

Electronic coarse graining: Predictive atomistic modeling of condensed matter

F. S. Cipcigan and J. Crain*

IBM Research UK, Daresbury Laboratory, Keckwick Lane,
Warrington WA4 4AD, United Kingdom

V. P. Sokhan

STFC Daresbury Laboratory, Keckwick Lane,
Warrington WA4 4AD, United Kingdom

G. J. Martyna†

Pimperl, Science Software and Materials Science, Westchester, New York 10598, USA

 (published 10 May 2019)

Atoms and molecules adapt to their environment through a rich hierarchy of electronic responses. These include dipolar many-body polarization contributions arising in the classical limit, many-body polarization beyond dipole order, as well as pair and many-body dispersion interactions and cross interactions at all orders arising from multipolar quantum fluctuations. Such fundamental phenomena give rise to emergent behavior across the physical and life sciences. However, their incorporation in simulations of large complex systems faces significant challenges as these are intrinsically many-body phenomena. Here the impetus for and development of a new class of molecular model employing embedded quantum Drude oscillators (QDO) as a coarse-grained but complete representation of electronic responses at long range within Gaussian statistics is given. The resulting level of completeness in physical description enables isolated molecule properties to define model parameters, thereby eliminating fitting to condensed phase data. This provides a physical and intuitive basis for predictive, next-generation simulation wherein all long-range diagrams emerge naturally from the model permitting the study of complex systems in novel environments. The model is derived from a many-body Hamiltonian and would afford no advantage without an $\mathcal{O}(N)$ scaling, *strong coupling* solution to avoid artificial truncation from perturbation theory and associated multipolar expansions which is possible due to the model's Gaussian structure. A scalar field theory and path integral form of the QDO Hamiltonian cast in such a way as to generate a strong coupling solution to the coarse-grained electronic structure are presented. Forces can be generated “on the fly” using modern adiabatic molecular dynamics methods with linear computational complexity. Thus, the approach is applicable to large condensed phase systems at finite temperature and pressure. Example applications and future perspectives are presented for key physical systems such as the phase diagram of water from ice to the supercritical regime.

DOI: [10.1103/RevModPhys.91.025003](https://doi.org/10.1103/RevModPhys.91.025003)

CONTENTS

I. Introduction	2	1. Adiabaticity: Decoupling of nuclear and electronic motion	4
A. The origin of intermolecular forces	2	2. Many-body effects and the Born-Oppenheimer surface	6
B. Challenges in modeling complex soft condensed matter at the molecular scale	3	3. Separation of interaction length scales	6
1. Fundamental compromises	3	B. Multipole electrostatics and induction in molecular systems	7
2. The need for new strategies	4	1. Electrostatic multipoles	7
II. Long-range Electronic Responses of Atoms and Molecules: Electrostatic, Inductive, and Dispersive Effects	4	2. Symmetry considerations	8
A. Many-body intermolecular forces and the emergence of natural length scales	4	3. Multipoles in nonuniform fields	10
		4. Induction and many-body linear responses	11
		C. Dispersion: Quantum mechanical prigns and many-body effects	12
		D. Breakdown of transferability and a pathway to next-generation strategies	14
		III. Electronic Coarse Graining with Quantum Oscillators	14
		A. A minimal model with complete long-range interactions	14

*jason.crain@ibm.com

†Some of the work incorporated into this review was performed at IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA.

1. Polarization responses	15
2. Zero-point fluctuations and the emergence of dispersion	16
B. Model invariants for the QDO and the behavior of real atoms and molecules	18
1. Relationships between electronic responses	18
2. Parametrization	18
3. Scaling of the response functions	19
C. A route to real systems: Damping and short-range repulsion	20
IV. Solving the Coarse-grained Electronic Problem	21
A. Path integral molecular dynamics for the quantum Drude oscillator	21
1. Condensed phases and the density matrix	21
2. Path integral formulation of the density matrix	22
3. Approximate density matrices, staging, and dynamics	23
V. Application Examples and Physical Insights	25
A. Noble gas solids and liquids as proof of concept	25
B. Liquid water	25
1. Existing water models	25
2. <i>Ab initio</i> models of water	26
3. Fixed-charge models	26
4. A new molecular model of water based on quantum Drude oscillators	27
5. Cluster energies of QDO water: A basic benchmark	28
6. Transferability of QDO water	28
7. Liquid-vapor coexistence	29
8. Ice II: A proton ordered ice	29
9. Supercritical water	30
10. Supercooled water	30
VI. Future Perspectives	31
A. Toward complex physical systems	31
Acknowledgments	31
Appendix A: Multipole Expansion	31
1. Approximate potentials at long range: Cartesian multipole expansion for a charge distribution	31
2. Energy of a charge distribution in an external field	32
3. Interactions between general multipoles	33
4. Spherical multipoles	33
Appendix B: Cluster Energies	36
References	39

I. INTRODUCTION

We begin by describing the fundamental challenges presented by intermolecular forces in condensed phase systems. We then propose and develop a solution based on a strong coupling solution of a model within Gaussian statistics to sum diagrams to all orders and allow high accuracy modeling of complex systems in nontrivial environments.

A. The origin of intermolecular forces

Atoms and molecules are composite objects consisting of localized heavy nuclei and delocalized light electrons which when assembled form complex charge distributions. Indeed, the rich phenomenology of intermolecular interactions emerges from this simple picture: The application of quantum theory to atoms and molecules (Heitler and London, 1927) quantified the phenomenology and led to an in-depth understanding. Within

the adiabatic limit or the Born-Oppenheimer (BO) approximation (Born and Oppenheimer, 1927), the province of this review, the Hellmann-Feynman theorem (Hellmann, 1937; Feynman, 1939) formally defines the prescription for determining intermolecular forces. Given the electron distribution on a single electronic surface within an assembly of molecules at some instant via the solution to the coupled electron-nuclear Schrödinger equation (i.e., in the adiabatic limit), the forces follow as analytic derivatives of the potential averaged over the electron distribution.

The consequences of connecting quantum theory to assemblies of atoms and molecules are far reaching. The time evolution of complex molecular systems under the action of such forces is governed by the classical mechanics of Newton and Hamilton (Goldstein, Poole, and Safko, 2001) or for light atoms, by the nuclear Schrödinger equation, with statistical properties available via the seminal work of Gibbs (1902). These properties can then be compared with and validated by experimental results as a test of the underlying model, thereby closing the loop. Importantly, such an approach allows us to understand how molecular properties and motion affect the behavior of condensed phases as well as to predict the properties of materials not yet experimentally synthesized.

In practice, however, exact solutions to the Schrödinger equation, even within the Born-Oppenheimer approximation, are not available for large systems. In fact, it could easily be argued that the “electronic structure problem” has been the most intensively investigated issue in 20th century condensed matter physics and theoretical chemistry as recognized by the Nobel Prize awarded jointly to John Pople and Walter Kohn (Kohn, 1999; Pople, 1999). Their efforts and those of others have led to the development of various approximate methods that retain elements of the essential physics at reduced computational cost.

In order to make progress, intermolecular forces are usually decomposed into a physically motivated hierarchy (Buckingham, 1967; Gray and Gubbins, 1984; Stone, 2013). This begins with the choice of reference system components, atoms, ions, and molecules, assuming here nonchemically reactive species. Forces arising from static charge distributions of the isolated components from monopoles to dipoles (giving rise to Keesom interactions) and higher-order multipoles. These forces are substantially modified by various forms of many-body electronic redistribution, i.e., from induced dipoles (Debye forces), a classical effect as well as induced higher-order multipoles. Furthermore, correlated quantum fluctuations of the electronic structure at the two-atomic center limit yield at second order the London dispersion interaction (induced dipole–induced dipole), which can be extended to many-body levels and higher multipolar interactions. On the other hand, short-range interactions arising mainly from exchange repulsion can typically be treated accurately at the pair level (Stone, 2013) except at high pressures or in complex ionic solids or salts (Madden and Wilson, 1996).

Although weaker than intramolecular chemical bonds which we consider as fixed in this review, the wide spectrum of long-range intermolecular forces is responsible for the stability of liquid phases, self-assembly, supramolecular

complexes, physisorption, as well as many other phenomena in biological and soft condensed matter systems. Often, simulation strategies involve *a priori* decisions as to which of these interactions are to be included or neglected in a particular instance. This step, which truncates the interaction hierarchy at certain order, compromises the predictive power of the conceptual advances given by the connection of quantum or classical statistical mechanics and electronic structure. Moving beyond this point requires understanding the nature of intermolecular forces and the physical principles underlying them, thus facilitating the development of simple yet rich models that can be solved nonperturbatively and subsequently employed to quantitatively interrogate complex systems without an *a priori* bias due to choice of truncation scheme. This sets high targets for new and improved models with fewer assumptions and expanded ranges of validity motivated by faster computers coupled to more efficient methods that permit facile sample phase space exploration. Here we review recent advances in intermolecular forces for nonreactive systems that do not involve charge transfer on the ground state Born-Oppenheimer surface.

B. Challenges in modeling complex soft condensed matter at the molecular scale

1. Fundamental compromises

Despite steady progress in computational power and phase space exploration methodology, in atomistic scale simulation inevitable trade-offs exist between the completeness of the description of interactions and the accessible time and length scales. Different requirements have driven the evolution of several distinct strategies which handle intermolecular forces at various levels of detail as illustrated in Fig. 1.

The lowest order description of molecular interactions at long range is the Coulomb interaction between molecules with noninteracting molecular species forming the reference system (Gray and Gubbins, 1984). The standard approach to model an isolated molecular charge distribution is to assign “point charges” to sites in the molecular frame as follows: Consider two atoms with different electronegativity (Pauling, 1960) forming a covalent bond. The resulting valence electron

distribution will be asymmetric and in localized areas surrounding the nuclei there will be a shift in electron density resulting in the assignment of *partial charges*: for two atoms a partial negative and a partial positive charge. Since molecules can have moments higher than a dipole moment, more partial charges can be assigned to various sites. More advanced methods to capture the charge density include distributing multipole moments in the frame (Millot and Stone, 1992; Stone, 2013) and utilizing bond charges (Kaminski *et al.*, 2002). We note that the fitting procedure can become ill posed for large molecules due to uniqueness problems in the mapping which can be circumvented by fitting fragments and allowing more complex interactions, polarization, etc., to redistribute the charge in the larger entity. This approach follows from the discussion of condensed systems consisting of small molecules given next.

In condensed phase systems containing interfaces and involving strong field gradients, the electrons surrounding molecular moieties redistribute (polarize) in response to the local environment affecting the electric fields of neighboring molecules (Gray, Gubbins, and Joslin, 2011). All materials are polarizable to some extent: A particularly well-known example is the strong enhancement of the dipole moment of the water molecule on moving from gas (1.85 D) to liquid phases (≈ 2.7 D), polarization in the dipole limit, which arises in the classical limit at second order in perturbation theory. Thus, the first order description utilizing the isolated molecular limit for the charge distribution can fail badly here. In general, and as we will see later, polarization modifies all electrostatic multipole moments with the responses determined by a series of well-defined, higher-order polarizabilities. A mean-field approach can be adopted where electrostatic moments of the constituent moieties are adjusted for the condensed phase; however, this limits transferability and fails at interfaces (e.g., between two media) where the “mean,” as characterized by the bulk environment, is not a good approximation. Furthermore, polarization induced between two molecules is affected by the presence of others. Thus, polarization is fundamentally a many-body phenomenon which cannot be fully described within the pairwise approach (Martyna and Berne, 1988, 1989; Stone, 2013) although mean-field approximations can be made as discussed herein.

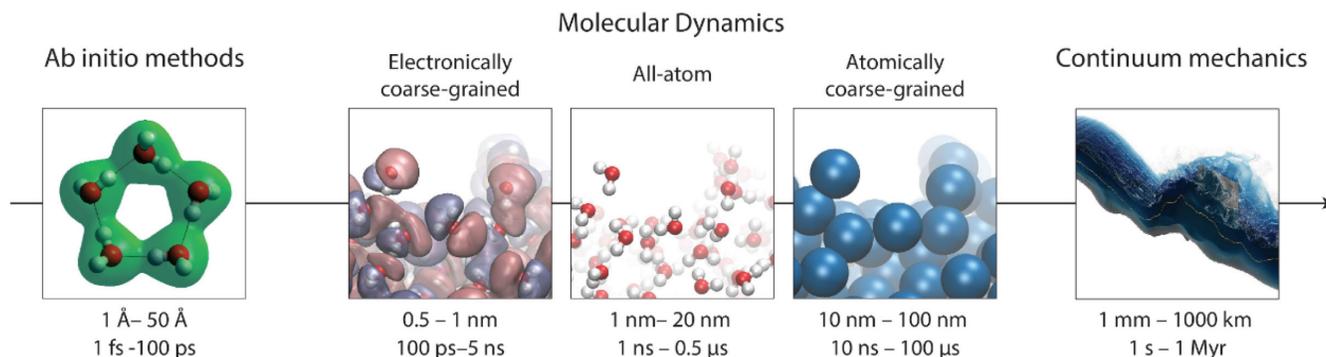


FIG. 1. Simulation methodologies arranged according to the time scales and length scales which they sample. This image focuses on electronic coarse graining occupying the range between so-called *ab initio* and classical molecular dynamics methods. From Cipcigan *et al.*, 2016.

Dispersion interactions have no classical origin. They are quantum mechanical phenomena arising from correlations between spontaneous electron density fluctuations. These include pairwise induced dipole–induced dipole interactions, which give rise to the familiar pairwise $1/R^6$ interaction of London (1937) again arising at second order in perturbation theory. However, higher-order fluctuations generate a full dispersion series (induced dipole–induced quadrupole, etc.) at the pair level. Since dispersion interactions arise from long-range correlations, they too are inherently many-body and nonpairwise as recognized in the dipole limit independently by Axilrod and Teller (1943) and Muto (1943), the well-known Axilrod-Teller-Muto (ATM) three-body dispersion term.

2. The need for new strategies

The inherent compromise between completeness of the physical description and accessible time and length scales means that the study of large, complex systems requires using highly simplified models which may not capture the fundamental interactions responsible for emergent behavior. Often this means taking mean-field classical models to treat electronic responses even for systems possessing highly polarizable entities and nontrivial interfaces.

In more detail, atomistic models of biological and materials systems are typically described via a force law with fixed functional form. Many approaches within this class of model include only pairwise electrostatic interactions between partial charges and pairwise van der Waals forces with parameters fit to data at some particular set of reference conditions, treating electronic responses within mean field (Oostenbrink *et al.*, 2004; Wang *et al.*, 2004; Vanommeslaeghe *et al.*, 2010). That is, all explicit polarization responses and many-body dispersion effects are neglected as are all contributions from higher-order terms in the dispersion series not captured well by mean-field theory. As alluded to above, aspects of these neglected interactions are represented in an average, mean-field sense by treating the charges and molecular geometry as adjustable parameters manipulated to match a set of reference properties that may or may not match the actual molecular state in the media. However, this procedure may produce significant distortion of molecular properties from their gas phase values if, for example, parameters are tuned to fit certain features of the liquid state. As an illustration, the molecular dipole moment of the TIP4P/2005 model of water (Abascal and Vega, 2005) is 2.305 D across the entire phase diagram leading to poor predictions of dielectric properties away from the fitting regime.

A key question, common to all empirically derived potentials, concerns the notion of *transferability* which we have presaged in the previous discussion: The extent to which the physics incorporated in a model applies outside the training conditions of parametrization or even depicts the true origin of system properties within the fitting window. Current approaches, based on multiparameter fitting, may obscure the physical origins of transferability or of its failure modes, in particular, circumstances.

In order to address this key question, upon which this review centers, we will explore in more detail the physical origins behind transferability and its breakdown in simplified treatments. We will develop the rationale for efficient, predictive models at the

molecular scale with complete electronic responses, reduced reliance on empirical input from the condensed phase, and the prospects for vastly improved transferability and predictive value, including the ability to observe emergent phenomena and elucidate the essential physics underlying it.

We review here a new class of molecular model for materials simulation containing a more complete description of the physical interactions than has been heretofore possible for large systems: The collective electronic responses are described by a set of charged oscillators embedded in the molecular frame, extending and adapting the work of Drude and Lorentz in the classical limit and London's in the quantum mechanical regime for noble gases. That is, a coarse-grained model within Gaussian statistics in which the electronlike particles are bound harmonically to sites within molecular frames. Mathematically, the on-site terms of the many-body Hamiltonian are simplified, preserving the full off-diagonal couplings, such that the model can be treated nonperturbatively. It has been shown that the model exhibits the full hierarchy of electronic responses at long range assuming the reference system of isolated molecules (moieties) does not undergo long-distance charge transfer (Jones, Crain, Sokhan *et al.*, 2013; Cipcigan *et al.*, 2016). Polarization to all orders can be induced by external fields (or fields arising from interfaces in the media itself) and zero-point fluctuations lead naturally to the full dispersion series, including many-body effects. Of course, the description will fail to describe some interactions or will not be able to achieve *ab initio* accuracy as must be the case for a model within Gaussian statistics, where sum rules which are only approximately obeyed by the real system of interest arise. However, we shall see that when solved nonperturbatively in strong coupling, including the whole hierarchy of responses, it can capture condensed phase behavior remarkably well (Sokhan *et al.*, 2015b), delivering a key message of theoretical physics over the last six decades. That is, a rich albeit simplified model solved in strong coupling (i.e., terms to all orders) is significantly more predictive and powerful than a perturbative approach wherein key interactions are neglected by choice of the modeler (i.e., through the order and/or symmetry of the selected truncation scheme). In the context of Fig. 1 this approach occupies a previously vacant space between the *ab initio* methods commonly used in materials, molecular and biomolecular simulation, and force-field driven all-atom molecular dynamics.

II. LONG-RANGE ELECTRONIC RESPONSES OF ATOMS AND MOLECULES: ELECTROSTATIC, INDUCTIVE, AND DISPERSIVE EFFECTS

A. Many-body intermolecular forces and the emergence of natural length scales

1. Adiabaticity: Decoupling of nuclear and electronic motion

The natural interaction-length scales which arise in atomic and molecular systems are rich and multimodal. An obvious fundamental scale is set by atomic size—the distance between nuclei and some measure of the diffuse boundary of the valence electron shell. A closely related concept is the van der Waals radius, defined from the closest approach distance between two atomic species which make contact but are not

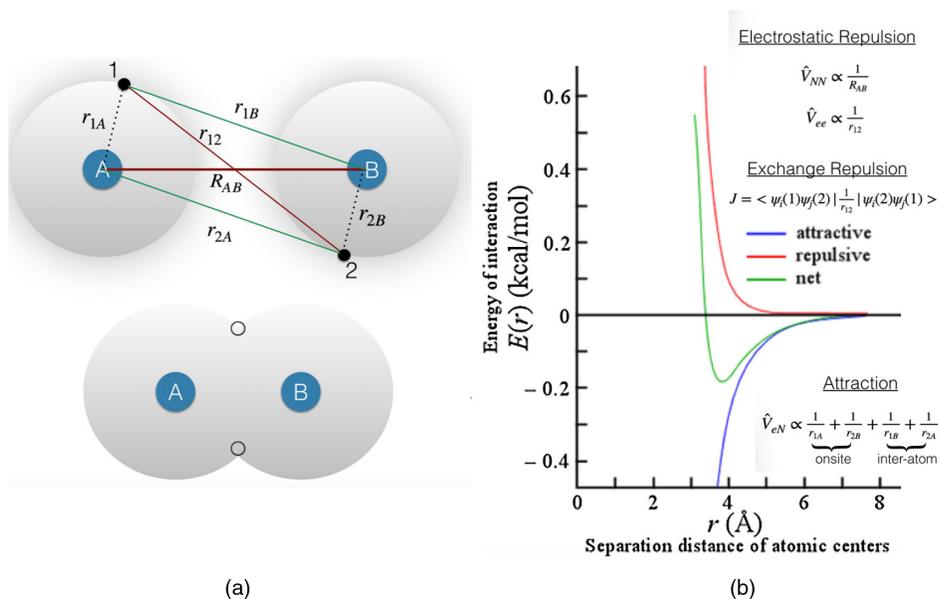


FIG. 2. Illustration of short- and long-range regimes determined by electron wave function overlap (a). Top left panel: The internuclear separation is assumed large compared to atomic dimensions resulting in negligible electron density overlap. The electrons (1 and 2) are therefore distinguishable particles. The relative electron-nuclear charges, positions, and polarizations result in electrostatic, inductive, and dispersive interactions. Short range is defined as the regime of electron overlap. Here electrons are indistinguishable and are described by antisymmetric wave functions and the emergence of exchange repulsion (in addition to classical electron-electron and internuclear repulsion). (b) A generic intermolecular potential separated into aggregate attractive and repulsive components.

chemically bound, typically wherein exchange repulsion effects become large.

Figure 2 illustrates a highly schematic representation of interactions between charge distributions. The “nuclei” (A and B) may refer to atoms or molecules with the position of two arbitrary electrons denoted as 1 and 2.

Of course, molecules evolve as a result of translation, rotation, and internal vibrational degrees of freedom driven either classically at higher temperature or quantum mechanically at lower temperature. However, the large mass disparity between nuclei and electrons means that the nuclear motion is much slower than electronic time scales in many circumstances of great interest in condensed phase molecular physics. Therefore, the electronic system sees a slowly varying nuclear potential which it can follow while remaining in its many-body ground state for any instantaneous nuclear configuration. It is analogous to the situation of a particle in a box for which the wave function remains in the ground state if the box dimensions are expanded gradually. This is the fundamental concept of adiabatic separation. In this context it means that there is no energy exchange between the nuclear and electronic subsystems thereby making contact to the thermodynamic notion of an *adiabatic process*. In dynamical systems which can be separated into slow and fast time scales the BO approximation allows for determination of energy eigenvalues where the slow degrees of freedom are deemed fixed. These considerations permit decoupling of the electronic and nuclear motion: We can then define a ground state electronic energy for each fixed nuclear configuration to map out the ground state BO potential energy surface. This is a statement of the physics underlying a many-body wave function of the form

$$\begin{aligned}
 \Psi(\mathbf{r}, \mathbf{R}) &= \sum_n \sum_k c_{nk} \psi_n(\mathbf{r}; \mathbf{R}) \chi_{nk}(\mathbf{R}), \\
 \hat{H}_{\text{el}}(\mathbf{r}; \mathbf{R}) \psi_n(\mathbf{r}; \mathbf{R}) &= \epsilon_n(\mathbf{R}) \psi_n(\mathbf{r}; \mathbf{R}), \\
 [\hat{H}_{NN}(\mathbf{R}) + \epsilon_n(\mathbf{R})] \chi_{nk}(\mathbf{R}) &= E_{nk} \chi_{nk}(\mathbf{R}), \\
 \hat{H}_{\text{el}} &= -\frac{\hbar^2}{2m_e} \nabla_{\mathbf{r}}^2 + V_{eN}(\mathbf{r}, \mathbf{R}) + V_{ee}(\mathbf{r}), \\
 \hat{H}_{NN} &= -\frac{\hbar^2}{2m_N} \nabla_{\mathbf{R}}^2 + V_{NN}(\mathbf{R}), \tag{1}
 \end{aligned}$$

where \mathbf{r} and \mathbf{R} denote electronic and nuclear degrees of freedom, correspondingly, and electronic and nuclear Hamiltonians are defined in the last two lines. For simplicity, the nuclear masses are represented by a single scalar. In the adiabatic limit, the $\psi_n(\mathbf{r}; \mathbf{R}) \chi_{nk}(\mathbf{R})$ become pure eigenstates with eigenvalue E_{nk} . Strictly, there is a family of such potential energy surfaces corresponding to ground state and excited electronic levels as shown in Fig. 3. However, here we concentrate on conditions wherein only the ground state surface is populated (zero electronic temperature).

If the nuclei can be treated classically and only the electronic ground state is occupied with nonzero probability, the force law is derived from the potential formed by $V_{NN}(\mathbf{R}) + \epsilon_0(\mathbf{R})$, the ground state BO surface, and the principles of classical statistical mechanics can be applied to study such systems. Similarly, in the limit the nuclei are light, quantum statistical mechanics can be applied on a single electronic surface, the $\chi_{0k}(\mathbf{R})$ given occupancy derived from the statistical mechanical ensemble of interest. When the BO approximation breaks down, multiple electronic surfaces couple and a full quantum treatment or approximate

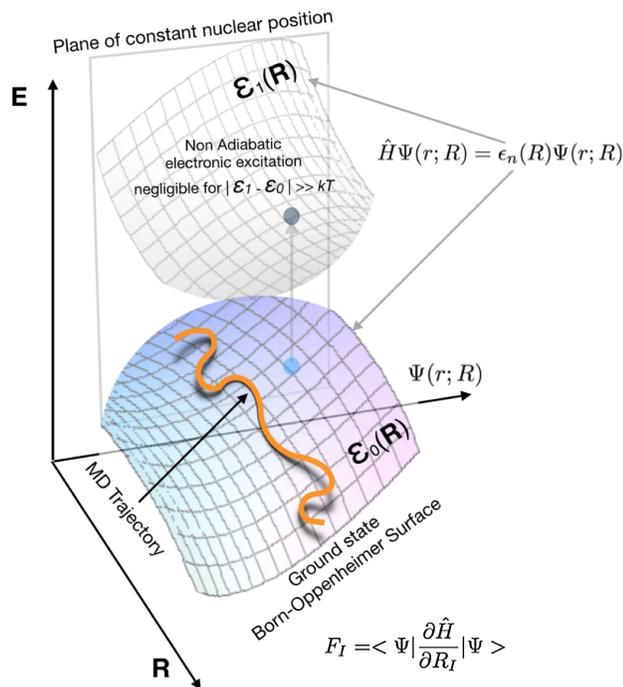


FIG. 3. Accurate forces can be computed from the Hellmann-Feynman theorem when the trajectory remains close to the ground state Born-Oppenheimer surface $\epsilon_0(\mathbf{R})$. Nonadiabatic excitations to higher electronic states are negligible (as assumed in the BO approximation) as long as $|\epsilon_1(\mathbf{R}) - \epsilon_0(\mathbf{R})| \gg kT$ is satisfied.

semiclassical methods can be brought to bear as reviewed elsewhere (Tully, 2012) and not discussed further herein.

2. Many-body effects and the Born-Oppenheimer surface

Since our objective is to reach long length and time scales in simulation studies, the BO surface must be described via a computationally efficient approximation which yet retains the essential physics of the intermolecular interactions. With this aim, various strategies at different levels of physical realism are presently employed in large scale atomistic simulation. However, the basic ingredients are uniformly similar: Nuclei repel each other; electrons are attracted to nuclei and are repelled by other electrons. This has been essentially captured by seminal early work on the hydrogen molecule and the hydrogen molecule ion—arguably the first applications of quantum mechanics to a molecular system (Heitler and London, 1927; James and Coolidge, 1933).

In principle, the ground state BO electronic wave function $\psi_0(\mathbf{r}; \mathbf{R})$ encodes all the information required to create forces derived from effects such as electrostatic multipole moments, polarizabilities, dispersion, and short-range repulsion at fixed nuclear position to *all orders*. The physical decomposition and hierarchy with which we are familiar arises from perturbation theory as emphasized above. According to the Hellmann-Feynman theorem, in the adiabatic limit, $\mathbf{F} = -\langle \psi_0(\mathbf{r}; \mathbf{R}) | \nabla_{\mathbf{R}} \hat{H}_{el} | \psi_0(\mathbf{r}; \mathbf{R}) \rangle - \nabla_{\mathbf{R}} V_{NN}(\mathbf{R})$.

As previously stressed, it is difficult to create a highly accurate BO surface from first principles and typically a reference of noninteracting molecules is adopted. Within this

approach, as we explore in the next sections, some contributions to the interaction energy emerge from permanent electrostatic interactions (pairwise additive) between the fixed reference states whereas others arise only because the electronic structure is responsive (polarizable) exhibiting many-body correlations (which are not pairwise additive). This motivates the idea of separating intermolecular interactions between N molecules into a sum having the general structure,

$$V(\mathbf{R}) = \sum_{i < j}^N V_2(\mathbf{R}_i, \mathbf{R}_j) + \sum_{i < j < k}^N V_3(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) + (\text{higher orders}). \quad (2)$$

The above decomposition arises naturally from perturbation theory around the reference system, as in the *virial expansion* of the equation of state for a homogeneous fluid (Hansen and McDonald, 2013) which measures (as powers of the density) progressive deviations in the pressure p of a gas from ideal (noninteracting) case at temperature T according to $\beta p = \rho + B_2(T)\rho^2 + B_3(T)\rho^3 + \dots$, where $\beta = 1/k_B T$. Historically, the virial expansion provided direct evidence for the impact of the three-body interactions which challenge atomic and molecular simulation even now.

