# Use of DFT in *ab initio* Molecular Dynamics

Jesus Alvarez, University of Chicago

## Abstract

Molecular dynamics (MD) simulations work by generating trajectories for particles calculated from the equations of motion that describe the system. In classical MD, we treat these systems with the Newtonian equations of motion along a potential energy landscape. Within this approximation various methods have arisen, namely from the use different approximations to the potential function that describes the potential landscape. However, this treatment does not allow for the inclusion of more complex electronic effects.

A way to address this problem is by the use of the Schrodinger equation of quantum mechanics instead of Newtonian mechanics as is done in *ab initio* Molecular Dynamics (AIMD). Using many of the approximations familiar to quantum mechanics, we are able to simulate phenomena that are much more difficult to describe using classical MD such as hydrogen bonding and chemical reactions. Using density functional theory (DFT) to help define the potential is one way that allows a more accurate description of these. Here we will explore the differences in the two approaches to Molecular Dynamics (MD and AIMD) by comparison of a typical force field used in classical MD and a DFT calculation that could be used in AIMD, both for water.

## Introduction

The computational study of molecular systems has historically been of great interest as a way to study systems that are hard to observe experimentally and in an attempt to understand mechanisms that are not directly macroscopically observable. One computational method for doing this is Molecular Dynamics (MD) simulations. These simulations model a system of molecules and their dynamic behavior in order to calculate their macroscopic properties. This is very popular for studying the behavior of complex biomolecules such as proteins and lipids. It is also used in materials science in studying the properties of novel materials.

There are many different varieties of MD methods that vary in how they design the system, how they treat the dynamics of the system, and the mathematical and computational approach to the solving the system. For example there are a great variety of numerical algorithms for solving the dynamics of the systems numerically. As the systems are generally composed of a great many atoms, it is extremely difficulty to solve the equations of motion for such large systems analytically. Many numerical integrators exist for this purpose. Another important variation in different MD methods is the treatment of the dynamic evolution of the system. Broadly speaking two main approaches to the treating the dynamics of the system are classical and quantum methods.

Classical Molecular Dynamics simulations treat a system of atoms as particles, restrict the electronic degrees of freedom and move the particles through time using the Newtonian equations of motions. We begin by defining an initial state of the system and certain boundary conditions. This can be done by specifying the initial positions/velocities of the particle(s) and careful choice of the external potential. From this point we allow a short time step, $\Delta t$, and arrive at a new potions and velocity based on the defined conditions. From this new potions and velocity the forces and accelerations can be updated to guide the next change in position after another $\Delta t$ we can repeat this process. This is done for the

duration of the simulation.

The restriction of the electronic degrees of freedom is accounted for in the use of a force field to address the potential of interaction between particles. These different force fields are usually parameterized to fit experimental data and can be tailored to address a specific type of problem. For example many models of proteins make use of the AMBER force field, whereas the CHARMM force field is a very versatile force field used for small and large macromolecules. Another example is the VMOS force field for use with organometallic networks. These force fields work under various assumptions about the system itself in order to model the system as efficiently as they can. While this allows mimicking of the real system, it is easy to overlook fine complex interactions that can have significant impacts on the values calculated from this method.
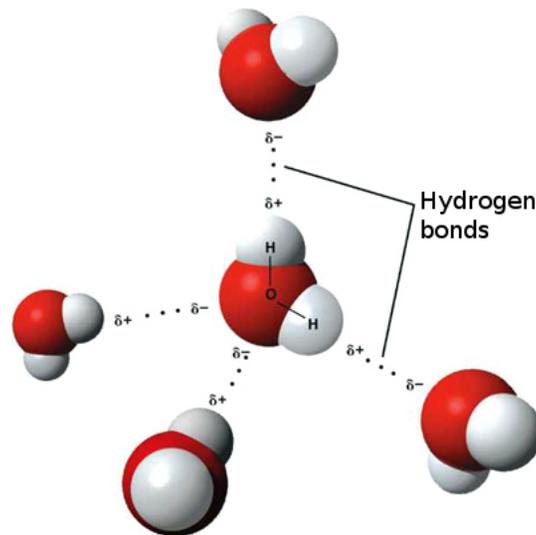


Source: https://en.wikipedia.org/wiki/Hydrogen_bond / CC BY-SA 3.0

An example of a system that suffers from these intricacies is liquid water. Water molecules in bulk experience a great variety of intermolecular forces such as van der Waals, regular electrostatic dipole-dipole interactions, and hydrogen bonding. Hydrogen bonding interactions are especially difficult to describe in classical MD methods. These interactions must all be explicitly expressed in the force field used to propagate the system in space and time.

