Taylor & Francis
Taylor & Francis Group

REVIEW

🔓 OPEN ACCESS   Check for updates

Interatomic potentials: achievements and challenges

Martin H. Müser[a], Sergey V. Sukhomlinov[a] and Lars Pastewka[b,c]

[a]Department of Materials Science and Engineering, Saarland University, Saarbrücken, Germany;
[b]Department of Microsystems Engineering, University of Freiburg, Freiburg, Germany; [c]Cluster of Excellence livMatS, Freiburg Center for Interactive Materials and Bioinspired Technologies, University of Freiburg, Freiburg, Germany

ABSTRACT

Interatomic potentials approximate the potential energy of atoms as a function of their coordinates. Their main application is the effective simulation of many-atom systems. Here, we review empirical interatomic potentials designed to reproduce elastic properties, defect energies, bond breaking, bond formation, and even redox reactions. We discuss popular two-body potentials, embedded-atom models for metals, bond-order potentials for covalently bonded systems, polarizable potentials including charge-transfer approaches for ionic systems and quantum-Drude oscillator models mimicking higher-order and many-body dispersion. Particular emphasis is laid on the question what constraints ensue from the functional form of a potential, e.g., in what way Cauchy relations for elastic tensor elements can be violated and what this entails for the ratio of defect and cohesive energies, or why the ratio of boiling to melting temperature tends to be large for potentials describing metals but small for short-ranged pair potentials. The review is meant to be pedagogical rather than encyclopedic. This is why we highlight potentials with functional forms sufficiently simple to remain amenable to analytical treatments. Our main objective is to provide a stimulus for how existing approaches can be advanced or meaningfully combined to extent the scope of simulations based on empirical potentials.
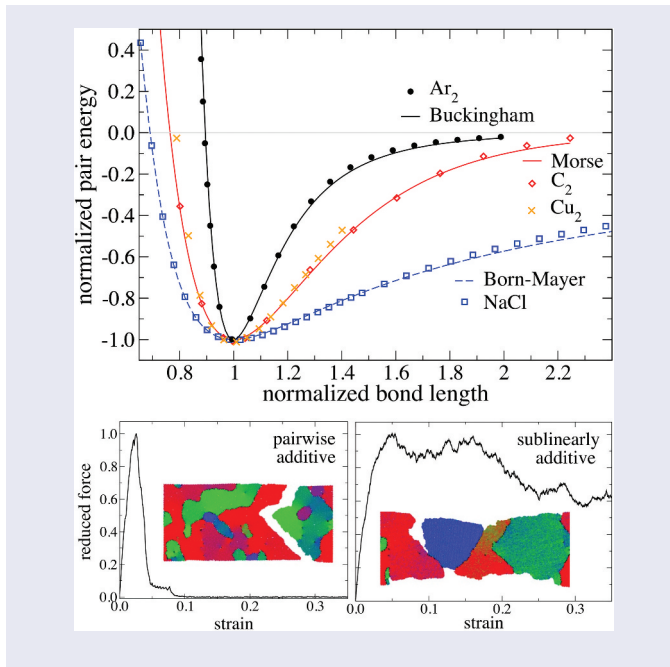

CONTACT Martin H. Müser ✉ martin.mueser@mx.uni-saarland.de 🖳 Department of Materials Science and Engineering, Saarland University, Saarbrücken, Germany

1. Introduction

Interatomic potentials are functions of nuclear coordinates approximating the electronic ground state energy, or for metals, the electronic free energy of a system. Forces on atoms, as needed in molecular-dynamics simulations, can be obtained by calculating the gradient of these functions with respect to the nuclear coordinates. Frequently, the terms *interatomic potential* and *force field* are used synonymously. There is, however, a subtle and sometimes important difference between the two. Many-body force fields are not necessarily gradients of a scalar function, in contrast to "real" interatomic forces when electrons are given enough time to equilibrate. In this review, we focus on force fields that can be represented as gradients of scalar functions approximating the exact interatomic potentials. Before getting started on details, let us take a step back.

In his *Lectures on Physics*, Richard Feynman wondered what statement would contain the most information in the fewest words, if all of scientific knowledge were to be destroyed in some cataclysm, and only one sentence could be passed to the next generations of scientists [1]: *I believe it is the atomic hypothesis that all things are made of atoms – little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another.* Two-

body potentials reflect this generic behavior, except those describing two ions carrying a charge of identical sign, since they repel each other even at large separation.

In fact, quite a bit can be learned from studying two-body potentials. For example, the crystalline structure of many elemental or binary crystals can be rationalized, as it is done in any better text book on solid state physics. Moreover, the functional form of *constitutive laws* is often independent of the precise nature of the potentials, such as Hooke's law, which states restoring forces of solids in equilibrium to be linear in small deformations. Even many non-linear constitutive equations do not depend on the details of the potential. The exponents with which compressibility or specific heat diverge as the temperature or pressure approaches the fluid-vapour critical point are identical for all substances irrespective of their specific interactions [2–4]. The way how the viscosity of polymers increases with molecular weight can also be described with the help of two-body potentials, as long as they prevent two polymers from crossing each other [5]. These are but a few examples for the success of two-body potentials. The only thing that does depend on chemical detail, it almost seems, are boring prefactors. Unfortunately, or, depending on your viewpoint, fortunately, this is not quite right.

Realistic parameterizations of two-body potentials, including Morse [6], Buckingham [7], or Lennard-Jones (LJ) [8], favor close-packed assemblies of atoms, e.g. face centered cubic or hexagonal closed packed lattices. Thus, neither molecular crystals as those formed by oxygen or nitrogen at small temperature, nor layered crystals like graphite at ambient conditions could be thermodynamically stable. Even those elements that do like to close pack may not be describable by two-body potentials. For example, the proper description of the elastic properties of metals and their ductility hinges on many-body terms [9]. The importance of directed interactions and thus of many-body terms is particularly apparent for carbon, where depending on the hybridization of individual carbon atoms, different interatomic forces ensue. In the worst case, history dependence of interatomic, or rather, interionic forces can arise. To illustrate this point, assume a NaCl molecule dissociates into ions in water and the water is evaporated later so that two isolated ions emerge. If, however, the NaCl molecule had been separated slowly in an inert gas environment before the nuclear degrees of freedom had been brought to their final destinations, two neutral atoms would have formed. This is because the ionization energy of sodium exceeds the electron affinity of chlorine. Real powerful potentials should mimic such a process correctly and account for charge transfer in the appropriate way, though the notion that forces arise as an unambigous function of nuclear coordinates would have to be abondened. From this discussion, we can see that no chemistry and none of its implication (batteries, laptops, soccer, or life in a more general sense) could arise if atomic systems could be accurately

described with two-body potentials. In brief, without many-body interactions, the world in general, and science in particular, would be much less exciting than it actually is.

The root of many-body potentials is that two atoms change their interaction when additional atoms are present, because their electronic structure changes. For instance, two hydrogen atoms prefer to form a strong covalent bond between them rather than to remain lonely. However, as soon as an oxygen enters the scene, the hydrogens stop being homo and happily form a heteronuclear water molecule with the oxygen.

Many-body potentials were much advanced over the last few decades. Realistic, large-scale simulations of several thousand atoms can nowadays be conducted even on commodity computers, at least for selected compounds such as simple metal alloys or hydrocarbons. Force-field based simulations of increasingly complex systems and processes become possible, including those during which atoms rehybridize in the course of a chemical reaction [10–12]. However, the need for further improvement remains, in particular for systems in which two or even more different bonding types (ionic, metallic, covalent, dispersive) combine. This text is meant to provide a rudimentary understanding for why specific functional forms for potentials were chosen and what a given (class of) potential may or may not achieve. With this, we hope to stimulate insight on how to combine or to generalize different potential classes so that systems held together by different bonding types can be better simulated in the future. Since an incredible number of papers on potentials has been published, we are certain to have missed important contributions, even if we did our best to identify the original key literature.

Focusing on concepts leaves us little room to provide readers with the best possible parameterization for a given substance and application. Performance evaluations of potentials and their parameterization containing information like *the specific parameterization by author A using potential B reproduces properties C and D of the material E but does a poor job on properties F and G* are probably better looked up in one of several excellent and important databases [13–18]. In addition, we wish to refer to many excellent reviews [19–33] or even text books [34,35] on interatomic potentials emphasizing specific parameterization of potentials more than this review. Here, we discuss materials- or element-specific numbers only when we explore the transferability of a potential, which is its ability to accurately predict properties, structures, and/or stoichiometries, which it was not adjusted for. In addition, one system is highlighted for each class of bonding type. To this end, we chose argon as the representative of systems bonding through van-der-Waals, because its dispersive interactions are half way between those of water and CH_2 or CH_3 units in hydrocarbon chains. Carbon, copper, and rocksalt (NaCl) complement the list to represent metals, covalently bonded materials, metals, and systems.

2. Classification, construction, and consequences of (many-body) potentials

2.1. Brief classification of interaction potentials

The way in which interactions are incorporated into potentials is so diverse that they can be classified according to several main criteria: the chemical nature of the bond, whether a bonding topology is prescribed or allowed to change, two-body versus many-body interactions, and the degree of empiricism or level of theory used to construct the functional form of the potential and to gauge adjustable parameters.

Atoms with open valence shells form covalent or metallic bonds, while atoms with closed valence shell interact through van der Waals forces. The latter include repulsion at short separation and dispersive forces, which result from the mutual induction of dipoles and other multipoles owing to quantum-mechanical fluctuations of the valence shell. Each bonding type has its own characteristics how interatomic forces change with interatomic separation and as a function of their environment, which motivates the distinction between open-shell and closed-shell potentials. Van der Waals or "non-bonded" interactions have energies on the order of thermal energy at room temperature and are thus weaker than covalent and metallic bonds. In addition, there can be Coulombic interactions due to partial or close-to-integer charges on atoms or ions. Interactions between atoms can also be a combination of all of the above, the most prominent example involving the hydrogen bond.

Potentials with a fixed bonding topology range from simple bead-spring models [5] to highly sophisticated, chemically realistic valence force fields containing explicit bond-angle or torsional terms [36–43]. Two atoms interact differently with each other, depending on whether they are considered bonded or non-bonded to each other, even if all involved atoms carry the same chemical symbol, or, if they correspond to the same coarse-grained entities, such as a so-called united atom reflecting, for example, a CH_2 or CH_3 group [39]. The design of valence force fields is in a rather mature state [36–43], which is why we will discuss them only peripherally. This does not prevent us from highly recommending their use for any situation, in which bond breaking and bond formation can be ruled out, simply because they are computationally lean while being sufficiently accurate for many purposes. Nonetheless, weaknesses remain, in particular the frequently poor treatment of higher-order and many-body dispersion as well as of charge transfer, which, however, are all best explained assuming free atoms or ions rather than bonded entities as reference. In a biophysical context, a proper treatment of dispersive interactions can be particularly important when molecules transfer between aqueous and hydrophobic environments, since the dispersive interactions with carbon-hydrogen

chains are stronger than with water molecules [44]. Adsorbing higher-order or many-body effects by tweaking pertinent Lennard-Jones interaction coefficients is then even more hazardous than for homogeneous phases. Thus, readers being interested in improving bonded potentials for biomolecular simulations have no excuse to stop reading here.

The distinction between two-body and many-body potentials should be self-explanatory. The latter can be further subdivided into explicit or implicit. An explicit many-body potential does not require on-the-fly adjustments of parameters like atomic charges or dipoles. In contrast, implicit many-body potentials necessitate the determination of such quantities. This is commonly achieved by minimizing expressions for the potential energy with respect to degrees of freedom representing a coarse-grained version of the full electronic response. Also explicit many-body potentials can be categorized into further subgroups. They can be cast in terms of analytical functions, or numerical tables, or they can be extrapolated from a large set of high-dimensional data as is done in machine-learned potentials [29,31–33].

Potential classes are sometimes also named differently depending on the quantum-mechanical framework from which they are motivated. For example, embedded-atom potentials [9,45] are best motivated from density-functional theory, while second-moment tight-binding potentials [46,47] arise, as their name says, quite naturally from the tight-binding approximation. Potentials lacking a direct theoretical justification are also called empirical. While we could elaborate much more on how to further classify potentials, we feel compelled to cut to the chase.

2.2. Constructing (many-body) potentials

Formally, it is possible to expand the potential energy $U(\{\mathbf{r}\})$ of a (classical) many-atom system into two-body, three-body, and higher-order contributions, where $\{\mathbf{r}\}$ represents the positions of atoms, i.e.

$$U(\{\mathbf{r}\}) = \sum_i U_1(\mathbf{r}_i) + \sum_{i<j} U_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i<j<k} U_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \cdots . \quad (1)$$

Here, the index in the $U(\ldots)$'s on the r.h.s. of the equation gives the order of the many-body expression. For non-elemental systems, the $U_n(\ldots)$ also depend explicitly on the atomic indices, or more precisely on the nature of the atoms interacting with each other. The underlying assumption of the expansion is that electrons are in a well-defined state, e.g. in their ground state or in thermal equilibrium.

As long as external fields are present, two-body and higher-order interactions generally do not obey Galilean or rotational invariance, which is illustrated in Figure 1. It depicts two atoms with full valence shells, which are polarized by the electrostatic field of an external dipole. Depending on the relative position of the external dipole to the two atoms, the atoms will experience an electrostatic repulsion, as in panel (a), or attraction, as in panel (b), in addition to their previous interaction.

In the absence of external fields, the two-body potential can be reduced to depend only on the interatomic distances r_{ij}, in which case Eq. (1) reduces to

$$U(\{\mathbf{r}\}) = \sum_{i<j} U_2(r_{ij}) + \sum_{i<j<k} U_3(r_{ij}, r_{jk}, r_{ki}) + \dots \tag{2}$$

Expansions of the form given by Eq. (2) are sometimes called *cluster potentials* [19,35]. Unfortunately, even this Galilean invariant expansion Eq. (1) is of limited (direct) use for systems of practical interest, because its convergence can be extremely slow. This is best seen when considering an ion in close proximity to a metallic surface. The ion induces a charge distribution in the metal that depends on the shape of the metal surface turning a formal expansion into explicit many-body terms useless. Likewise, comparing local structural motifs to an existing database, as is done with machine-learned potentials, does not appear to be promising in this example, although they frequently allow to effectively truncate an expansion similar to that outlined in Eq. (2), however, with atoms not being in vacuum but surrounded by other atoms.

Equation (2) converges relatively quickly only for closed-shell atoms, meaning atoms whose valence shell is filled, like noble gas atoms and singly charged alkali cations or halogen anions [48]. Even then convergence might not be truly impressive. Two-body interactions commonly used to simulate condensed phases of noble gas atoms are *effective* pair potentials rather than *true* pair potentials accurately reproducing the second virial coefficient [49]. The difference between effective and true potentials compensates to some
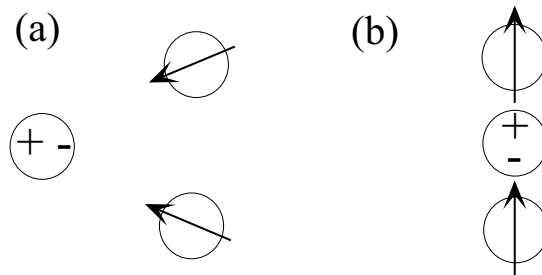

Figure 1. Effect of an external dipole, marked by a positive and negative sign, on secondary, induced dipoles indicated by arrows. Depending on the location and orientation of the central dipole, be it permanent or dynamic, the induced dipoles repel (a) or attract (b) each other.

extent for omitted many-body effects. Thus, pair and three-body potentials trained exclusively in the bulk risk to be inaccurate when applied to surfaces and gases. Despite its frequent ineffectiveness, Eq. (2) allows the difference between two opposite philosophies to the construction of potential function to be discussed: the bottom-up and the top-down design, which would lead to the construction of non-empirical and empirical potentials, respectively.

In the bottom-up approach, pen-on-paper quantum chemistry, density-functional theory (DFT), or any related electronic structure calculations provides analytical results and/or small-scale data, such as forces on individual atoms as a function of the atomic configurations $\{\mathbf{r}\}$ or energies of diatomic molecules as a function of bond length. Such information can motivate the functional form of a two-body potential or be used to construct soulless look-up tables. From a philosophical point of view, it could be argued that inputting experimental information on dimers, trimers, etc., falls into the bottom-up approach, as the electronic structure problem is solved by nature rather than by computers. Of course, the downside of letting nature do that job is that it is quite difficult to get enough experimental data on small molecules and clusters that would allow the higher-order terms in the series of Eq. (2) to be determined to any significant accuracy. Only pair potentials can probably be deduced to high precision from measurements of the vibrational and rotational spectra without having to make serious restrictions on their functional form.

In the top-down design, no atomic-scale information is provided, but instead collective, typically macroscopic properties, such as elastic properties, equations of state (EOS), surface energies, or the temperature dependence of the specific heat including heats of melting and evaporation. In the sense of an inverse problem, or reverse coarse-graining, potential energy functions are constructed such that the available information is reproduced. Historically, the first attempts to design potentials followed this approach. The arguably most systematic top-down design makes use of the virial expansion [50,51], in which the pressure p of a many-particle system in thermal equilibrium and in the absence of external fields is expanded into a power series of the number density ρ,

$$p = \rho k_B T \{ 1 + B_2(T)\rho + B_3(T)\rho^2 + \cdots \}, \tag{3}$$

where $k_B T$ is the thermal energy, while $B_2(T)$ and $B_3(T)$ denote the second and third virial coefficient, respectively. The so-called cluster expansion [52,53] allows the various virial coefficients to be related to two-body, three-body, and higher-order interactions. For example, knowledge of the second virial coefficient

$$B_2(T) = -\frac{1}{2} \int \mathrm{d}^3 r \left[\exp\left\{-\beta U_2(r)\right\} - 1\right] \qquad (4)$$

allows the pair potential to be reconstructed by an inverse procedure, though a numerically stable inversion works best if $U_2(r)$ is represented as a function with few adjustable parameters. The third virial coefficient, $B_3(T)$, depends on two- and three-body interactions, and so on. Thus, in principle, two substances may have similar second but different third virial coefficients so that knowledge of $B_3(T)$ allows deviations from pair additivity to be quantified and adjustable parameters of a three-body potential $U_3(r_{12}, r_{23}, r_{31})$ to be gauged.

Unfortunately, the virial expansion does not converge for every state point, due to the existence of thermodynamic discontinuities a.k.a. phase transformations [54]. This is one reason why the expansion of Eq. (2) cannot be fully parametrized from a cluster expansion after all. Yet, early calculations performed by van der Waals in the spirit of a cluster expansion, revealed that the leading-order corrections to the ideal-gas EOS requires the two-particle interaction in simple gases to obey [55]

$$U_2(r) = -\frac{C_6}{r^6}, \qquad (5)$$

at large distances. Today, C_6 is called the dipole-dipole dispersion coefficient.

About 50 years after van der Waals, London [56] rationalized why atoms and molecules with closed valence shells attract each other through pairwise additive $1/r^6$ interactions by coupling the quantum-mechanical ground state fluctuations of their dipoles in the far-field approximation. Extending London's treatment to higher-order electrostatic multipoles and beyond second-order perturbation theory leads to refinements of the theory, which are outlined in Sect. 4. Despite the existence of corrections, the bottom-up approach reveals quite clearly that the exponent in Eq. (5) is essentially exact, at least as long as electrostatic interactions can be treated as instantaneous (or the speed of light as infinitely large) [57,58]. Tweaking the exponent to better match reference data would quickly result in overfitting of the potential: a better match of the reference data would deteriorate the description of interatomic forces between two isolated particles at large distances. Thus, bottom-up and the top-down design of interatomic potentials are complementary to each other but should converge to similar results assuming sufficient and sufficiently accurate input.

2.3. Consequences of many-body interactions

A common way to assess the relevance of many-body interactions is to determine what percentage of the cohesive energy is related to the exact or to an effective two-body interaction. Such an analysis may be misleading,

because defect energies or elastic properties may not be described well even if the total energy of a crystal or the radial distribution function of a liquid reproduces *ab-initio* or experimental data. We can always adjust a pair potential that reproduces the cohesive energy of a specific crystalline phase (discussed extensively below) or that fits a specific liquid pair distribution function (see e.g. Refs. [59,60]). These properties often benefit from the annihilation of or from the insensitivity to many-body contributions, but the resulting potentials are then not *transferable* between crystal structures, temperatures, or other states of the system. Even three-body correlation functions may be poor properties on which to gauge potentials. Liquid copper and liquid argon just above their crystallization temperature both assemble in structures similar to that of random-sphere packing [61], despite their potentials being utterly different.

Induced dipoles or charge transfer decrease the total energy (and hence make the "bonds" stronger), or they would not occur. In contrast, most other many-body interactions weaken bonds, i.e. the energy per bond in metals decreases with increasing coordination number Z, typically with $1/\sqrt{Z}$ for metals and even more quickly for covalently bonded systems (see Sect. 5 or Refs. [46,62,63]). This scaling has important consequences not only on what crystal structure is assumed but also on the relative increase of boiling relative to the melting temperature, which is roughly 4% for argon but 100% for copper.

Another frequent consequence of many-body terms is that bond lengths become shorter with decreasing coordination or increasing bond order. One of the best known example may be the bond length between two carbon atoms: $a_0 \approx 1.54$ Å ($Z_H = 3$, ethane, C_2H_6), 1.34 Å ($Z_H = 2$, ethylene, C_2H_4) 1.20 Å ($Z_H = 1$, acetylene, C_2H_2), where Z_H states how many hydrogen atoms each carbon is bonded to. Similarly, the spacing between layers near a free metal surface contract [64], while common short-range pair potentials would predict them to expand.

Historically, the non-additivity of atomic potentials lead to the dismissal of the atomic hypothesis among many prominent scientists in the 19′th century, see Ref. [65] for a well written and very enlightening account of the debate. The assumption was that atoms, so they exist, should only interact through central potentials and be pairwise additive. Cauchy and Poisson [65] and also potentially St. Vernan [66] had shown that this reduced the number of independent elastic tensor elements, e.g. from 21 to 15 for a triclinic crystal and from three to two for a cubic crystal. Specifically, they found, see Sect. 3.5 for a derivation, that any permutation of the four indices of an elastic tensor element should leave it unchanged so that relations, now known as Cauchy relations, like, for example, $C_{1212} = C_{1122}$, or, $C_{12} = C_{66}$ (in cubic systems $C_{44} = C_{66}$) in Voigt or Nye notation, would

hold. It then came as a blow to the atomic hypothesis when Voigt noticed that none of the crystals he had investigated satisfied the Cauchy relation within experimental errors [67]. Figure 2 epitomizes results on the violation of the pair-potential assumption. The latter is also frequently stated in terms of the Cauchy pressure $P_C = (C_{12} - C_{44})/2$, which vanishes for athermal, inversion-symmetric, classical crystals interacting with pair potentials.

Rather than abandoning the assumption of the pairwise additivity of potentials, many scientists rejected the atomic hypothesis alltogether. However, Rutherford's scattering experiments in the early 20'th century removed all legitimate doubt about it, even if quite a few muddleheads keep taking issue with it up to this day in happy concert with deniers of human-made climate change. Finally, Born [75] reconciled the properties of the elastic tensor of real materials with the atomic hypothesis by showing that one (of several) reasons for the breakdown of Cauchy relations is the existence of many-body interactions.

It should certainly not be concluded that solids obeying the Cauchy relations automatically satisfy the pair-potential assumption, since there can be fortuitous or symmetry-induced cancellation of many-body effects. For example, the dipole polarizability of anions in the rocksalt structure cannot reveal itself in the elastic tensor for symmetry reasons.
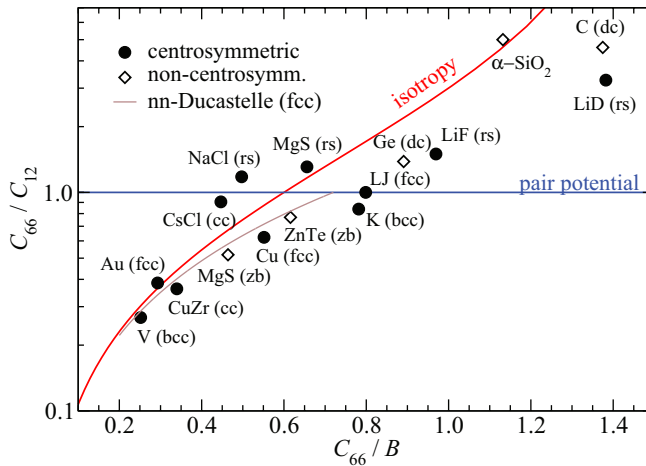

Figure 2. Ratio of elastic tensor elements C_{66} and C_{12} as a function of C_{66} in units of the bulk modulus B for a variety of cubic systems. ($C_{66} \equiv C_{44}$ in cubic systems.) The blue line holds for pair potentials, while the red line reflects elastically isotropic materials. The thin gray line assumes the generic model proposed by Ducastelle, introduced in Sect. 5.2, which has one dimensionless parameter affecting the C_{66}/C_{12} and C_{66}/B ratios in the nearest-neighbor approximation. Centrosymmetric lattices (closed circles) include face-centered cubic (fcc), body-centered cubic (bcc), rocksalt (rs), and caesium chloride (cc), while diamond cubic (dc) and zinc blende (zb) lack inversion symmetry (open diamonds). Experimental and simulated data are included for selected crystals (Cu [68], Au [68], K [68], V [68], C [68], Ge [68], NaCl [68], LiD [69], LiF [68], MgS(rs) [70], MgS(zb) [70], ZnTe [71], CsCl [72], $\alpha - SiO_2$ [73]) as well as calculated elastic constants for CuZr [74].

3. Two-body potentials

In this section, we introduce the most generic two-body potentials for pairs of atoms forming covalent and metallic bonds in the condensed phase as well as potentials for ionic and van der Waals interactions. We discuss these two classes separately, because atoms with open electron shells bond primarily through the sharing of electrons while closed-shell atoms interact predominantly through van der Waals or ionic forces. Both classes reflect Feynman's mantra that *[atoms attract] each other when they are a little distance apart, but repel upon being squeezed into one another* [1]. Interatomic potentials with well-motivated functional forms can reproduce the equation of state, reasonably well if their small number of adjustable parameters are gauged on a few accurate reference values. However, parameters are only transferable to other structures or properties if the pair-potential assumption is justified.

An element specific analysis of the potentials will be made for carbon (C), copper (Cu), sodium chloride (NaCl), and argon (Ar), which are representative of covalent, metallic, ionic, and van-der-Waals bonding, respectively. To set the stage for further discussion, their pair-potential energies are shown in Figure 3 as a function of the molecular bond length r. The reference data shown therein is not necessarily the most accurate on the market, but their overall errors should be minor compared to those of lean pair potentials, i.e. the data should be sufficiently accurate to test if a given functional form of the two-body potential is justified.

3.1. Morse potential

One of the early, great successes of quantum mechanics was Heitler and London's quantitative description of the chemical bond in a hydrogen molecule in 1927 [78]. Their theory is text-book material and will only be sketched here. Heitler and London linearly combined the atomic wave functions of the two hydrogen atoms forming a H_2 molecule, antisymmetrized the spin-wave function to reflect the Fermi principle, and evaluated the binding energy using first-order perturbation theory, thereby providing a lower bound for the molecular binding energy U_0. Sugiura [79] succeeded in identifying the correct solution for the exchange-energy integrals. Errors of order 40% for U_0 and bond length a_0 remain, see Figure 4. Treating a_B in the electronic ground-state wave functions $\Phi \propto \exp(-r/a_B)$ as a variational parameter, errors can be reduced by a factor close to two.

Morse [6] approximated the analytical expressions occurring in the Heitler-London treatment with
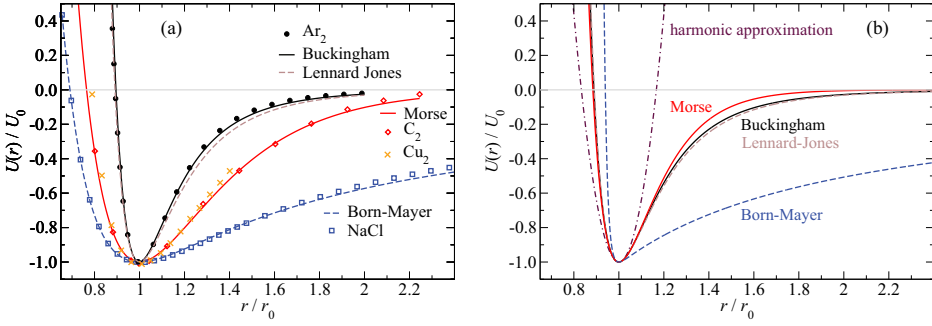
Figure 3. (a) Two-body potentials of different diatomic molecules, $U(r)$, normalized to the binding energy U_0 as a function of the bond length r in units of the equilibrium bond length r_0. Symbols reflect high-accuracy reference data, while lines show fits of pair potentials introduced in Sect. 3 to the energy minimum. Reference data include the Aziz potential for Ar_2 [76], and quantum chemical calculations for C_2 [77], Cu_2 [215], and NaCl [339]. Numbers used for the fits are listed in Table 1. (b) Two-body potentials with a curvature in the minimum corresponding to that of the Lennard-Jones potential.

Table 1. Comparison of pair-potential parameters obtained after fitting adjustable coefficients to accurate reference data on diatomic molecules (mol) and on crystals, where dc and rs are the diamond cubic and rock salt structure, respectively. Numbers indexed with a star are functions of the adjustable parameters. Used potentials are, Born-Mayer/Buckingham for argon and sodium chloride, and the Morse potential for carbon and copper. Plots of the respective potentials are shown in Figure 3. The first seven neighbor shells are included into the fit.

		U_0 (eV)	r_0 (Å)	n	r_0/n (Å)
Ar	(mol)	0.01234	3.757	14.6	0.26*
Ar	(fcc)	0.01155	3.811	14.1	0.27*
C	(mol)	6.318	1.250	5.82	0.21*
C	(dc)	2.868	1.804	5.34	0.34*
Cu	(mol)	1.823	2.275	6.80	0.33*
Cu	(fcc)	0.3964	2.717	7.81	0.35*
NaCl	(mol)	5.152*	2.442	7.91	0.31*
NaCl	(rs)	7.199*	2.309	7.47	0.31*

$$U(r) = U_0 \left(e^{n(1-r/r_0)} - 2 e^{n(1-r/r_0)/2} \right) , \qquad (6)$$

where U_0 is the molecular binding energy, also called dissociation energy, r_0 the equilibrium bond length, and n a dimensionless number. Historically, Morse did not introduce the parameter n but rather $a = n/(2r_0)$, which is still commonly used today. Moreover, he expressed the energy relative to the ground state, $\Delta U(r) := U(r) - U_0$,

$$\Delta U(r) = U_0 \left(1 - e^{-a(r-r_0)} \right)^2 . \qquad (7)$$

To ease comparisons between different potentials and to yield $U(r \to \infty) = 0$, we choose forms similar to Eq. (6) throughout this article.
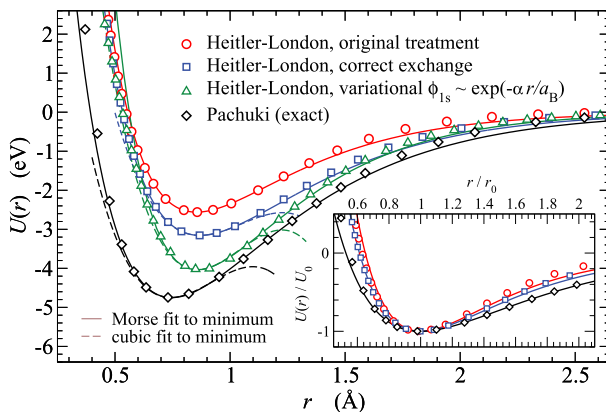
Figure 4. Binding energy assuming different quantum chemical approaches to the binding in H_2. Full lines are fits to that data near the minimum assuming the functional form proposed by Morse. Dashed lines represent cubic fits to the minimum. The inset shows the same data but normalized to the respective equilibrium bond lengths and binding energies.

Morse [6] recognized that the vibrational levels of many diatomic molecules, both homo- and hetero-nuclear, could be described quite accurately when treating U_0, r_0, and n as adjustable parameters. The parameter n allows the curvature of $U(r)$ in the minimum to be adjusted independently from the ratio U_0/r_0^2, however, no additional fine tuning of higher-order derivatives is possible. Given its simplicity, the Morse potential approximates reference data impressively well, as can be appreciated in Figure 3 for Cu_2 and C_2. It even represents the curves obtained for H_2 at different accuracy levels at which the hydrogen molecule is described and substantially better so than cubic fits to the energy minimum, see Figure 4.. Of course, there are also exceptions, where the Morse potential fails, e.g. Cr_2, which is described in Sect. 3.4.

An important result of the Heitler-London treatment is that short-range repulsion between atoms increases (approximately) exponentially with decreasing distance as two atoms approach each other. All highly-accurate potentials use repulsion that does not stray too far away from such a dependence as the pressure of (simple) bulk systems increases approximately exponentially at high compression with pressure for all bonding classes, at least as long as external pressures do not substantially exceed the bulk modulus B, which we will get back to in Sect. 6.1.

The simplest generalization of the Morse potential is a double exponential [62] of the form

$$U(r) = \frac{U_0}{n-m}\left\{ m\, e^{n(1-r/r_0)} - n\, e^{m(1-r/r_0)} \right\}, \tag{8}$$

which reduces to Morse if the additional parameter m is set to $n/2$. As Abell [62] noted, "*this choice is based in part on analytical convenience, but also on the physical grounds that atomic orbitals decay exponentially with r*". However, these long-range asymptotics are irrelevant at very small separations between nuclei, as briefly discussed in Sect. 3.2.2.

3.2. Mie/Lennard-jones and Born-Mayer/Buckingham potentials

The so-called Lennard-Jones potential [8] is the arguably most used potential in molecular simulations. It describes the interactions between entities with closed-electron shells quite well. This includes not only interactions between noble gas atoms, but also the intra- and intermolecular forces between, say, two CH_2 repeat units in polyethylene as long as the two units are not directly connected by a covalent bond. The functional form of the regular LJ potential is

$$U(r) = U_0 \left\{ \left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right\} . \qquad (9)$$

Here U_0 is the binding energy and r_0 the equilibrium bond length. The more common way to write the LJ potential is

$$U(r) = 4\epsilon \left\{ \left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right\} , \qquad (10)$$

where $\epsilon := U_0$ and $\sigma := r_0 / \sqrt[6]{2}$. However, comparison to other potentials and analytical manipulations are more easily done when starting from Eq. (9).

The attractive $1/r^6$ term is the leading-order dispersive interaction. The exponent can be motivated rigorously from perturbation theory, as sketched in Sect. 4.3. In contrast, the exponent from the repulsive $1/r^{12}$ term is a mixture of convenience and good luck. It results from the squaring of the $1/r^6$ term, which was particularly beneficial at times when numerics was not yet done by machines but by humans. This facilitated the work of those who were dismissively called *Rechenknechte* by theoretical physics professors in Germany in the beginning of the 20'th century. The term translates to *computing or arithmetic servants* and is nowadays used for computers. However, when treating the exponent as a fit parameter in a generalized Lennard-Jones potential

$$U(r) = \frac{U_0}{n-6} \left\{ 6 \left(\frac{r_0}{r} \right)^n - n \left(\frac{r_0}{r} \right)^6 \right\} \text{ with } n > 6 , \qquad (11)$$

the exponent n generally turns out close to its generic value of $n = 12$.

Replacing the exponent 6 in Eq. (11) with a number $0 < m < n$ yields the Mie potential [80]. The regular LJ potential can therefore be regarded as a limiting case of Mie's functional form, a so-called 12–6 Mie potential. In contrast, it may not be entirely appropriate to downgrade a Mie potential with arbitrary exponents as a generalized LJ potential, since Mie's work [80] predated those of Lennard-Jones [8,81] by more than two decades and in fact, it is not clear to us who was the first to use $n = 12$ as exponent in the repulsion. Jones, later known as Lennard-Jones, assumed it to be $n = 15$ [81] for argon.

Rather than tweaking the exponent of a Mie potential, it is probably more meaningful to replace it with an expression that somehow accounts for the Pauli repulsion between closed electron shells, whose density decreases as an exponential function of the interatomic distance. Slater [82] was the first to identify such a dependence for helium dimers by using a similar approach as Heitler and London [78]. The functional form he identified as being most appropriate for a limited range of interatomic distances was a single exponential, as in the Morse potential. Later, Born and Mayer [83] used the same functional form for the description of repulsion in simple ionic crystals without further theoretical justification but merely by arguing that it fits experimental results better than an inverse power as in Mie repulsion [83]. Buckingham [7] made a similar observation for noble-gas solids. Using exponential repulsion and $1/r^m$ attraction then leads to the two-body potential

$$U(r) = \frac{U_0}{n-m} \left\{ m \exp\left[n\left(1 - \frac{r}{r_0} \right) \right] - n \left(\frac{r_0}{r} \right)^m \right\}, \tag{12}$$

which is often called Buckingham potential ($m = 6$) but also Born-Mayer potential ($m = 1$) for two ions of dislike charges, in which case the potential parameters must be constrained to satisfy

$$\frac{n U_0 r_0}{n-1} = \frac{Q^2}{4\pi\varepsilon_0}, \tag{13}$$

where Q is the charge of one ion and ε_0 the vacuum permittivity. Born-Mayer potentials may also contain dispersive interactions in addition to exponential repulsion and Coulomb potentials. For two like ions, there is no bound state and repulsion is dominated by Coulomb interactions at typical interionic distances.

Later, Buckingham [84] found that the repulsion between hydrogen atoms in the lowest triple state and that between helium atoms are better described when replacing the prefactor of the exponential term with an appropriate polynomial. He motivated this with theoretical results obtained using Heitler-London type approaches, which is in line with more recent

comparisons [85]. Thus, the term Buckingham potential may also refer to a potential, in which the constant prefactor to the exponential repulsion is replaced with a polynomial in *r*. Van Vleet *et al.* [86] provide a contemporary discussion on how to construct the polynomials.

3.2.1. Higher-order dispersion and induction

Dispersive interactions are not limited to the leading-order dipole-dipole interactions, yielding the $1/r^6$ attraction. For example, dipole-quadrupole interactions cause a correction of

$$\Delta U(r) = -\frac{C_8}{r^8} \ , \tag{14}$$

while the quadrupole-quadrupole and the dipole-octopole interactions lead to corrections that asymptotically scale as $1/r^{10}$. Cipcigan *et al.* [87] recently summarized the hierarchy of dispersive interactions together with those resulting from electrostatic induction.

The most important correction to Coulomb interactions in ionic systems results from the large electrostatic, dipolar polarizability of anions, which was first considered by Rittner [88]. The energy gained by a dipole placed in an external electrostatic field $\mathbf{E}$ is $\Delta U_1 = -\mathbf{p} \cdot \mathbf{E}$, while the on-site energy required to create it is $\Delta U_2 = p^2/(2\alpha)$, where α is the polarizability of the anion. The dipole adjusts itself so that $\Delta U_1 + \Delta U_2$ is minimized. This yields a correction of

$$\Delta U(r) = -\frac{\alpha' Q^2}{8\pi\varepsilon_0 \, r^4} \tag{15}$$

to the Born-Mayer potential applied to a heteronuclear, diatomic molecule. Assuming bare Coulomb interactions, the polarizability of the cation is assumed to be negligible compared to that of the anion and $\alpha' := \alpha/(4\pi\varepsilon_0)$ has the unit of volume. As already recognized by Rittner [88], corrections to the correction in Eq. (15), arise due to the induction of the cation and the mutual induction of cation and anion. In addition, electrostatic field gradients induce quadrupoles and higher-order derivatives higher-order multipoles. Their incorporation may necessitate damping of the interactions [89,90] or further fine tuning of the on-site interaction between different multipoles beyond linear response to ensure a systematic improvements of the potential.

Rittner [88] found that adding this type of correction allowed him to reproduce important molecular properties of alkali halide molecules while assuming integer charges on the ions. He deemed Pauling's criterion for the fraction of ionic character of a bond (dipole moment divided by bond length times elementary charge) false, because it fails to include the far-from-

negligible polarization deformation on the ions. Madden and Wilson [91] provided much ammunition in favor of Rittner's conclusion by also considering crystalline structures of many ionic binaries.

It is important to note that the Rittner and related higher-order corrections are *not* pairwise additive. This is because an additional charge or charge distribution would change the induced dipole and thereby its interaction with the first charge. As a consequence, polarizability in many-atom systems is generally not solved in closed form, as described in more detail in Sect. 4.2.

