

Hauptseminar Modern Simulation Methods for Structure and Properties of Charged Complex Molecules

Reactive force fields: simulating chemical reactions

Ingo Tischler
Supervisor: Narayanan Krishnamoorthy Anand

1 Introduction

Reactive force fields are molecular dynamic force fields, with which it is possible to render chemical reactions. This handout, and the corresponding Hauptseminar talk, will provide you with some details on how these reactive force fields work. It will be discussed why we need these reactive force fields to simulate chemical reactions and why other kinds of force fields and simulation techniques are not qualified for those purposes. We will focus on the ReaxFF force field, a reactive force field developed in 2000, and how such a force field is able to render chemical reactions.

2 Motivation

Chemical reactions are an essential part of science and engineering, therefore is important how those chemical reactions work in detail. The principle behind chemical reactions, namely the forming and breaking of chemical bonds, is known and understood very well via quantum mechanics. But being aware of the principles is not quite satisfactory for bigger systems, like reactions of large organic molecules or catalyst reactions, because they are too complex to solve them with quantum mechanics. To get an experimental insight in chemical reactions is not that easy either, because of the small length and time scale in combination with thermal fluctuation of the reactants. Therefore it would be favourable if we could simulate chemical reactions.

One method to render such a simulation is the DFT (density functional theory) method. The disadvantage of this kind of simulation is that it scales with order $O(n^3)$. That means if the number of electrons in the system is doubled the computation time increases by a factor of eight. Therefore it would take a huge amount of computational power and time to simulate big reactions. The other candidate for simulating chemical reactions are the MD (molecular dynamics) simulation, they scale with order $O(n)$ and can be applied on much larger time and length scales. On the other hand the problem with classical MD is, that the electrons are not simulated and therefore it is not possible to achieve a chemical reaction within classical MD. An approach of rendering chemical reactions with MD simulations is the usage of a different kind of force field, namely the reactive force fields.

3 ReaxFF

3.1 What is ReaxFF?

Adri van Duin and William Goddard III et al. published their reactive force field in 2001[1]. This force field, which at that point could only handle hydrocarbons, was named ReaxFF. It is a MD force field and so is only able to compute atoms as a whole and does not render electrons. In contrast to a classical force field, like AMBER, ReaxFF uses *bond order* as a part of its energy calculations. An other difference is, that during chemical reactions there is a change in the partial charges of the system and ReaxFF needs to handle charge equilibration, too.

3.2 What is bond order

In chemistry bond order is defined as the number of chemical bonds between a pair of atoms[2]. Because of the coordination of the orbitals of the atoms, two atoms can only have three common bonds and thus the maximum bond order is also three. The higher the bond order between the two atoms is, the stronger they are connected. If there is a stronger connection between the two atoms the distance between them is shorter.

ReaxFF uses the distance between two atoms to define the bond order between those atoms[3].

$$BO_{ij}^{\alpha'} = \exp \left[a_{\alpha} \left(\frac{r_{ij}}{r_{0\alpha}} \right)^{b_{\alpha}} \right] \quad (1)$$

The atom types are indicated by the i and the j . The α names the bond type. The distance between the two atoms is r_{ij} and the $r_{0\alpha}$ is the optimal distance for this bond. The parameters a_{α} and b_{α} are specific to the bond type and the used elements. a_{α} is negative and the bond order is, hence, an exponential

decaying function and converges quickly to zero for greater distances.

There are three different types of bonds: σ -, π - and $\pi\pi$ -bonds (or in other notation sometimes $\sigma-\sigma$ -, $\sigma-\pi$ - and $\pi-\pi$ -bonds). A σ -bond is a bond in which the s-orbital is involved, regardless of its hybridisation[4]. And therefore it occurs in nearly every chemical bond, for example methane: The carbon atom has one covalent bond to each of the hydrogen atoms. Each of the bond is formed by an sp^3 -hybrid orbital. Thus in each of these bonds there is the s-orbital involved and each bond is a σ -bond. When there is a double bond, like in ethene, both carbon atoms have three sp^2 -hybridized and one p-orbital. The double bond is then formed by one sp^2 -orbital and the p-orbital. The bond between the two p-orbitals is called a π -bond. Hence ethene has a π - and a σ -bond. Within a triple bond there are one σ - and two π -bonds. The second π -bond is also called the $\pi\pi$ -bond.