The virial expansion depends on the range and many-body character of interactions where the virial coefficients B_n are each related to the terms (V_n) which appear in Eq. (2) corresponding to pair ($V_{n=2}$), triplet ($V_{n=3}$), and higher n -tuple contributions. In practice, it is far from straightforward to extract specific terms in the many-body expansion: Experiments measure all many-body terms and model pairwise potentials are normally fit to condensed phase data and therefore incorporate, to some extent, mean-field effects of the many-body terms at some particular thermodynamic state point (Sadus, 2002). In general, we note that only Coulomb interactions can be described without approximation by potentials of the V_2 type; inductive and dispersive effects require approaches beyond the pairwise-additive approximation.

3. Separation of interaction length scales

If the A and B atoms or molecules of Fig. 2 are charged, the familiar Coulomb interaction will dominate—the electrostatic interaction with the longest range. If there are no net charges but permanent asymmetric charge distributions occur we will have dipole (or motionally averaged Keesom forces) or higher-order multipole interactions. The charge or polarity of one molecule will also influence another, leading to induced interactions. Finally, since quantum mechanical charge fluctuations are present in all atoms and molecules they lead to attractive dispersion (London forces and higher-order effects attenuated by r^{-n} , $n = 6, 8, 10, \dots$ for spherical centers). Some examples of these forces are presented in Fig. 4.

When two atoms or molecular species approach sufficiently closely, a new phenomena emerges—exchange repulsion. For example, within a single particle picture of closed shell systems, the Pauli exclusion principle prohibits further

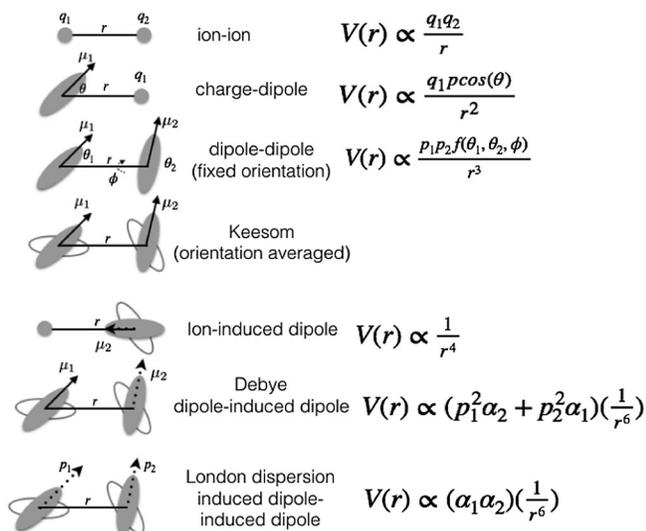


FIG. 4. Various forms of intermolecular interactions (electrostatic, inductive, dispersive) and their distance dependences. Ion-induced dipole and Debye forces are simple examples of inductive interactions where a permanent dipole induces a dipole moment in a neighboring molecule. London dispersion forces are the simplest example of the “dispersion series” which will be discussed more fully in Sec. II.C.

occupancy of the lowest occupied reference states upon close approach. To produce overlap some electrons must be promoted to unoccupied states (of the reference state); therefore electron overlap produces overall a steep increase in energy at very short range. More formally, in the overlap region, where the concept of polarizing the charge distribution within a tight binding picture breaks down, the electrons extend over both molecules. As the wave function must be antisymmetric under particle exchange, the effect is extra repulsive energy (over what would occur for bosons) called exchange repulsion.

For a two-electron system in which α and β denote distinct sets of atomic quantum numbers including spin (n, l, m_l, m_s) the antisymmetric linear combination of single particle products takes the form

$$\Psi(1, 2) = \frac{1}{2}[\psi_\alpha(1)\psi_\beta(2) - \psi_\alpha(2)\psi_\beta(1)]. \quad (3)$$

This has the property that $\Psi(1, 2) = -\Psi(2, 1)$ and can be generalized to N -electron systems with the use of Slater determinants (Stone, 2013). It is evident that in the case where the two quantum states are identical ($\alpha = \beta$) the wave function vanishes. This is the essence of the Pauli exclusion principle which precludes multiple occupation of quantum states and leads to an additional effective exchange repulsion which arises from the antisymmetry requirement of the two-particle wave function in addition to the electronic charge. The resulting interaction energies are described by two terms: First, the Coulomb integral $\langle \psi_i(1)\psi_j(2) | r_{ij}^{-1} | \psi_i(1)\psi_j(2) \rangle$ which represents the classical electrostatic repulsion between two interacting charge densities. In addition, the antisymmetry requirement for the wave function (or Slater determinant) produces an additional contribution—the exchange integral

$\langle \psi_i(1)\psi_j(2) | r_{ij}^{-1} | \psi_j(1)\psi_i(2) \rangle$, which has no classical counterpart.

These considerations therefore lead to a natural separation of interaction range into two basic parts: The short-range regime is defined as the region of close approach between two interaction sites where electron density overlap and exchange repulsion is not negligible. Roughly this occurs when the contact distance between two atoms reaches the sum of their van der Waals radii. At sufficiently small separations the exchange repulsion overwhelms the asymptotic behavior of the attractive terms to produce a steep repulsive potential.

At longer range, where overlap can be ignored, the electrons are distinguishable and can be assigned as belonging to one or other sets of nuclei (A or B). No antisymmetrization of the wave function is needed outside of the overlap region—the tight binding limit. The potential $V(\mathbf{R})$, arising from one of the charge distributions $\rho(\mathbf{r})$ determined at some instant, is given by

$$V(\mathbf{R}) = \int d\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{R} - \mathbf{r}|} = \int d\mathbf{r} \rho(\mathbf{r}) T(\mathbf{R} - \mathbf{r}), \quad (4)$$

which is recognized as the general solution to Poisson’s equation for an arbitrary charge distribution and is valid for all observation points \mathbf{R} outside the bodies. In the next section we explore how the potential can be decomposed into individual contributions with specific interaction ranges. Equation (4) gives rise to Coulomb interactions that are in some sense classical although $\rho(r)$ depends, implicitly, on \hbar .

B. Multipole electrostatics and induction in molecular systems

We establish here the formalism for the hierarchy of electronic responses of a molecule as a basis for future discussion and as a reference point for understanding the nature of approximations made in molecular-level simulations of materials. We consider first the description of an assembly of fixed charges in terms of electrostatic multipoles as a representation of an arbitrary molecule in which polarization is neglected. Next, we immerse this distribution in a static, homogeneous electric field F . Of course, neither the lack of polarization response nor field homogeneity is representative of real molecules or the variety of environments they encounter in condensed or vapor phases but it permits systematic development of the key ideas. Both assumptions will later be relaxed.

1. Electrostatic multipoles

Far from the charge distribution, where $|\mathbf{R}| > |\mathbf{r}|$ the potential introduced in Eq. (4) may be expanded in R^{-n} powers to decompose the contributions to the potential into a series of simpler, fundamental configurations—*multipoles* of ascending order (Stone and Alderton, 1985); see Appendix A for a formal discussion of the multipole expansion and development of the key concepts. The basic idea is that a molecule is considered an “ n -pole” object if it has an n -pole electrostatic moment but no lower moments. So a charged molecule is a monopole corresponding to $n = 0$ (zeroth moment). The charge is formally defined from a continuous distribution as $q = \int \rho(\mathbf{r}') d\mathbf{r}'^3$.

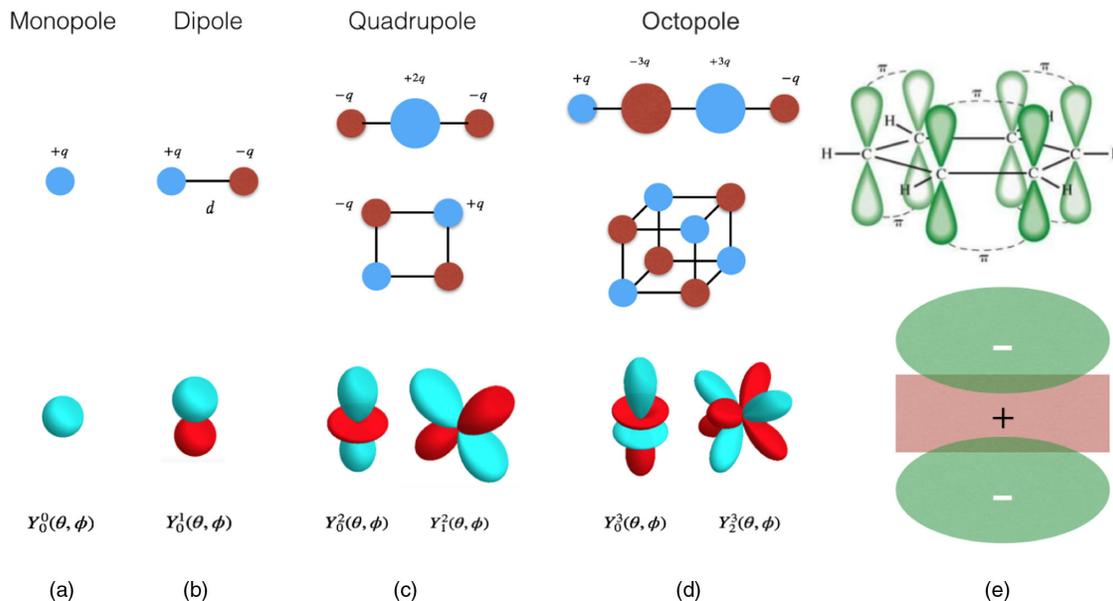


FIG. 5. Examples of elementary source configurations for the (a) electric monopole with far field $F \propto 1/R^2$, (b) dipole ($F \propto 1/R^3$), (c) quadrupole $F \propto 1/R^4$, and (d) octopole ($F \propto 1/R^5$). Note that for (c) and (d) the use of unequal charges can produce additional source configurations. In the lower panel are shown equivalent representations of multipole moments using spherical harmonics $Y_l^m(\theta, \phi)$ to which an introduction is given in Appendix A. (e) How a complex charge distribution arising from pi stacking in benzene exhibits quadrupolar symmetry.

A neutral molecule may however still exhibit various forms of charge separation or *polarity* which produces characteristic potential and field lines sufficiently far from the source. The simplest form of charge separation is measured by the dipole moment ($n = 1$). This can occur between two ions in an ionic bond or between atoms in a covalent bond where there is a difference in electronegativity. In its elementary form, an electric dipole consists of two equal and opposite electric monopoles ($\pm q$) (so that the total charge is equal to zero), separated by a small distance. In atomic units, the dipole is expressed in terms of ea_0 where e is the elementary charge and a_0 is the Bohr radius but debye (D) units are also used where $D = 10^{-18}$ esu. For a continuous charge distribution the general definition of the dipole is naturally extended as $\vec{\mu} = \int \rho(\mathbf{r}') \mathbf{r}' d\mathbf{r}'^3$ (Davis, 2011).

An electric quadrupole (Θ , $n = 2$) can be built from two electric dipoles, arranged so that the total charge and the total dipole moment cancel (Buckingham, 1959). The atomic unit of electric quadrupole moment is ea_0^2 . The so-called *primitive* form is also a direct extension from the dipole such that $q_{ij} = \int \rho(\mathbf{r}') r_i r_j d\mathbf{r}'^3$. The resulting quadrupole tensor (Hess, 2015) is clearly symmetric and often expressed according to its traceless definition here in the Cartesian frame: $\Theta_{\alpha\beta} \propto \int (3x_\alpha x_\beta - r^2) \rho(\mathbf{r})$ which measures the asphericity of the charge distribution and further reduces the number of independent components (see Appendix A for further discussion and the derivation of the traceless form).

Octopole moment tensors $\Omega_{\alpha\beta\gamma}$ and higher multipole source configurations (Stone, 2013) can be similarly constructed from self-canceling lower moments as shown in Fig. 5. The potential energy and field strength are attenuated more sharply with distance for higher multipoles (as discussed in the caption of Fig. 5) because the higher number of charges

required to build the multipole appear “more neutral” at shorter distances. At very large distances from the distribution the leading multipole will therefore dominate but at intermediate distances higher multipoles are not negligible. The field lines and equipotential surfaces for examples of dipole, quadrupole, and octopolar source configurations are shown in Figs. 6(a)–6(d). The questions of origin choice, translational invariance of the lowest multipole moment, and nonuniqueness of higher moments are discussed briefly in Appendix A.

2. Symmetry considerations

In the most general case, the n th multipole of a molecule with the lowest possible (C_1) point group symmetry can have up to $2n + 1$ independent components. Such a molecule can therefore have all three dipole moment components (μ_α , $\alpha \equiv \{x, y, z\}$) as distinct and independent.

The rank-two quadrupole tensor has nine components in total but even the lowest symmetry molecules can have at most five independent quadrupole tensor elements ($\theta_{x^2, y^2, xy, xz, yz}$) since the quadrupole is both symmetric and traceless [see Stone (2013) for a discussion of irreducible tensor notation].

Molecular symmetry further reduces the number of independent moments and determines which multipoles exist (Jeevanjee, 2015). For example, in molecules having a center of inversion, the dipole (and all higher odd multipoles) are zero (Davies, 2007). This includes, for example, the wide class of AB_n -type molecules, where A is the central atom and B are all equivalent. Here common molecular geometries such as linear, trigonal planar, tetrahedral, octahedral, and trigonal bipyramid have zero dipole moment even though the bonds are polar. Unlike the dipole, the quadrupole is unaffected by inversion ($r \rightarrow -r$) and therefore has even parity. Various so-called pi-stacked organic molecules have zero dipole with the quadrupole as the lowest moment; see Fig. 5(e).

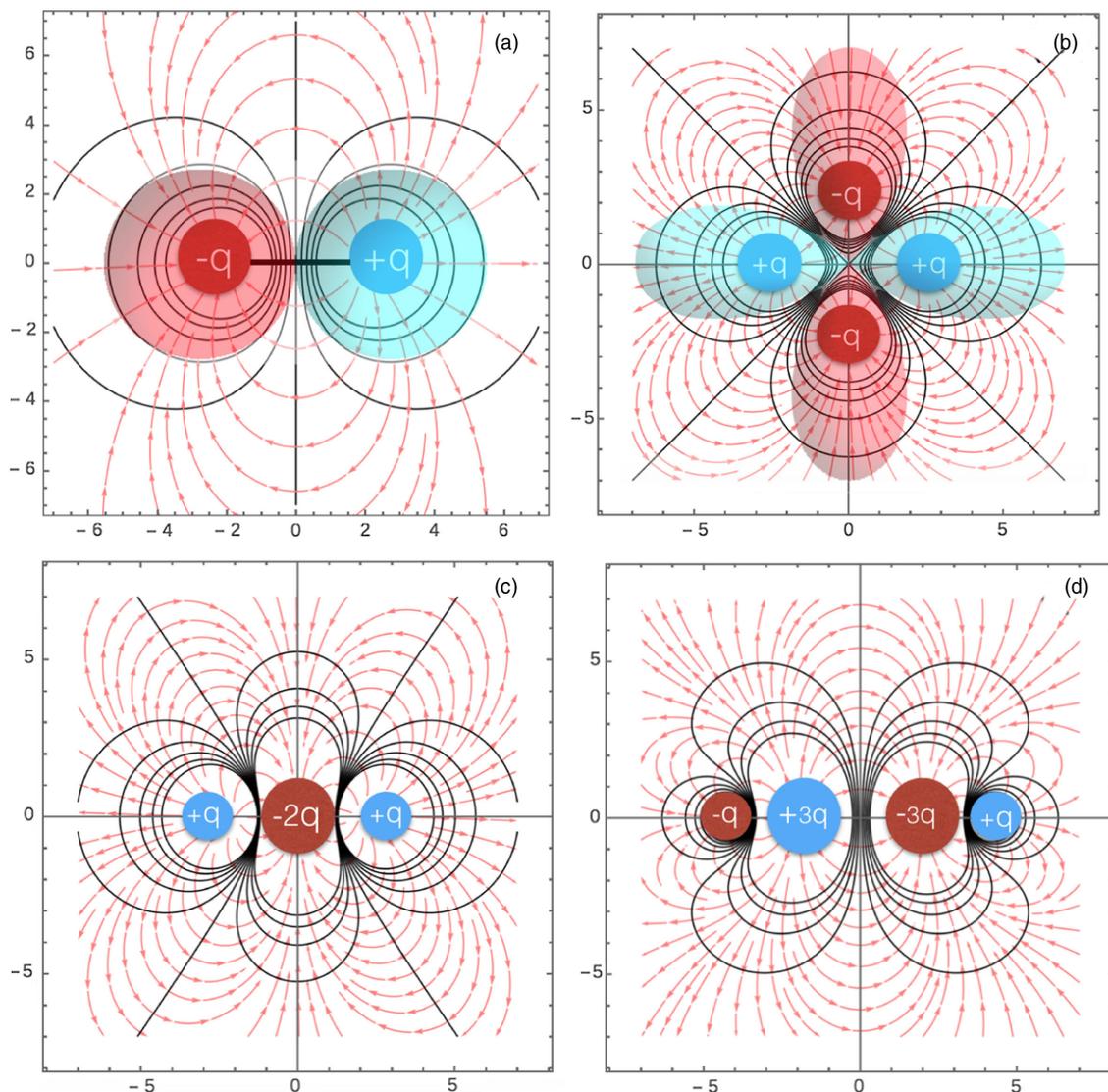


FIG. 6. Equipotential ($\phi = \text{const}$ solid lines) and field lines ($E = -\nabla\phi$) corresponding to (a) elementary dipole, (b) spatial quadrupole, (c) linear quadrupole, and (d) linear octopole source configurations. Note that, despite the differences in the source configurations in (b) and (c), the resulting far-field lines are similar. The spherical multipole representations corresponding to $Y_0^1(\theta, \phi)$ and $Y_1^2(\theta, \phi)$ are superposed in (a) and (b), respectively.

Higher symmetry tetrahedral molecules such as methane, white phosphorous, neopentane, adamantane, C_{28} fullerene, and many others with overall T_d symmetry have the octopole ($\Omega_{\alpha\beta\gamma}$) as the lowest nonzero moment. Molecules with octahedral (O_h) symmetry (e.g., sulfur hexafluoride and molybdenum hexacarbonyl) have the hexadecapole as the lowest moment (see Appendix A for a representation of the linear and spatially extended hexadecapole).

Linear versions (point group $C_{\infty,v}$) of all multipoles can be constructed. In this special case, there are many restrictions imposed by the axial (z -axis) symmetry which requires the lateral components of the dipole moment to be zero leaving only $\mu_z \neq 0$ if there is no inversion center. Also, all off-diagonal elements of the quadrupole must be zero. Axial symmetry further enforces $Q_{xx} = Q_{yy}$ leaving only Q_{zz} as the independent moment. For this symmetric case, in general, the distribution can be specified

with only one multipole component for all orders (Griffiths, 2017).

This axial form of the multipole (Zangwill, 2013) expansion can also naturally be expressed in terms of a single polar angle. Here symmetry requires that the dipole be described by odd functions of the polar angle to permit the sign change on reflection through the origin whereas the quadrupole is described by an even function as it exhibits reflection symmetry. More generally the idea can be extended to construct the multipole expansion using spherical harmonics of appropriate symmetry rather than the Cartesian form. Examples (including the axial cases) are shown in the lower panel of Figs. 5(a)–5(d) and also briefly discussed in Appendix A.

In the case of water, for example, we can determine the symmetry restrictions on the multipole moments from the characters of its C_{2v} molecular point group, which is shown in Fig. 7.

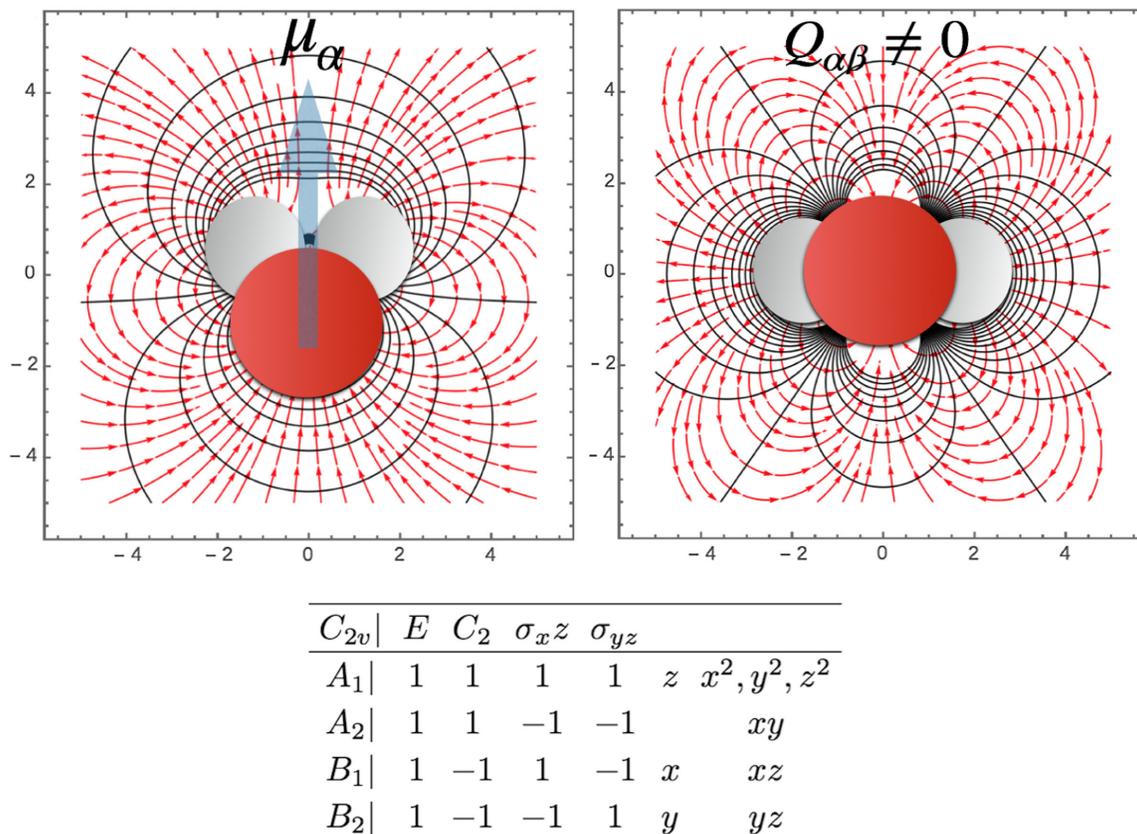


FIG. 7. (Top left) Water molecule illustrating the presence of a dipole moment along the H-O-H bisector, and a (top right) quadrupolar field as viewed along the dipolar axis. Both the dipole and quadrupole moments are permitted for molecules with C_{2v} point group symmetry for which the character table is also shown (bottom).

In this case the z component of the dipole moment transforms according to the totally symmetric (A_1) irreducible representation with the lateral components being zero. Since water is electrically neutral the dipole is the lowest nonzero (unique) multipole moment. This is illustrated in Fig. 7(a) showing the μ_z dipole moment and field lines; cf. Fig. 5(b).

The C_{2v} character table also shows that quadrupole tensor elements [bilinear forms $r_\alpha r_\beta = (x^2, y^2 \text{ and } z^2)$] transform according to the totally symmetric A_1 representation with all others being zero. Therefore the water molecule also exhibits a nonzero quadrupole moment; see Fig. 7 (top right panel) and Figs. 5(c) and 5(d) for comparison. Finally, since the quadrupole tensor $Q_{\alpha\beta}$ is traceless, there are only two independent elements ($Q_{zz}, Q_{xx} - Q_{yy}$) rather than three.

3. Multipoles in nonuniform fields

We next consider electrostatic interactions between two charge distributions (A and B) at long range; see, for example, Fig. 2. Here we express the potential energy of one system in the external field due to the other. If A and B are separated by a large distance then the potential due to the charges at B can be assumed to be slowly varying over the distribution at A . In that case, the potential can be expanded in a Taylor series and the interaction between each multipole with the field and its gradients can be separated. See Appendix A for a more complete discussion.

If we first consider the case of a neutral, polar molecule, the electrostatic energy E_{elec} in a *uniform* external electric field \vec{F} is simply given by $E_{\text{elec}} = -\vec{\mu} \cdot \vec{F} = -\mu_\alpha F_\alpha$ in tensor notation. This arises as the first term in the multipole expansion of the interaction potential and is the familiar orientational energy of a dipole. Here F denotes the field and α refers to the Cartesian components. A dipole therefore experiences a torque in a uniform field giving an orientational potential energy.

A pure quadrupole experiences no torque (no orientational energy) in a uniform field because the effects of the constituent dipole moments cancel. In fact, all higher multipoles enter the energy expression only through field gradients producing systematic additions to the orientational energy E_{elec} beyond the dipole term,

$$E_{\text{elec}} = -\mu_\alpha F_\alpha - \frac{1}{3}\Theta_{\alpha\beta} F_{\alpha\beta} - \frac{1}{15}\Omega_{\alpha\beta\gamma} F_{\alpha\beta\gamma} + \dots \quad (5)$$

Here the notation $F_{\alpha\beta} = \partial F_\alpha / \partial \beta$ is the field gradient. Additional indices imply higher field derivatives. The quadrupole moment then determines the energy of a quadrupolar charge arrangement in a smoothly nonuniform external field described by a constant field gradient. The terms of Eq. (5) (coupling of multipole moments to field and field gradients) are a consequence of the multipole expansion of the electrostatic potential (see Appendix A for further details). Finally Eq. (5) and the individual terms of the multipole expansion for

the potential make it possible to extract expressions for the interaction energy between electric multipoles of arbitrary order. This is discussed in Appendix A.3.

4. Induction and many-body linear responses

In “real” materials, electrons are bound to their nuclei with Coulomb forces which are not infinitely strong. Therefore atomic and molecular charge distributions can be distorted substantially by external fields. There are numerous examples where polarization effects are important such as in water, aqueous solutions, amine and amide hydration and other biomolecular contexts and interfaces. See, for example, the review by [Halgren and Damm \(2001\)](#) and references therein. Recent work showed that many-body polarization effects have a profound influence on the water contact angle ([Misra and Blankschtein, 2017](#)).

Pairwise potentials can approximate the average effects of polarization, in particular, environments by manipulation of atomic charges or fragment dipole moments as is normally done in conventional force fields ([Allen and Tildesley, 2017](#)). However, in contrast to the case of electrostatics, there is no superposition principle for polarization (induction): Induced multipole moments arise cooperatively from the action of all other molecules meaning that Eq. (2) cannot be truncated at the first term in the case of induction.

The redistribution of charge which occurs during polarization must also be accommodated within the electronic energy level structure of the atoms or molecules. Therefore we assume that it occurs by accessing orbitals which were unoccupied in the undistorted ground state.

The involvement of unoccupied states in the polarization response in atomic systems ([Mahan and Subbaswamy, 2013](#)) is illustrated in the textbook case of the second-order (quadratic) Stark shift in the hydrogen ground state ([Schwabl, 2007](#)). This gives the form of the contribution from the polarization energy (in the dipole limit, neglecting higher multipoles which we will explore later) in a weak homogeneous field \mathbf{F} as

$$E_{\text{ind}} = -\frac{1}{2}\alpha_{ij}F_iF_j, \quad (6)$$

thereby defining the dipole polarizability as

$$\alpha = 2e^2 \sum_{n>1} \frac{|\langle 1, 0, 0 | z | n, 1, 0 \rangle|^2}{E_n - E_1}, \quad (7)$$

which is interpreted as a measure of charge migration from the ground state to an excited state. This involves a sum over unoccupied states weighted by their energy relative to the ground state.

By extension, we infer that in molecular systems the electron distribution is more easily polarized if the gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital ($\Delta E^{\text{H-L}}$) is small so that the excited states are more accessible. This reasoning leads to an approximate extension to Eq. (7) for molecular systems which says that dipole polarizability increases with increasing molecular size and decreasing energy gap,

$$\alpha \approx \frac{e^2 R^2}{\Delta E^{\text{H-L}}}. \quad (8)$$

The dipole polarizability is therefore an example of a collective, coarse-grained description of the electronic responses to a weak external field.

For a polarizable molecule in a homogeneous field the components of the molecular dipole moment then acquire a field dependence according to $\mu_\alpha(\mathbf{F}) = \mu_\alpha + \alpha_{\alpha\beta}F_\beta$ where the second term is the *induced* moment and $\alpha_{\alpha\beta} = \partial\mu_\alpha/\partial F_\beta$ is the static linear, dipole-dipole polarizability tensor (of rank 2) with units of $\text{C}^2\text{m}^2\text{J}^{-1}$ or $e^2a_0^2E_h^{-1}$ in atomic units.

