

# The role of force field parameter uncertainty in the prediction of high pressure viscosities.

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

Uncertainty estimates increase substantially for high pressures at low reduced temperatures. Nevertheless, a prediction is made for the viscosity of 2,2,4, trimethylhexane at 293K and 1000 MPa, in compliance with the guidelines of the 10th IFPSC.

*Keywords:*

Uncertainty Quantification, Bayesian Inference, Thermophysical Properties, Shear Viscosity, Molecular Dynamics Simulation, Force Fields, Green-Kubo

---

## 1. Introduction

The Industrial Fluid Properties Simulation Challenge is an open competition organized by the Computational Molecular Science and Engineering Forum (CoMSEF) of the American Institute of Chemical Engineers (AIChE), the American Chemical Society (ACS), Army Research Lab, National Institute of Standards and Technology, The Dow

---

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

Chemical Company, 3M, and United Technologies Research Center. The goals of the competition are to drive improvements in the practice of molecular modeling, formalize methods for the evaluation and validation of simulation results with experimental data, and ensure relevance of simulation activities to industrial requirements. The Simulation Challenge was initiated by the workshop on "Predicting the Thermophysical Properties of Fluids by Molecular Simulation" ([link](#)) and is part of the overall vision of the Industrial Fluid Properties Simulation Collective.

The 10th Industrial Fluid Properties Simulation Challenge will test the capability of molecular dynamics simulation to provide the property of liquids most important to elastohydrodynamic lubrication (EHL), the pressure-viscosity relation. The temperature dependence at elevated pressure could be the subject of a future challenge.

A fundamental requirement of elastohydrodynamic lubrication (EHL) is a description of the viscosity of the liquid as a function of pressure [1]. The classical film thickness formulas all require a value for a property known as a pressure-viscosity coefficient; although the definition of this property is not always clear [2]. The shape of a traction (friction) curve has been the subject of much speculation for at least forty years [3]. The shape of a traction curve when plotted as friction coefficient or average shear stress versus the logarithm of sliding speed or slide-to-roll ratio depends strongly on the pressure dependence of viscosity at the Hertz pressure [4], more fragile liquids having a less steep logarithmic portion. Although the pressure dependence of viscosity is clearly essential to the field, until about ten years ago, experimentally measured values of this property were not a typical part of EHL analysis. The reasons for the previous neglect may be debated, but the demand for this information is now growing.

Molecular dynamics simulations have the promise of generating pressure-viscosity data for liquids which have not yet been synthesized but only if the accuracy of the method can be validated. There has been a claim of success in predicting the pressure dependence of viscosity for squalane [5], although the temperature dependence is not accurately recovered in this example [6]. There has been extensive experimental work

on squalane viscosity at elevated pressure [7,8] so that simulations have a known “target” value of viscosity. As of this time, there has been no success in recovering the super-Arrhenius pressure dependence that is important to friction [9].

A Lubricant Viscosity Simulation Challenge is now proposed to assess the possibility of employing molecular dynamics simulations to predict the pressure dependence of viscosity in a simple hydrocarbon molecule. This should be a material for which there is no presently published viscosity data, with the exception perhaps of viscosity at ambient pressure. It should, for simulation convenience, be composed of a minimum number of carbons. Linear alkanes are excluded because they are not glass-formers and would be crystallized at EHL pressures. The material should possess all of the pressure-viscosity trends of lubricating oil. A candidate fulfilling these requirements is 2,2,4 Trimethylhexane.

Prof. Scott Bair (Georgia Tech) has characterized the viscosity of 2,2,4 Trimethylhexane (>98%), Aldrich product number 92470, lot BCBR3588V. The viscometers were calibrated with di (2ethylhexyl) sebacate based on the correlation of Paredes et al. [10]. Estimated uncertainties are 3% for viscosity, 0.3°C for temperature and the greater of 1MPa and 0.4

Entrants are challenged to predict the viscosity at pressures of 0.1, 25, 50, 100, 150, 250, 400, 500, 600, 700, 800, 900, and 1000 MPa, all at temperature of 20°C (293K). Entries will be judged based on comparison to the benchmark data at those state conditions and to the pressure viscosity coefficient.

