

Forcefields for materials simulations

Contents

- [Introduction](#)
- [Dispersion and Repulsive Energy](#)
- [Coulomb Potential](#)
- [Valence Forcefields](#)
- [Non-valence Forcefields \(Interatomic Potentials\)](#)
- [Variable Charge Reactive Potentials](#)
- [Machine Learning Potentials](#)

1 Introduction

The classical methodology that describes inter- and/or intra-molecular interactions is called an interatomic potential in the convention of materials scientists and solid-state physicists, or a forcefield (or *force field*) in the convention of chemical and biological scientists. Empirical or classic potentials are a set of mathematical functional forms that describe inter-atomic interactions.

For classical forcefield-based simulations (molecular dynamics and Monte Carlo methods), the total energy E in the previous section is the sum of the potential energy U and the kinetic energy K as defined by the equation:

$$E = U + K \quad (1)$$

Molecular dynamics simulations utilize the above energy function to describe inter-atomic interactions. Forces are derived from the potential energy:

$$F_i = -\frac{\partial U}{\partial r_i} \quad (2)$$

to integrate Newton's equation of motion:

$$F = ma \quad (3)$$

Monte Carlo simulation, on the other hand, does not need to consider forces nor kinetic energy, since the influence of the kinetic energy is accounted for by the classical factorization of the partition function (see chapter III.E. Monte Carlo methods). However, the potential energy of many configurations needs to be evaluated in a reasonable computer time to build a statistically representative ensemble. This is the role assigned to the forcefield (also referred to as **molecular potential** or **interatomic potential**). The forcefield must predict the potential energy of the system as a function of atomic coordinates without having to solve the Schrödinger equation.

The following sections describe some of the forcefields (or interatomic potentials) provided in the *MedeA* Environment.

2 Dispersion and Repulsive Energy

Dispersion energy is often the main source of intermolecular attractive forces, which accounts for the cohesion of liquids. It is also often a significant part of the adsorption energy in microporous solids. Repulsion prevents molecules from overlapping, as they would do in liquids in the presence of attractive forces alone.

2.1 Origin of Dispersion and Repulsion Energy

A complete treatment of dispersion energy includes interactions between fluctuating dipoles, fluctuating quadrupoles, and higher moments of the electronic charge distribution. Also, dispersion energy not only includes pairwise interactions (i.e. involving two nuclei) but also three-body interactions, i.e. between three nuclei. In these contributions, the leading term (which corresponds to the interaction between fluctuating dipoles) decreases with the sixth power of the distance separating the related nuclei:

$$U_{disp} = -\frac{3}{2} \left(\frac{E_i E_j}{E_i + E_j} \frac{\alpha_i \alpha_j}{4\pi\epsilon_0} \frac{1}{r_{ij}^6} \right) \quad (4)$$

where α_i and α_j are the polarisabilities of the atoms, E_i and E_j the energies of first electronic transition in both atoms and r_{ij} the separation distance.

When fluctuating quadrupoles and higher moments are included, higher-order terms add to the expression although they are often neglected. Three-body contributions amount to 5-10% of the interaction in liquids [4], but they are generally not explicitly taken into account. These neglected terms are then implicitly taken into consideration through the empirical calibration of effective dispersion parameters, using the general dependence of the equation with separation distance. Determining dispersion parameters directly via quantum mechanical calculations is challenging because this component of the interaction energy is difficult to obtain with the desired accuracy.

The repulsive energy results from the overlap between electronic charge distributions and the Pauli exclusion principle. Numerous empirical potentials have been proposed for repulsion, such as exponential $\exp(-br_{ij})$ or power laws r_{ij}^{-n} with $n \geq 9$.

2.2 Lennard-Jones (LJ) Potential

The most well-established expression of dispersion-repulsion energy between two atoms i and j (or more generally, between two force centers) is the Lennard-Jones potential:

$$U_{LJ} = U_{rep} + U_{disp} = 4\epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right) \quad (5)$$

where σ_{ij} is the separation distance for which attractive and repulsive terms are exactly opposite making the Lennard-Jones interaction zero, and ϵ_{ij} is the depth of the minimum interaction energy which corresponds to a separation distance $r_{min} = 2^{1/6}\sigma_{ij} \approx 1.125\sigma_{ij}$. The 12th exponent of the repulsion, which was chosen arbitrarily to minimize computational expenses, appears to account fairly well for the behavior of simple fluids (Ar, Kr, ...) and molecular non-polar fluids with approximate spherical symmetry (methane, isopentane, etc.). This is the functional form selected in the forcefields OPLS-AA [5], UA-TraPPE [6], AUA [7] for instance.

[4] B M Axilrod and E Teller, "Interaction of the Van Der Waals Type Between Three Atoms," *Journal of Chemical Physics* 11, no. 6 (1943): 299.

[5] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," *Journal of the American Chemical Society* 118, no. 45 (January 1996): 11225-11236.

[6] MG Martin and IJ Siepmann, "Transferable Models for Phase Equilibria 1. United-Atom Description of N-Alkanes," *Journal of Physical Chemistry B* 102 (1998): 2569; MG Martin and IJ Siepmann, "Novel Configurational-Bias Monte Carlo Method for Branched Molecules. Transferable potential for Phase Equilibria. 2. United-Atoms Description of Branched Alkanes," *Journal of Physical Chemistry B* 103 (1999): 4508; Katie A Maerzke, Nathan E Schultz, Richard B Ross, and J Ilja Siepmann, "TraPPE-UA Force Field for Acrylates and Monte Carlo Simulations for Their Mixtures with Alkanes and Alcohols," *Journal of Physical Chemistry B* 113, no. 18 (May 7, 2009): 6415-6425.

