

# SIMPLE METADYNAMICS: THE DISSOCIATION OF NITRIC ACID ON GRAPHENE<sup>1</sup>

Molecular and Materials Modelling

Wednesday May 29, 2024

Daniele Passerone

daniele.passerone@empa.ch |

---

## 0 Introduction

In this exercise we will study the dissociation reaction of nitric acid on graphene described using coordination numbers. This simulation will illustrate a simple metadynamics run using the coordination numbers as variables. In order to examine these two phenomena, we will complete the following tasks:

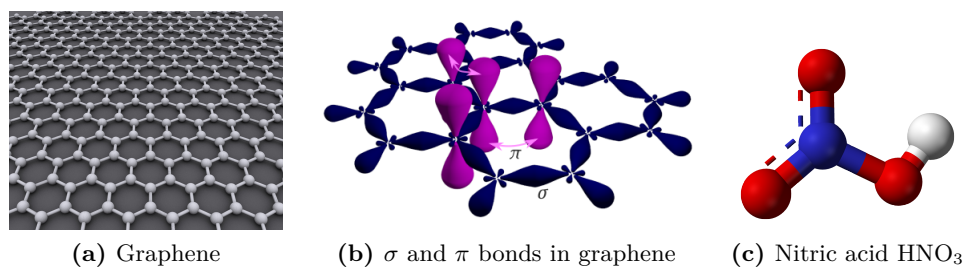
- Set up preliminary simulations to learn about the dynamics of nitric acid on graphene at the **density-functional tight binding** (DFTB) level of theory
- Restrict the configuration space to "interesting" situations by limiting the movement of the two  $\text{HNO}_3$  molecules via introduction of an external potential
- Establish a set of collective variables (CVs) to learn about the typical behaviour of any particular CV along an unbiased MD run
- Employ a metadynamics simulation to trigger the dissociation of nitric acid by following changes of three different coordination numbers

### 0.1 The density-functional tight binding method

The level of theory for the present exercise is that of **density-functional tight binding**, whereby the electronic density  $\rho(r) = \rho_0 + \delta\rho(r)$  is expanded about a reference density  $\rho_0(r)$ :

---

<sup>1</sup>The present exercise, WRITTEN BY RAYMOND AMADOR (Empa) has largely been adopted from [this](#) exercise by Prof. Marcella Iannuzzi at the UZH. In no way do the authors take credit for the originality of this content.



**Figure 1:** A pictorial representation of the systems we will examine in today's exercise. (a) Graphene on an atomic scale, seen here as being made of a hexagonal lattice of carbon atoms. (b) The  $\sigma$  and  $\pi$  bonds in graphene. The  $\sigma$  bonds result from an overlap of hybrid  $sp^2$  orbitals, while  $\pi$  bonds are due to tunneling between  $p_z$  orbitals. Each atom in graphene is connected to its three nearest neighbours by a  $\sigma$ -bond, thereby contributing one delocalized electron to a conduction band that extends over the whole sheet. Graphene conducts heat and electricity very well along its plane. Images and text adopted under courtesy from Wikipedia.

$$\begin{aligned}
 E[\rho(r)] &= E[\rho_0(r)] + \int \delta\rho(\mathbf{r}) \left. \frac{\delta E[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \right|_{\rho_0} \\
 &+ \frac{1}{2} \int \delta\rho(\mathbf{r}) \delta\rho(\mathbf{r}') \left. \frac{\delta^2 E[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} \right|_{\rho_0} + \dots \\
 &+ \frac{1}{p!} \int \dots \int \delta\rho(\mathbf{r}) \delta\rho(\mathbf{r}'') \dots \delta\rho(\mathbf{r}^{(p)}) \left. \frac{\delta^p E[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\dots\rho(\mathbf{r}^{(p)})} \right|_{\rho_0}.
 \end{aligned}$$

We refer to the relevant literature for a more thorough discussion on the theory. Nonetheless: DFTB presents a few computational advantages, in particular with regard to scaling, in comparison to standard DFT by (in some respects) combining the computational accuracy of DFT with the efficiency of TB, especially in systems where the electrons can be considered tightly bound.