The values of the bond order of each bond type can reach values between zero and one. Therefore The total bond order is the sum of the three individual bond orders.

$$BO_{ij}' = BO_{ij}^{\sigma\sigma'} + BO_{ij}^{\sigma\pi'} + BO_{ij}^{\pi\pi'} \quad (2)$$

The main difference within the simulation of the different bond types is the distance on where they apply. As can be seen in Figure 1 the σ -bond has the longest range, while the π - and the $\pi\pi$ -bonds only occur if both atoms are close to each other. The elements, which form these bonds is also of importance, considering for example hydrogen is only able to build σ -bonds. Consequently the bond orders of hydrogen for π - and $\pi\pi$ -bonds are zero for any distance.

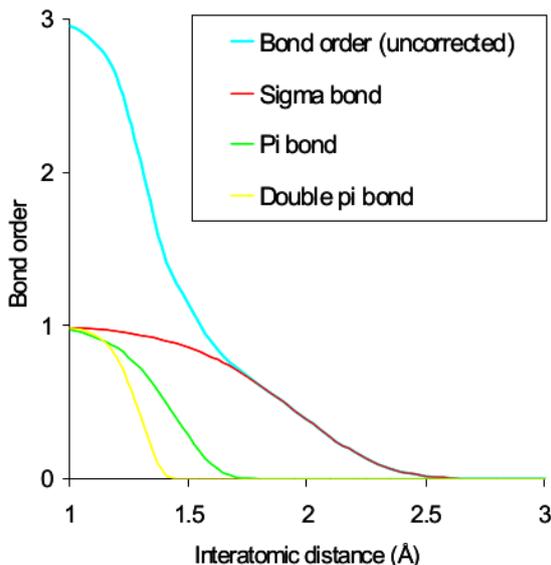


Figure 1: The bond order function (1) plotted against the distance between the atoms.[5]

3.2.1 Bond order correction

With the definition for the bond order up to this point there are two problems. Firstly the bond order has always the maximum value between the chosen atom types. And so two carbon atoms would always form a triple bond, however in chemistry this is rarely the case. This would also have the sum of all bond orders exceed four, if it is bonded to more than one partner. The other problem is, that atoms, which are closely together, would have a bond order, even if there should not be a bond between them.

Figure 2 shows these problems quite well. Ethane should have a single bond between the two carbon atoms and not a triple bond. On the other hand the hydrogen atoms should not have a bond to the carbon atom which is further away, because they already have a bond and hydrogen can only form one bond. With those two factors the bond order of the carbon atom exceeds four with the sum of all bond orders and under normal conditions carbon can have only four bonds. Therefore it is necessary to correct the bond order in a way, that the surrounding bonds are being considered.

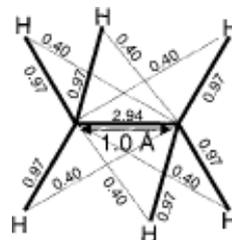


Figure 2: The uncorrected bond order of a ethane molecule [1]

At first a variable which includes all bond orders of the atom needs to be defined. [1]

$$\Delta_i' = \sum_{j=1}^{nbonds(i)} BO_{ij}' - Val_i \quad (3)$$

This Δ_i is the degree of deviation of the sum of the bond orders from its optimal state. The Val_i is this optimal state and is the valency, meaning the number of bonds an atom can have. The valency of a carbon atom is four and of a hydrogen atom is one.

With this degree of deviation and the total bond order between two atoms it is possible to correct this bond order.

$$BO_{ij} = BO_{ij}' \cdot f_1(\Delta_i', \Delta_j') \cdot f_2(\Delta_i', BO_{ij}') \cdot f_2(\Delta_j', BO_{ij}') \quad (4)$$

The BO_{ij} is now the corrected bond order of the bond. The function $f_1(\Delta_i', \Delta_j')$ takes only the surroundings of the two partners into account and hence solves the problem in which two atoms should not have a bond, but have a bond order. This function sets these bond orders to zero. The function $f_2(\Delta_i', BO_{ij}')$ is applied on each bond partner separately and corrects the problem of having always the maximum bond order between two atoms. This correction of the total bond orders can be seen in figure 3. Within this figure there is also a little problem, because if one sums up the bond order of one carbon atom it would be $BO_i \approx 4.6$. Thus the carbon atom is over coordinated. This is due to the length of the single carbon bond, which should be around $l = 1.5\text{\AA}$. Within a simulation this would be corrected via the *over coordination energy*, which will be explained within the next section.

