Supplementary Materials

Accurate Non-bonded Potentials based on Periodic Quantum Mechanics Calculations for use in Molecular Simulations of Materials and Systems

Saber Naserifar[1], Julius J. Oppenheim[1], Hao Yang[1], Tingting Zhou[1,2], Sergey Zybin[1], Mohamed Rizk[1], and William A. Goddard III[1]*

[1]*Materials and Process Simulation Center, California Institute of Technology, Pasadena, California 91125, United States*

[2]*Institute of applied physics and computational mathematics, Fenghao Donglu, Haidian District, Beijing, 100094.*

* Corresponding author: wag@caltech.edu

This document includes:

1. Computational Details

Optimization software. We used the SciPy 0.18.1 library of the Python 2.7.13 version package for all of the optimizations that performed in the present study. In particular, we used the Nelder-Mead[1] and Sequential Least Squares Programming (SLSQP)[2] optimizer methods. The Nelder-Mead algorithm is preferred in systems where the first and second derivative of the error function cannot be used. The SLSQP algorithm utilizes the first and second derivatives to find the minimum/maximum of a function and also allows for the addition of constraints on the error function (such as restricting a variable to a positive value). In our optimizations, we switched between these two methods in order to minimize the cost function in Equation 1.

Periodic QM calculations. All calculations in this section were performed using the VASP package (version 5.3.5).[3] The generalized gradient approximation (GGA) was used for the exchange–correlation energy[4] with a plane wave energy cutoff of 600 eV. We used the convergence criteria of 10^{-6} eV for energy and 10^{-4} eV Å^{-1} for force. The k-space sampling was the gamma point centered with full space group symmetry. The CRYSTAL package was used for the B3LYP calculations on oxygen crystal.

For all elements other than O, we used the Perdew–Burke-Ernzerhof (PBE)[4] flavor of DFT while for oxygen crystal we used the Becke, three-parameter, Lee-Yang-Parr (B3LYP)[5,6] function since PBE is inaccurate for describing the O_2 triplet state.[7,8]

Nonperiodic QM calculations. All non-periodic QM computations were performed with the Jaguar software by Schrodinger.[9] We used the PBE-D3 density functional method[4,10] and cc-pVQZ(-g)++ basis set[11] for the non-periodic calculation of noble gases. We used the PBE-D3 density functional method[4,10] and the 6-31G** basis set[12,13] for the calculation of Nelfinavir.

Periodic and Nonperiodic RexPoN calculations. We have integrated both $E_{PR\text{-}LD}$ and E_{UNB} functional forms in the RexPoN FF. The RexPoN FF has been integrated in LAMMPS[14] molecular dynamics package.

2. Description of the Training Set

The training set was constructed by subtracting the total DFT-D energy for each volume of the crystal from the fixed reference structure. The reference structure is similar for all cells. The optimization was performed to minimize the error (cost) function given by

$$Error = \sum_{i=0}^{N_t} \frac{1}{w_i^2} (E_{DFT-D} - E_{PR-LD})^2 \quad (1)$$

Where N_t is the number of items in the training set and w_i is the weight for each data point in the training set. The values of $E_{DFT\text{-}D}$ and $E_{PR\text{-}LD}$ were computed using

$$E_{DFT-D} = \frac{E_{DFT-D}^{crystal}}{N_{ref}} - E_{DFT-D}^{ref} \quad (2)$$

$$E_{PR-LD} = \frac{E_{PR-LD}^{crystal}}{N_{ref}} - E_{PR-LD}^{ref} \quad (3)$$

Where N_{ref} is the number of reference units in the crystal. The equilibrium crystal structures for different elements are provided in Figure 1 of the manuscript. The reference energy (E^{ref}) was computed by placing the reference structure in a large enough cell to exclude any interactions with the images in the periodic cells. For noble gases, we used a single atom in a large cell as the reference structure. For H, N (alpha), O, and F crystals we used a single X_2 molecule (X=H, N, O, F) in a large cell. For C, Cl, Br, and I we used a layer of atoms in *bc* plane and placed them in a cell with large vacuum in *c* direction. For P and N (BP) we used a layer of connected atoms corresponding to *ac* plane and placed them in a cell with large vacuum in *b* direction.

