
EXPERIMENTAL EVIDENCE OF QUANTUM DRUDE OSCILLATOR BEHAVIOR IN LIQUIDS REVEALED WITH PROBABILISTIC ITERATIVE BOLTZMANN INVERSION

A PREPRINT

B. L. Shanks^{1,2}, H. W. Sullivan^{2,3}, P. Jungwirth¹, and M. P. Hoepfner²

¹Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, CZH

²University of Utah, *Department of Chemical Engineering*, Salt Lake City, UT

³University of Minnesota, *Department of Chemical Engineering and Materials Science*, Minneapolis, MN

January 14, 2025

ABSTRACT

The first experimental evidence of quantum Drude oscillator behavior in liquids is determined using probabilistic machine learning-augmented iterative Boltzmann inversion applied to noble gas radial distribution functions. Furthermore, classical force fields for noble gases are shown to be reduced to a single parameter through simple empirical relations linked to atomic dipole polarizability. These findings highlight how neutron scattering data can inspire innovative force field design and offer insight into interatomic forces to advance molecular simulations.

Keywords Scattering · Force Fields · Polarizability · Molecular Simulation · Gaussian Processes

1 Introduction

Atomic organization, structure, and self-assembly are foundational concepts in understanding and modeling the behavior of liquids at the atomic scale [1–3]. Solvent structure is known to have a large impact on the behavior of complex liquid phase processes, ranging from protein folding dynamics [4–6], formation of lipid micelles in water [7, 8], to catalysis [9] and material separations [10]. Consequently, the need for molecular models to accurately model liquid structure has increased the significance of obtaining precise experimental measurements to serve as benchmarks or optimization targets, ideally through the gold-standard techniques of wide-angle X-ray and neutron scattering. Scattering experiments probe interatomic distances and interactions by detecting individual scattering events over time, allowing for the sampling of atomic positions within the system at the same length scale as interatomic forces. This capability makes scattering data extremely valuable for the training and refinement of force fields (FFs) [11].

Within a classical liquid state theory paradigm, learning interatomic forces from experimental scattering data directly - the inverse problem in statistical mechanics - is a promising approach to enforce structural behavior in a FF model. Of existing inverse problem methods [1, 11–22], iterative Boltzmann inversion (IBI) is one of the most widely used and easy-to-implement techniques. In a nutshell, IBI refines pair potentials based on their quality-of-fit to a target set of site-site partial radial distribution functions, $g_{\alpha,\beta}(r)$ (e.g., the O-O correlation in water). The algorithm proceeds iteratively until the target structure is reproduced to a pre-specified degree of accuracy.

Although experiencing a renewed interest for training coarse-grained FFs, IBI has its roots in a study on the uniqueness of the pair interaction potential in relation to the pair correlation functions from Henderson [23]. Based on this now dubbed "Henderson's inverse theorem", Wolfram Schommers designed a variational method to extract pair potentials for his studies of liquid gallium [24], which would later evolve into the numerical IBI type methods existing today. Applications of Schommer's algorithm range from neutron scattering analysis with empirical potential structure refinement [25], coarse-grained FF design [26, 27], to correcting structural predictions in machine learning potentials [28]. Recently, more rigorous mathematical formulations of the problem have been explored to try and elucidate fundamental properties of IBI within the framework of functional analysis [29–31].

Of course, IBI has well-documented limitations that restrict its transferability to complex model systems and experiments. For instance, it is generally observed that IBI potentials depend on the chemical environment and thermodynamic state of the training data [32], prompting the general shift towards multi-objective optimization over a range of thermodynamic state points [26]. In the context of the theory from which these algorithms arise, this state-dependence is not surprising. Explicitly, Henderson's inverse theorem only guarantees uniqueness of the pair potential for isotropic fluids with pairwise additive interactions at fixed density and pressure [30]. The former condition is violated for pairwise additive two component mixtures, while both conditions are generally violated for molecular liquids. Another problem is that, unlike atomistic simulations in which the site-site partial radial distribution functions can be computed directly from a trajectory, the experimental equivalent is not computable uniquely [33]. The fact that applying IBI to experimental data already violates assumptions of the theory, and that the target experimental data is potentially an inaccurate representation of the true atomic structure, leaves us in a difficult state-of-affairs for learning potentials in a rigorous and reproducible way.

However, it is now possible to mitigate many of the above challenges with the use of machine learning [34]. Specifically, we recently proposed a probabilistic IBI algorithm, i.e., structure optimized potential refinement (SOPR), to extract transferable forces from experimentally derived RDFs for fluid phase systems. The key advance that SOPR provides is a Gaussian process regression (GPR) step on the refinement equation aimed at mitigating numerical instability and overfitting to uncertain experimental data. GPR enforces continuity, differentiability, and can suppress spurious long-range potential oscillations. Conceptually, GPR takes the standard IBI algorithm and embeds rigorous uncertainty quantification [35] into its interatomic potential prediction. Application of SOPR to neutron scattering data of noble gases (Ne, Ar, Kr, and Xe) has reproduced both structural and thermophysical properties along the vapor-liquid line with accuracy that matches or exceeds state-of-the-art classical force fields [34].

The ability to compute accurate interaction potentials directly from scattering measurements opens the door to deeper insights into interatomic forces in liquids. SOPR potentials, serving as a two-body mean-field approximation of the quantum many-body interactions that govern liquid behavior, overcome the system size limitations of quantum models while remaining fully nonparametric and flexible. The Bayesian nature of the SOPR method also reduces model selection bias and parametric uncertainty, offering an experimentally validated method to quantify deviations from empirical functional forms like Lennard-Jones and to assess the magnitude of many-body effects against *ab initio* dimer potentials. Moreover, analyzing SOPR potentials across thermodynamic states or atomic numbers can reveal valuable links between fundamental atomic properties and classical representations of interatomic forces [36].

The analysis of noble gases provides a compelling test case for this approach. Noble gases largely satisfy the assumptions of the Henderson inverse theorem (excluding pairwise additivity), and their long-standing role in studies of

intermolecular interactions makes them ideal candidates for examining consistencies between classical and quantum models of interatomic forces. To this aim, we deconstructed noble gas SOPR potentials using classical parameterizations of atomic size and dispersion energy and compared them to the atomic dipole polarizability. The atomic dipole polarizability was selected as the variable quantity because it is known to be closely related to atomic size in quantum models of electron polarization [37] and is present in the leading term of the dispersion energy multipole expansion [38]. Empirical investigation of the atomic size derived from SOPR potentials revealed that both collision diameter (the σ parameter in classical force fields) and hard-particle diameter exhibit power law scaling that aligns precisely with the quantum Drude oscillator model of electronic polarization. This result, to our knowledge, is the first observation of quantum Drude oscillator type behavior in the liquid state.

