-OOMP2 and NEO-SOS′-MP2 correctly predict the relative energetic ordering, which is significantly influenced by the vibrational ZPE, and produce energies that agree with the NEO-CCSD results to within 0.2 kcal/mol. Finally, we computed the relative energies for protonated water hexamers and heptamers with all protons treated quantum mechanically. The NEO-SOS′-OOMP2 and NEO-SOS′-MP2 results disagree qualitatively with the results of conventional MP2 with harmonic ZPE but agree much better with the conventional MP2 results when an anharmonic correction is applied to the conventional harmonic ZPE. Thus, these calculations indicate that anharmonicity is crucial for accurately capturing the relative energies of geometries for large water clusters and illustrate that harmonic ZPE corrections can lead to qualitatively incorrect conclusions.

The NEO framework, specifically NEO-SOS′-OOMP2, provides a way to capture anharmonic ZPEs at modest computational cost for systems where conventional methods may be impractical. Moreover, the NEO approach provides energies that include anharmonic ZPEs associated with the quantum protons for any geometry, in contrast to conventional methods that are restricted to stationary points. With this efficient implementation of NEO-SOS′-OOMP2, we have demonstrated its place as the NEO wavefunction method with the ideal balance of efficiency and accuracy. We found that the SOS′ approach to correlation energy component scaling is the most important factor in obtaining accurate ground-state energies, as demonstrated by the close agreement between NEO-SOS′-OOMP2 and NEO-SOS′-MP2. However, orbital optimization is still crucial for computing accurate protonic densities and will be important for avoiding spin contamination in open-shell multicomponent systems. Future implementation of analytic nuclear gradients[41,42] will enable us to perform geometry optimizations and dynamics with NEO-SOS′-OOMP2, enabling studies of geometric isotope effects and proton transfer dynamics.

Acknowledgement The authors thank Dr. Qi Yu and Benjamin Rousseau for helpful discussions, Dr. Kurt Brorsen for useful discussions about electronic basis sets, and Dr. Joonho Lee for providing helpful details on the Q-Chem OOMP2 implementation. This work was supported by the National Science Foundation Grant No. CHE-1954348.

Supporting Information Available

The Supporting Information is available free of charge:

Error due to density fitting; proton affinity basis set effects; relative energies of protonated water tetramers with different electron–proton scaling factors; relative energies of full set of protonated water hexamers

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) 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.

(4) 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.

(5) 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.

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

(7) Swalina, C.; Pak, M. V.; Hammes-Schiffer, S. Alternative Formulation of Many-Body Perturbation Theory for Electron–proton Correlation. *Chem. Phys. Lett.* **2005**, *404*, 394–399.

(8) Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S. Multicomponent Coupled Cluster Singles and Doubles Theory Within the Nuclear-Electronic Orbital Framework. *J. Chem. Theory Comput.* **2019**, *15*, 338–347.

(9) Fajen, O. J.; Brorsen, K. R. Multicomponent MP4 and the Inclusion of Triple Excitations in Multicomponent Many-Body Methods. *J. Chem. Phys.* **2021**, *155*, 234108.

(10) Pavošević, F.; Hammes-Schiffer, S. Multicomponent Coupled Cluster Singles and Doubles and Brueckner Doubles Methods: Proton Densities and Energies. *J. Chem. Phys.* **2019**, *151*, 1–7.

(11) 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.

(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) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled Opposite-Spin Second Order Moller-Plesset Correlation Energy: An Economical Electronic Structure Method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

(14) Whitten, J. L. Coulombic Potential Energy Integrals and Approximations. *J. Chem. Phys.* **1973**, *58*, 4496–4501.

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

(16) Weigend, F.; Häser, M. RI-MP2: First Derivatives and Global Consistency. *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)* **1997**, *97*, 331–340.

(17) Distasio, R. A.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. an Improved Algorithm for Analytical Gradient Evaluation in Resolution-of-the-Identity Second-Order Møller-Plesset Perturbation Theory: Application to Alanine Tetrapeptide Conformational Analysis. *J. Comput. Chem.* **2007**, *28*, 839–856.

(18) Pavošević, F.; Tao, Z.; Hammes-Schiffer, S. Multicomponent Coupled

Cluster Singles and Doubles with Density Fitting: Protonated Water Tetramers with Quantized Protons. *J. Phys. Chem. Lett.* **2021**, *12*, 1631–1637.

(19) Pulay, P. Convergence Acceleration of Iterative Sequences. the Case of Scf Iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398.

(20) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of Orbital-Optimized, Spin-Component Scaled Second-Order Many-Body Perturbation Theory for Thermochemistry and Kinetics. *J. Chem. Theory Comput.* **2009**, *5*, 3060–3073.

(21) Grimme, S. Improved Second-Order Møller–Plesset Perturbation Theory by Separate Scaling of Parallel- and Antiparallel-Spin Pair Correlation Energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(22) Lochan, R. C.; Head-Gordon, M. Orbital-Optimized Opposite-Spin Scaled Second-Order Correlation: An Economical Method to Improve the Description of Open-Shell Molecules. *J. Chem. Phys.* **2007**, *126*, 164101.

(23) Epifanovsky, E.; Gilbert, A. T.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L. et al. Software for the Frontiers of Quantum Chemistry: An Overview of Developments in the Q-Chem 5 Package. *J. Chem. Phys.* **2021**, *155*, 084801.

(24) Lee, J.; Head-Gordon, M. Regularized Orbital-Optimized Second-Order Møller-Plesset Perturbation Theory: A Reliable Fifth-Order-Scaling Electron Correlation Model with Orbital Energy Dependent Regularizers. *J. Chem. Theory Comput.* **2018**, *14*, 5203–5219.

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

(26) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. the Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(27) Kendall, R. A.; Dunning, 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.

(28) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: Optimized Auxiliary Basis Sets and Demonstration of Efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(29) 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.

(30) Díaz-Tinoco, M.; Romero, J.; Ortiz, J. V.; 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.

(31) 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.

(32) Jolly, W. L. *Modern Inorganic Chemistry*; McGraw-Hill College: New York, NY, 1984.

(33) 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 Processes* **1990**, *96*, 181–198.

(34) 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.

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

(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) Jieli, M.; Aida, M. Classification of OH Bonds and Infrared Spectra of the Topology-Distinct Protonated Water Clusters $H_3O^+(H_2O)_{n-1}$ ($n \leq 7$). *J. Phys. Chem. A* **2009**, *113*, 1586–1594.

(38) 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*, 785–789.

(39) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XX. a Basis Set for Correlated Wave Functions. *J. Chem. Phys.* **1980**, *72*, 650–654.

(40) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. the 3-21+G Basis Set for First-Row Elements, Li–F. *J. Comput. Chem.* **1983**, *4*, 294–301.

(41) Head-Gordon, M. An Improved Semidirect MP2 Gradient Method. *Mol. Phys.* **1999**, *96*, 673–679.

(42) Bozkaya, U.; Sherrill, C. D. Analytic Energy Gradients for the Orbital-Optimized Second-Order Møller-Plesset Perturbation Theory. *J. Chem. Phys.* **2013**, *138*, 184103.