1. Introduce the industrial fluid properties simulation challenge
2. Discuss the details of the 10th challenge
3. Explain why this challenge is important/interesting:
  - (a) Viscosity is an important property for designing chemical systems
  - (b) Viscosity data typically do not cover the entire range of  $P\rho T$  of interest
  - (c) Prediction methods are typically quite poor for viscosity
  - (d) Molecular simulation is an attractive alternative, but two main challenges

- i. Difficulty of obtaining reproducible results from simulation
  - ii. Unreliable force fields
4. We performed a systematic investigation of several united-atom force fields and determined Potoff to be the most reliable
  5. Although Potoff over predicts viscosity and pressure with respect to density, it is quite reliable at predicting viscosity with respect to pressure
  6. The uncertainty in force field parameters is key for rigorously quantifying the uncertainty

One of the entry guidelines for the challenge is “an analysis of the uncertainty in the calculated results.” Traditional calculation uncertainties are limited to the random fluctuations of simulation output and/or the uncertainty related to the data post-processing. This class is referred to as “numerical uncertainty.” Two other classes of uncertainty exist which are often more important, namely, “parameter uncertainty” and “functional form uncertainty” (also referred to as “model uncertainty”). The latter refers to the uncertainty associated with the choice of force field functional form. Parameter uncertainty refers to the uncertainty in the force field parameters, for a given force field functional form.

Quantifying the force field functional form uncertainty is a difficult task, as it often requires considering and performing simulations with numerous functional forms. In this study, we investigate both numerical and parameter uncertainties, where the chosen functional form is the same as the Potoff force field, namely, a united-atom, fixed bond length, harmonic angular potential, cosine series torsional potential, and a Mie 16-6 non-bonded potential (see Section 2.3 for details). As previous studies demonstrate the high sensitivity of viscosity on the non-bonded and torsional potentials, we limit our parameter uncertainty investigation to the non-bonded and torsional parameters.

The design of efficient and reliable technical processes requires accurate estimates of thermophysical properties. Shear viscosity ( $\eta$ ) is an important property for characterizing flow, e.g., sizing pumps, assessing flow assurance in fossil fuel recovery, and lubricating bearings in tribological applications. There are primarily three different means by which

shear viscosity values are obtained: experimental measurement, semi-empirical prediction models, and molecular simulation (molecular dynamics, MD). Significant limitations exist for each of these methods.

For example, experimental measurements can be expensive, time-consuming, and challenging at extreme temperatures ( $T$ ) and pressures ( $P$ ). Experimental data tend to be distributed among several prototypes of linear, branched, ring, and polar molecules, with many gaps among a homologous series. Most experimental data are available below 200 MPa, while tribological applications may require estimates at pressures as high as 1000 MPa. Flow assurance applications are generally at pressures below 200 MPa, but at temperatures of 423 to 523 K. The ever expanding conditions of interest and economic constraints on new measurements foster increased research in predictive methods.

The National Institute of Standards and Technology (NIST) Reference Fluid Properties (REFPROP) database software provides “reference quality” viscosity correlations for experimentally well-studied compounds (around 100 species) [? ]. Most compounds, however, do not have sufficient *reliable* experimental data covering a wide range of temperatures, pressures, and densities ( $\rho$ ) for developing “reference quality” correlations. These less-studied compounds require predictive methods that pool together data from several related molecular species.

Semi-empirical prediction models are typically not reliable over the industrially relevant ranges of  $P\rho T$  [? ]. For example, corresponding states methods are recommended for vapors, dense fluids, and high temperature liquids. These methods rely on the similarity of trends in the properties relative to reference compounds, e.g., methane and *n*-octane. Corresponding states methods are less reliable for more complex molecular structures, e.g., branched compounds. Typical compilations indicate that deviations from experiment may vary by 5 to 50 %, with little guidance about when to expect lower or higher accuracy.

For low temperature liquids, group contribution schemes are favored, but these tend to extrapolate poorly when applied to compounds or conditions outside the training set.

