Mie 16-6 force field predicts viscosity with faster than exponential pressure dependence for 2,2,4-trimethylhexane

Richard A. Messerly

Thermodynamics Research Center, National Institute of Standards and Technology, Boulder, Colorado, 80305

Michelle C. Anderson

Thermodynamics Research Center, National Institute of Standards and Technology, Boulder, Colorado, 80305

S. Mostafa Razavi

Department of Chemical and Biomolecular Engineering, The University of Akron, Akron, Ohio, 44325-3906

J. Richard Elliott

Department of Chemical and Biomolecular Engineering, The University of Akron, Akron, Ohio, 44325-3906

Abstract

In response to the 10$^{\text{th}}$ Industrial Fluid Properties Simulation Challenge, we report viscosity (η) estimates of 2,2,4-trimethylhexane at 293 K for a range of pressures (P) from 0.1 MPa to 1000 MPa. The Potoff force field is utilized in this study, as a previous study demonstrated that it provides reliable estimates of η with respect to P. Whereas most studies report only the uncertainties associated with random fluctuations in the simulation output, we investigate the effect of uncertainties arising from the force field non-bonded and torsional parameters. The pressure-viscosity coefficient as a function of pressure is reported for several different empirical models, namely, McEwen-Paluch, Roelands, Roelands-Modified and Barus. Although the uncertainties increase substantially with increasing pressure, cross-validation model selection provides quantitative evidence supporting so-called super-Arrhenius behavior with an inflection point in a $\log_{10}(\eta)$-P plot around 200 MPa.

Email addresses: `richard.messerly@nist.gov` (Richard A. Messerly), `michelle.anderson@nist.gov` (Michelle C. Anderson), `sr87@zips.uakron.edu` (S. Mostafa Razavi), `elliot1@uakron.edu` (J. Richard Elliott)

Keywords:
Pressure-viscosity coefficient, Industrial Fluid Properties Simulation Challenge, Uncertainty Quantification, Molecular Dynamics Simulation

1. Introduction

The Industrial Fluid Properties Simulation Challenge (IFPSC) is an open international competition aimed at aligning the molecular simulation community, which is primarily academic, with the goals of industrial research. The present work is a submission to the 10^{th} Industrial Fluid Properties Simulation Challenge (IFPSC10), which challenges modelers to predict the viscosity (η) of 2,2,4-trimethylhexane (224TMH) over a wide range of pressures (P), specifically, from 0.1 MPa (atmospheric) to 1000 MPa, at a constant temperature (T) of 293 K.

The practical application of IFPSC10 is elastohydrodynamic lubrication (EHL), where knowledge of the pressure-viscosity relationship is paramount. The challenge compound was chosen as an ideal lubricating oil candidate for which no published experimental viscosity data are available above ambient pressure. New experimental measurements have been performed by Scott Bair of Georgia Tech but have not been made available at this time. The sample purity is greater than 98 % and the estimated experimental uncertainties for η, T, and P are, respectively, 3 %, 0.3 K, and the greater of 1 MPa and 0.4 %.

Classical film thickness formulas rely heavily on the so-called pressure-viscosity coefficient (α), which is essentially an Arrhenius-like activation parameter that is obtained from the slope of an $\log_{10}(\eta)$-P plot. However, faster-than-exponential, a.k.a. super-Arrhenius, dependence on pressure has been observed through experimental viscometry measurements for nearly a century [1]. This super-Arrhenius trend is typically manifest by an inflection point in the $\log_{10}(\eta)$-P plot at high pressures. While this behavior is common in experimental measurements, we are not aware of any rheological molecular simulation studies that have addressed this topic, as most simply assume an Arrhenius relationship when reporting α [2, 3, 4]. IFPSC10 is an ideal opportunity to demonstrate whether or not molecular simulation can provide evidence supporting or opposing the existence of super-Arrhenius behavior.

In a previous study [5], we investigated the accuracy of predicting viscosity with four different united-atom (UA) Mie λ-6 (generalized Lennard-Jones, LJ) force fields, namely, the Transferable Potentials for Phase Equilibria (TraPPE-UA [6, 7, 8]), Transferable Anisotropic Mie (TAMie) [9, 10], Potoff [11, 12], and fourth generation anisotropic-united-atom (AUA4) [13, 14]. The comparisons with experimental data were made for saturated liquid viscosity ($\eta_{\text{liq}}^{\text{sat}}$) over a wide temperature range and compressed liquid viscosity ($\eta_{\text{liq}}^{\text{comp}}$) at 293 K from atmospheric pressure to 1000 MPa, i.e., the same temperature and pressure range as the challenge. The compounds in question were *n*-alkanes ranging in length from ethane to *n*-docosane and branched alkanes ranging in size from 2-methylpropane to 2,2,4-trimethylpentane (224TMP). The 224TMP results at high pressures are especially useful as this compound is a close analogue to the challenge compound and, in contrast with 224TMH, 224TMP has been well studied experimentally.

While TraPPE and AUA4 (LJ 12-6 based potentials) under predict $\eta_{\text{liq}}^{\text{sat}}$ by 20 % to 50 % for all compounds studied, TAMie (Mie 14-6) and Potoff (Mie 16-6) predict $\eta_{\text{liq}}^{\text{sat}}$ within 10 % for most compounds. For $\eta_{\text{liq}}^{\text{comp}}$, TAMie is the most reliable at predicting the viscosity-density dependence, while Potoff significantly over estimates $\eta_{\text{liq}}^{\text{comp}}$ with respect to density. However, since Potoff also over estimates pressure at high densities [15], the viscosity-pressure trend for Potoff is remarkably accurate even at pressures approaching 1000 MPa. In particular, the Potoff force field predicts the viscosity-pressure trend for 224TMP to within 10 % accuracy. For this reason, we implement the Potoff Mie 16-6 force field to predict η and α for the challenge compound. We should note, however, that our previous study did not provide any definitive evidence that the Potoff force field could predict a super-Arrhenius trend for any compound studied, including for 224TMP.

One of the entry guidelines for IFPSC is "an analysis of the uncertainty in the calculated results." Traditionally, simulation uncertainties are limited to the random fluctuations of simulation output and/or the uncertainty related to data post-processing. This class of uncertainty is referred to as “numerical uncertainty” (frequently referred to as “statistical uncertainty”) [16, 17, 18, 19]. Two other classes of uncertainty, namely, “parameter uncertainty” and “functional form uncertainty” (also referred to as “model uncertainty”) are typically ignored in uncertainty quantification (UQ) due to the increased computational cost [16, 17, 18, 19]. The latter refers to the uncertainty associated with the choice of force field functional form, while the former refers to the uncertainty in the force field parameters for a given force field functional form.

Quantifying the functional form uncertainty is an extremely difficult task, as it often requires testing numerous force field functional forms. For this reason, we focus on numerical and parameter uncertainties without addressing functional form uncertainties. Specifically, we apply bootstrap re-sampling [20] and Bayesian inference Markov Chain Monte Carlo (MCMC) [16, 15] to quantify numerical and parameter uncertainties, respectively. The chosen functional form is the same as the Potoff force field, namely, a united-atom, fixed bond length, harmonic angular potential, Fourier series torsional potential, and a Mie 16-6 non-bonded potential (see Section 2.1 for details). As viscosity is highly sensitive to the non-bonded [21, 5] and torsional [22, 23] potentials, we limit our parameter uncertainty investigation to the non-bonded and torsional parameters.

The outline for the present work is the following. Section 2 explains the force field, parameter uncertainties, simulation methodology, data analysis, and empirical models. Section 3 presents the simulation results, with an emphasis on uncertainty quantification. Section 4 discusses some important observations and limitations. Section 5 recaps the primary conclusions from this work.