3. Comparison between Universal Non-bonded Potential of RexPoN and Other Potentials

In this section, we compare our universal nonbonded function (E_{UNB}) with other well-known potentials including Lennard-Jones 12-6 (LJ12-6), Lennard-Jones 9-6 (LJ9-6), Morse, and the modified Buckingham exponential-6 (Be-6) potentials,

$$E_{LJ12-6}(r) = \varepsilon_{LJ12-6}\left[\left(\frac{R_e}{r}\right)^{12} - 2\left(\frac{R_e}{r}\right)^{6}\right], \quad (4)$$

$$E_{LJ9-6}(r) = \varepsilon_{LJ9-6}\left[2\left(\frac{R_e}{r}\right)^{9} - 3\left(\frac{R_e}{r}\right)^{6}\right], \quad (5)$$

$$E_{Morse}(r) = \varepsilon_M\left[e^{2\alpha_M(R_e-r)} - 2e^{\alpha_M(R_e-r)}\right], \quad (6)$$

$$E_{Be-6}(r) = \frac{\varepsilon_B}{1-\frac{6}{A_B}}\left[\frac{6}{A_B}e^{A_B\left(1-\frac{r}{R_e}\right)} - C_B\left(\frac{R_e}{r}\right)^{6}\right], \quad (7)$$

where ε_{LJ12-6}, ε_{LJ9-6}, ε_M, ε_B are the well depths and α_M, A_B, and C_B are the parameters of the corresponding potentials.

The potential functions are transformed to the dimensionless forms by setting,

$$\varepsilon^* = \frac{E}{D_e} \quad (8)$$

$$r^* = \frac{r}{R_e} \quad (9)$$

where D_e is the well depth and R_e is the distance value at the minimum of E. Using the above transformations, the following conditions would be satisfied for the given functional forms

$$r^*|_{r=R_e} = 1, \tag{10}$$

$$\varepsilon^*(r^*)|_{r^*=1} = -1, \tag{11}$$

$$\lim_{r^*\to\infty} \varepsilon^*(r^*) = 0, \tag{12}$$

$$\frac{\partial \varepsilon^*}{\partial r^*}|_{r^*=1} = 0. \tag{13}$$

Using Equations 8 and 9, the dimensionless forms of the above potentials and E_{UNB} function are given by,

$$\varepsilon^*_{LJ12-6}(r^*) = \left[\left(\frac{1}{r^*}\right)^{12} - 2\left(\frac{1}{r^*}\right)^{6}\right], \tag{14}$$

$$\varepsilon^*_{LJ9-6}(r^*) = \left[2\left(\frac{1}{r^*}\right)^{9} - 3\left(\frac{1}{r^*}\right)^{6}\right], \tag{15}$$

$$\varepsilon^*_{Morse}(r^*) = \left[e^{2\alpha^*_M(1-r^*)} - 2e^{\alpha^*_M(1-r^*)}\right], \tag{16}$$

where $\alpha^*_M = \alpha_M R_e$,

$$\varepsilon^*_{Be-6}(r^*) = \frac{1}{1-\alpha^*_B}\left[\alpha^*_B e^{\frac{6(1-r^*)}{\alpha^*_B}} - \left(\frac{1}{r^*}\right)^{6}\right], \tag{17}$$

where $\alpha^*_B = 6/A_B$,

$$\varepsilon^*_{UNB}(r^*) = -e^{\beta^*(1-r^*)}\sum_{i=0}^{5}\alpha^*_i(r^*-1)^i \tag{18}$$

where $\beta^* = \beta R_e/L$ and $\alpha^*_i = \alpha_i(R_e/L)^i$.

