

SOFTWARE AND TOOLS FOR LIFE SCIENCE APPLICATIONS (VI-SEEM)

**TRAINING MATERIAL 1 FOR THE SCIENTIFIC APPLICATION IN LIFE
SCIENCE:**

CONVENTIONAL VS NOVEL COMPUTER ASSISTED DRUG DELIVERY

(CNCADD)

TITLE:

**CALCULATION OF LOCAL MODE FREQUENCIES (PARTIAL VIBRATIONAL
DENSITY OF STATES) FROM CLASSICAL OR AB INITIO MOLECULAR
DYNAMICS SIMULATIONS**

TASK:

TO EFFICIENTLY COMPUTE ACCURATE LOCAL MODE FREQUENCIES (PARTIAL
VIBRATIONAL DENSITY OF STATES) FROM CLASSICAL OR AB INITIO
MOLECULAR DYNAMICS SIMULATIONS FOR MOLECULAR SPECIES RELEVANT
TO LIFE SCIENCE

SCIENTIFIC RELEVANCE:

VIBRATIONAL SPECTROSCOPY HAS BECOME A RATHER POWERFUL TOOL TO
STUDY EVEN COMPLEX BIOMOLECULAR SYSTEMS. THE METHOD IS RATHER
SIMPLE FROM A TECHNICAL VIEWPOINT, AND, AT THE SAME TIME, PROVIDES
INFORMATION ABOUT THE STRUCTURE, INTRA-AND INTERMOLECULAR
INTERACTIONS OF A BIOMOLECULE, AS WELL AS OF ITS DYNAMICS. WHEN
CERTAIN CHARACTERISTIC GROUPS ARE PRESENT IN A BIOMOLECULAR
SYSTEM THAT HAVE SPECIFIC “ABSORPTION FINGERPRINTS” IN THE
VIBRATIONAL SPECTRA (THE SO-CALLED “VIBRATIONAL CHROMOPHORES),
THEY CAN SERVE AS NANOPROBES TO STUDY INTERACTIONS OF THE
PARTICULAR BIOMOLECULE WITH ITS ENVIRONMENT (THAT CAN BE

COMPLEX TO AN ARBITRARY DEGREE). ASIDE FROM ALL THESE ADVANTAGES, HOWEVER, EXPERIMENTAL VIBRATIONAL SPECTROSCOPIC TECHNIQUES LEAD ONLY TO INFORMATION CONCERNING THE DISTANCES BETWEEN VIBRATIONAL ENERGY LEVELS (ON AN ENERGETIC SCALE), NOT CONCERNING THE LEVELS THEMSELVES. THEREFORE, IN ORDER TO RATIONALIZE THE EXPERIMENTAL FINDINGS AND TO DISTINGUISH BETWEEN SEVERAL POSSIBILITIES, A SOLID THEORETICAL SUPPORT IS REQUIRED.

PARTICULAR EXAMPLES CONSIDERED:

CIS- ISOMER OF FORMIC ACID

THE APPROACH:

The computational approach is a nonstandard sequential one, i.e. it cannot be fully automated. One needs to use an electronic structure computational system, and couple it to our locally developed codes and other codes as well, that need to be modified depending on the particular system studied.

COMPUTATIONAL WORKFLOW (SEE THE END OF THIS DOCUMENT FOR SCHEMATIC PRESENTATION):

STEP 1:

Construct the initial Z-matrix; choose a computational methodology and make an input file for QM calculation (geometry optimization and subsequent harmonic vibrational frequency computations) with the electronic structure system (ESS). See figure 1 below as an example for a particular case when Gaussian09 is used as an ESS.

```
%chk=fa-cis-mp2
%mem=4000MB
%nprocshared=12
# mp2/6-311++g(3df,3pd) opt freq=numer

cis formic acid free, optimization and freq calculations

0 1
C1
```

```

H2,1,RH2C1
O3,1,RO3C1,2,AO3C1H2
O4,1,RO4C1,3,AO4C1O3,2,D0,0
H5,4,RH5O4,1,AH5O4C1,2,D,0
  Variables:
  RH2C1=1.00
  RO3C1=1.50
  RO4C1=1.50
  AO3C1H2=120.0
  AO4C1O3=120.0
  D0=180.0
  RH5O4=1.00
  AH5O4C1=118.0
  D=0.0

```

Fig. 1. Example Gaussian input file for simple geometry optimization and subsequent analytical harmonic frequency calculation starting from cis-conformer of the free formic acid molecule.