The empirically derived scaling laws were subsequently used to predict force field parameters for noble gases outside the training set, specifically helium (He), radon (Rn), and oganesson (Og). These predictions matched independently optimized force fields up to the estimated uncertainty in the pair potential reconstruction (excluding Og, for which no experimental data of liquid phase properties currently exists). Notably, the excellent quantitative agreement with a Rn force field optimized to critical point data [39] is striking. This suggests that SOPR can be leveraged to predict accurate force field parameters that are fully consistent with quantum representations of atomic dipole polarizability, despite being derived within a purely classical framework. Moreover, this result provides a clear example of dimensionality reduction in force field design, where the problem of developing classical potentials can be reduced to computing a single parameter and applying simple scaling relations learned from experimental structure data. These results demonstrate that nonparametric potentials derived from neutron scattering patterns not only align with quantum mechanical theory but also have the potential to guide both classical and next-generation quantum mechanical force field development in subtle yet powerful ways.

2 Theory and Computational Methods

In practice the pairwise additive approximation has been widely successful at modeling thermodynamic and dynamic properties of liquid state systems at a comparatively low computational cost. Under the additional assumption that the liquid is isotropic, the two-body effective potential is uniquely defined by the radial distribution function according to Henderson’s inverse theorem [23]. This theorem can be shown to be mathematically equivalent to the condition that for any point $r_i \in \mathbb{R}_0^+$, the product of the difference in the effective pair potential, $\Delta v_2(r_i)$, and radial distribution function, $\Delta g(r_i)$, must be non-positive,

$$\Delta v_2(r_i)\Delta g(r_i) \leq 0 \quad (1)$$

establishing a variational framework to refine a reference potential to experimentally derived radial distribution functions. The two-body potential derived from the experimental scattering data is therefore the unique potential that gives an equality in eq 1.

2.1 Structure Optimized Potential Refinement

Structure optimized potential refinement (SOPR) is a probabilistic iterative Boltzmann inversion algorithm designed to learn effective two-body potentials from neutron/X-ray scattering derived radial distribution functions [34]. Whereas most IBI methods are used for coarse-grained simulations [26], SOPR is uniquely designed to handle experimental data. In a nutshell, SOPR takes as inputs a reference pair potential, v_2^{ref} , and experimentally derived radial distribution function, g^{exp} , and computes an effective pair potential, v_2^{eff} , that reproduces the given experimental data (Figure 1). This computation is performed via iterative potential refinement, in which (1) a molecular simulation using the given reference potential is used to produce a simulated radial distribution function, g^{sim} , (2) the difference between the simulated and experimental radial distributions functions is calculated, $\Delta g = g^{sim} - g^{exp}$, and (3) the pointwise Henderson inverse theorem is applied to estimate an updated potential, $v_2^{up'}$, for the next iteration via a refinement equation,

$$v_2^{up'}(r) = v_2^{ref}(r) + \gamma(r)\beta^{-1} \sum_n \Delta g^{(n)'}(r) \quad (2)$$

where i is the radial index, n is the iteration number, β is the inverse thermal energy $(k_B T)^{-1}$, and $\gamma(r_i) > 0$ is an empirical scaling function used to dampen the potential correction. Note that eq (2) is just one refinement equation example that satisfies eq (1) and modifications to this function can lead to changes in convergence speed and stability.

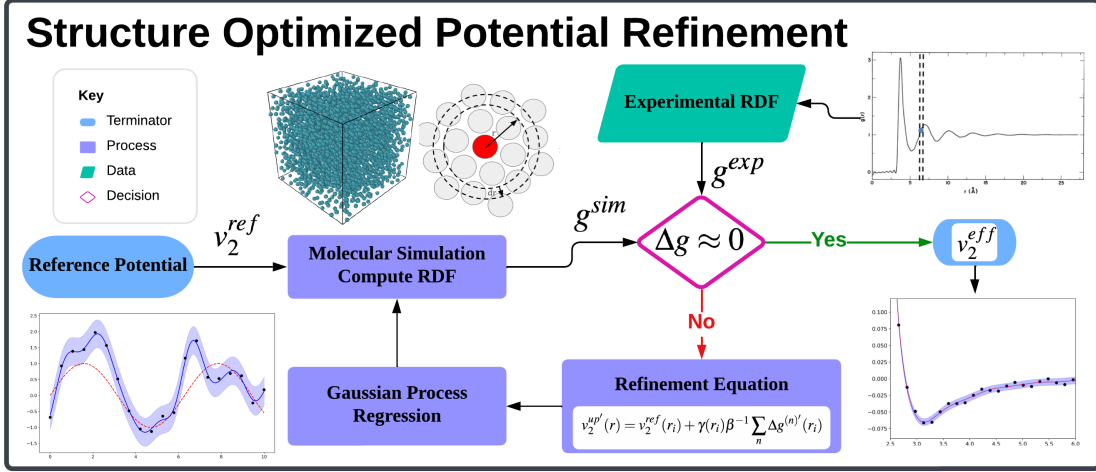


Figure 1: General overview of the structure optimized potential refinement algorithm. The loop between the molecular simulation, potential refinement equation, and GPR is always convergent to a unique potential for pairwise additive and isotropic systems according to Henderson’s inverse theorem.

At this stage, traditional IBI repeats steps (1)-(3) until the simulated and experimental radial distribution functions converge to each other ($\Delta g \approx 0$ for all r). However, SOPR introduces an additional machine learning step (4) Gaussian process regression (GPR) on the updated potential. Conceptually, we are assuming that the updated potential is a *sample* from a distribution of possible pair potential functions that are approximately distributed as an infinite dimensional multivariate Gaussian over its input parameters,

$$v_2^{up'} \sim \mathcal{N}(\mu(\mathbf{r}), \mathbf{K}(r_i, r_j)) \quad (3)$$

where r_i, r_j are any pair of radial positions, $\mu(\mathbf{r})$ is a mean function, and $\mathbf{K}(r_i, r_j)$ is a covariance function (or kernel) describing the relatedness of observations $v_2^{up'}(r_i)$ on $v_2^{up'}(r_j)$ [35]. The specification of the mean and kernel functions are where we can enforce physics based knowledge such as continuity and differentiability (i.e., smoothness) and long-range tail behavior ($\lim_{r \rightarrow \infty} v(r) = 0$). It is then a straightforward matrix calculation to compute the most probable potential given the iterative potential refinement estimate as,