It might be also useful to consider the value of the second derivative (ω*) of these scaled functions at r*:

$$\omega^*_{LJ12-6} = \frac{\partial^2 \varepsilon^*_{LJ12-6}}{\partial {r^*}^2}|_{r^*=1} = 72, \tag{19}$$

$$\omega^*_{LJ9-6} = \frac{\partial^2 \varepsilon^*_{LJ9-6}}{\partial {r^*}^2}|_{r^*=1} = 54, \tag{20}$$

$$\omega^*_{Morse} = \frac{\partial^2 \varepsilon^*_{Morse}}{\partial {r^*}^2}|_{r^*=1} = 2{\alpha^*_M}^2, \tag{21}$$

$$\omega^*_{Be-6} = \frac{\partial^2 \varepsilon^*_{Be-6}}{\partial r^{*2}}|_{r^*=1} = \frac{6(6-7\alpha^*_B)}{\alpha^*_B(1-\alpha^*_B)}, \tag{22}$$

$$\omega^*_{UNB} = \frac{\partial^2 \varepsilon^*_{UNB}}{\partial r^{*2}}|_{r^*=1} = -\alpha^*_0\beta^{*2} + \alpha^*_1\beta^* - 2\alpha^*_2 = \left(\frac{R_e}{L}\right)^2(-\alpha_0\beta^2 + \alpha_1\beta - 2\alpha_2) = \left(\frac{R_e}{L}\right)^2 \tag{23}$$

Thus, for LJ12-6 and LJ9-6 potentials the scaled second derivative are always equal to 72 and 54, respectively. However, for other potentials it depends on the selected parameters. Note that in Equation 23 the scaled second derivative is equal to $(R_e/L)^2$ since $-\alpha_0\beta^2 + \alpha_1\beta - 2\alpha_2$ is always 1. Based on R_e and L values given in Table 2 of the manuscript, the ω^* of the E_{UNB} potential varies from about 32 to 62 with the average value of ~48. This means that the curvature of E_{UNB} potential at R_e falls between LJ12-6 and LJ9-6 potential.

In Figure S1 we provide the dimensionless potentials curves for the functions given in Equations 14 to 18. Here we chose $\alpha^*_M = 2.0$ and $\alpha^*_B = 3.0$ according to reference[15]. For Equation 18, we use the average value $R_e/L \approx 48$ which gives $\beta^* = 60\beta$ and $\alpha^*_i = 60^i\alpha_i$.

Figure S1a to Figure S1c provide the comparison of the scaled potential functions for different ranges of r^*. As shown the ε^* for E_{UNB} is softer than LJ12-6 and LJ9-6 but stiffer than More and Buckingham exponential-6 (Be-6) potentials.