In the condensed phase of a real molecular system the field \mathbf{F} which appears here is the total field arising from both permanent and induced moments on all other molecules. The electronic configuration is then determined by many-body effects which require an iterative or self-consistent solution. In the dipole limit this can be expressed rigorously [following [Applequist, Carl, and Fung \(1972\)](#)] by defining the induced dipole as

$$\mu_{\text{ind}} = \alpha(\mathbf{F}_{\text{perm}} + \mathbf{F}_{\text{ind}}) = \alpha(\mathbf{F}_{\text{perm}} + \mathbf{T}\mu_{\text{ind}}). \quad (9)$$

Here the total field acting on a molecule is separated into permanent and induced fields from all surrounding polarizable molecules. The second equality arises from the use of the dipole interaction tensor [see Eq. (A11)]. The expression can then be solved for the induced moment in terms of the permanent field and a so-called “relay” tensor ([Applequist, 1985](#)),

$$\mu_{\text{ind}} = (\alpha^{-1} - \mathbf{T})^{-1}\mathbf{F}_{\text{perm}}. \quad (10)$$

We note that if the weak field assumption is relaxed, additional nonlinear ($F_\beta F_\gamma$) contributions to the induced moment would arise (from hyperpolarizabilities) in the strong field limit. However, as we are concerned with the interactions between molecules in condensed phases we will not consider effects beyond the linear response regime. So far, the discussion has been limited to dipole polarizability but higher multipoles will also respond to homogeneous fields: In the quadrupole case $\Theta_{\alpha\beta}(\mathbf{F}) = \Theta_{\alpha\beta} + A_{\alpha\beta,\gamma}F_\gamma$ where \mathbf{A} is the third-rank polarizability tensor the physical significance of which we will consider further in the next section. If we now also relax the uniform field assumption, the dipole, quadrupole, and higher moments are further modified by local field variations as determined by additional linear response tensors (multipole polarizabilities) which couple to higher field derivatives ([Maroulis, 2006](#)).

These effects lead to the complete form of the energy expression arising from the sum of electrostatic and induced contributions ($E_{\text{elec+ind}}$) relevant for a polarizable molecule in an inhomogeneous field,

$$\begin{aligned}
E_{\text{elec+ind}} = & -\mu_{\alpha}F_{\alpha} - \frac{1}{3}\Theta_{\alpha\beta}F_{\alpha\beta} - \frac{1}{15}\Omega_{\alpha\beta\gamma}F_{\alpha\beta\gamma} \\
& + (\text{higher-order electrostatic terms}) \\
& - \frac{1}{2}\alpha_{\alpha\beta}F_{\alpha}F_{\beta} - \frac{1}{3}A_{\alpha,\beta\gamma}F_{\alpha}F_{\beta\gamma} - \frac{1}{6}C_{\alpha\beta,\gamma\delta}F_{\alpha\beta}F_{\gamma\delta} \\
& + (\text{higher-order induction terms}). \quad (11)
\end{aligned}$$

Despite its complexity, Eq.(11) directly illustrates the consequences of various levels of approximation. In particular, it reduces to the familiar result that the energy of a neutral, nonpolarizable charge distribution in a homogeneous field is given simply by $-\mu_{\alpha}F_{\alpha}$ as given by the first term. All the other contributions arise from the combined effects of induction and field inhomogeneity which are fundamental features of real molecules and the environments they encounter in naturally occurring situations. Here the first three terms represent the orientational potential energy of the permanent multipoles (nonpolarizable case). We see directly that the permanent dipole contributes to the energy linearly in the field, the permanent quadrupole via the field gradient, and so on.

The second set of terms illustrates the structure of the energy contributions arising from induction. We can get some further insight into the nature of the response tensors by considering the derivatives of the energy with respect to components of the field and field gradient. In particular,

$$\partial E / \partial F_{\xi} = \mu_{\xi} - \alpha_{\xi\beta}F_{\beta} - \frac{1}{3}A_{\xi,\beta\gamma}F_{\beta\gamma}$$

and

$$\partial E / \partial F_{\xi\eta} = \frac{1}{3}\Theta_{\xi\eta} + \frac{1}{3}A_{\alpha,\xi\eta}F_{\alpha} + C_{\alpha\beta,\xi\eta}F_{\alpha\beta}.$$

From this we see that the derivative of energy with respect to field produces the permanent dipole moment which exists even in the limit of zero field as expected. The expressions then illustrate how a hierarchy of additional contributions to the dipole appear. First is the dipole polarizability (thereby defining α) which we already introduced. The next correction to the dipole moment arises from the field gradient and is determined by the response tensor \mathbf{A} and so on. By analogy, the first term in the derivative with respect to the field gradient is proportional to the permanent quadrupole moment with the first correction being linear in the field and also determined by the response tensor \mathbf{A} . Therefore, \mathbf{A} determines both the response of a dipole in a field gradient and the response of a quadrupole to a homogeneous field. It is therefore a mixed response tensor called the *dipole-quadrupole* polarizability which vanishes for molecules with inversion centers but is otherwise nonzero. The response tensor $C_{\alpha\beta,\gamma\delta}$ (defined as $\partial\Theta_{\alpha\beta}/\partial F_{\gamma\delta}$) determines the quadrupole induced by the field gradient (hence it is called the *quadrupole polarizability*). It is similarly possible to define higher multipole polarizabilities (e.g., octopole-octopole, etc.).

As the final point we note that the group theory considerations which determine whether particular multipole moments are nonzero can be extended to polarizabilities. In particular, the quantum mechanical expression for the dipole polarizability appears [according to Eq. (7)] as a square of electric dipole transition moments. Correspondingly, the tensor \mathbf{A} is defined from the product of transition dipole and

quadrupole moments. The forms and symmetry properties of these have been discussed in detail by Hirschfelder (2009) and Stone (2013).

C. Dispersion: Quantum mechanical prigns and many-body effects

Although comparatively weak, dispersion is the most ubiquitous of the intermolecular forces operating between all atoms and molecules including those that are neutral and nonpolar (for which there are neither electrostatic nor inductive effects). These interactions are inherently quantum mechanical arising from correlated fluctuations in electron density at long range (Mahanty and Ninham, 1976; Parsegian, 2005).

In the framework of Fig. 2 and Eq. (2), a neutral, nonpolar atom (such as a noble gas atom) has zero expectation value for all multipoles but nonzero variance. A schematic is shown in Fig. 8. This is analogous to the case of a quantum harmonic oscillator in its ground state which exhibits zero-point fluctuations. Here the fluctuating dipole (or higher multipole) moment of atom A creates a fluctuating field at atom B which thereby induces a transient polarization.

These interactions are directly responsible for condensed phases of nonpolar molecules and noble gases. And as a result of their generality they also influence diverse phenomena including liquid state physics, molecular crystal stability, low-dimensional systems (Cole *et al.*, 2009), hybrid organic-inorganic systems (Maurer, Ruiz, and Tkatchenko, 2015), interactions between atoms in cold optical traps, biophysical interactions such as drug-ligand binding, and other aspects of supramolecular associations.

The dispersion interaction is only approximately pairwise additive and always attractive, at second order in perturbation theory (London, 1937; Elrod and Saykally, 1994; Tkatchenko *et al.*, 2012). At third order, interactions between atom triplets arise (Axilrod, 1951) which can be attractive or repulsive depending on the interaction geometry. Further many-body effects occur for higher orders and the contributions to the interactions between atoms within molecules [$E_{\text{disp}}(R_i)$] can therefore be separated [in the spirit of Eq. (2) into two-body ($E^{(2)}$), three-body ($E^{(3)}$), and higher components (Donchev, 2006)].

In general we have

$$\begin{aligned}
E_{\text{disp}}(R_i) = & \frac{1}{2} \sum_{ij} E^{(2)}(R_i, R_j) + \frac{1}{6} \sum_{ijk} E^{(3)}(R_i, R_j, R_k) \\
& + (\text{higher orders}). \quad (12)
\end{aligned}$$

These contributions along with the effects contained in Eq. (11) represent the full hierarchy of electrostatic, inductive, and dispersive responses encountered in molecules. They also define the ingredients of the general many-body interaction decomposition introduced in Eq. (2) and a long-range representation of the terms which are captured in Eq. (2), leading to a realistic BO surface.

At long range (negligible charge overlap between interacting species) the two-body dispersion terms can be organized into the dispersion series as follows:

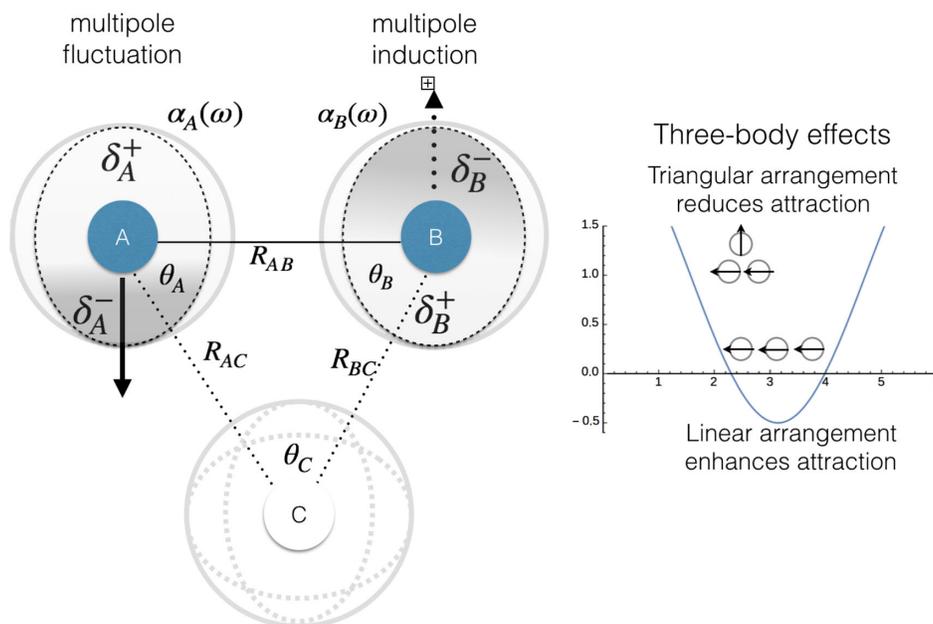


FIG. 8. Charge density fluctuations leading to transient multipoles produced by separated δ^\pm charges in atoms A and B. In the dipole limit, the effect of a third atom C is described by the Axilrod-Teller expression, Eq. (15), and can be attractive or repulsive depending on the relative orientations given by the three interior angles θ_A , θ_B , and θ_C . Higher-order multipoles and $n > 3$ -body interactions are also present.

$$E_{\text{disp}}^{(2)}(R_A, R_B) = -\frac{C_6}{R_{AB}^6} - \frac{C_8}{R_{AB}^8} - \frac{C_{10}}{R_{AB}^{10}} + (\text{higher orders}), \quad (13)$$

where $R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$ is the internuclear distance in Fig. 8.

The dispersion coefficient C_6 appearing in Eq. (13) is familiar from the Lennard-Jones two-body potential (of spherically symmetric centers) and arises from correlated dipole-dipole fluctuations. However, its attractive R^{-6} dependence is only the asymptotic behavior at large separations. Higher multipoles interact through shorter ranged forces. The first of these arises from dipole-quadrupole correlations (C_8 coefficient) and C_{10} mixes quadrupole-quadrupole and dipole-octopole correlations. Equation (13) therefore includes contributions from fluctuating multipoles. Measurements on ultracold atoms have now made it possible to obtain high accuracy values for the C_6 coefficients for several atomic species and provided constraints and reduced uncertainty on higher-multipole coefficients (Van Kempen *et al.*, 2002).

Although the fluctuating electrostatic moments leading to dispersion interactions are spontaneous distortions of the charge density, we may still surmise that they are related in some way to the same polarizabilities that govern induction [Eq. (11)]. In fact a rigorous relationship exists which links the C_6 coefficient (arising from dipole-dipole correlations) to the frequency-dependent generalization of the static polarizability tensor $\alpha(i\omega)$ of interacting molecules A and B evaluated at complex frequency $i\omega$. This is expressed by the Casimir-Polder equation (Casimir and Polder, 1948; Parrill and Lipkowitz, 2016) which establishes the linear response relationship as follows:

$$C_6 = \frac{3}{\pi} \int_0^\infty d\omega \alpha^A(i\omega) \alpha^B(i\omega). \quad (14)$$

Correspondingly, higher-order terms in the dispersion series can be computed from the Casimir-Polder expression using higher-multipole dynamic polarizabilities (Parrill and Lipkowitz, 2016). The extension of Eq. (7) to include the frequency-dependent linear response is discussed, for example, by Wheatley (2008).

The leading-order correction ($\propto R^{-9}$) for many-body effects arises from the three-body (triple-dipole) contribution (DDD) to the total energy. This was first reported independently by Axilrod and Teller (1943) and Muto (1943),

$$E_{\text{disp}}^{(3)}(R_{ij}) \propto \frac{C_9}{R_{AB}^3 R_{AC}^3 R_{BC}^3}. \quad (15)$$

Here the R_{ij} are permutations of the A, B, and C nuclear positions in Fig. 8 and $C_9 \propto \alpha^3$, where α is the molecular polarizability. Since then, further contributions to the non-additive third-order dipole term (DDD) arising from quadrupolar (Q) (e.g., DDQ, DQQ, QQQ, etc.) have been considered (Bell, 1970; Elrod and Saykally, 1994). In the case of solid argon, the dipole-dipole-quadrupole (DDQ) energy is one-third as large as the triple-dipole (DDD) energy (Johnson and Spurling, 1971).

In the case of weakly polarizable molecules, the pairwise-additive approximation is often sufficient to capture essential physics possibly including higher-order terms in the dispersion series. However, many molecular systems are not weakly polarizable making pairwise approximation less viable. For nonpolar systems dominated by dispersion forces (e.g., noble gas crystals) estimates of interatomic many-body

contributions to cohesive energies may reach 10% (von Lilienfeld and Tkatchenko, 2010). Also, several important classes of problem depend on energy differences between molecular systems in dissimilar environments. These might include, for example, energy ordering among polymorphs or self-assembled motifs. Here small contributions from non-additive terms become particularly significant and may lead to qualitative and quantitative failures of the pairwise approximation. Broadly, there is evidence that pairwise methods tend to overbind molecular crystals: They perform better for assemblies of relatively small molecules but fail to capture essential collective effects in larger systems. A particularly notable class of supramolecular assemblies is the wide range of host-guest systems (Risthaus and Grimme, 2013) which comprises some form of relatively large dock site and a smaller molecule where cohesion occurs entirely via non-covalent interactions (Ambrosetti *et al.*, 2014). Some of these minimal model systems for molecular recognition processes appear to exhibit extreme collective phenomena which involve many-body forces well beyond the Axilrod-Teller (three-body) level.

D. Breakdown of transferability and a pathway to next-generation strategies

We show in this section that molecules exhibit a rich and varied set of electronic responses even in environments often encountered in condensed phases. Across a wide range of commonly encountered conditions, the manifestations and relative influences of electrostatic, inductive, and dispersive interactions are rebalanced and the importance of treating many-body phenomena nonperturbatively is key. This presents a fundamental challenge for the development of model potentials where, normally, the force types included (or neglected) and their relative importance are decided in advance. The resulting (and often extreme) truncation of the response hierarchy therefore limits transferability of models between distinct conditions or environments such as between bulk liquids, surfaces, interfaces, clusters, along the liquid-vapor coexistence curve, or in supercooled or supercritical conditions. Such artificial truncation challenges all perturbation theory inspired approaches but modern physics presents a way forward—using an exactly solvable model within Gaussian statistics to generate approximate interactions to all orders.

In the next sections we introduce an alternative approach to the problem expressed in Eq. (2) and a general pathway to broad transferability in a practical model. In particular, we will describe a minimal molecular model capable of generating the full many-body inductive and dispersive response hierarchy that has been developed here. The “reduced” electronic structure arises from a quantum mechanical treatment of a set of distributed quantum (Drude) oscillators (QDOs) (Whitfield and Martyna, 2006, 2007). Unlike “full” electronic structure methods, e.g., with full configuration interaction (Martyun, 2004), this coarse-grained construction is simple enough to be applied to relatively large condensed phase systems—tractable low order scaling methods such as Car-Parrinello-type computations are based on approximate density functional theory methods and have higher prefactor (and most often scaling) than QDOs. Also, due to the necessary use of

approximate functionals dispersion is either neglected, included empirically at the pair level, and/or truncated at the dipole level. We note that our proposed approach in its current form is applicable to molecular systems without appreciable electron transfer only. In order to cover the charge-transfer processes significant extension of the model is required.

We describe how on-the-fly forces arising from the Born-Oppenheimer surface [$E_{\text{el}}(\mathbf{R})$ in Eq. (2)] may be generated with linear scaling in the number of atoms in the system using adiabatic path integral molecular dynamics (APIMD) to generate a nonperturbative solution within electronically coarse-grained framework (Cao and Martyna, 1996; Jones, Crain, Sokhan *et al.*, 2013). We demonstrate the applicability of the approach to challenging systems such as condensed phases of noble gases (where dispersion dominates) and to water (where polarization is significant and coexists with dispersion). Predictions of the model are achieved with substantially reduced empirical input suggesting that the framework is rich enough to embody the essential physics of emergent systems and yet simple enough to retain intuitive features.

III. ELECTRONIC COARSE GRAINING WITH QUANTUM OSCILLATORS

Here we introduce a new class of computational models for materials that is rich enough to capture the full hierarchy of inductive and dispersive responses at the atomic and molecular scale, yet simple enough to render it practical for application to large complex systems by employing Gaussian statistics, fast computers, and novel modern simulation.

A. A minimal model with complete long-range interactions

The objective is to develop a model with a rich BO surface capable of generating a complete description of long-range many-body forces and which can be sampled within an efficient computational framework.

This is achieved by embedding a charged quantum harmonic oscillator in the molecular frame with parameters chosen to reproduce reference properties of the isolated molecule. The classical version of this construction is far from new and was used by Paul Drude over a century ago (Drude, 1900) to explore the optical properties of materials.

For this classical Drude oscillator, an external electric field will shift the equilibrium position of the charge, to produce an induced dipole, but as the position of a classical particle is precisely defined at any instant, this point distribution is too limited to permit higher-multipole polarizations. By contrast, the quantum Drude oscillator is much richer, yielding an extended charge distribution. This profoundly changes the behavior and brings the model’s physics much closer to that of real atoms and molecules. Specifically, higher-order responses such as multipole induction arise, as shown in Sec. III.A.1.

The quantum treatment of the oscillator also brings about zero-point fluctuations of the multipole moments at all orders. The full dispersion series (see Sec. II.C) therefore emerges naturally within the same framework albeit within Gaussian statistics.

In the next sections we discuss the theoretical underpinnings of quantum Drude oscillators, construction of models of realistic molecules and parametrization strategies, including the handling of short-range repulsion. We then describe a path integral method that solves the coarse-grained electronic problem nonperturbatively, retaining all the responses of the theoretical model. We note here that nonlinear polarization effects, the subject of nonlinear optics (Bloembergen, 1996), are outside the current consideration but are in principle captured by the model with small modifications. Finally, we discuss ideas introduced previously, such as the concept of adiabaticity, calculation of forces via the Hellmann-Feynman theorem, and the generation of a realistic BO surface.

1. Polarization responses

A quantum Drude oscillator (London, 1937; Jones, Crain, Sokhan *et al.*, 2013), models a bound state of two point charges of opposite sign, a light particle of mass μ with a charge $-q$ that is harmonically tethered to a neutralizing charge $+q$ embedded in the nuclear framework of the molecule. Its angular frequency ω is related to force constant $k = m\omega^2$. Specifically, in this construction, the two charges do not “see” each other electrostatically, but interact with neighboring fields. This is one of the features that makes this model a type of tight binding model valid for fixed molecular connectivity. The “on-site” Coulomb attraction between electrons and nuclei of real materials is replaced by a simplified form—a tight binding approach.

The harmonic potential is an approximate representation of the bonding between valence electrons and nuclei, giving rise to a ladder of energy states. However, in the electric fields produced in condensed systems, electronic responses involve a limited range of orbitals. Intuitively, a major contribution to molecular polarization will be due to the occupation of the lowest unoccupied molecular orbital and a few low-lying empty orbitals by the valence electrons (see Sec. II.B.4). To reproduce these effects it is therefore not necessary to describe the full spectrum of atomic energy states. A simplified model of low-lying states may be sufficient for many practical problems. This is exactly what a quantum Drude oscillator framework affords—by replacing the on-site electron-nuclear interaction with a harmonic bond to create a coarse-grained model within Gaussian statistics.

When immersed in a uniform electric field F_x , this oscillator acquires a displacement x determined by the balance between the electrostatic and harmonic forces. Setting qF_x equal to $\mu\omega^2x$ gives a dipole moment $qx = (q^2/\mu\omega^2)F_x$ and thus a dipole polarization $\alpha_1 = q^2/\mu\omega^2$.

A classical calculation would end here. But, a quantum mechanical oscillator has a full hierarchy of responses due to the excitation of the ground state electron into excited states of the oscillator. These states are separable into normalized radial and angular parts $R_{kl}(r)$ and $Y_{lm}(\theta, \phi)$:

$$\psi_{klm}(r, \theta, \phi) = R_{kl}(r)Y_{lm}(\theta, \phi). \quad (16)$$

Here $Y_{lm}(\theta, \phi)$ are the spherical harmonics and $R_{kl}(r)$ are proportional to the generalized Laguerre polynomials. The radial quantum number k quantifies the distance between the negative particle and the center. The angular quantum number

l quantifies the angular momentum and the resulting electrostatic moment. Specifically, a spherical harmonic of angular momentum l corresponds to an electrostatic moment of order l . In other words, $l = 1$ causes an induced dipole moment, $l = 2$ a quadrupole moment, and so on. This can be seen in Fig. 6 and detailed in Appendix A.

The energy of the isolated oscillator (or reference system) depends on both the radial, k , and angular, l, m , quantum numbers as follows:

$$E(k, l, m) = \hbar\omega(2k + l + \frac{3}{2}). \quad (17)$$

The energy spectrum is simple—increasing in a ladderlike fashion by multiples of $\hbar\omega$ with each excitation up the ladder. The spectrum of real molecules is more complex; however, we shall show that this simple spectrum, with appropriate choice of parameters, can mimic the responses of complex assemblies surprisingly well—Gaussian statistics is generally observed to be quite robust belying its simplicity.

In order to better understand the limitations and successes of our Gaussian model we first perform a Raleigh-Ritz perturbation theory analysis following Stone (2013). We note that fields around molecules are strong and have complex symmetries. Thus, single particle oscillator states with high angular and radial quantum numbers will mix into the many-body wave function. Jones, Crain, Sokhan *et al.* (2013) provided a diagrammatic analysis that shows how high order dispersion, polarization, and cross responses emerge as a function of the number of bodies and powers in \hbar . Again, due to high field strength, even interactions with high powers in \hbar and many-body character and can become important in condensed phase systems, particularly at interfaces or near transitions where small terms of appropriate symmetry can effect the emergent physics.

We next consider the form of the higher-order responses (Jones, Crain, Sokhan *et al.*, 2013) by introducing spherical tensors, whose role in the expansion of electrostatic fields is detailed in Appendix A. For our discussion here, we need only to remember that higher-angular momentum spherical harmonics correspond to higher-multipole moments. This intuition leads to a natural definition of *spherical multipole moments* as follows (Stone, 2013):

$$Q_{lm} = \int \rho(\mathbf{r})r^l C_{lm}(\theta, \phi)d^3\mathbf{r}, \quad (18)$$

where

$$C_{lm}(\theta, \phi) = \left(\frac{4\pi}{2l+1}\right)^{1/2} Y_{lm}(\theta, \phi) \quad (19)$$

are rescaled versions of the spherical harmonics. Intuitively, the spherical multipole moments decompose the charge distribution into an orthonormal basis set defined by the spherical harmonics, with higher-order spherical harmonics corresponding to higher-order electrostatic moments. Q_{lm} are in general complex for $m \neq 0$ and we will use this form as it simplifies perturbation calculations. If needed, real forms can be defined, as detailed by Stone (2013). For our purpose, as polarizabilities of a spherical harmonic oscillator will be isotropic, we are mainly interested in the $m = 0$ component Q_{l0} , which is equal

to the diagonal component of the corresponding Cartesian moment: $Q_{00} = q$, $Q_{10} = \mu_z$, $Q_{20} = \Theta_{zz}$, etc.

This representation is a spherical *tensor* since it transforms like a spherical harmonic under a rotation described by three Euler angles α , β , and γ ,

$$R(\alpha, \beta, \gamma) Q_{lm} = \sum_{m'} Q_{lm'} D_{m'm}^l(\alpha, \beta, \gamma), \quad (20)$$

where $D_{m'm}^l(\alpha, \beta, \gamma)$ are the Wigner rotation matrices (Stone, 2013). More detail is given in Appendix A.

The spherical multipole moments are more natural quantities to work with in atomic and molecular physics than the Cartesian ones, as they are an irreducible representation. Molecular moments are not point charges aligned neatly in a line, but are extended, three-dimensional distributions of charge that are more naturally described by the spherical harmonics. Since the spherical harmonics are the natural basis in which to have a natural structure for describing electrostatic moments and make the calculations more natural due to the orthonormality of the spherical harmonics, we will use them from now on in our analysis.

The definition of Q_{lm} naturally gives rise to a spherical multipole operator:

$$\hat{Q}_{lm} = qr^l C_{lm}(\theta, \phi) = qr^l \left(\frac{4\pi}{2l+1} \right)^{1/2} Y_{lm}(\theta, \phi). \quad (21)$$

In terms of this operator, the spherical polarizabilities are defined perturbatively as (Stone, 2013)

$$\alpha_{lm,l'm'} = \sum_n \left(\frac{\langle 0 | \hat{Q}_{lm} | n \rangle \langle n | \hat{Q}_{l'm'}^\dagger | 0 \rangle}{E(n) - E(0)} + \frac{\langle 0 | \hat{Q}_{l'm'}^\dagger | n \rangle \langle n | \hat{Q}_{lm} | 0 \rangle}{E(n) - E(0)} \right), \quad (22)$$

where the quantum state $|n\rangle = |klm\rangle$ is described by the three quantum numbers previously introduced.

The matrix elements in this perturbation calculation are the following:

$$\begin{aligned} \langle 000 | \hat{Q}_{lm} | k'l'm'' \rangle &= \int_r \int_\Omega R_{00}^\dagger(r) Y_{00}^\dagger(\theta, \phi) \left[qr^l \left(\frac{4\pi}{2l+1} \right)^{1/2} Y_{lm}(\theta, \phi) \right] \\ &R_{k'l'm''}(r) Y_{l'm''}(\theta, \phi) dr d\Omega \\ &= \int_r R_{00}^\dagger(r) qr^l R_{kl'}(r) dr \\ &\int_\Omega Y_{00}^\dagger(\theta, \phi) \sqrt{\frac{4\pi}{2l+1}} Y_{lm}(\theta, \phi) Y_{l'm''}(\theta, \phi) d\Omega. \end{aligned} \quad (23)$$

In order to simplify these expressions, we use the following orthonormality and ladder properties of the harmonic oscillator eigenfunctions:

$$\int_\Omega Y_{lm}^\dagger(\theta, \phi) Y_{l'm''}(\theta, \phi) d\Omega = \delta_{l,l'} \delta_{m,-m''}, \quad (24)$$

$$\int_r R_{kl}^\dagger(r) R_{k'l'}(r) dr = \delta_{k,k'} \delta_{l,l'}, \quad (25)$$

$$Y_{00}(\theta, \phi) = \frac{1}{\sqrt{4\pi}}, \quad (26)$$

$$r^l R_{00}(r) = \xi_l R_{0l}(r), \quad (27)$$

$$\xi_l^2 = (2l+1)!! \left(\frac{\hbar}{2\mu\omega} \right)^l. \quad (28)$$

Substituting into the matrix elements (23) gives

$$\langle 000 | \hat{Q}_{lm} | k'l'm'' \rangle = \sqrt{\frac{q^2}{2l+1}} \xi_l \delta_{0,k} \delta_{l,l'} \delta_{m,-m''}. \quad (29)$$

Substituting back in Eq. (22) and simplifying gives

$$\alpha_{lm,l'm'} = \frac{q^2}{\mu\omega^2} \left[\frac{(2l-1)!!}{l} \right] \left(\frac{\hbar}{2\mu\omega} \right)^{l-1} \delta_{l,l'} \delta_{m,m'}. \quad (30)$$

Given that the polarizabilities are isotropic and the previously shown relationships between the Cartesian and spherical polarizabilities, we can define the *isotropic polarizabilities* as

$$\alpha_l = \frac{q^2}{\mu\omega^2} \left[\frac{(2l-1)!!}{l} \right] \left(\frac{\hbar}{2\mu\omega} \right)^{l-1}, \quad (31)$$

with α_1 representing the dipole polarizability [introduced earlier in the tensor form as $\alpha_{\alpha\beta}$ in Eq. (11)], α_2 the quadrupole (corresponding to the tensor $C_{\alpha\beta\gamma\delta}$), and so on for a full infinite series of electrostatic responses arising due to a perturbing electric field.

It is worth noting that $\alpha_l \sim \hbar^{l-1}$, meaning that only α_1 has no dependency on \hbar . This means that the classical limit of the model has only a dipole polarizability given by $\alpha_1 = q^2/\mu\omega^2$, exactly as the calculation presented at the start of this section.