$$v_2^{up}(r') = \mathbf{K}_{r',r}^T [\mathbf{K}_{r,r} + \sigma_{noise}^2 \mathbf{I}]^{-1} v_2^{up'}(r) \quad (4)$$

where $v_2^{up}(r')$ is the estimate for the structure-optimized potential at iteration n , σ_{noise}^2 is the variance in the function resulting from random noise, and $\mathbf{K}_{r',r}$ is shorthand notation for the kernel function. Generally, the σ_{noise}^2 hyperparameter can be inferred using hierarchical Bayesian inference for each regression step, but for our purposes it was sufficient to choose a value of 0.01 to capture known noise levels observed during refinement. The GPR step has been shown to reduce numerical instability and enforce physically justified behavior in the refined potentials, allowing SOPR to address well-known problems with IBI such as artifacts associated with intermediate and long-range structural correlations and potential non-uniqueness in heterogeneous systems (i.e., molecular liquids and mixtures).

In this work, SOPR was applied to neutron scattering data for four noble gas species (Ne, Ar, Kr, Xe) [40–42] to learn nonparametric potentials for subsequent analysis. Details of the molecular simulations, SOPR parameters such as the reference potentials and scaling function, mean and kernel selection for GPR regression, and validation of the resulting potentials can be found in the Appendix.

2.2 Short Range Interactions: Atom Size Estimation

In classical force field design, the atomic size parameter, often the "collision diameter" or σ , is used in empirical potentials like the (12-6) Lennard-Jones (LJ) and $(\lambda - 6)$ Mie potentials. The collision diameter represents the repulsive core of the atom, where Pauli exclusion causes strong repulsion. σ is typically defined as the diameter where the pair potential shifts from positive to negative at short range. Interestingly, σ can vary significantly across different FFs (c.f.

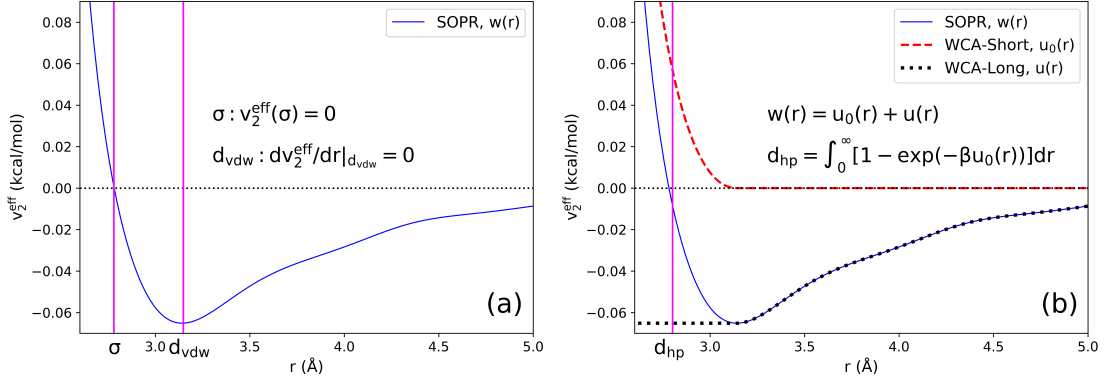


Figure 2: WCA separation of the SOPR potential into a short-range repulsive reference potential and long-range perturbation.

Madrid 2019 [43] and CHARMM [44]), which can easily result in a poor quality of fit to experimentally derived radial distribution functions. For instance, it has been estimated that the uncertainty in the σ parameter is less than 0.1 \AA when using structure factor measurements as an optimization target [45]. This means that two classical FFs with σ values differing by more than 0.1 \AA are likely to produce significantly different structure factor predictions, with at least one diverging notably from experiment.

The van der Waal (vdW) diameter is another important quantity in quantum mechanical calculations that often appears in density functional theory (DFT) [46]. Specifically, the vdW diameter is directly influential to the so-called "vdW force" present in DFT functionals [47]. The standard definition for the vdW diameter used today is the distance from the center of an atom at which Pauli exclusion and London dispersion forces are balanced. In classical pair potential models, this point occurs at the diameter where the force is zero, typically located at the bottom of the potential well. In DFT, the exchange-correlation energy functional, which describes interactions of many-particles in the standard Kohn-Sham DFT method, includes vdW forces in an approximate manner [47]. By directly computing nonparametric classical potentials from experimental data, we can compare vdW diameter estimates in many-body liquid-phase systems with those obtained from quantum mechanical calculations. While the idea to learn vdW diameters from scattering data dates back to the work of Bondi [48], to our knowledge, no attempts have been made so far to estimate them using nonparametric potentials derived from scattering.

Finally, for model developers interested in entropy driven self-assembly [49], the effective hard-particle diameter defines particle size. Here the hard-particle diameter was determined using the Weeks-Chandler-Andersen (WCA) perturbation theory [50], in which the total potential, $w(r)$, is separated into two parts: a short-range repulsive reference, $u_0(r)$ and a long-range attractive perturbation $u(r)$,

$$w(r) = u_0(r) + u(r). \quad (5)$$

such that,

$$u_0(r) = \begin{cases} w(r) + \varepsilon & \text{if } r < r_\varepsilon, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

$$u(r) = \begin{cases} -\varepsilon & \text{if } r < r_\varepsilon, \\ w(r) & \text{otherwise.} \end{cases} \quad (7)$$

where ε is the pair potential well depth and r_ε is the radial position corresponding to the potential energy ε . Eq (5) is shown graphically in Figure 2(b) for the SOPR potential of liquid neon.

Once the potentials are decomposed using the WCA separation, an equivalent hard-particle diameter, d_{hp} , can be estimated from the repulsive part of the potential [3, 50, 51] according to the following equation,

$$d_{hp} = \int_0^\infty [1 - \exp(-\beta u_0(r))] dr \quad (8)$$

such that the integrand $[1 - \exp(-\beta u_0(r))]$ is unity at low r since $\exp(-\beta u_0(r))$ is negligibly small due to the y -asymptotic behavior of $u_0(r)$. At longer range, $u_0(r)$ is zero by definition, which sends the integrand to zero. Note that this hard-particle diameter falls out of a second order approximation of the functional Taylor expansion of the reduced excess free energy density, $\phi = -\beta F^{ex}/V$, in powers of the blip function, $\Delta e(r)$, defined as the difference between Boltzmann factors of the short-range repulsive reference and a system of equivalent hard spheres,

$$\Delta e(r) = \exp(-\beta u_0(r)) - \exp(-\beta u_d(r)) \quad (9)$$

where $u_d(r)$ is an equivalent hard-particle reference potential. The beauty of performing this analysis on SOPR potentials is that the soft-core reference system, empirical hard-particle size, and blip function can now be estimated directly from experiments, allowing for comparisons between experimentally inferred potentials and those derived from liquid state theory.