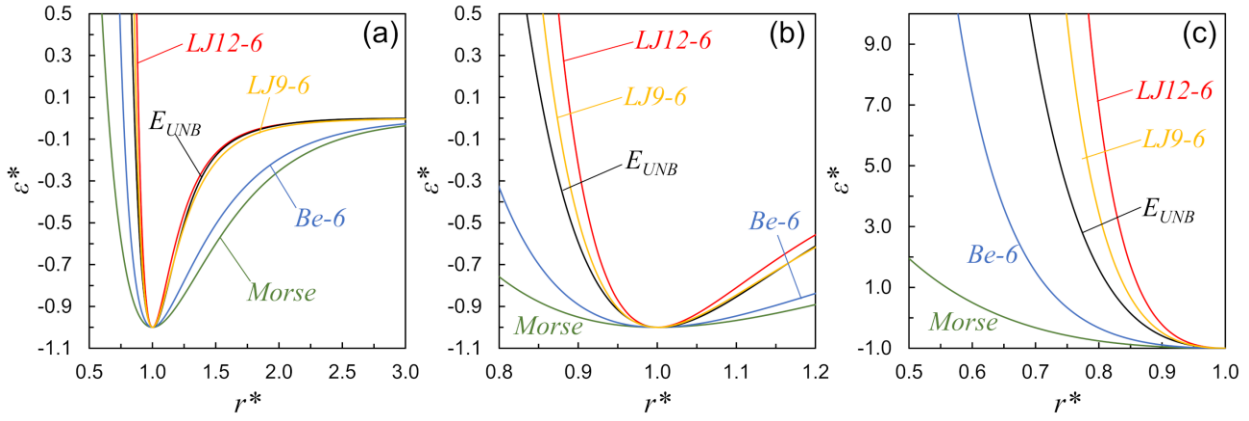

Figure S1. Comparison between the scaled energy (ε^*) of the Lennard-Jones 12-6 (LJ12-6), Lennard-Jones 9-6 (LJ9-6), Morse, modified Buckingham exponential-6 (Be-6), and universal non-bonded (E_{UNB}) potentials (Equations 14 to 18) for scaled distances (r^*) in the range of a) r^* = 0.5-3.0, b) r^* = 0.8-1.2, and c) r^* = 0.5-1.0.

4. Melting Point Calculations for Different Isotopes of Water

RexPoN provides very accurate value the melting point of normal water (H_2O) however due to isotope effects, heavy water (D_2O) has a melting temperature at 276.97 K (1 atm).[16] The question is whether RexPoN can cope with these isotope effects. Thus, we performed melting simulations of D_2O at 1 atm using the RexPoN FF. Here we used same melting simulation procedure as was explained in our in our previous work for normal H_2O.[17] Then we used the two-phase thermodynamics (2PT) methodology[18,19] to determine the melting point of heavy water. The change in the standard molar entropy (S^0) and free energy (A) as a function of temperature for normal and heavy water are shown in Figure S2. Since melting is a first-order transition it results in a sharp discontinuity and slope change in such thermodynamics properties as S^0 and A. For normal water RexPoN finds the sharp discontinuity in S^0 curve (Figure S2a) between 273.0 and 273.5 K (i.e. T_{melt} $H_2O \approx 273.3$ K) and for heavy water between 279.5 and 280.0 K (i.e. T_{melt} $H_2O \approx 279.75$ K). Also, the free energy curves of both normal and heavy water show sharp discontinuities at the melting points. These results are in good agreement with experimental values.

Standard classical simulations on water using simple potentials predict nearly identical structural and thermodynamic properties for both H_2O and D_2O, leading to similar crystal structures and melting points. We believe that the reason that RexPoN does not suffer from the usual problem in classical MD is that in PQEq the charges and polarization change every femtosecond. Since the masses are different between H_2O and D_2O they follow different pathways each femtosecond with the result that they see a different evolution in the potential surface during the dynamics. Thus, RexPoN is able to capture the fundamental physics underlying the difference in the melting points.