AIMD methods allow us to better encapsulate these effects by allowing them to be expressed implicitly in the calculation of the trajectory. One of the main differences in AIMD calculations is the use of the Schrodinger equation to calculate the trajectories of the particles, similarly to classical MD. An important extra step unique to AIMD is the performance of an electronic structure calculation at each state. This is typically done using density functional theory (DFT). Use of density functionals in AIMD was first used in Kohn-Sham MD, where the potential energy is composed of the interaction between the electron density and the external potential, and the interaction of the density with the self-created potential - similar to the interaction between electrons as explicit particles.

We can explore the different approach to these methods (classical/force field vs. quantum/ab initio) by examining thermodynamic variables that can be calculated for the same system from each. By comparing the calculated values to experimental values, we can see how they compare and discuss uses for the different methods.

# Application

# Quantum Chemistry

> $Digits := 10;$

$$Digits := 10 \tag{3.1.1}$$

> $with(QuantumChemistry);$

$[AOLabels, ActiveSpaceCI, ActiveSpaceSCF, AtomicData, BondAngles, BondDistances,$ $\qquad$ **(3.1.2)**
$\quad Charges, ChargesPlot, CorrelationEnergy, CoupledCluster, DensityFunctional,$
$\quad DensityPlot3D, Dipole, DipolePlot, Energy, FullCI, GeometryOptimization, HartreeFock,$
$\quad Interactive, Isotopes, MOCoefficients, MODiagram, MOEnergies, MOIntegrals,$
$\quad MOOccupations, MOOccupationsPlot, MOSymmetries, MP2, MolecularData,$
$\quad MolecularGeometry, NuclearEnergy, NuclearGradient, Parametric2RDM, PlotMolecule,$
$\quad Populations, RDM1, RDM2, ReadXYZ, SaveXYZ, SearchBasisSets, SearchFunctionals,$
$\quad SkeletalStructure, Thermodynamics, Variational2RDM, VibrationalModeAnimation,$
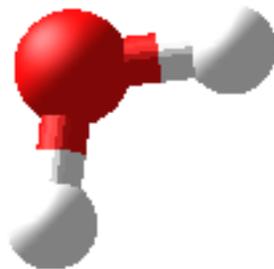$\quad VibrationalModes, Video]$

## Water DFT

Using the *MolecularGeometry* command, we are able to define the molecule (water) in its equilibrium geometry from a database of geometries. We can then compute the electron density for this geometry optimized molecule using the *DensityFunctional* command. Here we have calculated the electron density using the B3LYP density functional with the STO-3G basis set, both of which are the default options for the commands. Having defined the electron density for the molecule, we can then use the Quantum Chemistry package to calculate the thermodynamic quantities for this system using DFT.

$H2O := MolecularGeometry(water);$

$H2O := [["O", 0, 0, 0], ["H", 0.27740000, 0.89290000, 0.25440000], ["H", 0.60680000,$ $\qquad$ **(3.2.1)**
$\quad -0.23830000, -0.71690000]]$

> $PlotMolecule(H2O);$



> $output\_df := DensityFunctional(H2O);$

$$output\_df := \text{table}\left(\left[\left[mo\_symmetry = \begin{bmatrix} "A" \\ "A" \\ "A" \\ "A" \\ "A" \\ "A" \\ \vdots \end{bmatrix}, dipole = \begin{bmatrix} 1.23089830 \\ 0.91126926 \\ -0.64384926 \end{bmatrix}, rdm1 = [\,[ \right.\right.\right.$$ **(3.2.2)**