The reason for selecting these definitions of atomic size to describe short-range interactions are three-fold. First, the collision diameter, vdW diameter, and hard-particle diameter are common parameters across a wide range of computational models, covering classical FFs, DFT, and hard-particle Monte Carlo, respectively. In effect, any relationships derived from these three parameters will be widely applicable across numerous fields of computational molecular science. Second, we wanted to leverage the low computational cost of SOPR to see if certain size parameters follow known theoretical behaviors while others do not. Such an undertaking would be far more expensive for existing *ab initio* methods due to the large number of atomic clusters required to represent a liquid phase. Finally, it is not known whether or not any trends exist in empirical atomic size parameters like the collision diameter and hard-particle diameter. Elucidating these relationships could provide an experimentally validated FF design strategy or be used as initial guess for optimization.

2.3 Long Range Interactions

The treatment of long-range interactions differs substantially between classical and quantum mechanical models of interatomic forces, complicating direct comparisons between SOPR potentials and quantum models of dispersion interactions. In this work, we approximate the classical dispersion energy as the minimum energy of the SOPR pair potential, equivalent to the ϵ parameter in the familiar Lennard-Jones model, and perform linear regression on a power law relation of the form, $\epsilon = a\alpha^b$, where $a, b \in \mathbb{R}$. As demonstrated later in the manuscript, even this simplified approach consistently reproduces the ϵ parameter for He and Rn within the expected parametric uncertainty of structure-derived potentials.

2.4 Atomic Dipole Polarizability and Quantum Drude Oscillators

Atoms and molecules are electrically polarizable particles, and their degree of polarization is well-known to correlate with their effective size [52]. In condensed phases, where interatomic interactions significantly impact the static and dynamic behavior of the system, electronic polarization effects can become significant in response to the chemical environment. These effects become increasingly important in biomolecular systems, such as interactions between ions and proteins/lipids, charged amino acids, and stabilization of protein structures through salt bridges, among others. Therefore, there is currently a widespread effort to incorporate electronic polarization models in FFs using both explicit methods, such as fluctuating charge models or classical drude oscillators [53], and mean-field electron continuum corrections (ECC) [54–57].

Of the explicit models of electron polarization, the quantum Drude oscillator (QDO) is emerging as a popular method to describe electronic polarization from first principles [58–60]. In the QDO model, the interaction between the electrons and nucleus is represented as a harmonic term between a *drudon* quasiparticle with charge $-q$ and a classical quasinucleus of charge $+q$. Each quasiparticle interacts with other particles in the system through Coulombic forces. In a spherically symmetric QDO, polarizability can be calculated from second-order perturbation theory by constructing a test charge perturbed Hamiltonian and solving for the induced multipole moments [61]. Of particular interest to the biomolecular FF community is the construction of computationally efficient two-body interatomic potentials consistent with QDO models [62].

Another benefit of the QDO model is its simple mathematical structure, which makes it easy to derive relationships between atomic properties [63]. For example, Federov *et al.* [37] showed that the vdW diameter is related to atomic dipole polarizability via a power law,

$$R_{vdW} \approx a\alpha^{1/7} \quad (10)$$

where α is polarizability and $a \approx 2.54$ is a constant fit to target vdW diameter data. It has been argued that this relation, which touts reasonably accurate predictions for 72 elements, obviates the need to compute vdW diameters since the atomic dipole polarizability is easily measured experimentally [64–66] or computed [67, 68]. Note that this relation differs from the classical drude oscillator (CDO) solution, which scales with $R \propto \alpha^{1/3}$.

These scaling laws provide a convenient consistency check between SOPR potentials and quantum mechanical representations of electronic polarization. To this aim, we fit the QDO power law scaling in eq (10) to SOPR estimated vdW diameters, collision diameters, and hard-particle diameters to elucidate whether or not the structure-derived potentials are consistent with quantum mechanical behavior of electronic polarization. The resulting empirical QDO power relation was then used to estimate collision diameters for the noble gases He and Rn.

3 Results and Discussion

Our methodology combines the following four steps: (1) inference of the interatomic potentials using SOPR, (2) calculation of the σ parameter, vdW diameter, hard-particle diameter, and dispersion energy (3) linear regression of power law relations between these parameters and atomic dipole polarizability, and (4) application of the resulting relations to estimate σ and ϵ parameters in noble gases outside of the training set. In a nutshell, the results show that the SOPR-QDO line derived solely from liquid noble gas scattering data is quantitatively accurate at predicting optimal σ and ϵ parameters for classical models of noble gases.