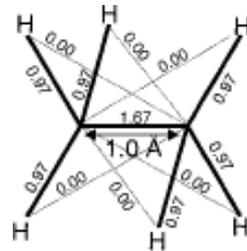


Figure 3: The corrected bond order of an ethane molecule [1]

3.3 Energies

Like in any force field one needs to get the potential energies of the atoms. With that the force on the atoms can be calculated. In a classical force field these energies can be parted in two groups: the bonded and the non bonded interaction energies. Within the topology of the simulation it is fixed which atoms are bonded and which are not. The bonded interaction energies are only calculated between the bonded atoms and on the other hand the non bonded interactions apply only on atoms which are not in a common compound. In ReaxFF there is no such topology which defines what atoms are bonded and which are not. The terms *bonded interactions* and *non bonded interactions* can still be used by the definition, that all bonded interactions need the bond order to be determined.

There are also more energy terms, that need to be taken into account in reactive force fields than in the classical force fields. The bonded interaction energies of a classical force field are the bond energy, the bond angle energy and the bond torsion energy. On the other hand there are the Van der Waals and the Coulomb interaction as the non bonded interaction. In ReaxFF all these interaction energies look as follows.

$$E_{sys} = E_{bond} + E_{lp} + E_{over} + E_{under} + E_{val} + E_{tors} + E_{hbond} + E_{VDW} + E_{Coulomb} \quad (5)$$

- E_{bond} (Bond energy): The bond energy uses the bond order to determine the energy of the bond.

$$E_{bond} = -D_e^{\sigma\sigma} \cdot BO_{ij}^{\sigma\sigma} \cdot \exp[p_{be1}(1 - BO_{ij}^{\sigma\sigma})^{p_{be2}}] - D_e^{\pi\sigma} \cdot BO_{ij}^{\pi\sigma} - D_e^{\pi\pi} \cdot BO_{ij}^{\pi\pi} \quad (6)$$

The D_e s are the dissociation energies of the corresponding bond. The exponential function of the σ bond ensures that the energy goes more smoothly to zero and hence has a better cut-off distance. The bond energy can be seen in figure 4.

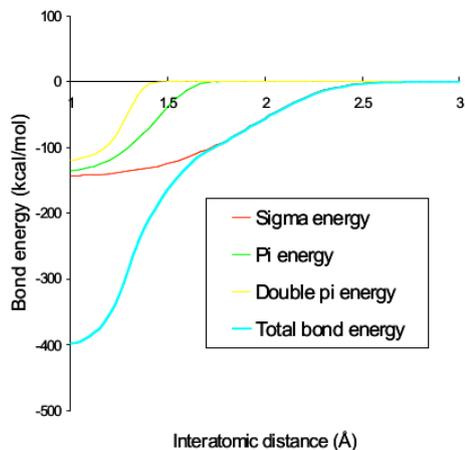


Figure 4: The bond energy plotted against the distance.[5]

- E_{lp} (Lone pair energy): This energy takes the unpaired electrons into account. This means it ensures that the unpaired electron orbitals have their given space around the atom. This leads to the water molecule being angled.
- $E_{over/under}$ (Over/under coordination energy): This energy solves the problem of atoms being over or under coordinated. If an atom is over coordinated there will be a repulsive force and if it is under coordinated there will be an attractive force.
- E_{val} (Valence angle energy): This energy provides the correct angles between two bonds. Therefore the corrected bond orders of these two bonds are needed.
- E_{tors} (Torsion angle energy): The torsion of a bond will be represented by this energy. The torsion is defined by four atoms and their corrected bond orders.
- E_{hbond} (Hydrogen bond energy): This is the energy associated with the hydrogen bond. The bond order of the covalent bond of the hydrogen atom and distance to the hydrogen bond acceptor is used for this calculation.
- E_{vdw} (Van der Waals energy): In the ReaxFF the Van der Waal potential is represented by a Morse potential [6]. However this potential is modified in a way that the strong repulsive forces, which normally occur in close range, can be handled. This needs to be done, because in a classical force field only the non bonded pairs could exchange Van der Waals interactions. In an reactive force field the Van der Waals interactions needs to be calculated for all atom pairs and therefore it is possible for these pairs to be close together. These pairs are bonded and should not have a strong repulsive force between them.
- $E_{Coulomb}$ (Coulomb energy): This energy is the electrostatic interaction potential. Like the Van der Waals potential this potential is also shielded against a divergence for a close range interaction pair. The shielded potentials can be seen in figure 5. The big problem with the electrostatics in ReaxFF is that chemical reactions change the partial charges within a system and this needs also to be calculated. This complication is solved in the next section.