3.2.2. Damping interactions at short distances

Obviously, higher-order dispersion and induction only matters significantly at small interatomic distances at which point the charge distributions of the interacting atoms start to overlap, which leads to a reduction or damping of the interactions. To prevent artificially large attractions from occurring at small separations, pertinent terms of the potentials are multiplied with damping functions $f_n(r)$. They are generally constructed such that the overall potential goes linearly to zero with r at small r, while the damping function quickly approaches unity with increasing r. To this end, Tang and Toennies [89] proposed to replace a dispersive interaction scaling with $1/r^n$ according to

$$\frac{1}{r^n} \to \frac{1}{r^n} f_n(r/b) \ , \tag{16}$$

where b is a parameter which depends on the nature of the two interacting atoms and potentially also on the index n, while

$$f_n(x) = 1 - \frac{\Gamma(n+1, x)}{n!} \ , \tag{17}$$

$\Gamma(n+1, x)$ being the incomplete Gamma function, which, in the case of a non-negative integer n, can be expressed as

$$\Gamma(n+1, x) = n! \, e^{-x} \sum_{k=0}^{n} \frac{x^k}{k!} \ . \tag{18}$$

The Tang and Toennies damping function is supposedly the most widely used one.

Not only dispersive but also regular Coulomb interactions between charges or dipoles can be damped, rewind, *should* be damped with functions like that defined in Eq. (16) to reflect the finite extent of electronic shells and to avoid grossly exaggerated attraction at small separation. In fact, as early as

1962, Dalgarno [92] discussed damping, or, "shielding" as it was called at the time, in the framework of shell potentials accounting for atomic polarizability.

3.2.3. Electrostatic screening at large distances

When atoms approach each other much more closely than their typical nearest-neighbor spacing, as it can happen during high-energy ion bombardment, it is more appropriate to assume the bare Coulomb repulsion, $U_C(r) = Z_1 Z_2 e^2 / (4\varepsilon_0 r)$ between the nuclei as asymptotic reference rather than the exponential repulsion reflecting the large-r asymptotics. Here, $Z_{1,2}$ are the nuclear charges of the interacting atoms or ions. Due to the propensity of matter to be locally charge neutral, electrostatic interactions are screened at "large" distances, i.e. at distances approaching equilibrium bond lengths, as is also the case, for example, between charged colloids in an electrolyte. To lowest order, the bare Coulomb interaction in a charge-neutral system is screened with an exponential $f(r) = \exp(-r/\lambda)$, where λ is called the screening length. The resulting potential $U(r) = f(r)U_C(r)$ is known as the Yukawa potential [93].

Ziegler, Biersack, and Littmark (ZBL) [94] proposed an empirical screening function $f(r) = \phi(r/\lambda)$ with

$$\phi(x) = 0.1818e^{-3.2x} + 0.5099e^{-0.9423x} + 0.2802e^{-0.4029x} \\ + 0.02817e^{-0.2016x} , \tag{19}$$

where only the screening length $\lambda = 0.46850$ Å$/(Z_1^{0.23} + Z_2^{0.23})$ depends explicitly on the nuclear charges. ZBL may still be the most used short-range repulsion, despite reported improvements [95], which, however, lack the mentioning of specific functional forms.

In practice, it is necessary to switch between short and long-distance asymptotics to predict repulsion at intermediate r. This is commonly done through switching functions as discussed in the appendix of Ref. [96].

3.3. Combining rules for two-body potentials

Bragg [97] observed that the lattice constants of many crystals can be reproduced quite accurately when atomic radii are assigned to individual elements and the assumption is made that two nearest-neighbors touch in ideal crystalline structures. This observation implies no rigorous but an approximate constraint for how (two-body) potentials parameterized for individual elements can be combined to mixed interactions. An arsenal of propositions was made in that regard. The simplest and most wide spread for the Lennard-Jones potential are the Lorentz-Berthelot [98,99] rules. They take the arithmetic mean of the

length scale parameter σ, whereby Lorentz predated Bragg's observations by 40 years, and, the geometric mean of the binding energies ϵ [98,99], expressed here using ϵ rather than U_0,

$$\sigma_{AB} = \frac{\sigma_A + \sigma_B}{2} \tag{20}$$

$$\epsilon_{AB} = \sqrt{\epsilon_A \epsilon_B} \; . \tag{21}$$

The major flaw of Lorentz-Berthelot is that it generally violates a rigorous result for the mixed dispersion coefficient, $C_{6,AB} := 4\epsilon_{AB}/\sigma_{AB}^6$, which is given further below in Eq. (50). An apparently reasonable approximation to Eq. (50) for combined dispersion coefficient reads [100]

$$C_{6,AB} \approx \frac{2\,\alpha_A\,\alpha_B\,C_{6,AA}\,C_{6,BB}}{\alpha_B^2 C_{6,AA} + \alpha_A^2 C_{6,BB}} \; , \tag{22}$$

where α_X is the polarizability of atom or ion X, while the geometric mean $\sqrt{C_{6,AA}\,C_{6,BB}}$ merely provides an upper bound for $C_{6,AB}$ [100]. We abstain from reviewing combining rules not reproducing Eq. (22) or better motivated combining rules, other than Lorentz-Berthelot. Instead, we content ourselves noting that Tang and Toennis [101] found the potential depths and locations of hetero nuclear noble-gas molecules to be well reproduced when using their equations (4) and (5), which satisfy Eq. (22) from this work, as combining rules. We would expect that reasonable answers can also be obtained for closed-shell (united) atoms when using Eq. (22) in conjunction with the Lorentz rule, from where the prefactor for the $1/r^{12}$ repulsion can be deduced.

A better model for repulsion than that used by Mie [80] or Lennard-Jones [8] is the exponential repulsion from Born and Mayer [83], originally Slater [102], which we write here as

$$U_{ij}(r_{ij}) = A_{ij}\, e^{-r_{ij}/\sigma_{ij}} \; , \tag{23}$$

where A_{ij} typically lies in the range of a few to several dozen keV [103]. Values for A and σ pertaining to our selected reference structure can be read off from Table. 1 by associating A with $U_0 \exp(n)$ and σ with r_0/n. Assuming that repulsion originates from he overlap of exponentially decaying charge densities yields the harmonic means for the length scale,

$$\sigma_{ij}^{-1} = \frac{\sigma_{ii}^{-1} + \sigma_{jj}^{-1}}{2} \; . \tag{24}$$

A standard assumption, similar to the Lorentz-Berthelot rule, is to apply a geometric mean for the energy prefactors, $A_{ij} = \sqrt{A_{ii}A_{jj}}$. Abrahamson [103] compiled a list for $1/\sigma_{ii}$ and A_{ii} for all neutral elements up to the atomic number 105.

An alternative combining rule for the prefactors A_{ij} was proposed by Smith [104] having in mind atoms or ions with noble-gas electronic configuration. He argued that repulsion is an energy excess related to the deformation of the electron density caused by the Fermi principle. He treated that deformation in a way as if each of the two interacting atoms or ions were in contact with an infinitesimally thin but impenetrable wall, which is positioned a distance $r_{ij}\sigma_{ii}/(\sigma_{ii} + \sigma_{jj})$ from atom i. The excess energy associated with each side of the wall can then be calculated from the pair repulsion. This nonrigorous, yet plausible model leads to the combining rule

$$A_{ij} = \left(A_{ii} \frac{\sigma_{ij}}{\sigma_{ii}} \right)^{\sigma_{ii}/(2\sigma_{ij})} \left(A_{jj} \frac{\sigma_{ij}}{\sigma_{jj}} \right)^{\sigma_{jj}/(2\sigma_{ij})} . \tag{25}$$

An identical rule had already been found empirically by Gilbert [105] for alkali halide monomers a few years earlier. Besides proposing his semi-empirical explanation, Smith [104] added to that list various pairs consisting of noble gas atoms, metal ions, and halogen anions, each with a complete shell. Böhm and Ahlrichs [106] found the combining rule for A_{ij} in Eq. (25) to be more accurate than others and further complemented the list of investigated pair repulsions to diatomic molecules beyond those consisting of atoms or ions with closed electron shell. Recent work [86], which is motivated by the analysis of electronic-charge density overlap between two atoms, proposes altered combining rules, in which σ_{ij} is the geometric mean of σ_i and σ_j. In addition, more complicated mixing rules apply than those "derived" by Smith and the prefactors have a certain r_{ij} dependence.

Given its success, it might be in place to further comment on the arguments leading to Eq. (25). First, it explains quite naturally that repulsion is not necessarily a pair-wise additive quantity. The electron distribution of a first atom is more easily squished by a second atom if there is no third atom already squishing the maltreated electron cloud of the first atom. The poor electrons simply have no more volume of refuge. Thus, the more atoms squeeze against a central atom, the larger the central atom should appear to be. Table 1 reflects this trend for single-component systems not only for the equilibrium lengths but also for $\sigma = r_0/n$. It is particularly strong for the covalently bonded carbon atoms and still quite noticeable for copper. In fact, from own unpublished work on alkali metals, our impression is that the assumption of pairwise additive repulsion is particularly poor for small coordination numbers. Second, potentials adjusting atomic charges on the fly might want to include the effect that this charge has on the atomic size and thus on its repulsive interaction. Neither of these two points appears to have attracted the attention it deserves.

3.4. Funky two-body potentials

Simple, computationally lean potentials generally lead to simple behavior without producing the intricate dynamics of highly viscous liquids or the complex structure of amorphous solids. The phase diagram of Lennard-Jonesium, like that of noble gas atoms other than helium consists of a closed-packed structure at low temperature, a low-viscosity fluid phase in a narrow temperature range at (what would be characteristic for) ambient pressure, and the gas phase at large temperature. While nature manages to produce elements with extended liquid phases at moderate pressure and temperature, like mercury or gallium, they do not become highly viscous. Even binaries often do not form highly viscous fluids and instead either phase separate or crystallize upon cooling from the liquid phase. Notable exceptions to this rule at elevated temperature are SiO_2 and CuZr, which both resist crystallization upon cooling to a significant degree while being describable with relatively simple potentials, e.g. by the silica potential proposed by van Beest, Kramer, and van Santen (BKS) [107] or by the embedded-atom model (EAM) for CuZr [108–110]. Unfortunately, BKS necessitates long-range electrostatic interactions to be evaluated while EAM lacks pairwise additivity, even if it can be computed at the cost of order $\mathcal{O}(Z_{loc}N)$, like pairwise additive potentials, where Z_{loc} is the number of atoms within the cutoff radius and N the number of atoms.

To study processes taking place on times exceeding vibrational periods by several decades and to address *generic* questions typical for disordered solids or highly viscous fluids, it can be advantageous to work with potentials producing the observed macroscopic behavior while being computationally as lean as possible. Such potentials may not be representative of any real material, but, so the hope, produce the correct qualitative behavior for the right reason. In the worst case, they provide a model for what nature could be and thereby allow theories for the fracture of disordered solids [111,112], the supercooling of liquids [113,114] or the rigidity of glasses [115,116] to be tested. Moreover, for many of the frequently qualitative questions to be answered, it can be advisable to sacrifice accuracy in the interactions rather than to make compromises in system size or cooling rate. For example, the anomaly of the specific heat in a bulk-metallic glass former can have serious artifacts when the linear system size is not at least twice the density correlation length [117]. Likewise, determining the proper scaling for the vibrational density of states with frequency in amorphous solids requires the use of large system sizes and long simulation times [118–121] and thereby the use of lean potentials.

The common recipe to construct simple potentials keeping systems from crystallizing quickly is to build frustration into them. Dzugutov [122] achieved this by introducing a hump in the pair potential. Its functional form is given by

$$U(r) = A\left(r^{-m} - B\right)e^{c/(r-a)}\Theta(r-a) + Be^{d/(r-b)}\Theta(r-d) , \qquad (26)$$

where the parameters A, B, a, b, and m used for Figure 5 are those from the original work. The most important feature of the Dzugatov potential certainly is that the hump is located near next-nearest neighbors spacings in a closed-packed structure if nearest neighbors settle in the vicinity of the energy minimum. As a consequence, atoms like to adopt local icosahedral structures, which can arrange similar to the order found in quasi crystals lacking long-range order, though large system sizes and small cooling rates may be required to prevent the crystallization into a bcc structure [123].

For mixtures of particles, frustration can be encoded to some degree through artificial combining rules in Lennard-Jones binaries, as for example, by setting all $\sigma_{AA} = 1.2$ while choosing $\sigma_{AA} = 1.2$, $\sigma_{AB} = 1.1$, $\sigma_{BB} = 1$ [113]. This first attempt by Wahnström is quite prone to crystallization as revealed by Kob and Anderson [114]. In an attempt to mimic a Stillinger-Weber potential designed to mimic amorphous Ni $_{80}$ P $_{20}$ [124], they suggested a binary Lennard-Jones potential in which dislike atoms have a large binding energy while being incompatible in size: $\epsilon_{AA} = 1$, $\epsilon_{BB} = 0.5$, $\sigma_{AA} = 1.0$, and $\sigma_{AA} = 1.0$, $\sigma_{AB} = 0.8$, and $\sigma_{BB} = 0.88$ [114]. To further reduce the risk of having unnoticed or even worse noticeable crystalline
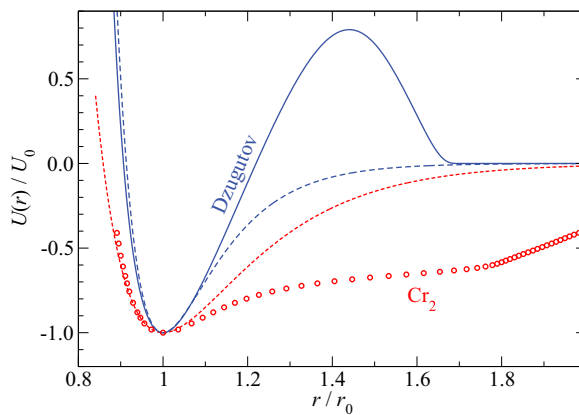

Figure 5. Two examples of funky pair potentials, $U(r)$, normalized to their binding energy U_0 and their bond distance r_0. The solid blue line represents a Dazugutov potential, which was constructed to make a mono-atomic system avoid crystallization. Circles depict experimental results on the pair potential of the chromium dimer, Cr_2. Dashed lines are Morse potential fits to the energy minimum.

reference phases, continuous distributions of Lennard-Jones radii can be used [115,116]. For a comparison of different glass-forming models, the reader is referred to a recent work by Ninarello, Berthier, and Coslovich [125].

We note in passing that pair potentials with humps similar to the one present in the Dzugutov potential are occasionally used to model mono-atomic bcc phases, in particular that of iron. This is because the third neighbor shell in bcc sits at $\sqrt{11/3} \approx 1.91$ times the radius of the nearest-neighbor shell and still $\sqrt{11/4} \approx 1.65$ times that of the second nearest-neighbor shell. Thus the first two shells comfortably fit into the potential well while the third shell has to pay no penalty. In contrast to bcc, the next-nearest neighbor shell in fcc or hcp would be utmost unhappy.

Interestingly, real pair potentials of bcc-forming metals can deviate substantially from the Morse potential, as is the case, for example for the chromium dimer, Cr_2, whose pair potential is included in Figure 5 and which appears to be challenging to compute accurately even with the most advanced post-Hartree Fock methods [126]. Despite the funky shape of "exact" pair potentials of some transition metal dimers, they still do not allow quantitative predictions to be made for elemental bcc metals, because real interactions simply happen not to be pairwise additive. It would lead to a sometimes substantial overestimation of the shear modulus, whereby dislocations are artificially suppressed in crystalline materials (their energy depends predominantly on the shear elastic constants) and thus brittle fracture be enhanced. Pair potentials also lack the important effect that a (metallic) bond between two atoms is strengthened when at least one of the two reduces its coordination as it would happen in a propagating crack. This lack biases material behavior toward brittleness.

Finally, pair potentials with a hump can also be designed to mimic chemical reactions as they occur, for example, during polymeric chemical reactions of step growth [127] and also radical-reaction based chain-growth [128]. It remains a philosophical question if such potentials should be deemed reactive, since the notion of a chemical reaction traditionally requires electron transfer or hybridization changes resulting in altered pair interaction, as is induced, for example between two hydrogen atoms due to the presence of an oxygen atom. Pair potentials can mimic neither one by definition so that they do not classify as reactive, in our opinion. Nonetheless, a well-designed bump-potential can reproduce steric effects while reproducing the large energy barrier that needs to be overcome for two monomers to bond, whereby the complex interplay of boundary condition, chain relaxation, and chain growth can be studied.

3.5. Relating potentials and elastic properties

Elastic properties are among the most important properties of solids. Any potential should therefore be tested for its ability to reproduce the elastic tensor of crystalline references. Ideally, they include not only thermodynamically stable but also hypothetical structures or finite stress to enhance transferability. Non-existing reference data can be obtained from DFT or related methods. One of the problems surfacing quickly is that elastic constants are not uniquely defined, except at zero stress and temperature. Different definitions lead to deviations of similar order as the external stress [129–136]. This can matter, for example, under geophysical conditions, so that properly converting between different elastic tensors may be critical, which we will come back to briefly at the end of this section.

To set the stage for many of the calculations throughout this article, we briefly review central aspects of the theory of elasticity, while deriving the Cauchy relations, whose violation has been and remains central to the development of potentials. We attempted our summary to be more condensed than other texts, while being palatable for readers not familiar with the topic.

We assume a deformation of a crystal, in which atomic positions after deformation are $r_{n\alpha} = (\delta_{\alpha\beta} + u_{\alpha\beta}) r_{n\beta}^0 + \chi_{n\alpha}$. Here, atoms are enumerated by n, Cartesian components are represented by Greek letters, and the Einstein summation convention applies. The upper index 0 in $r_{n\alpha}^0$ denotes an (equilibrium) position in a reference structure, the tensor u is the macroscopic displacement gradient, while $\chi_{n\alpha}$ is a local atomic displacement. The deformation is affine if all $\chi_{n\alpha}$ vanish. The Eulerian strain tensor ε is the symmetrized displacement gradient, $\varepsilon_{\alpha\beta} = (u_{\alpha\beta} + u_{\beta\alpha})/2$. From this definition, the squared distance between atoms i and j, $S_{ij} := r_{ij\alpha}r_{ij\alpha}$, after an affine deformation can be written as

$$S_{ij} = \left(\delta_{\alpha\beta} + 2\eta_{\alpha\beta} \right) r_{ij\alpha}^0 r_{ij\beta}^0 \; , \tag{27}$$

where $\eta_{\alpha\beta} := (u_{\alpha\beta} + u_{\beta\alpha} + u_{\gamma\alpha}u_{\gamma\beta})/2 = \varepsilon_{\alpha\beta} + \varepsilon_{\alpha\gamma}\varepsilon_{\gamma\beta}/2$ is called the Lagrangian strain tensor [137]. Other strain tensors exist but are not relevant here.

Stress tensor elements are generally defined as the first derivative of a thermodynamic potential w.r.t. a strain tensor element and normalized to the volume of the reference, e.g.

$$\sigma_{\alpha\beta} := \frac{1}{v_{\mathrm{pa}}^0} \frac{\partial U_{\mathrm{pa}}}{\partial \eta_{\alpha\beta}} \bigg|_{\eta=0} \; , \tag{28}$$

is called a Cauchy stress. Here, we have normalized energies and volumes per atom (pa). Note that the evaluation of the derivative at $\eta = 0$ only means that the derivative is taken at the reference but not that the stress disappears. Other stress tensor definitions exist, but they only differ when evaluated at nonzero or finite strain. For example, Eq. (28) defines the second Piola-Kirchhoff stress, if evaluated at nonzero η.

For systems interacting through (central) pair potentials, $\sigma_{\alpha\beta}$ can be easily computed. To this end, we first reexpress $U_2(r_{ij})$ as $\tilde{U}_2(S_{ij})$ so that

$$\sigma_{\alpha\beta} = \frac{1}{v_{\mathrm{pa}}^0} \sum_j \tilde{U}'_2(S_{ij}) r_{ij\alpha}^0 r_{ij\beta}^0 \ , \tag{29}$$

where the prime in $\tilde{U}'$ indicates the first derivative. For simple lattices under isotropic stress, the summation can be sorted according to neighbor shells, where $s = 0$ denotes the nearest-neighbor shell, $s = 1$ the next nearest-neighbor shell, and so on. Thus,

$$\sigma_{\alpha\beta} = \frac{1}{2v_{\mathrm{pa}}^0} \sum_s a_s U'_2(a_s) v_s^{\alpha\beta} \ . \tag{30}$$

Here a_s is the distance between a "central" atom i and an atom j in shell s and $v_s^{\alpha\beta}$ is the second rank shell tensor, whose elements are defined as

$$v_s^{\alpha\beta} = \sum_{j \in \text{shell } s} \frac{r_{ij\alpha}}{a_s} \frac{r_{ij\beta}}{a_s} \ . \tag{31}$$

For sufficiently symmetric shell structures, the second-rank shell tensor simply turns out to be Z_s/D, where Z_s is the number of atoms in neighbor shell s and D the spatial dimension [138]. In static equilibrium, the external hydrostatic pressure is nothing but $p = -\sigma_{\alpha\alpha}/D$.

Elastic constants are defined as the change of stress with strain, i.e. as second-order derivative of a thermodynamic potential w.r.t. to strain. This implies that results differ at non-zero stress for Eularian and Lagrangian strain tensors, see also Eq. (38). Moreover, it matters if atomic reference positions, i.e. the Wykhoff positions of atoms in a unit cell are kept fixed or change with strain, as could be expressed by assuming the $\chi_{n\alpha}$ to depend on strain whenever applicable. Such structural non-affine relaxation can happen when atomic sites lack inversion symmetry, as when shearing a diamond structure, or, as an extreme example, amorphous materials [139].

In-silico, it is an easy matter to constrain all $\chi_{n\alpha}$ to zero, experimentally less so. Elastic constants defined for $\chi_{n\alpha} \equiv 0$ are marked with an upper zero. In that case, taking the derivative of the two sides in Eq. (29) simply yields

$$C^0_{\alpha\beta\gamma\delta} = \frac{2}{v^0_{pa}} \sum_j \tilde{U}''_2(S_{ij}) r^0_{ij\alpha} r^0_{ij\beta} r^0_{ij\gamma} r^0_{ij\delta} \ . \tag{32}$$

By introducing the fourth-rank shell tensor,

$$v^{\alpha\beta\gamma\delta}_s := \sum_{j\in s} \frac{r^0_{ij,\alpha}}{r^0_{ij}} \frac{r^0_{ij,\beta}}{r^0_{ij}} \frac{r^0_{ij,\gamma}}{r^0_{ij}} \frac{r^0_{ij,\delta}}{r^0_{ij}} \ . \tag{33}$$

and by using $4S^2 \tilde{U}''_2(S) = \{r^2 U''_2(r) - r U'_2(r)\}$, the elastic tensor reads

$$C^0_{\alpha\beta\gamma\delta} = \frac{1}{2v^0_{pa}} \sum_s \{a^2_s U''_2(a_s) - a_s U'_2(a_s)\} v^{\alpha\beta\gamma\delta}_s = \frac{1}{2v^0_{pa}} \sum_s a^2_s k(a_s) v^{\alpha\beta\gamma\delta}_s \ . \tag{34}$$

Structural relaxation always reduces the energy so that C^0_{44} of diamond would violate the Cauchy relation even more than already evidenced in Figure 2.

The elastic tensor is invariant w.r.t. to any permutation of its indices, including for example, $C_{1122} = C_{1212}$, or, in Voigt notation $C_{12} = C_{66}$. It also includes implicitly the recipe for what elastic tensor obeys the Cauchy relations at finite hydrostatic pressure p, namely the second-order derivative of the internal energy w.r.t. the Lagrangian strain. However, most simple crystals violate it as is obvious from Figure 2.

Equation (33) allows the elastic tensor of simple crystals obeying pair potentials to be computed in a straightforward fashion, in particular when knowing the fourth-rank shell tensor. Its independent components are compiled in Tab. 2 for selected lattices along with other information useful to quickly compute properties of simple crystals.

For short-range potentials only the first two shells contribute substantially to the elastic tensors. Using the information collected in this section so far, we find for fcc, bcc, and sc, using Voigt instead of tensor notation,

$$C_{11} \approx \frac{1}{a_0} \begin{cases} \sqrt{2}k_0 + 2\sqrt{2}k_1 & \text{(fcc)} \\ k_0/\sqrt{3} + \sqrt{3}k_1 & \text{(bcc)} \\ k_0 + 2k_1 & \text{(sc)} \end{cases} \tag{35a}$$

$$C_{12} \approx \frac{1}{a_0} \begin{cases} k_0/\sqrt{2} & \text{(fcc)} \\ k_0/\sqrt{3} & \text{(bcc)} \\ k_1 & \text{(sc).} \end{cases} \tag{35b}$$

with the effective shell spring constant

$$k_s := k(a_s) = U''(a_s) - U'(a_s)/a_s \ . \tag{36}$$

Table 2. Useful dimensionless numbers for selected crystal systems using standard orientation of the lattices: simple cubic (sc), face-centered cubic (fcc), body-centered cubic (bcc), diamond cubic (dc), and hexagonal close packed (hcp). Additional elements for hcp are $v_0^{1133}/Z_0 = 1/18$, $v_0^{3333}/Z_0 = 2/9$, and $v_0^{112}/Z_0 = \mp 1/(24\sqrt{3})$, the negative sign applies to A-layer atoms and the negative for B-layer atoms if B is shifted by $(1/2, 1/\sqrt{12}, \sqrt{2/3})$ w.r.t. A. In diamond, each atom on $(0,0,0)$ and $(1/4,1/4,1/4)$ contribute $v_0^{123}/Z_0 = 1/(3\sqrt{3})$ and $v_0^{123}/Z_0 = -1/(3\sqrt{3})$, respectively.

	Z_0	Z_1	a_1/a_0	vpa/a_0^3	v_0^{11}/Z_0	v_0^{1111}/Z_0	v_0^{1122}/Z_0
sc	6	12	$\sqrt{2}$	1	1/3	1/3	0
fcc	12	6	$\sqrt{2}$	$1/\sqrt{2}$	1/3	1/6	1/12
bcc	8	6	$2/\sqrt{3}$	$4/\sqrt{27}$	1/3	1/9	1/9
dc	4	12	$\sqrt{8/3}$	$8/\sqrt{27}$	1/3	1/9	1/9
hcp	12	6	$\sqrt{2}$	$1/\sqrt{2}$	1/3	5/24	5/72

Note that for fcc, bcc and sc, $C_{\alpha\beta} \equiv C_{\alpha\beta}^0$ since their primitive unit cell contains only a single atom.

To obtain the relation between different elastic tensor definitions, it is useful to deduce from the definition of the Lagrangian strain that

$$\frac{\partial \eta_{\alpha\beta}}{\partial \varepsilon_{\gamma\delta}} = \delta_{\alpha\gamma}\delta_{\beta\delta} + \frac{1}{2}\left(\delta_{\alpha\gamma}\varepsilon_{\delta\beta} + \varepsilon_{\alpha\gamma}\delta_{\beta\delta}\right) . \tag{37}$$

Thus, for a second-order derivative of an arbitrary function f, it follows that

$$\frac{\partial^2 f}{\partial \varepsilon_{\alpha\beta}\partial \varepsilon_{\gamma\delta}} = \frac{\partial^2 f}{\partial \eta_{\alpha\beta}\partial \eta_{\gamma\delta}} + \frac{\partial f}{\partial \eta_{\gamma\alpha}}\delta_{\beta\delta} . \tag{38}$$

The differentiation rule of Eq. (38) also allows one to determine the second-order derivative of the term pv/v_0, where p is a constant (external) pressure. Given the relative change of a volume element, $\Delta v/v_0 = \varepsilon_{\alpha\alpha} + \left(\varepsilon_{\alpha\alpha}\varepsilon_{\beta\beta} - \varepsilon_{\alpha\beta}\varepsilon_{\beta\alpha}\right)/2 + O(\varepsilon^3)$ and using Voigt notation, $\Delta C_{ij} := p(\partial^2 v/\partial \eta_i \partial \eta_j)/v_0$ can be deduced to be

$$\Delta C_{ij} = p\Delta_{ij} \tag{39}$$

with $\Delta_{ii} = -1$ for all i, $\Delta_{ij \neq i} = 1$ for i and j both ≤ 3, and else $\Delta_{ij} = 0$.

Finally, we note that for a solid to be stable, the elastic tensor must be positive definite, which leads to the Born stability criteria [140,141], e.g. $C_{11} > C_{12}$, $C_{12} > -C_{11}/2$, and $C_{44} > 0$ in cubic materials. At constant pressure, $C_{ij} + \Delta C_{ij}$ must be positive definite. For a more general discussion on how imposed boundary conditions affect mechanical stability, we refer to an instructive work by Wang *et al.* [135], who studied a simple model for gold.

4. Many-body potentials for closed-shell systems

This section focuses on the description of many-body effects in systems composed of atoms and ions in which the constituents can be said to have a closed electron shell. Paradigms are noble gases and alkali metal halides, however, many of the methods and insights conveyed here pertain to a broader context such as valence force fields. The central goal of this section is to highlight the role that atomic dipoles play in the interaction between atoms. These dipoles may originate from quantum mechanical ground-state fluctuations or be induced by an electrostatic field, which arises, for example, on a chlorine atom in rock salt due to a lattice distortion breaking the local inversion symmetry.

Incorporating many-body effects of atomic dipoles—as well as higher-order electrostatic multipoles, which are somewhat meant to be referred to implicitly whenever the term dipole is mentioned—can be achieved in many different ways. For neutral atoms, their effect can be incorporated into potentials in the spirit of the expansion of Eq. (1). However, as soon as ions linger around, it is more efficient to model the dipoles explicitly. This can be done either by placing a formal dipole on the atom, or, by introducing a Drude model, that is, by coupling the center of mass of the displaced electron shell with a spring to a nucleus, or, to a given interaction site in a molecule. Finally, the dipole can be treated either classically, or, quantum mechanically. The latter appears to be a promising route to model many-body dispersion quite accurately, albeit at a large computational cost. For reasons of completeness, we state that Drude models would better be named after Lorentz, since Drude [142] considered free electrons in a metal, while Lorentz [143] attached (dissipative) springs between electrons and nuclei to model the optical response of bound charges in insulating matter.

4.1. Explicit many-body dispersion

Axilrod and Teller [144] and independently Muto [145] (ATM) extended the second-order perturbative treatment of the quantum dipole fluctuations leading to the leading-order $1/r^6$ London forces to third-order perturbation theory. The resulting three-body ATM potential reads (for three like atoms)

$$U_{\mathrm{ATM}}(\{\mathbf{r}\}) = C_9 \sum_{i<j<k} \frac{1 + 3\cos\vartheta_{ijk}\cos\vartheta_{jik}\cos\vartheta_{kij}}{r_{ij}^3 r_{jk}^3 r_{ki}^3} \ , \qquad (40)$$

where the sum runs over all unique triangles formed by three atoms i-j-k. Interior angles of the triangle formed by the three atoms i-j-k are denoted by the Greek letter ϑ_{ijk}, where the first index denotes the atom at the apex of the

angle. The cosines can be computed straightforwardly from the individual bond lengths, $\cos \vartheta_{ijk} = (r_{ij}^2 + r_{ik}^2 - r_{jk}^2)/(2r_{ij}r_{ik})$. A frequently used approximation for (mixed) dispersion coefficients is obtained from geometric averages in atomic units, i.e. $C_{9,ABC} = \sqrt{C_{6,A}C_{6,B}C_{6,C}}/\sqrt{H}$, where A, B, C are element specific and H is 1 Hartree. More rigorous results can be deduced from the Casimir-Ponder integral as described in Refs. [146,147].

The usual course of action when using ATM in conjunction with LJ or Buckingham is to explore improvements on predicted properties. Traditionally, ATM corrections to the binding energy for noble gas crystals like argon are estimated to be of order 10% [49]. For very polarizable systems, they can be much larger. Von Lilienfeld and Tkatchenko [146] found them to account for 50% of the binding between two graphene sheets. In addition, ATM interactions can correct for systematic deficiencies in the elastic and vibrational properties that cannot be overcome in the realm of pair potentials [148]. We see it as beneficial to analyze U_{ATM} as an independent quantity in terms of a lattice ATM constant, which can be defined in analogy to the Madelung constant as

$$\alpha_{\mathrm{ATM}} := a_0^9 \sum_{1 < j < k} \frac{1 + 3\cos\gamma_1 \cos\gamma_j \cos\gamma_k}{r_{1j}^3 r_{jk}^3 r_{k1}^3} \tag{41}$$

We obtain $\alpha_{\mathrm{ATM}} \approx 57.548$ for an ideal fcc crystal from which one sixth can be assigned to each atom. For hcp, we obtain $\alpha_{\mathrm{ATM}} \approx 57.563$, which can topple the balance in favor of fcc. The differences between these two numbers is small, yet, larger than between the results for C_6, i.e. $C_6 \approx 14.4538$, $C_{12} \approx 12.1319$ for fcc, versus, $C_6 \approx 14.4545$, $C_{12} \approx 12.1323$. In order for α_{ATM} to be larger for hcp than for fcc, interactions beyond next-nearest neighbors must be included. Unfortunately, we did not manage to find recent literature results on the simple-to-compute ATM lattice sums beyond relatively rough, initial estimates by Axilrod [149]. Nonetheless, different authors [150,151] have previously concluded that three-body dispersion favors fcc over hcp, although the corrections due to nuclear quantum fluctuations appears to be substantially more important in that regard [152].

4.2. Classical polarizable potentials

The electrostatic field on a central atom or molecule is produced by other charges, dipoles, or higher-order multipoles. This leads to a hierarchy of inductive interactions [90], which we have already touched upon in Sect. 3.2.1. It contains the $1/r^4$ attraction between a charge and an induced dipole. Next in line is the (asymptotic) $1/r^6$ attraction between a permanent dipole

and an induced dipole or that between a charge and an induced quadrupole. Unfortunately, the interactions related to induced dipoles or induced higher-order multipoles, which will be ignored for the most part in the following, are not pairwise additive. Two like charges placed at $\pm \Delta \mathbf{r}$ from an atom will not induce a dipole on that atom and thus not lead to twice the energy gained if only one of the two charges were present.

Since adjacent existing (static, molecular) dipoles try to align themselves to an external electrostatic field in an attempt to minimize the potential energy, induced dipoles tend to be parallel to pre-existing static dipoles. This effect is well known to increase the mean dipole moment of water molecules from 1.85 D in the gas phase to approximately 2.7 D in the liquid, whereby dipole-dipole interactions essentially double.

The polarizability of a homogeneous medium having cubic or higher symmetry is stated in terms of its dielectric constant ε_r, which implicitly reflects the feedback that dipoles have on each other. In atomic or molecular systems, ε_r can be well estimated through the Clausius-Mossotti relation [153]

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \sum_i \frac{\rho_i \alpha_i}{3\varepsilon_0} \; ,$$ (42)

where ρ_i is the number density of species i and α_i its (orientationally averaged) polarizability. Once the right-hand side of Eq. (42) is greater or equal one, ε_r is no longer a finite positive number so that the model has reached its physically meaningful limit. The dipoles "want" to grow *ad infinitum*, at which point the system becomes metallic. The underlying polarizable potential needs to be extended to either suppress this so-called polarization catastrophe, e.g. by shielding the Coulomb interaction at small distances [89], or, to actually allow the system to become conducting. In advanced shell potentials being a compromise between polarizable and charge-transfer models, this can be achieved by not constraining an electron cloud to one particular atom [154].

When modeling charged or polar systems by atomistic means, one would certainly want to reproduce the dielectric response function of a medium correctly, e.g. when simulating the condensation of water on a surface [155], the Helmholtz double layer on an electrode [156,157], or the damped Coulomb interaction of charged colloids in water. For the simulation of simple geometries and homogeneous media, it may be possible to achieve this with effective potentials [158] or with the concept of induced mirror charges [159]. However, many important problems lack the symmetry or isotropy required to pursue such approaches so that the dielectric response has to be solved on the fly. This can be achieved with molecular approaches encoding the polarizability into the potential energy surface.

One possibility is to induce *ideal* dipoles or higher-order multipoles on atoms or specific interaction sites in a molecule [160,161]. This is done by adding an energy contribution for each inducible multipole $\mathbf{p}_i$. The usual interatomic potential $U(\{\mathbf{R}\})$ is the one that minimizes $U(\{\mathbf{R}\}, \{\mathbf{p}\})$, e.g.

$$U(\{\mathbf{R}\}) = \min_{\{\mathbf{p}\}} \ U(\{\mathbf{R}\}, \{\mathbf{p}\}), \tag{43}$$

where

$$U(\{\mathbf{R}\}, \{\mathbf{p}\}) = U_{\mathrm{sr}}(\{\mathbf{R}\}) + U_{\mathrm{C}}(\{\mathbf{R}\}, \{\mathbf{p}\}) + \sum_i \frac{p_i^2}{2\alpha_i} \cdot \tag{44}$$

with respect to the dipoles. In Eq. (44), contributions to the short-range (sr) interaction are separated from those due to Coulomb (C) interactions, which may contain contributions from static dipoles (for molecular rather than atomic simulation) in addition to those from point charges. Since, Eq. (44) is a second-order polynomial in the induced dipoles, a well-defined minimum exists as long as the Hessian related to the dipoles is positive definite. A brute-force inversion of the Hessian to yield the exact minimum is generally inadvisable. Alternatives are conventional minimization techniques, such as those based on conjugate gradients [162] or extended Lagrangians [163,164]. In the latter case, the dipoles, or other adjustable degrees of freedom such as those describing the shape of the periodically repeated system [163] or prefactors to electronic wavefunctions [164], are assigned an inertia and propagated and relaxed along with the atomic coordinates. It is beyond the scope of this review to discuss the pros and cons of extended Lagrangians in detail. It suffices to say that their implementation is relatively simple. However, they introduce an effective delay on the molecular dynamics [165]. In addition, relaxation to the energy minimum can be slow when the Hessian has strongly differing eigenvalues, which automatically happens for large ε_{r}. Then, the dipolar response functions are "stiff" at small length scales but "soft" at the continuum scale.

One disadvantage of placing *ideal* dipoles on atoms or ions is that the evaluation of Coulomb interactions is substantially complicated, even if solutions exist to include their effect into the (fast) Ewald summation [166]. Another deficiency is that higher-order multipoles are ignored that are generated when an electron cloud displaces with respect to a nucleus. These drawbacks can be remedied with Drude particles, or "Drudes", in which a fixed Drude charge q_{D} is coupled harmonically with a spring constant k_{D} to an atom or ion, to which a charge $-q_{\mathrm{D}}$ is added [167–169]. The last summand on the r.h.s. of Eq. (44) must then be replaced with $\sum_i k_{D,i} u_i^2 / 2$, where $\mathbf{u}_i$ is the displacement of one Drude charge w.r.t. the atom or interaction site that it is bonded to. On-site Coulomb interactions

between all charges on the Drude, which despite similar spelling is not to be confused with *The Dude* from *The Big Lebowski*, must be switched off. Minimization of the total energy w.r.t. to the Drude displacements can be done in a similar fashion as for ideal dipoles.

In order for the Drude to reproduce the correct polarizability, the relation

$$\alpha = q_{\mathrm{D}}^2/k_{\mathrm{D}} \tag{45}$$

must be obeyed. The limiting case of ideal dipoles is obtained for $q_{\mathrm{D}}, k_{\mathrm{D}} \to \infty$ while keeping α constant. With an appropriate choice of k_{D} and sign for q_{D}, the correct quadrupolar induction of a spherically symmetric Drude can be matched to a homogeneous field. Traditionally, the three independent parameters of Drude oscillators, which in addition to q_{D} and k_{D} are also assigned an inertia or mass m_{D}, were chosen to best match the frequency dependence of the dielectric constant at high frequencies [167]. Ideally, those values would be close to results obtained from a fit to accurate reference data on forces and energies, for which k_{D} and possibly even q_{D} are treated as adjustable parameters. Large discrepancies between different parametrization procedures should probably be seen as a sign that something is missing or wrong with the potential.

A stringent test for the correctness but also for the relevance of dipolar polarizability can be obtained when comparing force-field-based infrared absorption spectra to reliable reference data. For example, in the case of amorphous silica, obtaining the correct number, positions, and intensities of peaks requires the polarizability of the oxygen atoms to be included [170] (Figure 6a). Accounting for electrostatic induction also appears necessary to accurately reproduce bond-angle distributions in silica and related systems [171]. While global bond-angle histograms may seem fine for pair potentials, deviations between symmetry-specific angles in crystalline structures from the ideal tetrahedral bond angle have the wrong sign for the most commonly used SiO_2 pair potential [107] but the correct sign and magnitude for the polarizable Tangney-Scandolo potential [171,172].