3.1 Analysis of SOPR Potentials and QDO Scaling Laws

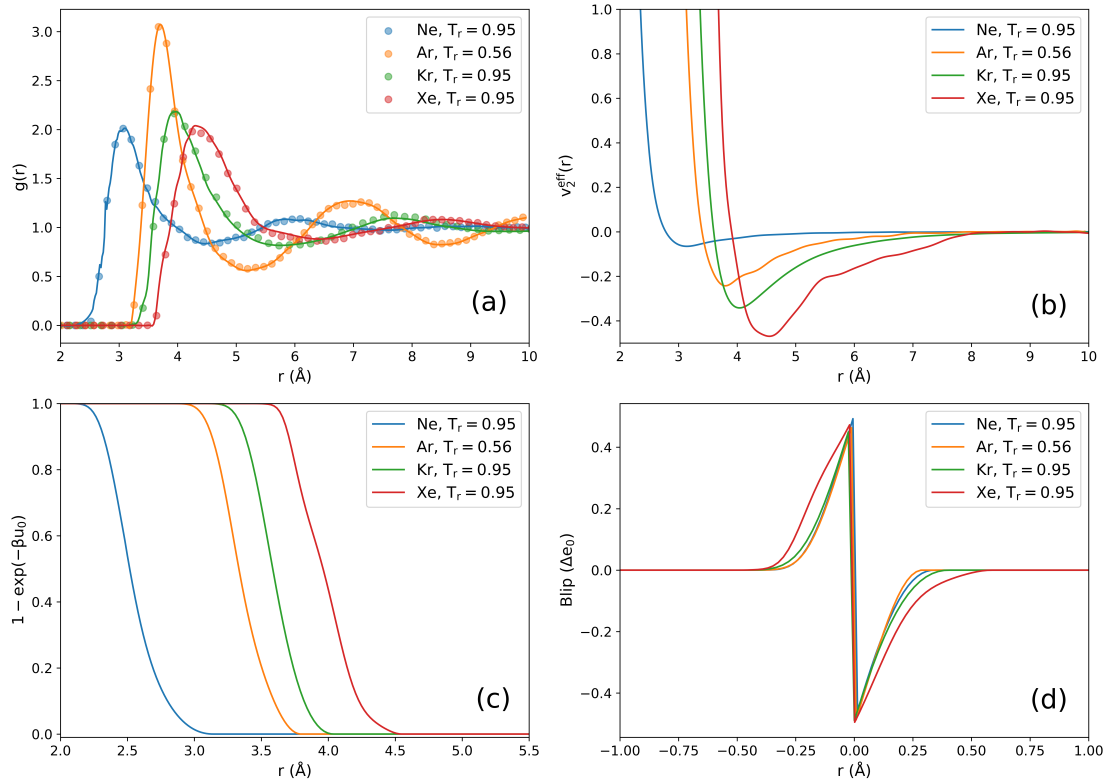


Figure 3: SOPR analysis for Ne, Ar, Kr, and Xe at a single state point. (a) Experimentally derived radial distribution functions compared to SOPR simulation. (b) Converged SOPR potentials. (c) Integrand of the soft-core repulsive potential after Weeks-Chandler-Andersen separation with the Barker-Henderson hard-particle diameter approximation. (d) Neutron scattering derived blip functions.

Radial distribution functions, SOPR potentials, hard-particle integrands, and blip functions calculated from neutron scattering data of Ne, Ar, Kr, and Xe are shown in Figure 3. The excellent quality of fit to the experimental radial

distributions is expected according to the Henderson inverse theorem. Also, the SOPR potentials themselves are well-behaved in comparison to other IBI potentials derived from experimental data [24, 69]. Figure 3 (c) shows the integrand of the Barker-Henderson hard-particle equation after WCA separation of the SOPR potential and (d) the blip function. The widening of the blip function with atom identity indicates a larger deviation from hard-particle behavior, consistent with the progressively larger and more polarizable electron clouds of Ne, Ar, Kr, and Xe, respectively. Atomic vdW diameters, collision diameters, and hard-particle diameters computed from the SOPR potentials are provided in Table 1.

Table 1: Summary of atomic size and dispersion energy parameters determined from SOPR potentials. All lengths are given in units of Å.

Element	σ	vdW	d_{hp}	ϵ (kcal/mol)	α (Å ³) [68]
Ne	2.77	3.15	2.80	0.065	0.396
Ar	3.40	3.79	3.51	0.242	1.661
Kr	3.58	4.05	3.64	0.342	2.528
Xe	3.91	4.55	3.96	0.469	4.118

A comparison between the SOPR potential atomic size predictions and atomic dipole polarizabilities is presented in Figure 4(a) while dispersion energy is presented in Figure 4(b). Power law constants, $a, b \in \mathbb{R}$, were computed using linear least-squares regression to the equation, $c = a\alpha^b$. For the atomic size predictions, the power law exponent, b , was fixed to the QDO value of $1/7$, whereas for the dispersion energy b was treated as an unknown. The power law relations and fitting parameters are recorded in Figure 4 (a-b).

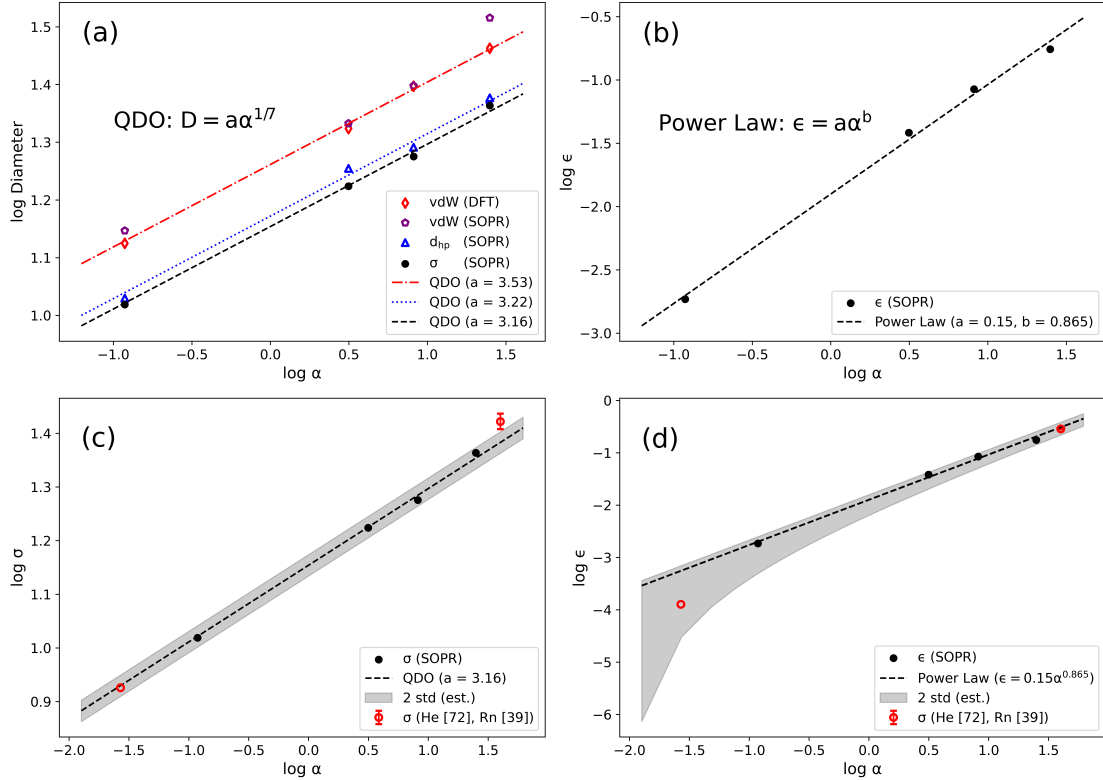


Figure 4: Interatomic potential parameters derived from SOPR compared to a QDO power law (for atomic size) and variable power law relation (for dispersion energy) in the noble gas series. (a) Empirical quantum Drude oscillator scaling law, $D \propto \alpha^{1/7}$, (dashed lines) compared to *ab initio* calculated vdW diameter (red diamonds) and neutron scattering derived vdW diameter (violet pentagons), effective hard particle diameter (blue triangles) and interaction potential collision diameter (black circles). (b) Dispersion energy at the pair potential minimum vs variable power law relation. (c-d) Extrapolation of QDO and variable power laws to atomic dipole polarizabilities of He and Rn parameters with uncertainty estimation. Red markers represent independently optimized force field parameters.