More recent advances such as machine learning [? ? ] and entropy scaling [? ] have shown great promise in prediction of historically challenging properties, such as viscosity, thermal conductivity, and surface tension. However, machine learning relies on large amounts of experimental data and often suffers from dubious extrapolation. While entropy scaling has a stronger theoretical basis, it requires a reliable reference viscosity and an adequate equation-of-state, which may not be readily available for the compound of interest.

As an alternative to experiment and semi-empirical prediction models, molecular simulation is an attractive means for estimating viscosity. However, there are two fundamental challenges impeding the use of molecular simulation as a mainstream chemical engineering tool for viscosity prediction. The first challenge is that obtaining reproducible results is more difficult for transport properties, such as viscosity, than for thermodynamic properties. Recently, a “Best Practices Guide” was developed to address this challenge, namely, to improve reproducibility of viscosity estimates [? ]. We apply these “Best Practices” and address some outstanding issues mentioned therein.

The second challenging aspect of obtaining accurate simulation estimates is that viscosity is extremely sensitive to the force field. In addition to the strong dependence on the non-bonded interactions, the bonded potential plays a much greater role for viscosity than for thermodynamic properties. For example, varying the torsional potential has a significant impact on viscosity [? ], while vapor-liquid coexistence is relatively unaffected [? ]. Therefore, the ability to predict viscosities with molecular simulation requires both robust methods and adequate force fields.

We investigate the accuracy of four force fields, namely, Transferable Potentials for Phase Equilibria (TraPPE-UA, also referred to simply as TraPPE [? ? ? ]), Transferable Anisotropic Mie (TAMie) [? ? ], Potoff [? ? ], and fourth generation anisotropic-united-atom (AUA4) [? ? ]. Each force field is a variation of the united-atom (UA) Mie  $\lambda$ -6 (generalized Lennard-Jones, LJ) model, a popular class designed for the engineering purpose of predicting thermophysical properties. However, the suitability of these force

fields for quantitative viscosity prediction, especially at high pressures, has been widely debated in the literature.

For example, depending on the compound structure and state conditions, some studies suggest that UA LJ 12-6 models (e.g., TraPPE) are inadequate for estimating viscosities and recommend the use of anisotropic-united-atom (AUA) or all-atom (AA) models for this purpose [? ? ? ? ]. Considering the significant increase in computational cost of AA simulations, two promising alternatives have been investigated, namely, the Mie  $\lambda$ -6 potentials and/or modified torsional models. For example, Nieto-Draghi et al. demonstrate significant improvement in viscosity prediction for the AUA4 model by increasing the torsional barriers [? ]. In addition, the UA Mie  $\lambda$ -6 model has been shown to accurately predict saturated liquid viscosity ( $\eta_{\text{liq}}^{\text{sat}}$ ) without significant degradation of other vapor-liquid saturation properties [? ], i.e., saturated liquid density ( $\rho_{\text{liq}}^{\text{sat}}$ ), saturated vapor density ( $\rho_{\text{vap}}^{\text{sat}}$ ), and saturated vapor pressure ( $P_{\text{vap}}^{\text{sat}}$ ).

Hoang et al. demonstrate that including viscosity data in the force field development can improve the identification of a unique set of transferable Mie  $\lambda$ -6 parameters, while simultaneously improving viscosity predictions [? ]. By contrast, the force fields compared in the present work were optimized solely with vapor-liquid coexistence data, i.e., dynamic properties, such as viscosity, were not included in their parameterization. Notwithstanding the potential benefits of including viscosity as a property of interest during force field development, we assess the accuracy of TraPPE-UA, TAMie, Potoff, and AUA4 for estimating viscosity as they currently stand, including their torsional potential models. While the Potoff and TAMie force fields have shown considerable promise in predicting static properties (in particular,  $P_{\text{vap}}^{\text{sat}}$ ), their ability to predict dynamic properties has not been investigated previously.

The outline for the present work is the following. Section 2 explains the force fields, simulation methodology, and data analysis. Section 3 presents the simulation results for each force field, compound, and state point studied. Section ?? discusses some important observations and limitations. Section 4 recaps the primary conclusions from this work.