7 element Vector[column]

$[\,2.10759927, -0.45522640, 0.08020918, 0.05938120, -0.04195521, -0.02338091, \ldots\,]$,

$[\,-0.45522640, 2.02493463, -0.45127962, -0.33409534, 0.23605206, -0.06540376, \ldots\,]$,

$[\,0.08020918, -0.45127962, 1.50079504, -0.12471125, 0.40340245, 0.24464532, \ldots\,]$,

$[\,0.05938120, -0.33409534, -0.12471125, 1.06677636, -0.42338145, 0.66848282, \ldots\,]$,

$[\,-0.04195521, 0.23605206, 0.40340245, -0.42338145, 1.36944451, 0.15522292, \ldots\,]$,

$[\,-0.02338091, -0.06540376, 0.24464532, 0.66848282, 0.15522292, 0.61449713, \ldots\,]$,

$[\,\vdots, \vdots, \vdots, \vdots, \vdots, \vdots, \;]\,]\,]$,

$$\left[\; 7 \times 7 \text{ Matrix} \;\right], mo\_energy = \begin{bmatrix} -18.83329244 \\ -0.92707944 \\ -0.43459524 \\ -0.22897837 \\ -0.14026787 \\ 0.34943936 \\ \vdots \end{bmatrix}, mo\_coeff = [\,[$$

7 element Vector[column]

$[\,0.99342469, -0.22775817, -6.17933348\ 10^{-8}, -0.12261008, -1.25096161\ 10^{-8}, -0.12864443, \ldots\,]$,

$[\,0.02971202, 0.80276922, 3.86586075\ 10^{-7}, 0.60592582, 8.98294654\ 10^{-8}, 0.84897321, \ldots\,]$,

$[\,0.00350462, 0.13865777, -0.13118151, -0.55626303, 0.63602053, 0.55849592, \ldots\,]$,

$[\,0.00259457, 0.10265216, 0.45049354, -0.41181756, -0.38769299, 0.41347397, \ldots\,]$,

$[\,-0.00183317, -0.07252812, 0.38681454, 0.29096570, 0.66721214, -0.29213199, \ldots\,]$,

$[\,-0.00705606, 0.17232479, 0.44499179, -0.28193128, -6.42465011\ 10^{-8},$

$$-0.78544308, \ldots \,],$$

$$[:, :, :, :, :, :, \; ]]],$$

$$\left[ \begin{array}{c} \\ \\ \\ 7 \times 7 \text{ Matrix} \\ \\ \\ \\ \end{array} \right], populations = \left[ \begin{array}{c} 1.99739099 \\ 1.85608024 \\ 1.64463427 \\ 1.31936610 \\ 1.53891948 \\ 0.82180475 \\ \vdots \end{array} \right], aolabels = \left[ \begin{array}{c} \text{"0 O 1s"} \\ \text{"0 O 2s"} \\ \text{"0 O 2px"} \\ \text{"0 O 2py"} \\ \text{"0 O 2pz"} \\ \text{"1 H 1s"} \\ \vdots \end{array} \right], group = \text{"C1"},$$

7 element Vector[column]     7 element Vector[column]

$$mo\_occ = \left[ \begin{array}{c} 2.00000000 \\ 2.00000000 \\ 2.00000000 \\ 2.00000000 \\ 2.00000000 \\ 0. \\ \vdots \end{array} \right], e\_tot = -75.27815045, converged = 1, charges$$

7 element Vector[column]

$$= \left[ \begin{array}{c} -0.35639107 \\ 0.17819525 \\ 0.17819582 \end{array} \right]$$