The incorporation of dipolar polarizability is particularly crucial for structures in which highly polarizable (united) atoms or ions, most notably anions, are located in sites deviating strongly from inversion symmetry. This concerns in particular oxygen whenever it is two-coordinated, as it is in water, but also in low-temperature tetrahedral network formers like silica. If competing local orders exist, neglecting the polarizability of anions will artificially favor the order with the smaller deviation from inversion symmetry, because the anion is shorn of its ability to reduce its energy in the larger electrostatic field of the symmetry-broken phase.
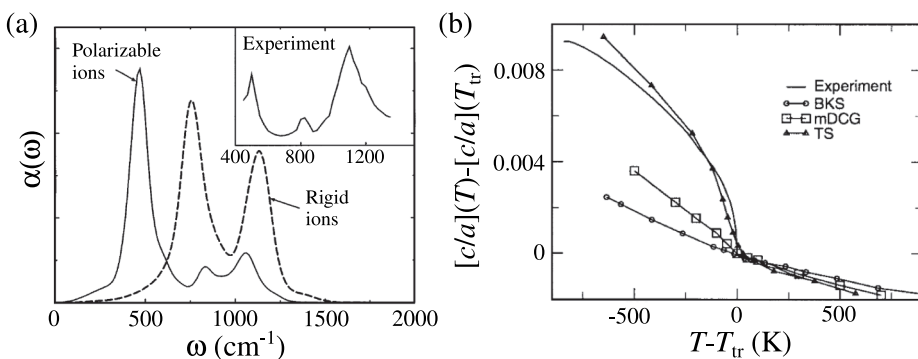
Figure 6. Influence of polarizability on properties of SiO_2. (a) Infrared absorption spectrum of amorphous SiO_2 (silica) computed one time with a rigid, fixed-charge ion model (dashed line) and one time with a dipol-polarizable anion (solid line). The polarizable model matches qualitatively the experimental results shown in the inset. (b) c/a ratio of crystalline SiO_2 (quartz) computed with the rigid ion potentials of van Beest, Kramer and van Santen (BKS) [107], the fluctuating-charge potential of Demiralp, Çağın and Goddard after modifying parameters (mDCG) [173], and the polarizable force field of Tangney and Scandolo (TS) [171]. Only the polarizable force field reproduces the c/a anomaly at the $α$-$β$-quartz transition with temperature T_{tr}. (a) is reprinted with permission from Wilson, M., Madden, P. A., Hemmati, M., and Angell, C. A. *Phys. Rev. Lett.* **77**, 4023 (1996) (Ref. [170]. Copyright (1996) by the American Physical Society. (b) is reprinted from Herzbach, D., Binder, K. and M. H. Müser, *J. Chem. Phys.* **123**, 124711 (2005) (Ref. [172]. with the permission of AIP Publishing.

Including dipolar polarizability can also be required to describe macroscopic structural changes occurring during the transition between high- and low-symmetry phases in a *qualitatively* correct fashion, as is the case for the $α − β$ transition in quartz. Using pair potentials, the c/a ratio shows no anomaly as a function of temperature at the transition temperature. For charge-transfer potentials, the slope of the c/a ratio changes at the transition temperature, while it is discontinuous, as in experiment, when dipolar polarizability on oxygen is added to the pair potential [172] (Figure 6b). This discussion may show that the analysis of phase transformations can benefit the validation of potentials, most notably polarizable potentials, even if applications will evolve primarily around liquid or amorphous phases.

An affine deformation of a highly symmetric ionic crystal, as for example, rocksalt (B1) or cesium chloride (B2), does not lead to an electrostatic field on a central lattice position created by the atoms residing on other positions. Thus, the central atom does not develop a dipole. However, a non-isotropic deformation, e.g. a uniaxial strain, can reduce the cubic symmetry, whereby the charge density on the anion may develop a quadrupole moment. Consequently, unlike dipolar polarizability, quadrupolar polarizability can induce a Cauchy violation in the said crystals, which is why it can be said to be more important than dipolar polarizability, at least from a continuum-mechanics point of view. Madden, Wilson, and coworkers demonstrated that inclusion of quadrupolar polarizability can substantially increase the

agreement between experimental and simulation results, at least in the case of simple ionic systems like AgCl [161] or MgO [174]. An important aspect of their work [174] is that they managed to gauge adjustable parameters independently from one another, even those describing a charge-density change on an anion due to electrostatic polarization from those being caused by short-range repulsion.

4.3. Quantum-Drude oscillators

The starting point for the description of dispersive interactions are two, or more, isolated atoms, described by their free-atom Hamiltonians $\hat{H}_0^{(1,2)}$. The atoms interact through their dipoles $\mathbf{p}_{1,2}$ via $\{\mathbf{p}_1 \cdot \mathbf{p}_2 r_{12}^2 - 3(\mathbf{p}_1 \cdot \mathbf{r}_{12})(\mathbf{p}_2 \cdot \mathbf{r}_{12})\}/(4\pi\varepsilon_0 r_{12}^5)$ plus potentially through higher-order multipoles. Second-order perturbation theory then leads to a pairwise-additive $1/r_{12}^6$ interactions between the atoms.

Asymptotic $1/r^6$ interactions are also obtained when replacing the atomic Hamiltonians with quantized Drude oscillators [90,175,176], as will be shown here below. Their three, free parameters per atom, q_D, k_D, and m_D can be chosen to match the correct polarizability $\alpha_\mathrm{D} = q_\mathrm{D}^2/k_\mathrm{D}$ and leading-order dispersive coefficient C_6, which allows dispersive and inductive interactions to be treated on a common footing. The justified hope is that higher-order and many-body dispersive interactions can be mimicked reasonably well simultaneously by properly selecting the third remaining parameter [87]. If the nucleus is also quantized, m_D must be replaced with a reduced mass μ_D.

The Hamiltonians of two identical, three-dimensional quantum Drudes can be decoupled into three pairs with

$$\hat{H}_k = \sum_{i=1}^{2} \left(\frac{\hat{p}_{ki}^2}{2\mu_\mathrm{D}} + \frac{k_\mathrm{D}}{2} u_{ki}^2 \right) + \frac{g_\alpha}{4\pi\varepsilon_0} \frac{q_\mathrm{D}^2 u_{k1} u_{k2}}{r_{k12}^3} . \tag{46}$$

The parameters g_k ($k = 1, \ldots, 3$, no summation convention) take the values 1 or -2 depending on whether the displacements $u_{1,2}$ are orthogonal or parallel to $\mathbf{r}_{12}$. The oscillators can be decoupled further through the transformation $u_{k\pm} = (u_{k1} \pm u_{k2})/\sqrt{2}$ leading to three isolated oscillators pairs described by Hamiltonians of the form

$$\hat{H}_{k\pm} = \frac{\hat{p}_\pm^2}{2\mu_\mathrm{D}} + \frac{1}{2} \left(k_\mathrm{D} \pm \frac{g_k}{4\pi\varepsilon_0} \frac{q_\mathrm{D}^2}{r_{k12}^3} \right) u_{k\pm}^2 . \tag{47}$$

In leading order, their combined excess ground-state energy w.r.t. that of two uncoupled oscillators, $U_0 = \hbar\omega_\mathrm{D}/2$ with $\omega_\mathrm{D} = \sqrt{k_\mathrm{D}/\mu_\mathrm{D}}$ per free quantum Drude pair is

$$\Delta U_{k0}(r_{12}) = -\frac{\hbar\omega_{\mathrm{D}}}{8}\left(\frac{g_k q_{\mathrm{D}}^2}{4\pi\varepsilon_0 k_{\mathrm{D}}}\right)^2 \frac{1}{r_{12}^6} \ . \tag{48}$$

Adding up the three Drude pairs, using $\alpha' = q_{\mathrm{D}}^2/(4\pi\varepsilon_0 k_{\mathrm{D}})$, and repeating the entire exercise for potentially dislike Drudes, A and B, then yields,

$$C_{6,AB} = \frac{3\hbar}{2}\frac{\alpha'_A \alpha'_B \omega_A \omega_B}{\omega_A + \omega_B}, \tag{49}$$

which is the same combination rule as Eq. (22). Using the true atomic reference Hamiltonians and placing the dimer on the z-axis, the correct mixed dispersion coefficient reads [177]

$$C_{6,AB} = 6\left(\frac{e}{4\pi\varepsilon_0}\right)^2 \sum_{n_A,n_B}{}' \frac{|z_{0\,n_A}|^2 |z_{0\,n_B}|^2}{E_{n_A}^{(A)} - E_0^{(A)} + E_{n_B}^{(B)} - E_0^{(B)}} \ , \tag{50}$$

where $z_{0\,n}$ is the matrix element $\langle 0|z|n\rangle$ of an atom and where the primed sum indicates that at least one of the two quantum numbers n_A and n_B must differ from zero.

Typical parametrizations of quantum Drudes [90] yield $\hbar\omega_{\mathrm{D}}$ of order 0.5 H to 1 H for hydrogen and noble gas atoms as well as for small, closed-shell molecules, q_{D}/e in between 0.7 and 1.4, and m_{D}/m_e from 0.1 to 0.6. Since a regular quantum Drude is fully defined by three parameters, many terms related to higher-order polarizability α_l ($l = 1$, dipole, 2 = quadrupole, 3 = octupole) are constrained to take fixed ratios. Some of them are given by [87]

$$\sqrt{\frac{20}{9}}\frac{\alpha_2}{\sqrt{\alpha_1\alpha_3}} = 1\,(\text{polarization}) \tag{51a}$$

$$\sqrt{\frac{49}{40}}\frac{C_8}{\sqrt{C_6 C_{10}}} = 1\,(\text{dispersion}) \tag{51b}$$

$$\frac{C_6 \alpha_1}{4 C_9} = 1\,(\text{mixed}) \ . \tag{51c}$$

Figure 7 reveals that these ratios tend to reproduce experimental results within 10% error.

Once a quantum Drude model is defined, various strategies exist to deduce the energy and interatomic forces following from it. Brute-force diagonalization is not advisable for reasons of computational efficiency but also because the model is no longer quadratic in the Drude displacements, as soon as they seize to be negligible compared to interatomic distances, i.e. at typical condensed-phase, nearest-neighbor spacings. Moreover, any solution strategy should allow for the possibility to replace the harmonic
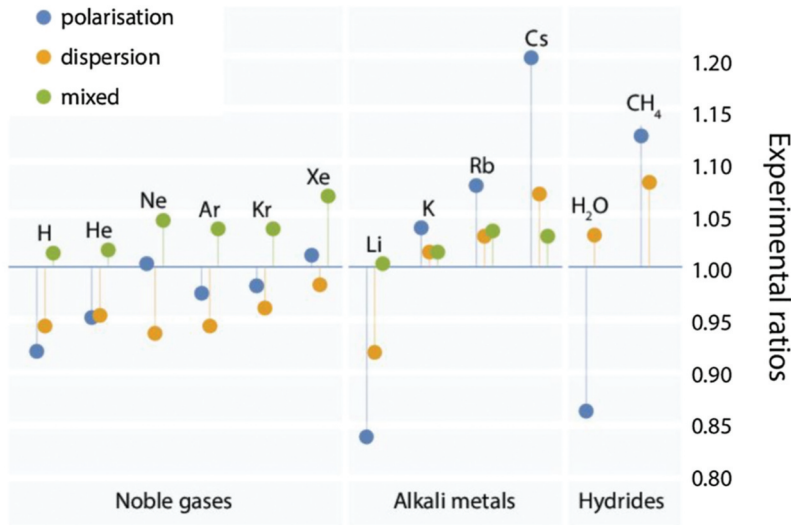
Figure 7. Experimental values for the left-hand sides of Eq. (51). Reprinted from *J. Comput. Phys.*, 326, Cipcigan, F. S., Sokhan V. P., Crain, J., and Martyna, G. J., Electronic coarse graining enhances the predictive power of molecular simulation allowing challenges in water physics to be addressed, 222–233, Copyright (2016), with permission from Elsevier.

coupling with one in the spirit of finitely-extensible non-linear coupling, as to potentially match simultaneously more than just one inductive or dispersive coefficient in addition to α and C_6.

A common strategy to handle quantum Drudes is the use of path-integral techniques [178–181]. The trouble with quantum Drudes is that the number of system replicas to be simulated is of order one Rydberg divided by thermal energy, which is a few hundred at room temperature. One can get away with slightly fewer replicas with diffusion Monte Carlo [182], which is closely related to path integrals, in conjunction with a diagrammatic expansion of interacting Drudes [183]. Unfortunately, improvements do not seem such that numerical costs fall below an approach, in which induction is handled classically and ATM interactions are included explicitly. This turns the quantum Drude oscillator approach into something like a sleeping beauty, which could be awaken if a handsome prince managed to cut down the computational overhead w.r.t. classical polarizable potentials plus two-body like short-range potentials to a factor of order ten or less.

4.4. Shell models and many-body repulsion

The interaction of atoms results in the deformation of (closed) electronic shells beyond the induction of electric multipoles through electrostatic fields, reflected, for example, by an approximately rigid translation of the valence shell with respect to the remaining ion. A reduction of the Coulomb

interaction between charges and dipoles due to the overlap of electronic shells can be described with damping functions, which were introduced in Sect. 3.2.2. The electrostatic field, its gradients but also the Fermi principle and hence repulsion deforms the electronic density compared to the superposition of the free-ion or free-atom references. These deformations were first approximated as being spherical in the so-called breathing-shell model [184] and described on a common footing with dipolar polarizability [167]. Later, non-spherical shell deformations were also considered with deformable-shell [185], de-formable-ion [161], or distortable-ion [174] potentials. Variables describing the state of deformation are the ionic radius, or rather, the deviation δ_i from the radius of the free ion, plus variables describing the shape of the deformation. In atomistic simulations, the shell variables, including the induced dipoles, are assumed to minimize the potential energy.

The starting point of much of the original literature [48,184,185] assumes a harmonic expansion of the potential energy in terms of small variables, i.e. displacements, dipoles, and breathing modes. However, for general situations, the potentials are better cast in terms of implicit many-body potentials, the ingredients of which are the repulsive interactions and on-site energy penalties for the deformation of the shell with respect to free ions. In such models, the short-range repulsion can be expressed, for example, in a breathing-shell potential, through an expression of the type [174,186,187]

$$U_{\mathrm{sr}} = \sum_i U_{\mathrm{bre},i} \cosh(\delta_i/\sigma_i) + \sum_{i,j>i} U_{\mathrm{rep},ij} e^{-(r+\delta_i+\delta_j)/(b_i+b_j)}. \tag{52}$$

Here, $U_{\mathrm{bre},i}$, σ_i, $U_{\mathrm{rep},ij}$, and b_i are constant coefficients, while the δ_i minimize U_{sr}. The first summand on the r.h.s. of Eq. (52) reflects the on-site coupling in a heuristic fashion, i.e. it can be parameterized to yield the correct shell stiffness at small δ_i while suppressing embarrassing ion shrinkage at large compression.

Accounting for anisotropic shell deformation requires tensors of rank one and higher to be included as arguments in the functions appearing on the r.h.s. of Eq. (52) such that they reduce to a scalar in a meaningful way, i.e. without violating the isotropy of space. Anisotropic shell distortions seem to be required to simultaneously improve phonon spectra and the binding energy differences of competing ionic structures compared to rigid-shell potentials merely reflecting dipolar polarizability [184,188]. We note in passing that aspherical atoms modeled with potentials assuming fixed bonding topography has been pioneered by Price, Stone, and coworkers [189]. Deviations from spherical symmetry then stems predominantly from intra-molecular interactions so that shape parameters can be treated as fixed relative to a molecular coordinate system.

A central motivation for introducing flexible shell models was to reproduce the experimentally observed violation of the Cauchy relation in simple salts and their phonon dispersion at high symmetry points [48,184]. Here, it might be worth observing that the Cauchy pressures tend to be minor for alkali halides but quite substantial for alkaline earth oxides like MgO. One central difference between these two classes of simple salts is that a free, singly-charged halogen anion is stable while a free, doubly charged chalcogen anion is not. Thus, assuming doubly charged ionic references without accounting for charge transfer upon a change in bond length could be difficult to justify.

Interestingly, charge-transfer and breathing-shell models have a similar effect on the Cauchy pressure. When the application of a stress in x-direction induces ions to shrink in all spatial directions, the external pressure required to maintain the crystal shape in the y or z direction is reduced, which implies that C_{12} is reduced compared to a rigid-shell model with fixed or zero δ_i. In contrast, when shearing, for example, a rock-salt structure, the nearest-neighbor bond lengths only change in order $\mathcal{O}(\varepsilon_{44}^2)$ so that the shear modulus is the same as that of a rigid-shell potential. In other words, $C_{44} > C_{12}$, is also unavoidably obtained for charge-transfer potentials applied to rocksalt crystals, irrespective of whether neutral atoms or ions are taken as reference [190]. Figure 8 pictures the just-made arguments for shell models schematically and explains qualitatively how anisotropic shell models can lead to a Cauchy pressure of either sign. Closed form expressions for the elastic constants in breathing-shell models are summarized in Ref. [191], which also contains a critical comparison of different breathing-shell potentials.

Ion shrinkage under compressive stress without charge transfer and the unavoidable sublinear scaling of repulsion with the coordination number Z_0 in breathing-shell models strikes us as implausible. It certainly violates the superlinear scaling of repulsion with Z_0 obtained in a first-order DFT-based perturbation theory, in which the kinetic energy density of the electrons increases with $\rho^{5/3}$, where ρ is the electronic density. Despite this argument being merely qualitative, it should reflect the proper trend as the trend is very clear. Our own quick and dirty analysis of the EOS of MgO in the NaCl and CsCl structures using exponential repulsion plus Coulomb interaction and the respective Madelung constants makes us believe that trends are complicated. Although the compressive part of the EOS can be described quite well with the most simple Born Mayer potential, the fitted charges turn out close to unity in both cases while Bader analysis finds more intuitive charges close to 1.7. When using the parameters deduced for the NaCl and CsCl-structure for the dimer, differences in the Coulomb energy turn out larger than for the repulsion. At the same time, the Bader charges barely change with the lattice constants in contradiction to our suspicion that
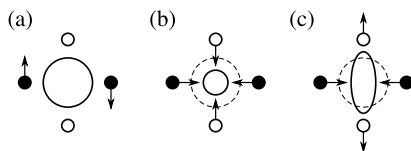
Figure 8. Deformation of an electron shell in response to a deformation of the macroscopic body. The electron shell is schematically depicted by the central circle. (a) During shear deformation, the displacement does not lead to a deformation of the shell. There is hence no influence of shell deformation on the shear modulus C_{44}. (b) If the shell deforms isotropically, then an inwards motion of the black atoms will lead to an inwards force on the white atoms, implying $C_{12} < C_{44}$. (c) For anisotropic shell deformation, an inwards motion of the black atoms will lead to an outwards force on the white atoms, leading to $C_{12} > C_{44}$. Adapted from Ref. [192].

breathing-shell models reflect charge-transfer effects to low order. Nonetheless, short-range repulsion certainly induces non-spherical distortions leading to electrostatic multipoles, which then must be included in the overall electrostatic interactions between atoms [161].

5. Many-body potentials for open-shell systems

By definition, open-shell atoms have an incomplete valence shell. Bonding between them occurs through the formation of either covalent, or, in condensed phases, also through metallic bonds. Interactions between open-shell atoms cannot be faithfully described with pair potentials without producing false trends, some of which are discussed in Sect. 2.3. The arguably most important many-body effect in open-shell systems is the weakening of a bond due to the presence of additional atoms. An extreme example is the onset of repulsion between hydrogen atoms, which attract each other as free radicals but repel each other in the presence of oxygen. A more subtle effect is the contraction of layers near unpassivated metal surfaces, where missing neighbors of the atoms in the surface layer strengthen their interaction with atoms in the layer underneath.

Figure 9 reveals this weakening for copper and carbon in different crystalline structures, including hypothetical structures. The energy per atom $U_{\mathrm{pa}}(a_0)$ as a function of nearest-neighbor bond length (Figure 9a – copper, Figure 9e – carbon) was obtained from the materials scientist's favorite electronic structure method, DFT [193,194] within the local density approximation [194] using projector-augmented waves [195]. The minimum of these curves is the cohesive energy of the crystal $U_{\mathrm{coh}} = U_{\mathrm{pa}}(a_0^{\mathrm{eq}})$, where a_0^{eq} is the equilibrium bond length shown in Figure 9b for copper and in Figure 9f for carbon. a_0^{eq} is approximately a logarithmic function of the coordination number Z_0 for both copper and carbon. The energy per bond in the equilibrium configuration, $U_{\mathrm{pb}}^{\mathrm{eq}}(Z_0) = 2U_{\mathrm{coh}}(Z_0)/Z_0$, clearly decreases with Z_0 and much more so for

carbon (once $Z_0 \geq 3$) than for copper. For both elements, the decrease of $U_{pb}^{eq}(Z_0)$ is approximately algebraic in Z_0, albeit with a steeper power law for carbon than for copper.

The results presented in Figure 9 imply that pair potentials cannot describe the energetics of these systems, although each individual $U_{pa}(a_0)$ at fixed Z_0 can be reproduced quite accurately using a regular Morse potential as evidenced by the solid lines in panels (a) and (e) of Figure 9. As Abell [62] emphasized, environment-dependent parameters are needed. Various families of many-body potentials for open-shell systems have been developed over the years to reflect these trends and, of course, more subtleties. In this section, we will outline their functional forms and the properties of the solids that they describe. We will also motivate some of them from quantum-mechanical considerations.

A commonality of the most popular, simple open-shell potentials, namely EAM, second-moment tight-binding expansion (TB2M), Stillinger-Weber (SW), and Tersoff, which will all be introduced in this section, is that their total potential energy can be written as a *cluster functional* [19]
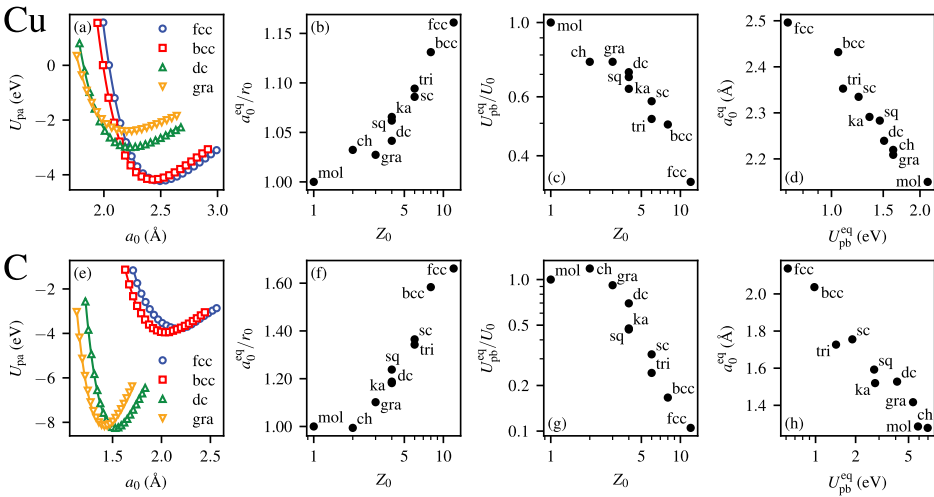

Figure 9. (a,e) Energy per atom U_{pa} in selected crystal structures as a function of the nearest-neighbor spacing a_0. Solid lines are fits of the data to a standard Morse potential assuming nearest-neighbor interactions. (b,f) Equilibrium bond length a_0^{eq} as a function of coordination number Z_0. (c,g) Equilibrium binding energy per bond U_{pb}^{eq} as a function of Z_0. (d,h) Equilibrium bond length a_0^{eq} as a function of equilibrium binding energy per bond U_{pb}^{eq} (Pauling plot). Crystal structures range from 1D to 3D include hypothetical ones and are abbreviated as follows: dimer/molecule (mol), chain (ch), graphene (gra), triangular lattice (tri), kagome lattice (ka), square lattice (sq), diamond cubic (dc), simple cubic (sc), body-centered cubic (bcc) and face-centered cubic (fcc). The top row shows results for copper, while the bottom row shows results for carbon. U_0 and r_0 are the binding energy and bond length of the dimer. Data, which were produced using consistent DFT-LDA methodology, are merely meant to convey trends.

$$U(\{\mathbf{r}\}) = \frac{1}{2}\sum_{ij}\left\{U_2(r_{ij}) + U_m(r_{ij}, \xi_{ij})\right\}, \tag{53}$$

where $U_2(r_{ij})$ is *a* pairwise additive interaction—to be distinguished from *the* pair potential—and $U_m(r_{ij}, \xi_{ij})$ is an effective pair interaction, which, however, depends on a bond-environment variable ξ_{ij}, which is not necessarily symmetric in the indices and may contain a non-zero on-site term ξ_{ii}. Its specific interpretation changes from one potential class to the next but it always depends on a sum over third atoms through

$$\xi_{ij} = \sum_k \Xi(r_{ij}, r_{ik}, r_{jk}) \ . \tag{54}$$

In a limiting case, $\Xi(r_{ij}, r_{ik}, r_{jk})$ simply corresponds to the third term, $U_3(r_{ij}, r_{ik}, r_{jk})$, of the expansion given in Eq. (2). However, $U_m(r_{ij}, \xi_{ij})$ can also implicitly include higher-order terms. The way how to express the various potentials in the generic form of Eq. (53) and (54) are detailed in Sect. 5.6. Those summaries may be useful to design prototype templates for functions or to work out elastic properties falling into this general category.

5.1. Density-dependent potentials

The alkali metals can be seen as a polar opposite to closed-shell systems, because their electrons are delocalized to the extent that their dispersion relation is close to that of free electrons up to the Fermi wave vector k_F, albeit with an effective mass. This is because there is only one electron per atom in the conduction band so that only its minimum is sampled. For this reason, early attempts of describing bonding in (alkali) metals assume the Jellium model [75,196,197], which expresses the energy of an electron gas as a function of its density $\rho = N/V$, where in the case of (neutral) alkali metals the number of atoms N is identical to that of valence electrons. The Jellium model is the DFT-based approach to matter, in which the main effect of ions is assumed to provide a charge-neutralizing background to the electrons without explicitly accounting for the relative positions between ions. Interestingly, the Jellium model leads to densities and bulk moduli that are relatively close to those of crystals formed by light alkali metals, even when using the true electronic mass [196]. However, Jellium does not stiffen sufficiently much with increasing density when the Coulomb repulsion between the ionic cores starts to matter.

The potential of the form [198,199],

$$U(\{\mathbf{r}_i\}) = Nu(\rho) + \frac{1}{2}\sum_{i,j} U_2(r_{ij}), \tag{55}$$

can certainly be motivated from the Jellium model. In this potential, which could be coined a global-glue potential, the Jellium model is augmented with a pair-wise additive repulsion, e.g. in the form of damped Coulomb interactions. In principle, not only the pair repulsion but also $u(\rho)$ can be made a function of the element, or, elements under consideration. While Eq. (55) can be generalized to non-homogeneous phases and alloys, e.g. via the construction of Wigner-Seitz cells and the request of local charge neutrality, surfaces would cause problems, since surface atoms "own" excessive volume, which makes a local density difficult to define. Thus, it is clear that potentials footing on Eq. (55), or, simple generalizations thereof, are not practicable. It is yet utmost instructive, to study their properties and further reasons for their failure.

Assuming interactions past the nearest-neighbor shell to be screened, see Sect. 9, severe restrictions for the elastic constants of simple solids ensue. The glue effectively acts as a positive, external pressure, p_g, squeezing the ions together. Combining Eqs. (35) and (39) then allows the elastic tensor elements of elementary fcc and bcc crystals at zero external stress to be approximated (assuming a smoothly changing p_g with density) with

$$
\begin{aligned}
C_{11} &= 2\sqrt{2}C_0 - p_g & C_{12} &= \sqrt{2}C_0 + p_g & C_{44} &= \sqrt{2}C_0 - p_g & \text{fcc} \\
C_{11} &= 2/\sqrt{3}C_0 - p_g & C_{12} &= 2/\sqrt{3}C_0 + p_g & C_{44} &= 2/\sqrt{3}C_0 - p_g & \text{bcc,}
\end{aligned}
$$

$$(56)$$

where C_0 would be an element-specific constant. Equation (56) predicts $C_{44} < C_{12}$. Check mark for simple metals! For fcc, the sum rule $c_{sr} := (C_{11} + C_{12})/(C_{12} + C_{44}) = 3/2$ can be read off. It is obeyed reasonably well by our favorite metal, $c_{sr}(\text{Cu}) \approx 1.46$ and other fcc metals, $c_{sr}(\text{Pb}) \approx 1.56$, $c_{sr}(\text{Ni}) \approx 1.46$, $c_{sr}(\text{Pd}) \approx 1.66$, $c_{sr}(\text{Ag}) \approx 1.54$, but less so by the favorite metals of our wives $c_{sr}(\text{Au}) \approx 1.73$ and $c_{sr}(\text{Pt}) \approx 1.85$.

Unfortunately, bcc turns out mechanically unstable, because the pertinent elastic tensor is not positive definite owing to $C_{12} > C_{11}$ for positive glue pressures. This inequality can be easily understood. The global glue attempts to increase density. This is done most effectively when nuclei keep a maximum mutual distance, which the most closed-packed structures achieve the best. But why then do alkali metals condense into bcc? Before addressing that question, we note that C_{11}/C_{12} tends to be quite close to unity for simple bcc alkaline metals but not for bcc metals with an approximately half-filled d-shell. Specifically, $C_{11}/C_{12} = 1.2$ within 2% for the alkali metals from lithium to cesium, while transition metals assume much larger ratios. For example, 1.95 (V), 4.37 (Cr), 2.60 (W), and 1.76 (Fe). Thus, density-dependent potentials fail distinctly more for transition metals than for alkali metals. Also, most alkali metals transform into fcc at relative moderate pressures less than 10 GPa, thus, check mark again, before they actually undergo a series of additional phase transformations at even higher pressures [200]. Thus, despite being highly

flawed, Eq. (55) contains a few elements of truth for alkali metals. However, the existence of non-closed-packed equilibrium structures at large compression is also a clear indication that the assumption of pairwise repulsion may not be particularly accurate.

In a seminal work, in which the term *electron correlation* appears to have been introduced, Wigner [196] demonstrated that free electrons at small density can lower their energy compared to the cheapest, simple constant-density solution of fermions, which is spanned by the product of two Slater determinants, one for spin up and one for spin down, in which all k-states are filled up to k_F. At electron densities characteristic for alkali metals and using second-order perturbation theory, he found them to condense into a bcc crystal. One possible conclusion is that a reliable potential for alkali metals should implicitly reflect higher-order gradients in the electronic charge density. Stabilizing the right phase for the wrong reason, e.g. with potentials merely depending on (estimates for) the local electronic charge density and/or by tweaking cut-off distances or functional dependencies, as done, for example by some funky potentials in Sect. 3.4, cannot lead to accurate, transferable potentials.

5.2. Embedded-atom method and second-moment tight-binding potentials

A simple generalization of Eq. (55) is

$$U(\{\mathbf{r}_i\}) = \frac{1}{2} \sum_{i,j} U_2(r_{ij}) + \sum_i F(\rho_i) \, , \qquad (57)$$

where ρ_i would be an estimate for the electronic charge density near atom i. Equation (57) forms the basis of a class of potentials, which has two names, embedded-atom method and tight-binding (TB) approximated to second-moment (2M) expansion. It has been historically pursued by two communities. The TB2M and EAM camps differ predominantly in how they motivate and later generalize $F(\rho_i)$ for alloys or encode directional bonding, which will be sketched further below. Yet, neither camp keeps the interpretation of ρ_i as the charge density of *all* valence or conduction-band electrons at or averaged over the vicinity of $\mathbf{r}_i$ and assumes instead

$$\rho_i = \sum_{j \neq i} f(r_{ij}), \qquad (58)$$

which excludes the contribution of the valence shell of the atom i itself. The function $f_j(r_{ij})$ is the square of a bond-integral in the TB interpretation and the charge density from atom j seen by atom i for the EAM camp. However, as Finnis and Sinclair [47] rightfully note in their pioneering study: "*the consequences of the form of the model [...] does not depend on the physical interpretation.*"

Ducastelle [46] appears to have been first to suggested Eq. (57) building on earlier work by Cyrot-Lackmann [201] and Friedel [202] on TB2M approaches to metals. He also proposed what could be seen as the most generic functional form of an EAM or TB2M potential, simple exponential repulsion and furthermore

$$F(\rho_i) = -A\sqrt{\rho_i} \tag{59a}$$

$$\rho_i = \sum_{j \neq i} e^{-r_{ij}/\sigma_a} \,. \tag{59b}$$

For alloys, A and σ_a depend on the atom type. Moreover, atom (EAM) or bond (TB2M) specific prefactors must be added to each summand on the r.h.s. of Eq. (59b). The central arguments leading to the square-root dependence of $F(\rho)$ are sketched in Sect. 5.8. The main reason for the exponential dependence is, as always, the exponential (Slater) type of the atomic orbitals, even though the full TB bond integrals have polynomial prefactors, see, e.g. Ref. [203]. This functional form was optimized for a wide variety of metals and alloys by Cleri and Rosato [204].

Recast into our generalized Morse form, Eq. (8), the potential energy expression becomes

$$U(\{r_i\}) = \frac{1}{2} \frac{U_0}{m-n} \sum_i \left\{ \sum_{j \neq i} m e^{n(1-r_{ij}/r_0)} - \left[\sum_{j \neq i} n^2 e^{2m(1-r_{ij}/r_0)} \right]^{1/2} \right\} \,. \tag{60}$$

We would like to argue that the usefulness of this expression for rationalizing trends in metallic bonding (see Sect. 6) is on par with that of the Lennard-Jones potential for noble gas atoms, which is why we feel that calling systems described by this potential *Ducastellium* is as appropriate as using the well-established term Lennard-Jonesium.

Historically, the EAM/TB2M potential described by Eq. (57) was independently discovered several times after Ducastelle's work. For example, Gupta [64] showed that it reproduces the lattice relaxation of metals near surfaces. Tománek, Mukherjee, and Bennemann revealed its appropriateness to describe the energetics of small metal clusters [205] as well as surface and vacancy energies of transition metals [206]. In thermodynamics, a formulation identical to EAM/TB2M has been employed by Pagonabarraga and Frenkel for coarse-grained particle dynamics calculations, where $F(\rho_i)$ becomes a free-energy that is adjusted to the equation of state of a liquid [207,208].

Equation (57) can also be motivated from the quasi atom theory of Stott and Zaremba [209] and from the effective-medium approximation theory of Nørskov and Lang [210,211]. In their approaches, the energy gained when

embedding an atom i into a given site $\mathbf{r}_i$, e.g. an interstitial or vacancy site, is argued to be a functional of the electronic density at the embedding site and thereby to depend in leading order on the charge density ρ_i that exists at $\mathbf{r}_i$ *before* the atom is embedded. Daw and Baskes [9,45] built on these ideas and approximated the charge density ρ_i at the embedding site $\mathbf{r}_i$ as a superposition of the charge density from neighboring atoms. They coined the term *embedded-atom method* and provided important insight as to what extent many-body terms, as well as hydrogen, affect materials behavior including ductility. Ercolessi, Parrinello and Tosatti renamed EAM into *glue potential* [212–214] while claiming that their specific realization of an EAM potential "*accounts for all known lattice properties of Au*" [212].

Modern EAM potentials use more complex functional forms than the simple exponentials of Ducastelle. Those are out of scope for the discussion in this review and the interested reader is referred to the original literature, see for example works by Mishin and coworkers [215–219]. In the spirit of early works by Foiles, Daw, and Baskes [220], parameterizations of such EAM potentials are distributed using tabulated data for $F(\rho)$, $f(r)$ and the pair-potential $U_2(r)$.[1] Care has to be taken when interpolating between these data points in the final implementation of a potential, as the choice of spline order is not to be cast aside as a technicality, because it affects the properties of the potential [221].

In our own experience on copper [222], the simple form first used by Ducastelle [46] and Gupta [64] performs best when testing structures with coordination numbers varying systematically from $Z = 2$ to $Z = 12$. The reason for this trend might be, as so often, that simple, physically well motivated functional forms are less prone to overfitting than elaborate functions having been tweaked to enforce right numbers for selected properties in one or few structures. To this we wish to add that we highly doubt the pair-additive repulsion in open-shell systems to be an accurate approximation. Thus, there is little reason to fiddle around with the 4′th digit of an embedding function, if "bond-order effects" on repulsive forces lead to errors in the second digit.

5.3. Modified-embedded-atom-method potentials

A deficiency of EAM potentials is their generic preference for closed-packed structures. While funky parametrizations allow bcc to be stabilized, we refer yet again to Wigner's work on alkali metals in Sect. 5.1 and repeat our claim that stabilizing the right phase for the wrong reason cannot yield a robust potential. To better encode directional bonding in EAM, Baskes [223,224], having silicon in mind, suggested to augment the computation of the electronic density with an angular three-body term. Baskes already hinted [224] that

his explicit expression can be interpreted as a dependence of the embedding function on the gradient and higher-order spatial derivatives of the density, i.e. to replace $F(\rho)$ with $F(\rho, \partial_\alpha \rho, \partial_\alpha \partial_\beta \rho, \ldots)$, where ∂_α is short-hand notation for $\partial/\partial r_\alpha$. Shortly after, Baskes [225,226] suggested to replace (the estimate for) the embedding density according to $\rho_i \to \bar{\rho}_i$ with

$$\bar{\rho}_i^{\,2} = \sum_l t_i^{(l)} \left(\rho_i^{(l)} \right)^2, \tag{61}$$

the first four $\rho_i^{(l)}$ satisfying

$$\rho_i^{(0)} = \sum_j f_j^{(0)}(r_{ij}) \tag{62a}$$

$$\rho_{i\alpha}^{(1)} = \sum_j n_{ij}^\alpha f_j^{(1)}(r_{ij}) \tag{62b}$$

$$\left\{ \rho_i^{(2)} \right\}^2 = \left\{ \sum_j n_{ij}^\alpha n_{ij}^\beta f_j^{(2)}(r_{ij}) \right\}^2 - \frac{1}{3} \left\{ \sum_j f_j^{(2)}(r_{ij}) \right\}^2 \tag{62c}$$

$$\rho_{i\alpha\beta\gamma}^{(3)} = \sum_j n_{ij}^\alpha n_{ij}^\beta n_{ij}^\gamma f_j^{(3)}(r_{ij}) \tag{62d}$$

where $t_i^{(l)}$ are weighting coefficients, $\mathbf{n}_{ij}$ is a unit vector parallel to bond i-j, and Einstein summation convention is implied on the Cartesian indices for the squared quantities. Here, the $f_j^{(l)}(r_{ij})$ reflect generalized, partial background electron density from atom j seen by atom i. Through this generalization, semi-explicit angular dependencies result, as for example, through the squaring of the $\rho_i^{(1)}$ term. It leads to a summand of the form $n_{ij}^\alpha n_{ik}^\alpha f_j^{(1)}(r_{ij}) f_k^{(1)}(r_{ik})$, which is proportional to $\cos \vartheta_{ijk}$. With each higher-order term, higher-order sinusoidal dependencies on bonding angles result. An appealing aspect of these modifications is that modified EAM (MEAM) potentials pick up information on the local order beyond the coordination number and germanium contain the square of the left-hand side of Eq. (62d). Disappears for any mirror-inversion crystal but not for the diamond lattice. Thus, MEAM-potentials for silicon and germanium contain an expression energetically favoring large absolute values of ρ_i^3.

We now argue how a MEAM potential can stabilize open crystal lattices. To this end we consider the expression

$$g_i = \sum_{j,k \neq i,j} \left\{ \cos \vartheta_0 - n_{ij}^\alpha n_{ik}^\alpha \right\}^2 f_j(r_{ij}) f_k(r_{ik}) \tag{63}$$

$$= \rho_i^2 \cos^2 \vartheta_0 - 2\rho_{i\alpha}\rho_{i\alpha} \cos \vartheta_0 + \rho_{i\alpha\beta}\rho_{i\alpha\beta} - (1 - \cos \vartheta_0)^2 \sum_j f_j^2(r_{ij}), \quad (64)$$

where the upper index on the partial embedding charge densities was omitted. g_i clearly has a minimum at the preferred bond angle ϑ_0, since $n_{ij}^\alpha n_{ik}^\alpha$ is the cosine of the angle j-i-k. Now let us proceed with an ideal crystal structure, in which only partial densities from nearest neighbors substantially contribute to the shell. In the given approximation, we can replace $f(r_{ij})$ with ρ/Z_0 so that after executing the restricted double sum correctly

$$g_i \approx \rho^2 \left(\cos^2 \vartheta_0 - 2 \cos \vartheta_0 v_0^\alpha v_0^\alpha + v_0^{\alpha\beta} v_0^{\alpha\beta} \right) - \frac{\rho^2}{Z_0} (1 - \cos \vartheta_0)^2 . \quad (65)$$

Assigning an energy penalty on g_i allows the degeneracy of different crystalline structures with identical coordination number but different bond angles to be lifted within a nearest-shell approximation. At the same time, open structures with small Z_0 and correct bond angles can be favored over closed-packed structures. It is clear that Eq. (65) can be constructed (with the exception of the additive constant) from a linear combination of the MEAM invariants like those given in Eq. (62).

