Description of the CVHD implementation in PLUMED

Kristof Bal

The *collective variable-driven hyperdynamics* (CVHD) method[1] is a powerful and flexible simulation method to study chemical transformations over long time scales. It combines metadynamics[2] with aspects from hyperdynamics.[3] The method has been successfully applied to catalysis,[4] pyrolysis and combustion,[5] graphite etching,[6,7] polymer degradation,[8] and rate calculations of organic reactions.[9] A detailed account of the method can be found in the literature[1] or my PhD dissertation, but this document will restate any key information along the way.

1 Getting and compiling the code

It useful to first get familiar with the PLUMED code[10,11], for which an excellent manual is available.[12] In particular, one should read up on its installation instructions, the procedure to patch an MD code of choice, and try to get a simple metadynamics simulation up & running.

Specifically for CVHD, one should not download the official plumed release, but get my own fork of the code. Check out the right branch using git:

```
mkdir plumed-cvhd
cd plumed-cvhd
git init
git remote add -t cvhd -f origin https://github.com/kbal/plumed2.git
git checkout cvhd
```

Then, PLUMED can be compiled and installed the regular way. However, all CVHD-specfic parts of the code are in a separate module, which has to be enabled through `--enable-modules`:

```
./configure --enable-modules=+cvhd <other flags>
```

2 CVHD-specific instructions

2.1 Global distortion

The first step of any CVHD simulation, is to identify the set of properties (or *local CVs*) that should be biased, collected into a vector $\mathbf{s}$. The goal of the CVHD simulation is to bias the *distortion* χ_i of any local CV s_i: if $\mathbf{s}$ is chosen properly, a large distortion in some χ_i corresponds to a chemical reaction. For the whole set of $\mathbf{s}$, all distortions can be combined into a global distortion χ_t:

$$\chi_t = \left(\sum_i \chi_i^p \right)^{1/p} . \tag{1}$$

Such a behavior can be implemented through the `GLOBALDISTORTION` collective variable.

2.1.1 Biasing bond lengths

In the original formulation of CVHD,[1] only *bond breaking* could be biased. This way, we compute χ_i for each bond r_i as:

$$\chi_i = \begin{cases} 0 & \text{if} \quad r_i < r^{\text{ref}} \\ \dfrac{r_i - r^{\text{ref}}}{r^{\text{ref}}} & \text{if} \quad r_i \geq r^{\text{ref}} \end{cases} . \tag{2}$$

In the code, this is realized as follows:

```
GLOBALDISTORTION ...
  GROUPA=1,2
  GROUPB=3
  P=8
  REF=1.55
  RESET_REF
  RESET_MAXDIST=0.5
  RESET_TIME=1000
  USE_CUTOFF
  REF_MIN=0.0
  REF_MAX=1.8
  LABEL=bdist
... GLOBALDISTORTION
```

This means that

1. We consider all bonds between atoms in `GROUPA` and `GROUPB`, use a power p defined by `P`, and set r^{ref} = `REF`.
2. We only consider contacts that are shorter than `REF_MAX` (and longer than `REF_MIN`) at the first step as real bonds.

3. We assume a chemical reaction happened if the total distortion remains larger than RESET_MAXDIST for RESET_TIME steps. When this happens, a new list of to-be-tracked bonds is created, as in the first step, using the same set of cutoffs REF_MAX (and REF_MIN).

(The values as used in the example might be rather suitable for C–H bonds.)

If one is only interested in interactions within *a single group*, GROUPB may be omitted and only GROUPA will be used.

2.1.2 Biasing bond breaking and/or formation

A new feature of the PLUMED implementation is that you can also bias bonded and non-bonded interactions alike. Instead of directly biasing the bond length, a SWITCHINGFUNCTION is used (the action of which can be found in the PLUMED manual). Because a switching function value close to 0 means no bond, and a value close to 1 points to a bonded interaction, such a choice of **s** allows one to treat bond breaking and formation in the same consistent framework. Here, we can simply take $\chi_i = |s_i - s^{\mathrm{ref}}|$, where s^{ref} can only be 0 or 1. An example:

```
GLOBALDISTORTION ...
  GROUPA=1,2
  GROUPB=3
  P=8
  REF=1.0
  RESET_REF
  RESET_MAXDIST=0.5
  RESET_TIME=1000
  SWITCH={RATIONAL R_0=2.5}
  LABEL=bdf
... GLOBALDISTORTION
```

where we have used the default rational switching function. Furthermore, we do not need to use cutoffs anymore (because all contacts are considered anyway).

Two additional remarks may be of use:

1. If many non-bonded interaction must be biased, calculation of the CV might be very slow. To circumvent this, a *neighbor list* may be used. This list is updated every NL_STRIDE steps, and makes sure that only interaction between REF_MIN and REF_MAX are considered (so these have to be defined again).

2. In practice, it can be very difficult to simultaneously bias bond breaking and formation in the same CV. Hence, it might be more convenient to bias each with a different CV. Through a proper choice of REF_MIN and REF_MAX, a dedicated CV can be made for either. A bond formation CV should operate on small s_i values, while a bond breaking CV should be applied at large values.