## 1 Task one: dynamics of two $\text{HNO}_3$ molecules over a graphene sheet

The goal of this first task is to simulate the dissociation of  $\text{HNO}_3$  molecules with formation of products like  $\text{H}_2\text{O}$  and/or  $\text{NO}$  or  $\text{NO}_2$  fragments. These reactions can occur in the gas phase; however, the reaction should be catalyzed in the presence of C atoms (engl. *soot*; dt. *Glanz- bzw. Hartruß*). In the proposed example, the molecules are located in the vicinity of graphene, which should mimic the role of soot. Graphene is indeed not very reactive, and so better models can be considered using defective or functionalized graphene.

We begin by logging into our aiidalab docker accounts in a similar manner as described in the previous exercises. Open a Terminal, update the course git and move to the `Exercise_13` directory. Copy the `Run_13` directory to daint:

```
$ scp -r Run_13 daint.cscs.ch/scratch/snx3000/YOURACCOUNTNAME
```

In the left panel, click on `scratch` to enter the directory. We move into the `SCRATCH` directory in the bash shell:

```
$ cd scratch
```

Now connect to daint:

```
$ ssh daint.cscs.ch  
course24@daint104: cd $SCRATCH/Run_13
```

The investigation begins with a simple **molecular dynamics** (MD) simulation at a constant temperature of 300 K to learn about the dynamics of the two molecules on graphene. The DFTB description is employed to speed up the simulation, even if this might not be the optimal choice to faithfully describe the dissociation reaction.

In the `gr2hno3_nvt.inp` input file, the section pertaining to DFTB is

```

.....
&DFT
  &QS
    METHOD DFTB
    &DFTB
      SELF_CONSISTENT      F
      DISPERSION            T
      ORTHOGONAL_BASIS     F
    &PARAMETER
      PARAM_FILE_PATH DFTB/nonscc
      SK_FILE   C C   cc
      SK_FILE   H H   hh
      SK_FILE   N N   nn
      SK_FILE   O O   oo
      SK_FILE   O H   oh
      SK_FILE   H O   oh
      SK_FILE   C H   ch
      SK_FILE   H C   ch
      SK_FILE   C N   cn
      SK_FILE   N C   nc
      SK_FILE   O N   on
      SK_FILE   N O   no
      SK_FILE   O C   oc
      SK_FILE   C O   co
      SK_FILE   H N   nh
      SK_FILE   N H   nh
    &END PARAMETER
  &END DFTB
&END QS
&POISSON
  PERIODIC XYZ
  &EWALD
    EWALD_TYPE SPME
    GMAX 25
    O_SPLINE 5
  &END EWALD
&END POISSON
.....

```

The system is fully periodic, and enough space is left above the graphene layer in order to avoid interactions with the images along  $z$ . This represents a very simple model of the type of particles which might trigger the dissociation reaction, and we are not interested in the dynamics of the layer itself. Therefore, a few atoms of the layer are constrained to fixed positions by

```
&CONSTRAINT
  &FIXED_ATOMS
    LIST      48 51 54 57 60 45 59 44 58 43
  &END
&END
```

Change the **RESTART** files as necessary. Otherwise, the **MOTION** section is quite standard for NVT simulations. Since the total number of degrees of freedom is small, even with the thermostat the equilibration to the desired temperature might prove to be difficult. In this light, the rescaling of the temperature is activated by the **TEMP\_TOL** keyword, whenever the difference between the instantaneous temperature and the desired value exceeds the given tolerance.

#### Assignment 1: Density-functional tight binding

Submit the calculation as usual:

```
$ sbatch -A crs01 run_nvt.sh
```

The output of this short 5ps simulation is stored in **DFTB\_NVT**. The kinetic energy (3rd column), temperature (4th), potential energy (5th), and total energy (6th) can be viewed in the **gr2hno3\_nvt-1.ener** file. Visualize the trajectory in a Python notebook. Example including the lines to copy the files back to your docker is in the exercise trajectory. By visualizing the short trajectory, it is observed that the two molecules move very fast, and explore a large portion of configuration space, being often far from each other and far from graphene.