2. Zero-point fluctuations and the emergence of dispersion

As in the previous section, we begin by looking at a simple system: two interacting 1D harmonic oscillators of identical mass μ , frequency ω , and charge q . The oscillators have a displacement x_1 and x_2 , are parallel to each other and separated by a distance R . Their interaction energy is, in atomic units,

$$E = \frac{(qx_1)(qx_2)}{R^3}. \quad (32)$$

Thus, the Hamiltonian of the system is

$$\hat{H} = \frac{\hat{p}_1^2}{2\mu} + \frac{\hat{p}_2^2}{2\mu} + \frac{1}{2}\mu\omega^2(x_1^2 + x_2^2 + \gamma x_1 x_2), \quad (33)$$

where

$$\gamma = 2 \frac{q^2}{\mu\omega^2} \frac{1}{R^3} = 2 \frac{\alpha_1}{R^3}. \quad (34)$$

Changing coordinates to

$$\hat{p}_\pm = \frac{1}{\sqrt{2}}(\hat{p}_1 \pm \hat{p}_2)$$

and

$$x_\pm = \frac{1}{\sqrt{2}}(x_1 \pm x_2)$$

decouples the system into two isolated oscillators

$$\hat{H} = \left(\frac{\hat{p}_+^2}{2\mu} + \frac{1}{2}\mu\omega^2(1+\gamma)x_+^2 \right) + \left(\frac{\hat{p}_-^2}{2\mu} + \frac{1}{2}\mu\omega^2(1-\gamma)x_-^2 \right). \quad (35)$$

The ground state energy of this system is, to leading order in γ ,

$$E_0 = \frac{1}{2}\hbar\omega(\sqrt{1+c} + \sqrt{1-c}) \approx \hbar\omega[1 - \frac{1}{8}\gamma^2 + O(\gamma^3)]. \quad (36)$$

The leading-order correction to the ground state energy is thus

$$\frac{\hbar\omega}{8}\gamma^2 = \frac{1}{2}\hbar\omega\alpha_1\alpha_1\frac{1}{R^6}. \quad (37)$$

Thus, the leading-order correction is the standard van der Waals interaction, proportional to R^{-6} and α_1^2 as first derived by London.

Having established the fact that two quantum harmonic oscillators can reproduce from physical principles the leading-order van der Waals interaction between neutral atoms, we now derive the interaction coefficients and energies of the full model. Given two three-dimensional quantum Drude oscillators in the same geometry, their interaction Hamiltonian is

$$H'_{AB} = Q_{lm}^A T_{lm,l'm'}^{AB} Q_{l'm'}^B, \quad (38)$$

where $T_{lm,l'm'}^{AB}$ is the spherical tensor version of the electrostatic interaction tensor.

The second-order perturbation energy is the following, with the n indexing the quantum states specified by the three indices k , l , and m :

$$E^{(2)} = -\sum_{n_A} \sum_{n_B} \frac{\langle 0, 0 | H'_{AB} | n_A, n_B \rangle}{E(n_A) + E(n_B)}. \quad (39)$$

As before, due to the orthonormality of R_{kl} , the sum simplifies to

$$E^{(2)} = -\sum_{l_A m_A} \sum_{l_B m_B} \frac{[T_{l_A m_A, l_B m_B}^{AB} \langle 000 | Q_{l_A m_A}^A | 0 l_A m_A \rangle \langle 000 | Q_{l_B m_B}^B | 0 l_B m_B \rangle]^2}{l_A \hbar\omega_A + l_B \hbar\omega_B}. \quad (40)$$

Since the quantum Drude oscillator model considered here is isotropic (anisotropy developed by interaction of multiple

spherical QDOs), $\langle 000 | Q_{lm}^A | 0lm \rangle = \langle 000 | Q_{l0}^A | 0l0 \rangle$ for all m . Thus, the energy reduces to

$$E^{(2)} = -\sum_{l_A l_B} |T_{l_A l_B}^{AB}|^2 \frac{\langle 000 | Q_{l_A 0}^A | 0 l_A 0 \rangle^2 \langle 000 | Q_{l_B 0}^B | 0 l_B 0 \rangle^2}{l_A \hbar\omega_A + l_B \hbar\omega_B}, \quad (41)$$

where

$$|T_{l_A l_B}^{AB}|^2 = \sum_{m_A m_B} [T_{l_A m_A, l_B m_B}^{AB}]^2. \quad (42)$$

Substituting the matrix elements from Eq. (23) gives

$$E^{(2)} = -\sum_{l_A l_B} |T_{l_A l_B}^{AB}|^2 \frac{q_A^2}{2l_A + 1} \frac{q_B^2}{2l_B + 1} \frac{\xi_A^2 \xi_B^2}{l_A \hbar\omega_A + l_B \hbar\omega_B}. \quad (43)$$

Simplifying,

$$E^{(2)} = -\sum_{l_A l_B} |T_{l_A l_B}^{AB}|^2 \alpha_{l_A}^A \alpha_{l_B}^B \left[\frac{\hbar}{4} \frac{l_A l_B \omega_A \omega_B}{l_A \omega_A + l_B \omega_B} \right]. \quad (44)$$

The m traces of the electrostatic interaction tensors are

$$\begin{aligned} |T_{11}^{AB}|^2 &= 6R^{-6}, \\ |T_{12}^{AB}|^2 &= 15R^{-8}, \\ |T_{13}^{AB}|^2 &= 28R^{-10}, \\ |T_{22}^{AB}|^2 &= 70R^{-10}. \end{aligned} \quad (45)$$

Substituting into Eq. (44)

$$E^{(2)} = -C_6 R^{-6} - C_8 R^{-8} - C_{10} R^{-10} + O(R^{-12}), \quad (46)$$

with the two-body dispersion coefficients given by

$$\begin{aligned} C_6^{AB} &= \frac{3}{2} \alpha_1^A \alpha_1^B \frac{\hbar\omega_A \omega_B}{\omega_A + \omega_B}, \\ C_8^{AB} &= \frac{15}{2} \left[\alpha_1^A \alpha_2^B \frac{\hbar\omega_A \omega_B}{\omega_A + 2\omega_B} + \alpha_2^A \alpha_1^B \frac{\hbar\omega_A \omega_B}{2\omega_A + \omega_B} \right], \\ C_{10}^{AB} &= \left[21 \alpha_1^A \alpha_3^B \frac{\hbar\omega_A \omega_B}{\omega_A + 3\omega_B} + 21 \alpha_3^A \alpha_1^B \frac{\hbar\omega_A \omega_B}{3\omega_A + \omega_B} \right. \\ &\quad \left. + 70 \alpha_2^A \alpha_2^B \frac{\hbar\omega_A \omega_B}{2\omega_A + 2\omega_B} \right]. \end{aligned} \quad (47)$$

If the quantum Drude oscillators are identical, these simplify to

$$\begin{aligned} C_6 &= \frac{3}{4} \alpha_1 \alpha_1 \hbar\omega, \\ C_8 &= 5 \alpha_1 \alpha_2 \hbar\omega, \\ C_{10} &= \left[\frac{21}{2} \alpha_1 \alpha_3 + \frac{35}{2} \alpha_2 \alpha_2 \right] \hbar\omega. \end{aligned} \quad (48)$$

The use of a Gaussian model has limitations in that there are obvious sum rules that will be only approximately obeyed by real systems as discussed next.

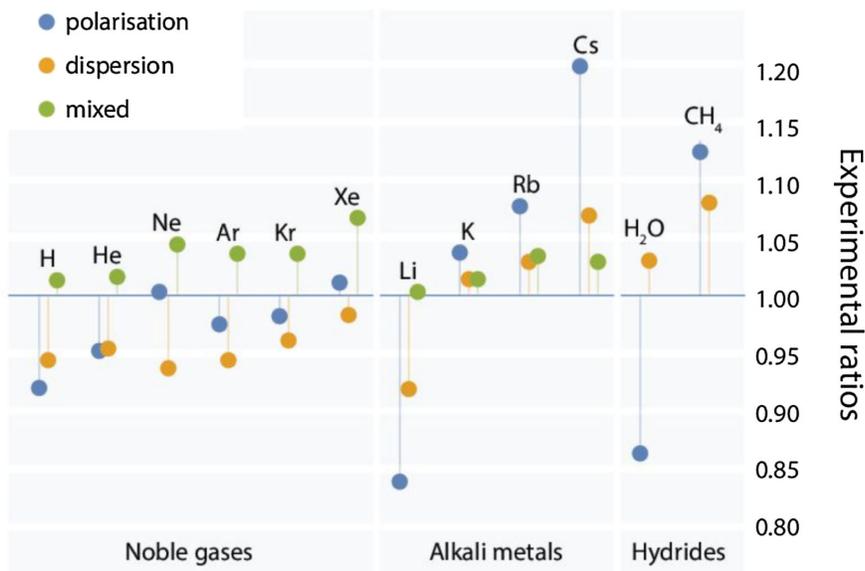


FIG. 9. Three types of invariant ratios between polarization and dispersion coefficients, which arise within the quantum Drude oscillator model as introduced here. Polarization ratios involve only polarization coefficients and analogously for dispersion ratios. Mixed ratios involve both polarization and dispersion coefficients. Deviation from theory is shown for three types of atoms and molecules: noble gases, alkali metals, and small hydrides. From Cipcigan *et al.*, 2016.

B. Model invariants for the QDO and the behavior of real atoms and molecules

1. Relationships between electronic responses

Since the polarization and dispersion coefficients depend only on three fundamental parameters of the oscillator (μ , q , and ω), these can be combined to form model *invariants*, closed expressions within QDO's Gaussian statistics. Specifically, combining the lowest-order response coefficients gives

$$\begin{aligned} \sqrt{\frac{20}{9}} \frac{\alpha_2}{\sqrt{\alpha_1 \alpha_3}} &= 1, \\ \sqrt{\frac{49}{40}} \frac{C_8}{\sqrt{C_6 C_{10}}} &= 1, \\ \frac{C_6 \alpha_1}{4C_9} &= 1. \end{aligned} \quad (49)$$

Figure 9 shows the values of these quantities for noble gases, alkali metals, and small hydrides. The agreement is within 15% for all, with some with a 10% or better agreement. This occurs because the outer shell of these elements is algebraically spherical for the noble gases (filled shell) and alkali metals (s orbitals) and close to spherical for hydrides (due to the electronegativity of the dominant atom).

Taken together, the foregoing results imply that the QDO construction contains a coarse-grained but complete (at long range) set of the electronic structure response of an atom or molecule to external influences. The reaction of this Gaussian charge distribution to external fields naturally leads to many-body polarization. Quantum correlations between QDOs lead to van der Waals interactions. This includes three-body and all higher level terms in the interaction energy.

The QDO invariants therefore appear to reflect the behavior fairly well and show promise to be transferable to complex systems. In the next section we demonstrate how the response functions within the QDO model can be independently scaled so as to allow exploration of how such electronic responses influence physical properties.

2. Parametrization

An alternative combination of the coefficients gives rules for parametrizing QDOs from the response coefficients:

TABLE I. The QDO parameters for various species determined using the QDO parametrization rules. $E_h \approx 27.211$ eV is the Hartree energy, $m_e \approx 9.11 \times 10^{-31}$ kg is the mass of an electron, and $e \approx 1.60 \times 10^{-19}$ C is the elementary charge. All of these parameters are electronic in magnitude, with the Hartree energy being twice the ionization energy of a hydrogen atom. From Jones, Crain, Sokhan *et al.*, 2013.

Species	$\omega/(E_h/\hbar)$	m/m_e	q/e
H	0.4273	0.6099	0.708
Li	0.0687	1.2545	0.9848
K	0.063	0.8101	0.967
Rb	0.0603	0.7343	0.9274
Cs	0.0531	0.6939	0.895
He	1.0187	0.5083	0.8532
Ne	1.2965	0.3491	1.2494
Ar	0.7272	0.302	1.3314
Kr	0.6359	0.2796	1.3741
Xe	0.5152	0.2541	1.357
H ₂ O	0.6287	0.3656	1.1973
NH ₃	0.5603	0.3541	1.2722
CH ₄	0.5794	0.2615	1.2313
BH ₃	0.8776	0.1165	1.0793

$$\begin{aligned}\omega &= \frac{1}{\hbar} \frac{4C_6}{3\alpha_1^2}, \\ \mu &= \frac{\hbar}{\omega} \frac{3\alpha_1}{4\alpha_2}, \quad \text{or} \quad \mu = \frac{\hbar}{\omega} \frac{5C_6}{C_8}, \\ q &= \sqrt{\mu\omega^2\alpha_1}.\end{aligned}\quad (50)$$

The resulting parameters are displayed in Table I and shown in Fig. 10. Similar species have similar parameters, with the exception of hydrogen, whose parameters are closer to noble gases than to group I metals. There is in general a larger variation of the parameters between the different categories than within a single one. This is indicative of the fact that only the outermost electrons contribute significantly to polarization

and dispersion. Hydrides and noble gases (with the exception of hydrogen and helium) have $q > e$ which indicates the fact that there is more than one electron in the outer shell. However, q is still close to a single electron's charge, varying from $\sim 0.7e$ for hydrogen to $\sim 1.4e$ for xenon. This indicates that most of the contribution from polarization can be effectively treated by the excitation of a single quasiparticle. Since group I metals only have a single electron in their unfilled shell, this might be expected intuitively.

3. Scaling of the response functions

Finally, we can derive scaling relationships between the parameters of the QDO that increase or decrease the polarization coefficients independent of each other. For example, we

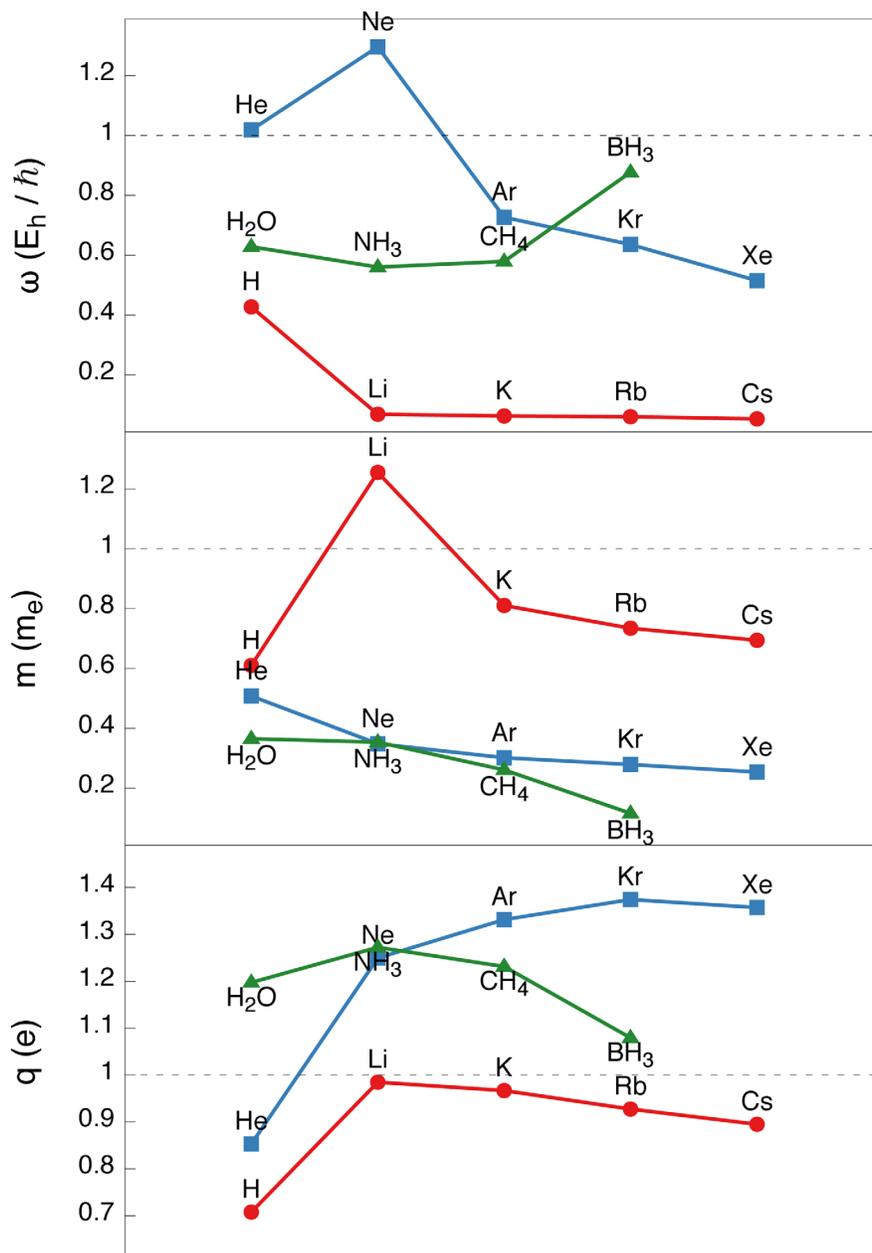


FIG. 10. Variation in QDO parameters for various noble gases, first row metals, and simple hydrides. From Jones, Crain, Sokhan *et al.*, 2013.

can increase or decrease all dispersion coefficients of a model, while keeping the polarization constant and vice versa. These tunable parameters allows us to interrogate both types of electronic responses separately.

To calculate these scaling relationships, we note that

$$\begin{aligned} \alpha_l &\sim \frac{q^2}{\mu\omega^2} \left(\frac{\hbar}{2\mu\omega} \right)^{l-1}, \\ C_n &\sim \sum \alpha_j \alpha_k \hbar\omega, \end{aligned} \quad (51)$$

with the sum over all possible j and k such that $n = 2(j + k + 1)$. Consider scaling $\{\mu, \omega, q\}$ to $\{\mu', \omega', q'\}$ such that $C'_n = \kappa C'_n$ and $\alpha'_l = \lambda \alpha_l$ for all n and l . This requires

$$\begin{aligned} (\lambda\alpha_j)(\lambda\alpha_k)\hbar\omega' &= \kappa\alpha_j\alpha_k\hbar\omega, \\ \mu'\omega' &= \mu\omega, \\ \frac{(q')^2}{\mu'(\omega')^2} &= \lambda \frac{q^2}{\mu\omega^2}. \end{aligned} \quad (52)$$

Solving them for $\{\mu'\omega'q'\}$ finally gives

$$\begin{aligned} \omega' &= \frac{\kappa}{\lambda^2} \omega, \\ \mu' &= \frac{\lambda^2}{\kappa} \mu, \\ q' &= \sqrt{\frac{\kappa}{\lambda}} q. \end{aligned} \quad (53)$$

Therefore, setting $\lambda = 1$ would scale all dispersion coefficients by κ leaving polarizabilities intact. Alternatively, setting $\kappa = 1$ would scale all polarizabilities only by λ . This is of high interest in studying model systems wherein polarization effects can be examined at fixed dispersion interaction strength and vice versa allowing a novel understanding to emerge.

C. A route to real systems: Damping and short-range repulsion

We have seen that at large intermolecular distances, where the interaction is well described in terms of monomer properties, the QDO model gives an accurate description including all many-body effects. At short range, where a significant overlap in electronic distributions of monomers starts to occur, two factors arise that affect the performance of the model.

The first is related to the fact that QDOs are distinguishable (spinless) quasiparticles which obey Boltzmann (distinguishable particle) statistics due to the replacement of the Coulomb electron-nuclear (on-site) interaction with a harmonic bond. Therefore, certain provisions should be made for the missing exchange repulsion (discussed in Sec. II.A.3), arguably the largest of the charge penetration effects. It is an important ingredient of any classical force-field recipe, where using the simplifying fact that all short-range energy contributions are approximately additive (Stone, 2013), it is often given in simple Lennard-Jones (A/r^{12}) or Born-Mayer [$A \exp(-r/b)$] forms, although more refined schemes exist (Van Vleet *et al.*, 2016). The problems with and controversies in accurate

determination of many-body contributions to the short-range interactions, both theoretically and experimentally, have been thoroughly analyzed and discussed by Elrod and Saykally (1994), in particular, the apparent contradiction in condensed phase properties described by pair potentials with the ATM term and small-cluster results, indicating that three-body exchange effects largely cancel the ATM term (in systems such as water but not in van der Waals systems). In the QDO approach, based on pairwise approximation to exchange, as demonstrated later for the case of liquid water (Sec. V.B) and clusters (Sec. V.B.5), for both systems the results are consistently accurate.

Because of the complex nature of intermolecular interactions at short range (Stone, 2013) there is no direct route to derive an analytic expression in this case. One way to grossly account for the missing interactions in the QDO approach is empirical: fitting the difference between a reference potential and the QDO energy to a suitable analytic expression using a single cut through the potential energy surfaces of a dimer, which also provides a good test of the accuracy of the pairwise approximation. It is believed that these interactions decay approximately exponentially (Stone and Tong, 1994). A triexponential decay, used in early QDO work (Whitfield and Martyna, 2006, 2007),

$$U_{\text{rep}}(r) = \sum_{i=1}^3 \kappa_i e^{-\lambda_i r}, \quad (54)$$

provides a suitable, albeit not unique solution.

In such an approach a high quality reference potential energy surface is required. It can be provided, e.g., by coupled-cluster (CC) calculations (Crawford and Schaefer, 2007) with single, double, and perturbative triple excitations included [CCSD(T)], the so-called ‘‘gold standard’’ in calculating energy surfaces for molecular complexes (Řezáč and Hobza, 2013). Currently, extensive data sets using this approach are under development (Patkowski, 2017), providing a reference quality benchmark database of nonbonded potentials taken at the complete basis set (CBS) limit. Application of this procedure to noble gases is straightforward and Fig. 11 shows excellent agreement in QDO *predicted* and experimental compressibility factors of krypton along the (supercritical) isotherm of $T = 251.4$ K. The inset illustrates *ab initio* ground state and QDO potential energies for the krypton dimer. *Ab initio* energy was calculated with augmented correlation-consistent Dunning basis sets (Dunning, 1989) extrapolated to CBS (Jäger *et al.*, 2016).

The second complicating factor is more technical and stems from the point charge approximation used for electrostatic interactions. Utilizing pure Coulomb interactions as effective interactions can fail at very short (intra-atomic) separations where charges no longer can be regarded as classical point objects. This is an abiding problem in atomistic simulation of condensed matter where to regularize both Coulomb and multipolar interactions various damping functions (Thole, 1981; Sprik and Klein, 1988; Whitfield and Martyna, 2006) have been used.

An efficient and convenient way to regularize the Coulomb interactions is to employ a smearing of charges (Sprik and Klein, 1988; Chialvo and Cummings, 1998) [see Thole (1981)

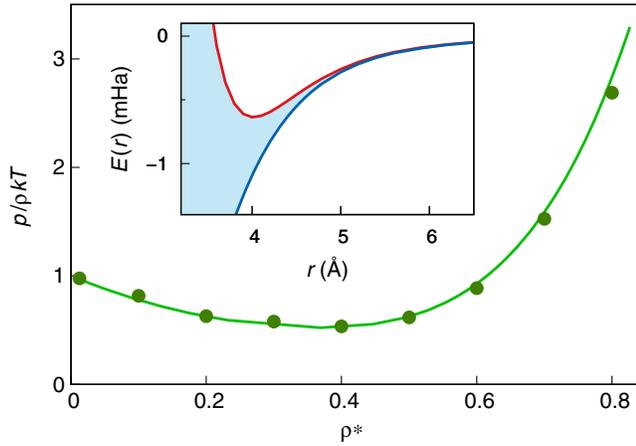


FIG. 11. Compressibility factor of krypton at 251.4 K as a function of reduced density. Solid line, empirical equation of state [Lemmon, McLinden, and Friend, 2018](#); points, QDO results. Reduced density ($\rho^* = \rho\sigma^3$) values are based on Lennard-Jones σ from the recent parametrization ([Vrabec, Stoll, and Hasse, 2001](#)). The inset shows a comparison of the reference *ab initio* dimer energy (red line) and neat QDO energy (blue line). The shadowed area illustrates the energy difference fit to an analytic expression.

for a general discussion of damping]. Interactions remain finite at all separations, which obviates the necessity of damping used in early work on QDO ([Whitfield and Martyna, 2007](#)) and can easily be incorporated in the framework of Ewald summation and its smooth particle mesh variant ([Essmann *et al.*, 1995](#); [Gingrich and Wilson, 2010](#)). One approach ([Jones, Cipcigan *et al.*, 2013](#); [Sokhan *et al.*, 2015b](#)) is to regularize using Gaussian charge distribution of width σ ,

$$\rho(r) = \frac{q}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{r^2}{2\sigma^2}\right). \quad (55)$$

For QDO quasiparticles it is convenient to set the Gaussian width of the center charge equal to the unperturbed oscillator width $\sigma = \hbar/\mu\omega$. The electrostatic energy between two Gaussian charges q_1 and q_2 with widths σ_1 and σ_2 is simply

$$V(r_{12}) = \frac{q_1 q_2}{r_{12}} \operatorname{erf}\left(\frac{r_{12}}{\sigma_{12}\sqrt{2}}\right), \quad (56)$$

where $\sigma_{12}^2 = \sigma_1^2 + \sigma_2^2$ and $r_{12} = |\mathbf{r}_2 - \mathbf{r}_1|$. [Thole \(1981\)](#) and [Madden and Wilson \(1996\)](#) have shown that the Gaussian approach is not necessarily the best.

IV. SOLVING THE COARSE-GRAINED ELECTRONIC PROBLEM

In the preceding sections we introduced the concept of the embedded quantum oscillator as a means of simplifying the electronic structure problem for atoms and molecules while retaining the full set of long-range responses. We next turn to strategies for computing physical properties within the electronically coarse-grained framework and for incorporating it into practical simulations of large systems.

We seek a nonperturbative solution strategy which preserves the rich physics of the model and all its responses without introducing further simplifications or assumptions. In particular, we seek to determine interactions between the nuclei directly from the Drude model electronic structure calculations, in the spirit of the Hellmann-Feynman theorem, thereby eliminating the need to devise condense phase empirical potential in advance.

A. Path integral molecular dynamics for the quantum Drude oscillator

1. Condensed phases and the density matrix

In the previous discussion, we employed perturbation theory to understand physically the long-range electronic responses at fixed molecular and atomic positions captured by the QDO model. Our ambition, however, is to move toward implementation in condensed phases where the constituent species evolve at finite temperature. As fields in condensed phases are large we want a nonperturbative solution designed to capture all the diagrams: continuum of environments with complex interfaces using a selected *a priori* subset of diagrams would almost certainly bias the results. We shall include all diagrams by developing a scalar field theory approach based on thermal density matrices with parameters selected for the physics desired—to generate the ground state properties of the quasiparticle QDO systems and develop highly accurate solutions that do not employ perturbation theory.

One advantage of the QDO model, although it has a limit, is its ability to model short-range interactions, is that it obeys distinguishable particle statistics—Fermi statistics is not required. The thermal density matrix $\rho^{(\text{QDO})}(\mathbf{r}, \mathbf{r}'; \beta_D; \mathbf{R})$ in the canonical ensemble at inverse temperature $T_D = 1/k_B\beta_D$ of the model is, therefore, positive semidefinite for all elements (diagonal and off-diagonal) at all temperatures, allowing it to be statistically sampled facily. The associated free energy is

$$\begin{aligned} F_D(\mathbf{R}) &= -\beta_D^{-1} \ln\{\operatorname{Tr}\hat{\rho}^{(\text{QDO})}(\beta_D, \mathbf{R})\} \\ &= E_0(\mathbf{R}) + \mathcal{O}(\exp(-\beta_D E_{10}(\mathbf{R}))), \end{aligned} \quad (57)$$

given a sufficiently low “electronic” temperature T_D is selected as desired here. In a typical simulation study the electronic temperature required to generate the ground state is $T_D = 1/k_B\beta_D \gg T$, where T is the true or molecular temperature. Thus, T_D is referred to as the “faux high electronic temperature”—for systems in their ground state, temperature is irrelevant. The approximation $\beta_D E_{10}(\mathbf{R}) \approx \beta_D \hbar\omega_D$ with ω_D setting the temperature $k_B T_D \approx \hbar\omega_D/8$. Thus, the way forward is to develop a formalism to statistically sample these QDO density matrices at a high electronic temperature T_D and then couple that to nuclear motion at temperature T . Surprisingly, using techniques developed over the last 25 years, these seemingly contradictory goals can be accomplished.

More formally, the evaluation of quantum mechanical observables in a statistical mechanical ensemble requires computation of both expectation values and statistical averages with appropriate weights determined by the physical conditions of the ensemble ([Gibbs, 1902](#)). In particular, if all objects in an ensemble are in the same pure state $|\psi\rangle$,

the expectation value for an observable (A) would be the elementary result $\langle A \rangle = \langle \psi | A | \psi \rangle$. With mixed states, however, the $|\psi_i\rangle$ occur only with some assigned probability p_i so the expectation value expression generalizes to $\sum_i p_i \langle \psi_i | A | \psi_i \rangle \equiv \text{tr}(\rho A) / \text{tr}(\rho)$, where $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$ defines the density matrix—again the expression for p_i depends upon ensemble.