Although MEAM potentials, similar in spirit to the ones just described, are routinely used in molecular simulations, it is easily argued that these MEAMs have issues. First, there is no physically motivated reason for the existence of the upper indices in the $f_j^{(l)}$ coefficients when justifying the modification from Baskes' argument that the embedding function should also depend on derivatives of the embedding density. Second, within this picture, the embedding function can generally depend on scalars that can be constructed from derivatives, which, however, must be constructed such that they are invariant w.r.t. a rotation of the coordinate system. There are many more scalars of a similar order as those contained in Eq. (62) [222]. For example, if $\rho_{\alpha_1 \ldots \alpha_l}(\mathbf{r}) := \partial_{\alpha_1} \ldots \partial_{\alpha_l} \rho(\mathbf{r})$, then in addition to $\rho_{\alpha\beta}^2$ there can be a contribution of $\rho_{\alpha\alpha}\rho_{\beta\beta}$, which, in principle, can be linearly independent of the former, rather than to be constrained to $\rho_{\alpha\beta}^2 - \rho_{\alpha\alpha}\rho_{\beta\beta}/3$, which can be loosely associated with the term of Eq. (62c). Yet another invariant, which can be constructed with four Cartesian indices and a square in the densities, would be $\rho_\alpha \rho_{\alpha\beta\beta}$. Thus, to flesh out our second point of criticism, the $l = 2$ and $l = 3$ terms in Eq. (62c) are far from complete and it is not clear if the most relevant ones have been considered.

The "expansion" of Eq. (62c) truncates at the third-order term, which is the lowest-order invariant that can discriminate between nearest-neighbor-shell energies of fcc and hcp. However, different cubic structures cannot be distinguished with shell tensors of rank three or less. Fourth, the coefficients $t_i^{(l)}$ can depend, in principle, on

$\rho_i^{(0)}$. It might be possible to obtain some dependencies through asymptotic, dimensional, or other systematic analyses. Guessing them correctly appears an almost impossible task. Thus, we feel that the already impressive performance of MEAMs should be further improvable. However, one potentially important ingredient missing is the environment dependence of the functions $f_j(r_{ij})$ beyond screening, e.g. an environment-dependence of σ_a. Lifting this restriction would allow mediated interaction to take place as they occur, for example, in breathing shell models or in TB potentials going beyond the second-moment approximation, albeit at a potentially much increased computational cost.

5.4. Stillinger-Weber potential and extensions

The binding energy in non-funky pair-, EAM, and TB2M potentials is largest when atoms are most closely packed. However, the three lightest group-14 elements (C, Si, and Ge) are (meta-) stable in the diamond lattice, which has a packing fraction less than 50% of that in fcc or hcp. To stabilize the *open* diamond lattice, Stillinger and Weber [227] proposed to add a rather simple angle-dependent term to a Mie pair potential, which, however, was multiplied with the rather clever cut-off function

$$f_c(r_{ij}, \gamma, r_c) = \exp\left[\gamma/(r_{ij} - r_c)\right] \Theta(r_c - r_{ij}) .\qquad(66)$$

It has the nice property that all derivatives continuously approach zero at the cut-off distance r_c. The angular add-on consisted of a penalty quadratic in the deviation from tetrahedral bonding on a given atom i, as originally proposed by Keating [228]. The extra-term is cut off when the length of one of the two bonds, r_{ij} and r_{ik}, forming a bond angle of ϑ_{ijk} on atom i, approaches r_c. Put together in the notation used throughout this article, the SW potential reads

$$\begin{aligned}
U(\{\mathbf{r}\}) = &\sum_{i,j>i} \frac{U_0}{n-m} \left\{ m\left(\frac{r_0}{r_{ij}}\right)^n - n\left(\frac{r_0}{r_{ij}}\right)^m \right\} f_c(r_{ij}, \gamma_1, r_c) \\
&+ \sum_{i,j\neq i, k\neq i,j} U_t \left\{ \cos\vartheta_{ijk} - \cos\vartheta_0 \right\}^2 f_c(r_{ij}, \gamma_2, r_c) f_c(r_{ik}, \gamma_2, r_c)
\end{aligned}$$

$$(67)$$

with $\cos\vartheta_0 = -1/3$ for the tetrahedral angle. Besides this angle, the SW potential has eight independently adjustable parameters: U_0, r_0, m, n, γ_1, r_c for the pair potential, and in addition U_t, γ_2 for the tetrahedral bonding part.

Although the SW potential is occasionally used to also model carbon, it should be kept in mind that the original SW potential does not recognize graphite as being energetically favorable at ambient pressure.

Despite its simplicity and empirical nature, the SW potential can be parametrized for silicon to reproduce (by construction) its elastic properties in the diamond structure, but also fairly well some high-pressure phases, and the melting temperature at ambient pressure, including the anomalous jump from small to large density during melting [227,229,230]. The latter goes hand in hand with a coordination change from $Z_0 = 4$ in the diamond structure to $Z \gtrsim 6$ in the liquid [231].

An appealing aspect of the SW potential is that the competition between tetrahedral and dense packing favored by the angular and the pair part, respectively, can be tuned through the ratio U_t/U_0 [232]. By increasing it from a value characteristic for germanium through that of silicon, it is possible to make the disordered phase adopt a local tetrahedral order at large U_t/U_0, which explains why SW is occasionally used for the united-atom modeling of water [233]. Using small U_t/U_0, SW can describe condensed phases of tin, including that condensing in the β-tin structure [234], which can be seen as a compromise between tetrahedral and close packing.

Of course, as Stillinger and Weber [227] admitted, their model has quantitative deficiencies. They were revealed most clearly by Biswas and Hamann [235], who demonstrated that the equation of state of any crystalline structure other than the diamond structure is highly flawed. However, even for the diamond structure, the deficiency of the SW can become qualitative. For example, SW fails to reproduce that silicon is brittle under tension, although SW can be used to mimic brittle (non-silicon) tetrahedral solids, see Refs. [236,237]. This weakness is shared by other empirical potentials and related to the finite range of the cutoff, which will be discussed in more detail in Sect. 9.

The SW potential formally looks like a cluster potential. Given that the three-body term has a zero-energy contribution in the ideal diamond structure, it must be concluded that the pair interaction is an effective $Z_0 = 4$ pair potential rather than the "true" pair potential. Improvements of the SW address this issue partially by augmenting SW with an environment dependence in terms of a coordination-dependent equilibrium bond angle ϑ_0. Such environment-dependent interaction potentials (EDIPs) were first developed for silicon [238,239] and later for other elements including carbon [240,241]. Due to its ability to reflect the quasi-degeneracy of graphite and diamond, an EDIP for carbon was the first empirical interatomic potential that correctly predicted the relation of 3-fold (graphite-like, sp^2-hybridized) and 4-fold (diamond-like, sp^3-hydridized) atoms in amorphous carbon obtained from liquid quenches at varying densities [241].

Another interesting generalization of the SW potential, which Biswas and Haman [235] proposed to model bond-angle energetics on atom i beyond the Keating model, reads

$$U_3(r_{ij}, r_{ik}, r_{jk}) = \sum_l C_l F_l(r_{ij}, r_{ik}) P_l(\cos \vartheta_l), \qquad (68)$$

where C_l are expansion coefficients, $P_l(\cos \vartheta)$ Legendre polynomials, and the $F_l(r_{ij}, r_{ik})$ functions, which can be defined to reproduce the exact or any general three-body potential. When assuming them to factorize as $F_l(r_{ij}, r_{ik}) = \phi_l(r_{ij})\phi_l(r_{ik})$, e.g. with simple exponential dependencies, Biswas and Hamann demonstrated the computation of U_3 to be reducible to order Z_{loc}, where Z_{loc} is the number of atoms within the interaction range of atom i, while irreducible potentials, such as the ATM potential, require a number of operations proportional to Z_{loc}^2.

5.5. Tersoff potentials

Dissatisfied by the poor transferability of the SW potential [235] and inspired by Abell's analysis of the sensitivity of bond strengths on the local environment [62], Tersoff presented a series of papers [242–246], in which he introduced the concept of bond order to the world of empirical potentials. According to Pauling [247], the bond order is the difference of the number of electrons in bonding and anti-bonding orbitals. It is a monotonically decreasing function of the coordination number. Tersoff's work was an attempt to construct functions that measure (effective) coordination numbers and thereby the bond order. To do so, he introduced a cut-off function, which is unity up to a distance r_{c1} meant to include nearest neighbors and which then quickly falls off to zero at a distance r_{r2} supposedly less than typical next-nearest neighbor distances:

$$f_c(r) = \Theta(r_{c1} - r) + \frac{\Theta(r - r_{c1})\Theta(r_{c2} - r)}{2} \left\{ 1 + \cos\left(\pi \frac{r - r_{c1}}{r_{c2} - r_{c1}}\right) \right\} . \quad (69)$$

With the help of this cut-off function, the bonds to atom i other than the ij bond can be characterized with

$$\xi_{ij} = \sum_{k \neq i,j} g(\cos \vartheta_{ijk}) f_c(r_{ik}) , \qquad (70)$$

where the angular dependence, originally chosen as

$$g(\cos \vartheta) = 1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + (\cos \vartheta_0 - \cos \vartheta)^2} \qquad (71)$$

will later help to make a crystal in the diamond structure resist shear stresses. Here c and d are adjustable parameters, while ϑ_0 is, for example, the ideal tetrahedral angle. The term ξ_{ij} can be interpreted as an effective coordination number (minus one, since atom j is excluded from the sum) under which atom i "sees" atom j. If all bond angles on atom i are equal to ϑ_0, $\xi_{ij} = Z_i - 1$ assuming the cut-off function to assume the values one and zero for nearest and next-nearest neighbors, respectively. However, "unhappy" angles formed by ij and ik bonds lead to an increase of ζ_{ij}.

Tersoff then constructed a prefactor, b_{ij}, to the attractive pair interaction as a function of the (effective) coordination number and called it a bond-order variable. It is defined heuristically as

$$b_{ij}(\xi_{ij}) = \left\{ 1 + (\beta_i \xi_{ij})^{\gamma_i} \right\}^{-1/(2\gamma_i)}, \tag{72}$$

where β_i, and γ_i are element-specific parameters. For large ξ_{ij}, b_{ij} scales with $1/\sqrt{\xi_{ij}}$, which was shown by Brenner [248] to be similar to EAM or TB2M potentials. With these ingredients, a general Tersoff potential reads

$$U(\{\mathbf{r}\}) = \frac{1}{2} \sum_{i,j \neq i} \left\{ u_r(r_{ij}) - b_{ij} u_a(r_{ij}) \right\} f_c(r_{ij}), \tag{73}$$

where the (pair) contributions $u_r(r_{ij})$ and $u_a(r_{ij})$ are strictly repulsive and strictly attractive, respectively. Typically, $u_{a,r}(r_{ij})$ are exponential functions to reflect the universal equation of state of crystals [62,249].

Using a symmetrized bond-order parameter, $\bar{b}_{ij} = (b_{ij} + b_{ji})/2$, the Tersoff potential can be cast as

$$U(\{r_i\}) = \sum_{i,j > i} \frac{U_0}{n - m} \left\{ m e^{n(1 - r_{ij}/r_0)} - \bar{b}_{ij}\, n\, e^{m(1 - r_{ij}/r_0)} \right\} f_c(r_{ij}). \tag{74}$$

Tersoff chose $n = 2m$, as to reproduce the original Morse potential [245]. From the Ducastelle-potential perspective, i.e. when setting $g(\cos \vartheta)$ to unity, this is precisely the choice of exponents, in which the binding energy is insensitive to the coordination number within the nearest-neighbor approximation, see our discussion on this in Sect. 6.1. Tweaking the exponents toward larger (smaller) n/m ratios biases structures in favor of larger (smaller) coordination numbers.

A point that might be particularly important and easy to make in the context of bond-order potentials is that the effect of screening is cast through a cutoff rather than through the analysis of local topology. For example, given fixed values for the two cut-off radii $r_{c1,2}$ in the Tersoff potential, it simply seems wrong that the variable ξ_{ij} can change from its generic value of 3 in the diamond structure to 15

upon a relatively minor isotropic compression or to 0 upon a hypothetical homogeneous decompression. A more elaborate discussion of this issue is given in Sect. 9.

5.6. Generic functional form

All open-shell potentials discussed in the preceding sections, and the closed-shell ATM potential, can be cast into the universal functional form given by Eq. (53) and (54). The universal form highlights similarities between the construction of these empirical potentials and simplifies the analytic manipulation, such as the computation of properties that depend on derivatives like forces,

$$f_i = -\frac{1}{2} \sum_{ij} \left(\frac{\partial U_2}{r_{ij}} + \frac{\partial Um}{r_{ij}} \right) \hat{n}_{ij} - \frac{1}{2} \sum_{ij} \frac{\partial Um}{\partial \xi_{ij}} \sum_{k \neq i,j} \sum_{X} \frac{\partial Um}{\partial r_X} \hat{n}X \tag{75}$$

with $\Xi_{ij} \equiv (\partial \Xi / \partial r_{ij}) r_{ij} / r_{ij}$, where r_{ij}, r_{ik} and r_{jk} are the first, second and third argument of $\Xi()$. Stresses or higher order derivatives like elastic constants are equally straightforward to evaluate.

We summarize the functional forms for the potentials discussed here in Tab. 2. Both ATM and SW potentials have a many-body contribution of the form $U_m(r_{ij}, \xi_{ij}) = \xi_{ij}$, which reduces the many-body contribution to a true three-body contribution, e.g. U_3 in Eq. (2). Carlsson [19] dubbed these types of expressions *cluster potentials*. Ducastelle and Tersoff introduce a nonlinear mapping for the functional dependency on ξ_{ij} in U_m, turning the cluster potential into what Carlsson termed a *cluster functional*. The critical difference is that while cluster potentials go to a finite order in the slowly converging series Eq. (2), cluster functionals implicitly include higher-order terms. This can be seen by expressing the many-body contribution as a Taylor series in ξ_{ij}. This Taylor series is truncated for cluster potentials but contains quadratic ξ_{ij}^2, cubic ξ_{ij}^3 and higher-order terms for cluster functionals. The quadratic term,

$$\xi_{ij}^2 = \sum_{k,l} \Xi(r_{ij}, r_{ik}, r_{jk}) \Xi(r_{ij}, r_{il}, r_{jl}), \tag{76}$$

clearly contains a four-body contribution to U_4 in Eq. (2). Similarly, the cubic terms contributes a five-body interaction and so on. This implicit incorporation of higher-order terms can alleviate the slow convergence of the formal series expansion given by Eq. (2), while avoiding computationally expensive, explicit calculations of higher-order terms.

$$U(\{r\}) = \frac{1}{2} \sum_{ij} \{U_2(r_{ij}) + U_m(r_{ij}, \xi_{ij})\} \quad \text{with} \quad \xi_{ij} = \sum_{k \neq i,j} \Xi(r_{ij}, r_{ik}, r_{jk})$$

$$U_2(r_{ij}) \propto \begin{cases} m - n \text{ Lennard} - \text{Jones} & \text{ATM, SW} \\ \exp(-r_{ij}/\sigma_r)f_c(r_{ij}) & \text{Ducastelle/Gupta, Tersoff} \end{cases}$$

$$U_m(r_{ij}, \xi_{ij}) \propto \begin{cases} \xi_{ij} & \text{ATM, SW} \\ \delta_{ij}\sqrt{\xi_{ij}} & \text{Ducastelle/Gupta} \\ (1 + \xi_{ij})^{-1/(2\gamma)} \exp(-r_{ij}/\sigma_a)f_c(r_{ij}) & \text{Tersoff} \end{cases}$$

$$\Xi(r_{ij}, r_{ik}, r_{jk}) \propto \begin{cases} (1 + 3\cos\vartheta_{ijk}\cos\vartheta_{jik}\cos\vartheta_{kij})/(r_{ij}^3 r_{jk}^3 r_{ki}^3) & \text{ATM} \\ \exp(-r_{ik}/\sigma_a) & \text{Ducastelle/Gupta} \\ g^{SW}(\cos\vartheta_{ijk})f_c^{SW}(r_{ij})f_c^{SW}(r_{ik}) & \text{SW} \\ g^{T}(\cos\vartheta_{ijk})f_c(r_{ik}) & \text{Tersoff} \end{cases}$$

Table 3. Selected many-body potentials as cast into the generic functional form, Eq. (53) and (54). We here present simplified expressions for unary systems and have omitted prefactors to highlight functional dependencies, but generalization to alloys and introduction of prefactors is straightforward. σ_r and σ_a are characteristic length-scales of repulsion and attraction, respectively. f_c are cutoff functions whose expressions differ between Stillinger-Weber and Tersoff, see Eq. (132). The functions g^{SW} and g^T encode the angular dependence of Stillinger-Weber and Tersoff, respectively. Divergent onsite terms ξ_{ii} are implicitly excluded from summation. Generic forms of the SW and Tersoff potentials differ from their original formulation by a (constant) onsite contribution that is inconsequential for forces.

5.7. Brenner potentials

Brenner reparameterized Tersoff's functional form, specifically with $\gamma_i = 1$, for modeling hydrocarbon chemistry [10,250]. This potential is now known as the *reactive empirical bond-order potential* (REBO). Brenner realized that Tersoff's potential essentially ignores the π bond in *sp*-valent system. Thus, in its raw form, it is unable to describe the complex chemistry of hydrocarbons, for example radicals or bond conjugation. To correct for these deficiencies, Brenner introduced lookup tables that correct Tersoff's total energy expression depending on the coordination of the atoms involved in the bond. The many-body term becomes

$$U_m(r_{ij}, \xi_{ij}) = -\left[b_{ij}(\xi_{ij}) + \Pi_{ij}(N_i, N_j, N_{ij}^{conj})\right]u_a(r_{ij})f_c(r_{ij}), \quad (77)$$

where N_i is the coordination of atom i obtained with a smooth cutoff function $f_c(r)$, $N_i = \sum_j f(r_{ij})$. The function Π_{ij} accounts for the missing π-bond and is a bicubic interpolation between values tabulated for specific coordination of the atoms. N_{ij}^{conj} is another integer variable that is ≥ 2 if the bond is conjugated, which is the case if a neighboring atom (of the bond i-j) has a coordination less than four. N_{ij}^{conj} increases the neighbor shell, which influences the energy of bond i-j, making the potential longer ranged than Tersoff's.

In addition to the correction for the π-bond, Brenner also corrects the σ-bond contribution, using the bond-order expression

$$b_{ij}(\zeta_{ij}) = \left[1 + \xi_{ij} + P_{ij}(N_i^{(\text{H})}, N_i^{(\text{C})})\right]^{-1/2}, \tag{78}$$

where P_{ij} is again a lookup table and $N_i^{(\text{H})}$ and $N_i^{(\text{C})}$ are coordination numbers of atom i, but only counting hydrogens or carbons, respectively, such that $N_i = N_i^{(\text{H})} + N_i^{(\text{C})}$. P_{ij} is only nonzero for C-C bonds if a hydrogen is present as a bonding partner, $N_i^{(H)} > 0$. The lookup tables were fitted to atomization energies of many different hydrocarbon molecules. More information can be found in Refs. [10,250], including nodal values for the lookup table Π_{ij} and P_{ij}.

A *second generation reactive empirical bond-order potential* (REBO2) was published twelve years later by Brenner *et al.* [11]. This potential improves upon the original formulation by adopting new functional forms for $u_r(r)$ and $u_a(r)$. Specifically, the repulsion in REBO2 contains a screened Coulomb (Yukawa) contribution, as discussed in Sect. 3.2.3, in addition to the purely exponential repulsion of REBO and Tersoff's formulation. REBO2 has two distinct angular functions $g(\vartheta_{ijk})$, depending on the coordination of atom i. The potential also added a (four-atom) dihedral potential, allowing proper modeling of rotations around carbon-carbon double bonds. The REBO2 potential was augmented with non-bonded (dispersion) interactions by Stuart, Tutein and Harrison [251] (a development that, despite being published earlier, occurred after the development of REBO2).

We here put forward the bold claim that REBO is the first successful machine-learned potential. Essentially, Brenner did away with the fixed functional form through his lookup tables to correct Tersoff's potential for a specific application. While the sophisticated interpolation techniques that are nowadays used for machine-learned potentials (neural networks, Gaussian processes [252]) were still under development at the time of Brenner's publications, the spline-interpolated lookup tables essentially fulfill a similar role of extrapolating in a high-dimensional space.

5.8. Beyond second moments

Before talking about potentials *beyond* second moments, we need to take a step back and elucidate the origins of the second-moment expansion. In particular, we want to emphasize that there is an atom-centered and a bond-centered expansion that lead to different functional forms for binding energies. We start our discussion with the atom-centered expansion.

The TB2M expressions can be rationalized from orthogonal tight-binding models as follows: The total (band) energy of the electronic system is given by

$$E_{\text{band}} = \text{tr } \hat{\rho}\hat{H} = \sum_n f(\varepsilon_n)\varepsilon_n = \rho_{i\alpha j\beta}H_{i\alpha j\beta} \tag{79}$$

where the $f(\varepsilon)$ is the Fermi-Dirac distribution, ε_n are the energy eigenvalues and $\hat{\rho} = f(\hat{H})$ is the density matrix.[2] At low electronic temperature, $f(\varepsilon)$ is simply a step function such that the sum in Eq. (79) runs over all states with energy below the Fermi level ε_F. The total energy of the system typically includes a pair-wise repulsive contribution in addition to E_{band}, which absorbs (and approximates) all the things that were implicity or explicitly neglected in the derivation of the orthogonal tight-binding band model, such as three-center terms, overlap repulsion, and others.

A crucial step in turning the tight-binding total energy picture into the atom or bond-centric pictures underlying interatomic potential was the formulation of the local density of states (DOS) based on moments of the Hamiltonian. In metals, bonding is determined by the shape of the electronic density of states around the Fermi level. The local density of states is given by [34]

$$n_{i\alpha}(\varepsilon) = \langle i\alpha | \delta(\varepsilon - \hat{H}) | i\alpha \rangle, \tag{80}$$

where $|i\alpha\rangle$ is orbital α on atom i. In terms of this local density of states, the band energy becomes

$$E_{\text{band}} = \sum_{i\alpha} \int \mathrm{d}\varepsilon\, \varepsilon f(\varepsilon) n_{i\alpha}(\varepsilon) . \tag{81}$$

The *moments* of the local density of states are given by [253,254],

$$\mu_{i\alpha}^{(p)} = \int \mathrm{d}\varepsilon\, \varepsilon^p n_{i\alpha}(\varepsilon) = \langle i\alpha | \hat{H}^p | i\alpha \rangle = H_{i\alpha j\beta}H_{j\beta k\gamma} \cdots H_{l\delta i\alpha}, \tag{82}$$

without a sum over $i\alpha$ (although that index is repeated). The right hand side of Eq. (82) has an intuitive interpretation. For the second moments,

$$\mu_{i\alpha}^{(2)} = H_{i\alpha j\beta}H_{j\beta i\alpha}, \qquad \text{(no summation over } i\alpha!) \tag{83}$$

we hop from atom i to all neighboring atoms j and back to atom i and the sum of the squared Hamiltonian elements given the moment. For the third moment,

$$\mu_{i\alpha}^{(3)} = H_{i\alpha j\beta} H_{j\beta k\gamma} H_{k\gamma i\alpha}, \tag{84}$$

we take all self-returning hopping paths involving three atoms, and so on. Since the Hamiltonian matrix is short ranged (remember, it decays roughly like an exponential!), the second moment only depends on the close vicinity of atom i. Higher moments "feel" the atomic structure out to further distances. The moments are useful because they characterize the shape of the density of states. If we know all moments, we know the density of states; conversely we can approximate the density of states using just the lowest moments.

Systematically approximating the density of states from the lowest moments is highly nontrivial, and a number of approaches have been suggested. These approaches include the recursion method [255,256] (that lead to the systematic development of *bond-order* potentials [257,258]), the maximum entropy method [259–262], or simply approximating the local DOS with a rectangle [202,263]. Let us now assume we have an unary (for simplicity s-valent) system (e.g. Cu) with bond integral $\beta_{ss\sigma}(r)$ for the ssσ-bond. There is no angular dependence and the second moment of the local DOS at atom i – see Eq. (83) – is given by

$$\mu_{is}^{(2)} = \sum_j \beta_{ss\sigma}^2(r_{ij}) . \tag{85}$$

Assuming a half-filled rectangular DOS and a single electron per atom in our s-valent metal, the contribution of atom i to the band *energy* (the first moment of the local DOS integrated up to the Fermi level) becomes

$$E_{\text{band},i} = -\sqrt{6\mu_{is}^{(2)}}, \tag{86}$$

which is the origin of the square-root in Eq. (59a). A maximum entropy estimate would lead to a Gaussian, which changes the numerical prefactor of Eq. (86) but not the functional form [264].

Semiempirical atom-centered higher-moment methods have also been developed. Carlsson considered fourth moments (that define something like the kurtosis of the DOS) for models of semiconductors [265] and transition metals [266]. While higher-moment atom-centered techniques have seen some development since [267,268], they have to the best of our knowledge not been widely employed in molecular

calculations. The breakthrough of higher-moment methods came with the development of systematic bond-centered techniques, briefly described in the following.

The arguments present above imply an atom-centered view, as the local density of states is a per-atom quantity. For a slight shift in perspective, we abandon this band-centric view and define $E_{\text{band}} = E_{\text{bond}} + E_{\text{prom}}$ where the *bond energy*

$$E_{\text{bond}} = \frac{1}{2} \sum_{i \neq j} 2\rho_{i\alpha j\beta} H_{i\alpha j\beta} \tag{87}$$

and the *promotion energy* E_{prom} contains the diagonal terms, i.e. the on-site energies of the tight-binding model. This construction is the basis of the tight-binding *bond* model [269]. Equation (87) for the bond energy is appealing, because it introduces the bond *i-j* as the central quantity: It has the appearance of a pair interaction with pair energy $E_{ij} = 2\rho_{i\alpha j\beta} H_{i\alpha j\beta}$; compare it for example to the attractive part of the Tersoff potential given by Eq. (73). The term $2\rho_{i\alpha j\beta}$ is called the *bond order*. As we know from Tersoff's potential, this bond order – or rather the bond order variable b_{ij}, Eq. (72) – depends on the environment of the specific bond and modulates its strength. The promotion energy is constant if local charge neutrality is imposed and can be ignored for the computation of forces and stresses [258].

We can write the density matrix as

$$\rho_{i\alpha j\beta} = \int \mathrm{d}\varepsilon f(\varepsilon) n_{i\alpha j\beta}(\varepsilon) \quad \text{with} \quad n_{i\alpha j\beta}(\varepsilon) = \langle i\alpha | \delta(\varepsilon - \hat{H}) | j\beta \rangle, \tag{88}$$

which is easily seen by inserting this expression into Eq. (79). Note that written this way, the density matrix has a form similar to Eq. (80) and (81). The onsite terms $\rho_{i\alpha i\alpha}$ correspond to the integral over the local density of states $n_{i\alpha}(\varepsilon) \equiv n_{i\alpha i\alpha}(\varepsilon)$ up to the Fermi level, which are the number of electrons $N_{i\alpha}$ in that orbital. We now consider the linear combination of orbitals $|\pm\rangle = (|i\alpha\rangle \pm |j\beta\rangle)/\sqrt{2}$ that correspond to bonding (+) and anti-bonding (−) states [203]. Then $n_{\pm} = \frac{1}{2}\{n_{i\alpha} + n_{j\beta}\} \pm n_{i\alpha j\beta}$ and hence $n_{i\alpha j\beta} = (n_+ - n_-)/2$. The bond-order can therefore be written as

$$2\rho_{i\alpha j\beta} = N_+ - N_-, \tag{89}$$

with $N_{\pm} = \int \mathrm{d}\varepsilon f(\varepsilon) n_{\pm}(\varepsilon)$. This is the difference of electrons in anti-bonding (N_-) and bonding (N_+) orbitals. We conclude that Pauling's interpretation [247] of the bond-order was indeed correct.

Exploiting the symmetry of the *i-j* bond, the off-diagonal Hamiltonian block H_{ij} can be diagonalized (by rotating into the frame of the local bond *i-j*), allowing the pair energy to be written as [270],

$$E_{ij} = 2H_\sigma(r_{ij})\rho_{\sigma,ij} + 2H_\pi(r_{ij})\rho_{\pi,ij} + 2H_\delta(r_{ij})\rho_{\delta,ij}, \qquad (90)$$

the sum of σ, π and δ contributions to the bond energy. Pettifor, Aoki, Horsfield and coworkers [257,258,271,272] developed a systematic expansion of the bond-order in the moments computed on atom i and j. To second order, they find $\rho_{\sigma,ij} \propto (\mu_{i\sigma}^{(2)} + \mu_{j\sigma}^{(2)})^{-1/2}$, leading to an expression similar to Tersoff's total energy (if π-bonds are neglected). Since this pioneering work, higher-order bond-order potentials have been developed for a number of elements and elemental combinations, such as hydrocarbons [273–275], molybdenum [276], gallium-arsenide [277], silicon [278], tungsten [279], iron [280], cadmiun-telleride [281], and aluminum-copper-hydrogen [282–284].

6. Properties of pair and many-body potentials

6.1. Binding energy, lattice constants, and equation of state

The properties of crystals with a single atom in the primitive cell that interact through a smooth, short-range two-body potential $U_2(r)$ [with a single (steep) minimum in $U_2(r)$ and a single extremum in relevant higher-order derivatives] are severely restricted. We start by discussing the cohesive energy, arguably the most important property of any structure. For a perfect mono-atomic and thus non-ionic crystal, in which each atom is equivalent, the energy per atom in the (effective) pair-potential approximation can be given by

$$U_{\mathrm{pa}}(a_0) = \frac{1}{2}\sum_{s=0} Z_s U_2(a_s) . \qquad (91)$$

Values for Z_0, Z_1, a_1/a_0 are stated in Table 2 for mono-atomic cubic crystals along with other characteristic, dimensionless numbers, which will be introduced further below. Given that $Z_0 = 12$ is the maximal coordination number, face-centered cubic (fcc) and hexagonal closed packed (hcp), which has the same Z_0, Z_1, and a_1/a_0 as fcc, are the generally preferred structures of mono-atomic systems with short-range interactions. What phase wins depends on the precise functional form of the potential but also where and how the potentials are cut off, as discussed in more detail in Sect. 9.1.

To lowest order, i.e. when only including interactions within the first shell, the nearest-neighbor spacing satisfies $a_0 = r_0$, where r_0 is the location of the energy minimum of the dimer. Like nearest neighbors, more distant neighbor shells are located at a negative potential energy. As a result, the total binding energy per atom in the solid will exceed Z_0 times the energy per atom in the dimer, which is $U_0/2$. Due to the further shells sitting in the

attractive part of the potential, the lattice will contract. As a consequence of this contraction, nearest neighbors are pushed more deeply into their repulsive interaction, which then leads to an increase of the bulk modulus B, despite counteracting effects of other shells. The inequalities

$$a_0^{\mathrm{eq}}/r_0 \lesssim 1 \tag{92a}$$

$$U_{\mathrm{pa}}^{\mathrm{eq}}(Z_0)/U_0 \gtrsim Z_0/2 \tag{92b}$$

$$B_{\mathrm{eq}} \gtrsim Z_0 \frac{k_{\mathrm{d}} r_0^2}{18\, v_{\mathrm{pa}}^{\mathrm{eq}}} \tag{92c}$$

encapsulate this discussion, $U_{\mathrm{pa}}^{\mathrm{eq}}$ being the binding energy per atom (in the ground state with nearest-neighbor distance a_0^{eq}) and $v_{\mathrm{pa}}^{\mathrm{eq}}$ is the volume per atom. Values of v_{pa} for different structures are tabulated in Tab. 2. These equations are expected to hold for non-ionic solids interacting through pair potentials. k_{d} is the curvature of the pair potential in the minimum, which could be called the dimer spring constant. It takes the value

$$k_{\mathrm{d}} = mn U_0/r_0^2 \tag{93}$$

for the pair potentials considered here, at least in their most simple form without amendments. The presented treatment is easily generalized to simple binaries as long as Coulomb interactions are negligible.

Ionic crystals with long-range Coulomb interaction must be treated separately, because Z_s can be of order $Z_0 a_s/a_0$ so that the sum over shells is not convergent. Summing up Coulomb interactions (in periodically repeated systems) is an entire industry by itself [285,286]. However, as long as crystallographic positions are fixed relative to the unit cell, it is clear that the sum over $1/r^m$ terms is a geometric factor depending on the lattice times $1/a_0^m$. To keep the treatment analytically amenable, we only consider nearest-neighbor repulsion plus Coulomb interaction and approximate the energy per atom as

$$U_{pa}(a_0) \approx Z_0 2 U_0 n - 1 e^{n(1-a_0/r_0)} - \alpha_M 8\pi\varepsilon_0 Q^2 a_0, \tag{94}$$

where α_{M} is the geometric factor for $m = 1$, the Madelung constant. Since $\alpha_{\mathrm{M}} < Z_0$ for all (simple) crystal structures, it is clear that the absolute ratio of Coulomb and repulsive potential is reduced compared to that of the dimer. This leads to a reduction of the energy per atom in the crystal and consequently to an inversion of all three inequalities listed in Eq. (92a). Eq. (94) is not necessarily sufficiently accurate to determine what phase an ionic crystal assumes as a function of the parameter n [48,150]. To this end, it is necessary to also include dispersive interaction and repulsion between

anions plus potentially many-body terms [150], which are frequently cast through so-called breathing and deformable shell models as described in Sect. 4.4.

Using Mie instead of exponential repulsion and restricting repulsion to nearest neighbors allows simple closed-form expressions to be derived [153],

$$a_0^{\text{eq}}/r_0 = \sqrt[n-1]{Z_0/\alpha_M} \tag{95a}$$

$$U_{\text{pa}}^{\text{eq}}/U_0 = \sqrt[n-1]{\alpha_M/Z_0}\,\alpha_M \tag{95b}$$

$$B^{\text{eq}} = \sqrt[n-1]{\alpha_M/Z_0}\,\alpha_M\,\frac{n}{n-1}\frac{k_d r_0^2}{18 v_{\text{pa}}}, \tag{95c}$$

where k_d follows from Eq. (93). As expected, the gain in energy and stiffness obtained through the condensation from molecules into crystals is much reduced compared to that of short-range potentials.

Given $U_{\text{pa}}(a_0)$, the equation of state (EOS) at small temperature follows from $p = -\partial U_{\text{pa}}/\partial v_{\text{pa}}$, which can be evaluated quite easily through

$$p(v_{\text{pa}}) = -\frac{\partial U_{\text{pa}}}{\partial a_0}\frac{\partial a_0}{\partial v_{\text{pa}}} \tag{96}$$

when using the volumes per atom $v_{\text{pa}}(a_0)$ listed in Tab. 2. The lean pair potentials introduced here allow the EOS of our four reference compounds to be reproduced quite accurately in their thermodynamically stable phase, as is revealed in Figure 10. However, for copper and carbon, the obtained parameters differ quite substantially from those describing the dimers, while those for NaCl and Ar differ only by roughly 10–20%. Pertinent numbers are listed in Tab. 1.

The apparent dimer binding energies U_0 deduced from the fits are decreased by a factor of approximately 2.2 for carbon ($Z_0 = 4$) and by a factor of 4.6 for copper ($Z_0 = 12$). Both numbers can be crudely rationalized using the EAM/TB2M potential proposed by Ducastelle (see Sect. 5.2), which predicts the binding energy per atom to increase a little less slowly than with $\sqrt{Z_0}$.

Frequently, a measured or computed EOS is also fitted to the so-called "universal equation of state" [249,293,294] given by

$$p(V) = 3 B_0 \frac{1-\eta}{\eta^2} e^{3(B'_0 - 1)(1-\eta)/2} \tag{97}$$

Here, B_0 is the bulk modulus at $p = 0$ and $\eta = \sqrt[3]{V/V(p = 0)}$, while B'_0 is the change of the bulk modulus $\partial B/\partial p$ at $p = 0$ and constant temperature, which is ideally zero temperature when gauging interaction parameters
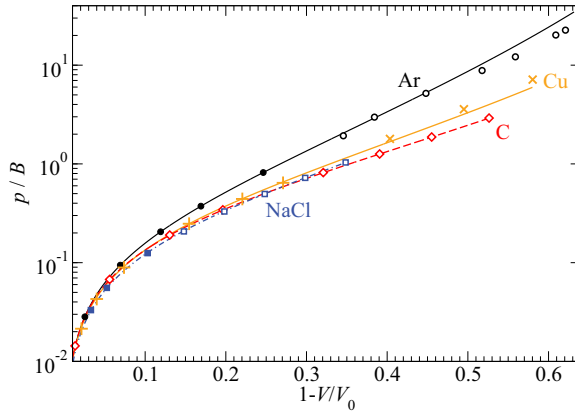
Figure 10. Equation of state, $p(V)$, for C, Cu, Ar, and NaCl as determined experimentally and by pertinent fits assuming pair-additivity of potentials of the same functional form as in Figure 3 and with parameters listed in Table 1. The pressure p and the volume v are normalized to the zero-pressure bulk modulus B and volume v_0, respectively, which the fits were constrained to reproduce. Black full circes [287] and open circles [288] show reference data for fcc argon, orange plus signs [289] and crosses [290] for fcc copper, red diamonds [295] for carbon in the cubic-diamond structure as well as blue, full [291] and open [292] squares for NaCl in the rock salt structure.

from lattice sums. In fact, the reference data for carbon [295] in Figure 10 was produced using Eq. (97) whose adjustable coefficients had been gauged on experimental data.

As can be seen, also an ionic solid satisfies the universal EOS under compressive stress reasonably well, although its two-body potential is marked-ly distinct from that of carbon or copper. Thus, inverting an EOS, in particular when obtained solely under compression, allows potentials to be tested but not to be parametrized. Using first-principle methods, the EOS can be extended to tensile stresses, thereby providing useful information also on the attractive part of the potential. Careful parametrizations of potentials make use of such data [296–299].

Even though pair potentials can be used to fit the EOS of individual crystalline structures, they are not transferable between them. As described in detail in Sect. 5, many-body terms are required for transferability. We will now analyze the Ducastelle potential as one of the simplest many-body formulations for metals. In the nearest-neighbor approximation, the per-bond energy of crystalline Ducastellium, Eq. (60), can be approximated as

$$U_{\text{pb}}(a_0, Z_0) = \frac{U_0}{m - n} \left\{ m e^{n(1 - a_0/r_0)} - \frac{1}{\sqrt{Z_0}} n e^{m(1 - a_0/r_0)} \right\} . \qquad (98)$$

It follows for the equilibrium bond length that

$$\frac{a_0^{\text{eq}}(Z_0)}{r_0} = 1 + \frac{1}{2}\frac{\ln Z_0}{n-m} \qquad (99)$$

and for binding energy

$$U_{\text{pa}}^{\text{eq}}(Z_0) = U_0 Z_0^{\mu}/2 \quad \text{or} \quad U_{\text{pb}}^{\text{eq}}(Z_0) = U_0 Z_0^{\mu-1} \qquad (100)$$

with $\mu = (n/2 - m)/(n - m) < 1/2$. Moreover, the generalized stiffness of the potential,

$$k^{\text{eq}}(Z_0) = mn U_{\text{pb}}^{\text{eq}}(Z_0)/r_0^2 , \qquad (101)$$

is identical to that of a pair-potential, see Eq. (93). Corrections due to next-nearest and more distant shells alter these results. However, all trends are robust and reflect the results for copper shown in Figure 9— as well as for a few other metals condensing in fcc. These trends are a logarithmic increase of a_0^{eq} with Z_0 and an algebraic increase of $U_{\text{pa}}^{\text{eq}}$ with a power less than $1/2$. It is important to note that for $n = 2m$ we obtain $\mu = 0$, and Eq. (100) shows that the cohesive energy is then independent of crystal structure.

Note that the same trends are also observed by carbon (also shown in Figure 9), despite the fact that it cannot be described by Ducastelle's potential, as carbon's ground-state structures are open crystal lattices. To understand covalent bonding, it is instructive to consider the energy per bond

$$U_{\text{pb}}(a_0) = \frac{U_0}{n-m}\left\{ m e^{n(1-a_0/r_0)} - \bar{b} n e^{m(1-a_0/r_0)} \right\} f_{\text{c}}(r_{ij}) \qquad (102)$$

of nearest-neighbor Tersoffium, Eq. (74). We directly see that we recover Eq. (98) of Ducastellium for $\bar{b} = 1/\sqrt{Z_0}$, illustrating the relationship between the bond-order variable and the coordination number, see also Ref. [248]. The bond-order variable $\bar{b}$ typically does not depend on bond-length, but only on crystal structure. We can therefore read-off the equilibrium bond length as,

$$\frac{a_0^{\text{eq}}}{r_0} = 1 - \frac{\ln \bar{b}}{n-m} . \qquad (103)$$

Similarly, the energy per bond scales as

$$U_{\text{pb}}^{\text{eq}}(\bar{b}) = U_0 \bar{b}^{2-2\mu}, \qquad (104)$$

explaining the rough structural trends in Figure 9 for carbon. We can combine Eq. (103) and (104) to eliminate $\bar{b}$, yielding

$$\frac{a_0^{\text{eq}}}{r_0} = 1 - \frac{1}{n} \ln \frac{U_{\text{pb}}^{\text{eq}}}{U_0}, \tag{105}$$

which is displayed in Figure 9h.