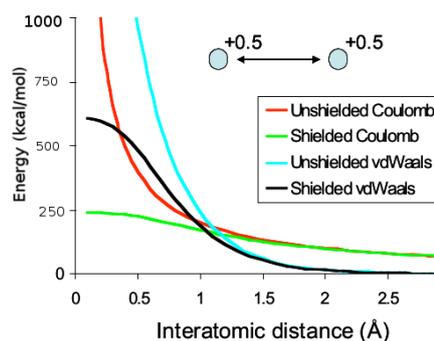
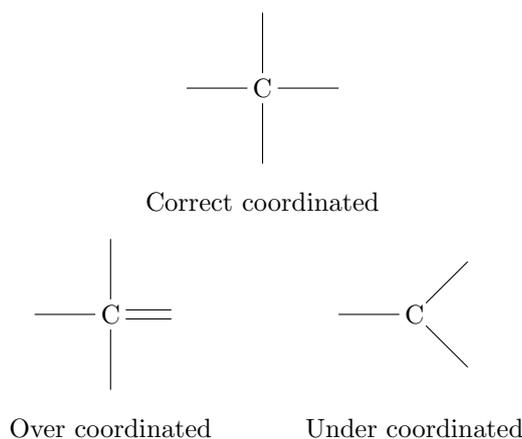


Figure 5: Comparison of the shielded and unshielded Coulomb and Van der Waals potential at close distances [5]

3.4 Charge equilibration

With a classical force field the partial charges of a system are fixed within the topology. Within a chemical reaction, however, these partial charges will change. For example an oxygen molecule has no partial charges, but after an oxidation of a carbon atom to a carbondioxyde both oxygen atoms have a negative partial charge, while the carbon atom is partial positively charged.

One could solve this problem by an *ab-initio* method, but this would take way too long to solve and would destroy the desired order $O(n)$.

The method used in ReaxFF is an approximation to the charge equilibration problem. Within this method [3] the total electrostatic energy is defined as :

$$E(q_1, \dots, q_N) = \sum_i E_i(q_i) + \sum_{i < j} J_{ij} q_i q_j \quad (7)$$

The first part of this equation sums up the atomic energies in respect to q_i . The second sum is the coulomb interaction energy of the system and depends on the positions of all atoms. The trick is that one could get the partial charges if the total electrostatic energy is minimized. To do this the Taylor expansion of the atomic energies is needed.

$$E_i(q_i) = E_{i0} + q_i \left(\frac{\partial E_i}{\partial q_i} \right)_{i0} + \frac{1}{2} q_i^2 \left(\frac{\partial^2 E_i}{\partial q_i^2} \right)_{i0} + \dots \quad (8)$$

The term E_{i0} identifies the energy of a isolated neutral atom. The q_i corresponds to the charge of the atom. Accordingly if the atom has a positive charge of one then there will be the ionisation potential (IP) added to the energy of the isolated neutral atom. This ionisation potential can be measured easily for each atom with experiments. On the other hand there will be the electron affinity (EA) added to the energy of the isolated neutral atom, if the atom has a negative charge of one. With these two assumptions the derivatives in equation (8) can be expressed via:

$$\left(\frac{\partial E_i}{\partial q_i} \right)_{i0} = \frac{1}{2} (IP + EA) = \chi_i^0 \quad (9)$$

$$\left(\frac{\partial^2 E_i}{\partial q_i^2} \right)_{i0} = IP - EA = J_{ii}^0 \quad (10)$$

Now these two equations are inserted in equation (7):

$$E(q_1, \dots, q_N) = \sum_i \left(E_{i0} + \chi_i^0 q_i + \frac{1}{2} J_{ii}^0 q_i^2 \right) + \sum_{i < j} J_{ij} q_i q_j \quad (11)$$

All that needs to be done is now to minimize this equation system to get the partial charges q_i . Sounds easy, however there is a problem and that is that under normal conditions the algorithm of finding a solution to such a problem is in order $O(n^3)$. How this is cut to order $O(n)$ is been described in the next section.