[7] Nicolas Ferrando, Véronique Lachet, Jean-Marie Teuler, and Anne Boutin, "Transferable Force Field for Alcohols and Polyalcohols," *Journal of Physical Chemistry B* 113, no. 17 (April 30, 2009): 5985-5995.

The Compass [8] and PCFF+ forcefields make use of a slightly different form of the Lennard-Jones potential, with a $\frac{1}{r^9}$ dependence for the repulsion energy. This functional form provides an improved representation of the repulsive interaction at somewhat increased computational effort because several additional multiplications are required to evaluate the '9/6' form. (The 12th power of the separation being simply the square of the 6th power.)

2.3 Combining Rules

The cross coefficients σ_{ij} and ϵ_{ij} depend on the pair of atoms considered. Although they may be calibrated separately in some instances, they are generally derived from parameters σ_i and ϵ_i through the application of combining rules (not be confused with the mixing rules used in classical thermodynamic models as they apply between atoms, while mixing rules apply between molecules) such as the Lorentz-Berthelot rules used in UA-TraPPE and AUA forcefields:

$$\begin{aligned}\sigma_{ij} &= \frac{\sigma_{ii} + \sigma_{jj}}{2} \\ \epsilon_{ij} &= \sqrt{\epsilon_{ii} \cdot \epsilon_{jj}}\end{aligned}\quad (6)$$

The rationale behind this rule being that it represents the diameter of dissimilar pairs in the limiting case of non-interpenetrating hard spheres. A number of forcefields use a geometric mean expression for Lennard-Jones interaction parameters. This is the case for PCC+ and OPLS-AA, for example. The rationale behind such expressions for cross energetic parameters is the fact that short range interactions are decidedly non-linear as a function of separation.

$$\epsilon_{ij}\sigma_{ij}^6 = \sqrt{\epsilon_i\sigma_i^6\epsilon_j\sigma_j^6}\quad (7)$$

This equation has been used in the combining rules of Kong [11] and Waldman & Hagler [12] and is also employed in *MedeA*. Compared with the ϵ obtained from the Lorentz-Berthelot combining rules, it results in lower ϵ for atoms with a substantial size difference but does not change significantly for atoms of equivalent size. Alternative geometric combining rules for σ have been selected by Kong [11]:

$$\epsilon_{ij}\sigma_{ij}^{12} = \frac{\epsilon_i\sigma_i^{12}}{2^{13}} \left[1 + \left(\frac{\epsilon_j\sigma_j^{12}}{\epsilon_i\sigma_i^{12}} \right)^{1/13} \right]^{13}\quad (8)$$

and by Waldman & Hagler [12]:

$$\epsilon_{ij} = \left(\frac{\epsilon_i^6 + \epsilon_j^6}{2} \right)^{1/6}\quad (9)$$

both resulting in diameters close to equation (6). Compass and PCFF make use of the Waldman & Hagler [12] combining rules, Eq. (7) and Eq. (9).

Combining rules are an important component of generic forcefields such as UA-TraPPE, AUA, OPLS-AA, PCFF+, and the *MedeA GIBBS* and *MedeA LAMMPS* implementations of these methods.

3 Coulomb Potential

Coulomb potential is an essential part of many forcefields that describes Coulomb interactions between particles carrying charges. On the basis of such discrete charge models, the electrostatic potential energy is computed using Coulomb's law:

[8] H Sun, "COMPASS: an Ab Initio Force-Field Optimized for Condensed-Phase -Overview with Details on Alkane and Benzene Compounds," *Journal of Physical Chemistry B* 102, no. 38 (September 1998): 7338-7364.

[11] Chang Lyoul Kong, "Combining Rules for Intermolecular Potential Parameters. II. Rules for the Lennard-Jones (12-6) Potential and the Morse Potential," *Journal of Chemical Physics* 59, no. 5 (1973): 2464.

[12] Marvin Waldman and A T Hagler, "New Combining Rules for Rare Gas Van Der Waals Parameters," *Journal of Computational Chemistry* 14, no. 9 (September 1993): 1077-1084.

$$U_{el} = \frac{1}{4\pi\epsilon_0} \sum_{i,j;i < j} \frac{q_i q_j}{r_{ij}} \quad (10)$$