## 2. Methods

### 2.1. Simulation set-up

Viscosity estimates can be obtained from both equilibrium molecular dynamics (EMD) and non-equilibrium molecular dynamics (NEMD) simulations. The “Best Practices Guide” is currently limited to EMD methods and purports that NEMD might be necessary for high viscosities (greater than 0.02 Pa-s). One purpose of the present work is to demonstrate that, by applying these guidelines, EMD can also provide meaningful estimates for highly viscous systems.

Equilibrium molecular dynamics simulations are performed using GROMACS version 2018 with “mixed” (single and double) precision [? ]. GROMACS is compiled using GNU 7.3.0, OpenMPI enabled, and GPU support disabled. Approximately three-fourths of the simulations are run using Linux 4.4.0-112-generic x86\_64 on an Intel(R) Xeon(R) CPU E5-2699 v4 @ 2.20GHz machine while the remaining one-fourth are run using Linux 4.15.0-22-generic x86\_64 on an Intel(R) Xeon(R) CPU E5-2450 0 @ 2.10GHz machine. 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.

Simulation specifications are provided in Table 1. Note that each force field utilizes a 1.4 nm non-bonded cut-off distance, with the exception of Potoff which employs a 1.0 nm cut-off (as recommended in Reference ? ) except for certain compounds (see Section ??). Analytical tail corrections are applied in all cases [? ]. For most systems, 1 ns is a sufficient production time, while longer simulations are required for the more viscous systems, e.g., 16 ns for 2,2,4-trimethylpentane at 1000 MPa. Sections ?? and ?? provide a detailed analysis for the number of molecules and the cut-off lengths, respectively. Sections ?? and ?? of Supporting Information investigate, respectively, the sensitivity of our results to the production time and constrained bonds. To validate our methods, a comparison with other reference simulation values [? ? ? ] is provided in Section ?? of Supporting Information.



Table 1: General simulation specifications.

Time-step (fs)	2
Equilibration time (ns)	1
Production time (ns)	1, 2, 4, 8, or 16
Cut-off length (nm)	1.4 (1.0 for Potoff)
Tail-corrections	$U$ and $P$
Constrained bonds	LINCS [? ? ]
LINCS-order	8
Number of molecules	400

When  $\eta$  is desired at a prescribed  $T$  and  $P$ , six simulation stages are required: energy minimization,  $NPT$  equilibration,  $NPT$  production, energy minimization,  $NVT$  equilibration, and  $NVT$  production. When  $\eta$  is desired at a prescribed  $T$  and  $\rho$ , the  $NPT$  stages are unnecessary and only three simulation stages are required: energy minimization,  $NVT$  equilibration, and  $NVT$  production. Note that the final production stage simulations are always performed using the  $NVT$  ensemble [? ]. Details regarding the thermostats and barostats employed are provided in Section ?? of Supporting Information.

We utilize 30 to 60 independent replicates to improve the precision and to provide more rigorous estimates of uncertainty [? ? ]. To ensure independence between replicates, the entire series of simulation stages are repeated for each replicate. Each energy minimization stage starts with a different pseudo-random configuration while the initial velocities are also randomized for each equilibration stage.

We investigate two different classes of viscosity, namely, saturated liquid viscosity ( $\eta_{\text{liq}}^{\text{sat}}$ ) and compressed liquid viscosities at a temperature of 293 K ( $\eta_{\text{liq}}^{\text{comp}}$ ). Saturated liquid viscosities are estimated by performing  $NVT$  ensemble simulations at the saturation temperature ( $T^{\text{sat}}$ ) and saturated liquid density ( $\rho_{\text{liq}}^{\text{sat}}$ ). The simulation densities correspond to the REFPROP  $\rho_{\text{liq}}^{\text{sat}}$ , which is admittedly not necessarily the same as the force

field  $\rho_{\text{liq}}^{\text{sat}}$ . This point is discussed in greater detail in Section ??.