2. Methods

2.1. Force field

We utilize the Potoff force field as it provides reliable estimates of the η-P dependence for normal and branched alkanes that are similar to the challenge compound [5]. In addition, we quantify the uncertainty in η that arises from uncertainties in the non-bonded Mie

16-6 and torsional parameters. The parameter uncertainties are obtained using Bayesian inference Markov Chain Monte Carlo (MCMC). This UQ analysis is performed sequentially. First, we account for only the non-bonded uncertainties (referred to as MCMC-nb). Then, we include both the non-bonded and torsional uncertainties (MCMC-nb-tors). This sequential approach provides insight into which source of uncertainty has a greater impact on η.

2.1.1. Potoff force field

The Potoff Mie λ-6 force field utilizes united-atom (UA) sites, where 2,2,4-trimethylhexane is represented with CH_3, CH_2, CH, and C UA sites. Neighboring UA sites are separated by a fixed 0.154 nm bond length. Note that we observed in our previous study that η increases by several percent when flexible bonds are employed instead of fixed bonds. Therefore, the choice of fixed bonds was not arbitrary and is a possible source of uncertainty for which we did not rigorously account. The primary reason we utilize fixed bonds is to reduce the fluctuations in the stress tensor and, thereby, improve the precision of the viscosity estimate.

The angular contribution to energy is computed using a harmonic potential:

$$u^{\mathrm{bend}} = \frac{k_\theta}{2} \left(\theta - \theta_0\right)^2 \tag{1}$$

where u^{bend} is the bending energy, θ is the instantaneous bond angle, θ_0 is the equilibrium bond angle (see Table 1), and k_θ is the harmonic force constant with $k_\theta / k_{\mathrm{B}} = 62500\ \mathrm{K/rad}^2$ for all bonding angles, where k_{B} is the Boltzmann constant.

Table 1: Equilibrium bond angles (θ_0) [7, 12]. CH_i and CH_j represent CH_3, CH_2, CH, or C sites.

Bending sites	θ_0 (degrees)
CH_i-CH_2-CH_j	114.0
CH_i-CH-CH_j	112.0
CH_i-C-CH_j	109.5

Dihedral torsional interactions are determined using a modified cosine series:

$$\begin{aligned} u^{\text{tors}} &= c_0 + c_1[1+\cos\phi] + c_2[1-\cos 2\phi] + c_3[1+\cos 3\phi] + A_{\text{s}}\sin^2\left[\frac{3}{2}(\phi+180^\circ)\right] \\ &= (c_0 - A_{\text{s}}) + c_1[1+\cos\phi] + c_2[1-\cos 2\phi] + \left(c_3 + \frac{A_{\text{s}}}{2}\right)[1+\cos 3\phi] \end{aligned} \tag{2}$$

where u^{tors} is the torsional energy, ϕ is the dihedral angle, c_n are the Fourier constants used in the Potoff force field and listed in Table 2, and $A_{\text{s}}\sin^2\left[\frac{3}{2}(\phi+180^\circ)\right]$ is an additional term proposed by Nieto-Draghi et al. to shift the torsional barrier heights for normal and branched alkanes [22, 14]. We follow a convention similar to that of the International Union of Pure and Applied Chemistry (IUPAC) such that $\phi = 180^\circ$ for the *trans* conformation [7], whereas Nieto-Draghi et al. define the *trans* conformation as 0° or 360° [22, 14], hence the $\phi + 180^\circ$ term in Equation 2. As $\sin^2\left[\frac{3}{2}(\phi+180^\circ)\right]$ has a maximum value of 1 at 0°, 120°, 240°, and 360°, the torsional barriers located at these dihedral angles increase by A_{s}. By contrast, this additional term does not shift u^{tors} for dihedral angles of 60°, 180°, and 300°, which correspond to the equilibrium conformations of *gauche*$^-$, *trans*, and *gauche*$^+$, respectively. Clearly, the non-shifted Potoff torsional potential is obtained only when $A_{\text{s}} = 0$. The actual reason we include this additional torsion term is to provide a simple method for quantifying the uncertainty in the torsional potential (see Section 2.1.2).

Table 2: Fourier constants (c_n/k_{B}) and shifting parameter $(A_{\text{s}}/k_{\text{B}})$ in units of K for Potoff force field [7, 12]. CH_i and CH_j represent CH_3, CH_2, CH, or C sites.

Torsion sites	c_0/k_{B}	c_1/k_{B}	c_2/k_{B}	c_3/k_{B}	$A_{\text{s}}/k_{\text{B}}$
CH_i-CH_2-CH-CH_j	-251.06	428.73	-111.85	441.27	0.0
CH_i-CH_2-C-CH_j	0.0	0.0	0.0	461.29	0.0

Non-bonded interactions between sites located in two different molecules or separated by more than three bonds within the same molecule are calculated using a Mie λ-6 potential (of which the traditional Lennard-Jones, LJ, 12-6 is a subclass) [24]:

$$u^{\mathrm{vdw}}(\epsilon, \sigma, \lambda; r) = \left(\frac{\lambda}{\lambda - 6}\right)\left(\frac{\lambda}{6}\right)^{\frac{6}{\lambda-6}} \epsilon \left[\left(\frac{\sigma}{r}\right)^{\lambda} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{3}$$

where u^{vdw} is the van der Waals interaction, σ is the distance (r) where $u^{\mathrm{vdw}} = 0$, $-\epsilon$ is the energy of the potential at the minimum $\left(\text{i.e., } u^{\mathrm{vdw}} = -\epsilon \text{ and } \frac{\partial u^{\mathrm{vdw}}}{\partial r} = 0 \text{ for } r = r_{\mathrm{min}}\right)$, and λ is the repulsive exponent.

The non-bonded Potoff Mie λ-6 force field parameters are provided in Table 3. Note that Potoff reports a "generalized" and "short/long" (S/L) CH and C parameter set. The "generalized" CH and C parameter set is an attempt at a completely transferable force field, while the "short" and "long" parameters are implemented when the number of carbons in the backbone is ≤ 4 and > 4, respectively. As the longest continuous carbon backbone for 2,2,4-trimethylhexane consists of six carbons, the Potoff results presented in this study are obtained with the "long" parameters.

Table 3: Non-bonded Potoff Mie λ-6 parameters [11, 12]. The CH and C values are the "long" Potoff parameters.

	Potoff		
United-atom	ϵ/k_{B} (K)	σ (nm)	λ
CH_3	121.25	0.3783	16
CH_2	61	0.399	16
CH	14	0.47	16
C	1.2	0.62	16

Non-bonded parameters between two different site types (i.e., cross-interactions) are determined using Lorentz-Berthelot combining rules [25] for ϵ and σ and an arithmetic mean for the repulsive exponent λ (as recommended in Reference 11):

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{4}$$

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{5}$$

$$\lambda_{ij} = \frac{\lambda_{ii} + \lambda_{jj}}{2} \tag{6}$$

where the ij subscript refers to cross-interactions and the subscripts ii and jj refer to same-site interactions.

2.1.2. MCMC parameter uncertainty

Nieto-Draghi et al. set A_s equal to 40% and 15% of the maximum dihedral barrier (the *cis* conformation) for the terminal and internal torsions, respectively [22, 14]. For example, this corresponds to $A_\mathrm{s}/k_\mathrm{B} \approx 1000$ K and ≈ 375 K for the CH_3-CH_2-CH_2-CH_2 and CH_2-CH_2-CH_2-CH_2 torsional potentials, respectively. The reason why Nieto-Draghi et al. increase the torsional barrier, i.e., $A_\mathrm{s} > 0$, is because AUA4 under predicts η by approximately 20 % to 40 %. However, despite the relatively large increase in the torsional barriers, the modified force field (AUA4m) typically provides only marginal improvement of 5 % to 15 % compared to AUA4 (see Tables 4 and 5 of Reference 14).