In the canonical ensemble, for a system in equilibrium at temperature T , governed by a Hamiltonian \hat{H} with eigenvalues E_i and eigenstates $|i\rangle$ the statistical weights p_i are proportional to the Boltzmann factor $\exp(-\beta E_i)$ where $\beta = 1/k_B T$ is the inverse temperature. The thermal density matrix thus takes the form

$$\hat{\rho}(\beta) = e^{-\beta \hat{H}} = \sum_i e^{-\beta E_i} |i\rangle \langle i|, \quad (58)$$

$$\hat{\rho}(\beta)(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | e^{-\beta \hat{H}} | \mathbf{r}' \rangle = \sum_i \langle \mathbf{r} | i \rangle \langle i | \mathbf{r}' \rangle e^{-\beta E_i} \quad (59)$$

diagonal in the basis of energy eigenstates. The trace gives the partition function for the system,

$$Z(\beta) = \text{tr} \hat{\rho}(\beta) \equiv \sum_i \exp(-\beta E_i), \quad (60)$$

from which all thermodynamic quantities can be derived. The matrix elements of the density matrix in the position representation, which will be useful in the next section, are

$$\rho(\mathbf{r}, \mathbf{r}'; \beta) = \langle \mathbf{r} | e^{-\beta \hat{H}} | \mathbf{r}' \rangle. \quad (61)$$

If the temperature is low, correlation distances are long and conversely if the temperature is high, the matrix is short range. Properties of density matrices have been extensively reviewed (Fano, 1957; Blum, 2013).

In the next section we introduce the path integral as a means of computing quantum statistical properties.

2. Path integral formulation of the density matrix

Feynman's path integral concept (Feynman, 1949) is an alternative formulation of quantum theory that uses sums over paths to express the density matrix. Importantly for the present application, path integrals allow for an efficient formulation of nonperturbative solutions to quantum mechanical problems. The method is therefore the strategy of choice for coping with the many-body character of the QDO model when applied to condensed phase situations enabled by the absence of the Fermi statistics (sign) problem in the QDO as presaged earlier.

The central idea is based upon the notion of the propagator (time evolution operator) defined as

$$K(x, x'; t) = \langle x | \exp\left(-\frac{i}{\hbar} H t\right) | x' \rangle \quad (62a)$$

$$\propto \sum_{\text{paths}} \exp\left(-\frac{i}{\hbar} S[x(t)]\right). \quad (62b)$$

The propagator gives the amplitude for detection of a particle at a measurement point x' given that it has evolved from point x . The path integral for the propagator appearing in Eq. (62b) arises from taking the limit of its discrete form,

$$K(x, x'; t) = \lim_{P \rightarrow \infty} \int dx_n \prod_{n=1}^{P-1} \langle x | e^{-i\hat{H}\epsilon/\hbar} | x_{P-1} \rangle \times \langle x_{P-1} | e^{-i\hat{H}\epsilon/\hbar} | x_{P-2} \rangle \cdots \langle x_1 | e^{-i\hat{H}\epsilon/\hbar} | x' \rangle, \quad (63)$$

in which the time is divided into P small intervals $\epsilon = t/P$. The sum over positions at each time slice forms a sum over all possible paths. The exponent becomes a time integral of the Lagrangian—the action S for each path as in Eq. (62b). As written, the expression is exact because the Hamiltonian operator commutes with itself. The formalism naturally emits the connection between classical and quantum propagation. Consider the classical “Hamilton's principle of least action” governing propagation between two points (Dirac, 1933). In the classical case only the stationary action paths contribute whereas in the quantum case a family of neighboring low-action (relative to \hbar) paths contributes giving rise to appropriate deviations from classical trajectories. Weighting the paths by the exponent of action gives identical results in solving the Schrödinger equation (Feynman, 1939). The fundamental methodology employed here has been reviewed many times: for excellent reviews see Herman, Bruskin, and Berne (1982), Ceperley (1995), Altland and Simons (2010), and Shankar (2017).

The transition to a path integral representation for the density matrix and partition function involves recognizing that the propagator [Eq. (62a)] and the density matrix [Eq. (61)] are related by analytic continuation. It is clear that a path integral form for the density matrix is available if we associate the temperature dependent factor β [appearing in Eq. (61)] with a time ($\beta = it/\hbar$). This step of replacing inverse temperature ($1/k_B T$) with an imaginary time (it/\hbar) is a *Wick rotation* (rotation by $\pi/2$ in the complex plane) (Baym and Mermin, 1961); see Fig. 12. Obtaining insight into quantum mechanics by invoking analytical continuation in path integral forms remains a topic of research today (Witten, 2010). Note that, unlike the propagator, the canonical density matrix within Boltzmann statistics is positive semidefinite and can be straightforwardly sampled using modern statistical methodology.

Conventional basis set approaches to determining the ground state energy of the quantized oscillator degrees of freedom, as opposed to the path integral approach championed herein, must explicitly include each individual term in the many-body energy, such as the dipole term, as an “excitation” in the basis set (Wang and Jordan, 2001). Such methods therefore become computationally intractable for large systems, scaling as $N!$ in general. Early attempts to include many-body dispersion in the dipole approximation (Cao and Berne, 1992) have $\mathcal{O}(N^3)$ scaling. Our path integral method, in contrast, allows for the inclusion of all many-body contributions (beyond the dipole limit) uniformly and can be solved with linear [$\mathcal{O}(N)$] or nearly linear scaling [$\mathcal{O}(N \times \log(N))$] with system size when solved using Monte Carlo or molecular

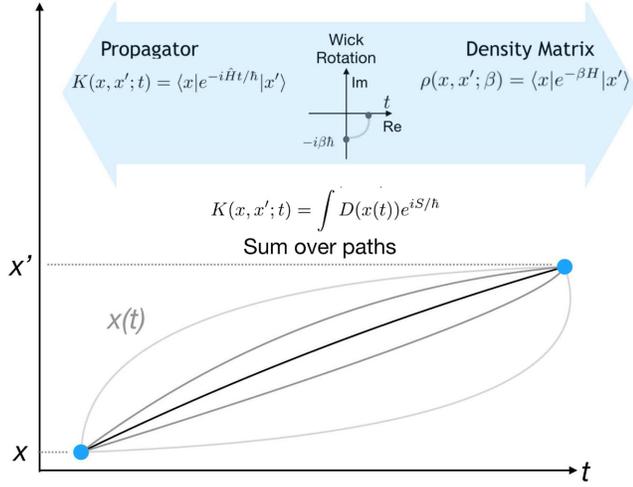


FIG. 12. The canonical density matrix can be obtained from the propagator at time $t = -i\beta\hbar$. The transformation is viewed as a $\pi/2$ Wick rotation in a complex-time plane.

dynamics techniques [most easily applicable to bosons or distinguishable quasiparticles (Boltzmann statistics) as herein].

Since the thermal density matrix has an exponential form, it is factorizable into products of higher temperature expressions. Thus,

$$Z(\beta) = \int dx \rho(x, x; \beta) = \int dx \langle x | e^{-\beta\hat{H}} | x \rangle. \quad (64)$$

Since \hat{H} commutes with itself, we can factorize and obtain

$$Z(\beta) = \int dx \langle x | [e^{-(\beta/P)\hat{H}}]^P | x \rangle \quad (65)$$

for any integer $P \geq 1$. Introducing a complete set of basis vectors $|x_i\rangle\langle x_i|$ in between the P density matrices gives

$$Z(\beta) = \int dx_0 dx_1 \cdots dx_{P-1} \langle x_0 | e^{-(\beta/P)\hat{H}} | x_1 \rangle \times \langle x_1 | e^{-(\beta/P)\hat{H}} | x_2 \rangle \langle x_2 | \cdots \langle x_{P-1} | e^{-(\beta/P)\hat{H}} | x_0 \rangle. \quad (66)$$

In the limit $P \rightarrow \infty$, Eq. (66) is a complete path integral representation of a thermal density matrix. However, even for finite P we can rewrite Eq. (66) as

$$Z(\beta) = \int dx_0 dx_1 \cdots dx_{P-1} e^{\hat{H}_P(\{x_i\}; \beta)}, \quad (67)$$

where

$$\hat{H}_P(\{x_i\}; \beta) = \sum_{i=1}^P \log \rho(x_i, x_{i+1}; \beta/P). \quad (68)$$

The advantage of the discrete path integral form is that increasing the temperature (corresponding to decreasing β) makes the density more “classical” and thus more accurate to approximate. What we essentially do is to start with a quantum mechanical system of N particles in contact with a heat bath at temperature T . Its state in the canonical ensemble is described by a density matrix $\rho(\beta)$, which encodes the state of the system. Equation (67) tells us that the partition function of this quantum system is equal to the partition function of a *classical* system with a Hamiltonian $-\hat{H}_P(\{x_i\}; \beta)/\beta_{\text{faux}}$, with an arbitrary inverse temperature β_{faux} that can be tuned to improve sampling.

3. Approximate density matrices, staging, and dynamics

The next step is to approximate the high temperature density matrices. For ease of notation, let $\tau = \beta/P$. Using a Trotter factorization (Trotter, 1959; Whitfield and Martyna, 2007), the Hamiltonian \hat{H} is split into a reference \hat{H}_0 , with known density matrix $\rho_0 = e^{-\tau\hat{H}_0}$, and a perturbation $V(\vec{x})$ such that

$$\rho(\vec{x}, \vec{x}'; \tau) \approx e^{-\tau V(\vec{x})/2} \rho_0(\vec{x}, \vec{x}'; \tau) e^{-\tau V(\vec{x}')/2} + \mathcal{O}(\tau^3). \quad (69)$$

In the case of QDOs it is convenient to set H_0 as the Hamiltonian of the isolated oscillator, with the following reference density matrix:

$$\begin{aligned} \rho_0(\vec{x}, \vec{x}'; \tau) &= \left[\frac{\alpha_P(\tau)}{\pi} \right] \exp \left[\alpha_P(\tau) (\vec{x} - \vec{x}')^2 + \frac{\lambda_P(\tau)}{2} (\vec{x}^2 + (\vec{x}')^2) \right], \end{aligned} \quad (70)$$

and coefficients $\alpha_P(\tau)$ and $\lambda_P(\tau)$ defined as

$$\begin{aligned} \alpha_P(\beta) &= \frac{m\omega}{2\hbar \sinh(f)}, \\ \lambda_P(\beta) &= \frac{2m\omega \tanh(f/2)}{\hbar}, \\ f &= \frac{\beta\hbar\omega}{P}. \end{aligned} \quad (71)$$

In order to motivate sampling schemes, we notice the strong nearest-neighbor coupling between the coordinates \vec{x} and \vec{x}' . To treat this coupling, the density matrix is diagonalized to the independent coordinates \vec{u}_i via a staging transformation with unit Jacobian (Whitfield and Martyna, 2007), leading to the following form for the partition function:

$$\begin{aligned}
Z(\beta_{\text{D}}) &= \prod_{i=1}^P \int d\vec{x}_i \rho_0(\vec{x}_i, \vec{x}_{i+1}; \tau) \exp\left(-\tau \sum_{i=1}^P V(\vec{x}_i)\right) \\
&= \prod_{i=1}^P \int d\vec{u}_i \underbrace{\left(\frac{1}{2\pi\sigma_i^2}\right)^{3N/2}}_{\text{staging coordinates}} \exp\left(-\frac{\vec{u}_i^2}{2\sigma_i^2}\right) \\
&\quad \times \underbrace{\exp\left(-\tau \sum_{i=1}^P V(\vec{x}_i(\vec{u}_i))\right)}_{\text{external potential}}.
\end{aligned} \tag{72}$$

The coefficients of the transformation are as follows:

$$\begin{aligned}
\vec{u}_1 &= \vec{x}_1, \\
\vec{u}_i &= \vec{x}_i - \vec{x}_i^*, \\
\vec{x}_i^* &= \frac{\sinh(\tau\hbar\omega)}{\sinh[i\tau\hbar\omega]} \vec{x}_1 + \frac{\sinh[(i-1)\tau\hbar\omega]}{\sinh[i\tau\hbar\omega]} \vec{x}_{i+1}, \\
x_{P+1} &= x_1, \\
\sigma_1^2 &= \frac{\hbar}{2m\omega \tanh(\beta\hbar\omega/2)}, \\
\sigma_i^2 &= \frac{\hbar \sinh[(i-1)\tau\hbar\omega] \sinh(\tau\hbar\omega)}{m\omega \sinh[i\tau\hbar\omega]}.
\end{aligned} \tag{73}$$

One step remains in order to make the partition function isomorphic to a fictitious classical system: the addition of conjugate momenta \vec{p}_i with corresponding faux masses \tilde{m}_i . Note that they are different from Drude oscillator mass μ , defined in Eq. (50). This transforms the partition function exactly to yield

$$\begin{aligned}
Z(\beta_{\text{D}}) &= \prod_{i=1}^P \int d\vec{u}_i \left(\frac{1}{2\pi\sigma_i^2}\right)^{3N/2} \exp\left(-\frac{\vec{u}_i^2}{2\sigma_i^2}\right) \\
&\quad \times \exp\left(-\tau \sum_{i=1}^P V(\vec{x}_i(\vec{u}_i))\right) \\
&\quad \times C \int d\vec{p}_i \exp\left(-\tau \frac{\vec{p}_i^2}{2\tilde{m}_i}\right),
\end{aligned} \tag{74}$$

where $C = \prod_i (2\pi\tilde{m}_i/\tau)^{-3/2}$ is a normalization constant required to make the momentum integral unity. This final transformation generates a form for the partition function with the effective classical Hamiltonian $H^{(\text{faux})}$, which can be sampled via existing methods used to sample classical systems (Tuckerman *et al.*, 1993; Whitfield and Martyna, 2006, 2007),

$$H^{(\text{faux})} = \sum_{i=1}^P \frac{\vec{p}_i^2}{2\tilde{m}_i} + \frac{\vec{u}_i^2}{2\sigma_i^2\tau} + \frac{V(\vec{x}_i(\vec{u}_i))}{P}. \tag{75}$$

Note that we now have free parameters: the faux masses \tilde{m}_i . These parameters will not affect the results of the calculations, but can be chosen to improve sampling efficiency. The masses are chosen to equalize the frequencies of the harmonic modes to a single sampling time scale $1/\omega_{\text{S}}$. A choice of $\tilde{m}_i = (\omega_{\text{S}}\sigma_i)^{-2}$ is appropriate (Jones, Crain, Cipcigan *et al.*, 2013). Using this approach, harmonic staging has been found to improve the computational efficiency of path integral sampling by 2–3 orders (Whitfield and Martyna, 2007) and the Trotter discretization numbers $P \approx 10^2$ are sufficient (Jones, Crain, Cipcigan *et al.*, 2013). The final free parameter ω_{S} can be selected equal to the fundamental oscillator frequency, sometimes called the Einstein frequency (Allen and Tildesley, 2017). In systems with the atoms frozen, the choice of ω_{S} simply sets the molecular dynamics time step [e.g., $\delta t \approx 1/(10\omega_{\text{S}})$ being a standard value].

We are primarily interested in coupling the nuclear motion at thermal temperature T to the Born-Oppenheimer surface generated by the path integral sampling of the QDO degrees of freedom at high temperature T_{D} selected as the lowest yet still generates the ground state. We begin by considering the equations of motion generated by independent heat baths coupled to the nuclei and the QDO path integral degrees of freedom. A variety of choices of heat bath are possible but for the path integral degrees of freedom the massive Nosé-Hoover method (Martyna, Klein, and Tuckerman, 1992; Tobias, Martyna, and Klein, 1993; Tuckerman *et al.*, 1993; Martyna, 1994) or variants thereof have thus far been our choice. For nuclei, if sampling of the ensemble and not nuclear dynamics is desired, a multiple thermostat approach is the most efficacious approach there as well. Typically, the heat bath time scale is set to the fundamental frequency of the associated degrees of freedom, here $1/\omega_{\text{N}}$, for nuclei and $1/\omega_{\text{S}}$ for the path integral. Such a set of equations of motion, with two temperatures, cannot be shown to sample any known ensemble except in the adiabatic limit when one set of degrees of freedom is much faster than the other. The result, as has been shown analytically (Cao and Martyna, 1996; Marx, Tuckerman, and Martyna, 1999), agrees with insight, that the slow nuclear degrees of freedom move, at temperature T on the potential of mean force generated by fast degrees temperature at temperature T_{D} as desired.

In practice, adiabatic separation is generated imperfectly through the introduction of a finite adiabaticity parameter γ , which relates the fast path integral frequency ω_S and nuclear characteristic frequency ω_N via $\omega_S = \gamma\omega_N$. Taking the nuclear time step $\Delta t_{\text{MD}} = 0.05/\omega_S$, which is a safe choice (Allen and Tildesley, 2017), and setting the adiabaticity factor $\gamma = 16$ (a compromise between the numerical accuracy and computational costs) (Jones, Crain, Sokhan *et al.*, 2013) defines the method. The rates of convergence for the APIMD-QDO simulated results to their thermodynamic values were studied for water (Jones, Crain, Cipcigan *et al.*, 2013); they define the typical values of the method. Multiple time step integration methods increase the time step and effective parallelization of both the nuclear and electronic degrees of freedom combine to permit the technique to generate high quality results using quite reasonable computer time allotments. As an example, on a modest Intel xeon-based system with 96 cores a system of 4000 noble gas atoms required 0.5 wall seconds per time step of 0.15 fs.

V. APPLICATION EXAMPLES AND PHYSICAL INSIGHTS

A. Noble gas solids and liquids as proof of concept

We have shown that the QDO model is, in principle, effective at reproducing the dispersion interaction behavior between mixed species of atoms (Sec. III.A.2). A significant consequence for materials simulation is that parameters fit for a single species will automatically generate correct long-range interactions when they interact with arbitrary new neighbors. This arises naturally, without the need to obtain distinct parameters for each possible pair, triplet or higher-order term.

The noble gases represent the simplest systems with which to implement the QDO method. While the electronic responses are confined to those arising from the dispersion series, the handling of short-ranged repulsion can be developed and the full machinery of molecular dynamics simulation illustrated. QDO models for noble gases have now been developed (Jones *et al.*, 2009; Jones, Crain, Sokhan *et al.*, 2013) using experimental data for the polarizabilities and dispersion coefficients to determine the three principle model parameters with Coulomb regularization and short-range internuclear pairwise repulsion potential. With no parametrization to the condensed phase, QDOs predict correct zero temperature properties of noble gas crystals (Ne, Kr, Ar, and Xe). Specifically, the crystal lattice constants are in good agreement and the energies are within 3% for all species except neon, which shows 8% deviation. The corresponding liquid phases (using APIMD-QDO) predict radial distribution functions close to experiment (see Fig. 13).

The success of the quantum Drude model in capturing the behavior of solid and liquid noble gases over a broad range of state points demonstrates its general effectiveness. Despite their simplicity they are nonetheless challenging in the sense that lowest-order induction responses present in systems with permanent multipole moments are absent leaving the leading-order dispersion response, a second-order perturbation effect.

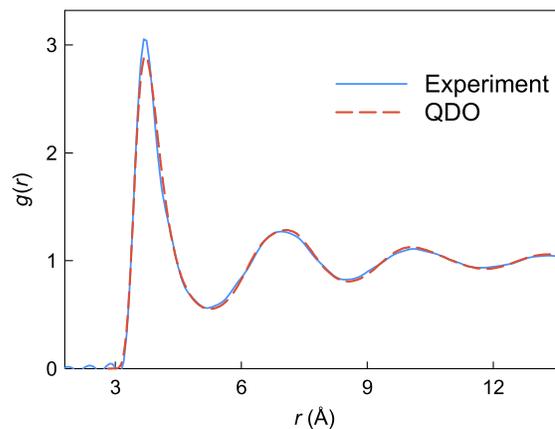


FIG. 13. The radial distribution function for argon at 85 K at the liquid coexisting density calculated using APIMD-QDO (dashed red line), and from the neutron scattering experiments (blue line). From Yarnell *et al.*, 1973.

B. Liquid water

Although water is among the simplest complex liquid. Its condensed phase behavior is essential both for life on Earth and for creating environmentally friendly technologies. However, even though it has only three constituent atoms its dynamics and thermodynamics are challenging to simulate. Hundreds of molecular models of water exist at many levels of physical description from all-electron simulations of small water clusters (Maheshwary *et al.*, 2001; Temelso, Archer, and Shields, 2011; Gillan *et al.*, 2012, 2013; Alfè *et al.*, 2013; Singh, Nandi, and Gadre, 2016) to density functional theory (DFT) simulations of small periodic systems (Schmidt *et al.*, 2009; Lin *et al.*, 2012; Gaiduk, Gygi, and Galli, 2015; Santra *et al.*, 2015; Ambrosio, Miceli, and Pasquarello, 2016; Gillan, Alfè, and Michaelides, 2016) to fixed-charge empirical models (Guillot, 2002; Abascal and Vega, 2005; Vega *et al.*, 2005; Vega and Abascal, 2011; Izadi, Anandakrishnan, and Onufriev, 2014; Cisneros *et al.*, 2016; Izadi and Onufriev, 2016) and coarse-grained simulations where each water molecule is treated as a single particle (Shell, 2008; Molinero and Moore, 2009; Lu *et al.*, 2014; Mashayak *et al.*, 2015; Vega, 2015; Zavadlav *et al.*, 2015; Wagner *et al.*, 2016). These models have given us insight into how water's condensed phase properties emerge from its molecular structure, emphasizing the importance of the competition between the tetrahedral structure favored by hydrogen bonds and the close-packed structure favored by dispersion interactions. The contention between these two effects, with strong many-body underpinnings, one driven by polarization and one driven by dispersion, combined with the lack of a near-quantitative molecular model for such an important substance inspired the creation of QDO water, a molecular model of water.

1. Existing water models

Unlike the case of noble gases, the water molecule is polar and the structure of its liquid form arises from competition between directional hydrogen bonds, which favor an open icelike local structure and more isotropic dispersion forces, which favor a close-packed local structure.

This competition makes modeling of water challenging since many-body polarization and dispersion interactions are key to distinguishing between competing motifs (Morawietz *et al.*, 2016) and have nontrivial many-body origins.

Hydrogen bonds are known to be cooperative, meaning that their interaction strength changes depending on environment thus leading to the emergence of a wide variety of motifs. A simple reporter of this cooperation is the molecular dipole moment, which changes from a value of 1.85 D in the gas phase to an estimated value of 2.5–3.6 D in the condensed phase (Kemp and Gordon, 2008; Jones, Cipcigan *et al.*, 2013), where four hydrogen-bonded motifs are dominant.

Dispersion interactions are also important in water: Including these responses is essential to generate even the basic structure of water at room temperature. An illustrative example of this balance is the overstructuring of room-temperature water by local approximations within DFT (Lin *et al.*, 2012), where including the electron correlations that lead to accurate van der Waals interactions is absent (although, there have been promising results using a similar technique of embedded quantum harmonic oscillators at the dipole level) (Tkatchenko *et al.*, 2012).

In addition, three-body effects account for as much as 25% of the binding energy of water, according to the estimates of Ojamae and Hermansson (1994) calculated at the second-order Moller-Plesset perturbation theory counterpoise-corrected level of theory. Thus, monatomic two-body potentials for the water molecule can only reproduce water's condensed phase properties if their parameters are allowed to vary with state point (Johnson, Head-Gordon, and Louis, 2007; Chaimovich and Shell, 2009). Even a full-atom description of water cannot be transferable if its electrostatics is fixed. Vega and Abascal (2005) showed that a nonpolarizable model of water cannot simultaneously reproduce the melting temperature and temperature of maximum density.

2. *Ab initio* models of water

High level *ab initio* simulations of water are challenged by the computational expense of treating electron correlation accurately. This has limited target systems to small clusters (Maheshwary *et al.*, 2001; Gillan *et al.*, 2012, 2013). Nonetheless, recent advances such as the use of embedded fragments (Gordon *et al.*, 2012) have enabled accurate coupled cluster calculations to be applied to the condensed phase (Willow *et al.*, 2015).

In the case of density functional theory, local, gradient corrected and hybrid functionals both overstructure liquid water (Lin *et al.*, 2012), with a melting point of around 400 K (Yoo, Zeng, and Xantheas, 2009). To achieve a more realistic structure dispersion interactions have to be included, using techniques such as van der Waals corrected functionals (Williams and Malhotra, 2006; Klimeš, Bowler, and Michaelides, 2009), the Tkatchenko-Scheffler embedded oscillators method (Santra *et al.*, 2008; Tkatchenko and Scheffler, 2009; Tkatchenko *et al.*, 2012), or hybrid functionals (Benighaus *et al.*, 2008).

Thus, despite the challenge, low level *ab initio* computations have revealed interesting insight into the physics of water. Cluster-based simulations are important in creating accurate benchmarks for the parametrization of empirical models

(Maheshwary *et al.*, 2001) and could even be used to interpret ultrafast experiments using nanodroplets consisting of a few thousand water molecules. Lower level DFT simulations are essential in interpreting experimental results of x-ray spectroscopy, which probes the change in electron ionization energies as a function of different environments (Nilsson, Schlesinger, and Pettersson, 2015; Amann-Winkel *et al.*, 2016).

Ab initio simulations also reveal insight about the nature of the hydrogen bond. For example, DFT simulations show, via an energy partition scheme, that water molecules have on average one strong acceptor and one strong donor bond, uncovering a fluctuating asymmetry in water's local structure (Kühne and Khaliullin, 2013).

3. Fixed-charge models

The majority of simulations of water models have used an approach where water molecules are modeled as a combination of three elements: fixed point charges arranged to reproduce the electrostatic moments of the molecule, Lennard-Jones R^{-6} interactions to reproduce dispersion, and an R^{-12} or exponential repulsion. This approach is conceptually simple, computationally cheap, and has reproduced the bulk properties of liquid water with good accuracy.

Such models can be characterized by the number of interaction sites, ranging from three to six, as summarized in Fig. 14. A few years after resolving the molecular structure of water, Bernal and Fowler, (1933) proposed the first molecular model of water. It was a four-site model, resembling the now widely used transferable intermolecular potential with four points (TIP4P) family, with two positive charges placed on the hydrogen atoms and a negative charge on the bisector of the HOH angle.

Computer simulation of water began in the late 1960s to the early 1970s, with Barker and Watts (1969) and Rahman (1971) performing the first Monte Carlo and molecular dynamics simulations of liquid water. The 1980s saw the creation and parametrization of an increasing number of water models. Some followed the philosophy of Bernal and Fowler (1933) and employed four interaction sites, leading to the TIP4P family of models (Jorgensen *et al.*, 1983). Others used only three sites, leading to the transferable intermolecular potential with three points (TIP3P) (Jorgensen *et al.*, 1983) and simple point charge (Berendsen, Grigera, and Straatsma, 1987) families. Others created a model incorporating the lone pair orbitals, resulting in five interaction sites and the Stillinger second model (ST2) family of models (Stillinger and Rahman, 1974).

The 2000s saw an increase in optimized parametrizations of these families of models, seeing the birth of models such as TIP4P/Ew (Horn *et al.*, 2004), optimized for use with Ewald summation techniques, TIP4P/Ice (Abascal *et al.*, 2005), optimized to reproduce the phases of ice and TIP4P/2005 (Abascal and Vega, 2005), optimized to give the best overall fit to the whole phase diagram of water.

All of these models assume that the interaction sites need to be placed on atomic sites. However, this choice is not optimum for reproducing the ground state electrostatic properties of real water. Recently, Izadi, Anandkrishnan, and Onufriev (2014) showed that relaxing this assumption leads to a better reproduction of the electrostatic moments of the water molecule and to two models, OPC (optimized point

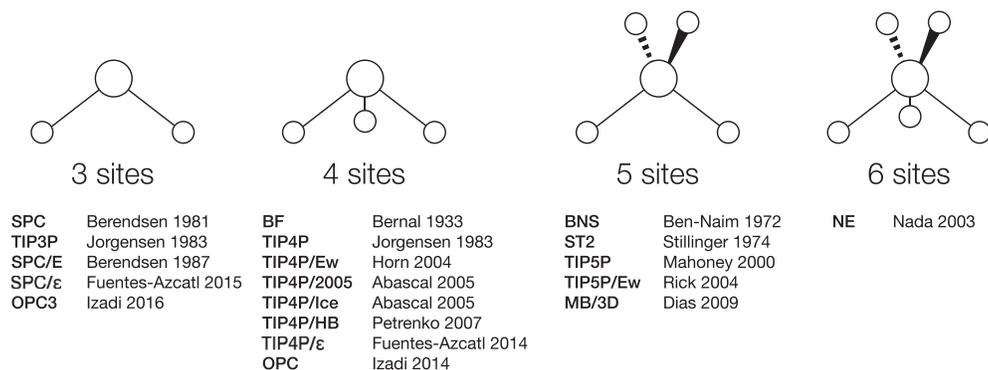


FIG. 14. Various parametrizations of classical, rigid, fixed-charge models using three sites (Berendsen *et al.*, 1981; Jorgensen *et al.*, 1983; Berendsen, Grigera, and Straatsma, 1987; Fuentes-Azcatl, Mendoza, and Alejandre, 2015; Izadi and Onufriev, 2016), four sites (Bernal and Fowler, 1933; Jorgensen *et al.*, 1983; Horn *et al.*, 2004; Abascal *et al.*, 2005; Abascal and Vega, 2005; Fuentes-Azcatl and Alejandre, 2014; Izadi, Anandakrishnan, and Onufriev, 2014), five sites (Ben-Naim, 1972; Stillinger and Rahman, 1974; Mahoney and Jorgensen, 2000; Rick, 2004; Dias *et al.*, 2009), and six sites (Nada and Eerden, 2003). From Molinero, 2013.

charge, 4 site) and OPC3 (optimized point charge, 3 site) that achieve a better fit to liquid water's condensed phase properties than the TIP4P and TIP3P families.