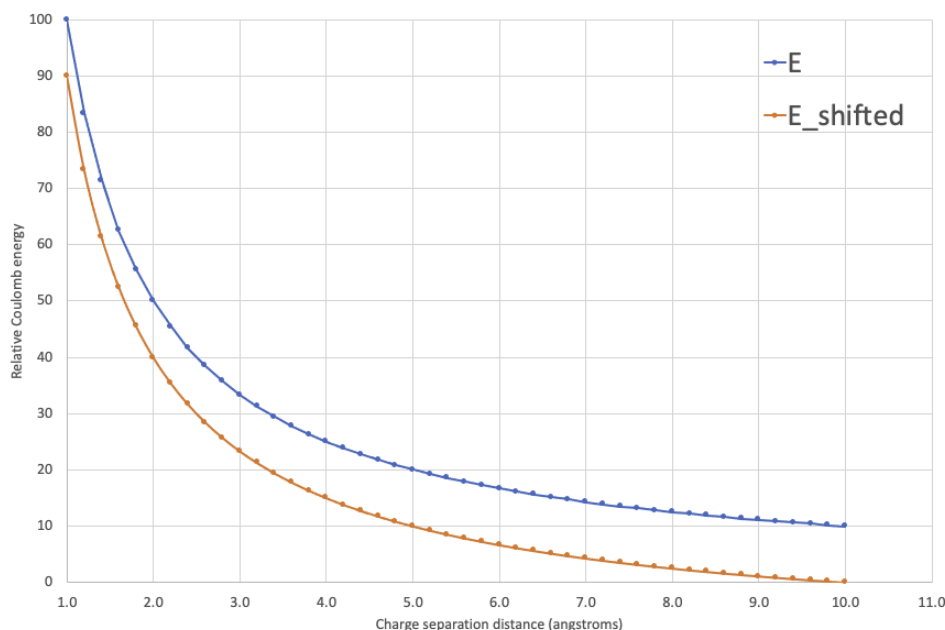
where the sum includes all possible pairs i, j of partial charges, r_{ij} is the separation distance and $\epsilon_0 = 8.85419 \cdot 10^{12} C^2 N^{-1} m^{-2}$.

Within different forcefields, charge can be represented by full valence charges (as in many Buckingham potential descriptions) or by partial charges (where less than complete electron transfer is considered) for organic valence forcefields. Variable charge forcefields determine atomic charge during the course of the simulation based on the electrostatic environment encountered by a given atom.

For valence forcefields, the location of these partial charges coincides with atomic nuclei. This is the case of the All Atom forcefields implemented in *MedeA* Forcefields: OPLS-AA [1]. Additional charges may be also placed at other sites to provide an improved representation of electron pairs, for example. This is particularly the case for small molecules like water, N_2 , O_2 , etc. [2]. The magnitude and location of these charges may be determined based on quantum mechanical calculations, and may also be obtained by reference to empirical properties such as the dipole and quadrupole moments, radial or angular distribution functions, cohesive energy, and similar properties. The quantum mechanical calculation may be made in a dielectric medium, to be more representative of the pure liquid when its dielectric constant is known.

3.1 Treating the slow-decaying $1/r$ term

Coulomb potential decays with $1/r$ which is still significant (~5-15%) when two charges are separated by 10 Å. Simply cutting the interactions off after the long-range cutoff (usually 9-12 Å) can lead to discontinuous energy and force descriptions. For example, the blue curve below has an abrupt and sudden change from 10.0 to 0.0 at the cutoff distance of 10.0 Å.



Therefore, one approach when employing the cutoff method is to shift the curve with the addition of a constant so that the interaction at the cutoff distance is zero, shown in the orange curve above.

Alternatively, in simulations with periodic boundary conditions, a convergence acceleration method, such as the Ewald method may be employed. The Ewald summation method divides the Coulomb potential into two parts: a short-range, real-space contribution within a finite cutoff, and a long-range, Fourier-space

[1] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," *Journal of the American Chemical Society* 118, no. 45 (January 1996): 11225-11236.

[2] William L Jorgensen, J Chandrasekhar, and J D Madura, "Comparison of Simple Potential Functions for Simulating Liquid Water," *Journal of Chemical Physics* 79 (1983): 926-935.

contribution. This combination results high accuracy at reasonable computational effort when computing long-range, Coulombic interactions for periodic systems.

An elaboration of the Ewald summation method is the particle-particle-particle-mesh (PPPM or P3M) summation method where particle positions are interpolated onto a mesh, and the potential is solved for this mesh, furthering improving the calculation speed.

4 Valence Forcefields

Valence forcefields have pre-defined bonds, angles, and torsions. Valence forcefields in MedeA include:

- pcff/pcff+
- compass/compass+
- opls-aa/opls-aa+
- AUA/AUA+
- gaff
- trape+
- clayff
- cvff_aug

4.1 Standard Decomposition of the Potential Energy

For valence forcefields, the potential energy of a group of molecules is classically subdivided according to two contributions:

$$U_{tot} = U_{inter} + U_{intra} \quad (11)$$

where U_{inter} is the intermolecular energy, also termed the *external energy*, i.e. the energy arising from the interaction between distinct molecules, and U_{intra} is the intramolecular energy, also termed the *internal energy*, which results from the interactions between the atoms belonging to the same molecule.

Classically, the intermolecular energy is divided into a sum of four terms, as a result of a perturbation expansion of electronic charge distribution in the interacting molecules:

$$U_{inter} = U_{el} + U_{pol} + U_{disp} + U_{rep} \quad (12)$$

where U_{el} is the electrostatic potential energy, which originates from the Coulomb forces between the permanent components of the electronic charge distribution. U_{pol} is the polarization energy, always attractive, which originates from the Coulombic interaction between the charge distribution induced in a molecule by the electric field created by the permanent charge distribution of the surrounding molecules and the permanent charge distribution of the surrounding molecules; U_{disp} is the dispersion energy, always attractive, which is the Coulombic interaction between the fluctuating components of the charge distribution of the molecules. Lastly, U_{rep} is the repulsive energy, which prevents the molecules from overlapping significantly, as a result of the impossibility of electrons to occupy the same state.