> $DensityPlot3D(H2O, output\_df\,);$
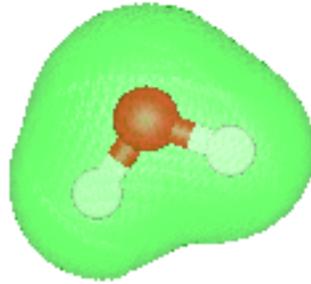
```
>  thermo := Thermodynamics(H2O, 'DensityFunctional');
```

$$thermo := \text{table}\left(\left[\, energy = -1.97510356 \cdot 10^8 \; \frac{J}{mol} \,,\, \theta_B = 20.64153340 \text{ K}, \, entropy \right.\right. \tag{3.2.3}$$

$$= 422.84807088 \; \frac{J}{mol\,K} \,,\, zpe = 63083.94033524 \; \frac{J}{mol} \,,\, free\_energy =$$

$$-1.97633949 \cdot 10^8 \; \frac{J}{mol} \,,\, heat\_capacity = 29.19465116 \; \frac{J}{mol\,K} \,,\, \theta_C = 13.38191298 \text{ K},$$

$$\left.\left. electronic\_energy = -1.97642757 \cdot 10^8 \; \frac{J}{mol} \,,\, enthalpy = -1.97507877 \cdot 10^8 \; \frac{J}{mol} \,\right]\right)$$

# Force Field Method

In 1973 Stillinger and Rahman performed calculations for liquid water in which they calculated various thermodynamic quantities using classical MD simulation. The model that they used treats the water molecule as composed of 4 point charges (two positive and two negative). This is done in order to reflect the dipole of water, wherein the positive point charges represent the partially deshieled protons of the hydrogen atoms and the negative point charges reflect the non-bonded electrons. These charges placed in a tetrahedral arrangement based on the optimal distribution of charges for intermolecular interactions between the molecules in this geometry.

The force field in this work is composed of a Lennard-Jones type potential, with the addition of a modulated Coulomb potential in order to reflect in the intermolecular forces between the water nearest neighbors. The potential used is of the form

$$V(1,2) = V_{LJ}(r_{12}) + S(r_{12})V_{el}(r_{12})$$

where $r_{12}$ is the separation between adjacent O atoms. The first term $[V_{LJ}(r_{12})]$ is a standard Lennard-Jones potential and the second term $[S(r_{12})V_{el}(r_{12})]$ is the modulated Coulomb interaction between molecules. The force field is dependent on $r_{12}$ and the following parameters; $\epsilon$, $\sigma$, $q$, $R_L$, and $R_U$. These parameters are optimized to best imitate the complex intermolecular interactions felt by the water molecules.

$$V_{LJ}(r_{12}) = 4\epsilon[(\sigma/r_{12})^{12} - (\sigma/r_{12})^6]$$

$$V_{el}(r_{12}) = q^2 \sum_{\alpha,\,\beta\,=\,1}^{4} (-1)^{\alpha+\beta}/d_{\alpha\beta}(1,2)$$

The modulation comes from $S(r_{12})$, which essentially turns off the Coulomb interaction at very close distances ($r_{12} < R_U$) and applies a weighting to the Coulomb term up to a defined separation passed which ($R_L < r_{12}$) the full interaction is felt ($S(r_{12}) = 1$).

$$S(r_{12}) = (r_{12} - R_L)^2(3R_U - R_L - r_{12})/(R_U - R_L)^2, \qquad \text{when } R_U < r_{12} < R_L$$

Using this method the constant volume heat capacity ($C_v$) was calculated at 273 K and found to be approximately 145 $\frac{J}{mol \cdot K}$ . This value was then compared to the experimental value that they reported (approximately 74 $\frac{J}{mol \cdot K}$) .

# Discussion

The different computational methods that exist each have their benefits and their costs. The $C_v$ calculated by the classical force field method was approximately 145 $\frac{J}{mol \cdot K}$ and the $C_v$ calculated using DFT was approximately 29 $\frac{J}{mol \cdot K}$ . In the example presented here, both methods differ from the experimental value of approximately 74 $\frac{J}{mol \cdot K}$ by a factor of about 2. However, the DFT value is still only about 35 units below the experimental value, whereas the force field result is about 70 units above. Most importantly the degree of computational efficiency involved in both methods and the way in which they arrive at this value is very different. We can look into a summary of what went into each method to try to compare the reasoning for the disparities.