As the Potoff Mie 16-6 potential is already quite reliable for predicting viscosity, we would expect significant over prediction of viscosity if we coupled the Potoff Mie 16-6 potential with $A_\mathrm{s}/k_\mathrm{B} \gg 0$. Thus, unlike Nieto-Draghi et al., we do not propose that the torsional barriers must be increased unilaterally. Instead, we assume that A_s follows a skewed distribution with a mean value near zero and the lower and upper 95 % confidence intervals correspond to -15 % and +40 % of the maximum barrier height for the non-shifted Potoff torsional potential. The MCMC-nb-tors parameters are sampled from this distribution. The rationale for the A_s distribution is presented in Supporting Information.

Figure 1 compares the non-shifted Potoff torsional potential, ± 40 % shift in barrier heights, and the MCMC-nb-tors potentials. The insets also depict the skewed distributions and the randomly sampled MCMC A_s sets. Note that the challenge compound consists of four $CH_i-CH_2-CH-CH_j$ torsions and three $CH_i-CH_2-C-CH_j$ torsions. Note that, unlike Nieto-Draghi et al., we make no distinction between internal and terminal torsions.

Figure 2 depicts the MCMC non-bonded parameters for CH_3, CH_2, CH, and C united-

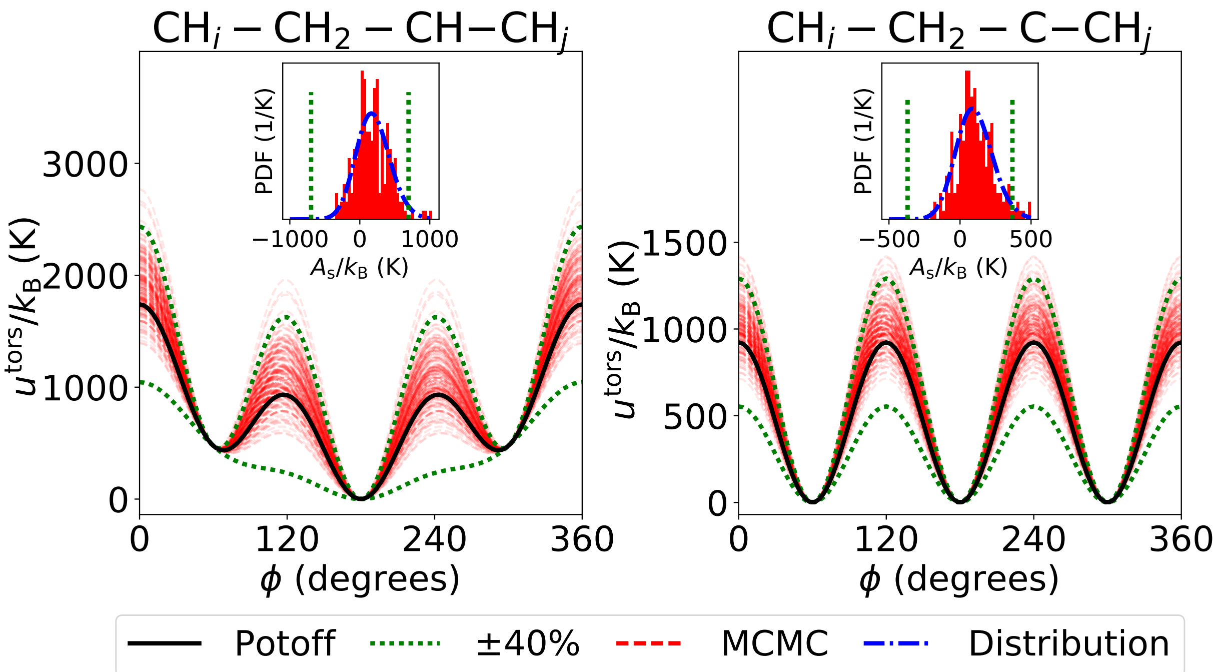

Figure 1: Comparison of Potoff (black solid lines), ± 40 % (green dotted lines), and MCMC-nb-tors (red dashed lines) torsional potentials. Insets show the distribution for A_s as blue dash-dotted lines. Left and right panels correspond to $CH_i{-}CH_2{-}CH{-}CH_j$ and $CH_i{-}CH_2{-}C{-}CH_j$ torsions, respectively. Both u^{tors}/k_B and A_s/k_B are expressed in units of K.

atom sites (ϵ_{CH_3}, σ_{CH_3}, ϵ_{CH_2}, σ_{CH_2}, ϵ_{CH}, σ_{CH}, ϵ_C, and σ_C) which are used for MCMC-nb and MCMC-nb-tors. Note that $\lambda_{CH_3} = \lambda_{CH_2} = \lambda_{CH} = \lambda_C = 16$. Parameters are assumed to be transferable, e.g., the CH_2 MCMC parameters do not depend on the CH_3 MCMC parameters. The MCMC non-bonded parameters for CH_3 and CH_2 sites were reported previously [15]. These parameters were obtained using a likelihood function based on saturated liquid density and saturated vapor pressure data for ethane, propane, *n*-butane, and *n*-octane. By contrast, the MCMC parameters for CH and C sites were obtained from the scoring function reported by Mick et al. [12] that depends on several vapor-liquid coexistence properties for a diverse set of branched alkanes. Details regarding the generation of MCMC parameter sets from the scoring function are found in Supporting Information. An important observation from Figure 2 is that the ranges of MCMC sampled CH and C non-bonded parameters are considerably larger (on a percent basis) than those for CH_3 and CH_2.

2.2. Simulation set-up

Historically, non-equilibrium molecular dynamics (NEMD) has been preferred for highly viscous systems [3, 4]. However, in our recent publication we successfully predicted the viscosity of 2,2,4-trimethylpentane at 293 K and 1000 MPa (the highest pressure required for the challenge) with equilibrium molecular dynamics (EMD). Consistent with our previous study, we perform EMD simulations using GROMACS version 2018 with "mixed" (single and double) precision [26]. Example GROMACS input files (.top, .gro. and .mdp) with corresponding shell and Python scripts for preparing, running, and analyzing simulations are provided as Supporting Information.

We utilize the same simulation specifications as our previous study [5]. The general simulation specifications are provided in Table 4. Note that we employ a 1.4 nm non-bonded cut-off with analytical tail corrections (where GROMACS only includes the contribution from the r^{-6} term) instead of the 1.0 nm cut-off utilized by Potoff's group [11, 12]. Although a 1.0 nm cut-off is reliable for most compounds, it can lead to spurious viscosity estimates for larger molecules. For example, our previous study demonstrates that, with a 2 fs time-step, a cut-off distance of 1.4 nm or greater is required for *n*-alkanes longer than *n*-dodecane. For this reason, we incur the additional computational cost with the longer cut-off so as to avoid any unforeseen simulation anomalies. Our previous study also shows that finite size effects are negligible for a 200 or 400 molecule system. We utilize the larger system size only for $P \leq 500$ MPa, while the smaller system size is favored for the longer (higher pressure) simulations.

We perform a sequence of six simulation stages: energy minimization, NPT equilibration, NPT production, energy minimization, NVT equilibration, and NVT production. The average box size from the NPT production stage is utilized in the second energy minimization and subsequent NVT stages. Table 5 lists the integrators, thermostats, barostats, and simulation time used for each NPT and NVT equilibration and production stage. These specifications are also the same as our previous study, with the exception

Table 4: General simulation specifications.

Time-step (fs)	2
Cut-off length (nm)	1.4
Tail-corrections	U and P
Constrained bonds	LINCS [27, 28]
LINCS-order	8
Number of molecules	200 or 400

of the NVT production simulation times, which are state point dependent. The specific production times for the NVT production stage are provided in Table 6.