STEP 2:

Run the Gaussian jobs. The Gaussian jobs are run on a HPC system in a standard way (depending on the particular workflow management system, architecture and user policies; in certain cases, modifications may be required in the input file especially regarding the amount of RAM allocated and the number of processors used through shared memory).

When the jobs had completed, check out the output files carefully. Have the stationary points on the molecular particular potential energy surfaces (PES) been located? Analyze the results of the subsequent harmonic vibrational analyses and check out if the stationary points correspond to true minima (*i.e.* there are no negative eigenvalues of the mass-weighted Hessian matrices).

If the located stationary points on the considered PESs correspond to true minima, proceed to the next step (3). If not, consider relaxing any symmetry restrictions that may have previously been imposed, reconstruct the initial z-matrix (if necessary), consider starting the geometry optimization with initial computation of the Hessian, or even applying the GDIIS algorithm for geometry optimization.

STEP 3:

If true minimum had been located in the previous step, sequentially generate input files for the BOMD simulation. For this step, one can either pick up the optimized geometry of the cis-formic acid from the Gaussian output file or read the optimized geometry from the checkpoint file from the previous run (using `geom=checkpoint` keyword in Gaussian series of codes), possibly along with the initial guess. A typical simple input file of this type is shown in Fig. 2, given below. Note that prior to doing “production” BOMD runs, one needs to equilibrate the system. For such a simple system, however, in gas phase, equilibration is achieved rather quickly. BOMD simulation is here carried out at $T = 10$ K, computing a total of 10 000 points.

```
%chk=fa-cis-mp2-bomd-10k
%mem=40000MB
%nprocshared=12
# mp2/6-311++g(3df,3pd) bomd=(RTemp=10, MaxPoints=10000)

cis formic acid free, bomd simulation at 10 K

0 1
C
H,1,RH2C1
O,1,RO3C1,2,AO3C1H2
O,1,RO4C1,3,AO4C1O3,2,D0,0
H,4,RH5O4,1,AH5O4C1,2,D,0
  Variables:
RH2C1=1.09898909
RO3C1=1.19547279
RO4C1=1.34976192
AO3C1H2=124.05877934
AO4C1O3=122.40720509
D0=180.
RH5O4=0.96277669
AH5O4C1=108.91540741
D=0.
```

Fig. 2. A simple example Gaussian input file for BOMD simulation of cis-conformer of the free formic acid molecule (starting from the optimized structure at the corresponding level of theory).

STEP 4:

Extract the computed time-series of O–H distances in the course of the production BOMD simulations. This can be done in many ways, using various tools or home-made

codes. For example, one may rather straightforwardly use the awk language (which is a part of every standard Linux distribution) to locate the O–H distance as a function of time directly from the Gaussian output file. Alternatively, a simple code in C or FORTRAN can be written to extract the coordinates of the hydroxyl oxygen and hydrogen atoms during the simulation and subsequently compute the distance as a function of time.

STEP 5:

Compute the time-derivatives of the O–H distance at each time step. This can be done using e.g. a spreadsheet program or a locally developed code.

STEP 6:

From the time series of bond length derivatives, compute the partial autocorrelation function $C_{\dot{\zeta}}(t) = \langle \dot{\zeta}(0)\dot{\zeta}(t) \rangle$. This can be done rather straightforwardly with our locally-developed simple FORTRAN code named “pacf”. This code uses the file with time series of the bond length derivatives, and writes an output file containing $\langle \dot{\zeta}(0)\dot{\zeta}(t) \rangle$.

STEP 7:

Calculate the Fourier transform of $\langle \dot{\zeta}(0)\dot{\zeta}(t) \rangle$. This can be done either with OriginLab, Wolfram Mathematica, R, Matlab, Octave, or any other package, or with locally developed codes.

The calculated pACF corresponding to the O-H stretching vibrational coordinate (the “local mode”) from the computed BOMD trajectories of the cis- rotamer of the free (gas-phase) formic acid is shown in Fig. 3. In Fig. 4, on the other hand, the corresponding Fourier transform of this correlation function is shown (*i.e.* the real part thereof), which actually corresponds to the contribution of the O-H stretching local mode to the overall IR spectrum in the case of this particular formic acid conformer. It actually shows the computed pVDOS corresponding to the O-H stretching mode of the formic acid cis- conformer.