The intramolecular U_{intra} energy is expressed as the sum of the following main contributions:

$$U_{intra} = U_{str} + U_{bend} + U_{tors} + U_{improper} + U_{nb} \quad (13)$$

where U_{str} is the stretching energy, associated with the variations of bond length, U_{bend} is the bending energy, arising from the variations of the angle formed by two successive chemical bonds U_{tors} is the torsion energy caused by the variations of the dihedral angles formed by four successive atoms in a chain, $U_{improper}$ is the improper torsion used to describe inversion barriers or out-of-plane deformation energies in planar molecules, and U_{nb} is the non-bond energy resulting from the interaction between atoms (or electrostatic charges/dipoles) separated by more than three chemical bonds.

4.2 Units for Energy

The energy per molecule or bond for valence forcefields is often expressed per mole of substance, i.e. kJ/mol . An alternative approach is to express energy in terms of the Boltzmann constant. For instance, the energy attraction parameter ϵ of the Lennard-Jones potential is often given as ϵ/k_B , which has the dimension of temperature (K).

4.3 Inter-Molecular Energies

Dipole Moment and Quadrupole Moments

For a set of electrostatic charges q_i with coordinates $(r_{\alpha i})$, the dipole moment vector is defined as the first-order moment of the charge distribution [3]:

$$\mu_{\alpha} = \sum_i q_i r_{\alpha i} \quad (14)$$

The dipole moment is the magnitude of this vector. For instance, in the case of a molecule bearing opposite charges q and $-q$ separated by a distance d it is:

$$|\mu| = qd \quad (15)$$

The classical unit for dipole moment is the Debye ($1 \text{ D} = 3.334 \cdot 10^{-30} \text{ C.m}$). The components of the quadrupole tensor are defined as:

$$Q_{\alpha\beta} = \frac{1}{2} \sum_i q_i (3r_{\alpha i} r_{\beta i} - r^2 \delta_{\alpha\beta}) \quad (16)$$

where α and β may be either of the axis x , y , or z , and $\delta_{\alpha\beta}$ is the Kronecker symbol. The quadrupolar moment is:

$$|Q|^2 = Q_{zz}^2 + \frac{4}{3}(Q_{xy}^2 + Q_{xz}^2 + Q_{yz}^2) + \frac{1}{3}(Q_{xx} - Q_{yy})^2 \quad (17)$$

In the case of a system of three aligned charges $(-q, 2q, -q)$ separated by a distance d the main component of the quadrupole tensor is:

$$Q_{zz} = -\frac{1}{2}qd^2 \quad (18)$$

For instance, CO_2 electrostatics may be represented by three partial charges as shown below, where $Q = -14.3 \cdot 10^{-40} \text{ Cm}^2$.

Polarizability Energy

If a molecule is placed within an electric field \mathbf{E} created by other molecules in the system, its internal charge distribution will change so that an additional-or induced-dipole moment is created:

$$\mu_P = \alpha_P E \quad (19)$$

Where α_P is the polarizability of the molecule, and \mathbf{E} the electric field. As a general rule, polarizability increases with the diameter of the electron cloud, i.e. with molecular size. Within a first-order approximation, the electric field may be obtained by deriving:

$$E = \sum_j \frac{1}{4\pi\epsilon_0} \frac{q_j}{r_j^2} \quad (20)$$

where the summation covers all the system charges apart from those of the molecule being considered, r_i being the distance between the polarized molecule and charge q_i .

[3] JL Rivail, *Éléments De Chimie Quantique à L'usage Des Chimistes, Savoirs actuels.* (EDP Sciences, 2000). William C Swope, Hans W Horn, and Julia E Rice, "Accounting for Polarization Cost When Using Fixed Charge Force Fields. I. Method for Computing Energy," *Journal of Physical Chemistry B* 114, no. 26 (July 8, 2010): 8621-8630.

The interaction energy of the induced dipole moment with the electrostatic field provides the polarization energy:

$$U_{pol} = -\frac{1}{2}\alpha_P E^2 \quad (21)$$

Potential energy calculations may be time-consuming when polarizability effects are significant because the electric field determination must account of the influence of induced dipoles (the correction is often called back polarizability).

As a general rule, the polarization energy is significant only when polar molecules or solids are present in the system because otherwise, the electric field is not strong enough. In this case, non-polar molecules may also contribute to polarization energy because their polarizability is significant. For instance, the polarizability of methane (2.59 Å³), which has no dipole moment, is greater than the polarizability of water (1.45 Å³), which has a large dipole moment.

The polarizability of many small molecules can be simply taken from experimental measurements. For complex molecules, polarizability may be often determined from group contributions.

4.4 Intra-Molecular Energies

Internal conformation changes cannot be neglected in flexible molecules like alkanes or polymers. To consider such changes in molecular simulation, we need to consider the internal potential energy of the molecule. Conceptually, the simplest way to achieve this would be to sum the values for dispersion, repulsive and Coulombic energy between the force centers belonging to the same molecule. However, this procedure inadequately reproduces major features like average bond angles and equilibrium molecular conformations. Hence, the classical dispersion and repulsion potentials between neighboring atoms are replaced by more appropriate terms, namely stretching, bending and torsion interactions. These successive terms correspond to increases in the number of atoms involved, i.e. two in the case of stretching, three in bending and four in torsion.

Stretching Energy

The stretching energy U_{str} is the potential energy associated with the variation of bond length l around its mean value l_0 . In general it is described by a harmonic potential:

$$U_{str} = \frac{1}{2}k_{str}(l - l_0)^2 \quad (22)$$

where k_{str} is the stiffness of the bond. At normal temperatures (say up to 700K), this term is neglected in some forcefields like UA-TraPPE and AUA, because the related vibrations are of small amplitude and simulations are sufficiently representative if bond lengths are set to mean values l_0 . A consequence of this approximation is that the contribution of bond stretching to internal energy and heat capacity is neglected.