**Estimated runtime on one node: 75min**

## 2 Task two: the external potential

In order to restrict the exploration to regions where the dissociation catalyzed by graphene might occur (and thereby avoiding configurations that are not interesting for this specific study), it is necessary to limit the movement of the two molecules. To this end, an external potential defining a spherical potential centered on the center of the system coordinates is added, which acts only on the two molecules and creates a wall around the "interesting" part of the energy. In order to simplify the definition of the external potential, the coordinates are first centered at zero; these new coordinates can be found in the **grly5x3\_2hno3\_cc.xyz** file. The new updated input file invoking the interaction with this spherical potential is **gr2hno3\_nvt\_epot.inp**, in which the only difference to the first input file is to be found in the **&FORCE\_EVAL** section:

```
&EXTERNAL_POTENTIAL
  ATOMS_LIST  61..70
  FUNCTION    0.000000000001*(Z^2)^4
&END

&EXTERNAL_POTENTIAL
  ATOMS_LIST  61..70
  FUNCTION    0.000000000001*(X^2)^4
&END

&EXTERNAL_POTENTIAL
  ATOMS_LIST  61..70
  FUNCTION    0.000000000001*(Y^2)^4
&END
```

#### Assignment 2: Using the external potential

Change the RESTART files as necessary and run the simulation. Now, along this resulting 10ps trajectory, the two molecules remain close to graphene, where they should be. Visualize the trajectory and comment on the results.

**Estimated runtime on one node: 75min**

## 3 Choosing the CVs

The next step is to set up the CVs which will be later used for the **metadynamics** (MTD) simulations. As we had seen last week, it is important to select good CVs that can describe the relevant configuration along the reaction pathway. Moreover, it is useful to learn about the typical behaviour of the selected CV along an unbiased MD run. Hence, after selecting a set of CVs, preliminary runs should be performed in order to monitor the dynamics of these variables. This can be done setting up MTD simulations, where in fact, no bias is added. The evolution of the variables is then monitored while the system explores the configurations around the initial one, i.e., belonging to the same (initial) basin of attraction on the **free energy surface** (FES), which we also met last week.

The evaluation of typical fluctuation amplitudes of the CVs is particularly important in order to set the width of the Gaussian beads which are going to build up the penalty potential along the "real" MTD run. Moreover, it is important to learn which variations in the CV can occur spontaneously, i.e., those that do not require any bias, and where the CV cannot move without activation.

### 3.1 The C–O coordination number

The input file `gr2hno3_mtd_4cv_h0_p1.inp` has been prepared, establishing and defining our four CVs. Open this input file using `vim`. The first CV is the **coordination number** (CN) of O to graphene:

```

&COLVAR
  &COORDINATION
    KINDS_FROM 0
    KINDS_TO   C
    R_0 [angstrom] 1.8
    NN 8
    ND 14
  &END COORDINATION
&END COLVAR

```

where both NN and ND determine the curvature of the function used to compute the CN, and R\_0 is the reference C–O distance. Then this CV, denoted  $CN_{CO}$  and given by

$$CN_{CO} = \frac{1}{N_O} \sum_{i_O} \sum_{j_C} \frac{1 - \left(\frac{r_{ij}}{R_0}\right)^{NN}}{1 - \left(\frac{r_{ij}}{R_0}\right)^{ND}}$$

describes the interaction between those O atoms which might dissociate from N, and graphene, where adsorption might occur. Moreover,  $CN_{CO}$  should be approximately zero when the molecules are far from the layer; it becomes larger than zero, but always smaller than unity, when one or more O gets closer to the layer. With our given parameters, the CN starts being larger than zero for C–O distances below 4 Å.

### 3.2 The N–O coordination number

The second CV is the CN of N to O, which is approximately 3 for non-dissociated molecules, and smaller when the dissociation begins.

```

&COLVAR
  &COORDINATION
    KINDS_FROM N
    KINDS_TO   O
    R_0 [angstrom] 1.8
    NN 8
    ND 14
  &END COORDINATION
&END COLVAR

```

### 3.3 The H–C coordination number

The third CV is the coordination number between H and graphene, since also H can be lost from the molecules and subsequently adsorbed onto graphene:

```

&COLVAR
&COORDINATION
  KINDS_FROM  H
  KINDS_TO    C
  R_0 [angstrom]  1.6
  NN  8
  ND  14
&END COORDINATION
&END COLVAR