Two different simulation protocols are implemented for estimating compressed liquid viscosities ( $\eta_{\text{liq}}^{\text{comp}}$ ). Specifically, we perform simulations with each force field either at the same  $\rho$  or the same  $P$ . For the purpose of comparing the accuracy of force fields, these two methods are essentially equivalent. From a practical standpoint, estimating  $\eta$  at a given  $P$  requires performing preliminary  $NPT$  ensemble simulations to determine the corresponding box size.

## 2.2. Data analysis

The analysis for the Potoff Mie  $\lambda$ -6 force field simulation results is identical to that prescribed in our previous study BLANK. In brief, we implement the Green-Kubo “time-decomposition” analysis [? ? ]

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

where  $t$  is time,  $V$  is volume,  $N_{\text{reps}}$  is the number of independent replicate simulations,  $\alpha$  and  $\beta$  are  $x, y$ , or  $z$  Cartesian coordinates,  $\tau_{\alpha\beta,n}$  is the  $\alpha$ - $\beta$  off-diagonal stress tensor element for the  $n^{\text{th}}$  replicate, and  $\langle \cdots \rangle_{t_0, \alpha\beta}$  denotes an average over time origins ( $t_0$ ) and  $\tau_{\alpha\beta}$ .

$\tau_{\alpha\beta,n}$  is recorded every 6 fs (3 time-steps) for accurate integration of Equation 1. To improve precision, Equation 1 averages several (between 30 and 60) independent replicate simulations, twelve different time-origins, and all three unique off-diagonal stress tensor components.

The “true” viscosity, i.e., the infinite-time-limit viscosity, is obtained by evaluating Equation 1 as  $t \rightarrow \infty$ . However, the long-time tail of the Green-Kubo integral is often quite noisy and does not converge smoothly. For this purpose, we fit the “running integral” to a double-exponential function

$$\eta(t) = A\alpha\tau_1 (1 - \exp(-t/\tau_1)) + A(1 - \alpha)\tau_2 (1 - \exp(-t/\tau_2)) \quad (2)$$

where  $A, \alpha, \tau_1$ , and  $\tau_2$  are fitting parameters and  $\eta^\infty = A\alpha\tau_1 + A(1 - \alpha)\tau_2$  is the infinite-time-limit viscosity.

Since the Green-Kubo “running integral” suffers from extreme fluctuations at long times, Equation 2 is fit by minimizing a weighted sum-squared error objective function. Weights are equal to the inverse of the standard deviation ( $\sigma_\eta$ ) of the replicate simulations. The time dependence of  $\sigma_\eta$  is modeled with  $At^b$ , where  $A$  and  $b$  are fitting parameters.

Despite weighting the data by  $\sigma_\eta$ , it can be challenging to obtain good fits of Equation 2 when all the data are included. For this reason, data are excluded where  $\sigma_\eta > 0.4 \times \eta^\infty$  [? ?]. Occasionally, this heuristic results in a cut-off that is too short, in particular for systems with slow dynamics, which also leads to very poor fits. In such cases, it is necessary to modify the heuristic, e.g., exclude data where  $\sigma_\eta > 0.8 \times \eta^\infty$ . As erroneously large fluctuations also exist at very short times, only data for  $t > 3$  ps are included in the fitting of Equation 2 [? ?].

The uncertainty in  $\eta^\infty$  is obtained by bootstrap re-sampling. Specifically, the fitting process described previously is repeated hundreds of times using randomly selected subsets of replicate simulations. Furthermore, each bootstrap repetition uses a randomly selected long-time cut-off between  $0.35 \times \eta^\infty$  and  $0.45 \times \eta^\infty$  to account for uncertainty in the  $0.4 \times \eta^\infty$  heuristic. A 95 % confidence interval is obtained from the distribution of bootstrap estimates for  $\eta^\infty$ . An example of this process is provided in Section ?? of Supporting Information.

