Ab initio Molecular Dynamics

Jürg Hutter

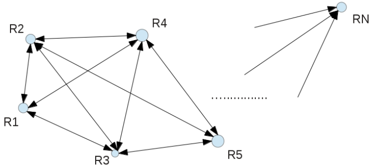
Department of Chemistry, University of Zurich

Overview

- Equations of motion and Integrators
- General Lagrangian for AIMD
- Car–Parrinello MD and Born–Oppenheimer MD
- *ab initio* Langevin MD
- Stability and efficiency
- Thermostats
- External Drivers

Equations of Motion (EOM)

Newton's EOM for a set of classical point particles in a potential.



$$M_I \ddot{\boldsymbol{R}}_I = -\frac{\mathrm{d}V(\boldsymbol{R})}{\mathrm{d}\boldsymbol{R}_I}$$

These EOM generate for a given number of particles *N* in a volume *V* the micro canonical ensemble (NVE ensemble).

The total energy *E* is a constant of motion!

Total Energy

Total Energy $=$ Kinetic energy $+$ Potential energy

$$\text{Kinetic energy} = T(\dot{\boldsymbol{R}}) = \sum_{I=1}^{N} \frac{M_I}{2} \dot{\boldsymbol{R}}^2$$

$$\text{Potential energy} = V(\boldsymbol{R})$$

We will use the total energy as an indicator for the numerical accuracy of simulations.

Lagrange Equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{R}}} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{R}}$$

with

$$\mathcal{L}(\boldsymbol{R}, \dot{\boldsymbol{R}}) = T(\dot{\boldsymbol{R}}) - V(\boldsymbol{R})$$

Equivalent to Newton's EOM in Cartesian coordinates, but is more general and flexible.

- Extended systems
- Constraints

Integration of EOM

Discretization of time

$$\boldsymbol{R}(t) \to \boldsymbol{R}(t+\tau) \to \boldsymbol{R}(t+2\tau) \to \cdots \to \boldsymbol{R}(t+m\tau)$$
$$\boldsymbol{V}(t) \to \boldsymbol{V}(t+\tau) \to \boldsymbol{V}(t+2\tau) \to \cdots \to \boldsymbol{V}(t+m\tau)$$

- Efficiency: minimal number of force evaluations, minimal number of stored quantities
- Stability: minimal drift in constant of motion (energy)
- Accuracy: minimal distance to exact trajectory

Sources of Errors

- Type of integrator
  predictor-corrector, time-reversible, symplectic
- Time step τ
  short time accuracy measured as $\mathcal{O}(\tau^n)$
- Consistency of forces and energy
  e.g. cutoffs leading to non-smooth energy surfaces
- Accuracy of forces
  e.g. convergence of iterative force calculations (SCF, constraints)

Velocity Verlet Integrator

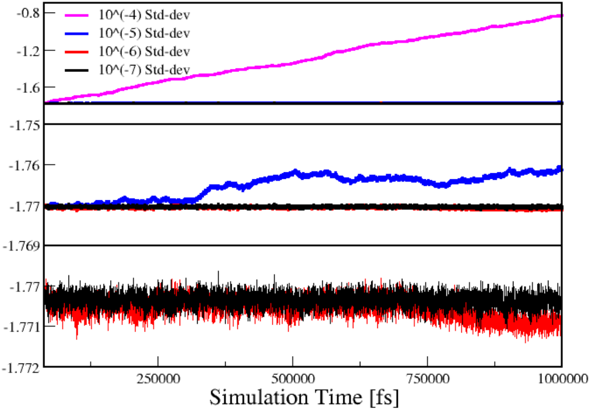
$$\boldsymbol{R}(t+\tau) = \boldsymbol{R}(t) + \tau \boldsymbol{V}(t) + \frac{\tau^2}{2M}\boldsymbol{f}(t)$$

$$\boldsymbol{V}(t+\tau) = \boldsymbol{V}(t) + \frac{\tau}{2M}[\boldsymbol{f}(t) + \boldsymbol{f}(t+\tau)$$

- Efficiency: 1 force evaluation, 3 storage vectors
- Stability: time reversible
- Accuracy: $\mathcal{O}(\tau^2)$
- Simple adaptation for constraints (shake, rattle, roll)
- Simple adaptation for multiple time steps and thermostats