Table 5: Simulation specifications for equilibration (Equil.) and production (Prod.) stages. t_{sim} is the simulation time, τ_T is the thermostat time-constant, τ_P is the barostat time-constant, and ζ_P is the barostat compressibility.

	NPT Equil.	NPT Prod.	NVT Equil.	NVT Prod.
t_{sim} (ns)	1	1	1	1 to 48
Integrator	Leap frog [30]	Leap frog	Velocity Verlet [29]	Velocity Verlet
Thermostat	Velocity rescale [31]	Nosé-Hoover [32, 33]	Nosé-Hoover	Nosé-Hoover
τ_T (ps)	1.0	1.0	1.0	1.0
Barostat	Berendsen [34]	Parrinello-Rahman [35, 36]	N/A	N/A
τ_P (ps)	1.0	5.0	N/A	N/A
ζ_P (1/bar)	4.5E-5	4.5E-5	N/A	N/A

A large number of replicate simulations are required at each state point to improve the precision and to provide more rigorous estimates of uncertainty [37, 38]. We utilize between 40 and 80 independent replicates, where a greater number of replicates are needed for more viscous systems, i.e. at higher pressures. Each Potoff replicate simulation utilizes the same force field parameters, whereas MCMC-nb and MCMC-nb-tors utilize different parameter sets for each replicate (see Figures 1 and 2). To ensure independence between

Table 6: State point specific production times. Pressure is prescribed only in NPT equilibration and production stages.

Pressure (MPa)	NVT Prod. time (ns)
0.1	1
25	1
50	1
100	1
150	1
250	2
400	4
500	8
600	8
700	16
800	32
900	32
1000	48

replicates, the entire series of simulation stages is repeated. The energy minimization stages start with a different pseudo-random configuration and the initial velocities are randomized for the NPT and NVT equilibration stages. The replicates, therefore, sample from a distribution of densities (i.e., box sizes) for a given pressure.

2.3. *Data analysis*

The post-simulation data analysis is identical to that prescribed in our previous study [5]. In brief, we implement the Green-Kubo "time-decomposition" analysis [37, 38]

$$\eta(t) = \frac{V}{k_{\mathrm{B}} T N_{\mathrm{reps}}} \sum_{n=1}^{N_{\mathrm{reps}}} \int_0^t dt' \, \langle \tau_{\alpha\beta,n}(t') \tau_{\alpha\beta,n}(0) \rangle_{t_0,\alpha\beta} \tag{7}$$

where t is time, V is volume, N_{reps} is the number of independent replicate simulations, α and β are x, y, or z Cartesian coordinates, $\tau_{\alpha\beta,n}$ is the α-β off-diagonal stress tensor element for the n^{th} replicate, and $\langle \cdots \rangle_{t_0,\alpha\beta}$ denotes an average over twelve different time origins (t_0) and all three unique off-diagonal $\tau_{\alpha\beta}$ components. For a precise integration of Equation 7, $\tau_{\alpha\beta,n}$ is recorded every 6 fs (3 time-steps).

The force field viscosity is obtained by evaluating Equation 7 as $t \to \infty$, i.e., the infinite-time-limit viscosity (η^{∞}). As the long-time tail does not converge, we fit the "running integral" to a double-exponential function

$$\eta(t) = A\alpha\tau_1 \left(1 - \exp\left(-t/\tau_1\right)\right) + A(1-\alpha)\tau_2 \left(1 - \exp\left(-t/\tau_2\right)\right) \tag{8}$$

where A, α, τ_1, and τ_2 are fitting parameters and $\eta^{\infty} = A\alpha\tau_1 + A(1-\alpha)\tau_2$. See our previous work for details regarding the fitting procedure [5].

The uncertainty in η is obtained by bootstrap re-sampling and reported at the 95 % confidence level. Specifically, the fitting of Equation 8 is repeated hundreds of times using randomly selected subsets of replicate simulations from which we obtain a distribution of η^{∞} estimates. Section SI.III of Supporting Information validates this process for the MCMC non-bonded CH_3 parameter sets with ethane.

2.4. *Pressure-viscosity coefficient*

The simulated η values with respect to P are fit to four different empirical models from which the pressure-viscosity coefficient (α) is calculated

$$\alpha \equiv \frac{d\ln(\eta)}{dP} \tag{9}$$

As viscosity ranges over several orders of magnitude, the objective function for fitting is the sum-squared-error of $\log_{10}(\eta)$. Only the MCMC-nb-tors values are included in the fit, as these simulation results account for more sources of uncertainty than the Potoff and MCMC-nb results.

The first empirical model we implement is the traditional Barus expression [39]

$$\eta = \eta_0 \exp(\alpha P) \tag{10}$$

where the fitting parameters η_0 and α correspond to the zero pressure viscosity and the pressure-viscosity coefficient, respectively. The second model is the popular Roelands equation [40]

$$\eta = \eta_p \left(\frac{\eta_0}{\eta_p}\right)^{\left(\frac{P_p - P}{P_p}\right)^Z} \tag{11}$$

where η_0 and Z are fitting parameters and $\eta_p = 6.31 \times 10^-5$ Pa-s and $P_p = -0.196$ GPa. The third model is an alternative form of Equation 11 where η_p and P_p are additional fitting parameters (rather than fixed constants). This four parameter Roelands equation, which we refer to as "Roelands-Modified," is more flexible than the standard Roelands model and can predict super-Arrhenius behavior. The fourth model is a hybrid McEwen-Paluch expression [1]

$$\eta = \eta_0 \left(1 + \frac{a_0}{q} P\right)^q \exp\left(\frac{C_\mathrm{F} P}{P_\infty - P}\right) \tag{12}$$

where η_0, a_0, q, C_F, and P_∞ are fitting parameters. Note that, although the Roelands-Modified equation can predict super-Arrhenius behavior, only the McEwen-Paluch model is capable of representing an inflection point, i.e., a transition from Arrhenius to super-Arrhenius behavior. By contrast, the Barus and Roelands models are only capable of fitting sub-Arrhenius and Arrhenius-like data.

3. Results

Table 7 provides tabulated values for viscosity (η) and density (ρ) for Potoff, MCMC-nb, and MCMC-nb-tors at the prescribed temperature (T) and pressure (P). The uncertainties (expressed at the 95 % confidence level) for T, ρ, and P are computed with the standard deviation of the replicate simulation averages, while those for η are obtained from bootstrap re-sampling. Note that the average temperature slightly decreases with increasing pressure, but T is always within the 0.3 K uncertainty stipulated by the IF-PSC10 experimental measurements. Also, the average pressures tend to be lower than the prescribed pressure. For most state points, however, the combined uncertainty in pressure from experiment (the greater of 1 MPa and 0.4 %) and simulation is sufficient to account for this discrepancy.

Table 7: Simulation results for Potoff, MCMC-nb, and MCMC-nb-tors. Subscripts and superscripts (when present) correspond to the lower and upper 95 % confidence intervals, respectively. If only a subscript is present the lower and upper uncertainties are approximately equal (i.e., the bootstrap distribution is essentially symmetric) and only the average uncertainty is reported.