Bending Energy

The bending energy U_{bend} is the potential energy associated with the angle θ between two successive bonds. As a consequence, it involves three atoms. It is generally treated using a harmonic potential of the same type as bond stretching:

$$U_{bend} = \frac{1}{2}k_{bend}(\theta - \theta_0)^2 \quad (23)$$

where θ_0 is the angle of minimum potential energy, k_{bend} a spring constant. Both parameters are obtained from molecular structure and infrared spectroscopy. In some cases, the variations of θ are sufficiently small that they do not significantly influence simulation results. Under such circumstances, it may be appropriate to impose constant bond angles and to neglect the associated potential energy. This is, for instance, the case of CO₂, which may be assumed to be linear (in this example $\theta = 180^\circ$).

An alternative formula for bending energy is that from Toxvaerd [13]:

[13] Soren Toxvaerd, "Equation of State for Alkanes II," Journal of Chemical Physics 107 (1997): 5197.

$$U_{bend} = \frac{1}{2}k'_{bend} (\cos\theta - \cos\theta_0)^2 \quad (24)$$

This expression may save computer time without introducing major changes compared to equation (23). However, care must be taken that first-order equivalence requires a different spring constant than equation (23).

Torsion Energy

Torsion energy is related to the dihedral angle ϕ defined from the coordinates of four successive atoms.

$$U_{tors} = \sum_{n=0}^p a_n \cos n\phi \quad (25)$$

This series comprises three or four terms so that the torsion potential exhibits several minima between 0 and 360 ° [14]. A formally equivalent form of equation (25), which is used in the AUA potential for instance, is:

$$U_{tors} = \sum_{n=0}^p a_n (\cos n\chi)^n \quad (26)$$

where the torsion angle χ is defined differently from the dihedral angle ϕ , differing by 180 °. In the example of linear alkanes, the minimal torsion energy occurs at a torsion angle ϕ of 180 ° (i.e. a *trans* conformation) while two secondary minima occur at 60 ° and -60 ° (i.e. *gauche* conformations). The *trans* configurations are thus favored over the *gauche* conformations.

Distant Neighbor Internal Energy

The distant neighbor (or non-bonded) intramolecular interaction U_{dn} corresponds to the usual pair interactions between atoms that are not interacting through either stretching, bending or torsion, i.e. interactions between atoms separated by more than three bonds.

Distant neighbor energy may include dispersion, repulsion, electrostatic and polarizability components, exactly in the same way, as is the case between separated molecules. For minimally polarized molecules like the alkanes, the distant neighbor interaction is generally reduced to the Lennard-Jones energy. In multifunctional polar molecules, internal electrostatic forces may also be considered. Alternatively, several forcefields make use of an empirical factor. [15]

4.5 All Atom, United Atom and Anisotropic United Atom Forcefields

When representing polyatomic molecules, two approaches may be used to represent the dispersion and repulsive forces:

- Assignment of a separate force center to each atom located on its nucleus. Such methods are referred to as “All Atom” or AA.
- Assignment of a force center for a group of atoms, with groups such as:
 - CH, CH₂ or CH₃. This technique, referred to as “United Atom”, can be divided into classical United Atom (UA) and Anisotropic United Atom (AUA) depending on the position of the force centers.
 - Multiple heavy atoms such as a phenyl ring. This approach is called “Mesoscale Simulation”.

[14] William L Jorgensen, JD Madura and Carol J Swenson, “Optimized Intermolecular Potential Functions for Liquid Hydrocarbons,” *Journal of the American Chemical Society* 106 (1984): 6638.

[15] Nicolas Ferrando, Véronique Lachet, Jean-Marie Teuler, and Anne Boutin, “Transferable Force Field for Alcohols and Polyalcohols,” *Journal of Physical Chemistry B* 113, no. 17 (April 30, 2009): 5985-5995.

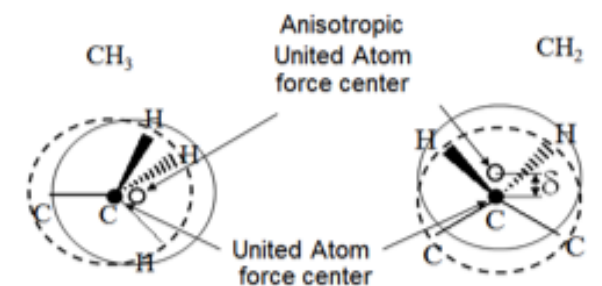
All Atom

The advantage of All Atom models is that they give a good account of molecular geometry and structure. The disadvantage is that they require greater computer time because of the large number of force centers, bending angles and torsion angles involved.

United Atom

United Atom methods generally neglect the hydrogen atoms other than those involved in polar groups while other atoms such as carbon, oxygen and sulfur are represented by separate force centers. The separation distances of equations (5) and (27) are then counted between the nuclei of these major atoms, as if hydrogen atoms did not exist. The influence of hydrogen atoms is considered through the parameterization of potential parameters. For instance, the Lennard-Jones diameter σ assigned to CH_2 or CH_3 . United Atoms is somewhat larger than the diameter of carbon in All Atom methods. United Atom methods are often used for complex molecules because, with only one-third or one-quarter the number of force centers, they need 5-10 times less computer time than All Atom methods.