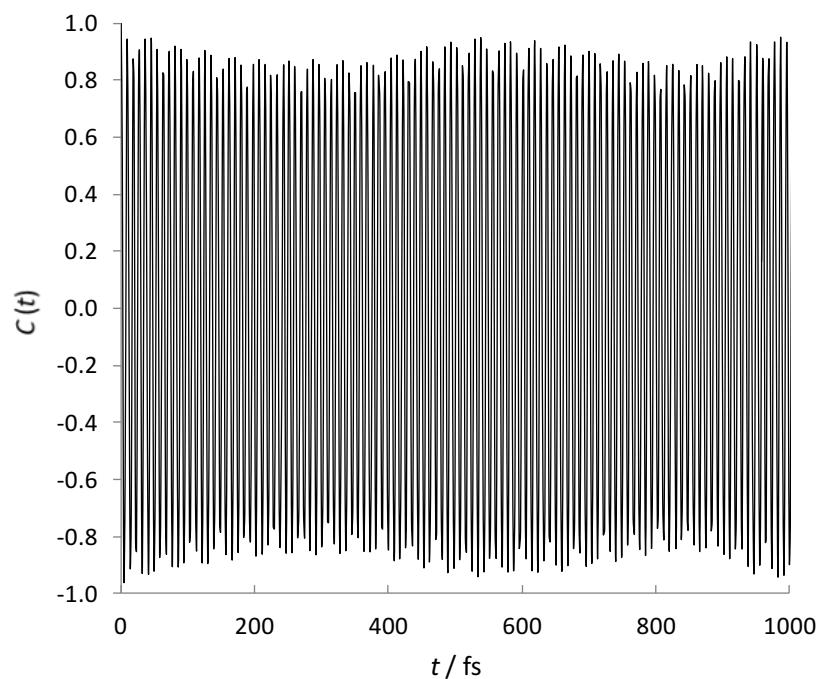


Fig. 3. The pACF for the O-H stretching coordinate time derivative computed from BOMD/MP2/6-311++G(3df,3pd) trajectory for cis- conformer of formic acid

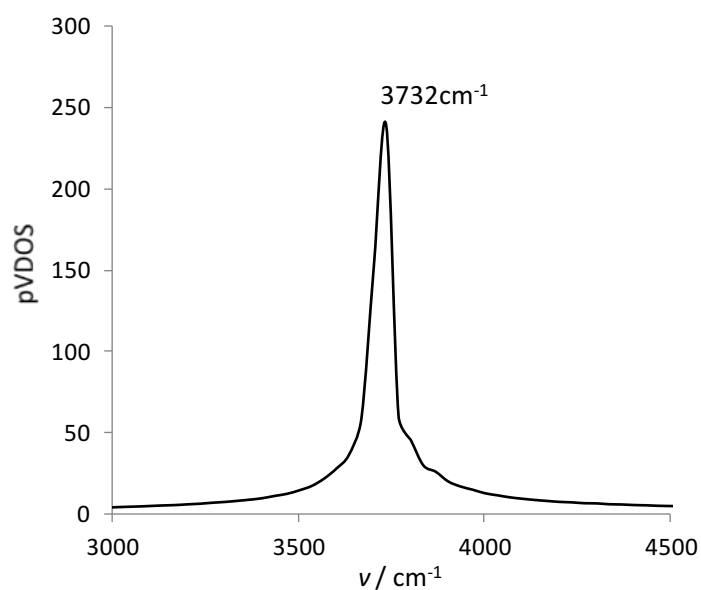


Fig. 4. The pVDOS computed by sequential Fourier transformation of the pACF for the O-H stretching coordinate time derivative computed from BOMD/MP2/6-311++G(3df,3pd) trajectory for cis- conformer of formic acid.

A schematic presentation of the previously outlined algorithm is shown in Fig. 5, given on the next page.

Prepare the initial z-matrix for the studied system, choose the computational method and prepare the input file for the electronic structure system (ESS) code.



Submit the job on the HPC system. When the job is finished, check out the results. Has the stationary point on the explored PES been located? Does it correspond to a real minimum?

No

Consider reconstructing the initial z-matrix, relax the symmetry constraints; also, consider the need to use e.g. the GDIIS algorithm for geometry optimizations, or even initial computation of the Hessian.



Yes

Generate input files for the BOMD simulation. Equilibrate the system. Run the production phase of the simulation.



Extract the computed time-series of O–H distances in the course of the production BOMD simulations.



Compute the time-derivatives of the O–H distance at each time step.



From the time series of bond length derivatives, compute the partial autocorrelation function $C_{\zeta}(t) = \langle \dot{\zeta}(0) \dot{\zeta}(t) \rangle$.



Calculate the Fourier transform of $\langle \dot{\zeta}(0) \dot{\zeta}(t) \rangle$.