These analytical equations are useful in parameterization of these potentials. Albe [296,297,299] developed a parameterization procedure, based on data of ground-state crystalline structures such as the one shown in Figure 9. The first step is fitting the parameters n to the universal energy-bond relation given by Eq. (105). Albe calls the corresponding plot, such as Figure 9d for copper and 9h for carbon, a Pauling plot. The remaining (pair) parameters U_0, m and r_0 are then adjusted to the binding energy, bond length and vibrational frequency of the dimer. In a final step, the parameters that enter the specific calculation of $\bar{b}$ are fitted to reproduce detailed structural information in the database.

6.2. Defects, melting, and boiling

Reproducing defect energies correctly constitutes an important benchmark for potentials, as their proper description is required for the atomistic modeling of thermodynamic and non-elastic mechanical properties. Defects are broadly characterized into point, line, and planar defects. Reproducing their energies accurately is more difficult than reproducing an EOS, because atoms near defects lack the symmetry of ideal crystallographic positions. Parameters having no effect on the EOS for reasons of symmetry and a small effect on lattice vibrations can become important for the proper description of defects, mostly because defects imply a coordination change and might be accompanied by the generation of radicals or ions in covalently bonded systems.

Vacancies are the most important point defects in simple crystals. Two values for their energy are associated with it, i.e. the vertical vacancy energy, $U_{\text{vac}}^{(0)}$, for which the lattice in the vicinity of the removed atom is kept undistorted, and the true defect energy, U_{vac}, which is obtained after lattice relaxation. The $\mathcal{O}(10\%)$ reduction of the defect energy due to relaxation can be deemed irrelevant when studying trends, in contrast to the 10 to 10^5 fold increase in computing and human time to estimate that correction to within 10%. Different system sizes would have to be selected, each structure be relaxed to the energy minimum, and results be extrapolated to the thermodynamic limit, while assessing a vertical defect energy requires merely one calculation in addition to the ones finding the crystal's ground state.

For a short-range potential without explicit angular dependence, the energy per atom can be estimated with $U_{\text{pa}}^{\text{eq}} = U_0 Z_0^{\mu}/2$ – see also Eq. (100) –, where, however, $\mu = 1$ for pair potentials and $\mu \lesssim 1/2$ for

Ducastellium. Removing an atom from an ideal lattice site and placing it hypothetically into another ideal lattice site makes its Z_0 ditched neighbors lose one bond each, leads to an estimate of, $U_{vac} = Z_0\{U_{pa}^{eq}(Z_0) - U_{pa}^{eq}(Z_0 - 1)\}$, which becomes

$$U_{vac}^{(0)} = \frac{U_0}{2} Z_0 \{Z_0^\mu - (Z_0 - 1)^\mu\} \approx \mu U_{pa}^{eq} \qquad (106)$$

where the latter approximation applies if $(1 - \mu)/Z_0 \ll 1$. Thus, $U_{vac}^{(0)} = U_{pa}^{eq}$ for a pair potential ($\mu = 1$), while for $\mu < 1$, our back-of-the envelope calculation provides a lower, yet good estimate for $U_{vac}^{(0)}$. Taking into account that the bond lengths of the ditched neighbors is not at the perfect bond length for their new $Z_0 - 1$ coordination would lead to a relative increase of $U_{vac}^{(0)}$ of typically less than 5% in a Ducastelle potential. Figure 11 confirms the predicted trend that $\mu' = U_{vac}^{(0)}/U_{pa}^{eq}$ is distinctly less than unity for simple metals, in fact their μ' falls slightly below 0.5, which is the upper bound for μ approached in the limit n, mentioned in the text below Eq. (100). Group 14 elements have substantially larger μ' than metals, however, μ' and μ cannot be expected to be similar, whenever bond-angle corrections matter. Only argon out of the elements considered in Figure 11 has a vertical defect energy reasonably consistent with the pair-potential assumption.

One reason why defect energies in metals are important is that they are believed to correlate with the melting temperature T_m: Górecki found the vacancy concentration just below T_m in many different metals to be $n_{vac}^* \approx$ 0.0037 [303] . Estimating the relative number of vacancies with $n_{vac} \approx \exp(-\beta U_{vac})$ would yield $k_B T_m / U_{vac} = -1/\ln n_{vac}^* \approx 2/11$, which roughly overestimates the melting temperatures reported in Figure 11 by a factor of two. Thus, the simple picture laid out here is only half the story, or, depending on viewpoint, it is already half the story. Given the utmost simplicity of the arguments, we would consider the melting-temperature glass to be half full.

To extend the discussion to the boiling (or sublimation) temperature, T_b, the temperature has to be identified at which the Gibb's free energy G of the condensed phase equals that of the gas. At low temperature, G can be crudely estimated with the cohesive energy of the crystal and at high temperature with $-T_b$ times the entropy of an ideal gas. This gives a quasi-linear relation between U_{pa}^{eq} and T_m with corrections logarithmic in T_m and p. When checking data for metals, we found them to obey a $k_B T_b \approx 0.06 U_{pa}^{eq}$, surprisingly well. Unfortunately, we did not find a name for this relation, although a similar rule certainly exists in the literature. Even water does not stray too far from this finding with $k_B T_b / U_{pa}^{eq} \approx 0.054$, when expressing its cohesive energy of $U_{pm} \approx 0.6\,eV$ [304] in ice Ih per molecule (pm) rather than per atom.
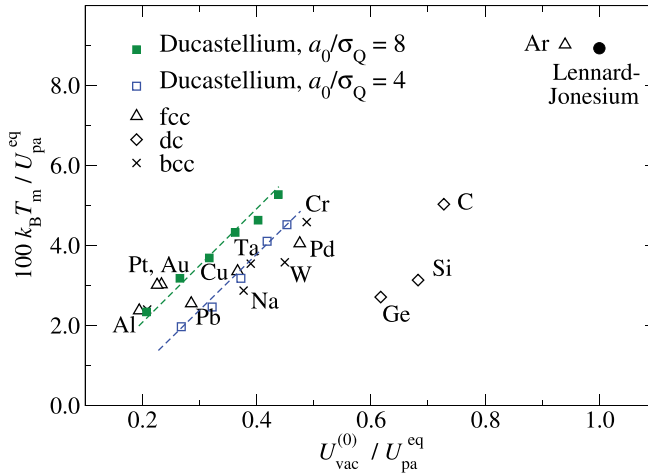
Figure 11. Melting temperature in units of cohesive energy as a function of $U_{vac}^{(0)}/U_{pa}^{eq}$ for a variety of elemental systems condensing in bcc (crosses), fcc (triangles), and diamond-cubic structure (dc, diamonds). Melting temperatures and cohesive energies were taken from Kittel's book [300]. Vertical defect energies are provided for Al [301], Au [301], Pt [301], Cu [301], Pb [301], Pd [301], Na [302], Ta [301], W [301], Cr [301]. For C, Si and Ge, U_{pa}^{eq} values were calculated using DFT-LDA. In addition, the vertical defect energy reduction of Ar was deduced from the ATM parameters from Ref. [147].

Given the discussions on melting and boiling temperatures, their ratio $r_{mb} := T_m/T_b$ must be expected to correlate with the exponent μ. For our representative pair-potential element argon, melting and boiling temperature differ only by 4% at atmospheric pressure. In contrast, $r_{mb}(\text{alkali}) \approx 0.3$, $r_{mb}(\text{earth alkali}) \approx 0.6$, and most transition metals being somewhere in between.

Of course, much more detailed analyses are required to estimate reliably melting temperatures and boiling points from defect and cohesive energies than the crude arguments presented above. The trouble is, as with potentials, that improving estimates even marginally requires a substantially increased effort. Thus, for the sake of brevity, we cannot go into the required level of detail and instead invite the reader to conduct the fun/instructive exercise of melting and vaporizing a two-dimensional crystal in a simulation cell having many times times the crystal's volume: one time with a simple pair potential and one time with a generic Ducastelle or related potential. If done correctly, interesting insights can be gained, for example, on long-range positional and orientational order in two spatial dimensions, or the exceedingly small densities required to make the metal gas consist of monomers rather than of dimers, revealing one more reason why the above presented reason for the correlation between T_b and U_c should be bad. For reasons of brevity, we do not only abstain from a more detailed

analysis of the effects of the $U_{\mathrm{pa}}^{\mathrm{eq}} \propto Z^{\mu}$ scaling, but also from reporting trends for surface or grain boundary energies, although both are central to the properties of materials.

6.3. Elastic properties

Deforming a solid from its equilibrium geometry requires the application of external stress. For a macroscopically affine deformation it can be deduced from elastic tensors. A brief introduction to them including their relation to pair potentials is given in Sect. 3.5. While we have already elaborated on the Cauchy relations, it seems much less appreciated that the pair-potential approximation also places rather tight bounds on C_{11}/B.

The bulk modulus B is easy to evaluate given the equation of state, such as Eq. (96),

$$B = -v_{\mathrm{pa}} \frac{\partial p}{\partial v_{\mathrm{pa}}}. \tag{107}$$

The first constraint to be discussed is the ratio C_{44}/B for fcc crystals interacting with short-range potentials. Using the results compiled in Eq. (35), a first estimate of $C_{44}/B = 3/4$ is obtained using the Cauchy relation $C_{44} = C_{12}$ and $B = (C_{11} + 2C_{12})/3$ valid for cubic systems interacting through pair potentials. This is already close for to the value of $C_{44}/B \approx 0.8$ for fcc Lennard-Jonesium and, by fortuitous error cancellation, even closer to the experimental value for argon.

There are two leading-order corrections to the nearest-shell approximation, which both decrease C_{11} relative to C_{12} and thus C_{44} so that the exact value for C_{44}/B is increased compared to the nearest-shell approximation. First, the effective spring constant k_1 associated with the next-nearest neighbor shell is slightly negative, because next-nearest neighbors are located well past the inflection point of the LJ potential at zero external pressure and even further beyond the point at which $k(r)$ becomes negative. Numerically, we obtain $k_0 \approx 129\, U_0/r_0^2$ and $k_1 \approx -4.8\, U_0/r_0^2$ when using $a_0/r_0 = 0.971$ for fcc Lennard-Jonesium. This accounts for more than 50% of the difference between the nearest-shell approximation and the exact value for fcc Lennard-Jonesium. Second, the dominant effect of more distant neighbors is similar to an isotropic compressive stress p_{c}, pushing the nearest neighbors more deeply into the repulsion and leading to a reduction of C_{11} by p_{c} and an increase of C_{12} and thus C_{44} by the same amount, see also Eq. (39).

The second constraint to be discussed for a short-range potential is the stability of the bcc phase. In the nearest-neighbor approximation, $C_{11} = C_{12}$, which means that the tensor of elastic constants is not positive definite thereby violating the Born mechanical stability criteria [140].

Including the second shell helps stabilizing bcc, but only if the interaction is sufficiently long ranged so that k_1 is positive. For a Morse potential, we find C_{12} to be less than C_{11}, once $n \lesssim 9$ and the binding energy per atom $E_0(\text{bcc}) > E_0(\text{fcc})$ for $n \lesssim 6.1$, at which point not even 20% of the interaction energy is associated with the nearest-neighbor shell. Thus, stabilizing bcc over fcc using a two-body potential straying not too far away from a physically meaningful, *atomistic* potential does not appear to be possible. This can differ for positively charged colloids repelling each other through Yukawa potentials [158].

Some solid-state physics textbooks argue that a positive next-nearest neighbor coupling k_1 stabilizes the simple cubic structure against shear, polonium being the single element in the periodic table adopting it as ground state. However, for k_1 to be positive at a next-nearest-neighbor distance of $r_1 = \sqrt{2}\,a_0$, the Morse exponent has to be as small as $n \approx 2.34$, which is much smaller than any reported value for a Morse type potential that we have come across.

Given the general form of an EAM or TB2M potential of Eq. (57), elastic constants of simple crystals can be found quite easily with the formalism laid out in Sects. 3.5,

$$v_{\text{pa}} C_{\alpha\beta\gamma\delta}^{U\eta} = \sum_s \left\{ 2\tilde{U}_2''(a_s^2) + F'(\rho_0)\tilde{f}''(a_s^2) \right\} a_s^4 v_s^{\alpha\beta\gamma\delta}$$

$$+ F''(\rho_0) \sum_{s,t} \tilde{f}'(a_s^2)\tilde{f}'(a_t^2) a_s^2 a_t^2 v_s^{\alpha\beta} v_t^{\gamma\delta}, \tag{108}$$

where $F'(\rho_0)$ and $F''(\rho_0)$ are the first and second derivative of $F(\rho)$ evaluated at ρ_0, respectively, $\tilde{U}_2(S) = U_2(r)$ and $\tilde{f}(S) = f(r)$. In the nearest-neighbor approximation, this expression simplifies to

$$v_{\text{pa}} C_{\alpha\beta\gamma\delta}^{U\eta} = \left\{ 2\tilde{U}_2''(a_0^2) + F'(\rho_0)\tilde{f}''(a_0^2) \right\} a_0^4 v_0^{\alpha\beta\gamma\delta} + F''(\rho_0) \left(\tilde{f}'(a_0^2) a_0^2 \right)^2 v_0^{\alpha\beta} v_0^{\gamma\delta}$$

$$\tag{109}$$

Similar results for elastic tensor elements have been identified numerous times [45,46] including a derivation using the shell-tensor concept [190], however, mostly for stress-free references. The violation of the Cauchy relations results exclusively from the last summand on the rhs of Eq. (109), e.g. for systems with inversion symmetry,

$$C_{12}^{U\eta} - C_{66}^{U\eta} \equiv C_{1122}^{U\eta} - C_{1212}^{U\eta} = \frac{F''(\rho_0)}{v_{\text{pa}}} \left\{ \tilde{f}'(a_0^2) a_0^2 v_0^{11} \right\}^2 \tag{110}$$

in the nearest-neighbor approximation. Thus, if the embedding function $F(\rho)$ has a positive curvature at ρ_0, as is the case for a Ducastelle potential with $F(\rho) \propto -\sqrt{\rho}$, then $C_{12} > C_{44}$ follows for cubic systems. The symmetrized shear modulus, $G(C_{44} + C_s)/2$, where the so-called tetragonal shear modulus is defined as $C_s(C_{11} - C_{12})/2$, can be shown to obey $G = \alpha\mu B$ in a nearest-shell Ducastelle potential, where $\alpha = 9/8$ for fcc and $\alpha = 1$ for bcc [138]. Thus, the most generic EAM or TB2M potential predicts G/B to scale linearly with μ, just like the vacancy energy. In fact, simple metals obey this trend, as demonstrated in Figure 12. Nonetheless, this correlation must be taken with a grain of salt, since bcc is not elastically stable in the nearest-neighbor approximation.

In contrast to the embedding term, angular terms as those used by Keating, Stillinger-Weber, or Tersoff, counteract shear, so that they lead to a Cauchy violation having the opposite sign as that in Ducastellium and related potentials.

6.4. Plasticity

Plasticity and ductility hinge on the ability of solids to introduce low-dimensional structural defects, in particular dislocations, whose presence is needed to prevent metals from brittle fracture under loading. The more easily (non-planar) defects are inserted into a crystal, the more ductile it will be. In a very microscopic picture, metals can be argued to be ductile, because *"delocalized electrons allow metal atoms to slide past one another without being subjected to strong repulsive forces that would cause other materials to shatter"* [306]. However, going directly from the delocalized nature of the electrons in metallic bonds is somewhat of a big step and does not provide any guidelines allowing the degree of ductility to be assessed.

In this section and the literature reviewed therein, vacancy energies in units of the cohesive energy and shear elastic constants in units of the bulk modulus are found to be relatively small if interactions are isotropic and bonds weaken with increasing coordination number. Both ratios turn out to be of similar order, and, generally less than 1/2 for most metals. Given that the Poisson's ratio of an isotropic material satisfies $\nu = (3/2 - G/B)/(3 + G/B)$, it can be seen that the low limit of $G/B \to 0$ bring the Poisson's ratio close to that of a liquid, i.e. to $\nu = 1/2$, while the lower limit for ν is just a little above the value valid for pair potentials $\nu = 1/4$. Thus, the smaller G/B, the more "liquid-like" the elastic tensor, which, of course, also implies that the energy to introduce a dislocation, which is linear in G, is relatively small. In fact, the elemental solids with a small $\mu' := G/B$ ratio, but also the ones with a small U_{vac}/U_c, are by and large more ductile than the ones with a larger μ'. A nice example is niobium, which happens to be quite ductile despite condensing in
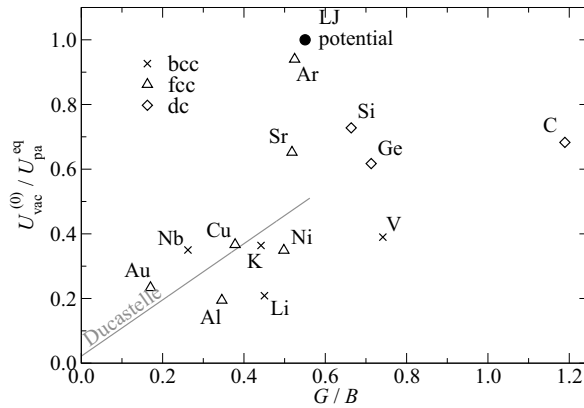
Figure 12. Vertical defect energy in units of cohesive energy as a function of the symmetrized, dimensionless shear modulus G/B for a variety of elemental systems condensing in bcc (crosses), fcc (triangles), and diamond-cubic structure (dc, diamonds). Elastic constants data for all chosen elements but Ar were taken from Ref. [68]. Data for C_{44}/C_{12} for Ar from Ref [305].. Vertical defect energies are provided for Al [301], Au [301], Ni [301], Cu [301], Li [302], K [301], V [301], Nb [301]. The vertical defect energy reduction of Ar was deduced from the ATM parameters from Ref. [147]. For C, Si, Ge and Sr, vertical defect energies were produced using DFT-LDA, with $4 \times 4 \times 4$ super-cell to minimize the interaction between a vacancy site and its image. The full and dashed gray lines show the prediction of the Ducastelle potential for fcc and bcc, respectively.

bcc, which, unlike fcc, lacks close-packed glide planes. Of course, assessing the effect of defects on the plasticity of a material requires the knowledge of many more properties than the dimensionless ratios G/B, $k_B T/U_{vac}$, and so on [307,308], however, the trends reflected in our examples, which were not cherry picked, can certainly be understood from these numbers.

The objective of much atomistic simulation work on metals is to reproduce thermal and mechanical properties of specific metals and their alloys [35,308]. The attempt to rationalize trends has become a secondary objective, which, however, is central to our overview article. Since we are not aware of a work asking the question how the choice of the potential affects the outcome of a typical tensile-load experiment given a certain micro structure, we took the liberty and produced such data. The purpose is to demonstrate that the trends, which are discussed in regard to ductility, are indeed borne out. To this end, we melted two-dimensional Ducastellium ($\mu = 0.36$, typical for copper, with $N \approx 33,000$ atoms) and quenched it to small temperature, whereby crystalline grains of different orientation were produced. The such produced samples, were then subjected to a tensile loading, where the velocity of a few hundred atoms in the outermost layer were constrained to assume a constant value in the direction of loading, and a uniform deformation was simultaneously applied to the rest of the sample. The exercise was repeated for Lennard-Jonesium, in which case the

configurations were rescaled and allowed to relax before loading. The orientation of the graines remained unaltered during the relaxation. Results are shown in Figure 13.

Figure 13 reveals very clearly that the pair-additive Lennard-Jonesium shows the characteristics of brittle fracture, while the isotropic, non-pair-additive Ducastellium produces a force strain relation akin of a ductile metal. In addition, the pair-additive Lennard-Jonesium breaks along existing grain boundaries without a change of internal grain boundary structure, while Ducastellium undergoes massive reconstruction, the precise nature of which cannot be predicted, as it changes with the random seed for the thermostat. Of course, these two-dimensional simulations are merely a cartoon for real solids, among other reasons because stress-induced recrystallization in two dimensions occurs much more easily than in three, due to a small (hyper-) surface to volume ratio.

7. Charge-transfer potentials

It is sometimes insufficient to treat atoms as if they had a constant (integer) charge. This calls for methods allowing charge assignments to be made on the fly. Adjustable coefficients entering the charge-transfer part of a potential must be gauged on given reference data as any other force-field parameter. One possibility is to identify them such that inter-molecular forces or interaction energies from a training set are matched as closely as possible, which, however, is rarely advisable since errors in certain parts of the regular part of the potential may be compensated by false charge assignments, which then jeopardizes the transferability. This is why reasonable, directly determined reference charges are needed in the training of charge-transfer potentials.

One reason why knowing partial charges accurately – whatever this means – is important, is that ionic interactions, so they are present in a system, are on par with covalent interactions. This is easily seen by comparing their magnitude. The covalent bond in H_2 has a binding energy of 4.74 eV. The Coulomb interaction in a NaCl molecule with a bond length of $a_0 = 2.36$ Å is about 6.1 eV, after subtracting sodium's ionization energy, $I_{Na} = 5.14$ eV, and adding chlorine's electron affinity, $A_{Cl} = 3.61$ eV, an energy gain of 4.57 eV remains. This needs to be contrasted to typical LJ interaction parameters of order 0.01 eV. Thus, even minor changes in partial charges easily alter energies more than relatively large changes of LJ prefactors.

7.1. Determination of reference charges

In the absence of periodic boundary conditions, a charge distribution can be rigorously characterized in terms of its net charge, dipole moment, quadrupole tensor, and so on. Nonetheless, there is no unique way to assign *atomic*
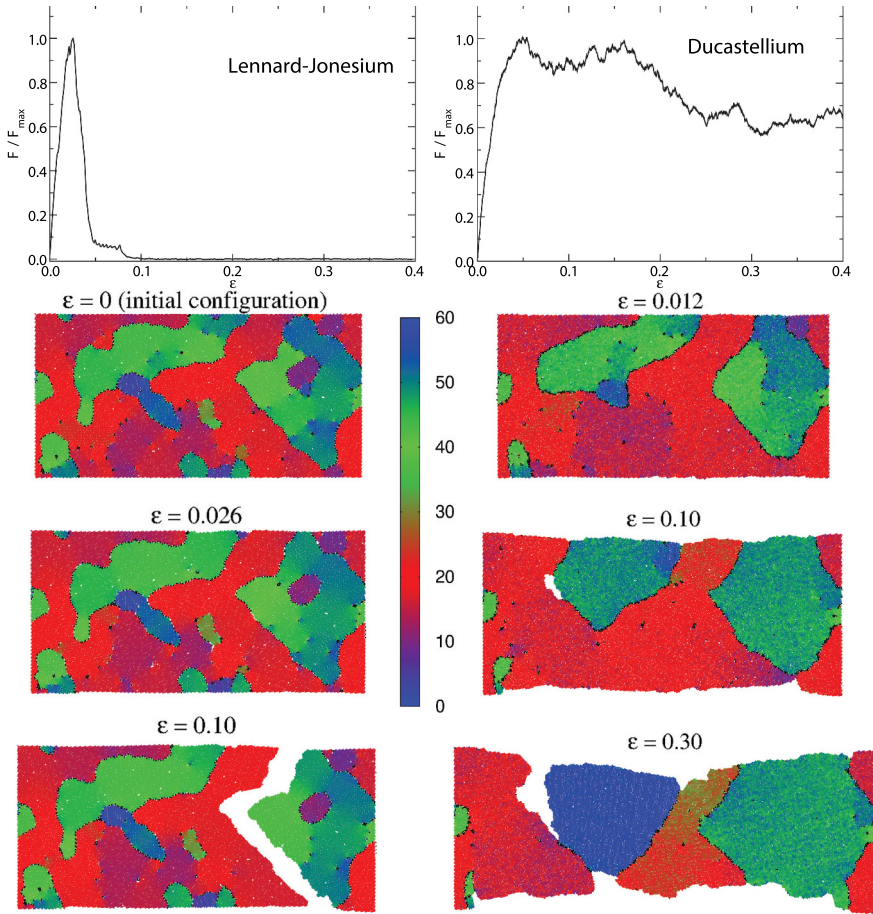
Figure 13. Comparison of two *in-silico* tensile tests. Both simulations started from identical microstructures. The left row assumes Lennard-Jones interactions, the right - a generic model originally proposed by Ducastelle, which later formed the basis for embedded-atom and Finnis-Sinclair potentials. Parameters of Ducastelle potential were taken from Ref. [222]. Simulations were run using LAMMPS [380,381]. Colors indicate the orientation of grains, i.e. the angle ϕ_i obtained when averaging $\Phi_i \exp(i6\phi_i) := \langle \exp(i6\varphi_{ij}) \rangle$ over all atoms within a cutoff, where φ_{ij} denotes the angle of a bond w.r.t. *x*-axis. Atoms, for which, $\Phi_i < 0.75$, are depicted in black.

charges, *atomic* dipoles, etc., in a many-body system. Yet, the concept of formal and partial charge is central in chemical physics and physical chemistry. As a consequence, many different charge-assignment schemes have arisen to achieve the ill-defined. In fact, any pragmatist simply *must* attempt to assign atomic charges allowing interatomic forces to be computed accurately while providing values that satisfy chemical intuition. For example, when trying to match the vibrational frequency, the bond length, and the binding energy of a NaCl molecule with a three-parameter potential, be it Mie or simple Buckingham with two parameters for repulsion plus one independent partial charge, the latter turn out in the vicinity of unity so that

chemical intuition is satisfied. Even more, the such obtained parameters allow quite reasonable predictions to be made, for example, for the lattice constant, bulk modulus and in fact, all three independent elastic tensor elements of NaCl in the rocksalt structure, even if truly accurate potentials require (many-body) corrections. We note in passing that producing numbers supporting this claim is a great exercise for students and instructors alike. At the same time, deducing charges from molecular dipoles, which yields $Q_{Na} = 0.75$ e when ignoring the induced dipole on the Cl$^-$ ion and making students deduce that a bond in NaCl to be 25% covalent is a task, which we deem halfway between useless and insane, even if countless chemistry lectures and quite a few text books provide discussions along such lines. The above back-of-the-envelope calculation simply neglects the high polarizability of anions [88,91].

Complications in assigning partial charges arise as soon as partial charges are not close to an integer value, which happens when atoms with electronegativity differences of, say, $\Delta I \gtrsim 0.5$ eV, are present, simple alkali halides being the major exception to this rule. Unfortunately, very few partial charges are assigned as easily as those for sodium or chlorine atoms in the NaCl molecule. As argued at the beginning of Sect. 7, it is usually beneficial to separate charge assignment from the construction of the potential. Existing methods to compute reference charges can be crudely categorized into those in which atomic charges are deduced from an analysis of (i) a representation of the wave function (Mulliken [309] and Löwdin [310]), (ii) the electronic density (Bader [311] and Hischfeld [312]), (iii) the electrostatic potential (ESP), (iv) displacement-induced changes in polarization (Born [75]), and frequently forgotten (v) experiment [313]. These methods will be sketched next. For more detailed overviews, see Ref. [313–315].

The Mulliken [309] definition of the partial charge starts from the representation of the total wave function as a linear combination of atomic orbitals. The charge associated with each orbital is assigned to the atom, however, results are far from unique, in particular when using large basis sets. Löwdin [310] removed a certain degree of ambiguity from Mulliken's definition that originated from the non-orthogonality of the functions spanning the basis set. However, uncertainties remain. Bader [311] partitions the space into regions such that each atom is assigned the charge density in its vicinity up to the points at which the charge density passes through a local minimum. However, this procedure ignores the possibility of atomic charge distributions to interpenetrate. It yields finite charges for the promolecules, whose charge density is defined to be the superposition of the atomic charge densities [316]. Hirshfeld [312] originally assumed each atom to own space with a weight proportional to the electronic density of the neutral atom, which, however is clearly inappropriate for ions. For example, in an alkali hydride, H$^-$ would be assigned a smaller radius than

the metal cation. Thus, in more advanced methods, the atomic charge densities are replaced with radially symmetric weighting function that are determined self-consistently [312,317]. In such "iterative stockholder approaches", it remains challenging to identify good trade-offs between the ESP accuracy and the transferability of the charges [318]. In conventional ESP schemes, atomic charges are adjusted to best match the ESP a save distance away from the nuclei, e.g. outside the atomic van-der-Waals radii. This becomes problematic for condensed phases, in which such safe distances are sparse or even non-existent, but also for large molecules with hidden atoms. Even worse, the ESP appears to be ill-defined up to a constant in periodic systems so that partial changes can only be gauged on ESP differences [319]. Born effective charges reflect the change of polarization occurring in response to atomic displacements, which, for periodic structures would be analyzed in terms of the Fourier transform of the dipole moment at small wave numbers. However, since both dipoles and displacements are vectors, Born charges are not scalars but tensors of rank two, unless the system is so highly symmetric that all eigenvalues of the tensor are identical. In addition, Born charges can cast changing atomic dipoles as effective charges, which can lead to charges exceeding the formal oxidation number. Central ways to deduce partial charges from experiment [313] include (i) the matching of electrostatic multipoles, which ultimately is a limiting case of an ESP method, (ii) Bader-type analysis of the charge density deduced from x-ray diffraction patterns, but also (iii) a combined analysis of phonons and dielectric response functions.

Quite a few authors claim to have identified the most meaningful way to assign partial charges. This can scarcely be true, since partial charges are ill-defined so that the optimal charge-assignment scheme depends on the relative weight of target properties. Moreover, the optimum method can also depend on the symmetry of the problem. Problems arising due to "falsely" assigning an induced dipole on an atom as a transferred charge vaporize for highly symmetric structure, in which the lowest-order allowed multipole can be a hexadecupole. Unfortunately, separating atomic dipoles from charge transfer is even difficult for diatomic molecules, which could otherwise serve as a well-defined reference system on which charge-assignment schemes can be gauged.

An optimum choice for a charge to be used in a force-field based simulation may also depend on whether polarizable dipoles are included or neglected so that the optimum reference-charge-determination method may depend on the precise nature of the force field wanting to be parametrized. As a consequence, we are neither willing nor able to express an opinion if any of the sketched methods is all in all the optimum choice and believe that it has to be made case by case.

7.2. Regular charge-equilibration approaches

In the ground-state of a full quantum-mechanical system, the functional derivative of the energy E w.r.t. the electron density $'$ disappears, because $\rho(\mathbf{r})$ minimizes E. In a coarse-grained description of the electron density in terms of a set of atomic charges, $\{Q_i\}$, this minimization principle translates into what Mortier and varying co-authors [320,321] called the electronegativity equalization principle. Assuming that E can be expanded into powers of atomic charges, the ground state energy can be written as

$$U = U(\mathbf{Q} = 0) + \sum_i \frac{\partial U}{\partial Q_i} Q_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 U}{\partial Q_i \partial Q_j} Q_i Q_j + \dots \qquad (111a)$$

$$= U(\mathbf{Q} = 0) + \sum_i \chi_i Q_i + \frac{\kappa_i}{2} Q_i^2 + \sum_{i,j \neq i} \frac{J_{ij}}{2} Q_i Q_j + \dots, \qquad (111b)$$

while nuclear coordinates remain fixed. Charges adopt themselves such that they minimize U usually with the constraint of a fixed total charge. In Eq. (111), the electron density still minimizes the total energy for any $\{Q\}$, albeit with the constraint imposed through the assigned partial charges. Thus, the parameters κ_i, χ_i, and J_{ij} are implicitly defined from a mathematical point of view in Eq. (111), even if their numerical values depend on the method used to relate charge density and partial charges.

The chemical meaning of the parameters κ_i, χ_i, and J_{ij} is best ascertained in limiting cases. Consider atom i to be neutral and isolated. If one electron is added to it so that $Q_i = -e$, the energy decreases by its electron affinity A_i, while it increases by its ionization energy I_i when one electron is taken away from it. Thus,

$$I_i = \frac{\kappa_i}{2} e^2 + \chi_i e \qquad (112a)$$

$$A_i = \frac{\kappa_i}{2} e^2 - \chi_i e \qquad (112b)$$

so that the so-called chemical hardness [322] turns out to be $\kappa_i = (I_i + A_i)/e^2$ and the electronegativity [323,324] $\chi_i = (I_i - A_i)/e$. Meanwhile, J_{ij} can be associated with the Coulomb potential given the charges Q_i and Q_j of two distant atoms i and j:

$$J_{ij} = \frac{1}{4\pi\varepsilon_0 r_{ij}} \quad \text{for} \ \ r_{ij} \to \infty \ . \qquad (113)$$

The effect of an external electrostatic potential would have to be added to χ_i. The bold hope now is that partial atomic charges can be predicted using Eq. (111) and parameterizations, which do not stray too far away from Eqs. (112) and (113).

Rappé and Goddard [325] pioneered the use of the electronegativity equalization principle for the on-the-fly determination of atomic charges in molecular simulation and renamed the approach to charge-equilibration (QEq) method. To keep charges in (partially) ionic systems from blowing up, they suggested to shield the Coulomb interactions at short distances through two-electron Coulomb integrals of the form

$$J_{ij} = \frac{1}{4\pi\varepsilon_0} \int d^3r_i d^3r_j \frac{\rho(\mathbf{r}_i)\rho(\mathbf{r}_j)}{r_{ij}}, \tag{114}$$

where the charge densities $\rho(\mathbf{r})$ are those associated with s-type Slater orbitals. This way, the Hessian, i.e. the matrix formed by $\partial^2 U/\partial Q_i \partial Q_j$ remains positive definite even when atoms approach each other closely. Nonetheless, their claim that QEq "*leads to charges in excellent agreement with experimental dipole moments and with the atomic charges obtained from the electrostatic potentials of accurate ab initio calculations*" can be seen skeptically: Can Coulomb shielding be responsible for charges in the alkali metal halides to be so darn close to unity? And how can QEq reproduce the experimental observation that partial charges of atoms and ions in the gas phase adopt integer multiples of the elementary charge? In fact, QEq obviously predicts the ions of a neutral, dissociated NaCl molecule to each carry a charge of magnitude $(\chi_{Cl} - \chi_{Na})/(\kappa_{Cl} + \kappa_{Na}) \approx 0.4$ elementary charges when truncating the expansion of Eq. (111) after the quadratic term [326].

QEq has multiple practical shortcomings, which all originate from the ease with which *partial* charge can be transferred from one atom to another one even over large distances. To name a few, it leads to the wrong dissociation limit of molecules [326], DFT having related issues, QEq makes the polarizability of polymers grow superlinearly rather than linearly in the chain length [327,328], and, similarly, solids adopt the dielectric response function of a metal [329], i.e. the dielectric constant of a QEq solid is infinitely large. Finally, QEq makes the dipole of alcohols be linear in the length of the hydrocarbon chain [330].

One fundamental problem of Eq. (111) is that the curve of lowest average energy of an *isolated* system versus the number of electrons turns out to be a series of straight lines when generalizing the Hohenberg-Kohn theorem to fractional particle numbers [331]. This property should be reflected in any coarse-grained approximation to DFT, which is built on the Hohenberg-Kohn theorem. However, Eq. (111) fails to do so.

Despite the fundamental issues of QEq, metals can be discussed, since partial charges can be readily transferred over large distances. In a parallel-plate capacitance geometry, the length over which an external-field induced charge density decays into the solids scales approximately with $\sqrt{\kappa}$ [329]. In other words, QEq implicitly contains corrections to the behavior of ideal metals. Thus, mirror charges, or, rather the charge distributions producing the same electrostatic field outside of the metal as mirror charges do, are more smeared out in a QEq solid than in an ideal metal. To reflect this, the term κ_i to be used in a QEq simulation of metals may best be parameterized through an appropriate choice of the Thomas-Fermi screening length, which can benefit, e.g. the description of the interfacial interactions between a metal and an electrolyte [332]. For nanostructured electrodes, the finite density of states at the Fermi level leads to an additional contribution to the capacity of the device [333], which can be modeled by replacing the quadratic term in Eq. (111) with an appropriate nonlinear function [334].

7.3. Charge-equilibration methods for non-metallic systems

Various ideas have been pursued to suppress the non-local charge transfer in QEq for dielectric systems. One of the first and most influential approaches, coined the fluctuating-charge (fluc-Q) model [335], imposes a charge neutrality constraint on individual molecules. Unfortunately, this bug fix does not remedy the superlinear scaling of the polarizability with the chain length of polymers. Moreover, the bonding topology of molecules needs to be defined making the simulation of bond breaking and formation difficult to describe.

To suppress long-range charge transfer, Chelli *et al.* [327] introduced atom-atom charge transfer (AACT) variables $q_{ij} = -q_{ji}$ so that the charge of an atom is given by

$$Q_i = \sum_j q_{ji}, \tag{115}$$

where q_{ji} is the charge transferred from atom j to atom i. In the AACT method, the bond hardness, κ_{ij}, replace the atomic hardness used in regular QEq so that a charge transfer costs a potential energy of $\kappa_{ij} q_{ij}^2/2$. The AACT model remedies all shortcomings in QEq approaches originating from non-local charge transfer, since the AACT method allows the latter to be (completely) suppressed with (infinitely) large bond hardnesses. However, for the Hessian of the AACT model to be positive definite, the bond hardness terms have to be made so large that the dielectric constants of condensed phases, ε_r, turn out to be barely exceed unity [336]. Moreover, AACT suppresses the observed chain-length dependent polarizability of short

oligomers [328,329] and, similarly, produces a zero screening length in dielectrics [329]. Last but not least, AACT cannot be used to model the polarization of metals that occurs in response to an electric field.

Reintroducing the atomic hardness term to the AACT model leads to a new model, which Nistor *et al.* [336] first meant to call fluc-Q 2. Ultimately, they found the term split-charge equilibration (SQE) to be more appropriate, since the model is a split between two models and a "split charge", q_{ij}, formerly known as AACT variable, is split between two atoms. Thus, their model reads

$$U = U(\text{QEq}) + \sum_{i,j>i} \frac{\kappa_{ij}}{2} q_{ij}^2, \tag{116}$$

where the Q_i entering $U(\text{QEq})$, i.e. Eq. (111), are defined as in Eq. (115). Including finite bond hardness suppresses the superlinear scaling of molecules with linear size [337], as is shown in Figure 14. The observable underestimation of the polarizability is supposedly due to the neglect of atomic polarizability.

Other charge-transfer models exist ensuring neutral separation limits. In one line of approach, chemical potential differences between distant atoms are screened with functions motivated from overlap integrals between orbitals on different atoms [326]. However, this procedure implicitly claims that the altitude difference of two objects becomes less when moving them apart laterally, which is not meaningful. In contrast, a very appealing approach is a model in which electrons are treated as shells, which are not tight to one individual atom [154], whereby, in principle, history dependence can be mimicked along the lines of the NaCl dissociation described in the introduction. There are also approaches using non-local softness matrices, the arguably most advanced one being the atom-condensed Kohn-Sham approximated to second order (ACKS2) [338]. Its purpose is twofold: First, defining the expressions for the softness matrix, which is the inverse of the hardness matrix, from expressions accessible from DFT calculations, and second to employ the result in simulations as a mean to deduce partial charges. Ultimately, ACKS2 contains SQE model as a limiting case. The downside of ACKS2 is that a $2N \times 2N$ matrix must be defined, where N is the number of charges, while the number of split charges is only of order ZN. A frequently raised argument against SQE is that the large bond hardnesses occurring at large separation cause numerical difficulties. However, they can be addressed using appropriate pre-conditioners in square-gradient minimization or large split-charge inertia in extended Lagrangian approaches.
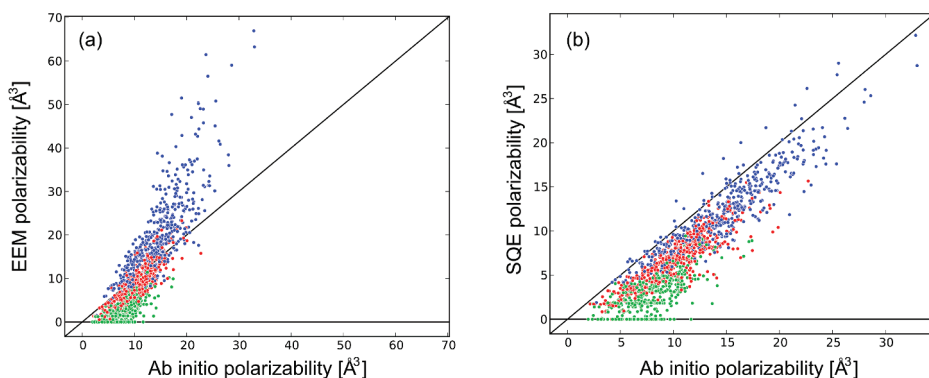
Figure 14. Polarizability of (a) electronegativity equilization model (EEM) (equivalent to the QEq model) and (b) SQE as a function of the polarizability deduced from quantum mechanical calculations on a variety of molecules. Green, red, and blue data points reflect the smallest, the medium, and the largest eigenvalue of the polarizability tensor for different molecules (including varying conformations). Reprinted from Verstraelen, T., Van Speybroeck, V., and Waroquier, M. *J. Chem. Phys.* **131**, 044127, (2009), with the permission of AIP Publishing.