Anisotropic United Atom



To take better stock of the influence of hydrogen atoms in United Atom potential, the force center may be shifted to an intermediate position between the major atom and the related hydrogen atoms [16]. In this approach, referred to as “Anisotropic United Atom” or AUA, the CH_2 force center is thus located on the external bisector of the C-C-C angle and the CH_3 force center is located on the C-C axis. The distance from the force center to the major atom is a parameter specific to the group under consideration. The separation distances in equations (5) and (27) are then counted between these force centers. The positions of the atomic nuclei are constructed in exactly the same way as with classical United Atom models, with the same bending and torsion potential, here shown for a CH_3 group (left) and a CH_2 group (right).

4.6 Mesoscale Forcefields

Combining multiple atoms in “beads” has the advantage that the number of centers for which interactions must be calculated is further reduced compared to all atom or united atom force fields. Since the vibrations of hydrogen atoms are also the fastest movements in a molecule, additionally the elimination of explicit hydrogen atoms in mesoscale dynamics simulations means, that the time step in the numerical integration of the equations of motion that underlie molecular dynamics can be substantially larger. Hence, mesoscale simulations calculations access length and time scales that cannot be reached with all atom or united atom force fields, albeit with a reduction in atomic detail..

[16] Soren Toxvaerd, “Molecular Dynamics Calculation of the Equation of State of Alkanes,” *Journal of Chemical Physics* 93, no. 6 (1990): 4290. Soren Toxvaerd, “Equation of State for Alkanes II,” *Journal of Chemical Physics* 107 (1997): 5197.

5 Non-valence Forcefields (Interatomic Potentials)

Non-valence forcefields, or interatomic potentials, do not have pre-defined bonds, angles, and torsions. Instead, all intra-molecular interactions are included in the inter-molecular interactions.

Typically an interatomic potential is designed and parameterized to describe one type of chemical bonding. Over time, this has led to several categorizations of potentials based on the type of bonding for which they are applicable: Buckingham for ionic bonding, Lennard–Jones (LJ) potentials for dispersions and van der Waals (vdW), Tersoff potentials for covalent bonding, and the embedded atom method (EAM) for metallic bonding are well-documented examples. Broadly, such interatomic potentials may be considered **reactive potentials** as they allow the rearrangement of atoms with the evolution of time.

Some interatomic potentials in *MedeA* include:

- Buckingham
- EAM/MEAM
- Stillinger-Weber
- Tersoff
- Streitz-Mintmire
- ReaxFF
- COMB
- SNAP
- NNP

5.1 Units for Energy

The energy per atom or formula unit for interatomic potentials is often expressed in eV.

5.2 Buckingham Potential

The Buckingham (also known as *exp-6*) potential is based on an exponential repulsion term and a dispersion term in $\frac{1}{r^6}$. This functional form is more flexible and may provide a more accurate description of fluid behavior [9] or of cation location in zeolites [10]. It is available in *MedeA* under the form:

$$U_{exp-6} = U_{rep} + U_{disp} = A_{ij} \exp\left(\frac{r_{ij}}{\rho_{ij}}\right) - \frac{C_{ij}}{r_{ij}^6} \quad (27)$$

Combined with a Coulomb potential, $\frac{q_i q_j}{r_{ij}}$, this becomes a well-established potential for describing ionic solids. Charges on atoms are usually full charges (e.g. 4+ for Si and 2- for O) for Buckingham potentials.

5.3 Embedded Atom Method (EAM) Potential

Simulations for metallic systems often employ the *embedded atom model* (EAM) description [17] where the energy of the system is represented by both two-body and electron density-related terms.

[9] Jeffrey R Errington and Athanassios Z Panagiotopoulos, "Phase Equilibria of the Modified Buckingham Exponential-6 Potential From Hamiltonian Scaling Grand Canonical Monte Carlo," *Journal of Chemical Physics* 109, no. 3 (1998): 1093.

[10] Angela Di Lella, N Desbiens, Anne Boutin, I Demachy, Philippe Ungerer, et al., "Molecular Simulation Studies of Water Physisorption in Zeolites," *Physical Chemistry Chemical Physics* 8, no. 46 (2006): 11.

[17] Murray Daw and M Baskes, "Embedded-Atom Method: Derivation and Application to Impurities, Surfaces, and Other Defects in Metals," *Physical Review B* 29, no. 12 (June 1984): 6443-6453.

$$U_{metallic} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N V(r_{ij}) + \sum_{i=1}^N F(\rho_i) \quad (28)$$

$$\rho_i = \sum_{j=1}^N \phi(r_{ij})$$

$U_{metallic}$ is the potential energy, i and j indicate each of the N atoms in the system, and r_{ij} is their interatomic separation. $V(r_{ij})$ is a pairwise potential, $F(\rho_i)$ is the embedding function which depends on the electron density experienced by a given atom, and $\phi(r_{ij})$ is an additional pairwise function, which represents the electron density at any atomic site based on its environment.

Modified EAM (MEAM)

The MEAM potential [18] is an extension to the original EAM potentials which adds angular forces. It is thus suitable for modeling metals and alloys with fcc, bcc, hcp and diamond cubic structures, as well as covalently bonded materials like silicon and carbon.

EAM and MEAM potentials do not include charges and therefore do not have the Coulomb term.