T (K)	ρ (kg/m^3)	P (MPa)	η (10^{-3} Pa-s)		
			Potoff	MCMC-nb	MCMC-nb-tors
$292.854_{0.095}$	$716.46_{0.46}$	$-2.08_{0.83}$	$0.605_{0.045}$	$0.555_{0.016}$	$0.572_{0.025}$
$292.851_{0.085}$	$734.67_{0.44}$	$22.9_{1.1}$	$0.753^{0.078}_{0.098}$	$0.743_{0.029}$	$0.751_{0.028}$
$292.847_{0.089}$	$749.22_{0.45}$	$47.9_{1.2}$	$0.957^{0.076}_{0.061}$	$0.970_{0.037}$	$0.932_{0.033}$
$292.85_{0.10}$	$772.24_{0.37}$	$98.3_{1.3}$	$1.60_{0.14}$	$1.43^{0.066}_{0.21}$	$1.426_{0.065}$
$292.838_{0.088}$	$790.25_{0.35}$	$148.3_{1.6}$	$2.18^{0.24}_{0.19}$	$2.134_{0.098}$	–
$292.819_{0.096}$	$818.22_{0.31}$	$248.5_{2.0}$	$4.56_{0.45}$	$4.28^{0.34}_{0.52}$	–
$292.816_{0.090}$	$849.17_{0.38}$	$398.7_{2.7}$	$13.5_{1.2}$	$12.61_{0.79}$	–
$292.790_{0.099}$	$865.64_{0.49}$	$498.4_{4.6}$	$28.0_{2.9}$	–	–
$292.774_{0.055}$	$879.9_{1.2}$	$598.5_{9.5}$	70_{10}	–	–
$292.763_{0.041}$	$892.6_{1.2}$	698_{10}	164_{21}	–	–
$292.752_{0.026}$	$903.5_{1.1}$	795_{11}	450^{42}_{25}	–	–
$292.738_{0.022}$	$912.7_{1.0}$	882_{11}	1096^{165}_{125}	–	–
$292.725_{0.020}$	$921.2_{1.1}$	973_{16}	3200_{669}	–	–

The values from Table 7 are also depicted in Figure 3 along with the model fits to the MCMC-nb-tors values and the available experimental viscosity at atmospheric pressure [41]. An inflection point, suggesting super-Arrhenius behavior, is observed in the top panel of Figure 3 around 200 MPa. The bottom panel of Figure 3 presents the percent deviation between the McEwen-Paluch empirical model fit and the simulation results, experimental data point, and the other three empirical model fits. This panel is useful for comparing the Potoff, MCMC-nb, and MCMC-nb-tors uncertainties and for quantifying the discrepancy between the various empirical models.

Recall that the Potoff results only account for numerical uncertainties, MCMC-nb accounts for numerical and non-bonded parameter uncertainties, and MCMC-nb-tors accounts for numerical, non-bonded and torsional parameter uncertainties. Notice in Table 7 and Figure 3 that the Potoff, MCMC-nb, and MCMC-nb-tors uncertainties are approximately the same. This somewhat surprising result supports the conclusion that the non-bonded and torsional uncertainties are negligible compared to the numerical uncertainties in the Green-Kubo viscosity.

Figure 4 presents the predicted pressure-viscosity coefficient (α), as determined by fitting the MCMC-nb-tors results to Equations 10, 11, and 12. The uncertainties in α are obtained with bootstrap re-sampling for the empirical model fits. Note that the α magnitudes for all empirical models are reasonable (i.e., similar in magnitude to other lubricants [2, 3, 4]) over the entire range of pressures. As expected, the traditional Barus α value is constant with respect to pressure. By contrast, the Roelands α value decreases with increasing pressure, while the Roelands-Modified α value increases with respect to pressure but without a change from negative to positive slope. Only the McEwen-Paluch α value shows the marked change in slope which corresponds to an inflection point in the $\log_{10}(\eta)$-P plot.

Although the hybrid McEwen-Paluch model clearly reproduces the simulation results with lower deviations than those of the Roelands and Barus models (see Figures 3 and 5), this should be anticipated considering the McEwen-Paluch model has five fitting parameters while the Barus and Roelands models only have two. Note that the four parameter Roelands-Modified model also has lower deviations than the Roelands and Barus models. Therefore, it is possible that the McEwen-Paluch model is actually over fit to our simulation results.

To assess this possibility, Figure 5 presents the cross-validation results for each model. Specifically, we implement a Monte Carlo cross-validation scheme where thousands of random sub-samples are selected for the training and testing set. Approximately 70 % of the MCMC-nb-tors simulation results (9 state points) are included in the training set while 30 % (4 state points) are excluded for the testing set. The complete set of η values also varies for each round of cross-validation according to the η bootstrap uncertainties.

The left panel of Figure 5 demonstrates that the mean-squared-error (MSE) for the training set is approximately equal to the MSE for the testing set of each model, suggesting that none of the models are over fit to the data. Note that, although somewhat counter intuitive, $\mathrm{MSE}_{\mathrm{test}} < \mathrm{MSE}_{\mathrm{train}}$ typically denotes that the training set was "easier" than the testing set. The right panel shows that only the Roelands-Modified model has a similar MSE to that of the McEwen-Paluch model for the same testing set. As both the Roelands-Modified and McEwen-Paluch models predict super-Arrhenius behavior, there is strong statistical evidence that the Potoff force field predicts super-Arrhenius behavior. However, as the Roelands-Modified model does not predict that the slope of α with respect to P changes signs, we cannot conclude whether an inflection point precedes the super-Arrhenius region.

Table 8 is included to facilitate scoring our entry for the 10^{th} Industrial Fluid Properties Simulation Challenge. Table 8 provides "smoothed" η and α values calculated with the McEwen-Paluch fit to our simulation results. The uncertainties reflect both the simulation and experimental pressure uncertainties. Note that the uncertainty in η at 1000 MPa is considerably larger than other state points. As the simulation pressure is nearly 30 MPa lower than the prescribed 1000 MPa (see Table 7), the increased uncertainty in the McEwen-Paluch fit is caused by this extrapolation to 1000 MPa.

4. Discussion

It is surprising that both the non-bonded and torsional parameter uncertainties are negligible compared to the numerical uncertainties in η (see Figure 3). A possible explanation for why the non-bonded parameter uncertainty has a negligible impact on η is that the CH_3 uncertainties are considerably smaller than those for CH_2, CH, and C (see Figure 2). As 224TMH consists primarily of CH_3 sites, the larger uncertainties in CH_2, CH, and C appear to not affect the results significantly.

Table 8: Smoothed simulation results for the purpose of scoring our entry to the 10$^{\text{th}}$ Industrial Fluid Properties Simulation Challenge. Uncertainties are expressed at the 95 % confidence level. Pressure uncertainties are those reported for the experimental measurements.

P (MPa)	η (10^{-3} Pa-s)	α (1/GPa)
$0.1_{1.0}$	$0.605_{0.045}$	$13.66_{0.32}$
$25.0_{1.0}$	$0.790_{0.011}$	$10.380_{0.055}$
$50.0_{1.0}$	$1.004_{0.012}$	$8.964_{0.088}$
$100.0_{1.0}$	$1.517_{0.020}$	$7.769_{0.072}$
$150.0_{1.0}$	$2.209_{0.033}$	$7.326_{0.036}$
$250.0_{1.0}$	$4.539_{0.069}$	$7.174_{0.047}$
$400.0_{1.6}$	$13.68_{0.28}$	$7.623_{0.073}$
$500.0_{2.0}$	$30.08_{0.86}$	$8.160_{0.074}$
$600.0_{2.4}$	$70.3_{2.6}$	$8.851_{0.070}$
$700.0_{2.8}$	$177.6_{8.0}$	$9.70_{0.10}$
$800.0_{3.2}$	493_{27}	$10.74_{0.18}$
$900.0_{3.6}$	1531_{114}	$11.99_{0.31}$
$1000.0_{4.0}$	5466_{587}	$13.51_{0.49}$

By contrast, no clear explanation exists for why the torsional parameter uncertainties do not affect the overall uncertainty in η. Previous studies suggest that a 15 % to 40 % increase in the torsional barriers should increase the viscosities appreciably for similar compounds [22, 14]. However, we did not observe such a strong dependence. Specifically, the required increase in the torsional barriers was between 80 % and 100 % to achieve an increase of approximately 10 % in viscosity (see Supporting Information). We propose that the discrepancy between our findings and those of Nieto-Draghi et al. arises as a consequence of differences between the united-atom sites and Mie 16-6 potential employed in this work and the anisotropic-united-atom sites and Lennard-Jones 12-6 potential employed by Nieto-Draghi et al.