Collision and hard-particle diameters show an excellent linear fit in the log-log plot, with a power law exponent of $1/7$, providing the first experimental evidence of QDO-like behavior experimental data in the liquid-state. In contrast, the SOPR vdW diameter deviates from the prediction by Fedorov *et al.* [37] for liquid Xe, which is expected since the QDO relation, derived in vacuum, does not account for the many-body interactions in the dense polarizable liquid. This discrepancy underscores the limitations of naively applying QDO scaling to classical force fields: while short-range interactions such as collision and hard-particle diameters align with the QDO model, the vdW diameter, influenced by many-body effects, does not.

Furthermore, this deviation may reveal deeper complexities in QDO behavior under liquid-state conditions. One possible explanation is that the solution to the perturbed QDO Hamiltonian for many-body systems at liquid densities does not follow a power law in atomic dipole polarizability, unlike the symmetric single QDO with a test charge perturbation. In this case, the SOPR vdW data may suggest the presence of another term in the solution that grows faster than the $\propto \alpha^{1/7}$ scaling, with this non-power-law term manifesting at polarizabilities between those of Kr and Xe.

Figure 4(c-d) compares empirical scaling relations with classical force field parameters for He and Rn (red markers) with uncertainty quantification estimation (gray credibility interval). Credibility intervals were calculated by real unit scaling of the two standard deviation estimates on the reduced LJ units ($\pm 0.02\sigma$ and $\pm 0.1\epsilon$) reported in our recent study on classical force field uncertainty propagated from noisy scattering data [45]. This gives upper and lower credibility interval bounds for the power law scaling relations of $\sigma_{int} : [3.097\alpha^{1/7}, 3.223\alpha^{1/7}]$ and $\epsilon_{int} : [0.135\alpha^{0.865}, 0.165\alpha^{0.865}]$. Note that these uncertainty estimates assume that the scattering data has a baseline noise level of 0.005 or less out to 30 \AA^{-1} . Error bars for Rn were reported according to Mick *et al.* [39], but we speculate that the true uncertainty is larger due to the multimodal nature of their likelihood function response surface.

Table 2 presents noble gas σ and ϵ parameters estimated using SOPR-derived scaling relations, compared with existing force fields. Percent error (%) is calculated in the usual way, $100(x_{pred} - x_{true})/x_{true}$, where x_{pred} is the SOPR estimated parameter and x_{true} is the independently optimized parameter. The estimates show errors below 3.9% in all cases except for He’s dispersion energy, where the inherent limitations of nuclear quantum effects and the high uncertainty in the estimation make pinpointing a single value difficult. Speculative estimates for the parameters of liquid-phase Og are also included, though they cannot be validated by thermodynamic predictions since Og can only be synthesized one atom at a time [70]. Furthermore, accurate modeling of Og’s interatomic forces requires considering relativistic effects via the Dirac equation [71], suggesting its properties may differ significantly from other noble gases. Despite these challenges, the parameters presented offer a theoretical foundation for investigating liquid-phase Og until experimental data becomes available.

Table 2: Estimated σ and ϵ parameters for the noble gas series using the SOPR-QDO scaling relations. Deviations from previously reported force field parameters are given in percent error.

Element	σ (SOPR)	σ	Error (%)	ϵ (SOPR)	ϵ	Error (%)	α (\AA^3)
He	2.533	2.524 [72]	0.4	0.039	0.020 [72]	89	0.208 [66]
Ne	2.768	2.794 [73]	-0.9	0.065	0.064 [73]	1.5	0.396 [68]
Ar	3.393	3.405 [73]	-0.4	0.242	0.242 [73]	-0.1	1.661 [68]
Kr	3.600	3.645 [73]	-1.2	0.342	0.350 [73]	-2.3	2.528 [68]
Xe	3.858	3.964 [73]	-2.7	0.469	0.484 [73]	-3.1	4.118 [68]
Rn	3.986	4.145 [39]	-3.8	0.600	0.580 [39]	3.4	4.965 [68]
Og	4.297	-	-	0.964	-	-	8.590 [71]

Our results reveal a clear connection between structure-derived potentials and fundamental quantum mechanical theories of interatomic forces. Indeed, the consistency between SOPR-derived scaling relations and the QDO model challenges the traditional view that structure-derived potentials are mere empirical representations of interatomic interactions [14]. Looking forward, physical insights gained from scattering analysis could be crucial for classical FF design, particularly in complex biomolecular systems where standardized prescriptions and benchmarks for interatomic interactions are currently lacking [74].

4 Conclusions

In summary, collision and hard-particle diameters inferred from neutron scattering experiments in noble liquids were found to follow a quantum Drude oscillator scaling law with atomic dipole polarizabilities. This finding represents the first experimental evidence of QDO-type behavior in the liquid state. Additionally, the classical description of noble gas interatomic forces was shown to depend essentially on a single parameter: the atomic dipole polarizability. Empirical

scaling laws on the classical σ and ϵ parameters derived from SOPR potentials were then shown to closely reproduce state-of-the-art classical models of noble gases ranging from helium to radon.

More broadly, the methodology presented in this work has the potential to advance research on inferring interatomic potentials from structural data. For many years, structure inversion techniques have been regarded as unable to yield physically accurate representations of interatomic interactions. Additionally, the focus on many-body and quantum mechanical methods has caused interest in developing approaches for investigating classical pairwise representations of interatomic forces to decline. Nevertheless, this study demonstrates that nonparametric inference of interaction potentials from high-resolution neutron scattering data can reveal subtle and complex aspects of intermolecular forces. It also highlights why machine learning interpretations of scattering experiments are instrumental in this respect.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

This study is supported by the EFRC-MUSE, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award No. DE-SC0019285.

Author Contributions

B.L.S: conceptualization, algorithm development and implementation, data analysis/visualization, and writing. H.W.S: algorithm implementation. P. Jungwirth: conceptualization, writing. M.P.H: conceptualization, funding acquisition.

Competing Interest

The authors declare no conflicts or competing interests.