Models such as OPC have reached the limits of what is achievable with fixed-charge models. In order to go farther in reproducing both the properties of bulk water and those of water in heterogeneous environments, the effects of many-body polarization and dispersion have to be taken into consideration.

4. A new molecular model of water based on quantum Drude oscillators

Quantum Drude oscillators unite the missing elements from fixed-charge models into a single framework. They closely reproduce the electrostatic responses including polarization beyond the dipole approximation and many-body dispersion interactions between molecules at a computational cost that scales near linearly with the size of the system. Thus, QDOs can provide the missing building block to water models that brings their properties in near-quantitative agreement to those of real water.

To construct a water model using a QDO, one needs three elements: a rigid molecular frame with embedded point charges to replicate the lowest-order electrostatic moments, a QDO to replicate electronic responses, and a short-range, pairwise repulsion potential to treat Pauli repulsion. These elements are shown in Fig. 15 with their respective parameters given in Table II.

The molecular frame is fixed in the experimental geometry of the isolated molecule, with an HOH angle of 104.52° and an OH distance of 0.9572 \AA (Abascal and Vega, 2005). Two positive charges of magnitude $q_H = +0.605e$ are placed on the H site and a negative $-2q_H$ charge is placed on the M site, on the bisector of the HOH angle at a distance $R_{OM} = 0.2667 \text{ \AA}$ from the oxygen. The distances and the charges are fixed to generate the dipole moment of the isolated molecule (1.85 D) and give the best fit to the quadrupole moment components.

To create an electronically responsive model, a QDO is centered on the M site. Its parameters are determined using the dipole and quadrupole polarizabilities and the C_6 dispersion

coefficient using Eq. (50). The reference values of the multipole moments and dispersion coefficients are from Millot *et al.* (1998), who reported a mix of *ab initio* calculations and experiment.

The repulsive potential is fit by calculating one cut through the dimer energy surface using the coordinates shown in Fig. 16. *Ab initio* calculations were done at the CCSD(T) level, with the aug-cc-pVTZ basis set using ACESIII 3.0.7 (Lotrich *et al.*, 2008) giving the energy shown in Fig. 16. The equivalent energy of the repulsion-free model (frame + QDO) was calculated using the norm-conserving diffusion Monte Carlo method (Jones *et al.*, 2009) with 1000 walkers. The differences between *ab initio* energies and repulsion-free models were fit to a double exponential of the form given in Eq. (54).

The Coulomb potential between charges was damped by replacing point charges with a Gaussian distribution of width σ_i as per Eq. (55).

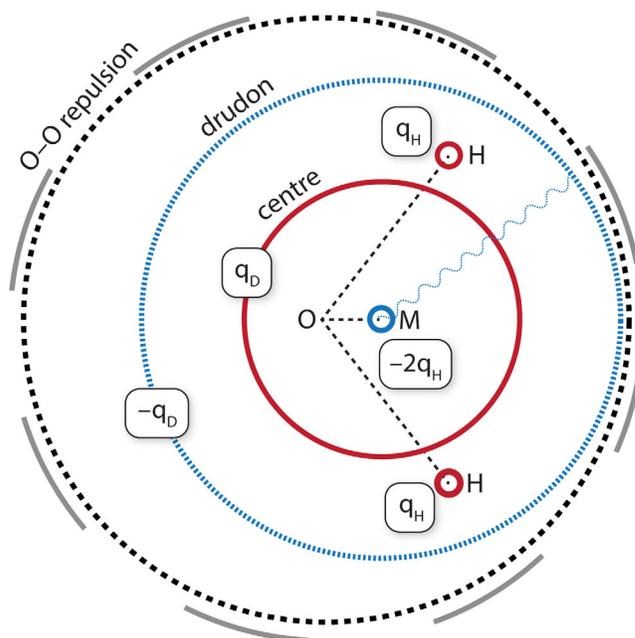


FIG. 15. QDO-water model with spherical regularizing Gaussian surfaces shown in proportion.

TABLE II. The free parameters of QDO water. $E_h \approx 27.211$ eV is the Hartree energy, $a_0 \approx 0.5292$ Å is the Bohr radius, and $e \approx 1.60 \times 10^{-19}$ C is the elementary charge.

Parameter	Value	Parameter	Value
Molecular geometry		Coulomb damping	
R_{OH}	0.9572 Å	$\sigma_D = \sigma_H = \sigma_M$	$0.1a_0$
$\angle HOH$	104.52°	σ_C	$1.2a_0$
Ground state electrostatics		Short-range repulsion	
q_H	$0.605e$	κ_1	$613.3E_h$
R_{OM}	0.2667 Å	λ_1	$2.3244a_0^{-1}$
Quantum Drude oscillator		κ_2	$10.5693E_h$
m_D	$0.3656m_e$	λ_2	$1.5145a_0^{-1}$
ω_D	$0.6287E_h/\hbar$		
q_D	$1.1973e$		

The Gaussian widths of each charge are given in Table II. For the bead (the path integral molecular dynamics sampler), hydrogen and the M site are 0.1 bohr, small in comparison to the dimensions of the molecule. For the center charge of the QDO, the Gaussian width is 1.2 bohr, which is comparable with the size of the molecule. This choice was made so that the center charge provides a background of positive charge inside the molecule, making the ground state of the QDO more uniformly neutral and reproducing some of the screening.

5. Cluster energies of QDO water: A basic benchmark

In order to assess how the energetics of QDO water extrapolate beyond the single cut through the dipole energy surface, a simple quantity to calculate is the energies of various water clusters. The reference energies and geometries were taken from the Benchmark Energy and Geometry DataBase (Řezáč *et al.*, 2008) and were calculated at the CCSD(T)/CBS noCP level of theory by Temelso, Archer, and Shields (2011), which is a comparable level of theory to what was used to parametrize the model.

Since the reference geometries were optimized via *ab initio* calculations (at the MP2/aug-cc-pVDZ noCP level of theory),

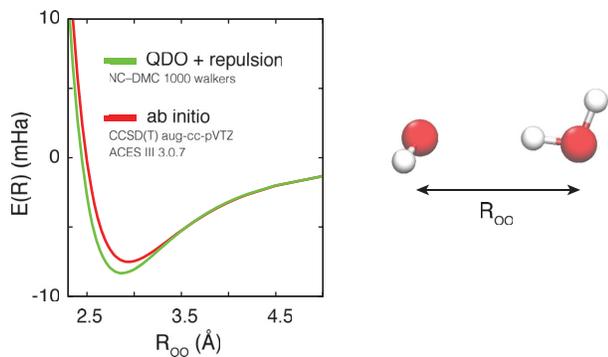


FIG. 16. (Left) A cut through the dimer energy surface calculated via *ab initio* methods and QDO with added repulsion. The figure is adapted from Sokhan *et al.* (2015b) and corrects the error there by showing the distance between the oxygen atoms in both cases. (Right) A sketch of the orientation used in calculations with the coordinate defining the abscissa in the dimer energy graph.

they had to be adapted to fit the rigid frame of QDO water. To do so, the oxygen atom of QDO water was placed in the position of the oxygen atom in the reference geometry. The bisector of the model's frame was then aligned with the bisector of the reference geometry and the positions of the hydrogen atoms were then generated. The model's hydrogen atoms were close to the reference ones, with the OH distances changing by a few percent. This is consistent with a fairly rigid OH bond. The rearrangement had to be done since the energetics of QDO water were parametrized with reference to a fixed geometry. The *ab initio* energies were not recomputed and hence quantitative agreement cannot be expected.

Figures 25–28 in Appendix B show a comparison of the binding energies per molecule for 38 clusters between QDO water and *ab initio* calculations. The agreement between QDO water and *ab initio* is excellent. The energies differ by a few percent, with 68% of the QDO-water clusters having energies 2% away from the *ab initio* reference and almost all (97%) with an energy 5% away from the *ab initio* reference. The only large deviation was the 3UUU trimer (three molecules in a triangle, with all free hydrogen atoms oriented in the same direction), with a deviation of 9.3%. It would be interesting to add the CP correction to the *ab initio* results and estimate the contribution of three-body repulsion, which is neglected in the QDO model.

6. Transferability of QDO water

The behavior of QDO water has been investigated in multiple environments, illustrated in Fig. 17, spanning the phase diagram of water. In a series of papers, it was shown to correctly predict the following properties:

Sokhan *et al.* (2015b): The lattice constants and bulk modulus of ice II, the liquid-vapor coexistence densities above 300 K, the temperature of maximum density, and the radial distribution function at 300 K.

Cipcigan *et al.* (2015): Surface tension, a positive surface expansion, orientation of water molecules at the liquid-vapor interface, and the radial distribution function at the liquid-vapor interface.

Sokhan *et al.* (2015a): Pressures of supercritical water at 673, 773, and 873 K.

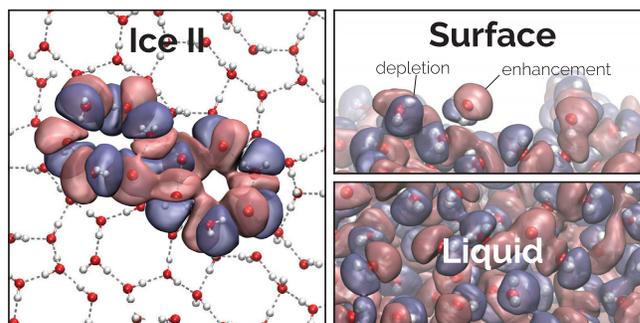


FIG. 17. Examples of quantum Drude oscillators in different environments: ice II, ambient temperature liquid water, and the surface of liquid water. The images illustrate their electronic responses, with red and blue isosurfaces corresponding to regions of enhancement and depletion of electronic density, respectively.

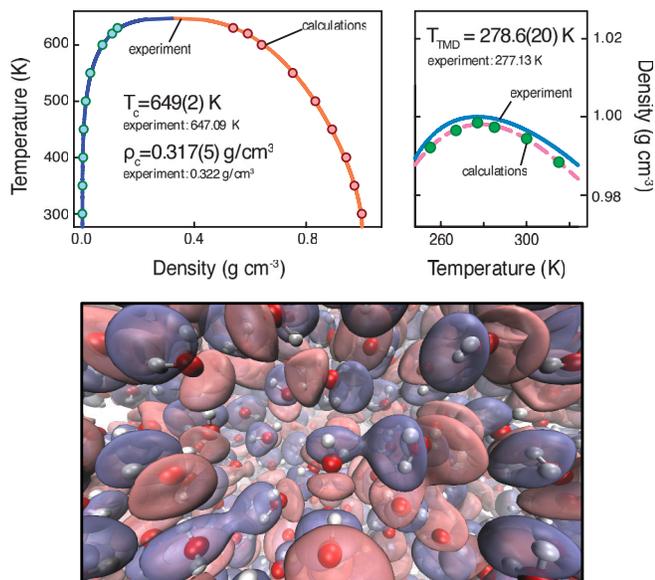


FIG. 18. The liquid-vapor coexistence curve, temperature of maximum density, and the polarization density of water molecules in the liquid phase. Blue regions represent depletion of electrons while red regions represent enhancement.

Cipcigan *et al.* (2018): Pressure of supercooled water at 51.8 and 55.3 mol/l above 220 K, and a radial distribution function of supercooled water.

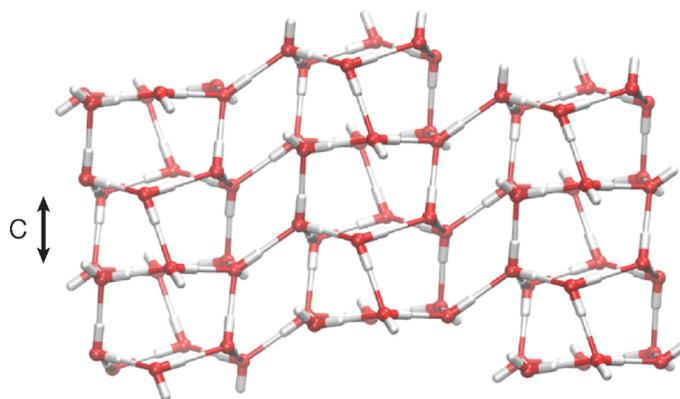
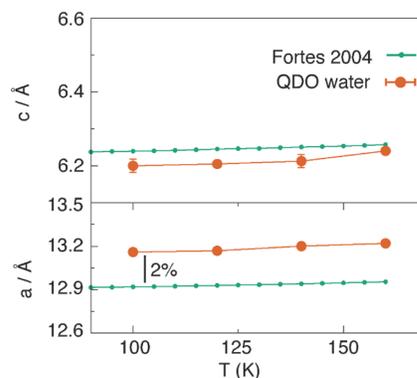
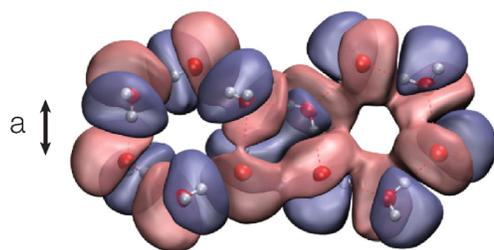


FIG. 19. The unit cell of ice II and predicted unit cell parameters compared to Fortes *et al.* (2005). The image in the top-left corner shows the polarization density of QDO water on the two rings of the ice II unit cell. Blue regions represent depletion of electrons while red regions represent enhancement.

7. Liquid-vapor coexistence

Fundamentally, the liquid-vapor interface is among the simplest heterogeneous systems water can form in nature. At this interface, the hydrogen-bonded network that binds the liquid is abruptly truncated in a manner that is still of debate for both simulation and experiment. The prediction of an accurate liquid-vapor critical point, a temperature of maximum density, and coexistence densities is an important benchmark for a water model. This remains challenging due to the change in molecular properties between the gas phase and the liquid phase water, with the dipole moment enhancing from 1.85 to around 2.6 D.

Sokhan *et al.* (2015b) reported a realistic enhancement of the dipole and higher-order moments between the gas and the liquid phases, leading to an accurate prediction of the coexistence properties (summarized in Fig. 18). The complex polarization produced by QDO water can be visualized by rendering the difference in density between the charge distribution of a QDO and its ground state. Figure 18 shows such a visualization for bulk liquid water.

Cipcigan *et al.* (2015) reported an accurate surface tension and novel insight into the molecular structure of the interface, linking asymmetries in hydrogen bonding to the measured surface orientation in surface-specific sum-frequency generation spectroscopy.

8. Ice II: A proton ordered ice

Ice II is a proton ordered ice with the structure illustrated in Fig. 19. Its unit cell consists of two hexagonal rings with

distinct hydrogen bonding motifs (Kamb, 1964). These rings are staggered and connected by hydrogen bonds both vertically and horizontally, resulting in a structure resembling nanotubes of water molecules. The structure is determined by two parameters: a is proportional to the size of each ring and c is proportional to the separation between the rings.

Sokhan *et al.* (2015b) reported the lattice parameters of ice II predicted by QDO, as a function of temperature, compared to the experimental references of Fortes *et al.* (2005) (presented in Fig. 19). These show a good match: c is less than 1% above experiment while a is 2% lower than experiment. Taking into consideration that the experimental literature shows a similar spread in values, this agreement is good.

9. Supercritical water

When water is heated above its critical temperature of 647 K, the distinction between liquid and vapor vanishes, resulting in a single, *supercritical* phase. The existence of this phase has been known almost for two centuries (Cagniard de la Tour, 1822) and has excited both fundamental and technological interest. At the molecular level, supercritical water contains small transient water clusters embedded in a gaslike phase. Locally, it maintains the structure of liquid water yet globally it expands to fill a space like a gas would.

Despite the lack of a phase transition, supercritical water contains regions of liquidlike and gaslike properties (Fisher, 1969; Brazhkin *et al.*, 2012; Fomin *et al.*, 2015). These have been separated using various boundaries in the phase diagram: the critical isochore, the maxima in thermodynamic responses (also known as the Widom lines) (Fisher, 1969) and a dynamic crossover (known as the Frenkel line) (Brazhkin *et al.*, 2012). These boundaries coincide close to the critical point (in the scaling region) and diverge at higher pressures and temperatures.

Sokhan *et al.* (2015a) reported a new such crossover between gaslike and liquidlike regions in supercooled water, marked by a cusp in the molecular dipole moment. This is the first report of such a *molecular signature* of the gas-liquid transition in supercritical water. This study extends the transferability of QDO water to supercritical water, where its predictions again match experiment, as reproduced in Fig. 20. Note that percolation transition in the hydrogen bond network and the Widom line in general do not coincide (Strong, Shi, and Skinner, 2018).

10. Supercooled water

A supercooled liquid is formed when the liquid is cooled below its freezing temperature. In this region, the crystal is the thermodynamically stable phase. However, without a nucleation site, crystallization faces a large energy barrier since it requires a global rearrangement from a low symmetry to a high symmetry phase. The energy barrier makes supercooled water metastable on a time scale influenced by a competition of two effects: the relaxation times of liquid water, whose rate of increase with cooling increases below the freezing point, and the similarity of the local environment to the stable phase (Limmer and Chandler, 2015).

Upon supercooling, the response functions of water such as heat capacity and isothermal compressibility increase rapidly

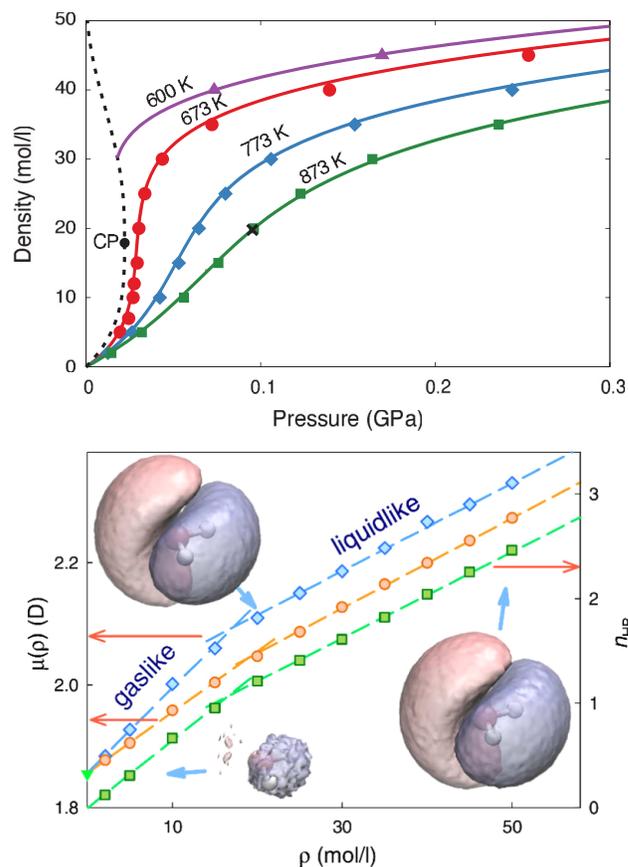


FIG. 20. (Top) Isotherms of supercritical QDO water at 600 (triangles), 673 (circles), 773 (diamonds), and 873 K (squares). The points are calculated values while the lines are NIST/IAPWS-95 experimental data (Wagner and Pruß, 2002; Lemmon, McLinden, and Friend, 2018). The dotted line is the liquid-vapor coexistence curve of water (Wagner and Pruß, 2002; Lemmon, McLinden, and Friend, 2018) ending in the critical point, labeled CP. The errors are calculated using bootstrapping (Allen and Tildesley, 2017) and are smaller than the size of the points. (Bottom) Dipole moment of water molecule as a function of water density at $T = 673$ K (blue diamonds) and 873 K (orange circles). Green squares and right scale, average number of hydrogen bonds per molecule at $T = 673$ K. The lines are linear fits to the data. Insets illustrate the corresponding variation in electron density—pink and blue isosurfaces correspond to gain and loss of electron density, correspondingly.

and appear to diverge just below the temperature of homogeneous nucleation (DeBenedetti, 2003). This is one of water's *anomalies* whereby below 4°C cooling causes expansion.

Cipcigan *et al.* (2018) reported a molecular-level anomaly driving this expansion. Below the temperature of maximum density, the first coordination shell of water molecules (i.e., the first four neighbors) still contracts upon cooling, even if the liquid expands. To counteract, the second coordination shell (the next eight molecules) expands (Fig. 21). Further, Cipcigan *et al.* (2018) compared predictions of QDO water in the supercooled region, showing again a good match with experiment.

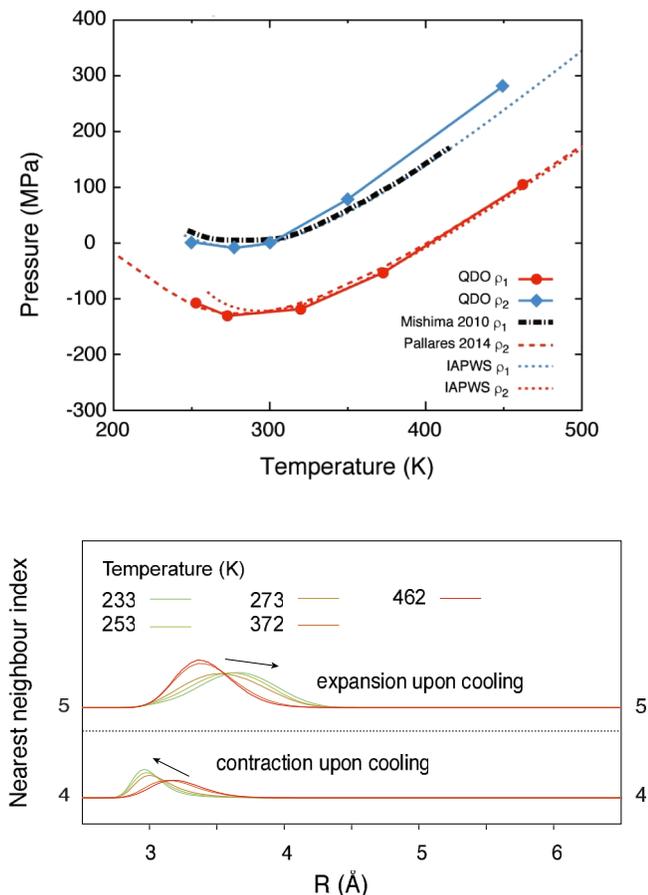


FIG. 21. (Top) The pressure of 300 QDO-water molecules as a function of temperature at two constant volumes: $\rho_1 = 51.8002$ mol/l (red) and $\rho_2 = 55.3173$ mol/l (blue). The dashed line represents equivalent data digitized from Pallares *et al.* (2014) (ρ_1 , experimental estimates via simulations of TIP4P/2005). The dotted lines represent the corresponding isotherms of the IAPWS-95 reference equation of state for water (Wagner and Pr u , 2002). The black dot-dashed line corresponds to an interpolated isochore based on the experimental data measured by Mishima (2010). (Bottom) Probability distribution of each neighbor (arbitrary units) as a function of temperature, showing the range of distances from a central molecule the neighbor occupies.

VI. FUTURE PERSPECTIVES

A. Toward complex physical systems

The electronically coarse-grained strategy described here points to a promising and versatile new approach to the simulation of nonreactive condensed matter systems. The results validate the basic premise that a complete model, sans *a priori* perturbative cutoffs selected by the modeler, allows complex systems to be described and their emergent properties to be gleaned.

While the systems studied so far have been limited to water and noble gases these basic ingredients lay the foundation for a rapid expansion in the space of applications on the path toward the simulation of biomolecules within the overall framework. The case of water is particularly significant as its physics includes the coupled effects of polarity, polarizability, dispersion, and network formation via hydrogen bonds.

Already within reach, as the next steps, are electronically coarse-grained models for hydrophobic hydration under ambient and extreme conditions. In the simplest cases of hydrated noble gases it will be possible to make direct structural comparisons between model predictions and available x-ray or neutron diffraction experiments. It then becomes possible to examine largely unexplored questions about the effects of electronic responses on hydration shell structure and self-assembly. Conclusions drawn here would likely have wider significance and guide thinking regarding the fundamentals of hydrophobic hydration at the small length scale limit.

Beyond single component systems, but using the same basic building blocks, it is possible to imagine assembling simple extended interfaces between water and nonpolar surfaces which may include apertures and other forms of nanoconfinement. Such constructs are rudimentary models for hydrophobic interactions on large length scales and for complex biological systems. Phenomena such as “hydrophobic gating” (Aryal, Sansom, and Tucker, 2015) and liquid-vapor oscillations in pores including the effects of long-range electronic responses become accessible. There are early indications of the importance of many-body electronic effects at the interface between water and graphitic surfaces and that polarization and dispersion interactions contribute differently to wetting phenomena (Misra and Blankschtein, 2017).

ACKNOWLEDGMENTS

This work was supported by the STFC Hartree Centre’s Innovation Return on Research program, funded by the Department for Business, Energy and Industrial Strategy. Pimpernel Science, Software and Information Technology provided partial support for this work. The views expressed here are those of the authors and do not reflect Pimpernel policy.

APPENDIX A: MULTIPOLE EXPANSION

1. Approximate potentials at long range: Cartesian multipole expansion for a charge distribution

Here we set out the principles and key steps leading to the multipole expansion for the potential of a charge distribution. In general, a multipole expansion is a description of the potential which generally converges to the exact potential when the charges are spatially confined and the observation point is far from the sources.

We start with the fundamental definition of the electrostatic potential $V(\mathbf{R})$ arising from an arbitrary charge distribution $\rho(\mathbf{r})$ [see Fig. 22 and Eq. (4)] as

$$V(\mathbf{R}) = \int d\mathbf{r} \rho(\mathbf{r}) T(\mathbf{R} - \mathbf{r}). \quad (\text{A1})$$

This is recognizable as the general solution for the Poisson equation where $T(\mathbf{R} - \mathbf{r}) = |\mathbf{R} - \mathbf{r}|^{-1}$ is the Green’s function (Tang, 2006). Since the Poisson equation is linear in both the potential and the source term its solutions are completely superposable.

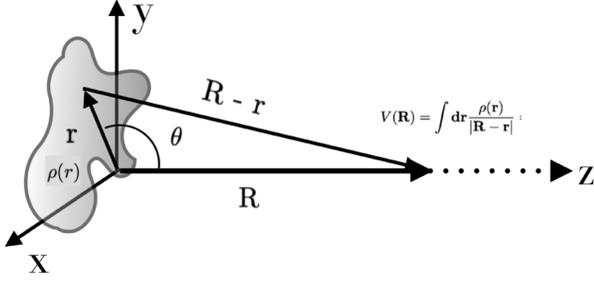


FIG. 22. An arbitrary charge distribution used to define the multipole expansion of the electric potential $V(\mathbf{R})$ for $R \gg r$.