Although the Potoff force field demonstrates super-Arrhenius behavior, we should caution that this could be an anomaly of the force field. Since the Mie 16-6 potential is known to be overly repulsive at short distances [15, 5], it is possible that this causes the rapid increase in η at high pressures. Furthermore, as observed in our previous study [15], the densities reported in Table 7 are expected to deviate strongly from the experimental values.

Other studies [4] correct for systematic errors in viscosity by normalizing η with respect to an experimental viscosity value at low pressure. This approach would be possible for the challenge compound since a single experimental data point is available at saturation pressure. Although this may provide a more accurate prediction, we prefer not to use an empirical correction, especially from a single data point. Our goal, rather, is to truly test the force field's predictive capabilities.

The slow system dynamics (i.e. long rotational relaxation times) at high pressures require extremely long simulations and a large number of replicates. An attractive alternative is the so-called time-temperature superposition method, where simulations are performed at higher temperatures (to enhance the configurational sampling) and the viscosity at 293 K is obtained through extrapolation. Despite some obvious benefits, we are wary of the inordinately large uncertainties that this method can produce (see Figure 11 of Ref. 4). Determining the existence of super-Arrhenius behavior necessitates manageable uncertainties at high pressures. For this reason, we choose the more arduous brute-force approach.

5. Conclusions

Previous work demonstrated that the Potoff force field provides reliable viscosities (typically within 10 %) for well-studied *n*-alkane and branched alkanes both at saturation and elevated pressures. For this reason, the Potoff force field was chosen to predict the viscosity-pressure relationship of 2,2,4-trimethylhexane as part of the 10^{th} Industrial Fluid Properties Simulation Challenge. In addition, we investigate the parameter uncertainty in the simulation results with Bayesian inference. Specifically, the non-bonded and torsional potentials are varied from run to run according to a Markov Chain in force field parameter space. Surprisingly, the non-bonded and torsional parameter uncertainties are typically negligible compared to the numerical fluctuations in simulation output. Furthermore, we use cross-validation model selection to verify the existence of so-called super-Arrhenius behavior at high pressures.

Supporting Information

Section SI.I provides GROMACS input files. Section SI.II describes how the CH and C non-bonded MCMC parameter sets are obtained. Section SI.III validates the MCMC-nb uncertainty quantification approach. Section SI.IV develops the A_s skewed distribution used for the MCMC-nb-tors torsional potentials. Section SI.V presents the average Green-Kubo integrals for each state point.

Acknowledgments

We would like to acknowledge Jeffrey J. Potoff and Mohammad Soroush Barhaghi for their invaluable contributions to this study. We are also grateful for the internal review provided by Alta Y. Fang and BLANK of the National Institute of Standards and Technology (NIST).

This research was performed while Richard A. Messerly held a National Research Council (NRC) Postdoctoral Research Associateship at NIST and while Michelle C. Anderson held a Summer Undergraduate Research Fellowship (SURF) position at NIST.

References

References

[1] Scott Bair, Laetitia Martinie, and Philippe Vergne. Classical EHL versus quantitative EHL: A perspective part II-super-Arrhenius piezoviscosity, an essential component of elastohydrodynamic friction missing from classical EHL. *Tribology Letters*, 63(3):37, 2016.

[2] Christopher J. Mundy, Michael L. Klein, and J. Ilja Siepmann. Determination of the pressure-viscosity coefficient of decane by molecular simulation. *The Journal of Physical Chemistry*, 100(42):16779–16781, 1996.

[3] Clare McCabe, Shengting Cui, Peter T. Cummings, Peter A. Gordon, and Roland B. Saeger. Examining the rheology of 9-octylheptadecane to giga-pascal pressures. *The Journal of Chemical Physics*, 114(4):1887–1891, 2001.

[4] Pinzhi Liu, Hualong Yu, Ning Ren, Frances E. Lockwood, and Q. Jane Wang. Pressure-viscosity coefficient of hydrocarbon base oil through molecular dynamics simulations. *Tribology Letters*, 60(3):34, 2015.

[5] Richard A. Messerly, Michelle C. Anderson, S. Mostafa Razavi, and J. Richard Elliott. Improvements and limitations of Mie λ-6 potential for prediction of saturated and compressed liquid viscosity. *Fluid Phase Equilibria*, Pending publication, 2018.

[6] M. G. Martin and J. I. Siepmann. Transferable potentials for phase equilibria. 1. United-atom description of *n*-alkanes. *The Journal of Physical Chemistry B*, 102(14):2569–2577, 1998.

[7] Marcus G. Martin and J. Ilja Siepmann. Novel configurational-bias monte carlo method for branched molecules. Transferable Potentials for Phase Equilibria. 2. United-Atom Description of Branched Alkanes. *The Journal of Physical Chemistry B*, 103(21):4508–4517, 1999.

[8] Mansi S. Shah, J. Ilja Siepmann, and Michael Tsapatsis. Transferable potentials for phase equilibria. Improved united-atom description of ethane and ethylene. *AIChE Journal*, 63(11):5098–5110, 2017.

[9] Andrea Hemmen and Joachim Gross. Transferable anisotropic united-atom force field based on the Mie potential for phase equilibrium calculations: *n*-alkanes and *n*-olefins. *The Journal of Physical Chemistry B*, 119(35):11695–11707, 2015.

[10] Dominik Weidler and Joachim Gross. Transferable anisotropic united-atom force field based on the Mie potential for phase equilibria: Aldehydes, ketones, and small cyclic alkanes. *Industrial & Engineering Chemistry Research*, 55(46):12123–12132, 2016.

[11] J. J. Potoff and D. A. Bernard-Brunel. Mie potentials for phase equilibria calculations: Applications to alkanes and perfluoroalkanes. *The Journal of Physical Chemistry B*, 113(44):14725–14731, 2009.

[12] Jason R. Mick, Mohammad Soroush Barhaghi, Brock Jackman, Loren Schwiebert, and Jeffrey J. Potoff. Optimized Mie potentials for phase equilibria: Application to branched alkanes. *Journal of Chemical & Engineering Data*, 62(6):1806–1818, 2017.

[13] Philippe Ungerer, Christele Beauvais, Jerome Delhommelle, Anne Boutin, Bernard Rousseau, and Alain H. Fuchs. Optimization of the anisotropic united atoms intermolecular potential for *n*-alkanes. *The Journal of Chemical Physics*, 112(12):5499–5510, 2000.

[14] Carlos Nieto-Draghi, Anthony Bocahut, Benoît Creton, Pascal Have, Aziz Ghoufi, Aurélie Wender, , Anne Boutin, Bernard Rousseau, and Laurent Normand. Optimisation of the dynamical behaviour of the anisotropic united atom model of branched alkanes: application to the molecular simulation of fuel gasoline. *Molecular Simulation*, 34(2):211–230, 2008.

[15] Richard A. Messerly, Michael R. Shirts, and Andrei F. Kazakov. Uncertainty quantifi-

cation confirms unreliable extrapolation toward high pressures for united-atom Mie λ-6 force field. *Journal of Chemical Physics*, Pending publication, 2018.

[16] F. Rizzi, H. N. Najm, B. J. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson, and O. M. Knio. Uncertainty quantification in MD simulations. Part II: Bayesian inference of force-field parameters. *Multiscale Modeling & Simulation*, 10(4):1460–1492, 2012.