One of the main limitations of the original SQE model is that it cannot produce isolated ions as the κ_{ij} diverges at large r_{ij}. This limitation can be overcome by introducing oxidation numbers n_i so that an atomic charge is given by

$$Q_i = n_i e + \sum_j q_{ji} \qquad (117)$$

with $n_i \in \mathbb{Z}$ in a so-called redox-SQE formalism [157,339]. The remaining part of the SQE formalism is kept, except that force-field parameters, including chemical hardness and electronegativity, must now be assigned individually for each oxidation number. The redox-SQE approach allows a system to effectively move on different potential energy surfaces (if you're a chemist) or different Landau-Zener levels (if you're a physicist). Jumping between them requires the oxidation number of one atom to be increased by an integer, typically by one, while that of another, nearby atom must be decreased by the same number. The corresponding discrete dynamics can be implemented in practice, for example, in terms of Monte Carlo dynamics.

7.4. Properties resulting from charge-transfer potentials

A central aspect of polarizable and charge-transfer potentials is to reproduce the dielectric response of (condensed) media correctly, the most important quantity from a continuum perspective being the dielectric constant ε_r. The relation between point-dipole polarizability

and ε_r is text-book material and reflected in the Clausius-Mossotti (CM) relation, Eq. (42), while that of a bond-polarizable model without point dipole polarizability was worked out by Nistor and Müser [329]. Combining the treatment of point and charge-transfer dipoles does not appear to have been presented in the literature hitherto. However, it can be easily achieved when following Hannay's derivation of the CM relation [340]. It starts from the insight that the full electrostatic field of a dipole contains a contribution $\mathbf{E}_{int} = -\mathbf{p}\,\delta(\mathbf{r})/(3\varepsilon_0)$ acting inside of it in addition to its usually stated far field.

Since each charge leads to a mean electrostatic field of zero, the mean electrostatic field of a charge distribution is also zero, however, outside the dipoles, it is reduced by $\mathbf{E}_{int}$ averaged over the volume that each point dipole occupies. Thus, the average field that a test charge sees outside of dipoles is the external electric field $\mathbf{E}_{ext}$ (which is assumed to be constant) minus $\langle\mathbf{E}_{int}\rangle = -\rho\mathbf{p}/(3\varepsilon_0)$, where ρ is the density of point dipoles. For a rocksalt lattice, or, when discretizing a homoegeneous material into cubes of linear size a_0 and arranging the local charge transfer variables, or, split charges in a vector $\mathbf{q}$, the dipoles obey associated with the charge-transfer, $a_0\mathbf{q}$ and the point dipoles obey

$$\frac{a_0\mathbf{q}}{\kappa_s} = \mathbf{E}_{ext} + \frac{\rho}{3\varepsilon_0}\,\mathbf{p} \tag{118a}$$

$$\alpha\mathbf{p} = \mathbf{E}_{ext} + \frac{\rho}{3\varepsilon_0}\,\mathbf{p}, \tag{118b}$$

respectively. Here, q_x would be the split charge donated from an atom or grid point to its right neighbor, and so on, while κs is the corresponding bond hardness. The total polarizability is nothing but the total dipole density $\mathbf{P} = \rho(\mathbf{p} + a_0\mathbf{q})$, while ε_r is defined through $\mathbf{P} = (\varepsilon_r - 1)\varepsilon_0\mathbf{E}_{loc}$ within linear response. After some minor algebra

$$\varepsilon_r - 1 = \frac{\rho/(a_0\kappa_s) + \rho\alpha}{\varepsilon_0 - \rho\alpha/3} \tag{119}$$

can be obtained, which contains the CM relation in the limit of $\kappa_s \to \infty$ as well as $\varepsilon_r - 1 = \rho/(a\kappa_s)$ valid for bond-polarizable models ignoring point dipoles [329]. Once bond polarizability is ignored in a charge-transfer model, ε_r diverges, which means that the system effectively turns metallic.

Continuum corrections to the above treatment would require the wave-number dependence of the charge-dipole and dipole-dipole coupling but also the chemical hardness coupling to be taken into account. The latter would augment the bond hardness in the presented treatment with $\kappa(qa_0)^2$. A wave-number dependent dielectric constant and thus a dielectric smearing or decay length δ ensues. For a pure bond-polarizable model and the

geometry considered above, $\delta = \sqrt{\varepsilon_0 a_0 \kappa / (1 + \varepsilon_0 a_0 \kappa_s)}$ [336]. One of the consequences of this result is that ideal mirror charges do not accurately represent the dielectric response of real metals [332].

The potential importance of modeling dielectric response correctly is also demonstrated in Figure 15. It shows two pairs of two dislike solids, or, clusters, which can exchange charge, specifically, oxidation numbers, when they are sufficiently close to each other. In reality, this distance would be a weak function of the approach and retraction velocity. In one pair, both materials are metals, in the other, they are insulators. A lot of "electrons" transfer in the metallic couple and swapped charge delocalizes over the entire samples, mostly on their surfaces. In the insulating couple, only one charge exchanges, which remains localized near the former contact points after retraction.

8. Machine-learned potentials

In the past two decades, machine learning has conquered the world of interatomic potentials, a development that is ongoing and accelerating. The main idea is to equip a machine learning (ML) technique with a large database of structures for which energies, forces, stresses, etc. are computed using well-converged first-principle methods.[3] Energies (or sometimes forces) are then extrapolated (or "predicted") based on (local) structural similarity. Structural similarity is either encoded by directly comparing to configurations stored in this database, or it can be implicit in the weights of a neural network or the points of a sparse Gaussian process. Recent methods supplement such computational approaches with experimental data [342]. Additionally, predictions are not limited to energies and forces. Examples of other properties obtained from ML regression are core energies and X-ray photoelectron spectra [343–346].

Machine-learned potentials (MLPs) are sometimes characterized as being parameter-free, however, one may also claim that the number of parameters is gigantic. Occam's razor suggests that models should have few parameters.
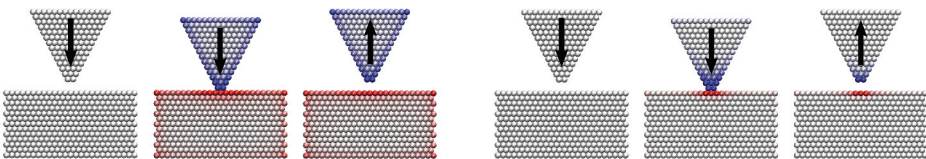

Figure 15. Contact electrification with history-dependent potentials between two metals (**left**) and two dielectrics (**right**). Reprinted by permission from Springer: *Eur. Phys. J.* **86**, 337, Towards time-dependent, non-equilibrium charge-transfer force fields, Dapp, W. B. and Müser M. H., [COPYRIGHT] (2013) Note that Figure 15 is a proof-of-principle demonstrator for the redox-SQE method [339]. It does not allow one to conclude that contact electrification between dielectrics primarily occurs through electron transfer. There are many good reasons to believe that ion exchange matters most [341].

Yet, overfitting of high-dimensional models can be avoided by using para-meterization procedures that are based on sound statistical principles, such as Bayesian inference. Additionally, all physically-justified potential forms (with few parameters) are based on a chain of approximations, e.g. going from the Born-Oppenheimer approximation, to DFT, neglecting three-body integrals in the tight-binding approximation, neglecting nonorthogonality and so on in the bottom-up construction of bond-order potentials for semiconductors. It is often unclear to what extend the final form is still a good approximation for interatomic interactions. To say it with the words of Bazant, Kaxiras and Justo [238]: "*In the case of Si, the abundance of potentials in the literature illustrates the difficulty of the problem and lack of specific theoretical guidance. In spite of the wide range of functional forms and fitting strategies, all proposed models possess comparable and insufficient overall accuracy. It has proved almost impossible to attribute the successes or failures of a potential to specific features of its functional form.*"

MLPs target reproducing the high-fidelity calculations used to construct the database. In that sense, they are higher accuracy (with respect to the reference DFT database) than more empirical or semiempirical models, but this may mean a lack of transferability to structures not explicitly considered in the database. This pitfall is of course not limited to MLPs: As an example, the first incarnation of the Tersoff potential for silicon [242] turned out to have bcc and not diamond cubic as the ground state [347]. Yet, potentials constructed on carefully curated databases give excellent transferability even to obscure crystal phases, see e.g. Ref. [348] for a current example, again for elemental silicon. This highlights the specific importance of database quality for MLPs.

Most MLPs consists of three basic components: First, a database of structures, already outlined above. Second, a "descriptor" that encodes the chemical environment of an atom into a unique fingerprint. In the absence of external fields, this fingerprint must be objective, i.e. it must be invariant with respect to translation and rotation. It must also be invariant with respect to the exchange of atoms of the same species. This allows the data set to be much reduced compared to the situation, where the ML technique has to learn objectivity by itself. An additional complication is that the database itself may not be perfectly objective, due to numerical errors in the underlying electronic structure calculations. Since we want forces, fingerprints also must be continuously differentiable at least once, and ideally, they also have continuous second derivatives. The third component is a regression framework, the "learning" aspect of the potential. Two broad directions of development have occurred here: Using neural networks, which can be regarded as a model having many parameters, or, Gaussian processes, which can be regarded as a scheme directly extrapolating between points in the database. The regression framework takes the fingerprints as input and predicts per-atom energies as output.

Before continuing with more details, we would like to point out that many reviews by researchers more qualified than us to discuss MLPs have appeared recently, see Refs. [29,31,32,349,350]. This review focuses on simple functional forms that allow analytical calculations to be made. Yet, not discussing at least the basic ideas behind MLPs could be a severe omission localized near the former contact points after retraction discussion highlighting similarities between classical and MLPs might help the development of either one.

8.1. Chemical fingerprints

Chemical fingerprints come in many forms, and we here discuss just a select few. The interested reader is referred to other literature, e.g. Refs. [350,351], for more comprehensive descriptions of the wide variety of fingerprints.

The arguably first and one of the simplest example of a chemical fingerprint consists of the coordination number and the bond-conjugation variable employed by Brenner in his potential for hydrocarbons [10,11,250]. Given a bond i-j, Brenner used the coordination numbers N_i, N_j and a bond conjugation parameter N_{ij}^{conj} as a fingerprint for the state of the bond (see Sect. 5.7 for more details). These variables are made continuous by computing the coordination numbers with a smooth cutoff, $N_i = \sum_j f_c(r_{ij})$ where $f_c(r)$ is one of the cutoff functions discussed in Sect. 9.1. This fingerprint is then used to correct the bond-energy of a Tersoff potential through a lookup table. The lookup procedure continuously interpolates between integer values of the triplet $(N_i, N_j, N_{ij}^{\text{conj}})$ using splines. The table entries were fitted to a small database of experimental properties of hydrocarbon molecules.

Brenner's tables are used to correct a baseline potential, a procedure that is sometimes called a Δ-potential, e.g. in Ref. [32]. Other descriptors are needed for full-fledged machine-learned energies. As a simple instructive example, we consider two and three-body terms in the sense of Eq. (2). An objective descriptor for two-body interaction is simply the bond distance r_{ij}. For a three-body term, we can use the triplet of distances (r_{ij}, r_{ik}, r_{jk}) as a descriptor. While this descriptor satisfies objectivity, it still has permutational symmetries with respect to the exchange of the positions of like atoms k and j. Bartók and Csányi [349] suggested the use of the descriptor

$$\mathbf{q}_{ijk} = (r_{ik} + r_{ij}, (r_{ik} - r_{ij})^2, r_{jk}) \tag{120}$$

which is equivalent to (r_{ij}, r_{ik}, r_{jk}) but symmetric with respect to the permutation $j \leftrightarrow k$.

We note that common potentials, like Axilrod-Teller-Muto (Sect. 4.1) and Stillinger-Weber (Sect. 5.4), which in their original formulation depend on the (r_{ij}, r_{ik}, r_{jk}) triplet, can also be written in terms of the descriptor $\mathbf{q}_{ijk}$. These analytical potentials obey the above permutational symmetries by construction. However, in a machine-learning context, where an energy is essentially obtained by extrapolating from a given set of database entries, there is no guarantee that such permutational symmetry will be fulfilled by the extrapolated quantities. This requires reformulation in a form such as Eq. (120).

The specific descriptor of Eq. (120) was used in the Gaussian approximation potential (GAP) [352] for amorphous carbon [353], later extended for improved accuracy of crystalline phases [354]. This potential followed a specific three-step construction: In the first step, a (nonparametric, i.e. machine learned) pair-potential was fitted to a database using the scalar r_{ij} as a descriptor. The second step involved correcting this potential using a three-body term employing the descriptor given by Eq. (120). Finally, this potential was again corrected using the many-body SOAP descriptor described further below. Incorporating simple descriptors (including a pair descriptor) into the construction of the potential gave improved transferability. This basic potential without the SOAP descriptor can be regarded as a machine-learned variant of the cluster expansion, see Eq. (2), in an embedding medium.

A successful strategy for obtaining more generic descriptors is the use of atom-centered symmetry functions introduced by Behler and Parrinello [355,356]. The atom-centered fingerprints consist of radial symmetry functions

$$q_i^{(1,\alpha)} = \sum_{j \neq i} e^{-\eta^{(\alpha)} \left(R_{ij} - R_s^{(\alpha)} \right)^2} f_c^{(\alpha)} \left(r_{ij} \right) . \tag{121}$$

These functions are Gaussians centered at a pair distance $R_s^{(\alpha)}$ that are forced smoothly to zero at a cutoff distance via $f_c^{(\alpha)}(r)$. They are similar to coordination numbers, but put specific weight on atoms at the distance $R_s^{(\alpha)}$. In addition to these radial functions, there are angular functions

$$q_i^{(2,\beta)} = 2^{1-\varsigma^{(\beta)}} \sum_{j,k \neq i} \left(1 + \lambda^{(\beta)} \cos v_{ijk} \right)^{\varsigma^{(\beta)}}$$
$$e^{-\eta^{(\beta)}\left(r_{ij}^2 + r_{ik}^2 + r_{jk}^2\right)} f_c^{(\beta)} \left(r_{ij} \right) f_c^{(\beta)} \left(r_{ik} \right) f_c^{(\beta)} \left(r_{jk} \right). \tag{122}$$

We have added the superscripts (α) and (β) to emphasize, that those functions are not evaluated for a single parameter vector $(\eta, R_s, \lambda, \zeta)$ and cutoff function but for a set of these parameters. The fingerprint vector $\mathbf{q}_i$ is then constructed by combinations of $q_i^{(1,\alpha)}$ and $q_i^{(2,\beta)}$ over these parameters

sets. This allows one to construct fingerprint vectors of arbitrary dimension but those fingerprints may contain redundant information. In the following, we will refer to the components of the fingerprint vector as $q_i^{(\mu)}$ and drop the double notation $(1, \alpha)$, $(2, \beta)$, implying that the index (μ) also refers to the type of symmetry function and not just its parameters.

To emphasize connection to classical potentials, the $q_i^{(1)}$ look like the onsite ξ_{ii} terms occurring in our general functional form, Eq. (54), while the $q_i^{(2)}$ look like offsite ξ_{ij} terms. Indeed, if we replace the Gaussian in Eq. (121) with an exponential, the expression becomes identical to the local density of Ducastellium, see Eq. (59b). Similarly, $q_i^{(2)}$ are constructed like the many-body contributions in the SW and MEAM potentials.

Another important strategy for obtaining chemical fingerprint is the smooth overlap of atomic positions (SOAP) descriptor, introduced by Bartók, Kondor and Csányi [357]. It can be regarded as a generalization of the bond-orientational order parameter of Steinhardt, Nelson and Ronchetti [358]. The idea is to map the atomic configurations onto an atomic density field $\rho(r)$, which is given by atom-centered functions,

$$\rho(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i), \tag{123}$$

where δ is broadened to a Gaussian with a characteristic length ℓ_{at}. Note that $\rho(\mathbf{r})$ is invariant to permutations of atoms by construction and resembles the density constructed for EAM potentials. To obtain angular information, this density field is then expanded into a set of spherical harmonics $Y_{lm}(\theta, \phi)$ for each atom i,

$$\rho_i(\mathbf{r}) = \sum_{nlm} C_{i,nlm} g_n(r) Y_{lm}(\theta_i, \phi_i) \tag{124}$$

with local angles θ_i and ϕ_i of the vector $\mathbf{r} - \mathbf{r}_i$. The functions $g_n(r)$ are orthogonal radial basis functions. The power-spectrum of the angular coefficients $C_{i,nlm}$,

$$p_{i,nn'l} \propto \sum_m C_{i,nlm}^* C_{i,n'lm} \tag{125}$$

is objective and hence fulfills symmetry considerations. Selected elements of this power spectrum, typically chosen up to an upper value of n, form the SOAP descriptor.

SOAP taken to higher body order – and with $\ell_{at} \to 0$ as in Eq. (123) – forms the basis of the atomic cluster expansion (ACE) [359,360]. We recommend study of Ref. [359] also because it outlines the relationship

between current machine-learning approaches and more traditional EAM and bond-order potentials. First parameterizations of (performant) ACE potentials are starting to appear in the literature [361].

8.2. Regression

Once we have a fingerprint vector, we need to interpolate (or since we are in a high-dimensional space, rather extrapolate) atomic energies. This is the task of the regression framework. Regression frameworks take as input a fingerprint vector and predict *per-atom* energies U_i, such that the total energy is simply $U = \sum_i U_i$. This decomposition corresponds to the one used in all other potentials discussed in this review.

Regression can be as simple as a linear (ridge) regression. This is for example employed in spectral neighbor analysis potential [362], which uses the fingerprints introduced in the original GAP [352]. A more flexible regression strategy is the use of Gaussian process regression [252], which lends its name to the GAP potential [352]. An advantage of Gaussian processes is that there is a clear statistical interpretation of Gaussian process regression, which allows one to predict confidence intervals on energies in addition to expectation value or maximum probability estimates. In an active learning framework, this can help to identify regions in a database for which additional first-principles calculations (or experiments) are necessary [363]. The final expression for the per atom energies takes the simple form,

$$U_i = \sum_t \alpha_t K(\mathbf{q}_i, \mathbf{q}_t), \tag{126}$$

where the sum over t runs over all elements of the training database. The fingerprints $\mathbf{q}_i$ of the current environment of atom i are compared to all fingerprints $\mathbf{q}_t$ in the training database. The function $K(\mathbf{q}_i, \mathbf{q}_t)$ is the kernel that encodes a measure of distance in fingerprint space (and implicitly includes a mapping into a high-dimensional representation, typically called "feature space" in the ML literature [252]). From Eq. (126) it becomes immediately clear that the numerical effort of a GAP is linear in the number of points in the reference database. It is therefore crucial to *sparsify* the database, i.e. to remove entries that are close together in fingerprint space and do not improve the predictive power of the potential. Sparsification is highly nontrivial and we refer to the original literature on details [32].

A key feature of Gaussian process regression is that a notation of smoothness of the potential energy landscape is build-in. A common choice for the kernel function K is a squared exponential (Gaussian) [252,353], which containsa set of distance scales ℓ_α as a hyper parameter,

$$K(\mathbf{q}_i, \mathbf{q}_t) \propto \exp\left\{ \frac{(q_{i,\alpha} - q_{t,\alpha})^2}{2\ell_\alpha^2} \right\} \qquad (127)$$

The final potential energy is smoothed over ℓ_α, which avoids parasitic local energy minima or artifacted phonon spectra, which are sometimes found when regressing with neural networks [364], or even when simply interpolating with splines [221]. GAPs additionally make use of a dot-product kernel for the many-body (SOAP) contribution to the energy [353]. The dot-product kernel has no intrinsic length-scale, but the SOAP descriptors contain intrinsic smoothing through the width ℓ_{at} of the underlying atom-centered Gaussians. The expansion coefficients α_t are obtained using a Bayesian regression scheme that is described in detail in the pertinent literature [252].

An interesting use of Gaussian process regression is the on-the-fly parameterization [365] of a force field based on electronic structure calculations carried out on snapshots throughout a molecular dynamics trajectory. Li, Kermode and de Vita [366] used a small database of reference configurations, generated from past simulation steps of a molecular dynamics trajectory, to extrapolate forces to the future. Their scheme differs qualitatively from the ones discussed above by predicting forces rather than energies. The reason for this is that unlike the per-atom energy, a per-atom force is an observable that can be computed straightforwardly within first-principles techniques for a subset of atoms in a system. This leads to a nonconservative force field, yet use of this method to accelerate a quantum-mechanical region in concurrent multi-scale modeling [365], e.g. of fracture [367,368], introduces no additional complication as the coupling scheme itself is typically non-conservative.

We note that GAPs are linear models (in feature space), as the final energies are linear combinations of Gaussians. ACE potentials [359,369] introduce an additional nonlinear mapping. Using a square-root (as in EAM or TB2M potentials) reduced the number of basis functions required for high-accuracy regression for copper [369]. Fully nonlinear regression can be achieved with neural networks. Behler and Parrinello [355] employed the specific form

$$U_i = f_a^{(2)}\left\{ w_{01}^{(2)} + \sum_v w_{v1}^{(2)} f_a^{(1)}\left(w_{0v}^{(1)} + \sum_\mu w_{\mu v}^{(1)} q_i^{(\mu)} \right) \right\}, \qquad (128)$$

for regression in their neural network potential. Here, $f_a^{(1)}(x)$ and $f_a^{(2)}(x)$ are sigmoidal "activation functions" and $w_{v1}^{(1)}$ and $w_{v1}^{(2)}$ are weights. The weights are numbers that are adjusted in the process of fitting the neural network. The sum over μ runs over all entries of the feature vector of atom $\mathbf{q}_i$. This

network has a single "hidden layer" with activation function $f_a^{(1)}(x)$ and the index v runs over the outputs of this layer. The original fit for silicon by Behler and Parrinello employed three nodes in the hidden layer and 48 entries in the fingerprint vector, amounting to a total of 196 weights. Equation (128) resembles the many-body contribution U_m of classical potentials for open-shell systems, see Eq. (53).

9. Limiting the interaction range

Chemistry is local for the most part [370]. As a consequence, forces between distant atoms are minor, unless they are charged or carry dipoles. By neglecting interactions with distant atoms, much computing time can be saved at the expense of minor systematic errors, which can often be made marginal with mean-field or other coarse-graining corrections [371–373]. Unfortunately, the locality of chemistry cannot necessarily be specified in terms of cut-off radii, because to what extent two atoms "see" each other depends on various factors. Screening and cutting are somewhat related but the bonding topology and density must be known to express screening through cutting.

A metal atom will scarcely be exposed to the charge density of another atom if a third atom is sandwiched in between them. Baskes incorporated such screening conceptually into MEAM potentials [225,374] motivated by the observation that "*the second-nearest neighbors do not play a part even in the bcc structure,*" despite the ratio of nearest to next nearest-neighbor shell being as small as 1.15. He continued that "*this surprising result gives some support to the procedure we have used here in ignoring all but first-neighbor interactions*" [225] (in the description of crystals). Defects and disorder, both central to materials properties, place large demands on the assessment of bonding topology and make it necessary to define gray shades between bonded and non-bonded. In realistic descriptions, those cannot be judged using only the distance between two atoms.

Dispersive interactions between noble-gas atoms are not screened so that limiting their interaction range (without accounting for long-range corrections) is merely an exercise in finding a good compromise between accuracy and computational efficiency. Therefore, the task of cutting potentials might appear technical or even trivial at first sight. Doubling the cutoff radius r_c beyond which interactions are ignored requires roughly an eightfold computational effort in three spatial dimensions to construct neighbor lists and to evaluate pair potentials, which can be unjustifiable if the total dispersive interaction is already accurate. However, undesired *qualitative* artifacts can arise when cutting improperly. Brittle materials can turn ductile [375,376] and thereby make it on the cover of contemporary text books [377], or, nano-tube-

based water pumps work endlessly without external driving due to improper cut-offs [378], the report of which [379] having made it through the exquisite reviewing process of Nature journals. Nonetheless, cutting potentials is a prerequisite for the linear scaling of numerical effort per time step with the number of atoms N and for the efficient parallelization on parallel, high-performance computing systems using domain-decomposition [380–386]. Even long-range Coulomb potentials can be computed with a numerical effort scaling linearly or quasi-linearly with N [387–389], which requires atom-atom distance evaluations to be cut off at a finite distance.

Unfortunately, choosing optimum cut-off radii and procedures depends not only on the potential but also on the property of interest. Someone studying phase equilibria between different thermodynamic phases wants the potential depths to be reproduced as accurately as possible, which can be achieved with a brutal truncation of the potential at a given r_c without shifting. Implementing such a potential is readily done for Monte Carlo, however, for a molecular dynamics simulation, one would have to account for the δ-function singularity in the interaction force at r_c, which no sympletic integrator in the world would be pleased about. On the other hand, someone interested in mechanical properties can care less about the depth of energy minima but should instead be careful about potential curvature and the smoothness of the force at the cutoff. A bond whose tensile force disappears only linearly at r_c cannot be broken adiabatically, which risks to exaggerate loss moduli or viscosities: a spring attached to a bond would snatch into or out of bonding whenever the bond length r passes through r_c, whereby energy is dissipated [390].

Three more general points are worth discussing before cutting to the chase: First, when two phases like fcc and hcp have a similar thermodynamic potential, the choice of cutting can affect what phase is preferred [391]. However, at that point it might always be wise to consider other effects like many-body interactions and thermal or quantum fluctuations of the nuclei. Second, changing the cut-off procedure or cut-off radius may require a new parameterization of the potential. Third, and perhaps most importantly, not only potentials can be cut or smoothed but all operations discussed here below can also be applied to forces. However, cutting forces generally leads to non-conservative forces unless they originate from central, two-body potentials.

9.1. Cutting

9.1.1. Pair functions

Not only pair potentials must be cut off but also expressions that many-body potentials depend on, such as the charge density in EAM potentials, or the interaction between point multipoles. Cutting procedures for pair functions can be broadly categorized into cut-and-shift and smoothing.

Cut-and-shift procedures [392] redefine a pair potential $U_2(r)$ such that all its derivatives up to l'th order approach zero continuously at r_c and remain zero for $r > r_c$:

$$U_{2,l}^{\text{shift}}(r) = \{U_2(r) - \mathcal{T}[U_2(r), r_c, l]\}\Theta(r_c - r) \tag{129}$$

Here, $\mathcal{T}[U_2(r), r_c, l]$ is the l'th order Taylor expansion of $U_2(r)$ about r_c.

The effect of cut-and-shift procedures on $U_2(r)$ and on the equation of state is shown exemplarily for a Morse potential ($n = 6$, $r_c/r_0 = 2$, $l = 0, \ldots, 2$) in Figure 16a. It becomes evident that "large" l lead to a poor representation of the binding energy unless r_c is large, which, however costs computing time. However, serious artifacts can occur when potentials are truncated but also when they are merely shifted with $l = 0$. The EOS of an athermal crystal is discontinuous even for $l = 0$ as revealed in Figure 16b. Thermal fluctuations smear out and potentially eliminate this discontinuity. Nonetheless, caution remains advisable. In principle, a l'th-order shifted potential can have a discontinuity in $\partial^l p/\partial V^l$ so that a force-shifted potential risks to cause a discontinuity in the compressibility or any elastic tensor element including its imaginary, i.e. viscous part, as just discussed.

An alternative to cut-and-shift potentials is the use of cut-off or smoothing functions $f_c(r)$, which are multiplied with the true pair potential to yield

$$U_2^{\text{cut}}(r) = U_2(r)f_c(r). \tag{130}$$

The general property of smoothing functions is that they are unity below an inner radius, r_i, which may be zero or even negative. They decrease past r_i, approach zero at r_c, and remain zero ever after. Formally they can be expressed as

$$f_c(r) = \Theta(r_i - r) + \varphi_c(x)\Theta(r - r_i)\Theta(r_c - r) \tag{131}$$

with $x = (r - r_i)/(r_c - r_i)$. Many of the lessons learned for cut-and-shift applies to smoothing, in particular w.r.t. what artifacts are produced depending on the order l with which derivatives of the potential disappears at r_c. However, an additional risk of smoothing functions is that using "large" l does not guarantee artifacts to be suppressed in particular when r_i is close to r_c. Prominent smoothing functions $\varphi_c(x)$ are [216,217,224,227,242,393]:

$$\{1 + \cos(\pi x)\}/2 \quad l_i = 1; \quad l_c = 1 \quad \text{Tersoff} \tag{132a}$$

$$(1 - x)^3(1 + 3x + 6x^2) \quad l_i = 2; \quad l_c = 2 \quad \text{MEAM} \tag{132b}$$

$$\{8 + 9\cos(\pi x) - \cos(3\pi x)\}/16 \quad l_i = 2; \quad l_c = 2 \quad \text{Murty water} \tag{132c}$$
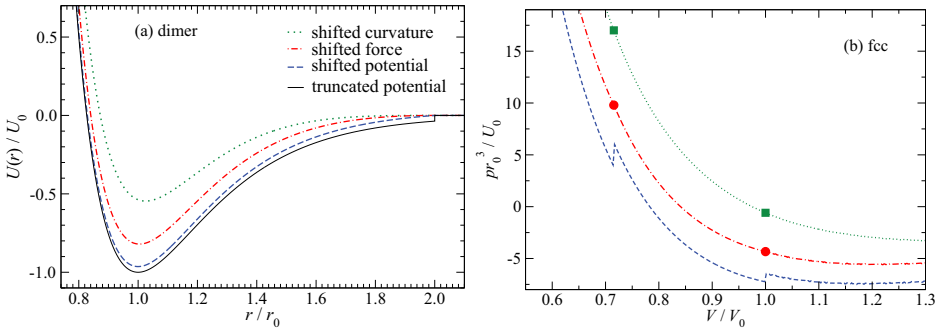
Figure 16. Left panel: Truncated (black), cut-and-shift (blue), shifted-force (red), and shifted-curvature (green) Morse potential with exponent $n = 8$ and a cut-off distance of $r_c = 2.0 r_0$. The cut-and-shift procedures include $l = 0$ (blue), $l = 1$ (red), and $l = 2$ (green). Right panel: energy per atom of an ideal fcc crystal as a function of the nearest-neighbor distance a_0/r_0 using the shifted potential (blue), shifted force (red), and shifted curvature (green) for the Morse potential with $n = 8$ and a cutoff radius $r_c = 2.0 r_0$. Dots and squares denote reduced volumes at which a neighbor-shell radius is equal to r_c.

$$\exp\{-(2x)^m\} \qquad l_i = m - 1 \ (\text{unknown}), \qquad (132\text{d})$$

$$\exp(\lambda + \lambda/(x - 1)) \quad l_i = 0; \quad l_c = \infty \quad \text{Stillinger \& Weber} \quad (132\text{e})$$

$$2(1 - x)^m/(1 + (1 - x)^n) \quad l_i = 0; \quad l_c = m - 1 \quad \text{Mishin} \quad (132\text{f})$$

where integer Mishin exponents are commonly selected as $m = n = 4$. In Eq. (132), l_i and l_c specify which highest l'th-order derivative of the smoothed function approaches zero at r_i and r_c, respectively. Moreover, $l_i = 0$ means that $\varphi_c(r_i) = 1$ while $\varphi'_c(0) \neq 0$, while omitting l_c means that $\varphi_c(x = 1^-) \neq 0$. Recently, it was found that cutting with polynomials constructed such that $l_i = l_c + 1$ can simultaneously improve binding energies while reducing cut-off artifacts [394]. This is because cut-off induced discontinuities of (the derivatives of) short-range potentials are distinctly larger at r_i than at r_c.

9.1.2. Cutting Coulomb potentials

Wolf *et al*. [395] suggested that the Coulomb interaction in an overall neutral system can be cut off without producing systematic errors. The idea is to represent the Coulomb interaction through the Ewald summation [396], keep the $q = 0$ contribution (which could be seen as a kind of mean-field correction), and to ignore any contribution at non-zero wave vector $\mathbf{q}$. Some authors find support for the correctness of Wolf's optimistic assessment [285,397], while others don't [398,399]. In fact, small cutoff radii can be used for homogeneous systems when using appropriate cut-off functions, e.g. clearly less than $10 \mathring{A}$ in silica melts [394]. However, problems arise when structural heterogeneities extend on length scales $\lambda > r_c$ [394]. They induce

undulations in chemical potentials with wave vectors of order $q = 2\pi/\lambda$, which favor a charge separation. Coulomb interactions suppress that separation with terms scaling as $1/q^2$. The pertinent restoring forces are neglected for the most part in the Wolf summation.

Another argument for why the Wolf summation cannot be cutoff in heterogeneous systems can be made when considering a gold nugget separated by a large distance from a chunk of lithium. If electrons can equilibrate, gold will acquire a negative charge due to its larger electronegativity. While the charge density on each lump will decrease with their size, the absolute force attracting them will increase. This mechanism also matters for discharge simulations of batteries or related devices [157], despite the presence of a Helmholtz double layer near the electrodes. If a Maxwellian demon were to keep all electrolyte atoms in place and closed an electrical switch allowing some charge to flow between anode and cathode, a current would flow, which would extend much beyond an elementary charge. Thus, local charge neutrality cannot be obeyed during the entire time of the discharge process. Once the demon releases the electrolyte particles, discharge can continue after the electrolyte rejuvenated the double layers. Modeling the described process with a Wolf summation properly would require one to use a cut-off exceeding the anode-cathode separation. Thus, the Wolf summations, or better generalizations thereof avoiding a discontinuous force at r_c [285,394], can be made for homogeneous systems, but steer clear of it otherwise, in particular in any biophysical context [286].

9.1.3. Many-body potentials

Smoothing functions are not only applied to pair potentials but also in many-body potentials, sometimes with the goal to imitate screening. In a crystalline structure, interactions with atoms in the nearest-neighbor shell should be unmodified but those with the next-nearest neighbors should be screened – often those are approximated to be fully screened. Figure 17 illustrates the effect of cutting the interaction in a Ducastelle potential at various distances. When cutting with a distance-based criterion, energies (Figure 17a) are forced to drop to zero somewhere between first and second neighbor shell of the equilibrium crystal (since we do not want interaction with second neighbors at equilibrium - they are screened). This leads to a force or pressure overshoot during homogeneous volumetric dissociation of the crystal, illustrated in Figure 17b. Attempting to break such a crystal with an external stress overestimates the critical stress for fracture by a factor of four. For this reason, some works have adopted cutting forces rather than energies for modeling fracture [400], despite the fact that this leads to a nonconservative force-field. The proper solution to

this issue is the use of screening functions, discussed in the next section. They yield the true nearest-neighbor potential (shown by the black line in Figure 17) and hence smooth dissociation of the crystal.

9.2. Screening

In order to reflect screening through screening rather than through cutting, Baskes *et al.* [225,374] proposed to replace cutoff functions with environment-dependent screening functions, S_{ij}, which themselves are products of three-body functions $0 < S_{ijk} < 1$,

$$S_{ij} = \prod_{k i,j} S_{ijk} \tag{133}$$

The S_{ijk} are constructed such that k when atom k screens the interaction between i and j completely, while $S_{ijk} = 1$ when atom k does not screen the ij interaction at all. Thus, the screening function acts like a topological cutoff function, $0 \leq S_{ij} \leq 1$. Baskes proposed different functional forms for S_{ijk}. The refined expression is [374]

$$S_{ijk} = \begin{cases} 1 & \text{if} \quad C_{ijk} \leq C_{\min} \\ \exp\left\{-\left[(C_{\max} - C_{ijk})/(C_{ijk} - C_{\min})\right]^2\right\} & \text{if} \quad C_{\min} < C_{ijk} < C_{\max}, \\ 0 & \text{if} \quad C_{ijk} \geq C_{\max} \end{cases} \tag{134}$$

which drops from unity to zero between $C_{\max}$ and $C_{\min}$. The quantity C_{ijk} characterizes the geometry of the ellipse that passes through the three atoms i-j-k. It is given by
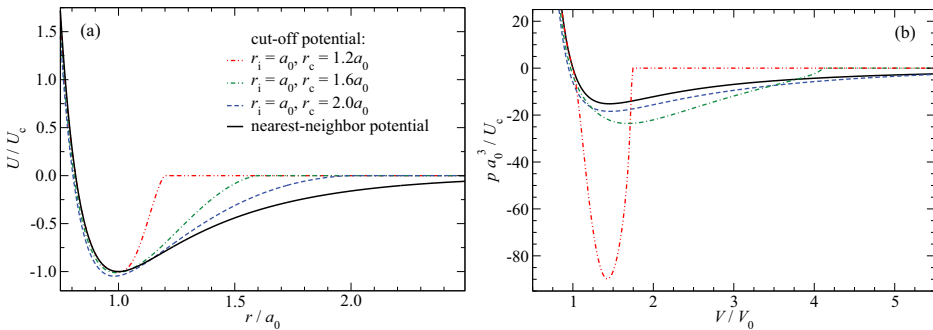
Figure 17. (a) Energy per atom as a function of distance and (b) pressure as a function of volume for a Cu fcc crystal described by a Ducastelle potential, respectively [222]. The MEAM cutoff function given by Eq. (132b) was applied to the repulsive potential and the density. U_c and a_0 are the cohesive energy and the nearest-neighbor distance for an fcc crystal described by Decastelle potential, respectively.

$$C_{ijk} = \frac{2(X_{ik} + X_{jk}) - (X_{ik} - X_{jk})^2 - 1}{1 - (X_{ik} - X_{jk})^2} \ , \tag{135}$$

the square of the ratio of the lengths of the two half-axes of that ellipse. Here $X_{ik} = r_{ik}/r_{ij}$ is the distance between atoms i-k normalized by the length of the central i-j bond.

Figure 18 illustrates this concept. The bond is unscreened ($S = 1$) if there is no third atom in its vicinity (Figure 18a). As a third atom moves into the shaded ellipsis (Figure 18b – the geometry of the ellipsis is described by the values of C_{min} and C_{max}) the bond becomes screened ($S = 0$, Figure 18c). It becomes clear, that if we simply rescale all coordinates (as in the homogeneous, volumetric dissociation of a crystal), S remains constant for each bond and an energy would follow the true nearest neighbor curve, see the black line in Figure 17(a). The screening function S is hence a scale-invariant formulation of a "cutoff" procedure.

The Baskes screening functions [374] were applied to empirical bond-order potentials independently by Pastewka *et al.* [25,376,401] and Kumagai *et al.* [402]. Both groups emphasized that the screened potentials significantly improved the properties of amorphous carbon modeled with REBO2 or Tersoff-type potentials. In addition, the screening functions served to overcome the issue with dissociation of a bond under external stress discussed in Refs. [25,375,376,401]. This enabled modeling of fracture in crystalline and amorphous carbon systems [403]. Perriot *et al.* [404] presented a slightly different screening concept requiring the REBO potential to be refitted.

Screening functions can be rationalized as originating from nonorthogonality in a tight-binding framework. This nonorthogonality leads to an environment-dependence of the bond-integrals, when the nonorthogonal tight-binding is "coarse-grained" to an equivalent orthogonal tight-binding model. Nguyen-Manh, Pettifor and Vitek [405,406] showed, that the theory of the bond-order expansion, briefly touched upon in Sect. 5.8, can be used to derive screening functions from a nonorthogonal tight-binding model. This first-principles construction lends additional support to Baskes' screening functions and other screening approaches, such as empirical environment-dependence introduced in the context of orthogonal tight-binding shortly after Baskes work [407–410].

10. Summary and perspectives

Eugene Wigner allegedly said: "*It is nice to know that the computer understands the problem. But I would like to understand it too.*" One main motivation for writing this review was to assist people with similar
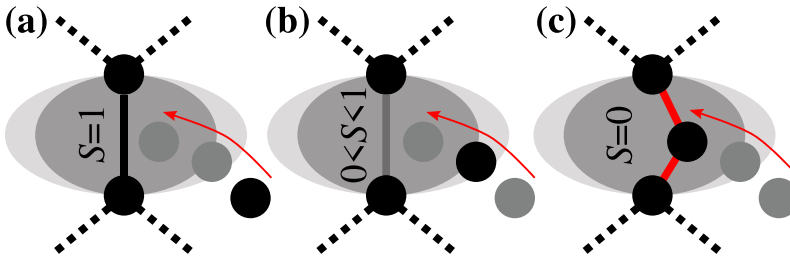
Figure 18. Illustration of the screening concept. The black atom moves along the hypothetical trajectory indicated by the shaded atoms and the red arrow. (a) Initially, the thick, black bond is unscreened, but as (b) the atom enters the region of influence (light gray ellipses) the bond weakens. (c) The atom has moved into the close vicinity of the bond (dark gray region), effectively disabling it while creating two new bonds (red lines). The Baskes screening functions [374] are defined the aspect ratio of the light gray region (C_{max}) and the dark gray (C_{min}), defining a measure of bond screen that is independent of the absolute lengths of the bonds in the system.

ambitions as Wigner. To this end, we summarized our understanding of what properties in condensed-matter systems can be induced by the functional form of the potentials used for their description. In our endeavor, we felt compelled to create much own data and new figures with the purpose to create insight and to convey trends and differences between potential classes rather than to produce numbers for a specific system. When doing so, we did our best to embed anything written into a historical context, which is summed up in Figure 19.