3.5 Computational optimisation

Under non optimized conditions any MD simulation would at least have an computational order of $O(n^2)$. This is because each atom interacts with every other atom. Most of these interactions are small and negligible. For this reason it is important to have a good neighbouring search. Most neighbouring search algorithms combine the use of Verlet-lists and cell lists. In these lists it is noted which atoms are the non negligible interaction partners of a given atom. Therefore only the necessary interactions needs to be calculated. This reduces the computational order to $O(n)$ for the interaction calculation. ReaxFF also uses these neighbouring search techniques for the energy estimation.

It is relevant to find a fast solver to the charge equilibration problem in ReaxFF. Their solution is a Krylow subspace method with an incomplete LU-decomposition. The Krylow subspace method [7] is an iterative sparse matrices [8]. In each step of the iteration the Krylow method uses an approximated solution to get a better approximation to the solution. When using this method in a simulation there will be one iteration each time step. The time step in a reactive force field simulation is smaller than with a classical force field, because the temperatures are higher. Due to this smaller time step the change in the partial charges are also small. Ergo the solution of the last step of the Krylow method can be used as an approximation for the next time step. The Krylow method combined with the incomplete LU-decomposition lead to an order of $O(n)$ for the charge equilibration. How good these approximations for the charge equilibration are in comparison to DFT methods will be a topic in the next section.

With these optimisation steps ReaxFF finally has the desired order $O(n)$. This can be seen in figure 6. The CPMD method is an *ab-initio* method. With this one can see that ReaxFF is one million times faster in simulating a chemical reaction than any DFT method.

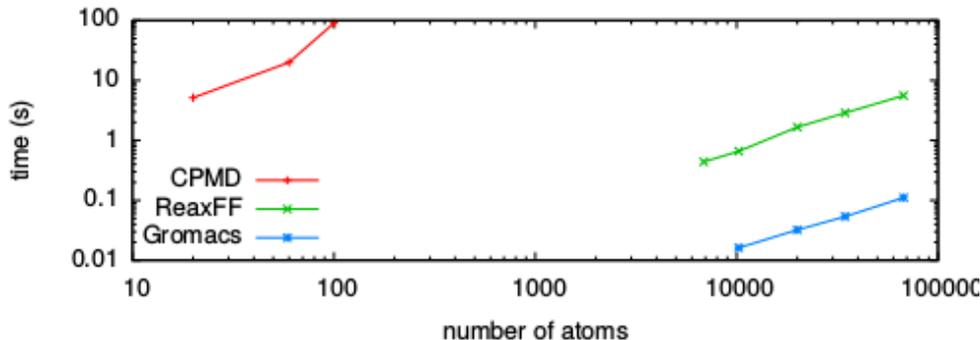


Figure 6: Comparison of the computation time of CPMD, Gromacs and ReaxFF [3]

3.6 Parametrisation

That ReaxFF is fast in comparison to DFT methods was shown, but it is not only the speed of an method that is important. It is the precision and validity of the results that counts. And to do so a good parametrisation is necessary.

To achieve a good parametrisation in ReaxFF they use a successive one-parameter search technique. This method is pretty straightforward and easy to use, but it is possible to get stuck in a local energy minima. Their goal with this parametrisation is the reproduction of data obtained by DFT simulation. One of the optimized values is the bond length of different bonds, this can be seen in table 1.

property	ReaxFF	ab-initio	experimental ¹
C-H bond	1.09 ± 0.01	1.100	1.118 ± 0.006
C-C bond	1.57 ± 0.01	1.533	1.533 ± 0.003

Table 1: Comparison of bond length [\AA] of ethane bonds [1]

But there are many more aspects within a chemical bond, that needs to be considers for a parametrisation. The dissociation process is one of those aspects, and can be seen for ethane in figure 7. One can notice that the energy for the ReaxFF is a good approximation for the DFT method for small distances. If the bond gets stretched further then there is a difference between the two methods and the ReaxFF dissociation energy in the end is less than for the DFT method. This is not that big of a deal, because in ReaxFF there is the under coordination energy that corrects these energy differences. This is due to the under coordination of the carbon atom at the point of dissociation.