[17] Panagiotis Angelikopoulos, Costas Papadimitriou, and Petros Koumoutsakos. Bayesian uncertainty quantification and propagation in molecular dynamics simulations: A high performance computing framework. *The Journal of Chemical Physics*, 137(14):144103, 2012.

[18] Fabien Cailliez and Pascal Pernot. Statistical approaches to forcefield calibration and prediction uncertainty in molecular simulation. *The Journal of Chemical Physics*, 134(5):054124, 2011.

[19] Richard A. Messerly, Thomas A. Knotts IV, and W. Vincent Wilding. Uncertainty quantification and propagation of errors of the Lennard-Jones 12-6 parameters for *n*-alkanes. *The Journal of Chemical Physics*, 146(19):194110, 2017.

[20] B. Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

[21] Peter A. Gordon. Development of intermolecular potentials for predicting transport properties of hydrocarbons. *The Journal of Chemical Physics*, 125(1):014504, 2006.

[22] Carlos Nieto-Draghi, Philippe Ungerer, and Bernard Rousseau. Optimization of the anisotropic united atoms intermolecular potential for *n*-alkanes: Improvement of transport properties. *The Journal of Chemical Physics*, 125(4):044517, 2006.

[23] Carlos Braga and Karl P. Travis. Computer simulation of the role of torsional flexibility on mass and momentum transport for a series of linear alkanes. *The Journal of Chemical Physics*, 137(6):064116, 2012.

[24] Carmelo Herdes, Tim S. Totton, and Erich A. Müller. Coarse grained force field for the molecular simulation of natural gases and condensates. *Fluid Phase Equilibria*, 406:91 – 100, 2015.

[25] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids*. Oxford University Press, Oxford England, second edition, 2017.

[26] M.J. Abraham, D. van der Spoel, E. Lindahl, B.Hess, and the GROMACS development team. *GROMACS User Manual version 2018, www.gromacs.org (2018)*.

[27] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, 1998.

[28] Berk Hess. P-LINCS: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, 2008.

[29] William C. Swope, Hans C. Andersen, Peter H. Berens, and Kent R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649, 1982.

[30] R.W Hockney, S.P Goel, and J.W Eastwood. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*, 14(2):148 – 158, 1974.

[31] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101, 2007.

[32] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31:1695–1697, 1985.

[33] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519, 1984.

[34] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690, 1984.

[35] Shuichi Nosé and M.L. Klein. Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5):1055–1076, 1983.

[36] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, 1981.

[37] Edward J. Maginn, Richard A. Messerly, Daniel J. Carlson, Daniel R. Roe, and J. Richard Elliott. Best practices for computing transport properties 1. Self-diffusivity and viscosity from equilibrium molecular dynamics v1. *Living Journal of Computational Molecular Science*, Pending publication, 2018.

[38] Yong Zhang, Akihito Otani, and Edward J. Maginn. Reliable viscosity calculation from equilibrium molecular dynamics simulations: A time decomposition method. *Journal of Chemical Theory and Computation*, 11(8):3537–3546, 2015.

[39] C. Barus. Isothermals, isopiestics and isometrics relative to viscosity. *American Journal of Science*, 45:87–96, 1893.

[40] C.J.A. Roelands. *Correlational Aspects of the Viscosity-Temperature-Pressure Relationships of Lubricating Oils*. PhD thesis, TU Delft, Delft University of Technology, 1966.

[41] Michael Frenkel, Robert D. Chirico, Vladimir Diky, Xinjian Yan, Qian Dong, and Chris Muzny. Thermodata engine (TDE): software implementation of the dynamic data evaluation concept. *Journal of Chemical Information and Modeling*, 45(4):816–838, 2005.

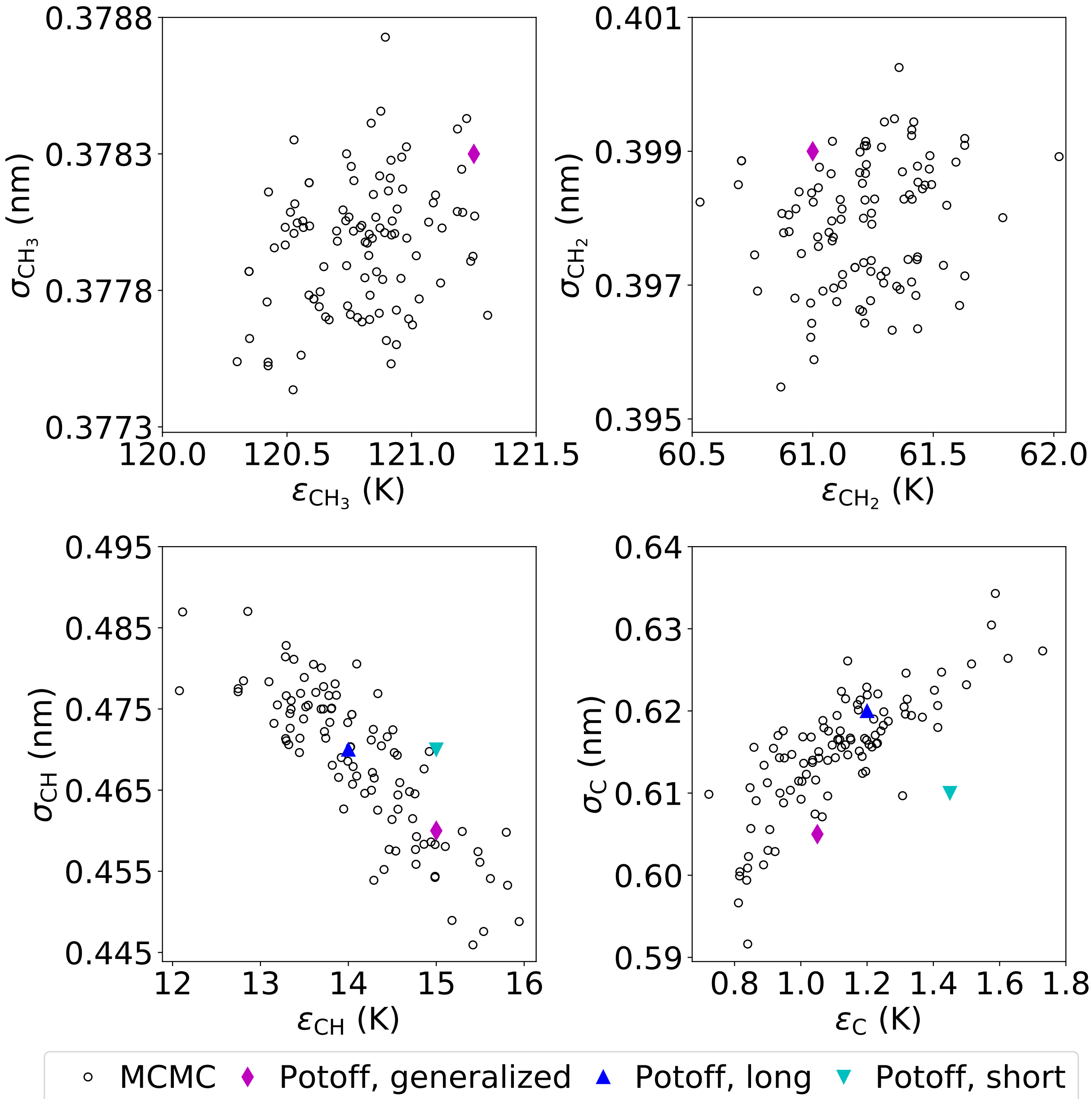

Figure 2: Uncertainty in non-bonded parameters determined with Markov Chain Monte Carlo (MCMC). The Potoff generalized and S/L parameters are also included as a reference [11, 12]. Top left, top right, bottom left, and bottom right panels correspond to CH_3, CH_2, CH, and C parameters, respectively. Note in Reference 12 that a small discrepancy exists between the generalized C parameters in Table 1 and the optimal region depicted in Figure 1. The generalized C parameters shown here were obtained from private communication of tabulated scoring function values and are, thus, slightly different than those reported in Reference 12.