5 Appendix

A Structure Optimized Potential Refinement Settings and Molecular Dynamics

The molecular simulation corrector is a Canonical (NVT) molecular dynamics (MD) simulation performed in HOOMD-Blue [75]. MD simulations were initiated with $N = 2916$ atoms on a cubic lattice of initial length of $3 * N^{1/3}$ Å and performed in the NVT ensemble using HOOMD’s native ConstantVolume integrator and Nose-Hoover thermostat with a default coupling constant of $\tau = 100\delta t$. The timestep, δt , was chosen to be 0.5 femtoseconds and pair potential interactions were truncated at 3σ with analytical tail corrections. MD simulations proceeded in four steps (1) thermalization of particle momenta for 5000 timesteps, (2) box size ramping to the experimental density using the `hoomd.variant.Ramp` command for 10,000 timesteps, (3) a 0.5 ns equilibration at the experimental density, and (4) a 0.5 ns production run. Radial distribution functions were calculated with Freud [76] sampled at 100 timestep intervals. Convergence is checked against the average squared error between the simulated and experimental radial distribution function such that $\langle [\Delta g^{(n)}(r)]^2 \rangle < 5 * 10^{-4}$, which generally took 15-20 iterations at SOPR scaling constant $\gamma = 0.2$.

Table 3: Reduced temperature ($T_r = T/T_c$) and atomic density (ρ) are listed for the neutron scattering experimental conditions. 12-6 Lennard-Jones potentials with parameters ($\sigma_{ai}, \epsilon_{ai}$) were used as reference potentials.

Element	T_r	ρ ($1/\text{Å}^3$)	σ_{ai} (Å)	ϵ_{ai} (kcal/mol)
Ne	0.95	0.02477	2.76	0.122
Ar	0.56	0.02125	3.35	0.287
Kr	0.95	0.01187	3.58	0.582
Xe	0.95	0.00881	3.89	0.811

The Gaussian process regression step used mean zero, $\mu = 0$, and a squared-exponential kernel with width parameter, $w = 0.05$ kcal/mol and length-scale parameter, $\ell = 1$ Å. GPR regression was initiated at 0.9σ which is well within the region of the potential that influences the MD. Experimental neutron scattering data and SOPR reference potentials are the same as the SOPR method paper [34] and reproduced below in Table 3. Source code for SOPR is available on GitHub here.

References

- (1) Ornstein, L. S.; Zernike, F. *Proc. Acad. Sci. Amsterdam*. **1914**, *17 II*, 793–806.
- (2) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 774–777.
- (3) Hansen, J.; McDonald, I. R., *Theory of Simple Liquids: With Applications to Soft Matter*; Academic Press: 2013.
- (4) Florová, P.; Sklenovský, P.; Banáš, P.; Otyepka, M. *J. Chem. Theory Comput.* **2010**, *6*, 3569–3579.
- (5) Smith, M. D.; Rao, J. S.; Segelken, E.; Cruz, L. *J. Chem. Inf. Model.* **2015**, *55*, 2587–2595.
- (6) Anandakrishnan, R.; Izadi, S.; Onufriev, A. V. *J. Chem. Theory Comput.* **2019**, *15*, 625–636.
- (7) Brodskaya, E. N.; Vanin, A. A. *J. Chem. Phys.* **2015**, *143*, 044707.
- (8) Deshmukh, S. A.; Solomon, L. A.; Kamath, G.; Fry, H. C.; Sankaranarayanan, S. K. R. S. *Nat. Commun.* **2016**, *7*, 12367.
- (9) Li, P.; Jiang, Y.; Hu, Y.; Men, Y.; Liu, Y.; Cai, W.; Chen, S. *Nat. Catal* **2022**, *5*, 900–911.
- (10) Kingsbury, R. S.; Zhu, S.; Flotron, S.; Coronell, O. *ACS Appl. Mater. Interfaces* **2018**, *10*, 39745–39756.
- (11) Shanks, B. L.; Sullivan, H. W.; Shazed, A. R.; Hoepfner, M. P. *J. Chem. Theory Comput.* **2024**, *20*, 3798–3808.
- (12) Levesque, D.; Weis, J. J.; Reatto, L. *Phys. Rev. Lett.* **1985**, *54*, 451–454.
- (13) Tóth, G. *J. Chem. Phys.* **2001**, *115*, 4770–4775.
- (14) Tóth, G. *J. Phys. Condens. Matter* **2007**, *19*, 335220.
- (15) Percus, J. K.; Yevick, G. J. *Phys. Rev.* **1958**, *110*, 1–13.
- (16) Wang, Q.; Keffer, D. J.; Nicholson, D. M.; Thomas, J. B. *Phys. Rev. E* **2010**, *81*, 061204.
- (17) A. Goodall, R. E.; A. Lee, A. *Soft Matter* **2021**, *17*, 5393–5400.
- (18) Mullinax, J. W.; Noid, W. G. *Phys. Rev. Lett.* **2009**, *103*, 198104.
- (19) Dunn, N. J. H.; Noid, W. G. *J. Chem. Phys.* **2016**, *144*, 204124.
- (20) DeLyser, M.; Noid, W. G. *J. Chem. Phys.* **2020**, *153*, 224103.
- (21) Lyubartsev, A. P.; Laaksonen, A. *Phys. Rev. E* **1995**, *52*, 3730–3737.
- (22) Tóth, G. *J. Chem. Phys.* **2003**, *118*, 3949–3955.
- (23) Henderson, R. L. *Phys. Lett. A* **1974**, *49*, 197–198.
- (24) Schommers, W. *Phys. Rev. A* **1983**, *28*, 3599–3605.
- (25) Soper, A. K. *Chem. Phys.* **1996**, *202*, 295–306.
- (26) Moore, T. C.; Iacovella, C. R.; McCabe, C. *J. Chem. Phys.* **2014**, *140*, 224104.
- (27) Wang, K. W.; Wang, Y.; Hall, C. K. *Fluid Ph. Equilib.* **2020**, *521*, 112704.
- (28) Matin, S.; Allen, A. E. A.; Smith, J.; Lubbers, N.; Jadrach, R. B.; Messerly, R.; Nebgen, B.; Li, Y. W.; Tretiak, S.; Barros, K. *J. Chem. Theory Comput.* **2024**, *20*, 1274–1281.
- (29) Hanke, M. *J. Stat. Phys.* **2018**, *170*, 536–553.
- (30) Frommer, F.; Hanke, M.; Jansen, S. *J. Math. Phys.* **2019**, *60*, 093303.
- (31) Frommer, F.; Hanke, M. *Lett. Math. Phys.* **2022**, *112*, 71.
- (32) Rosenberger, D.; Hanke, M.; van der Vegt, N. F. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 1323–1345.
- (33) Soper, A. K. *J. Phys. Condens. Matter* **2007**, *19*, 415108.
- (34) Shanks, B. L.; Potoff, J. J.; Hoepfner, M. P. *J. Phys. Chem. Lett.* **2022**, *13*, 11512–11520.
- (35) Rasmussen, C. E.; Williams, C. K. I., *Gaussian Processes for Machine Learning*; MIT Press: 2006.
- (36) Shanks, B. L. Uncertainty-aware liquid state modeling from experimental scattering measurements, arXiv:2408.00236, 2024.
- (37) Fedorov, D. V.; Sadhukhan, M.; Stöhr, M.; Tkatchenko, A. *Phys. Rev. Lett.* **2018**, *121*, 183401.
- (38) London, F. *Trans. Faraday Soc.* **1937**, *33*, 8b.
- (39) Mick, J. R.; Soroush Barhaghi, M.; Potoff, J. J. *J. Chem. Eng. Data* **2016**, *61*, 1625–1631.
- (40) Yarnell, J. L.; Katz, M. J.; Wenzel, R. G.; Koenig, S. H. *Phys. Rev. A* **1973**, *7*, 2130–2144.