### 2.3. Force fields

As we demonstrated in our previous study BLANK, the Potoff Mie  $\lambda$ -6 force field provides reliable estimates of the  $\eta$ - $P$  dependence for normal and branched alkanes. In particular, this force field predicts the viscosity within 10 % for 2,2,4-trimethylpentane up to 200 MPa and for propane up to 1000 MPa. By contrast, the TraPPE and TAMie models are c appears to ex being the to

For these reasons, we utilize the Potoff Mie  $\lambda$ -6 force field as the basis for our predictions of  $\eta$ . In addition, we quantify the uncertainties in the non-bonded and torsional parameters. We subsequently propagate these force field parameter uncertainties when predicting  $\eta$ .

The Potoff Mie  $\lambda$ -6 force field utilizes united-atom (UA) sites, where 2,2,4-trimethylhexane is represented with CH<sub>3</sub>, CH<sub>2</sub>, 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 the use of flexible bonds can impact  $\eta$  by several percent.

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

$$u^{\text{bend}} = \frac{k_\theta}{2} (\theta - \theta_0)^2 \quad (3)$$

where  $u^{\text{bend}}$  is the bending energy,  $\theta$  is the instantaneous bond angle,  $\theta_0$  is the equilibrium bond angle (see Table 2), and  $k_\theta$  is the harmonic force constant with  $k_\theta/k_B = 62500 \text{ K/rad}^2$  for all bonding angles, where  $k_B$  is the Boltzmann constant.

Table 2: Equilibrium bond angles ( $\theta_0$ ) [°]. CH<sub>*i*</sub> and CH<sub>*j*</sub> represent CH<sub>3</sub>, CH<sub>2</sub>, CH, or C sites.

Bending sites	$\theta_0$ (degrees)
CH <sub><i>i</i></sub> -CH <sub>2</sub> -CH <sub><i>j</i></sub>	114.0
CH <sub><i>i</i></sub> -CH-CH <sub><i>j</i></sub>	112.0
CH <sub><i>i</i></sub> -C-CH <sub><i>j</i></sub>	109.5

Dihedral torsional interactions are determined using a modified cosine series:

$$\begin{aligned} u^{\text{tors}} &= c_0 + c_1[1 + \cos \phi] + c_2[1 - \cos 2\phi] + c_3[1 + \cos 3\phi] + A_s \sin^2 \left( \frac{3\phi}{2} + 180^\circ \right) \\ &= (c_0 - A_s) + c_1[1 + \cos \phi] + c_2[1 - \cos 2\phi] + \left( c_3 + \frac{A_s}{2} \right) [1 + \cos 3\phi] \quad (4) \end{aligned}$$

where  $u^{\text{tors}}$  is the torsional energy,  $\phi$  is the dihedral angle,  $c_n$  are the Fourier constants used in the Potoff force field and listed in Table 3, and  $A_s \sin^2 \left( \frac{3\phi}{2} + 180^\circ \right)$  is an additional term proposed by ? to shift the torsional barrier heights. Note that  $\phi$  is defined using a convention similar to IUPAC where  $\phi = 180^\circ$  for the *trans* conformation, whereas ? defines the *trans* conformation as  $0^\circ$  or  $360^\circ$  deg, hence the  $+180^\circ$  term. As  $\sin^2 \left( \frac{3\phi}{2} + 180^\circ \right)$  has a maximum value of 1 at  $0^\circ$ ,  $120^\circ$ ,  $240^\circ$ , and  $360^\circ$ ,  $u^{\text{tors}}$  is shifted by  $A_s$  at these dihedral angles. By contrast, this additional term does not shift  $u^{\text{tors}}$  for dihedral angles of  $0^\circ$ ,

180°, and 300°, which correspond to the equilibrium conformations of *gauche*<sup>-</sup>, *trans*, and *gauche*<sup>+</sup>, respectively. Clearly, the non-shifted Potoff torsional potential is obtained only when  $A_s = 0$ .

Table 3: Fourier constants ( $c_n/k_B$ ) in units of K [? ?].  $CH_i$  and  $CH_j$  represent  $CH_3$ ,  $CH_2$ ,  $CH$ , or  $C$  sites.