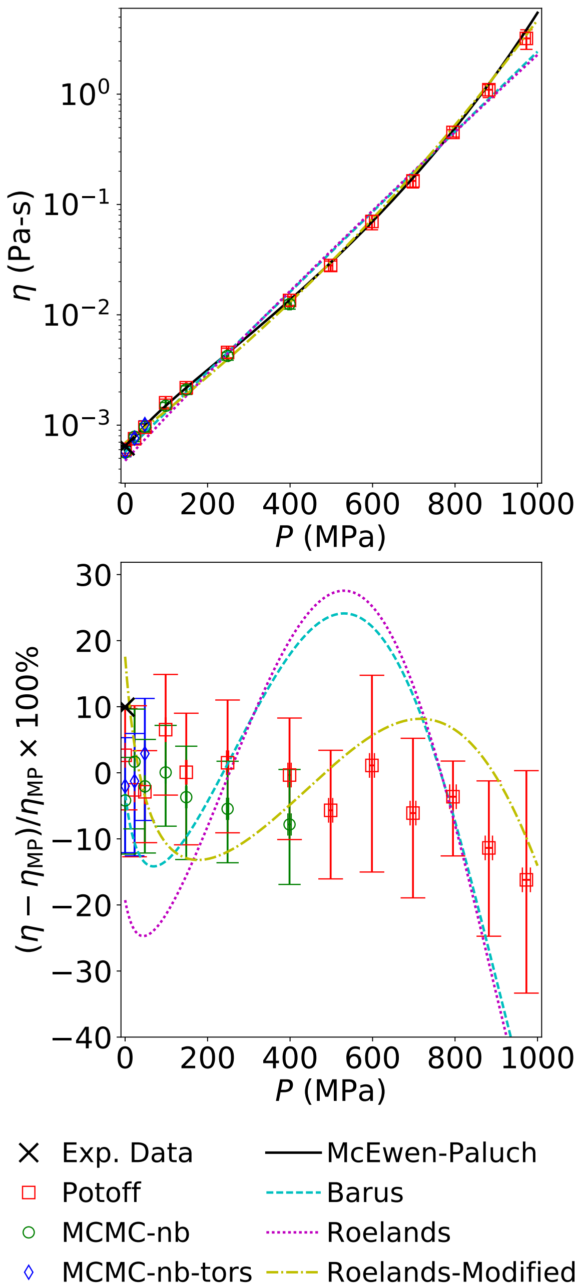

Figure 3: Viscosity-pressure results for Potoff (red squares), MCMC-nb (green circles), and MCMC-nb-tors (blue diamonds). Top panel plots $\log_{10}(\eta)$-P where different line colors and styles represent different empirical model fits (Equations 10, 11, and 12) to MCMC-nb-tors values. Bottom panel is a percent deviation plot relative to the McEwen-Paluch fit. Experimental viscosity point at atmospheric pressure is included as a reference [41].

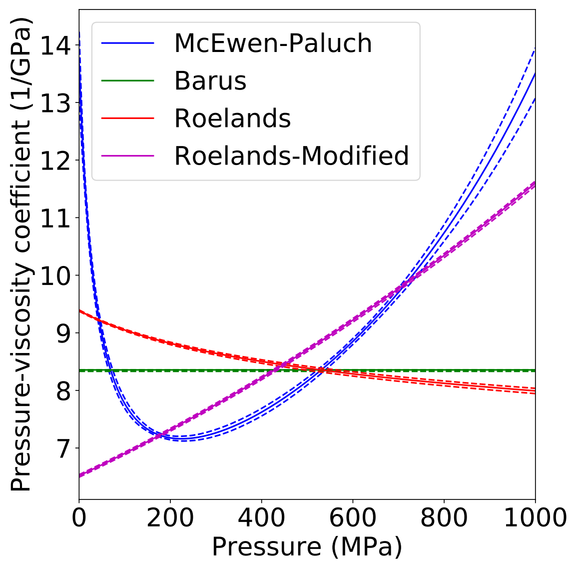

Figure 4: Pressure-viscosity coefficient predicted with empirical model fits (Equations 10, 11, and 12). Dashed lines represent 95 % confidence intervals obtained from bootstrap re-sampling.

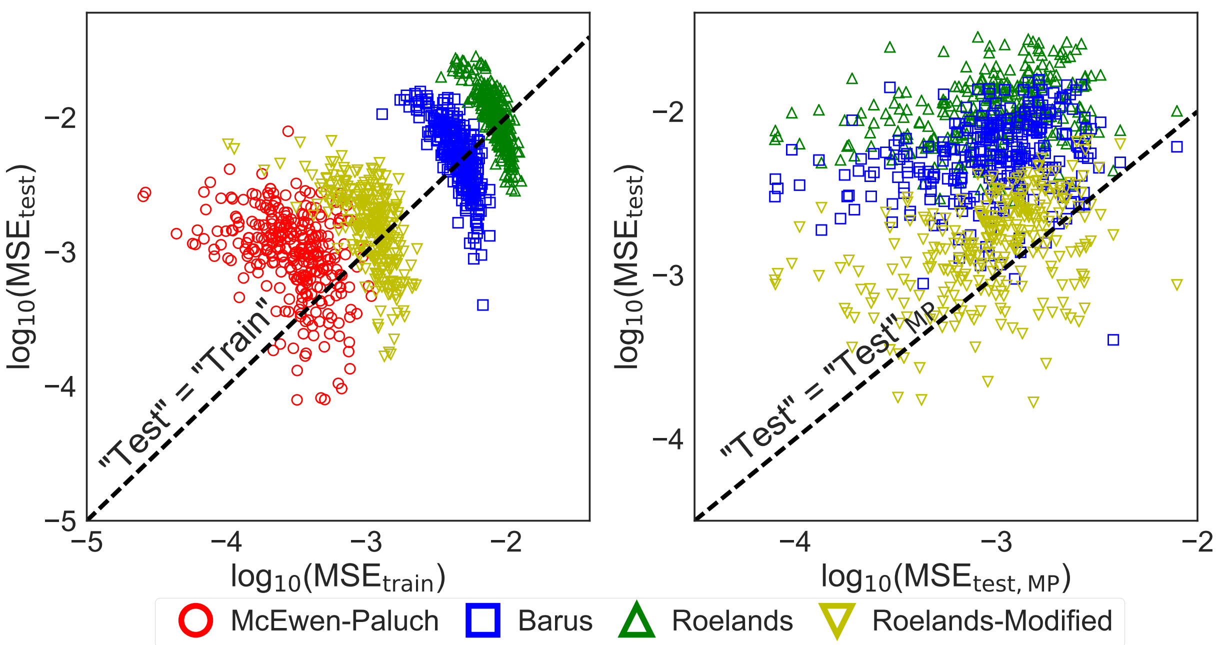

Figure 5: Monte Carlo cross-validation for empirical model fits (Equations 10, 11, and 12). $\mathrm{MSE}_{\mathrm{train}}$ and $\mathrm{MSE}_{\mathrm{test}}$ are the mean-squared-error for the "training" and "testing" sets, respectively. Left panel compares $\mathrm{MSE}_{\mathrm{train}}$ and $\mathrm{MSE}_{\mathrm{test}}$ for each model, while the right panel compares the McEwen-Paluch $\mathrm{MSE}_{\mathrm{test}}$ with the $\mathrm{MSE}_{\mathrm{test}}$ for the other three models.