Assuming we are in the far field (exterior multipole) range where $|\mathbf{r}| \ll |\mathbf{R}|$, $T(\mathbf{R} - \mathbf{r})$ can be expanded in a Taylor series about the origin:

$$T(\mathbf{R} - \mathbf{r}) = \frac{1}{R} - r_\alpha \nabla_\alpha \left(\frac{1}{R} \right) + \frac{1}{2} r_\alpha r_\beta \nabla_\alpha \nabla_\beta \left(\frac{1}{R} \right) - \frac{1}{3!} r_\alpha r_\beta r_\gamma \nabla_\alpha \nabla_\beta \nabla_\gamma \left(\frac{1}{R} \right) + \dots \quad (\text{A2})$$

From here the electrostatic potential can be expressed [following its definition in Eq. (A1)] in terms of the expansion

$$V(\mathbf{R}) = T(\mathbf{R}) \int \rho(\mathbf{r}) d\mathbf{r} - T_\alpha(\mathbf{R}) \int r_\alpha \rho(\mathbf{r}) d\mathbf{r} + \frac{1}{2} T_{\alpha\beta}(\mathbf{R}) \int r_\alpha r_\beta \rho(\mathbf{r}) d\mathbf{r} - \frac{1}{6} T_{\alpha\beta\gamma}(\mathbf{R}) \int r_\alpha r_\beta r_\gamma \rho(\mathbf{r}) d\mathbf{r} + \dots \quad (\text{A3})$$

The various terms here define the multipole moments as follows: The first term $\int \rho(\mathbf{r}) d\mathbf{r} = q$ defines the overall charge (monopole) within the distribution. The second term includes the general definition of the dipole arising from a continuous distribution as $\int r_\alpha \rho(\mathbf{r}) d\mathbf{r} = \mu_\alpha$. The third term defines the quadrupole (symmetric with nonzero trace) and so on. Together these terms define the multipole expansion of the potential as

$$V(\mathbf{R}) = qT(\mathbf{R}) - \mu_\alpha T_\alpha(\mathbf{R}) + \frac{1}{2} Q_{\alpha\beta} T_{\alpha\beta}(\mathbf{R}) + \dots, \quad (\text{A4})$$

where we define $T(\mathbf{R}) = R^{-1}$, $T_\alpha(\mathbf{R}) = \nabla_\alpha R^{-1}$, $T_{\alpha\beta}(\mathbf{R}) = \nabla_\alpha \nabla_\beta (R^{-1})$, and so on. The first term is the electrostatic potential arising from a point charge. Note that the multipoles are properties of the charge distribution while $T(\mathbf{R})$ and its gradients determine the distance dependences of the contribution from each moment (properties of the observation point \mathbf{R}).

There is one subtlety concerning the choice of origin which can be illustrated by returning to the case of a dipole configuration formed by two charges (q_1 and q_2) positioned on the x axis at x_1 and x_2 , respectively. If we imagine translating

the origin to the left along the x direction by a , we would compute a dipole moment in the translated coordinate system $\mu_T = q_1(x_1 + a) + q_2(x_2 + a) = q_1 x_1 + q_2 x_2 + a(q_1 + q_2)$. Thus,

$$\mu_T = \mu + a(q_1 + q_2). \quad (\text{A5})$$

Therefore, only in the case where $q_1 = -q_2$ (system has no net charge) do we find that the dipole moment in the original and translated coordinate systems are the same. In the case where the two charges do not exactly cancel (i.e., the lowest-order multipole is the monopole) the dipole moment will not be invariant under translations. This is a special case of a general result that only the lowest-order moment in the multipole expansion is unique and independent of origin choice.

As a final comment in this section we note that while the charge and dipole have familiar forms, there are several conventions for the quadrupole and higher multipoles. Most common among these alternatives is a traceless form of the tensor which reduces further the number of independent components:

$$\Theta_{\alpha\beta} = \frac{1}{2} \int d\mathbf{r} (3r_\alpha r_\beta - r^2 \delta_{\alpha\beta}) \rho(\mathbf{r}). \quad (\text{A6})$$

The freedom to choose such an alternative form arises because the quadrupole potential is unaltered by the addition of a term which satisfies the Laplace equation.

The trace-canceling term is interpreted as the spherically averaged second moment. Similar traceless constructions exist for higher moments.

2. Energy of a charge distribution in an external field

We next consider the general case where the charge distribution is placed in an external potential $\phi(\mathbf{r})$. We will see that the conclusions drawn here apply to the situation where the external field arises from a neighboring molecule and can be used to determine the interaction energy between arbitrary charge distributions. We begin by performing a similar expansion of the potential about the origin (chosen to be inside the charge distribution),

$$\phi(\mathbf{r}) = \phi(0) + r_\alpha \phi_\alpha(0) + \frac{1}{2} r_\alpha r_\beta \phi_{\alpha\beta}(0) + \frac{1}{3!} r_\alpha r_\beta r_\gamma \phi_{\alpha\beta\gamma}(0) + \dots, \quad (\text{A7})$$

where $\phi_\alpha(0) = \partial\phi(0)/\partial r_\alpha$, from which the electrostatic interaction energy follows as

$$E_{\text{elec}} = \int d\mathbf{r} \rho(\mathbf{r}) \phi(\mathbf{r}) = \int d\mathbf{r} \rho(\mathbf{r}) \phi(0) + \int d\mathbf{r} r_\alpha \rho(\mathbf{r}) \phi_\alpha(0) + \frac{1}{2} \int d\mathbf{r} r_\alpha r_\beta \rho(\mathbf{r}) \phi_{\alpha\beta}(0) + \frac{1}{6} \int d\mathbf{r} r_\alpha r_\beta r_\gamma \rho(\mathbf{r}) \phi_{\alpha\beta\gamma}(0) + \dots \quad (\text{A8})$$

Using the definition of the multipole moments we obtain

$$E_{\text{elec}} = q\phi + \mu_\alpha \phi_\alpha + \frac{1}{3}\Theta_{\alpha\beta} \phi_{\alpha\beta} + \frac{1}{15}\Omega_{\alpha\beta\gamma} \phi_{\alpha\beta\gamma} \cdots \quad (\text{A9})$$

The prefactors of 1/3 and 1/15 arise from using the traceless forms of the quadrupole and octopole tensors. Finally we can now make contact with the expression in Eq. (5) by introducing the electric field components as gradients of the potential according to $F_\alpha = -\nabla_\alpha \phi(\mathbf{R})$,

$$E_{\text{elec}} = q\phi - \mu_\alpha F_\alpha - \frac{1}{3}\Theta_{\alpha\beta} F_{\alpha\beta} - \frac{1}{15}\Omega_{\alpha\beta\gamma} F_{\alpha\beta\gamma} + \cdots, \quad (\text{A10})$$

which is equivalent to Eq. (5) with the exception that this expression includes the monopole source term $q\phi$.

The resulting energy is then the scalar product of the field with dipole, field gradient with quadrupole, and higher field gradients with the higher multipole tensors.

3. Interactions between general multipoles

The multipole expansion for the potential of an arbitrary charge distribution and the expression for the interaction energy given in Eqs. (A4) and (A10), respectively, allow us to determine forms for the interaction energy between multipoles of arbitrary order. To illustrate we consider first the trivial case of two point charges. The interaction energy is given by the first term of Eq. (A10). Here ϕ corresponds to the potential of another point charge as given by the first term of Eq. (A4). Taken together these produce the elementary result that the interaction potential is $q_1 q_2 / R_{12}$ the gradient of which gives Coulomb's force law, etc.

Interaction tensors for higher multipoles emerge systematically. In the case of two interacting dipoles we refer to the second term of Eq. (A10). Here the relevant energy is that of the dipole moment in the field of another dipole. We then need to obtain the field by taking the gradient of the corresponding dipole potential [the second term in Eq. (A4)]. This procedure gives the components of the dipole field tensor \mathbf{T}_{ij} defined as

$$\mathbf{T}_{ij} = \nabla_i \nabla_j \frac{1}{|R|} = \frac{3R_i R_j}{R^5} - \frac{\delta_{ij}}{R^3}. \quad (\text{A11})$$

The dipole-dipole energy for a fixed orientation is then expressed in a compact tensor notation as

$$E_{\text{dip}} = \mu_A \mathbf{T}^{AB} \mu_B. \quad (\text{A12})$$

The interaction energy between multipoles of arbitrary order can be similarly constructed with the basic trend being that higher multipoles give rise to shorter ranged interactions. The form of the interaction tensors also means that some multipole pairs have similar functional forms and distance dependences. For example, it is evident from Eqs. (A4) and (A10) that the form of the dipole-dipole interaction tensor just considered is the same as that for a quadrupole in the field of a point charge. The dipole-quadrupole interaction tensor is also the same as that for a charge-octopole pair, etc.

4. Spherical multipoles

We have seen that the potential of a distribution can be expanded in powers of $1/R$ in the far field as a Taylor series in the Cartesian coordinates. An alternative expansion can be performed in terms of spherical harmonics which depend on spherical polar coordinates; see Fig. 23.

The spherical multipole expansion includes terms which capture progressively finer angular features: By analogy with the Cartesian case, we expect the first term to be a monopole—a constant, independent of angle. The next term should have dipolar symmetry—varying once in angle from positive to negative around the sphere. Higher-order terms (such as the quadrupole and the octopole) exhibit greater angular modulation.

To develop these ideas rigorously we express the denominator of $T(\mathbf{R} - \mathbf{r})$ as

$$\sqrt{(\mathbf{R} - \mathbf{r}) \cdot (\mathbf{R} - \mathbf{r})} = [R^2 + r^2 - 2Rr \cos(\theta)]^{1/2}.$$

Here θ is the angle between the position vector of the charge element and the observation direction (see Fig. 22). In the far field case where $r < R$ this can be arranged as

$$\begin{aligned} T(\mathbf{R} - \mathbf{r}) &= \frac{1}{R} \left[1 + \left(\frac{r}{R}\right)^2 - 2\left(\frac{r}{R}\right) \cos(\theta) \right]^{-1/2} \\ &= \frac{1}{R} \sum_{l=0}^{\infty} P_l(\cos \theta) \left(\frac{r}{R}\right)^l, \end{aligned} \quad (\text{A13})$$

where $P_l(\cos \theta)$ is the Legendre polynomial and the expression in brackets is recognized as the generating function. If axes are chosen such that the observation point is located on the z axis, the angle θ corresponds to the normal polar angle in spherical coordinates.

If the radius of the observation point satisfies $R > r$, the potential of the distribution can be expanded as

$$\begin{aligned} V(\mathbf{R}, \theta) &= \sum_{l=0}^{\infty} \frac{1}{R^{l+1}} \int \rho(\mathbf{r}) P_l(\cos \theta) r^l d^3\mathbf{r} \\ &= \sum_{l=0}^{\infty} \frac{M_l}{R^{l+1}} P_l(\cos(\theta)), \end{aligned} \quad (\text{A14})$$

which is a form of multipole expansion in the polar angle only with multipole moments M_l defined exactly as in the Cartesian case, that is $M_0 = q$ —the net charge or zeroth moment for which the potential is attenuated by R^{-1} and $P_0(x) = 1$. The next term in the sum gives the first moment, defining M_1 as the dipole moment modulated by $P_1(\cos(\theta)) = \cos(\theta)$ which changes sign once over the interval 0 to π (see Fig. 23). This expansion applies to the axially symmetric distribution since there is no azimuthal angle dependence. The most general expansion involves spherical harmonics which are related to the Legendre functions by the addition theorem $P_n(\cos \theta) \propto \sum_m Y_l^m(\theta, \phi) Y_l^{*m}(\theta, \phi)$. The spherical multipole moments of the source are then defined as

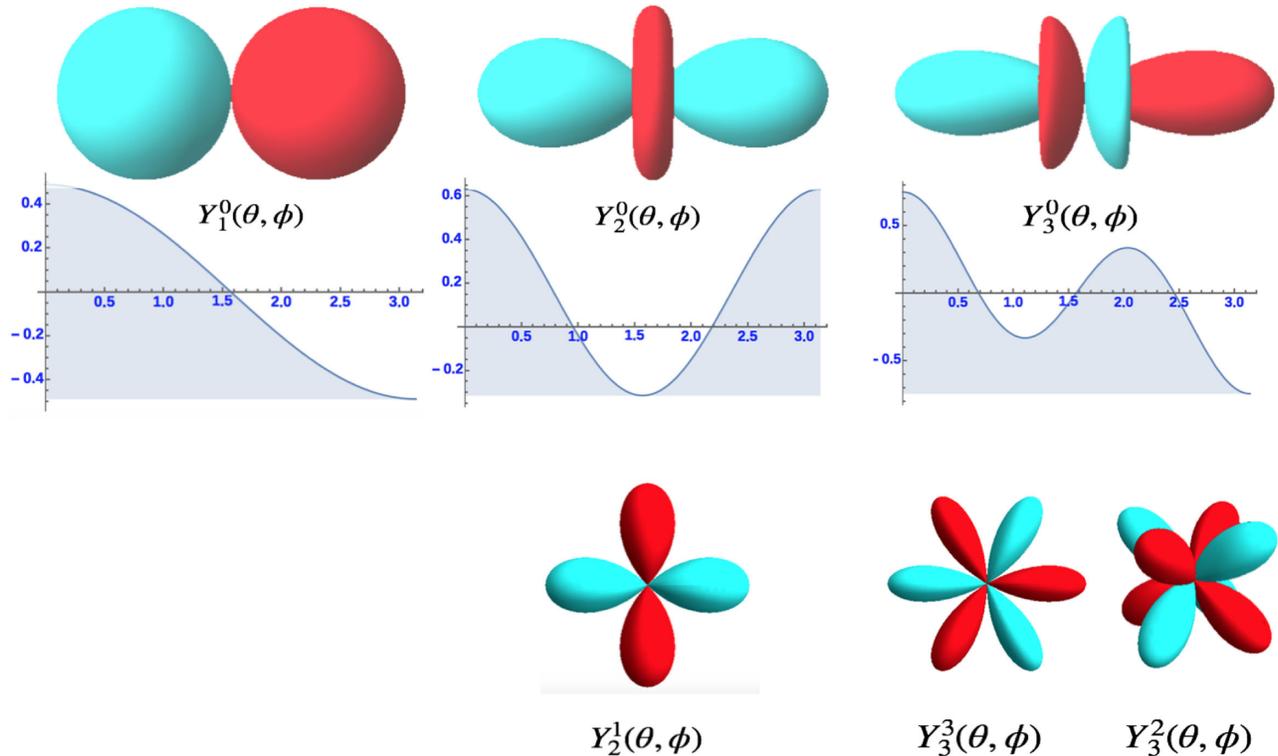


FIG. 23. Illustration of the spherical harmonics representation for multipole moments corresponding to (a) dipole, (b) quadrupole, and (c) octopole symmetries. The plots in the top three panels show the behavior of the corresponding Legendre function which determines the sign in the case where there is full axial symmetry corresponding to $m = 0$ (see the discussion in Sec. II.B.1 which introduced the idea of axially symmetric charge distributions). General examples of cases for $m \neq 0$ are also shown as viewed along the azimuthal (ϕ) axis.

$$q_{lm} = \int \rho(r) r^l Y_l^m(\theta, \phi) d^3 r. \quad (\text{A15})$$

By analogy with the Cartesian case, $l = 1$ corresponds to the dipole for which there are three forms arising from $m = -1, 0, 1$ giving the $\mu_{\alpha=x,y,z}$ Cartesian orientations.

The $l = 2$ form has quadrupolar symmetry with five ($m = \pm 2, \pm 1, 0$) independent components; the $l = 3$ octopole has seven, etc.; therefore the proper number of independent components for each multipole arises naturally in the spherical expansion. More generally, we can again make contact with group theory considerations as follows: the existence of a

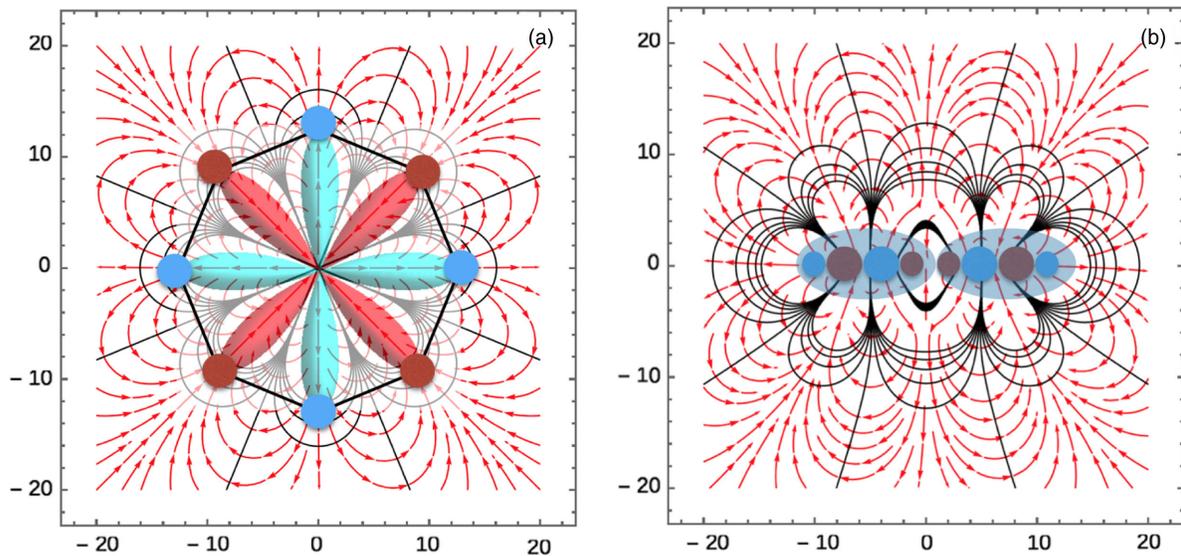


FIG. 24. A pure hexadecapole corresponding to $Y_4^4(\theta, \phi)$ (a) showing field lines and equipotentials. The field distribution is equivalent to that of alternating charges on the vertices of a regular octagon. (b) Two linear octopoles with charges reversed and displaced by a small distance to illustrate the similarity in the far field.

multipole there to be at least one nonvanishing multipole moment. We therefore require that the integrand transform according to the totally symmetric point group irreducible representation. Because the total charge distribution and the factor r^l transform according already satisfy this requirement

the symmetry of the integrand is completely determined by the spherical harmonic.

The spherical forms also permit natural extensions to higher multipole moments. For example, beyond the octopole the hexadecapole moment corresponds to the $l = 4$ index. The

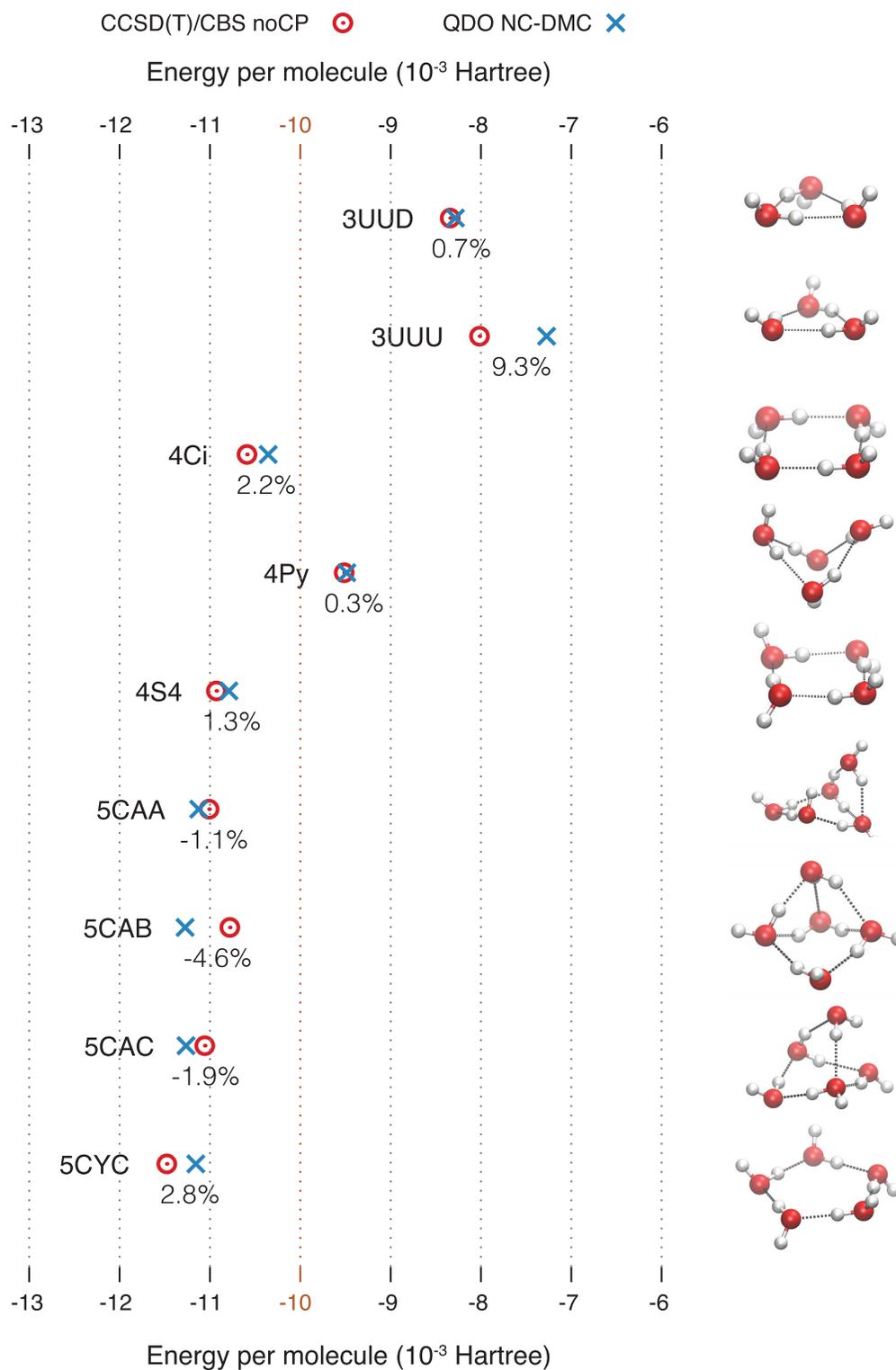


FIG. 25. (multipart figure 1/4) Energies per molecule of water clusters calculated using QDO water, compared with a CCSD(T)/CBS noCP. (Note that the axes in the first two parts of this figure are different than in the last two parts. The change in position of the orange line at 10 mHa indicates this change. From [Řezáč *et al.*, 2008](#) and [Temelso, Archer, and Shields, 2011](#)).

$Y_4^4(\theta, \phi)$ term is shown illustrating the azimuthal modulation in Fig. 24(a). The charge distribution shown is that of equal alternating charges on the vertices of a regular octagon. A Cartesian hexadecapole comprised of two canceling, linear octopoles with unequal charges is also shown for comparison.

APPENDIX B: CLUSTER ENERGIES

This Appendix contains Figs. 25–28 showing cluster energies.

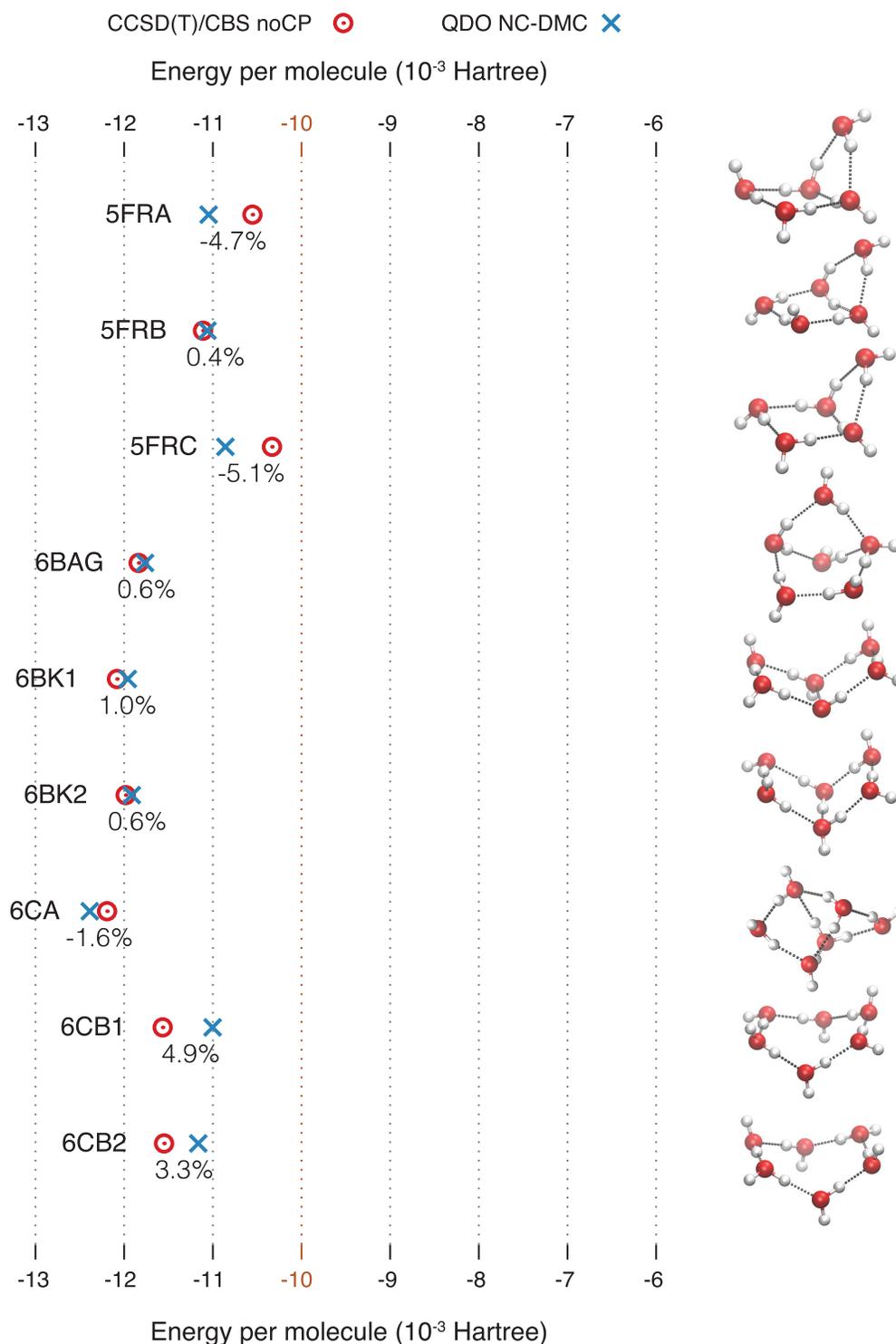


FIG. 26. (multipart figure 2/4) Energies per molecule of water clusters calculated using QDO water, compared with a CCSD(T)/CBS noCP. Note that the axes in the first two parts of this figure are different than in the last two parts. The change in position of the orange line at 10 mHa indicates this change. From Řezáč *et al.*, 2008 and Temelso, Archer, and Shields, 2011.

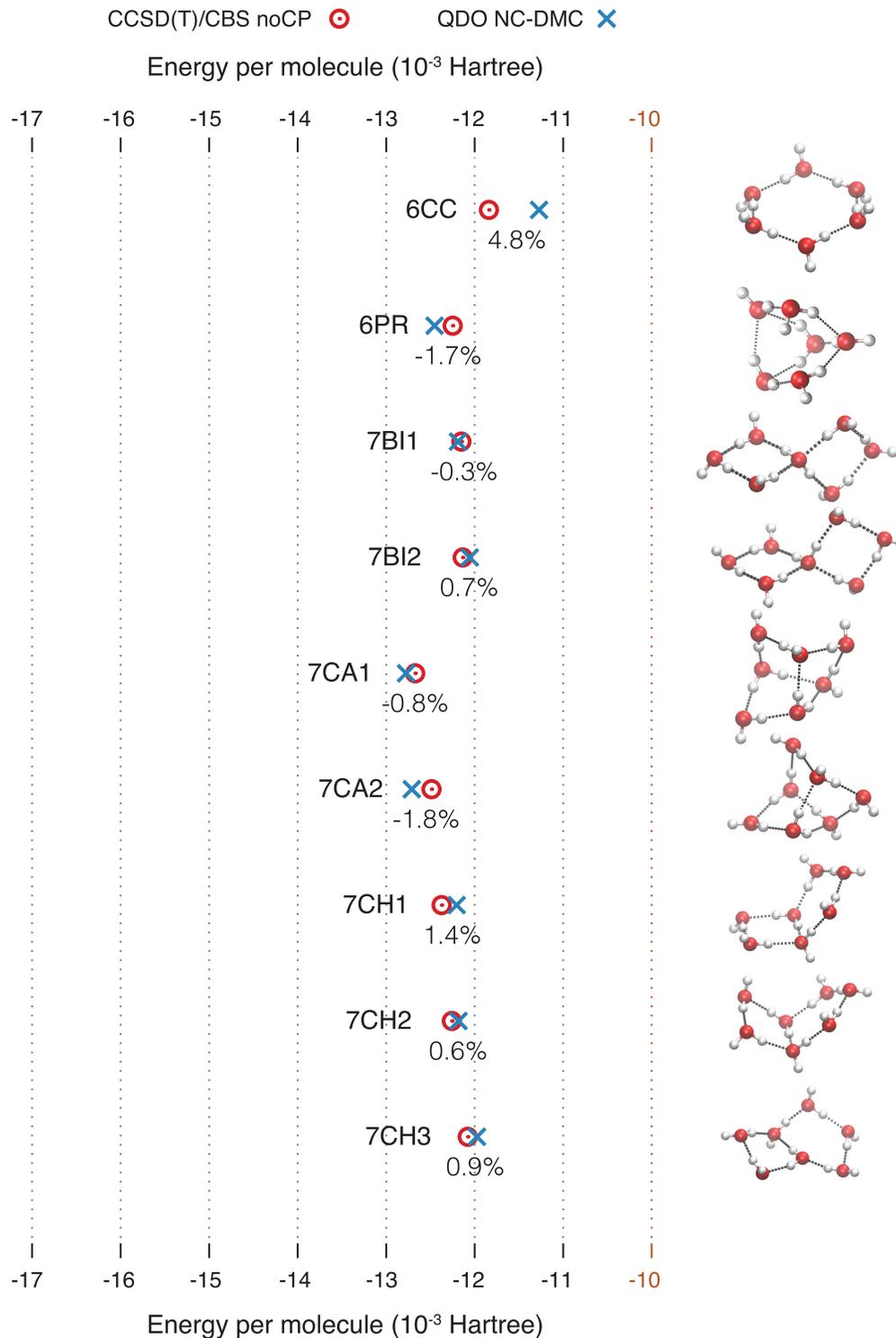


FIG. 27. (multipart figure 3/4) Energies per molecule of water clusters calculated using QDO water, compared with a CCSD(T)/CBS noCP. Note that the axes in the first two parts of this figure are different than in the last two parts. The change in position of the orange line at 10 mHa indicates this change. From [Řezáč *et al.*, 2008](#) and [Temelso, Archer, and Shields, 2011](#).