One other factor that is important, especially for ReaxFF, is the parametrisation of the correct charge equilibration. A comparison between a DFT method and the method imple-

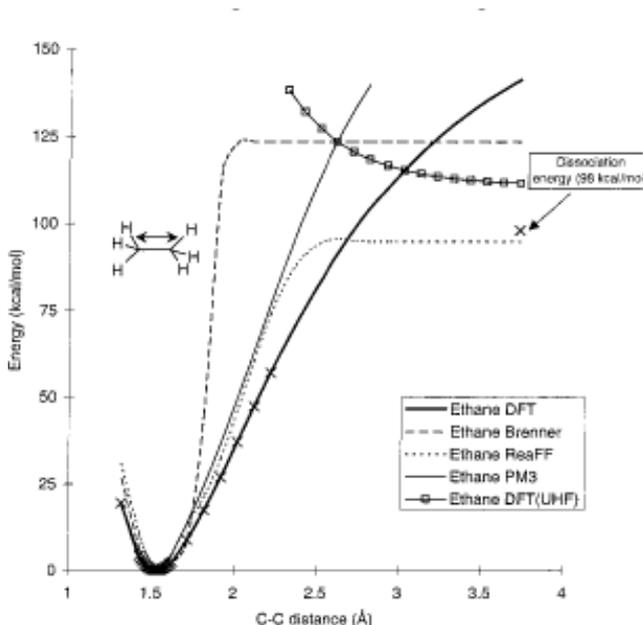


Figure 7: Comparison of the bond dissociation of the C-C ethane bond computed by different methods [1]

mented in ReaxFF can be seen in figure 8. In this figure a snapshot of a simulation of an oxidation process of an hydrocarbon can be seen. In principle the ReaxFF method fits the DFT method well, but in detail there are some differences. For the top picture, for example, the upper hydrogen atom has a charge of 0.10 in the DFT method, but only has the half of it in the ReaxFF. However, these small differences are easily acceptable according to the enormous speed up the ReaxFF has.

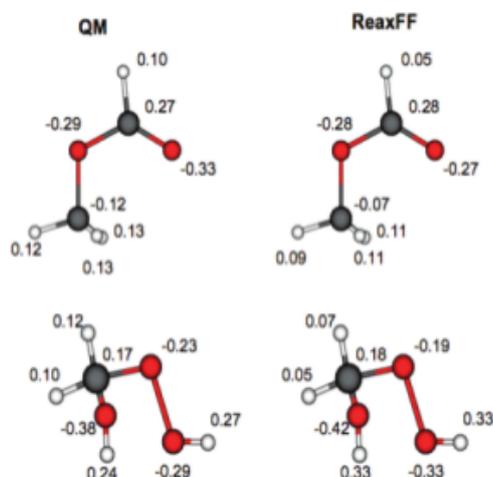


Figure 8: Partial charge comparison between ReaxFF and DFT [9]

3.7 Application

Up until now it was shown how ReaxFF works, but where can it be applied? At first it started as an force field only able to simulate hydrocarbons, but much has changed since then. Up until now there a about 50 different elements, which are possible to simulate. The implemented elements can be seen in figure 9.

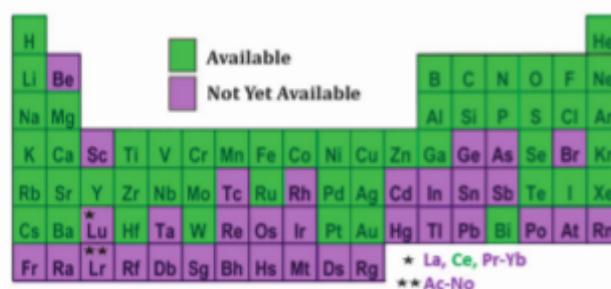


Figure 9: Different elements that can be used in ReaxFF [10]

However, depending on what system one want to simulate, one should use a different distribution of ReaxFF, which fits the system best. Mainly there are currently two different branches in ReaxFF: the combustion branch and the aqueous branch. The combustion branch is mainly used for oxidation processes, while the aqueous branch is used for reactions in water. Therefore it should be clear, why there are different parametrisations of oxygen and the other elements. The different distributions can be seen in figure 10.