What happens if multiple interactions of the same type should be tracked? For example, what should we do if both C–H and C–C bonds should be broken? This cannot be realized in a single `GLOBALDISTORTION` instance, and hence multiple CVs must be defined. They can, however, still be biased through a single CV, by using the clamping function of the next section.

2.2 Clamping function

The second crucial feature of CVHD is its clamping or cutoff function. It deliberately limits the range of the distortion CV and, consequently, of the bias. This way, it can be ensured that no transition states are biased, and that the overall dynamics is uncorrupted. In PLUMED, it is implemented as

$$\eta = \begin{cases} \frac{1}{2}\left(1 - \cos\pi\left(\frac{\chi_t}{\chi^{\text{cut}}}\right)^2\right) & \text{if} \quad \chi_t \le \chi^{\text{cut}} \\ 1 & \text{if} \quad \chi_t > \chi^{\text{cut}} \end{cases} . \tag{3}$$

The choice of χ^{cut} is critical: little biasing will be done if it is too small, but too large a value could result in incorrect dynamics. Depending on the system, values between 0.3 to 0.6 have been used.[1,4–7]

In the following example, we wish to bias two different bond types (say, C–H and C–C) using a single clamped CVHD CV. It works like this:

```
GLOBALDISTORTION ...
  label=cc
  GROUPA=1-10
  (...)
... GLOBALDISTORTION

GLOBALDISTORTION ...
  label=ch
  GROUPA=1-10
  GROUPB=11-32
  (...)
... GLOBALDISTORTION

CVHD ...
  LABEL=cv
  ARG=cc,ch
  P=8
  CUTOFF=0.5
  PERIODIC=NO
... CVHD
```

The two separately defined global distortion cc and ch are combined, again using (1). This is why the power p has to be supplied through the P keyword.

2.3 Biasing

The actual biasing is done using a standard metadynamics routine, accessible in the form of the METAD bias. Next to the standard parameters in the biasing, two additional CVHD-related switches have to be added, and calculation of Tiwary's acceleration factor[13] has to be enabled.

```
METAD ...
  ARG=cv
  SIGMA=0.025
  HEIGHT=0.2
  PACE=500
  BIASFACTOR=20
  ACCELERATION
  CVHD
  CVHD_RESETTIME=1000
  LABEL=mtd
... METAD
```

Unlike the initial implementation of CVHD,[1] more than one CV (as computed with an instance of CVHD) can be biased. It is a particularly convenient way to simultaneously bias bond breaking and bond formation.

Two extremely useful pieces of output are central to any CVHD simulation, and can be extracted from the METAD bias as output components. The first is the acceleration factor acc (or *boost factor*) which measures the average acceleration thanks to the bias ΔV, proportional to $\langle e^{\beta \Delta V} \rangle$. The second is the number of events (reactions) event as detected from the resetting of the bias during the simulation. The following line could be used to write all of this information (and also instantaneous values of CVs and bias) to a file called colvar:

```
FLUSH STRIDE=1000
PRINT ARG=cc,ch,cv,mtd.bias,mtd.acc,mtd.event STRIDE=100 FILE=colvar
```

Here FLUSH STRIDE controls the frequency of writing to the file. Of course, one should carefully check if all variable names are correct.

2.4 Putting everything together properly

Because a CVHD simulation has to be realized through the combination of several CVs, functions and biasing algorithm, some duplicate input cannot be avoided.

GLOBALDISTORTION and CVHD share both the power p that has to be supplied through the keyword P, and the the cutoff χ^{cut}, which is called RESET_MAXDIST in the former and CUTOFF in the latter.

Moreover, `GLOBALDISTORTION` and `METAD` share the the waiting time t_w in the form of `RESET_TIME` and `CVHD_RESETTIME`, respectively.

References

[1] K. M. Bal and E. C. Neyts, *J. Chem. Theory. Comput.* **11**, 4545–4554 (2015).

[2] A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci.* **99**, 12562–12566 (2002).

[3] A. F. Voter, *J. Chem. Phys.* **106**, 4665–4677 (1997).

[4] K. M. Bal, S. Huygh, A. Bogaerts and E. C. Neyts, *Plasma Sources Sci. Technol.* **27**, 024001 (2018).

[5] K. M. Bal and E. C. Neyts, *Chem. Sci.* **7**, 5280–5286 (2016).

[6] D. U. B. Aussems, K. M. Bal, T. W. Morgan, M. C. M. van de Sanden and E. C. Neyts, *Chem. Sci.* **8**, 7160–7168 (2017).

[7] D. U. B. Aussems, K. M. Bal, T. W. Morgan, M. C. M. van de Sanden and E. C. Neyts, *Carbon* **137**, 527–532 (2018).

[8] B. Arash, B. J. Thijsse, A. Pecenko and A. Simone, *Polym. Degrad. Stab.* **146**, 260–266 (2017).

[9] C. D. Fu, L. F. L. Oliveira and J. Pfaendtner, *J. Chem. Theory Comput.* **13**, 968–973 (2017).

[10] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello, *Comput. Phys. Commun.* **180**, 1961–1972 (2009).

[11] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni and G. Bussi, *Comput. Phys. Commun.* **185**, 604–613 (2014).

[12] `http://www.plumed.org/documentation`.

[13] P. Tiwary and M. Parrinello, *Phys. Rev. Lett.* **111**, 230602 (2013).