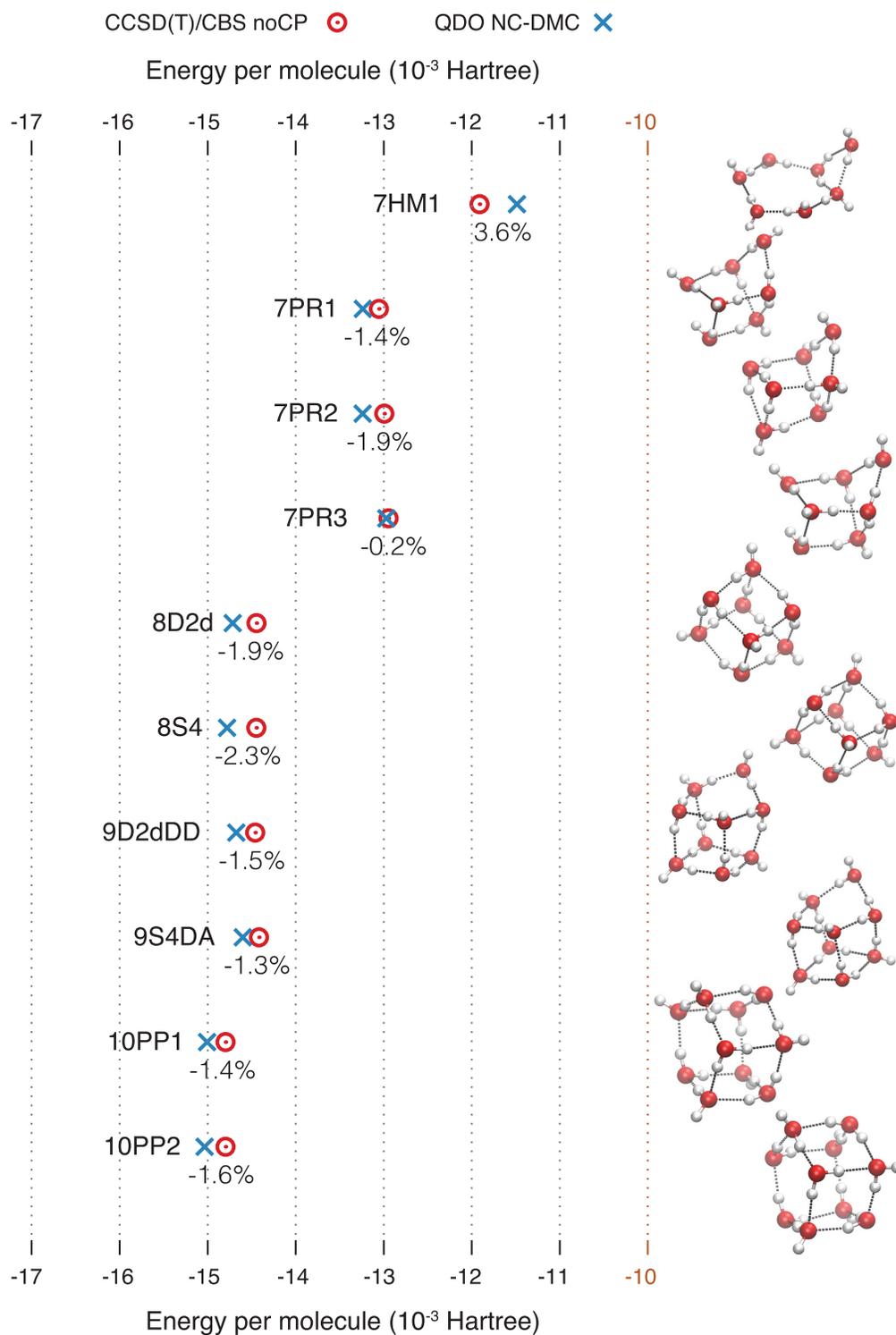


FIG. 28. (multipart figure 4/4) Energies per molecule of water clusters calculated using QDO water, compared with a CCSD(T)/CBS noCP. Note that the axes in the first two parts of this figure are different than in the last two parts. The change in position of the orange line at 10 mHa indicates this change. From Řezáč *et al.*, 2008; Temelso, Archer, and Shields, 2011.

REFERENCES

- Abascal, J. L. F., E. Sanz, R. G. Fernández, and C. Vega, 2005, *J. Chem. Phys.* **122**, 234511.
- Abascal, J. L. F., and C. Vega, 2005, *J. Chem. Phys.* **123**, 234505.
- Alfè, D., A. P. Bartók, G. Csányi, and M. J. Gillan, 2013, *J. Chem. Phys.* **138**, 221102.
- Allen, M. P., and D. J. Tildesley, 2017, *Computer Simulation of Liquids* (Oxford University Press, Oxford).
- Altland, A., and B. Simons, 2010, *Condensed Matter Field Theory*, Cambridge books online (Cambridge University Press, Cambridge).
- Amann-Winkel, K., M.-C. Bellissent-Funel, L. E. Bove, T. Loerting, A. Nilsson, A. Paciaroni, D. Schlesinger, and L. Skinner, 2016, *Chem. Rev.* **116**, 7570.
- Ambrosetti, A., D. Alfè, R. A. DiStasio, Jr., and A. Tkatchenko, 2014, *J. Phys. Chem. Lett.* **5**, 849.
- Ambrosio, F., G. Miceli, and A. Pasquarello, 2016, *J. Phys. Chem. B* **120**, 7456.
- Applequist, J., 1985, *J. Chem. Phys.* **83**, 809.
- Applequist, J., J. R. Carl, and K.-K. Fung, 1972, *J. Am. Chem. Soc.* **94**, 2952.
- Aryal, P., M. S. Sansom, and S. J. Tucker, 2015, *J. Mol. Biol.* **427**, 121.
- Axilrod, B. M., 1951, *J. Chem. Phys.* **19**, 719.
- Axilrod, B. M., and E. Teller, 1943, *J. Chem. Phys.* **11**, 299.
- Barker, J. A., and R. O. Watts, 1969, *Chem. Phys. Lett.* **3**, 144.
- Baym, G., and N. D. Mermin, 1961, *J. Math. Phys. (N.Y.)* **2**, 232.
- Bell, R. J., 1970, *J. Phys. B* **3**, 751.
- Benighaus, T., R. A. DiStasio, R. C. Lochan, J.-D. Chai, and M. Head-Gordon, 2008, *J. Phys. Chem. A* **112**, 2702.
- Ben-Naim, A., 1972, *Mol. Phys.* **24**, 705.
- Berendsen, H. J. C., J. R. Grigera, and T. P. Straatsma, 1987, *J. Phys. Chem.* **91**, 6269.
- Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, 1981, *Intermolecular Forces* (Springer, New York), pp. 331–342.
- Bernal, J. D., and R. H. Fowler, 1933, *J. Chem. Phys.* **1**, 515.
- Bloembergen, N., 1996, *Nonlinear Optics* (World Scientific, Singapore).
- Blum, K., 2013, *Density Matrix Theory and Applications*, Physics of atoms and molecules (Springer, New York).
- Born, M., and R. Oppenheimer, 1927, *Ann. Phys. (Berlin)* **389**, 457.
- Brazhkin, V. V., Y. D. Fomin, A. G. Lyapin, V. N. Ryzhov, and K. Trachenko, 2012, *Phys. Rev. E* **85**, 031203.
- Buckingham, A. D., 1959, *Q. Rev. Chem. Soc.* **13**, 183.
- Buckingham, A. D., 1967, “Permanent and induced molecular moments and long-range intermolecular forces,” in *Intermolecular Forces*, Advances in Chemical Physics, Vol. 12, edited by J. O. Hirschfelder, Chap. 2 (John Wiley & Sons, New York), pp. 107–142.
- Cagniard de la Tour, C., 1822, *Ann. Chim. Phys.* **21**, 127.
- Cao, J., and B. J. Berne, 1992, *J. Chem. Phys.* **97**, 8628.
- Cao, J., and G. J. Martyna, 1996, *J. Chem. Phys.* **104**, 2028.
- Casimir, H. B. G., and D. Polder, 1948, *Phys. Rev.* **73**, 360.
- Ceperley, D. M., 1995, *Rev. Mod. Phys.* **67**, 279.
- Chaimovich, A., and M. S. Shell, 2009, *Phys. Chem. Chem. Phys.* **11**, 1901.
- Chialvo, A. A., and P. T. Cummings, 1998, *Fluid Phase Equilib.* **150–151**, 73.
- Cipcigan, F. S., V. P. Sokhan, J. Crain, and G. J. Martyna, 2016, *J. Comput. Phys.* **326**, 222.
- Cipcigan, F. S., V. P. Sokhan, A. P. Jones, J. Crain, and G. J. Martyna, 2015, *Phys. Chem. Chem. Phys.* **17**, 8660.
- Cipcigan, F. S., V. P. Sokhan, G. J. Martyna, and J. Crain, 2018, *Sci. Rep.* **8**, 1718.
- Cisneros, G. A., K. T. Wikfeldt, L. Ojamäe, J. Lu, Y. Xu, H. Torabifard, A. P. Bartók, G. Csányi, V. Molinero, and F. Paesani, 2016, *Chem. Rev.* **116**, 7501.
- Cole, M. W., D. Velegol, H.-Y. Kim, and A. A. Lucas, 2009, *Mol. Simul.* **35**, 849.
- Crawford, T. D., and H. F. Schaefer, 2007, “An introduction to coupled cluster theory for computational chemists,” in *Reviews in Computational Chemistry*, Vol. 14, Chap. 2 (John Wiley & Sons, Inc., New York), pp. 33–136.
- Davies, M., 2007, *Dielectric and Related Molecular Processes*, Specialist Periodical Reports No. v.1 (CRC Press, Boca Raton, FL).
- Davis, W., 2011, *Physical Chemistry: A Modern Introduction, Second Edition* (Taylor & Francis, London).
- Debenedetti, P. G., 2003, *J. Phys. Condens. Matter* **15**, R1669.
- Dias, C. L., T. Ala-Nissila, M. Grant, and M. Karttunen, 2009, *J. Chem. Phys.* **131**, 054505.
- Dirac, P. A. M., 1933, *Phys. Z. Sowjetunion* **3**, 64.
- Donchev, A., 2006, *J. Chem. Phys.* **125**, 074713.
- Drude, P., 1900, *Lehrbuch der Optik* (S. Hirzel, Leipzig, Germany).
- Dunning, T. H., 1989, *J. Chem. Phys.* **90**, 1007.
- Elrod, M. J., and R. J. Saykally, 1994, *Chem. Rev.* **94**, 1975.
- Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, 1995, *J. Chem. Phys.* **103**, 8577.
- Fano, U., 1957, *Rev. Mod. Phys.* **29**, 74.
- Feynman, R. P., 1939, *Phys. Rev.* **56**, 340.
- Feynman, R. P., 1949, *Phys. Rev.* **76**, 769.
- Fisher, M. E., 1969, *J. Chem. Phys.* **50**, 3756.
- Fomin, Y. D., V. N. Ryzhov, E. N. Tsiok, and V. V. Brazhkin, 2015, *Sci. Rep.* **5**, 14234.
- Fortes, A. D., I. G. Wood, M. Alfredsson, L. Vočadlo, and K. S. Knight, 2005, *J. Appl. Crystallogr.* **38**, 612.
- Fuentes-Azcatl, R., and J. Alejandre, 2014, *J. Phys. Chem. B* **118**, 1263.
- Fuentes-Azcatl, R., N. Mendoza, and J. Alejandre, 2015, *Physica A (Amsterdam)* **420**, 116.
- Gaiduk, A. P., F. Gygi, and G. Galli, 2015, *J. Phys. Chem. Lett.* **6**, 2902.
- Gibbs, J. W., 1902, *Elementary principles in statistical mechanics* (Yale University Press, New Haven, CT).
- Gillan, M. J., D. Alfè, A. P. Bartók, and G. Csányi, 2013, *J. Chem. Phys.* **139**, 244504.
- Gillan, M. J., D. Alfè, and A. Michaelides, 2016, *J. Chem. Phys.* **144**, 130901.
- Gillan, M. J., F. R. Manby, M. D. Towler, and D. Alfè, 2012, *J. Chem. Phys.* **136**, 244105.
- Gingrich, T. R., and M. Wilson, 2010, *Chem. Phys. Lett.* **500**, 178.
- Goldstein, H., C. Poole, and J. Safko, 2001, *Classical Mechanics* (Addison-Wesley Boston, Boston), 3rd ed.
- Gordon, M. S., D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, 2012, *Chem. Rev.* **112**, 632.
- Gray, C., K. Gubbins, and C. Joslin, 2011, *Theory of Molecular Fluids: Vol. 2: Applications*, International Series of Monographs on Chemistry (Oxford University Press, Oxford).
- Gray, C. G., and K. E. Gubbins, 1984, *Theory of Molecular Fluids: I: Fundamentals*, International Series of Monographs on Chemistry (Oxford University Press, Oxford).
- Griffiths, D., 2017, *Introduction to Electrodynamics* (Cambridge University Press, Cambridge, England).
- Guillot, B., 2002, *J. Mol. Liq.* **101**, 219.
- Halgren, T. A., and W. Damm, 2001, *Curr. Opin. Struct. Biol.* **11**, 236.

- Hansen, J., and I. McDonald, 2013, *Theory of Simple Liquids* (Academic Press, New York).
- Heitler, W., and F. London, 1927, *Z. Phys.* **44**, 455.
- Hellmann, H., 1937, *Einführung in die Quantenchemie* (Franz Deuticke, Leipzig und Wien).
- Herman, M. F., E. J. Bruskin, and B. J. Berne, 1982, *J. Chem. Phys.* **76**, 5150.
- Hess, S., 2015, *Tensors for Physics*, Undergraduate Lecture Notes in Physics (Springer International Publishing Switzerland, Basel).
- Hirschfelder, J., 2009, *Intermolecular Forces*, Advances in Chemical Physics (Wiley, New York).
- Horn, H. W., W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, 2004, *J. Chem. Phys.* **120**, 9665.
- Izadi, S., R. Anandkrishnan, and A. V. Onufriev, 2014, *J. Phys. Chem. Lett.* **5**, 3863.
- Izadi, S., and A. V. Onufriev, 2016, *J. Chem. Phys.* **145**, 074501.
- Jäger, B., R. Hellmann, E. Bich, and E. Vogel, 2016, *J. Chem. Phys.* **144**, 114304.
- James, H. M., and A. S. Coolidge, 1933, *J. Chem. Phys.* **1**, 825.
- Jeevanjee, N., 2015, *An Introduction to Tensors and Group Theory for Physicists* (Springer International Publishing Switzerland, Basel).
- Johnson, C. H. J., and T. H. Spurling, 1971, *Aust. J. Chem.* **24**, 2205.
- Johnson, M. E., T. Head-Gordon, and A. A. Louis, 2007, *J. Chem. Phys.* **126**, 144509.
- Jones, A., F. Cipcigan, V. P. Sokhan, J. Crain, and G. J. Martyna, 2013, *Phys. Rev. Lett.* **110**, 227801.
- Jones, A., A. Thompson, J. Crain, M. H. Müser, and G. J. Martyna, 2009, *Phys. Rev. B* **79**, 144119.
- Jones, A. P., J. Crain, F. Cipcigan, V. Sokhan, M. Modani, and G. Martyna, 2013, *Mol. Phys.* **111**, 3465.
- Jones, A. P., J. Crain, V. P. Sokhan, T. W. Whitfield, and G. J. Martyna, 2013, *Phys. Rev. B* **87**, 144103.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, 1983, *J. Chem. Phys.* **79**, 926.
- Kamb, B., 1964, *Acta Crystallogr.* **17**, 1437.
- Kaminski, G. A., H. A. Stern, B. J. Berne, R. A. Friesner, Y. X. Cao, R. B. Murphy, R. Zhou, and T. A. Halgren, 2002, *J. Comput. Chem.* **23**, 1515.
- Kemp, D. D., and M. S. Gordon, 2008, *J. Phys. Chem. A* **112**, 4885.
- Klimeš, J., D. R. Bowler, and A. Michaelides, 2009, *J. Phys. Condens. Matter* **22**, 022201.
- Kohn, W., 1999, *Rev. Mod. Phys.* **71**, 1253.
- Kühne, T. D., and R. Z. Khalullin, 2013, *Nat. Commun.* **4**, 1450.
- Lemmon, E. W., M. O. McLinden, and D. G. Friend, 2018, *NIST Chemistry webbook, NIST Standard reference database*, Vol. 69, edited by P. J. Linstrom and W. G. Mallard (National Institute of Standards and Technology, Gaithersburg, MD) (retrieved November 12, 2018).
- Limmer, D. T., and D. Chandler, 2015, *Mol. Phys.* **113**, 2799.
- Lin, I.-C., A. P. Seitsonen, I. Tavernelli, and U. Rothlisberger, 2012, *J. Chem. Theory Comput.* **8**, 3902.
- London, F., 1937, *Trans. Faraday Soc.* **33**, 8b.
- Lotrich, V., N. Flocke, M. Ponton, A. D. Yau, A. Perera, E. Deumens, and R. J. Bartlett, 2008, *J. Chem. Phys.* **128**, 194104.
- Lu, J., Y. Qiu, R. Baron, and V. Molinero, 2014, *J. Chem. Theory Comput.* **10**, 4104.
- Madden, P. A., and M. Wilson, 1996, *Chem. Soc. Rev.* **25**, 339.
- Mahan, G., and K. Subbaswamy, 2013, *Local Density Theory of Polarizability*, Phys. Solids Liq. (Springer, New York).
- Mahanty, J., and B. W. Ninham, 1976, *Dispersion forces*, Vol. 5 (Academic Press, London).
- Maheshwary, S., N. Patel, N. Sathyamurthy, A. D. Kulkarni, and S. R. Gadre, 2001, *J. Phys. Chem. A* **105**, 10525.
- Mahoney, M. W., and W. L. Jorgensen, 2000, *J. Chem. Phys.* **112**, 8910.
- Maroulis, G., 2006, *Atoms, Molecules and Clusters in Electric Fields: Theoretical Approaches to the Calculation of Electric Polarizability*, Series in computational, numerical, and mathematical methods in sciences and engineering (Imperial College Press, London).
- Martyn, R. M., 2004, *Electronic Structure. Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, England).
- Martyna, G. J., 1994, *Phys. Rev. E* **50**, 3234.
- Martyna, G. J., and B. J. Berne, 1988, *J. Chem. Phys.* **88**, 4516.
- Martyna, G. J., and B. J. Berne, 1989, *J. Chem. Phys.* **90**, 3744.
- Martyna, G. J., M. L. Klein, and M. Tuckerman, 1992, *J. Chem. Phys.* **97**, 2635.
- Marx, D., M. E. Tuckerman, and G. J. Martyna, 1999, *Comput. Phys. Commun.* **118**, 166.
- Mashayak, S. Y., M. N. Jochum, K. Koschke, N. R. Aluru, V. Rühle, and C. Junghans, 2015, *PLoS One* **10**, e0131754.
- Maurer, R. J., V. G. Ruiz, and A. Tkatchenko, 2015, *J. Chem. Phys.* **143**, 102808.
- Millot, C., J.-C. Soetens, M. T. C. Martins Costa, M. P. Hodges, and A. J. Stone, 1998, *J. Phys. Chem. A* **102**, 754.
- Millot, C., and A. J. Stone, 1992, *Mol. Phys.* **77**, 439.
- Mishima, O., 2010, *J. Chem. Phys.* **133**, 144503.
- Misra, R. P., and D. Blankschtein, 2017, *J. Phys. Chem. C* **121**, 28166.
- Molinero, V., 2013, “Minimalistic models of water,” *Proceedings of the International School of Physics “Enrico Fermi”* (IOP Press, Amsterdam).
- Molinero, V., and E. B. Moore, 2009, *J. Phys. Chem. B* **113**, 4008.
- Morawietz, T., A. Singraber, C. Dellago, and J. Behler, 2016, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8368.
- Muto, Y., 1943, *Proc. Phys. Math. Soc. Jpn.* **17**, 629.
- Nada, H., and J. P. J. M. van der Eerden, 2003, *J. Chem. Phys.* **118**, 7401.
- Nilsson, A., D. Schlesinger, and L. G. M. Pettersson, 2015, “Water: Fundamentals as the Basis for Understanding the Environment and Promoting Technology,” in *Proceedings of the International School of Physics “Enrico Fermi,”* Vol. 187 (IOS Press Amsterdam, SIF Bologna), pp. 77–135.
- Ojamae, L., and K. Hermansson, 1994, *J. Phys. Chem.* **98**, 4271.
- Oostenbrink, C., A. Villa, A. E. Mark, and W. F. Van Gunsteren, 2004, *J. Comput. Chem.* **25**, 1656.
- Pallares, G., M. E. M. Azouzi, M. A. González, J. L. Aragones, J. L. F. Abascal, C. Valeriani, and F. Caupin, 2014, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7936.
- Parrill, A., and K. Lipkowitz, 2016, *Reviews in Computational Chemistry*, Vol. 29 (Wiley, New York).
- Parsegian, V., 2005, *Van der Waals Forces: A Handbook for Biologists, Chemists, Engineers, and Physicists* (Cambridge University Press, Cambridge, England).
- Patkowski, K., 2017, “Benchmark databases of intermolecular interaction energies: Design, construction, and significance,” in *Annual Reports in Computational Chemistry*, Vol. 13, edited by D. A. Dixon, Chap. 1 (Elsevier, New York), pp. 3–91.
- Pauling, L., 1960, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, NY), 3rd ed.
- Pople, J. A., 1999, *Rev. Mod. Phys.* **71**, 1267.
- Rahman, A., 1971, *J. Chem. Phys.* **55**, 3336.
- Řezáč, J., and P. Hobza, 2013, *J. Chem. Theory Comput.* **9**, 2151.

- Řezáč, J., *et al.*, 2008, *Collect. Czech. Chem. Commun.* **73**, 1261.
- Rick, S. W., 2004, *J. Chem. Phys.* **120**, 6085.
- Risthaus, T., and S. Grimme, 2013, *J. Chem. Theory Comput.* **9**, 1580.
- Sadus, R., 2002, *Molecular Simulation of Fluids: Theory, Algorithms, and Object-orientation* (Elsevier, New York).
- Santra, B., R. A. DiStasio, F. Martelli, and R. Car, 2015, *Mol. Phys.* **113**, 2829.
- Santra, B., A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, and M. Scheffler, 2008, *J. Chem. Phys.* **129**, 194111.
- Schmidt, J., J. VandeVondele, I.-F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy, 2009, *J. Phys. Chem. B* **113**, 11959.
- Schwabl, F., 2007, *Quantum Mechanics* (Springer, Berlin/Heidelberg).
- Shankar, R., 2017, *Quantum Field Theory and Condensed Matter: An Introduction* (Cambridge University Press, Cambridge, England).
- Shell, M. S., 2008, *J. Chem. Phys.* **129**, 144108.
- Singh, G., A. Nandi, and S. R. Gadre, 2016, *J. Chem. Phys.* **144**, 104102.
- Sokhan, V. P., A. Jones, F. S. Cipcigan, J. Crain, and G. J. Martyna, 2015a, *Phys. Rev. Lett.* **115**, 117801.
- Sokhan, V. P., A. P. Jones, F. S. Cipcigan, J. Crain, and G. J. Martyna, 2015b, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6341.
- Sprik, M., and M. L. Klein, 1988, *J. Chem. Phys.* **89**, 7556.
- Stillinger, F. H., and A. Rahman, 1974, *J. Chem. Phys.* **60**, 1545.
- Stone, A., and M. Alderton, 1985, *Mol. Phys.* **56**, 1047.
- Stone, A. J., 2013, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford), 2nd ed.
- Stone, A. J., and C.-S. Tong, 1994, *J. Comput. Chem.* **15**, 1377.
- Strong, S. E., L. Shi, and J. L. Skinner, 2018, *J. Chem. Phys.* **149**, 084504.
- Tang, K., 2006, *Mathematical Methods for Engineers and Scientists 3: Fourier Analysis, Partial Differential Equations and Variational Methods*, *Mathematical Methods for Engineers and Scientists* (Springer, Berlin/Heidelberg).
- Temelso, B., K. A. Archer, and G. C. Shields, 2011, *J. Phys. Chem. A* **115**, 12034.
- Thole, B. T., 1981, *Chem. Phys.* **59**, 341.
- Tkatchenko, A., R. A. DiStasio, R. Car, and M. Scheffler, 2012, *Phys. Rev. Lett.* **108**, 236402.
- Tkatchenko, A., and M. Scheffler, 2009, *Phys. Rev. Lett.* **102**, 073005.
- Tobias, D. J., G. J. Martyna, and M. L. Klein, 1993, *J. Phys. Chem.* **97**, 12959.
- Trotter, H. F., 1959, *Proc. Am. Math. Soc.* **10**, 545.
- Tuckerman, M. E., B. J. Berne, G. J. Martyna, and M. L. Klein, 1993, *J. Chem. Phys.* **99**, 2796.
- Tully, J. C., 2012, *J. Chem. Phys.* **137**, 22A301.
- Van Kempen, E. G. M., S. J. J. M. F. Kokkelmans, D. J. Heinzen, and B. J. Verhaar, 2002, *Phys. Rev. Lett.* **88**, 093201.
- Vanommeslaeghe, K., *et al.*, 2010, *J. Comput. Chem.* **31**, 671.
- Van Vleet, M. J., A. J. Misquitta, A. J. Stone, and J. R. Schmidt, 2016, *J. Chem. Theory Comput.* **12**, 3851.
- Vega, C., 2015, *Mol. Phys.* **113**, 1145.
- Vega, C., and J. L. F. Abascal, 2005, *J. Chem. Phys.* **123**, 144504.
- Vega, C., and J. L. F. Abascal, 2011, *Phys. Chem. Chem. Phys.* **13**, 19663.
- Vega, C., J. L. F. Abascal, E. Sanz, L. G. MacDowell, and C. McBride, 2005, *J. Phys. Condens. Matter* **17**, S3283.
- von Lilienfeld, A. O., and A. Tkatchenko, 2010, *J. Chem. Phys.* **132**, 234109.
- Vrabec, J., J. Stoll, and H. Hasse, 2001, *J. Phys. Chem. B* **105**, 12126.
- Wagner, J. W., J. F. Dama, A. E. P. Durumeric, and G. A. Voth, 2016, *J. Chem. Phys.* **145**, 044108.
- Wagner, W., and A. Pruß, 2002, *J. Phys. Chem. Ref. Data* **31**, 387.
- Wang, F., and K. D. Jordan, 2001, *J. Chem. Phys.* **114**, 10717.
- Wang, J., R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, 2004, *J. Comput. Chem.* **25**, 1157.
- Wheatley, R. J., 2008, *J. Comput. Chem.* **29**, 445.
- Whitfield, T. W., and G. J. Martyna, 2006, *Chem. Phys. Lett.* **424**, 409.
- Whitfield, T. W., and G. J. Martyna, 2007, *J. Chem. Phys.* **126**, 074104.
- Williams, R. W., and D. Malhotra, 2006, *Chem. Phys.* **327**, 54.
- Willow, S. Y., M. A. Salim, K. S. Kim, and S. Hirata, 2015, *Sci. Rep.* **5**, 14358.
- Witten, E., 2010, [arXiv:1009.6032](https://arxiv.org/abs/1009.6032).
- Yarnell, J. L., M. J. Katz, R. G. Wenzel, and S. H. Koenig, 1973, *Phys. Rev. A* **7**, 2130.
- Yoo, S., X. C. Zeng, and S. S. Xantheas, 2009, *J. Chem. Phys.* **130**, 221102.
- Zangwill, A., 2013, *Modern Electrodynamics*, *Modern Electrodynamics* (Cambridge University Press, Cambridge, England).
- Zavadlav, J., M. N. Melo, S. J. Marrink, and M. Praprotnik, 2015, *J. Chem. Phys.* **142**, 244118.