```

### 3.4 The distance between point and plane

The fourth and final CV is the distance between a point and a plane, whereby the point is the geometric center between the two N atoms of the system, and the plane is determined (uniquely!<sup>2</sup>) by the coordinates of three C species of graphene:

```

&COLVAR
&DISTANCE_POINT_PLANE
  &POINT
    TYPE GEO_CENTER
    ATOMS 1
  &END
  &POINT
    TYPE GEO_CENTER
    ATOMS 48
  &END
  &POINT
    TYPE GEO_CENTER
    ATOMS 60
  &END
  &POINT
    TYPE GEO_CENTER
    ATOMS 69 70
  &END
  ATOMS_PLANE 1 2 3
  ATOM_POINT 4
&END DISTANCE_POINT_PLANE
&END COLVAR

```

This last CV restrains the distance of the molecules from the layer, which is an important factor to determine whether dissociation is somehow favoured by the presence of graphene. The output of the unbiased MD run that monitors the behaviour of these 4 CV is contained in DFTB\_MTD\_4CV\_H0, and is obtained by invoking an MTD run where no penalty potential is added.

Hence, the `&FREE_ENERGY` subsection is added within the section `&MOTION`; the MTD

---

<sup>2</sup>This seemingly trivial and almost taken-for-granted factum has its origins (of course) in Euclid's *Elements*. In particular, his proof of Prop. 5 in Book IV can be generalized to the theorem *any three noncollinear points determine a plane*.



run is controlled from the `&METADYN` subsection. In the present case, where no bias has to be added, the `&METADYN` section contains comparatively few parameters:

```
&FREE_ENERGY
  &METADYN
    DO_HILLS  .FALSE.

    &METAVAL
      SCALE 0.08
      COLVAR 1
    &END METAVAL

    &METAVAL
      SCALE 0.3
      COLVAR 2
    &END METAVAL

    &METAVAL
      SCALE 0.08
      COLVAR 3
    &END METAVAL

    &METAVAL
      SCALE 1.5
      COLVAR 4
    &END METAVAL

    &PRINT
      &COLVAR
        COMMON_ITERATION_LEVELS 3
        &EACH
          MD 1
        &END
      &END
    &END METADYN
  &END FREE_ENERGY
```

Change the `RESTART` files as necessary. With the `.FALSE.` argument fed to the `DO_HILLS` parameter, it is required that no bias is added. Then, for each defined `COLVAR`, a MTD variable is initialized. The `PRINT%COLVAR` section controls the printing of the `COLVAR` output file, containing the instantaneous values of the CV as well as other parameters when needed. For the run without bias, no other information is needed, and the only interesting data in the `gr2hno3_mtd_4cv_h0_p1-COLVAR.metadynLog` are the second, third, fourth, and fifth columns, which are instantaneous values of the CV at the respective time (in fs, first column).

**Assignment 3: Time dependence of various CVs**

By plotting the CVs against time, the amplitude of the equilibrium fluctuations can be evaluated and then used to set up the size of the Gaussian hills that build up the biasing potential. In this exercise, we will generate a total of three plots.

The first CV (C–O coordination number) fluctuates close to zero, with fluctuations smaller than 0.2; the second CV (N–O coordination number) is around 2.8, with fluctuations smaller due to the stiffness of the three NO bonds. Generate this plot.

The third CV (H–C coordination number) is also typically zero, yet can change quite significantly when the molecules approach the layer, even if there is no dissociation of H, and no binding to C. This indicates that this variable is difficult to control and might turn out to be tricky to use to distinguish amongst different states of the reaction process. Generate this plot as well.

Finally, generate the plot of the fourth CV (point-plane distance). This plot shows quite large fluctuations and is clearly not suited to distinguish a specific state along the reaction path. Moreover, its minima, when the two molecules are closer to the layer, correspond to the maxima of the third CV, i.e., to the CN of H–C. At least prior to dissociation, the information that this variable provides is redundant.

Nonetheless, it might be interesting to run again this preliminary simulation after modifying the definition of the CV. For example, by changing the two exponents or even the reference distance of the CN, the range of the function can be made shorter or longer. It is maybe important to remind ourselves that the function defining the CV must have a gradient different from zero to affect the behaviour of the system in a MTD run. In more exact words: the MTD force term affecting the dynamics of the atoms involved in the definition of the CV is proportional to the CV function.