Torsion sites	$c_0/k_B$	$c_1/k_B$	$c_2/k_B$	$c_3/k_B$
$CH_i-CH_2-CH-CH_j$	-251.06	428.73	-111.85	441.27
$CH_i-CH_2-C-CH_j$	0.0	0.0	0.0	461.29

? set  $A_s$  equal to 40% and 15% of the maximum dihedral barrier for the  $CH_3-CH_2-CH_2-CH_2$  and  $CH_2-CH_2-CH_2-CH_2$  torsional potentials, respectively. This corresponds to  $A_s/k_B \approx 1000$  K and  $\approx 375$  K for the  $CH_3-CH_2-CH_2-CH_2$  and  $CH_2-CH_2-CH_2-CH_2$  torsional potentials, respectively. The primary reason why ? introduced this additional term was to increase the torsional barriers and, thereby, increase the viscosity obtained with the AUA4m force field. This methodology works fairly well for Lennard-Jones 12-6 force fields, which systematically under predict viscosity by greater than 30 %. However, since 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  $A_s/k_B \gg 0$  K.

The reason we include the additional term, however, is to provide a simple method for quantifying the uncertainty in the torsional potential. Specifically, we assume that  $A_s$  follows a normal distribution with a mean value of zero and a standard deviation equal to  $0.075 \times \max(u_{A_s=0}^{\text{tors}})$ . The standard deviation is assigned such that the 95% confidence interval is equal to 15% the maximum barrier height for the non-shifted Potoff torsional potential. We use a normal distribution such that the uncertainty in the dihedral barriers is symmetric, i.e., unlike ? we do not assume that the dihedral barriers must be increased unilaterally.

Non-bonded interactions between sites located in two different molecules or sepa-

Figure 1: Uncertainty in dihedral potentials. Black line is the non-shifted Potoff torsional potential. Red lines are the 200 MCMC sampled parameter sets used in this study. Insets show the distribution for  $A_s/k_B$  in units of K.

rated by more than three bonds within the same molecule are calculated using a Mie  $\lambda$ -6 potential (of which the Lennard-Jones, LJ, 12-6 is a subclass) [? ]:

$$u^{\text{vdw}}(\epsilon, \sigma, \lambda; r) = \left( \frac{\lambda}{\lambda - 6} \right) \left( \frac{\lambda}{6} \right)^{\frac{6}{\lambda - 6}} \epsilon \left[ \left( \frac{\sigma}{r} \right)^\lambda - \left( \frac{\sigma}{r} \right)^6 \right] \quad (5)$$

where  $u^{\text{vdw}}$  is the van der Waals interaction,  $\sigma$  is the distance ( $r$ ) where  $u^{\text{vdw}} = 0$ ,  $-\epsilon$  is the energy of the potential at the minimum (i.e.,  $u^{\text{vdw}} = -\epsilon$  and  $\frac{\partial u^{\text{vdw}}}{\partial r} = 0$  for  $r = r_{\text{min}}$ ), and  $\lambda$  is the repulsive exponent. The non-bonded Potoff Mie  $\lambda$ -6 force field parameters are provided in Table 4.

Table 4: Non-bonded Potoff Mie  $\lambda$ -6 parameters. The “short/long” Potoff CH and C parameters are included in parentheses.

	Potoff (S/L)		
United-atom	$\epsilon/k_B$ (K)	$\sigma$ (nm)	$\lambda$
CH <sub>3</sub>	121.25	0.3783	16
CH <sub>2</sub>	61	0.399	16
CH	15 (15/14)	0.46 (0.47/0.47)	16
C	1.2 (1.45/1.2)	0.61 (0.61/0.62)	16

Potoff reports a “generalized” and “short/long” (S/L) CH and C parameter set. The “short” and “long” parameters are implemented when the number of carbons in the backbone is  $\leq 4$  and  $> 4$ , respectively. Due to their superior accuracy, our reported Potoff results only implement the S/L parameter set.

The uncertainty in the CH<sub>3</sub> and CH<sub>2</sub> non-bonded Mie 16-6 parameters were obtained previously. Reference BLANK assumed that the CH<sub>3</sub> parameters were transferable from ethane for longer  $n$ -alkanes. The CH<sub>2</sub> parameters were obtained from propane,  $n$ -butane,

and *n*-octane. The data included in the analysis were saturated liquid densities and saturated vapor pressures over a reduced temperature range of 0.45 to 0.85, as available in ThermoData Engine (TDE).