Note: In these potentials, the atomic masses are defined for each atom type within the files containing the interaction parameters. Hence changing, for example, the mass of an atom type in the builder remains ineffective in subsequent LAMMPS runs. If simulations are desired with different isotopes (masses) for the same chemical element, then different atom types have to be specified in the `.frc` files.

5.4 Stillinger-Weber Potential

The simulation of semiconducting materials has led to the development of potential forms that can represent both bond angle deformation and the changing coordination of atoms. Here the simulated material possesses considerable covalent character, favoring the potential form typified by covalent descriptions, but additionally, there is interest in surface restructuring and additional bonding subtleties, prompting interest in the development of potentials able to handle variation in coordination numbers. The Stillinger & Weber [19] potential is an example of such potential combining elements of the EAM approach with the deformation from standard bond lengths employed in the covalent potential. The Stillinger-Weber potential form is summarized by Equations (29) - (31), where $A, \epsilon, B, \sigma, p, q, a$ and λ are adjustable potential parameters, there being 8 such parameters for each distinct atom type covered by the potential.

$$E = \sum_{ij} V_2(R_{ij}) + \sum_{jk} V_3(R_{ij}, R_{ik}, \theta_{ijk}) \quad (29)$$

$$V_2 = A\epsilon \left(B_{ij} \left(\frac{\sigma}{r_{ij}} \right)^p - \left(\frac{\sigma}{r_{ij}} \right)^q \exp \left(\frac{\sigma}{r_{ij} - a\sigma} \right) \right) \quad (30)$$

$$V_3 = \lambda (\cos\theta - \cos\theta_0)^2 \exp \left(\frac{\lambda\sigma}{r_{ij} - a\sigma} \right) \exp \left(\frac{\lambda\sigma}{r_{ik} - a\sigma} \right) \quad (31)$$

As can be seen from Equation (29) the essential form of the Stillinger-Weber potential is the summation of two-body interactions, across all pairs of atoms in the system, augmented with three-body terms, which encompass all triplets of atoms in the system. Both the two-body term (30) and the three-body term, Eq. (31) are damped by exponential functions according to the separation of the atoms involved in the interaction, which implicitly incorporates an effect the system density on the potential. This damping term also makes it possible to truncate interactions to zero at a separation distance governed by the a parameter for both two-body and three-body terms. Additionally, the Stillinger-Weber potential does not require fixed connectivity for the constituent atoms in the system, as the sums of Eq. (29) extend to cover all possible pairs and triplets of atoms within the system.

[18] Baskes, Phys Rev B, 46, 2727-2742 (1992).

[19] Frank H Stillinger and Thomas A Weber, "Computer Simulation of Local Order in Condensed Phases of Silicon," Physical Review B 31, no. 8 (1985): 5262-5271.

Stillinger-Weber does not include charges and therefore does not have the Coulomb term.

5.5 Tersoff Potential

Another well-established potential for covalent systems is the Tersoff potential.[#40_tersoff]_ The Tersoff description incorporates a “bond order” component, a measure of the local bonding state, which can regulate the strength of a bond according to its local environment. The bond order term includes many-body interactions and scales the attractive term thus allowing the breaking of existing bonds and the formation of new bonds. In its simplest form, Tersoff potential has the following terms:

$$E = \sum_{ij} E^R(r_{ij}) + b_{ij} * E^A(r_{ij}) \quad (32)$$

where E^R and E^A are the repulsive and attractive terms, respectively, and b_{ij} is the **bond order** term which takes the form:

$$b_{ij} = (1 + \beta^n \zeta_{ij}^n)^{-\frac{1}{2n}} \quad (33)$$

where β is a scaling factor and ζ includes coordination and angular terms.

Tersoff does not include charges and therefore does not have the Coulomb term.

6 Variable Charge Reactive Potentials

The obvious limitation with the above scheme of development (EAM for metals, Tersoff/SW for covalent materials, and Buckingham for ionic solids) is that none of these methods can seamlessly model complex bonding environments, such as those that occur in a heterogeneous interface between dissimilar materials. This limitation has been overcome by several advances in potential development. While the goal of a universal method applicable to all bond types has not yet been realized, significant progress has been made based on the notion of self-consistent variable charge equilibration.

While the analytical bond order formalism has the demonstrated flexibility to model many systems, the obvious limitation of traditional reactive potentials is the lack of explicit electrostatic effects such as those present in some ionic or heterogeneous systems. Fixed charge Coulomb interactions have been added to several different types of empirical potentials, such as the Buckingham potentials for metal oxides, in addition to extended SW and Tersoff potentials for silica. Although fixed charge schemes work well for bulk-like systems, they are generally less robust in environments that are far from bulk-like or that require charge redistribution in response to changing system conditions; these can include surfaces and other complex microstructural elements. The self-consistent charge equilibration approach, also called the variable or dynamic charge scheme, has addressed this limitation.

6.1 Variable Charge Equilibration

Self-consistent charge equilibration is a scheme by which each atom can dynamically and autonomously determine its charge according to its local environment based on the principle of electronegativity equalization (EE). This is driven by the thermodynamic requirement that the electronegativity, which is equal to the negative value of the electrochemical potential, should be equal at all atomic sites in a closed system at chemical equilibrium [26], [27].

$$\chi_i = -\mu_i = -\frac{\partial E(\rho)}{\partial \rho} = e \frac{\partial E(q_i)}{\partial q_i} \quad (34)$$

where χ is the electronegativity, μ the chemical potential, ρ the electron density, and q the charges. At equilibrium, electron density will transfer between atoms so that chemical potential (electronegativity) at all atomic sites is equalized.