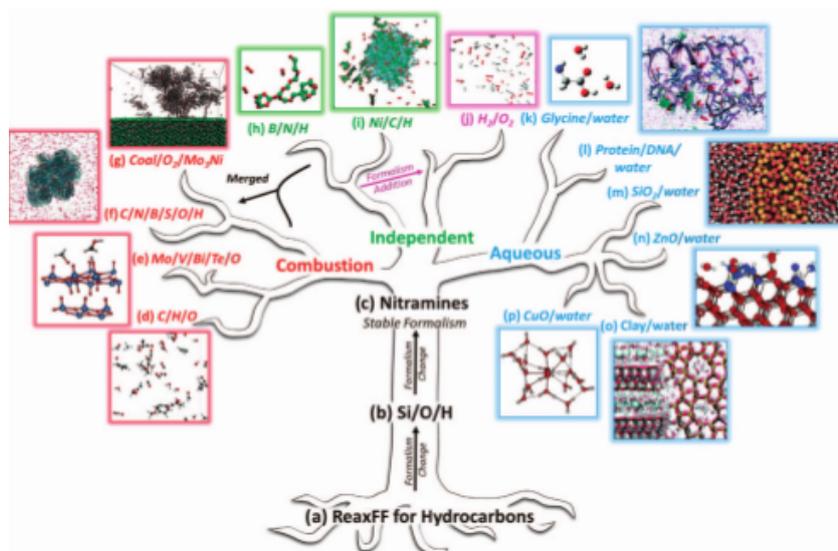


Figure 10: Visualisation of the different branches of ReaxFF [10]

3.8 Future development

The ReaxFF force field is constantly under development in terms of parameter optimisation and the inclusion of new elements. An other field of development is the parallelisability and the introduction of GPU support. This will be an enormous speed up of the simulations, because of the better usage of the computational resources. A code which will implement this features already exists, but it is not publicly available yet [10].

There is also a group that tries to implement a explicit electron description for ReaxFF called eReaxFF. The idea behind this is the introduction of explicit electron an hole-like particles. Off course not all electrons will be rendered in such a model. The goal behind this method to get a better insight on systems such as rechargeable battery interfaces, ferroelectric and piezoelectric materials [10].

4 Conclusion

Within this handout and the corresponding talk the topic of reactive force fields was discussed. It was shown that DFT methods are too slow to render bigger chemical reactions, while a classical MD approach on this problem is not even able to render this problem. The introduction of bond orders concept in the reactive force fields made it possible to achieve the breaking and combining of chemical bonds. The different parts of the energy calculation were outlined and it was explained how the bond orders effect these energies. The problem of an effective method for the charge equilibration was also elucidated. With all this in mind it is now possible to render a chemical reaction a million times faster than with the DFT methods prior to this.

References

- [1] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III, "Reaxff: A reactive force field for hydrocarbons," *The Journal of Physical Chemistry*, 2001.
- [2] https://en.wikipedia.org/wiki/Bond_order (viewed: 16.06.2016).
- [3] H. M. Aktulga, S. A. Pandit, A. C. van Duin, and A. Y. Grama, "Reactive molecular dynamics: Numerical methods and algorithmic techniques."
- [4] https://en.wikipedia.org/wiki/Orbital_hybridisation (viewed: 16.06.2016).
- [5] S. Srinivasan, M. Raju, A. C. van Duin, J. Yeon, S.-Y. Kim, T. Senftle, and K. Joshi, "Presentation: Development and application of the reaxff reactive force field method."

- [6] https://en.wikipedia.org/wiki/Morse_potential (viewed: 23.06.2016).
- [7] <https://en.wikipedia.org/wiki/Krylov-Unterraum-Verfahren> (viewed: 20.06.2016).
- [8] https://en.wikipedia.org/wiki/Sparse_matrix (viewed: 20.06.2016).
- [9] K. Chenoweth, A. C. T. van Duin, and W. A. Goddard III, "Reaxff reactive force field for molecular dynamics simulation of hydrocarbon oxidation," *The Journal of Physical Chemistry*, 2008.
- [10] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin, "Reaxff reactive force-field: development, applications and future directions," *nature partner journals*, 2016.