In the DFT calculation, we had to make specific decisions concerning the definition of our system. First, we defined the geometry of our molecule. Second we performed the calculation of the electron density. In this step, a decision needed to be made with respect to the choice of basis and the density functional to be used. Here we chose to use the default options, a STO-3G basis and the B3LYP density functional. These choices are computationally less intensive, but also allow for a larger margin of error. The calculations could be run using other bases and density functionals that, while potentially much more computationally expensive, could help yield a much more accurate answer.

The force field calculation was found in a study performed by Stillinger and Rahman. A model was used in which the charge distribution about the molecule was formed using point charges with the molecules organized into a tetrahedral arrangement. The potential was composed of a Lennard-Jones potential with a modulated Coulomb term. The parameters in this method had to be optimized to be mimic the intermolecular forces between adjacent water molecules. Previous models upon which the authors were basing their work started from a model in which the point charges were not tetrahedrally arranged and did not consider the modulated Coulomb term. A previous iteration of the potential used in the study worked with the same model and potential, but different parameters. These parameters were then modified for the study.

As the intermolecular forces expressed explicitly in the force field method, the calculations of the trajectories in the MD simulations are more straight-forward and thus computationally less intensive. The implicit nature of the intermolecular forces in the DFT method are more sensitive to the system as a whole at each time step, but are thus more computationally expensive. This results in constraints of the sizes of systems that can be used in each method. Large systems for which there is good experimental data available can be observed over longer time scales using classical MD methods. While the systems for which AIMD can be used must be smaller and run for shorter time scales, the ability for studying more nuanced effects is greater due to the implicit inclusion of the intermolecular interactions calculated at each time step. As with all things, especially in chemistry, there is a fit for purpose and while different each method can be extremely useful in computational modeling of a variety of systems.

# References

(1) Bristow, J. K.; Skelton, J. M.; Svane, K. L.; Walsh, A.; Gale, J. D. A General force field for Accurate Phonon Properties of Metal–Organic Frameworks. Phys. Chem. Chem. Phys. 2016, 18 (42), 29316–29329. https://doi.org/10.1039/C6CP05106E.
(2) Bayly, C. I.; Gould, S. I. R.; Merz, K. M.; Ferguson, D. M.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules.
(3) Iftimie, R.; Minary, P.; Tuckerman, M. E. Ab Initio Molecular Dynamics: Concepts, Recent Developments, and Future Trends. Proceedings of the National Academy of Sciences 2005, 102 (19), 6654–6659. https://doi.org/10.1073/pnas.0500193102.
(4) Hassanali, A. A.; Cuny, J.; Verdolino, V.; Parrinello, M. Aqueous Solutions: State of the Art in Ab Initio Molecular Dynamics. Phil. Trans. R. Soc. A 2014, 372 (2011), 20120482. https://doi.org/10.1098/rsta.2012.0482.
(5) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. J Comput Chem 2009, 30 (10), 1545–1614. https://doi.org/10.1002/jcc.21287.
(6) Paquet, E.; Viktor, H. L. Computational Methods for Ab Initio Molecular Dynamics. Advances in Chemistry 2018, 2018, 1–14. https://doi.org/10.1155/2018/9839641.
(7) Stillinger, F. H.; Rahman, A. Improved Simulation of Liquid Water by Molecular Dynamics. The Journal of Chemical Physics 1974, 60 (4), 1545–1557. https://doi.org/10.1063/1.1681229.
(8) Jensen, F. Introduction to Computational Chemistry, 2nd ed.; John Wiley & Sons: Chichester, England ; Hoboken, NJ, 2007.
(9) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A New Local Exchange-Correlation Functional for Kohn–Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids. J. Chem. Theory Comput. 2016, 12 (3), 1280–1293. https://doi.org/10.1021/acs.jctc.5b01082.
(10) Paesani, F.; Iuchi, S.; Voth, G. A. Quantum Effects in Liquid Water from an Ab Initio-Based Polarizable Force Field. J. Chem. Phys. 2007, 127 (7), 074506. https://doi.org/10.1063/1.2759484.