## 4 Task four: metadynamics of the dissociation

Now, for the metadynamics of the dissociation of  $\text{HNO}_3$  on a graphene sheet, we remove the point-plane distance as a CV, and thereby employ only the three CNs as CVs as described above. The input file is `gr2hno3_mtd_3cv_p1.inp`, whose input file has already been modified to activate the MTD algorithm:

```
&METADYN
  DO_HILLS
  NT_HILLS 100
  WW 3.0e-3
&METAVAL
  SCALE 0.2
  COLVAR 1
&END METAVAL
&METAVAL
  SCALE 0.3
  COLVAR 2
```

```

&END METAVAR
&METAVAR
  SCALE 0.2
  COLVAR 3
&END METAVAR
&PRINT
  &COLVAR
    COMMON_ITERATION_LEVELS 3
    &EACH
      MD 1
    &END
  &END
  &HILLS
    COMMON_ITERATION_LEVELS 3
    &EACH
      MD 1
    &END
  &END
&END
&END METADYN

```

One Gaussian hill is dropped for every number of steps specified in the `NT_HILLS` parameter; the height of the hill (in Ha) is given by `WW`. These parameters, together with the width of the Gaussian hills, are important to determine the accuracy of the description of the FES via the MTD biasing potential. Since each variable has, in principle, different dimensions and different dynamics, the shape of the hills filling up the  $N_{CV}$ -dimensional configuration space, as defined by the selected CVs themselves, is *not* isotropic: rather, the parameter `SCALE` associated to the  $i$ th MTD variable determines the amplitude of the Gaussian in the  $i$ th space direction of the  $N_{CV}$ -dimensional configuration space.

This parameter, as well as the hill's height and the frequency of collocation, can be changed along the same MTD run by restarting with different values in the input. This feature is useful when the dynamics of some variable changes after some event has occurred (e.g., the fluctuation of a distances becomes larger after a bond breaks), or to modulate the resolution of the biasing potential (coarser or finer) in different regions of the phase space (e.g., coarser at the bottom of the FES basin, and finer in the vicinity of the transition state).

For an efficient exploration of configuration space, it is crucial to spawn hills that are neither too big, lest important features of the topography of the FES might not be sufficiently well resolved, or even the MTD trajectory could follow the wrong path, thereby missing the **minimum energy path** (MEP). On the other hand, filling up the whole space with hills which are too small might require excessively long simulation time. Given the hill size (height and width), and knowing (approximately) the size of the space spanned by the CVs and the barrier height, it is possible to estimate the number of hills needed to fill the basin of the FES and move to the next minimum.

The `HILLS` keyword controls the printing of the `HILLS` file, where the information on the spawned hills is stored: timestep, coordinates of the center in the space of CVs (three CVs means three columns), followed by the width of the hill in each Cartesian direction of space (three more columns), followed by the height (last column).

**Assignment 4: Analysis of the MTD output files**

Submit the calculation and then view the trajectory of the output. The provided trajectory is about 100 ps, and indeed shows the dissociation of the two molecules into  $\text{NO}_2$  and OH, whereby the OH fragments tend to interact with the graphene layer. From the behaviour of the first and second CNs, the evolution of the system can be somehow deduced. In particular, there are clear changes in the N–O CN, which becomes larger when the two molecules get closer to each other, and becomes smaller when the OH is dissociated.

Soon after the dissociation, the coordination is again close to 3, because the lost O is compensated by the fact that the two  $\text{NO}_2$  fragments stay close together, i.e., each N sees the O of the other fragments. The C–O CN becomes larger when the molecules are closer to the layer, and fluctuates quite a lot due to the rapid movement of the molecules. After the dissociation, higher values of the CN are kept for longer time, indicating some more stable interaction of O with C.

A better choice of the parameters defining the C–O CN might help resolve more clearly the two states: O interacting and O not interacting with the layer. As predicted, the very large fluctuations of the H–C CN are difficult to interpret and make this variable not particularly useful for the description of the process. In this light, a CN of N–H, describing the dissociation of H from  $\text{HNO}_3$ , might have been resp. might be a better choice as a third CV.

Other quantities that can be monitored from the COLVAR output, besides the instantaneous values of the three CVs (2nd, 3rd, 4th columns, resp.) are: the instantaneous gradient of the bias potential computed with respect to the CV (5th, 6th, 7th col.); the gradients with respect to the CVs of the wall potentials, if present (8th, 9th, 10th col.); the instantaneous value of the bias potential (11th col.), and the instantaneous values of the wall potentials (12th col.).

**Estimated runtime on one node: 100min**