The uncertainty in the CH and C parameter sites were obtained from Reference BLANK. Rather than apply

With the assumption of sequential transferability from CH<sub>3</sub>

The uncertainty for the non-bonded Mie 16-6 parameter sets in  $\epsilon_{CH_3}$  non-bonded Mie 16-6 parameters

To simplify the uncertainty analysis, we assume the correlation between Mie parameters of different UA sites is zero. Therefore, we only account for the correlation between  $\epsilon$  and  $\sigma$  of a given UA site type.

The Potoff “generalized” CH and C parameter set is an attempt at a completely transferable set. However, since the “generalized” parameters performed poorly for some compounds, the S/L parameter set was proposed, where the “short” and “long” parameters are implemented when the number of carbons in the backbone is  $\leq 4$  and  $> 4$ , respectively.

Figure 2: Uncertainty in non-bonded potentials. Black line is the Potoff non-bonded potential. Red lines correspond to the 200 MCMC sampled parameter sets used in this study. Insets show the distribution for  $\epsilon$  and  $\sigma$ .

Non-bonded parameters between two different site types (i.e., cross-interactions) are determined using Lorentz-Berthelot combining rules [?] for  $\epsilon$  and  $\sigma$  and an arithmetic mean for the repulsive exponent  $\lambda$  (as recommended in Reference ?):

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \quad (6)$$

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad (7)$$

$$\lambda_{ij} = \frac{\lambda_{ii} + \lambda_{jj}}{2} \quad (8)$$

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

### 3. Results

### 4. Conclusions

This study demonstrates the improvement that has taken place over the past two decades for predicting viscosity with molecular simulation. First, following “Best Practices” for EMD lead to more reproducible results. Second, the state-of-the-art Mie  $\lambda$ -6 force fields are significantly more accurate than the traditional Lennard-Jones 12-6 force fields for viscosity, as well as for vapor-liquid coexistence properties. More specifically, the Potoff and TAMie force fields typically predict saturated liquid viscosities for *n*-alkanes to within 10 % of the REFPROP values. By contrast, the TraPPE and AUA4 models under predict saturated liquid viscosities by 20 % to 50 %, where the deviations are largest at lower temperatures. While Potoff and TAMie are also more reliable for branched alkanes, deviations are larger and demonstrate a similar temperature dependence.

The key limitation of the Potoff force field is that the choice of  $\lambda = 16$  is too repulsive at close distances, which causes the viscosity to be over estimated at high densities. Due to a fortuitous cancellation of errors, the Potoff potential does provide a reliable  $\eta$ - $P$  trend. Since TAMie uses  $\lambda = 14$ , the  $\eta$ - $\rho$  trend is slightly more reliable than that of Potoff. It is important to emphasize that transport properties were not included in the training set for parameterizing the Potoff and TAMie force fields. Therefore, the results from this study demonstrate that the improved prediction of static vapor-liquid coexistence properties obtained with Mie  $\lambda$ -6 potentials also results in improved prediction of a dynamic property, liquid viscosity.

### Supporting Information

Section ?? provides GROMACS input files. Section ?? enumerates all systems simulated. Section ?? details the MD integrator, thermostat, and barostat. Section ?? compares simulation results with varying production times. Section ?? determines the sensitivity of our results to the use of fixed bonds. Section ?? outlines the Green-Kubo analysis process.



Section ?? validates our methods by comparing with reference simulation values. Section ?? contains tabulated simulation values.

## **Acknowledgments**

We are grateful for the internal review provided by Andrei F. Kazakov and Alta Y. Fang 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.

Commercial equipment, instruments, or materials are identified only in order to adequately specify certain procedures. In no case does such identification imply recommendation or endorsement by NIST, nor does it imply that the products identified are necessarily the best available for the intended purpose.

Partial contribution of NIST, an agency of the United States government; not subject to copyright in the United States.

## **References**