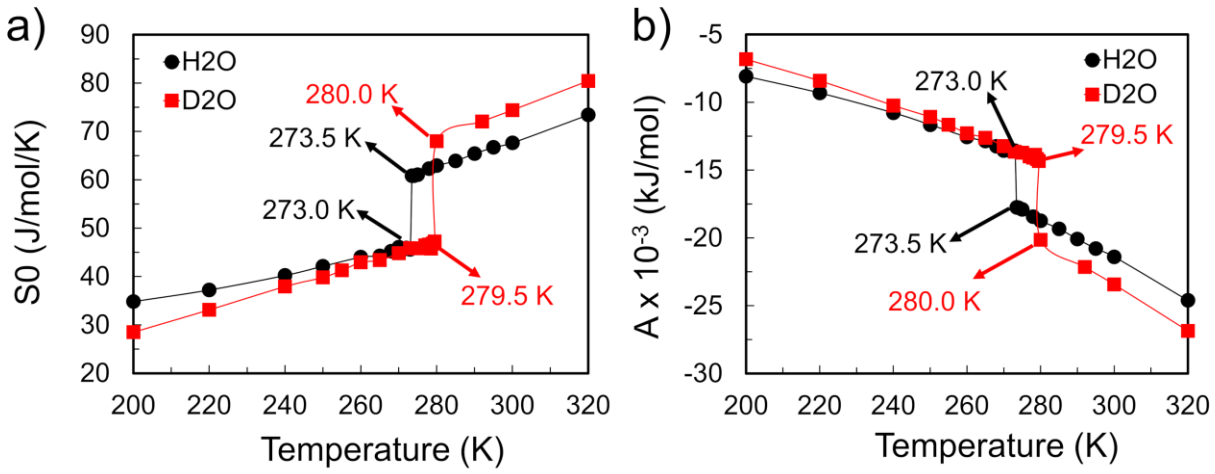

Figure S2. The (a) absolute entropy and (b) free energy of water as a function of temperature from the 2PT analysis of the MD trajectory for normal water (H_2O) and heavy (D_2O) water. We find for H_2O a sharp change between 273.0 and 273.5 K and for D_2O between 279.5 and 280.0 K, corresponding to the melting point.

5. References:

(1) Gao, F.; Han, L. Implementing the Nelder-Mead Simplex Algorithm with Adaptive Parameters. *Computational Optimization and Applications* **2012**, *51* (1), 259–277.

(2) Kraft, D. A Software Package for Sequential Quadratic Programming. *Forschungsbericht-Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt* **1988**.

(3) Kresse, G.; Furthmüller, J.; Hafner, J. *Vienna Ab Initio Simulation Package (VASP), Version 5.3. 5*.

(4) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical review letters* **1996**, *77* (18), 3865.

(5) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *The Journal of Chemical Physics* **1993**, *98* (7), 5648–5652. https://doi.org/10.1063/1.464913.

(6) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Physical review B* **1988**, *37* (2), 785.

(7) Barrett, C.; Meyer, L.; Wasserman, J. Antiferromagnetic and Crystal Structures of Alpha-Oxygen. *The Journal of Chemical Physics* **1967**, *47* (2), 592–597.

(8) Freiman, Y. A.; Jodl, H.-J. Solid Oxygen. *Physics Reports* **2004**, *401* (1–4), 1–228.

(9) Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. Jaguar: A High-Performance Quantum Chemistry Software Program with Strengths in Life and Materials Sciences. *Int. J. Quantum Chem.* **2013**, *113* (18), 2110–2142. https://doi.org/10.1002/qua.24481.

(10) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *The Journal of chemical physics* **2010**, *132* (15), 154104.

(11) 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* (2), 1007–1023. https://doi.org/10.1063/1.456153.

(12) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *The Journal of Chemical Physics* **1971**, *54* (2), 724–728.

(13) Hehre, W.; Pople, J. Self-Consistent Molecular Orbital Methods. XIII. An Extended Gaussian-Type Basis for Boron. *The Journal of Chemical Physics* **1972**, *56* (8), 4233–4234.

(14) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics* **1995**, *117* (1), 1–19. https://doi.org/10.1006/jcph.1995.1039.

(15) Xantheas, S. S.; Werhahn, J. C. Universal Scaling of Potential Energy Functions Describing Intermolecular Interactions. I. Foundations and Scalable Forms of New Generalized Mie, Lennard-Jones, Morse, and Buckingham Exponential-6 Potentials. *The Journal of chemical physics* **2014**, *141* (6), 064117.

(16) O'Neil, M. J. *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*; RSC Publishing, 2013.

(17) Naserifar, S.; Goddard III, W. A. The Quantum Mechanics-Based Polarizable Force Field for Water Simulations. *The Journal of chemical physics* **2018**, *149* (17), 174502.

(18) Lin, S.-T.; Maiti, P. K.; Goddard III, W. A. Two-Phase Thermodynamic Model for Efficient and Accurate Absolute Entropy of Water from Molecular Dynamics Simulations. *The Journal of Physical Chemistry B* **2010**, *114* (24), 8191–8198.

(19) Pascal, T. A.; Lin, S.-T.; Goddard III, W. A. Thermodynamics of Liquids: Standard Molar Entropies and Heat Capacities of Common Solvents from 2PT Molecular Dynamics. *Physical chemistry chemical physics* **2011**, *13* (1), 169–181.