- (41) Bellissent-Funel, M. C.; Buontempo, U.; Filabozzi, A.; Petrillo, C.; Ricci, F. P. *Phys. Rev. B* **1992**, *45*, 4605–4613.
- (42) Barocchi, F.; Chieux, P.; Magli, R.; Reatto, L.; Tau, M. *Phys. Rev. Lett.* **1993**, *70*, 947–950.
- (43) Zeron, I. M.; Abascal, J. L. F.; Vega, C. *J. Chem. Phys.* **2019**, *151*, 134504.
- (44) Mackerell Jr., A. D. *J. Comput. Chem.* **2004**, *25*, 1584–1604.
- (45) Shanks, B. L.; Sullivan, H. W.; Hoepfner, M. P. *J. Phys. Chem. Lett.* **2024**, 12608–12618.
- (46) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515–524.
- (47) Berland, K.; Cooper, V. R.; Lee, K.; Schröder, E.; Thonhauser, T.; Hyldgaard, P.; Lundqvist, B. I. *Rep. Prog. Phys.* **2015**, *78*, 066501.
- (48) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.
- (49) Glotzer, S. C.; Solomon, M. J. *Nat. Mater.* **2007**, *6*, 557–562.
- (50) Weeks, J. D.; Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1971**, *54*, 5237–5247.
- (51) Barker, J. A.; Henderson, D. *J. Chem. Phys.* **1967**, *47*, 4714–4721.
- (52) Mandel, M.; Mazur, P. *Physica* **1958**, *24*, 116–128.
- (53) Lamoureux, G.; MacKerell, A. D.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185–5197.
- (54) Leontyev, I. V.; Stuchebrukhov, A. A. *J. Chem. Theory Comput.* **2012**, *8*, 3207–3216.
- (55) Kirby, B. J.; Jungwirth, P. *J. Phys. Chem. Lett.* **2019**, *10*, 7531–7536.
- (56) Cruces Chamorro, V.; Jungwirth, P.; Martinez-Seara, H. *J. Phys. Chem. Lett.* **2024**, *15*, 2922–2928.
- (57) Nencini, R.; Tempa, C.; Biriukov, D.; Riopedre-Fernandez, M.; Cruces Chamorro, V.; Polák, J.; Mason, P. E.; Ondo, D.; Heyda, J.; Ollila, O. H. S.; Jungwirth, P.; Javanainen, M.; Martinez-Seara, H. *J. Chem. Theory Comput.* **2024**, *20*, 7546–7559.
- (58) Whitfield, T. W.; Martyna, G. J. *J. Chem. Phys.* **2007**, *126*, 074104.
- (59) Jones, A.; Cipcigan, F.; Sokhan, V. P.; Crain, J.; Martyna, G. J. *Phys. Rev. Lett.* **2013**, *110*, 227801.
- (60) Hermann, J.; DiStasio, R. A.; Tkatchenko, A. *Chem. Rev.* **2017**, *117*, 4714–4758.
- (61) Jones, A. P.; Crain, J.; Sokhan, V. P.; Whitfield, T. W.; Martyna, G. J. *Phys. Rev. B* **2013**, *87*, 144103.
- (62) Khabibrakhmanov, A.; Fedorov, D. V.; Tkatchenko, A. *J. Chem. Theory Comput.* **2023**, *19*, 7895–7907.
- (63) Tkatchenko, A.; Fedorov, D. V.; Gori, M. *J. Phys. Chem. Lett.* **2021**, *12*, 9488–9492.
- (64) Hohm, U.; Kerl, K. *Mol. Phys.* **1990**, *69*, 803–817.
- (65) Minemoto, S.; Sakai, H. *J. Chem. Phys.* **2011**, *134*, 214305.
- (66) Gaiser, C.; Fellmuth, B. *Phys. Rev. Lett.* **2018**, *120*, 123203.
- (67) Mitroy, J.; Safronova, M. S.; Clark, C. W. *J. Phys. B* **2010**, *43*, 202001.
- (68) Gobre, V. V. Efficient Modelling of Linear Electronic Polarization in Materials Using Atomic Response Functions, PhD Dissertation, Berlin, Germany: Technische Universitaet Berlin, 2016.
- (69) Soper, A. K. *Phys. Rev. B* **2005**, *72*, 104204.
- (70) Smits, O. R.; Düllmann, C. E.; Indelicato, P.; Nazarewicz, W.; Schwerdtfeger, P. *Nat. Rev. Phys.* **2024**, *6*, 86–98.
- (71) Smits, O. R.; Mewes, J.; Jerabek, P.; Schwerdtfeger, P. *Angew. Chem. Int. Ed.* **2020**, *59*, 23636–23640.
- (72) Tchouar, N.; Benyettou, M.; Kadour, F. O. *Int. J. Mol. Sci.* **2003**, *4*, 595–606.
- (73) Mick, J. R.; Soroush Barhaghi, M.; Jackman, B.; Rushaidat, K.; Schwiebert, L.; Potoff, J. J. *J. Chem. Phys.* **2015**, *143*, 114504.
- (74) Van der Spoel, D. *Curr. Opin. Struct. Biol.* **2021**, *67*, 18–24.
- (75) Anderson, J. A.; Glaser, J.; Glotzer, S. C. *Comput. Mater. Sci.* **2020**, *173*, 109363.
- (76) Ramasubramani, V.; Dice, B. D.; Harper, E. S.; Spellings, M. P.; Anderson, J. A.; Glotzer, S. C. *Comput. Phys. Commun.* **2020**, *254*, 107275.