Yet many times, we could not find appropriate references or quotes, which we are certain do exist. As one of numerous examples, we found many papers computing shear and bulk moduli of metals or vacancy-defect and cohesive energies, but always missed the argument why their respective ratios are correlated and how they relate to the ratio of melting and boiling temperature. We expect our discussion to have satisfied Wigner's desire for understanding the correlation between these ratios. Despite certainly having missed well known studies, we did find some old works, which may have been underappreciated, such as Slater's paper on the interaction between helium atoms [82]. As mentioned earlier, Slater derived the exponential repulsion between atoms with closed valence shell, which promoted Born and Mayer [83] as well as Buckingham [7] to use this or slightly modified forms for the repulsion in the potentials now carrying their names. However, Slater also derived the dispersive coefficient for helium to within 15% accuracy, two years before London [56] generalized the results to other closed-valence shells.

Despite the length of this article, we could only scratch the surface of the large field of interatomic potentials. Many central aspects were not touched upon. Most importantly, we barely discussed how to adjust parameters, in
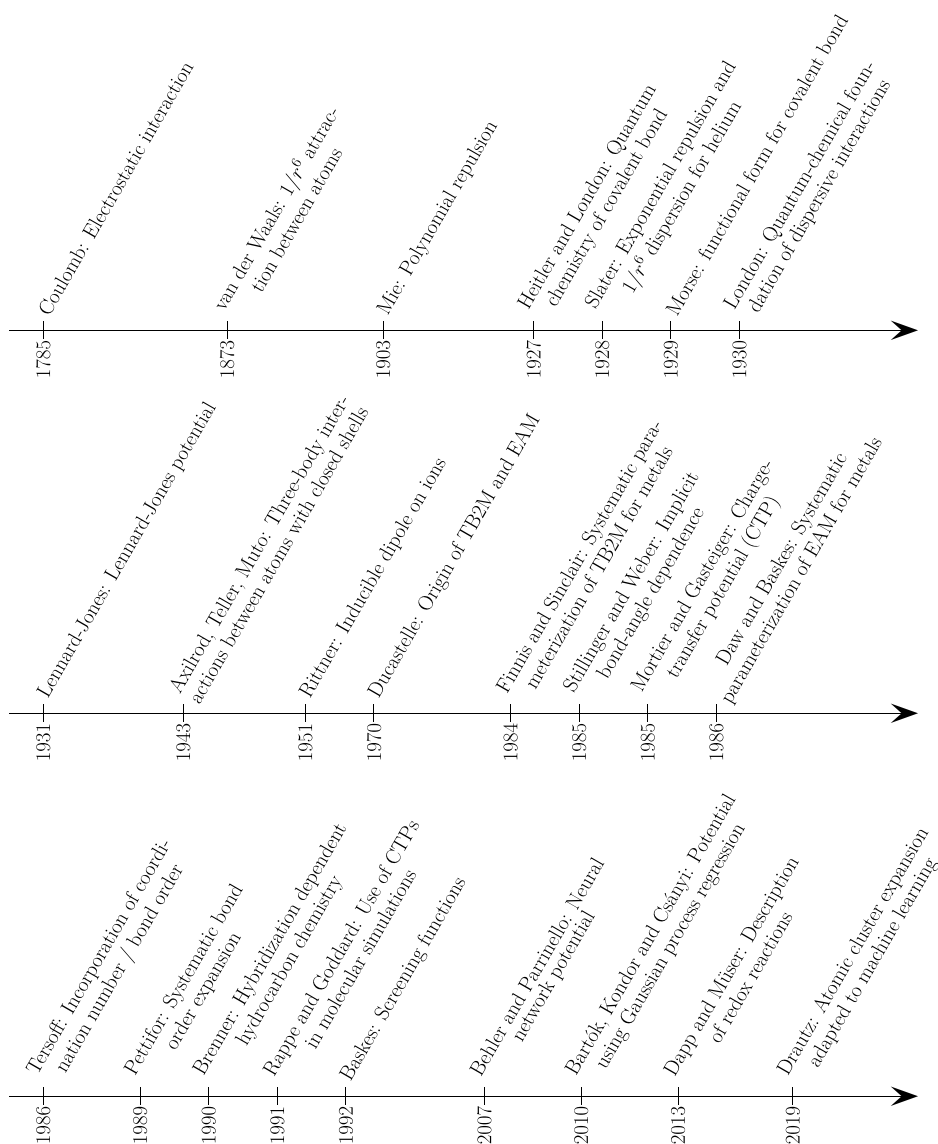
Figure 19. Selected highlights of the development of interaction potentials.

particular the pros and cons to fit to experimental or to *in-silico* data. There are good reasons to follow the main-stream opinion that the quality of a potential increases the less empirical the data on which the potential is parameterized. Computer-generated reference data is much more versatile than that provided from experiments. Forces on individual atoms can be used for characteristic bonding situations or rare but important configurations like a transition state occurring during a chemical reaction or a collective phase transformation [411]. Moreover, *in-silico* data does not contain quantum effects, which frequently need to be accounted for when

comparing computer-generated data to experiments. Describing how to do that properly would have required us to outline how to approximately subtract the nuclear quantum effects from experimental data or how to incorporate quantum fluctuations into the simulations, e.g. through path-integral techniques, or, by encoding their effect into effective temperature-dependent, many-body potentials, which would have been beyond the scope of this review.

Thus, there is scarcely any argument to gauge parameters on experiments, if there was not the small but important detail that experimental data is by and large more accurate than density-functional theory, which cannot be deemed exact, as long as the exact functional has not been identified. It could be argued that we base one theory or potentials with uncontrolled approximations on another one with uncontrolled approximations, which has trouble to predict two dislike molecules or clusters separated by a large distance to each acquire an integer charge [412,413]. When dismissing empiricism as fundamentally problematic, one may also keep in mind that one of the greatest theoretical achievements in chemistry, arguably in all of science, was the construction of the periodic table by Mendeleev. He even predicted the existence of unknown elements including some of their physical and chemical properties with an accuracy that people using potentials or even DFT might have a hard time to match if they did not know what they had to predict, or, rather postdict. Moreover, the amount of data that Mendeleev could build on was noticeably less than what is required in machine learning.

The potentials discussed in this review pertain mostly to situations, in which bonds can be clearly classified as dispersive, metallic, covalent, or ionic. For situations, where this simple categorization cannot be made, different potential classes are combined in a mix-and-match fashion into compound potentials. Prominent examples are the adaptive intermolecular REBO (AIREBO) [251] (combining Brenner's potential with nonbonded Lennard-Jones interaction), the Streitz-Mintmire potential [414] (combining EAM with charge transfer), the charge-optimized many-body potential (COMB) [415–417] (combining Tersoff's potential with charge transfer), a merger of REBO with split-charge equilibration [330,418], as well as early combinations of Keating-type with charge-transfer potentials potentials [419,420]. Of course, the widely-used ReaxFF potential [12,421], which merges a bond-order approach (different in nature than the approaches discussed in this review) with non-bonded interactions and charge transfer, must also be mentioned.

While compound approaches can be extremely powerful, many of them simply add different energy terms. This can be problematic even for seemingly simple alloys or intermetallics formed by elements of large electronegativity difference. Put simply, negatively charged atoms grow in size

while positively charged atoms shrink. This symmetry breaking between negative and positive charge is not reflected when simply adding charge equilibration to a (post) Ducastelle potential. Yet, it is supposedly responsible for why the negatively charged atoms in intermetallics have the tendency to close pack while positive atoms occupy interstitial positions, as it happens, for example, for Al_2Au, also called the purple plague: Au atoms form an fcc lattice while Al atoms assume interstitial positions. Although promising steps toward true compound potentials have been taken [422], e.g. by augmenting or reducing the valence density of a neutral atom with a term proportional to its partial charge, systematically merged potential remain a dream.

Novel paths that are taken with machine learning potentials seem extremely promising. However, a puzzling question is why machine-learned potentials outperform parameterized potentials. The claim that they are parameter free or free of functional constraints is not entirely justified. Many of the local descriptors are suspiciously close to what is used in potentials, as indeed they are often "physics-inspired" [350]. However, the big advantage of MLPs is that they do not make strong assumptions like pair-wise additive repulsion, which might be one of the most important sources of error in classical interatomic potentials.

A show-stopping problem central to all potentials is the curse of dimensionality. Fitting multi-species (or alloy) potential requires a number of pair-parameters that scales asymptotically as N_s^2 with the number of atomic species N_s. The scaling becomes even less favorable if we need specific parameters for triplets ($\propto N_s^3$), quadruplets ($\propto N_s^4$) and so on, quickly becoming intractable for a large number of species. The compression of chemical fingerprints has recently been proposed to circumvent the curse of dimensionality for MLPs [423]. Using explicit functional forms, it can be possible to circumnavigate the curse of dimensionality with combining rules. However, they are only available for few interactions types and may be plagued with poor transferability.

As a final note, we would like to point out that despite the fact that (with the exception of bare Coulomb interaction) all potentials discussed here are local, chemistry can be quite non-local. By non-locality we do not mean the range of the bare interaction, such as the range of the bond-integrals in a tight-binding formulation. We mean the non-locality intrinsic in the diagonalization of the quantum mechanical Hamiltonian. In hydrocarbon chemistry, the non-locality manifests itself for example in bond conjugation and in metals through an algebraic decay of the density matrix [424], while in group 15–17 in the periodic table, it is reflected in the Peierls deformation causing

elemental crystals to reduce from the simple cubic to less symmetric structures [425]. As another example, carbon chains – also called carbynes – can exist in a polyynic form of alternating single and triple bonds or a cumulenic form of repeating double bonds [426–428]. Which form is chosen depends on whether the chain is odd or even numbered and how it is terminated. This crucially affects how they interact with their environment, for example with oxygen [429,430]. Such non-local effects even manifest in bulk materials: Force-locality tests on amorphous carbon by Deringer and Csányi showed that chemistry in low-density, graphite-like amorphous carbon is much longer ranged than in denser more diamond-like carbon [353]. Approaches for incorporating true quantum non-locality into potentials currently do not appear to exist. Modeling it appears to require new classes of potential, e.g. the ability of an EAM or MEAM potential to make atoms adjust their donating charge density in response to the environment in a fashion that allows for multistability.

We hope that this review was successful in highlighting the incredible achievements throughout the last century in understanding the bonding of matter, and molding these insights into simple analytical expressions. The wide availability of high-accuracy electronic structure calculations and advances in statistical modeling have moved the field into exciting new directions. We would also like to add that the wide availability of present-day interatomic potentials in the form of open-source software, ideally embedded in a standard database [13] or a standard code [380,381], is accelerating quick adoption of potentials into practice — not to mention the savings in students' lifetimes, by not having to dissect which of the 50 parameters just manually copied from printed publication XYZ is missing a 0 in print. (Yes, we are thinking about our own PhD theses.) Of course, significant challenges remain, both for traditional fixed-form as well as machine-learned interatomic potentials, of which we believe the curse of dimensionality and the coupling of electron transfer and Coulomb interaction to the electronic bond as most crucial [431].

Notes

1. A large set of these tabulated potentials files can be found in the NIST interatomic potentials repository at https://www.ctcms.nist.gov/potentials/.
2. Functions of the Hamitonian are defined through their action on eigenfunctions, $f(\hat{H})|n\rangle = f(\varepsilon_n)|n\rangle$, or through a power series expansion of f.
3. See https://libatoms.github.io/GAP/data.html for links to high-quality *ab-initio* databases for C, Si, and other materials.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the author(s).

Nomenclature/Notation

Abbreviations
ACE atomic cluster expansion
ATM Axilrod-Teller-Muto interaction potential
DFT density-functional theory
EAM embedded-atom method
EOS equation of state
GAP Gaussian approximation potential
LJ Lennard-Jones
MEAM modified EAM
ML machine learning
MLP machine-learned potential
QEq charge equilibration
REBO reactive empirical bond-order potential
REBO2 second generation REBO
SQE split-charge equilibration
SW Stillinger-Weber
TB, TBnM tight-binding, TB n-th order moment expansion

Symbols
α, α' polarizability in SI and atomic units
α_{M} Madelung constant
$\delta_{\alpha\beta}$ Kronecker delta
ϵ Lennard-Jones energy parameter
ε_0 vacuum permittivity
ε_{r} dielectric constant
$\varepsilon_{\alpha\beta}$ element of the Eulerian strain tensor
$\eta_{\alpha\beta}$ element of the Lagrangian strain tensor
ρ charge density, number density
σ length scale parameter
$\sigma_{\alpha\beta}$ element of the Cauchy stress tensor
A electron affinity
B_n n-th order virial coefficient
$C_{\alpha\beta\gamma\delta}$ element of elastic tensor
C_{ij} element of elastic tensor in Voigt notation
C_n dispersion coefficient of order n

$\hat{H}$ Hamilton operator

$H_{i\alpha j\beta}$ Hamiltonian integral between orbital α on atom i and orbital β on atom j

I ionization energy

B bulk modulus

N particle number

Q_i charge of atom i

S_{ij} square of the distance between atoms i and j

U interaction energy

U_0 dimer/molecular binding energy

U_{pa} potential energy per atom

$U_{\mathrm{pa}}^{\mathrm{eq}}$ U_{coh} equilibrium potential energy per atom (cohesive energy)

U_{pb} potential energy per bond

$U_{\mathrm{pb}}^{\mathrm{eq}}$ equilibrium potential energy per bond

U_1 $U_1^{(i)}$ single body interaction energy

U_2 U_3 pair and triplet interaction energy

V volume

Z_0 coordination number

Z_s number of atoms in s'th nearest-neighbor shell

a_n distance between an atom with a $(n+1)$-nearest neighbor

a_0^{eq} equilibrium bond length

a_{B} Bohr radius

f_{c} cutoff function

$k_B T$ thermal energy

m mass

$n(\varepsilon)$ density of states

$n_{i\alpha}(\varepsilon)$ local density of states of orbital $i\alpha$

$\hat{p}_i$ momentum operator

$\mathbf{p}_i$ p_i dipole moment of species i and its magnitude

p pressure

$\mathbf{q}_i$ descriptor of the environment of atom i

q_{ij} bond charge donated from atom i to atom j

$r,$ r_{ij}(pair) distance

r_0 equilibrium distance in a diatomic molecule

r_{c} cutoff radius

$v_s^{\alpha\beta},\ v_s^{\alpha\beta\gamma\delta}$ second- and fourth-rank shell tensor for the s'th nearest-neighbor shell

v_{pa} volume per atom

References

[1] Feynman RP, Leighton RB, Sands M. The Feynman lectures on physics. New York: Addison-Wesley; 1964.

[2] Guggenheim EA. The principle of corresponding states. J Chem Phys. 1945;13:253–261.

[3] Kadanoff LP, Götze W, Hamblen D, et al. Static phenomena near critical points: Theory and experiment. Rev Mod Phys. 1967;39:395–431.

[4] Fisher ME. The theory of equilibrium critical phenomena. Rep Prog Phys. 1967;30:615–730.

[5] Kremer K, Grest GS. Dynamics of entangled linear polymer melts: a molecular-dynamics simulation. J Chem Phys. 1990;92:5057–5086.

[6] Morse PM. Diatomic molecules according to the wave mechanics. II. Vibrational levels Phys Rev. 1929;34:57–64.

[7] Buckingham RA. The classical equation of state of gaseous helium, neon and argon. Proc Roy Soc Lond Math Phys Sci. 1938;168:264–283.

[8] Lennard-Jones JE. Cohesion. Proc Phys Soc. 1931;43:461–482.

[9] Daw MS, Baskes MI. Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals. Phys Rev Lett. 1983;50:1285–1288.

[10] Brenner DW. Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. Phys Rev B. 1990;42:9458–9471.

[11] Brenner DW, Shenderova OA, Harrison JA, et al. A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons. J Phys Condens Matter. 2002;14:783–802.

[12] van Duin ACT, Dasgupta S, Lorant F, et al. ReaxFF: a reactive force field for hydrocarbons. J Phys Chem A. 2001;105:9396–9409.

[13] Tadmor EB, Elliott RS, Sethna JP, et al. The potential of atomistic simulations and the knowledgebase of interatomic models. JOM. 2011;63:17.

[14] Becker CA, Tavazza F, Trautt ZT, et al. Considerations for choosing and using force fields and interatomic potentials in materials science and engineering. Curr Opin Solid State Mater Sci. 2013;17:277–283.

[15] Trautt ZT, Tavazza F, Becker CA. Facilitating the selection and creation of accurate interatomic potentials with robust tools and characterization. Model Simulat Mater Sci Eng. 2015;23:074009.

[16] de Tomas C, Suarez-Martinez I, Marks NA. Graphitization of amorphous carbons: A comparative study of interatomic potentials. Carbon. 2016;109:681–693.

[17] Hale LM, Trautt ZT, Becker CA. Evaluating variability with atomistic simulations: the effect of potential and calculation methodology on the modeling of lattice and elastic constants. Model Simulat Mater Sci Eng. 2018;26:055003.

[18] de Tomas C, Aghajamali A, Jones JL, et al. Transferability in interatomic potentials for carbon. Carbon. 2019;155:624–634.

[19] Carlsson A. Beyond pair potentials in elemental transition metals and semiconductors. Solid state physics. Elsevier; 1990. p. 1–91.

[20] Brenner D. The art and science of an analytic potential. Phys Status Solidi (b). 2000;217:23–40.

[21] Ackland GJ. Comprehensive nuclear materials. In: Konings RJ, editor. 1.10 - interatomic potential development. Oxford: Elsevier; 2012. p. 267–291.

[22] Sinnott SB, Brenner DW. Three decades of many-body potentials in materials research. MRS Bull. 2012;37:469–473.

[23] Finnis M. Concepts for simulating and understanding materials at the atomic scale. MRS Bull. 2012;37:477–484.

[24] Foiles SM, Baskes MI. Contributions of the embedded-atom method to materials science and engineering. MRS Bull. 2012;37:485–491.

[25] Pastewka L, Mrovec M, Moseler M, et al. Bond order potentials for fracture, wear, and plasticity. MRS Bull. 2012;37:493–503.

[26] Shin YK, Shan TR, Liang T, et al. Variable charge many-body interatomic potentials. MRS Bull. 2012;37:504–512.

[27] Plimpton SJ, Thompson AP. Computational aspects of many-body potentials. MRS Bull. 2012;37:513–521.

[28] Liang T, Shin YK, Cheng YT, et al. Reactive potentials for advanced atomistic simulations. Annu Rev Mater Res. 2013;43:109–129.

[29] Behler J. Perspective: Machine learning potentials for atomistic simulations. J Chem Phys. 2016;145:170901.

[30] Harrison JA, Schall JD, Maskey S, et al. Review of force fields and intermolecular potentials used in atomistic computational materials research. Appl Phys Rev. 2018;5:031104.

[31] Deringer VL, Caro MA, Csányi G. Machine learning interatomic potentials as emerging tools for materials science. Adv Mater. 2019;31:1902765.

[32] Deringer VL, Bartók AP, Bernstein N, et al. Gaussian process regression for materials and molecules. Chem Rev. 2021;121:10073–10141.

[33] Mishin Y. Machine-learning interatomic potentials for materials science. Acta Mater. 2021;214:116980.

[34] Finnis M. Interatomic forces in condensed matter. Oxford: Oxford University Press; 2005.

[35] Tadmor EB, Miller RE. Modeling materials. Cambridge University Press; 2009.

[36] Weiner PK, Kollman PA. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. J Comput Chem. 1981;2:287–303.

[37] MacKerell AD, Wiorkiewicz-Kuczera J, Karplus M. An all-atom empirical energy function for the simulation of nucleic acids. J Am Chem Soc. 1995;117:11946–11975.

[38] Pearlman DA, Case DA, Caldwell JW, et al. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Comm. 1995;91:1–41.

[39] Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc. 1996;118:11225–11236.

[40] MacKerell AD, Bashford D, Bellott M, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B. 1998;102:3586–3616.

[41] Mackerell AD, Feig M, Brooks CL. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem. 2004;25:1400–1415.

[42] Case DA, Cheatham TE, Darden T, et al. The amber biomolecular simulation programs. J Comput Chem. 2005;26:1668–1688.

[43] Salomon-Ferrer R, Case DA, Walker RC. An overview of the amber biomolecular simulation package. Wiley Interdiscip Rev Comput Mol Sci. 2012;3:198–210.

[44] Walters ET, Mohebifar M, Johnson ER, et al. Evaluating the london dispersion coefficients of protein force fields using the exchange-hole dipole moment model. J Phys Chem B. 2018;122:6690–6701.

[45] Daw MS, Baskes MI. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. Phys Rev B. 1984;29:6443–6453.

[46] Ducastelle F. Modules élastiques des métaux de transition. J Phys. 1970;31:1055–1062.

[47] Finnis MW, Sinclair JE. A simple empirical N-body potential for transition metals. Phil Mag A. 1984;50:45–55.

[48] Sangster M, Dixon M. Interionic potentials in alkali halides and their use in simulations of the molten salts. Adv Phys. 1976;25:247–342.

[49] Barker J, Pompe A. Atomic interactions in argon. Aust J Chem. 1968;21:1683.

[50] Thiesen M. Untersuchungen über die Zustandsgleichung. Ann Phys. 1885;260:467–492.

[51] Masters AJ. Virial expansions. J Phys Condens Matter. 2008;20:283102.

[52] Mayer JE, Montroll E. Molecular distribution. J Chem Phys. 1941;9:2–16.

[53] Chandler D, Andersen HC. Optimized cluster expansions for classical fluids. II. Theory of molecular liquids. J Chem Phys. 1972;57:1930–1937.

[54] Lebowitz JL, Penrose O. Convergence of virial expansions. J Math Phys. 1964;5:841–847.

[55] van der Waals JD. Over de Continuiteit van den Gas- en Vloeistoftoestand [dissertation]. Leiden: Leiden University; 1873.

[56] London F. Zur Theorie und Systematik der Molekularkräfte. Z Phys. 1930;63:245–279.

[57] Israelachvili JN, Tabor D. Measurement of van der Waals dispersion forces in the range 1.4 to 130 nm. Nat Phys Sci. 1972;236:106.

[58] Israelachvili JN, Tabor D. The measurement of van der Waals dispersion forces in the range 1.5 to 130 nm. Proc R Soc A: Math Phys Eng Sci. 1972;331:19–38.

[59] Reith D, Pütz M, Müller-Plathe F. Deriving effective mesoscale potentials from atomistic simulations. J Comput Chem. 2003;24:1624–1636.

[60] Lyubartsev AP, Laaksonen A. Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. Phys Rev E. 1995;52:3730–3737.

[61] Sukhomlinov SV, Müser MH. A mixed radial, angular, three-body distribution function as a tool for local structure characterization: Application to single-component structures. J Chem Phys. 2020;152:194502.

[62] Abell GC. Empirical chemical pseudopotential theory of molecular and metallic bonding. Phys Rev B. 1985;31:6184–6196.

[63] Heine V, Robertson IJ, Payne MC, et al. Many-atom interactions in solids. Philos Trans R Soc A. 1991;334:393–405.

[64] Gupta RP. Lattice relaxation at a metal surface. Phys Rev B. 1981;23:6265–6270.

[65] Capecchi D, Ruta G, Trovalusci P. From classical to Voigt's molecular models in elasticity. Arch Hist Exact Sci. 2010;64:525–559.

[66] Love AEH. Mathematical theory of elasticity. Cambridge UK: Cambridge University Press; 1954.

[67] Voigt W. Lehrbuch der Kristallphysik. Leipzig: B.G. Teubner; 1910.

[68] Simmons G, Wang H. Single crystal elastic constants and calculated aggregate properties: a handbook [by] Gene Simmons and Herbert Wang. 2nd ed. Cambridge Mass: M.I.T. Press; 1971.

[69] James BW, Kheyrandish H. The low-temperature variation of the elastic constants of lithium hydride and lithium deuteride. J Phys C Solid State Phys. 1982;15:6321–6337.

[70] Fu H, Liu W, Gao T. Atomistic simulation of MgS polymorphs. Phys Status Solidi (b). 2010;247:48–53.

[71] Lee BH. Elastic constants of ZnTe and ZnSe between 77°–300°K. J Appl Phys. 1970;41:2984–2987.

[72] Slagle OD, McKinstry HA. Temperature dependence of the elastic constants of the alkali halides. III. CsCl, CsBr, and CsI. J Appl Phys. 1967;38:451–458.

[73] Carpenter MA, Salje EKH, Graeme-Barber A, et al. Calibration of excess thermodynamic properties and elastic constant variations associated with the alpha <−> beta phase transition in quartz. Am Mineral. 1998;83:2–22.

[74] Du J, Wen B, Melnik R, et al. Phase stability, elastic and electronic properties of Cu–Zr binary system intermetallic compounds: A first-principles study. J Alloy Comp. 2014;588:96–102.

[75] Born M, Huang K. Dynamical theory of crystal lattices. London: Oxford Clarendon Press; 1954.

[76] Aziz RA, Slaman M. The argon and krypton interatomic potentials revisited. Mol Phys. 1986;58:679–697.

[77] Booth GH, Cleland D, Thom AJW, et al. Breaking the carbon dimer: The challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods. J Chem Phys. 2011;135:084104.

[78] Heitler W, London F. Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. Z Phys. 1927;44:455–472.

[79] Sugiura Y. Über die Eigenschaften des Wasserstoffmoleküls im Grundzustande. Z Phys. 1927;45:484–492.

[80] Mie G. Zur kinetischen Theorie der einatomigen Körper. Ann Phys. 1903;316:657–697.

[81] Jones JE. On the determination of molecular fields. I. From the variation of the viscosity of a gas with temperature. Proc R Soc Lond A. 1924;106:441–462.

[82] Slater JC. The normal state of helium. Phys Rev. 1928;32:349–360.

[83] Born M, Mayer JE. Zur Gittertheorie der Ionenkristalle. Z Phys. 1932;75:1–18.

[84] Buckingham RA. The repulsive interaction of atoms in S states. Trans Faraday Soc. 1958;54:453.

[85] O'Connor TC, Andzelm J, Robbins MO. AIREBO-m: A reactive model for hydrocarbons at extreme pressures. J Chem Phys. 2015;142:024903.

[86] Vleet MJV, Misquitta AJ, Stone AJ, et al. Beyond Born-Mayer: Improved models for short-range repulsion in ab initio force fields. J Chem Theor Comput. 2016;12:3851–3870.

[87] Cipcigan FS, Sokhan VP, Crain J, et al. Electronic coarse graining enhances the predictive power of molecular simulation allowing challenges in water physics to be addressed. J Comput Phys. 2016;326:222–233.

[88] Rittner ES. Binding energy and dipole moment of alkali halide molecules. J Chem Phys. 1951;19:1030–1035.

[89] Tang KT, Toennies JP. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. J Chem Phys. 1984;80:3726–3741.

[90] Cipcigan F, Crain J, Sokhan V, et al. Electronic coarse graining: Predictive atomistic modeling of condensed matter. Rev Mod Phys. 2019;91:025003.

[91] Madden PA, Wilson M. 'Covalent' effects in 'ionic' systems. Chem Soc Rev. 1996;25:339–350.

[92] Dalgarno A. Atomic polarizabilities and shielding factors. Adv Phys. 1962;11:281–315.

[93] Yukawa H. On the interaction of elementary particles. I Proc Phys Math Soc Jpn 3rd Series. 1935;17:48–57.

[94] Ziegler JF, Biersack JP, Littmark U. The stopping and range of ions in matter. New York: Pergamon; 1985.

[95] Nordlund K, Runeberg N, Sundholm D. Repulsive interatomic potentials calculated using Hartree-Fock and density-functional theory methods. Nucl Instrum Methods Phys Res B. 1997;132:45–54.

[96] Juslin N, Erhart P, Träskelin P, et al. Analytical interatomic potential for modeling nonequilibrium processes in the W–C–H system. J Appl Phys. 2005;98:123520.

[97] Bragg WL XVIII. The arrangement of atoms in crystals. Lond Edinb Dublin Philos Mag J Sci. 1920;40:169–189.

[98] Lorentz HA. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. Ann Phys. 1881;248:127–136.

[99] Berthelot D. Sur le mélange des gaz. Compt Rendus. 1898;126:1703–1855.

[100] Zeiss G, Meath WJ. Dispersion energy constants C_6 (A, B), dipole oscillator strength sums and refractivities for Li, N, O, H_2, N_2, O_2, NH_3, H_2O, NO and N_2O. Mol Phys. 1977;33:1155–1176.

[101] Tang KT, Toennies JP. The van der Waals potentials between all the rare gas atoms from He to Rn. J Chem Phys. 2003;118:4976–4983.

[102] Slater JC, Kirkwood JG. The van der Waals forces in gases. Phys Rev. 1931;37:682–697.

[103] Abrahamson AA. Born-Mayer-type interatomic potential for neutral ground-state atoms with $Z=2$ to $Z=105$. Phys Rev. 1969;178:76–79.

[104] Smith FT. Atomic distortion and the combining rule for repulsive potentials. Phys Rev A. 1972;5:1708–1713.

[105] Gilbert TL. Soft-sphere model for closed-shell atoms and ions. J Chem Phys. 1968;49:2640–2642.

[106] Böhm HJ, Ahlrichs R. A study of short-range repulsions. J Chem Phys. 1982;77:2028–2034.

[107] van Beest Bwh, Kramer GJ, van Santen RA, et al. Force fields for silicas and aluminophosphates based on ab initio calculations. Phys Rev Lett. 1990;64:1955–1958.

[108] Mendelev MI, Sordelet DJ, Kramer MJ. Using atomistic computer simulations to analyze x-ray diffraction data from metallic glasses. J Appl Phys. 2007;102:043501.

[109] Mendelev M, Kramer M, Ott R, et al. Development of suitable interatomic potentials for simulation of liquid and amorphous Cu–Zr alloys. Phil Mag. 2009;89:967–987.

[110] Cheng Y, Ma E. Atomic-level structure and structure–property relationship in metallic glasses. Progr Mater Sci. 2011;56:379–473.

[111] Shi Y, Falk ML. Atomic-scale simulations of strain localization in three-dimensional model amorphous solids. Phys Rev B. 2006;73. DOI:10.1103/PhysRevB.73.214201

[112] He Y, Yi P, Falk ML. Critical analysis of an FeP empirical potential employed to study the fracture of metallic glasses. Phys Rev Lett. 2019;122. DOI:10.1103/PhysRevLett.122.035501

[113] Wahnström G. Molecular-dynamics study of a supercooled two-component Lennard-Jones system. Phys Rev A. 1991;44:3752–3764.

[114] Kob W, Andersen HC. Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture I: The van Hove correlation function. Phys Rev E. 1995;51:4626–4641.

[115] Lerner E. Mechanical properties of simple computer glasses. J Non Cryst Solids. 2019;522:119570.

[116] Rainone C, Bouchbinder E, Lerner E. Pinching a glass reveals key properties of its soft spots. Proc Natl Acad Sci Unit States Am. 2020;117:5228–5234.

[117] Sukhomlinov SV, Müser MH. Anomalous system-size dependence of properties at the fragile-to-strong transition in a bulk-metallic-glass forming melt. Comput Mater Sci. 2019;156:129–134.

[118] Laird BB, Schober HR. Localized low-frequency vibrational modes in a simple model glass. Phys Rev Lett. 1991;66:636–639.

[119] Schober HR, Ruocco G. Size effects and quasilocalized vibrations. Phil Mag. 2004;84:1361–1372.

[120] Lerner E, Düring G, Bouchbinder E. Statistics and properties of low-frequency vibrational modes in structural glasses. Phys Rev Lett. 2016;117. DOI:10.1103/PhysRevLett.117.035501

[121] Mizuno H, Shiba H, Ikeda A. Continuum limit of the vibrational properties of amorphous solids. Proc Natl Acad Sci Unit States Am. 2017;114:E9767–E9774.

[122] Dzugutov M. Glass formation in a simple monatomic liquid with icosahedral inherent local order. Phys Rev A. 1992;46:R2984–R2987.

[123] Dzugutov M. Formation of a dodecagonal quasicrystalline phase in a simple monatomic liquid. Phys Rev Lett. 1993;70:2924–2927.

[124] Weber TA, Stillinger FH. Local order and structural transitions in amorphous metal-metalloid alloys. Phys Rev B. 1985;31:1954–1963.

[125] Ninarello A, Berthier L, Coslovich D. Models and algorithms for the next generation of glass transition studies. Phys Rev X. 2017;7.

[126] Vancoillie S, Malmqvist PÅ, Veryazov V. Potential energy surface of the chromium dimer re-re-revisited with multiconfigurational perturbation theory. J Chem Theor Comput. 2016;12:1647–1655.

[127] Deng B, Shi Y. A reactive coarse-grained model for polydisperse polymers. Polymer. 2016;98:88–99.

[128] Deng B, Palermo EF, Shi Y. Comparison of chain-growth polymerization in solution versus on surface using reactive coarse-grained simulations. Polymer. 2017;129:105–116.

[129] Birch F. Finite elastic strain of cubic crystals. Phys Rev. 1947;71:809–824.

[130] Birch F. Elasticity and constitution of the Earth's interior. J Geophys Res. 1952;57:227–286.

[131] Wallace DC. Lattice dynamics and elasticity of stressed crystals. Rev Mod Phys. 1965;37:57–67.

[132] Barron THK, Klein ML. Second-order elastic constants of a solid under stress. Proc Phys Soc. 1965;85:523–532.

[133] Wallace DC. Thermoelasticity of stressed materials and comparison of various elastic constants. Phys Rev. 1967;162:776–789.

[134] Wallace DC. Thermoelastic theory of stressed crystals and higher-order elastic constants. Solid state physics. Elsevier; 1970. p. 301–404.

[135] Wang J, Yip S, Phillpot SR, et al. Crystal instabilities at finite strain. Phys Rev Lett. 1993;71:4182–4185.

[136] Wang J, Li J, Yip S, et al. Mechanical instabilities of homogeneous crystals. Phys Rev B. 1995;52:12627–12635.

[137] Landau LD, Lifshitz EM. Theory of elasticity. Butterworth-Heinemann; 1986.

[138] Sukhomlinov SV, Müser MH. Constraints on phase stability, defect energies, and elastic constants of metals described by EAM-type potentials. J Phys Condens Matter. 2016;28:395701.

[139] Lemaître A, Maloney C. Sum rules for the quasi-static and visco-elastic response of disordered solids at zero temperature. J Stat Phys. 2006;123:415–453.

[140] Born M. On the stability of crystal lattices. I Math Proc Camb Phil Soc. 1940;36:160–172.

[141] Mouhat F, Coudert FX. Necessary and sufficient elastic stability conditions in various crystal systems. Phys Rev B. 2014;90. DOI:10.1103/PhysRevB.90.224104

[142] Drude P. Zur Elektronentheorie der Metalle. Ann Phys. 1900;306:566–613.

[143] Lorentz HA. The theory of electrons and its applications to the phenomena of light and radiant heat. Leipzig: Teubner; 1909.

[144] Axilrod BM, Teller E. Interaction of the van der Waals type between three atoms. J Chem Phys. 1943;11:299–300.

[145] Muto Y. Force between nonpolar molecules. J Phys Math Soc Jpn. 1943;17:629.

[146] von Lilienfeld OA, Tkatchenko A. Two- and three-body interatomic dispersion energy contributions to binding in molecules and solids. J Chem Phys. 2010;132:234109.

[147] Tang LY, Yan ZC, Shi TY, et al. The long-range non-additive three-body dispersion interactions for the rare gases, alkali, and alkaline-earth atoms. J Chem Phys. 2012;136:104104.

[148] Barker JA, Klein ML, Bobetic MV. Elastic constants and phonon dispersion curves for solid argon near 0 °K. Phys Rev B. 1970;2:4176–4179.

[149] Axilrod BM. Triple-dipole interaction. II. Cohesion in crystals of the rare gases. J Chem Phys. 1951;19:724–729.

[150] Jansen L. Quantum chemistry and crystal physics. In: Advances in quantum chemistry volume 2. Elsevier; 1966. p. 119–194.

[151] Lotrich VF, Szalewicz K. Three-body contribution to binding energy of solid argon and analysis of crystal structure. Phys Rev Lett. 1997;79:1301–1304.

[152] Rościszewski K, Paulus B, Fulde P, et al. *Ab initio* coupled-cluster calculations for the fcc and hcp structures of rare-gas solids. Phys Rev B. 2000;62:5482–5488.

[153] Ashcroft NW, Mermin ND. Solid State Physics. Holt-Saunders; 1976.

[154] Leven I, Head-Gordon T. C-GeM: Coarse-grained electron model for predicting the electrostatic potential in molecules. J Phys Chem Lett. 2019;10:6820–6826.

[155] Ranathunga DTS, Shamir A, Dai X, et al. Molecular dynamics simulations of water condensation on surfaces with tunable wettability. Langmuir. 2020;36:7383–7391.

[156] Reed SK, Lanning OJ, Madden PA. Electrochemical interface between an ionic liquid and a model metallic electrode. J Chem Phys. 2007;126:084704.

[157] Dapp WB, Müser MH. Redox reactions with empirical potentials: Atomistic battery discharge simulations. J Chem Phys. 2013;139:064106.

[158] Robbins MO, Kremer K, Grest GS. Phase diagram and dynamics of Yukawa systems. J Chem Phys. 1988;88:3286–3312.

[159] Arnold A, Breitsprecher K, Fahrenberger F, et al. Efficient algorithms for electrostatic interactions including dielectric contrasts. Entropy. 2013;15:4569–4588.

[160] Buckingham AD. Permanent and induced molecular moments and long-range intermolecular forces. In: Advances in chemical physics. John Wiley & Sons, Inc.; 1967. p. 107–142.

[161] Wilson M, Madden PA, Costa-Cabral BJ. Quadrupole polarization in simulations of ionic systems: application to AgCl. J Phys Chem. 1996;100:1227–1237.

[162] Nocedal J, Wright SJ. Numerical optimization. 2nd ed. Springer series in operations research. New York: Springer; 2006.

[163] Parrinello M, Rahman A. Crystal structure and pair potentials: A molecular-dynamics study. Phys Rev Lett. 1980;45:1196–1199.

[164] Car R, Parrinello M. Unified approach for molecular dynamics and density-functional theory. Phys Rev Lett. 1985;55:2471–2474.

[165] Tangney P, Scandolo S. How well do Car-Parrinello simulations reproduce the Born-Oppenheimer surface? theory and examples. J Chem Phys. 2002;116:14.

[166] Hummer G, Pratt LR, García AE. Molecular theories and simulation of ions and polar molecules in water. J Phys Chem A. 1998;102:7885–7895.

[167] Dick BG, Overhauser AW. Theory of the dielectric constants of alkali halide crystals. Phys Rev. 1958;112:90–103.

[168] de Leeuw Nh, Parker SC, de Leeuw NH. Molecular-dynamics simulation of MgO surfaces in liquid water using a shell-model potential for water. Phys Rev B. 1998;58:13901–13908.

[169] Lamoureux G, Roux B. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. J Chem Phys. 2003;119:3025–3039.

[170] Wilson M, Madden PA, Hemmati M, et al. Polarization effects, network dynamics, and the infrared spectrum of Amorphous SiO2. Phys Rev Lett. 1996;77:4023–4026.

[171] Tangney P, Scandolo S. An ab initio parametrized interatomic force field for silica. J Chem Phys. 2002;117:8898–8904.

[172] Herzbach D, Binder K, Müser MH. Comparison of model potentials for molecular-dynamics simulations of silica. J Chem Phys. 2005;123:124711.

[173] Demiralp E, Çağin T, Goddard WA. Morse stretch potential charge equilibrium force field for ceramics: Application to the quartz-stishovite phase transition and to silica glass. Phys Rev Lett. 1999;82:1708–1711.

[174] Rowley AJ, Jemmer P, Wilson M, et al. Evaluation of the many-body contributions to the interionic interactions in MgO. J Chem Phys. 1998;108:10209–10219.

[175] London F. The general theory of molecular forces. Trans Faraday Soc. 1937;33:8b.

[176] Whitfield TW, Martyna GJ. A unified formalism for many-body polarization and dispersion: The quantum Drude model applied to fluid xenon. Chem Phys Lett. 2006;424:409–413.

[177] Eisenschitz R, London F. über das Verhältnis der van der Waalsschen Kräfte zu den homöopolaren Bindungskräften. Z Phys. 1930;60:491–527.