[26] R.G. Parr, et al. Journal of Chemical Physics 68 (8) (1978) 3801–3807.

[27] R.T. Sanderson, Journal of the American Chemical Society 105 (8) (1983) 2259–2261.

In atomistic simulations, the EE principle has encompassed a wide range of successive approximations from density functional theory (DFT). The pioneering work to derive the geometry and connectivity dependent charges under the framework of the EE principle includes the electronegativity equalization method (EEM) [28], [29] proposed by Mortier et al. and the electronegativity equilibration (QEq) method proposed by Rappe and Goddard [30].

6.2 Streitz-Mintmire Potential

The Streitz-Mintmire potential is one of the earliest developments that combines variable charge equilibration (QEq) scheme with other types of potentials. [31] Streitz and Mintmire added QEq to EAM potentials for Al and Al₂O₃, therefore allowing the Al charges to change according to their atomic environment: neutral in bulk-like Al, fully ionized in Al₂O₃, and for those in between Al and Al₂O₃ intermediate charges were found.

6.3 Charge-Optimized Many Body (COMB) Potential

The COMB potential builds on the Tersoff potential and the Streitz-Mintmire potential, by combining the Tersoff potential with variable charge equilibration. [21], [22] Some modifications were introduced, including the addition of dispersion interactions (vdW) and several correction terms.

6.4 Reactive Force Field (ReaxFF)

The ReaxFF potential [23] took a different approach than the COMB potential. Starting from the valence forcefield, ReaxFF added **bond order** terms to stretching, bending, and torsion interactions. Though it does not have pre-defined bonds, bonds, angles, and dihedrals are identified and calculated on the fly. Dispersion (vdW) and Coulomb interactions were also added and are computed for all atom pairs (no exclusions). ReaxFF includes an EEM scheme for variable charge equilibration.

7 Machine Learning Potentials

Machine learning-based methods for energy and force calculation have been used for several years. For example, Blank et al. in 1995 [32] employed a neural net-based methodology to probe the energetics of CO on the Ni(111) surface. Such methods, employing novel descriptors, and machine learning methods, can yield exceptionally accurate reproduction of quantum mechanical training data at substantially reduced computational cost. [33]

MedeA supports two machine learning potentials: SNAP and NNP.

7.1 Spectral Neighbor Analysis Potential (SNAP)

Like the GAP framework of Bartok et al. [34], SNAP uses bispectrum components to characterize the local neighborhood of each atom in a general manner. [24] The total energy is decomposed into a sum over

[28] W.J. Mortier, S.K. Ghosh, S. Shankar, *Journal of the American Chemical Society* 108 (1986) 4315–4320.

[29] G.O.A. Janssens, et al. *Journal of Physical Chemistry* 99 (10) (1995) 3251–3258.

[30] A.K. Rappe, W.A. Goddard III, *Journal of Physical Chemistry* 95 (8) (1991) 3358.

[31] F.H. Streitz, J.W. Mintmire, *Physical Review B* 50 (16) (1994) 11996–12003.

[21] T.-R. Shan, B. D. Devine, T. W. Kemper, S. B. Sinnott, and S. R. Phillpot, *Phys. Rev. B* 81, 125328 (2010)

[22] T. Liang, T.-R. Shan, Y.-T. Cheng, B. D. Devine, M. Noordhoek, Y. Li, Z. Lu, S. R. Phillpot, and S. B. Sinnott, *Mat. Sci. & Eng: R* 74, 255-279 (2013).

[23] Chenoweth, van Duin and Goddard, *Journal of Physical Chemistry A*, 112, 1040-1053 (2008).

[32] Blank, T. B., Brown, S. D., Calhoun, A. W., Doren, D. J. *Neural network models of potential energy surfaces*, *J. Chem. Phys.*, 103(10), 4129 (1995)

[33] Behler, J. and Parrinello, M. *Generalized neural-network representation of high-dimensional potential energy surfaces*, *Phys. Rev. Lett.* 98:146401 (2007)

[34] Bartok, Payne, Risi, Csanyi, *Phys Rev Lett*, 104, 136403 (2010).

[24] Thompson, Swiler, Trott, Foiles, Tucker, *J Comp Phys*, 285, 316 (2015).

atom energies. The energy of atom i is expressed as a weighted sum over bispectrum components, which characterize the radial and angular distribution of neighbor atoms.

With the GAP potentials, Bartok et al. [34] proposed mapping this 3D ball onto the 3-sphere, the surface of the unit ball in a four-dimensional space. In this way, all possible neighbor positions are mapped onto a subset of the 3-sphere. The SNAP potential, assumes a linear relationship between atom energy and bispectrum components. The linear SNAP coefficients are determined using weighted least-squares linear regression against the full QM training set. This allows the SNAP potential to be fit in a robust, automated manner to large QM data sets using many bispectrum coefficients

7.2 Neural Network Potential (NNP)

NNP potentials are be constructed from atom centered symmetry functions (ACSF) by using a neural network approach. Neural networks were proposed in the 1940s to model the function of the human brain and have been developed substantially since their introduction. Neural network potentials are based on feed-forward neural networks which take a set of descriptors G_i as input and predict the potential energy as their output.

For more details on machine learning potentials, refer to the MedeA MLPG: Generating Machine Learning Potentials from First Principles Data section [25].

[25] Behler, J.; Parrinello, M. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces." Phys. Rev. Lett. 2007, 98 (14), 146401