[178] Tuckerman ME, Berne BJ, Martyna GJ, et al. Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals. J Chem Phys. 1993;99:2796–2808.

[179] Ceperley DM. Path integrals in the theory of condensed helium. Rev Mod Phys. 1995;67:279–355.

[180] Habershon S, Manolopoulos DE, Markland TE, et al. Ring-polymer molecular dynamics: Quantum effects in chemical dynamics from classical trajectories in an extended phase space. Annu Rev Phys Chem. 2013;64:387–413.

[181] Herrero CP, Ramírez R. Path-integral simulation of solids. J Phys Condens Matter. 2014;26:233201.

[182] Anderson JB. Quantum chemistry by random walk. $h^2 P$, h_3^+ D_{3h} $^1A_1$, h_2 $^3\sigma^+_u$, h_4 $^1\sigma^+_g$, be ^1S. J Chem Phys. 1976;65:4121–4127.

[183] Jones A, Thompson A, Crain J, et al. Norm-conserving diffusion Monte Carlo method and diagrammatic expansion of interacting Drude oscillators: Application to solid xenon. Phys Rev B. 2009;79. DOI:10.1103/PhysRevB.79.144119.

[184] Schröder U. A new model for lattice dynamics (breathing shell model). Solid State Commun. 1966;4:347–349.

[185] Basu AN, Sengupta S. A deformable shell model for the alkali halides. Phys Status Solidi (b). 1968;29:367–375.

[186] Matsui M. Breathing shell model in molecular dynamics simulation: Application to MgO and CaO. J Chem Phys. 1998;108:3304–3309.

[187] Tangney P, Scandolo S. A many-body interatomic potential for ionic systems: Application to MgO. J Chem Phys. 2003;119:9673–9685.

[188] Wilson M, Madden PA, Pyper NC, et al. Molecular dynamics simulations of compressible ions. J Chem Phys. 1996;104:8068–8081.

[189] Price SL, Stone AJ, Lucas J, et al. The nature of –Cl . . . Cl– intermolecular interactions. J Am Chem Soc. 1994;116:4910–4918.

[190] Sukhomlinov SV, Müser MH. Charge-transfer potentials for ionic crystals: Cauchy violation, LO-TO splitting, and the necessity of an ionic reference state. J Chem Phys. 2015;143:224101.

[191] Basu AN, Roy D, Sengupta S. Polarisable models for ionic crystals and the effective many-body interaction. Phys Status Solidi (a). 1974;23:11–32.

[192] Sangster M. Interionic potentials and force constant models for rocksalt structure crystals. J Phys Chem Solid. 1973;34:355–363.

[193] Hohenberg P, Kohn W. Inhomogeneous electron gas. Phys Rev. 1964;136:B864–B871.

[194] Kohn W, Sham LJ. Self-consistent equations including exchange and correlation effects. Phys Rev. 1965;140:A1133–A1138.

[195] Blöchl PE. Projector augmented-wave method. Phys Rev B. 1994;50:17953–17979.

[196] Wigner E. On the interaction of electrons in metals. Phys Rev. 1934;46:1002–1011.

[197] Ceperley DM, Alder BJ. Ground state of the electron gas by a stochastic method. Phys Rev Lett. 1980;45:566–569.

[198] Ascarelli P, Harrison RJ. Density-dependent potentials and the hard-sphere model for liquid metals. Phys Rev Lett. 1969;22:385–388.

[199] Vitek V. Pair potentials in atomistic computer simulations. MRS Bull. 1996;21:20–23.

[200] Degtyareva O. Crystal structure of simple metals at high pressures. High Pres Res. 2010;30:343–371.

[201] Cyrot-Lackmann F. On the calculation of surface tension in transition metals. Surf Sci. 1969;15:535–548.

[202] Friedel J. The physics of clean metal surfaces. Ann Phys. 1976;1:257–307.

[203] Pettifor DG. Bonding and structure of molecules and solids. Oxford University Press; 1995.

[204] Cleri F, Rosato V. Tight-binding potentials for transition metals and alloys. Phys Rev B. 1993;48:22–33.

[205] Tománek D, Mukherjee S, Bennemann KH. Simple theory for the electronic and atomic structure of small clusters. Phys Rev B. 1983;28:665–673.

[206] Tománek D, Bennemann K. Electronic model for energies, relaxations and reconstruction trends at metal surfaces. Surf Sci. 1985;163:503–515.

[207] Pagonabarraga I, Frenkel D. Dissipative particle dynamics for interacting systems. J Chem Phys. 2001;115:5015–5026.

[208] Warren PB. Vapor-liquid coexistence in many-body dissipative particle dynamics. Phys Rev E. 2003;68. DOI:10.1103/PhysRevE.68.066702

[209] Stott MJ, Zaremba E. Quasiatoms: An approach to atoms in nonuniform electronic systems. Phys Rev B. 1980;22:1564–1583.

[210] Nørskov JK, Lang ND. Effective-medium theory of chemical binding: Application to chemisorption. Phys Rev B. 1980;21:2131–2136.

[211] Nørskov JK. Covalent effects in the effective-medium theory of chemical binding: Hydrogen heats of solution in the 3d metals. Phys Rev B. 1982;26:2875–2885.

[212] Ercolessi F, Tosatti E, Parrinello M. Au (100) surface reconstruction. Phys Rev Lett. 1986;57:719–722.

[213] Ercolessi F, Parrinello M, Tosatti E. Au(100) reconstruction in the glue model. Surf Sci. 1986;177:314–328.

[214] Ercolessi F, Parrinello M, Tosatti E. Simulation of gold in the glue model. Phil Mag A. 1988;58:213–226.

[215] Mishin Y, Mehl MJ, Papaconstantopoulos DA, et al. Structural stability and lattice defects in copper: Ab initio, tight-binding, and embedded-atom calculations. Phys Rev B. 2001;63:224106.

[216] Mishin Y, Farkas D, Mehl MJ, et al. Interatomic potentials for monoatomic metals from experimental data and *ab initio* calculations. Phys Rev B. 1999;59:3393–3407.

[217] Mishin Y, Mehl MJ, Papaconstantopoulos DA. Embedded-atom potential for *B*2-NiAl. Phys Rev B. 2002;65. DOI:10.1103/PhysRevB.65.224114

[218] Williams PL, Mishin Y, Hamilton JC. An embedded-atom potential for the Cu–Ag system. Model Simulat Mater Sci Eng. 2006;14:817–833.

[219] Pun GPP, Mishin Y. Embedded-atom potential for hcp and fcc cobalt. Phys Rev B. 2012;86. DOI:10.1103/PhysRevB.86.134116

[220] Foiles SM, Baskes MI, Daw MS. Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys. Phys Rev B. 1986;33:7983–7991.

[221] Wen M, Whalen SM, Elliott RS, et al. Interpolation effects in tabulated interatomic potentials. Model Simulat Mater Sci Eng. 2015;23:074008.

[222] Jalkanen J, Müser MH. Systematic analysis and modification of embedded-atom potentials: case study of copper. Model Simulat Mater Sci Eng. 2015;23:074001.

[223] Baskes MI. Application of the embedded-atom method to covalent materials: A semiempirical potential for silicon. Phys Rev Lett. 1987;59:2666–2669.

[224] Baskes MI, Nelson JS, Wright AF. Semiempirical modified embedded-atom potentials for silicon and germanium. Phys Rev B. 1989;40:6085–6100.

[225] Baskes MI. Modified embedded-atom potentials for cubic materials and impurities. Phys Rev B. 1992;46:2727–2742.

[226] Baskes MI, Johnson RA. Modified embedded atom potentials for HCP metals. Model Simulat Mater Sci Eng. 1994;2:147–163.

[227] Stillinger FH, Weber TA. Computer simulation of local order in condensed phases of silicon. Phys Rev B. 1985;31:5262–5271.

[228] Keating PN. Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure. Phys Rev. 1966;145:637–645.

[229] Broughton JQ, Li XP. Phase diagram of silicon by molecular dynamics. Phys Rev B. 1987;35:9120–9127.

[230] Kluge MD, Ray JR. Velocity versus temperature relation for solidification and melting of silicon: A molecular-dynamics study. Phys Rev B. 1989;39:1738–1746.

[231] Waseda Y, Suzuki K. Structure of molten silicon and germanium by x-ray diffraction. Z Phys B. 1975;20:339–343.

[232] Molinero V, Sastry S, Angell CA. Tuning of tetrahedrality in a silicon potential yields a series of monatomic (metal-like) glass formers of very high fragility. Phys Rev Lett. 2006;97. DOI:10.1103/PhysRevLett.97.075701

[233] Dhabal D, Chakravarty C, Molinero V, et al. Comparison of liquid-state anomalies in Stillinger-Weber models of water, silicon, and germanium. J Chem Phys. 2016;145:214502.

[234] Hujo W, Jabes BS, Rana VK, et al. The rise and fall of anomalies in tetrahedral liquids. J Stat Phys. 2011;145:293–312.

[235] Biswas R, Hamann DR. Interatomic potentials for silicon structural energies. Phys Rev Lett. 1985;55:2001–2004.

[236] Holland D, Marder M. Ideal brittle fracture of silicon studied with molecular dynamics. Phys Rev Lett. 1998;80:746–749.

[237] Holland D, Marder M. Erratum: Ideal brittle fracture of silicon studied with molecular dynamics Phys. Rev. Lett. [80, 746 (1998)]. Phys Rev Lett. 1998;81:4029.

[238] Bazant MZ, Kaxiras E, Justo JF. Environment-dependent interatomic potential for bulk silicon. Phys Rev B. 1997;56:8542–8552.

[239] Justo JF, Bazant MZ, Kaxiras E, et al. Interatomic potential for silicon defects and disordered phases. Phys Rev B. 1998;58:2539–2550.

[240] Marks NA. Generalizing the environment-dependent interaction potential for carbon. Phys Rev B. 2000;63. DOI:10.1103/PhysRevB.63.035401

[241] Marks N. Modelling diamond-like carbon with the environment-dependent interaction potential. J Phys Condens Matter. 2002;14:2901–2927.

[242] Tersoff J. New empirical model for the structural properties of silicon. Phys Rev Lett. 1986;56:632–635.

[243] Tersoff J. New empirical approach for the structure and energy of covalent systems. Phys Rev B. 1988;37:6991–7000.

[244] Tersoff J. Empirical interatomic potential for carbon, with applications to amorphous carbon. Phys Rev Lett. 1988;61:2879–2882.

[245] Tersoff J. Empirical interatomic potential for silicon with improved elastic properties. Phys Rev B. 1988;38:9902–9905.

[246] Tersoff J. Modeling solid-state chemistry: Interatomic potentials for multicomponent systems. Phys Rev B. 1989;39:5566–5568.

[247] Pauling L. The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Mode. Third edition ed. Ithaca New York: Cornell University Press; 1960.

[248] Brenner DW. Relationship between the embedded-atom method and Tersoff potentials. Phys Rev Lett. 1989;63:1022.

[249] Rose JH, Smith JR, Ferrante J. Universal features of bonding in metals. Phys Rev B. 1983;28:1835–1845.

[250] Brenner DW. Erratum: Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. Phys Rev B. 1992;46:1948.

[251] Stuart SJ, Tutein AB, Harrison JA. A reactive potential for hydrocarbons with intermolecular interactions. J Chem Phys. 2000;112:6472–6486.

[252] Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Cambridge Mass:MIT Press;2006. Adaptive computation and machine learning.

[253] Cyrot-Lackmann F. Sur le calcul de la cohésion et de la tension superficielle des métaux de transition par une méthode de liaisons fortes. J Phys Chem Solid. 1968;29:1235–1243.

[254] Ducastelle F, Cyrot-Lackmann F. Moments developments and their application to the electronic charge distribution of d bands. J Phys Chem Solid. 1970;31:1295–1306.

[255] Haydock R, Heine V, Kelly MJ. Electronic structure based on the local atomic environment for tight-binding bands. J Phys C Solid State Phys. 1972;5:2845–2858.

[256] Haydock R, Heine V, Kelly MJ. Electronic structure based on the local atomic environment for tight-binding bands. II J Phys C Solid State Phys. 1975;8:2591–2605.

[257] Pettifor D. New many-body potential for the bond order. Phys Rev Lett. 1989;63:2480–2483.

[258] Horsfield AP, Bratkovsky AM, Fearn M, et al. Bond-order potentials: Theory and implementation. Phys Rev B. 1996;53:12694–12712.

[259] Jaynes ET. Information theory and statistical mechanics. Phys Rev. 1957;106:620–630.

[260] Jaynes ET. Information theory and statistical mechanics. II. Phys Rev. 1957;108:171–190.

[261] Mead LR, Papanicolaou N. Maximum entropy in the problem of moments. J Math Phys. 1984;25:2404–2417.

[262] Brown RH, Carlsson AE. Critical evaluation of low-order moment expansions for the bonding energy of lattices and defects. Phys Rev B. 1985;32:6125–6130.

[263] Friedel J. On the band structure of transition metals. J Phys F Met Phys. 1973;3:785–794.

[264] Ackland GJ, Finnis MW, Vitek V. Validity of the second moment tight-binding model. J Phys F Met Phys. 1988;18:L153–L157.

[265] Carlsson AE, Fedders PA, Myles CW. Generalized embedded-atom format for semiconductors. Phys Rev B. 1990;41:1247–1250.

[266] Carlsson AE. Angular forces in group-VI transition metals: Application to W(100). Phys Rev B. 1991;44:6590–6597.

[267] Hansen L, Stoltze P, Jacobsen KW, et al. Self-diffusion on copper surfaces. Phys Rev B. 1991;44:6523–6526.

[268] Foiles SM. Interatomic interactions for Mo and W based on the low-order moments of the density of states. Phys Rev B. 1993;48:4287–4298.

[269] Sutton AP, Finnis MW, Pettifor DG, et al. The tight-binding bond model. J Phys C Solid State Phys. 1988;21:35–66.

[270] Pettifor DG From exact to approximate theory: The tight binding bond model and many-body potentials. In: Many-atom interactions in solids. (Springer Proceedings in Physics; Vol. 48). Berlin, Heidelberg: Springer; 1990. p. 64–84.

[271] Pettifor DG, Aoki M, Murrell JN, et al. Bonding and structure of intermetallics: a new bond order potential. Philos Trans R Soc A. 1991;334:439–449.

[272] Nishitani SR, Alinaghian P, Hausleitner C, et al. Angularly dependent embedding potentials and structural prediction. Phil Mag Letters. 1994;69:177–184.

[273] Oleinik II, Pettifor DG, Oleinik II. Analytic bond-order potentials beyond Tersoff-Brenner. II. Application to the hydrocarbons. Phys Rev B. 1999;59:8500–8507.

[274] Mrovec M, Moseler M, Elsässer C, et al. Atomistic modeling of hydrocarbon systems using analytic bond-order potentials. Progr Mater Sci. 2007;52:230–254.

[275] Zhou XW, Ward DK, Foster ME. An analytical bond-order potential for carbon. J Comput Chem. 2015;36:1719–1735.

[276] Mrovec M, Nguyen-Manh D, Pettifor DG, et al. Bond-order potential for molybdenum: Application to dislocation behavior. Phys Rev B. 2004;69. DOI:10.1103/PhysRevB.69.094115.

[277] Murdick DA, Zhou XW, Wadley HNG, et al. Analytic bond-order potential for the gallium arsenide system. Phys Rev B. 2006;73. DOI:10.1103/PhysRevB.73.045206.

[278] Gillespie BA, Zhou XW, Murdick DA, et al. Bond-order potential for silicon. Phys Rev B. 2007;75. DOI:10.1103/PhysRevB.75.155207.

[279] Mrovec M, Gröger R, Bailey AG, et al. Bond-order potential for simulations of extended defects in tungsten. Phys Rev B. 2007;75. DOI:10.1103/PhysRevB.75.104119.

[280] Mrovec M, Nguyen-Manh D, Elsässer C, et al. Magnetic bond-order potential for iron. Phys Rev Lett. 2011;106. DOI:10.1103/PhysRevLett.106.246402.

[281] Ward DK, Zhou XW, Wong BM, et al. Analytical bond-order potential for the cadmium telluride binary system. Phys Rev B. 2012;85. DOI:10.1103/PhysRevB.85.115206.

[282] Zhou XW, Ward DK, Foster M, et al. An analytical bond-order potential for the copper–hydrogen binary system. J Mater Sci. 2015;50:2859–2875.

[283] Zhou X, Ward D, Foster M. An analytical bond-order potential for the aluminum copper binary system. J Alloy Comp. 2016;680:752–767.

[284] Zhou XW, Ward DK, Foster ME. A bond-order potential for the Al–Cu–H ternary system. New J Chem. 2018;42:5215–5228.

[285] Fennell CJ, Gezelter JD. Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics. J Chem Phys. 2006;124:234104.

[286] Cisneros GA, Karttunen M, Ren P, et al. Classical electrostatics for biomolecular simulations. Chem Rev. 2013;114:779–814.

[287] Anderson M, Swenson C. Experimental equations of state for the rare gas solids. J Phys Chem Solid. 1975;36:145–162.

[288] Ross M, Mao HK, Bell PM, et al. The equation of state of dense argon: A comparison of shock and static studies. J Chem Phys. 1986;85:1028–1033.

[289] Dewaele A, Loubeyre P, Mezouar M. Equations of state of six metals above 94 GPa. Phys Rev B. 2004;70:094112.

[290] Wang Y, Chen D, Zhang X. Calculated equation of state of Al, Cu, Ta, Mo, and W to 1000 GPa. Phys Rev Lett. 2000;84:3220–3223.

[291] Boehler R, Kennedy GC. Equation of state of sodium chloride up to 32 kbar and 500°C. J Phys Chem Solid. 1980;41:517–523.

[292] Dorogokupets PI, Dewaele A. Equations of state of MgO, Au, Pt, NaCl-b1, and NaCl-B2: Internally consistent high-temperature pressure scales. High Pres Res. 2007;27:431–446.

[293] Ferrante J, Smith JR, Rose JH. Diatomic molecules and metallic adhesion, cohesion, and chemisorption: A single binding-energy relation. Phys Rev Lett. 1983;50:1385–1386.

[294] Vinet P, Rose JH, Ferrante J, et al. Universal features of the equation of state of solids. J Phys Condens Matter. 1989;1:1941–1963.

[295] Occelli F, Loubeyre P, LeToullec R. Properties of diamond under hydrostatic pressures up to 140 GPa. Nat Mater. 2003;2:151–154.

[296] Albe K, Nordlund K, Averback RS. Modeling the metal-semiconductor interaction: Analytical bond-order potential for platinum-carbon. Phys Rev B. 2002;65. DOI:10.1103/PhysRevB.65.195124

[297] Albe K, Nordlund K, Nord J, et al. Modeling of compound semiconductors: Analytical bond-order potential for Ga, As, and GaAs. Phys Rev B. 2002;66. DOI:10.1103/PhysRevB.66.035205.

[298] Nord J, Albe K, Erhart P, et al. Modelling of compound semiconductors: analytical bond-order potential for gallium, nitrogen and gallium nitride. J Phys Condens Matter. 2003;15:5649–5662.

[299] Erhart P, Albe K. Analytical potential for atomistic simulations of silicon, carbon, and silicon carbide. Phys Rev B. 2005;71. DOI:10.1103/PhysRevB.71.035211

[300] Kittel C. Introduction to solid state physics. 8th ed. New York: Wiley; 2005.

[301] Schaefer HE. Investigation of thermal equilibrium vacancies in metals by positron annihilation. Phys Status Solidi (a). 1987;102:47–65.

[302] March NH. Displaced charge and formation energies of point defects in metals. J Phys F Met Phys. 1973;3:233–247.

[303] Górecki T. Comments on vacancies and melting. Scr Metall. 1977;11:1051–1053.

[304] Thierfelder C, Hermann A, Schwerdtfeger P, et al. Strongly bonded water monomers on the ice ih basal plane: Density-functional calculations. Phys Rev B. 2006;74. DOI:10.1103/PhysRevB.74.045422.

[305] Fujii Y, Lurie NA, Pynn R, et al. Inelastic neutron scattering from solid ^{36}Ar. Phys Rev B. 1974;10:3647–3659.

[306] [https://en.wikipedia.org/wiki/Ductility]; 2022. Online; accessed on April 1, 2022.

[307] Hirth JP, Lothe J. Theory of dislocations. 2nd ed. Malabar Florida: Krieger Publishing Company; 1982.

[308] Cai W, Nix WD. Imperfections in crystalline solids. Cambridge University Press; 2016.

[309] Mulliken RS. Electronic population analysis on LCAO-MO molecular wave functions. i. J Chem Phys. 1955;23:1833–1840.

[310] Löwdin PO. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. J Chem Phys. 1950;18:365–375.

[311] Bader RFW. A quantum theory of molecular structure and its applications. Chem Rev. 1991;91:893–928.

[312] Hirshfeld FL. Bonded-atom fragments for describing molecular charge densities. Theor Chim Acta. 1977;44:129–138.

[313] Meister J, Schwarz WHE. Principal components of ionicity. J Phys Chem. 1994;98:8245–8252.

[314] Mao JX. Atomic charges in molecules: A classical concept in modern computational chemistry. PostDoc J Rev. 2014;2.

[315] Choudhuri I, Truhlar DG. Calculating and characterizing the charge distributions in solids. J Chem Theor Comput. 2020;16:5884–5892.

[316] Maslen E, Spackman M. Atomic charges and electron density partitioning. Aust J Phys. 1985;38:273.

[317] Lillestolen TC, Wheatley RJ. Redefining the atom: atomic charge densities produced by an iterative stockholder approach. Chem Comm. 2008;;45:5909–5911.

[318] Verstraelen T, Ayers PW, Speybroeck VV, et al. Hirshfeld-e partitioning: AIM charges with an improved trade-off between robustness and accurate electrostatics. J Chem Theor Comput. 2013;9:2221–2225.

[319] Campañá C, Mussard B, Woo TK. Electrostatic potential derived atomic charges for periodic systems using a modified error functional. J Chem Theor Comput. 2009;5:2866–2878.

[320] Mortier WJ, Genechten KV, Gasteiger J. Electronegativity equalization: application and parametrization. J Am Chem Soc. 1985;107:829–835.

[321] Mortier WJ, Ghosh SK, Shankar S. Electronegativity-equalization method for the calculation of atomic charges in molecules. J Am Chem Soc. 1986;108:4315–4320.

[322] Parr RG, Pearson RG. Absolute hardness: companion parameter to absolute electronegativity. J Am Chem Soc. 1983;105:7512–7516.

[323] Mulliken RS. A new electroaffinity scale; together with data on valence states and on valence ionization potentials and electron affinities. J Chem Phys. 1934;2:782–793.

[324] Parr RG, Donnelly RA, Levy M, et al. Electronegativity: The density functional viewpoint. J Chem Phys. 1978;68:3801–3807.

[325] Rappe AK, Goddard WA. Charge equilibration for molecular dynamics simulations. J Phys Chem. 1991;95:3358–3363.

[326] Chen J, Martínez TJ. QTPIE: Charge transfer with polarization current equalization. a fluctuating charge model with correct asymptotics. Chem Phys Lett. 2007;438:315–320.

[327] Chelli R, Procacci P, Righini R, et al. Electrical response in chemical potential equalization schemes. J Chem Phys. 1999;111:8569–8575.

[328] Warren GL, Davis JE, Patel S. Origin and control of superlinear polarizability scaling in chemical potential equalization methods. J Chem Phys. 2008;128:144110.

[329] Nistor RA, Müser MH. Dielectric properties of solids in the regular and split-charge equilibration formalisms. Phys Rev B. 2009;79. DOI:10.1103/PhysRevB.79.104303

[330] Mikulski PT, Knippenberg MT, Harrison JA. Merging bond-order potentials with charge equilibration. J Chem Phys. 2009;131:241105.

[331] Perdew JP, Parr RG, Levy M, et al. Density-functional theory for fractional particle number: Derivative discontinuities of the energy. Phys Rev Lett. 1982;49:1691–1694.

[332] Scalfi L, Dufils T, Reeves KG, et al. A semiclassical thomas-fermi model to tune the metallicity of electrodes in molecular simulations. J Chem Phys. 2020;153:174704.

[333] Pomorski P, Pastewka L, Roland C, et al. Capacitance, induced charges, and bound states of biased carbon nanotube systems. Phys Rev B. 2004;69. DOI:10.1103/PhysRevB.69.115418.

[334] Pastewka L, Järvi TT, Mayrhofer L, et al. Charge-transfer model for carbonaceous electrodes in polar environments. Phys Rev B. 2011;83. DOI:10.1103/PhysRevB.83.165418.

[335] Rick SW, Stuart SJ, Berne BJ. Dynamical fluctuating charge force fields: Application to liquid water. J Chem Phys. 1994;101:6141–6156.

[336] Nistor RA, Polihronov JG, Müser MH, et al. A generalization of the charge equilibration method for nonmetallic materials. J Chem Phys. 2006;125:094108.

[337] Verstraelen T, Speybroeck VV, Waroquier M, et al. The electronegativity equalization method and the split charge equilibration applied to organic systems: Parametrization, validation, and comparison. J Chem Phys. 2009;131:044127.

[338] Verstraelen T, Ayers PW, Speybroeck VV, et al. ACKS2: Atom-condensed Kohn-Sham DFT approximated to second order. J Chem Phys. 2013;138:074108.

[339] Dapp WB, Müser MH. Towards time-dependent, non-equilibrium charge-transfer force fields. Eur Phys J B. 2013;86. DOI:10.1140/epjb/e2013-40047-x

[340] Hannay JH. The Clausius-Mossotti equation: an alternative derivation. Eur J Phys. 1983;4:141–143.

[341] Lacks DJ, Shinbrot T. Long-standing and unresolved issues in triboelectric charging. Nat Rev Chem. 2019;3:465–476.

[342] Sivaraman G, Gallington L, Krishnamoorthy AN, et al. Experimentally driven automated machine-learned interatomic potential for a refractory oxide. Phys Rev Lett. 2021;126. DOI:10.1103/PhysRevLett.126.156002.

[343] Aarva A, Deringer VL, Sainio S, et al. Understanding x-ray spectroscopy of carbonaceous materials by combining experiments, density functional theory, and machine learning. part II: Quantitative fitting of spectra. Chem Mater. 2019;31:9256–9267.

[344] Aarva A, Deringer VL, Sainio S, et al. Understanding x-ray spectroscopy of carbonaceous materials by combining experiments, density functional theory, and machine learning. Part I: Fingerprint spectra. Chem Mater. 2019;31:9243–9255.

[345] Aarva A, Sainio S, Deringer VL, et al. X-ray spectroscopy fingerprints of pristine and functionalized graphene. J Phys Chem C. 2021;125:18234–18246.

[346] Golze D, Hirvensalo M, Hernández-León P, et al. Accurate computational prediction of core-electron binding energies in carbon-based materials: A machine-learning model combining DFT and GW. Journal of chemical theory and computation. 2021 ArXiv:2112.06551;17:1662–1677. DOI:10.1021/acs.jctc.0c01282.

[347] Dodson BW. Development of a many-body Tersoff-type potential for silicon. Phys Rev B. 1987;35:2795–2798.

[348] Bartók AP, Kermode J, Bernstein N, et al. Machine learning a general-purpose interatomic potential for silicon. Phys Rev X. 2018;8.

[349] Bartók AP, Csányi G. Gaussian approximation potentials: A brief tutorial introduction. Int J Quant Chem. 2015;115:1051–1057.

[350] Musil F, Grisafi A, Bartók AP, et al. Physics-inspired structural representations for molecules and materials. Chem Rev. 2021;121:9759–9815.

[351] Himanen L, Jäger MO, Morooka EV, et al. DScribe: Library of descriptors for machine learning in materials science. Comput Phys Comm. 2020;247:106949.

[352] Bartók AP, Payne MC, Kondor R, et al. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. Phys Rev Lett. 2010;104. DOI:10.1103/PhysRevLett.104.136403.

[353] Deringer VL, Csányi G. Machine learning based interatomic potential for amorphous carbon. Phys Rev B. 2017;95. DOI:10.1103/PhysRevB.95.094203

[354] Rowe P, Deringer VL, Gasparotto P, et al. An accurate and transferable machine learning potential for carbon. J Chem Phys. 2020;153:034702.

[355] Behler J, Parrinello M. Generalized neural-network representation of high-dimensional potential-energy surfaces. Phys Rev Lett. 2007;98. DOI:10.1103/PhysRevLett.98.146401

[356] Behler J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. J Chem Phys. 2011;134:074106.

[357] Bartók AP, Kondor R, Csányi G. On representing chemical environments. Phys Rev B. 2013;87. DOI:10.1103/PhysRevB.87.184115

[358] Steinhardt PJ, Nelson DR, Ronchetti M. Bond-orientational order in liquids and glasses. Phys Rev B. 1983;28:784–805.

[359] Drautz R. Atomic cluster expansion for accurate and transferable interatomic potentials. Phys Rev B. 2019;99. DOI:10.1103/PhysRevB.99.014104

[360] Dusson G, Bachmayr M, Csányi G, et al. Atomic cluster expansion: Completeness, efficiency and stability. J Comput Phys. 2022;454:110946.

[361] Lysogorskiy Y, van der Oord C, Bochkarev A, et al. Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. Npj Comput Mater. 2021;7. DOI:10.1038/s41524-021-00559-9.

[362] Thompson A, Swiler L, Trott C, et al. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. J Comput Phys. 2015;285:316–330.

[363] Vandermause J, Torrisi SB, Batzner S, et al. On-the-fly active learning of interpretable bayesian force fields for atomistic rare events. Npj Comput Mater. 2020;6. DOI:10.1038/s41524-020-0283-z.

[364] Wen M, Tadmor EB. Uncertainty quantification in molecular simulations with dropout neural network potentials. Npj Comput Mater. 2020;6. DOI:10.1038/s41524-020-00390-8

[365] Csányi G, Albaret T, Payne MC, et al. "Learn on the Fly": A hybrid classical and quantum-mechanical molecular dynamics simulation. Phys Rev Lett. 2004;93. DOI:10.1103/PhysRevLett.93.175503.

[366] Li Z, Kermode JR, Vita AD. Molecular dynamics with on-the-fly machine learning of quantum-mechanical forces. Phys Rev Lett. 2015;114. DOI:10.1103/PhysRevLett.114.096405

[367] Kermode JR, Albaret T, Sherman D, et al. Low-speed fracture instabilities in a brittle crystal. Nature. 2008;455:1224–1227.

[368] Kermode J, Ben-Bashat L, Atrash F, et al. Macroscopic scattering of cracks initiated at single impurity atoms. Nat Comm. 2013;4. DOI:10.1038/ncomms3441.

[369] Lysogorskiy Y, van der Oord C, Bochkarev A, et al. Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. Npj Comput Mater. 2021;7.

[370] Anderson P. Chemical pseudopotentials. Phys Rep. 1984;110:311–319.

[371] Allen MP, Tildesley DJ. Computer simulation of liquids. Oxford University Press; 1989.

[372] in 't Veld, T Veld Pj AE, Grest GS, et al. Application of Ewald summations to long-range dispersion forces. J Chem Phys. 2007;127:144711.

[373] Isele-Holder RE, Mitchell W, Hammond JR, et al. Reconsidering dispersion potentials: Reduced cutoffs in mesh-based ewald solvers can be faster than truncation. J Chem Theo and Comp. 2013;9:5412–5420.

[374] Baskes MI, Angelo JE, Bisson CL. Atomistic calculations of composite interfaces. Model Simulat Mater Sci Eng. 1994;2:505–518.

[375] Mattoni A, Ippolito M, Colombo L. Atomistic modeling of brittleness in covalent materials. Phys Rev B. 2007;76. DOI:10.1103/PhysRevB.76.224103

[376] Pastewka L, Pou P, Pérez R, et al. Describing bond-breaking processes by reactive potentials: Importance of an environment-dependent interaction range. Phys Rev B. 2008;78. DOI:10.1103/PhysRevB.78.161402.

[377] Griebel M, Knapek S, Zumbusch G. Numerical simulation in molecular dynamics. Berlin Heidelberg: Springer; 2007.

[378] Wong-Ekkabut J, Miettinen MS, Dias C, et al. Static charges cannot drive a continuous flow of water molecules through a carbon nanotube. Nat Nanotechnol. 2010;5:555–557.

[379] Gong X, Li J, Lu H, et al. A charge-driven molecular water pump. Nat Nanotechnol. 2007;2:709–712.

[380] Plimpton S. Fast parallel algorithms for short-range molecular dynamics. J Comput Phys. 1995;117:1–19.

[381] Thompson AP, Aktulga HM, Berger R, et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. Comput Phys Comm. 2022;271:108171.

[382] Brugè F, Fornili S. Concurrent molecular dynamics simulation of spinodal phase transition on transputer arrays. Comput Phys Comm. 1990;60:31–38.

[383] Liem S, Brown D, Clarke J. Molecular dynamics simulations on distributed memory machines. Comput Phys Comm. 1991;67:261–267.

[384] Chynoweth S, Klomp U, Scales L. Simulation of organic liquids using pseudo-pairwise interatomic forces on a toroidal transputer array. Comput Phys Comm. 1991;62:297–306.

[385] Pinches MRS, Tildesley DJ, Smith W. Large scale molecular dynamics on parallel computers using the link-cell algorithm. Mol Simulat. 1991;6:51–87.

[386] Brown D, Clarke JH, Okuda M, et al. A domain decomposition parallelization strategy for molecular dynamics simulations on distributed memory machines. Comput Phys Comm. 1993;74:67–80.

[387] Rokhlin V. Rapid solution of integral equations of classical potential theory. J Comput Phys. 1985;60:187–207.

[388] Greengard L, Rokhlin V. A fast algorithm for particle simulations. J Comput Phys. 1987;73:325–348.

[389] Greengard L. Fast algorithms for classical physics. Science. 1994;265:909–914.

[390] Prandtl L. Ein Gedankenmodell zur kinetischen Theorie der festen Körper. Z Angew Math Mech. 1928;8:85–106.

[391] Smit B. Phase diagrams of Lennard-Jones fluids. J Chem Phys. 1992;96:8639–8640.

[392] Toxvaerd S, Dyre JC. Communication: Shifted forces in molecular dynamics. J Chem Phys. 2011;134:081102.

[393] Murty MVR, Atwater HA. Empirical interatomic potential for Si-H interactions. Phys Rev B. 1995;51:4889–4893.

[394] Müser MH. Improved cutoff functions for short-range potentials and the Wolf summation. Mol Simulat. 2022.

[395] Wolf D, Keblinski P, Phillpot SR, et al. Exact method for the simulation of coulombic systems by spherically truncated, pairwise r^{-1} summation. J Chem Phys. 1999;110:8254–8282.

[396] Ewald PP. Die Berechnung optischer und elektrostatischer Gitterpotentiale. Ann Phys. 1921;369:253–287.

[397] Rahbari A, Hens R, Jamali SH, et al. Effect of truncating electrostatic interactions on predicting thermodynamic properties of water–methanol systems. Mol Simul. 2018;45:336–350.

[398] Patra M, Karttunen M, Hyvönen M, et al. Molecular dynamics simulations of lipid bilayers: Major artifacts due to truncating electrostatic interactions. Biophys J. 2003;84:3636–3645.

[399] Vlugt TJH, Garca-Pérez E, Dubbeldam D, et al. Computing the heat of adsorption using molecular simulations: The effect of strong coulombic interactions. J Chem Theo and Comp. 2008;4:1107–1118.

[400] Shenderova OA, Brenner DW, Omeltchenko A, et al. Atomistic modeling of the fracture of polycrystalline diamond. Phys Rev B. 2000;61:3877–3888.

[401] Pastewka L, Klemenz A, Gumbsch P, et al. Screened empirical bond-order potentials for Si-C. Phys Rev B. 2013;87. DOI:10.1103/PhysRevB.87.205410.

[402] Kumagai T, Hara S, Choi J, et al. Development of empirical bond-order-type interatomic potential for amorphous carbon structures. J Appl Phys. 2009;105:064310.

[403] Khosrownejad SM, Kermode JR, Pastewka L. Quantitative prediction of the fracture toughness of amorphous carbon from atomic-scale simulations. Phys Rev Mater. 2021;5. DOI:10.1103/PhysRevMaterials.5.023602

[404] Perriot R, Gu X, Lin Y, et al. Screened environment-dependent reactive empirical bond-order potential for atomistic simulations of carbon materials. Phys Rev B. 2013;88. DOI:10.1103/PhysRevB.88.064101.

[405] Nguyen-Manh D, Pettifor DG, Vitek V. Analytic environment-dependent tight-binding bond integrals: Application to $MoSi_2$. Phys Rev Lett. 2000;85:4136–4139.

[406] Nguyen-Manh D, Pettifor DG, Cockayne DJH, et al. Environmentally dependent bond-order potentials: New developments and applications. Bull Mater Sci. 2003;26:43–51.

[407] Tang MS, Wang CZ, Chan CT, et al. Environment-dependent tight-binding potential model. Phys Rev B. 1996;53:979–982.

[408] Tang MS, Wang CZ, Chan CT, et al. Erratum: Environment-dependent tight-binding potential model. Phys Rev B. 1996;54:10982.

[409] Haas H, Wang CZ, Fähnle M, et al. Environment-dependent tight-binding model for molybdenum. Phys Rev B. 1998;57:1461–1470.

[410] Wang CZ, Pan BC, Ho KM. An environment-dependent tight-binding potential for Si. J Phys Condens Matter. 1999;11:2043–2049.

[411] Ercolessi F, Adams JB. Interatomic potentials from first-principles calculations: The force-matching method. Europhys Lett. 1994;26:583–588.

[412] Bally T, Sastry GN. Incorrect dissociation behavior of radical ions in density functional calculations. J Phys Chem A. 1997;101:7923–7925.

[413] Zhang Y, Yang W. A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons. J Chem Phys. 1998;109:2604–2608.

[414] Streitz FH, Mintmire JW. Electrostatic potentials for metal-oxide surfaces and interfaces. Phys Rev B. 1994;50:11996–12003.

[415] Yu J, Sinnott SB, Phillpot SR. Charge optimized many-body potential for the Si/SiO_2 system. Phys Rev B. 2007;75. DOI:10.1103/PhysRevB.75.085311

[416] Shan TR, Devine BD, Hawkins JM, et al. Second-generation charge-optimized many-body potential for Si/SiO_2 and amorphous silica. Phys Rev B. 2010;82.

[417] Martinez J, Liang T, Sinnott SB, et al. A third-generation charge optimized many body (COMB3) potential for nitrogen-containing organic molecules. Comput Mater Sci. 2017;139:153–161.

[418] Knippenberg MT, Mikulski PT, Ryan KE, et al. Bond-order potentials with split-charge equilibration: Application to C-, H-, and O-containing systems. J Chem Phys. 2012;136:164701.

[419] Martin RM. Elastic properties of ZnS structure semiconductors. Phys Rev B. 1970;1:4005–4011.

[420] Vashishta P, Kalia RK, Rino JP, et al. Interaction potential for SiO: A molecular-dynamics study of structural correlations. Phys Rev B. 1990;41:12197–12209.

[421] Senftle TP, Hong S, Islam MM, et al. The ReaxFF reactive force-field: development, applications and future directions. Npj Comput Mater. 2016;2:094106. DOI:10.1038/npjcompumats.2015.11.

[422] Bhattarai H, Newman KE, Gezelter JD. Polarizable potentials for metals: The density readjusting embedded atom method (DR-EAM). Phys Rev B. 2019;99. DOI:10.1103/PhysRevB.99.094106

[423] Darby JP, Kermode JR, Csányi G, Compressing local atomic neighbourhood descriptors. Npj Comput Mater. 2022;8:166. DOI:10.1038/s41524-022-00847-y.

[424] Goedecker S. Linear scaling electronic structure methods. Rev Mod Phys. 1999;71:1085–1123.

[425] Gaspard JP, Pellegatti A, Marinelli F, et al. Peierls instabilities in covalent structures I. Electronic structure, cohesion and the $Z=8N$ rule. Phil Mag B. 1998;77:727–744.

[426] Smith PPK, Buseck PR. Carbyne forms of carbon: Do they exist?. Science. 1982;216:984–986.

[427] Heimann RB, Kleiman J, Salansky NM. A unified structural approach to linear carbon polytypes. Nature. 1983;306:164–167.

[428] Whittaker AG. Carbyne forms of carbon: Evidence for their existence. Science. 1985;229:485–486.

[429] Moras G, Pastewka L, Walter M, et al. Progressive shortening of sp-hybridized carbon chains through oxygen-induced cleavage. J Phys Chem C. 2011;115:24653–24661.

[430] Moras G, Pastewka L, Gumbsch P, et al. Formation and oxidation of linear carbon chains and their role in the wear of carbon materials. Tribol Lett. 2011;44:355–365.

[431] Enkovaara J, Rostgaard C, Mortensen JJ, et al. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. J Phys Condens Matter. 2010;22:253202.