Test on Required Accuracy of Forces

Classical Force Field Calculations, 64 molecules, 330 K
TIP3P (flexible), SPME ($\alpha = 0.44$, GMAX $= 25$),



Stability: Accuracy of Forces

Stdev. Δf Hartree/Bohr	Stdev. Energy μHartree	Drift μHartree/ns	 Kelvin/ns
–	170.35	35.9	0.06
10^{-10}	179.55	-85.7	-0.14
10^{-08}	173.68	6.5	0.01
10^{-07}	177.83	-58.2	-0.10
10^{-06}	—	-385.4	-0.63
10^{-05}	—	9255.8	15.21
10^{-04}	—	972810.0	1599.31

Born–Oppenheimer MD: The Easy Way

$$M_I \ddot{\boldsymbol{R}}_I = -\frac{\mathrm{d}V(\boldsymbol{R})}{\mathrm{d}\boldsymbol{R}_I} \qquad \text{EOM}$$

$$V(\boldsymbol{R}) = \min_{\Phi} \left[E_{\mathrm{KS}}(\{\Phi(\boldsymbol{r})\};\boldsymbol{R}) + \text{const.}\right] \qquad \text{Kohn–Sham BO potential}$$

Forces

$$\begin{aligned} f_{\mathrm{KS}}(\boldsymbol{R}) &= \frac{\mathrm{d}\min_{\Phi} E_{\mathrm{KS}}(\{\Phi(\boldsymbol{r})\};\boldsymbol{R})}{\mathrm{d}\boldsymbol{R}} \\ &= \frac{\partial E_{\mathrm{KS}}}{\partial \boldsymbol{R}} + \frac{\partial \text{const.}}{\partial \boldsymbol{R}} + \sum_i \underbrace{\frac{\partial(E_{\mathrm{KS}} + \text{const.})}{\partial \Phi_i}}_{=0} \frac{\partial \Phi_i}{\partial \boldsymbol{R}} \end{aligned}$$

Computational Details

- System
    - 64 water molecules
    - density $1 gcm^{-3}$
    - Temperature $\approx 330K$
    - Timestep $0.5fs$

- DFT Calculations
    - GPW, TZV2P basis (2560 bsf), PBE functional
    - Cutoff 280 Rydberg, $\epsilon_{\text{default}} = 10^{-12}$
    - OT-DIIS, Preconditioner FULL_SINGLE_INVERSE
    - Reference trajectory ($1ps$), $\epsilon_{\text{SCF}} = 10^{-10}$

Stability in BOMD

Unbiased initial guess; $\Phi(t) = \Phi_0(\boldsymbol{R}(t))$

ϵ_{SCF}	MAE E_{KS} Hartree	MAE f Hartree/Bohr	Drift Kelvin/ns
10^{-08}	$1.2 \cdot 10^{-11}$	$5.1 \cdot 10^{-09}$	0.0
10^{-07}	$9.5 \cdot 10^{-10}$	$5.6 \cdot 10^{-08}$	0.1
10^{-06}	$6.9 \cdot 10^{-08}$	$4.8 \cdot 10^{-07}$	0.4
10^{-05}	$7.4 \cdot 10^{-06}$	$5.6 \cdot 10^{-06}$	2.3
10^{-04}	$3.3 \cdot 10^{-04}$	$5.9 \cdot 10^{-05}$	50.0

Consistent with results from classical MD
Note accuracy of forces!

Efficiency: Initial Guess of Wavefunction

4th order Gear predictor (PS extrapolation in CP2K)

Method	ϵ_{SCF}	Iterations	Drift (Kelvin/ns)
Guess	10^{-06}	14.38	0.4
Gear(4)	10^{-07}	6.47	5.7
Gear(4)	10^{-06}	5.22	11.8
Gear(4)	10^{-05}	4.60	86.8

What is the problem?

Time reversibility has been broken!

Generalized Lagrangian

$$\mathcal{L}(\mathbf{q},\dot{\mathbf{q}},\mathbf{x},\dot{\mathbf{x}}) = \frac{1}{2}M\dot{\mathbf{q}}^2 + \frac{1}{2}\mu\dot{\mathbf{x}}^2 - E(\mathbf{q},\mathbf{y}) + k\mu G(\|\mathbf{x}-\mathbf{y}\|)$$

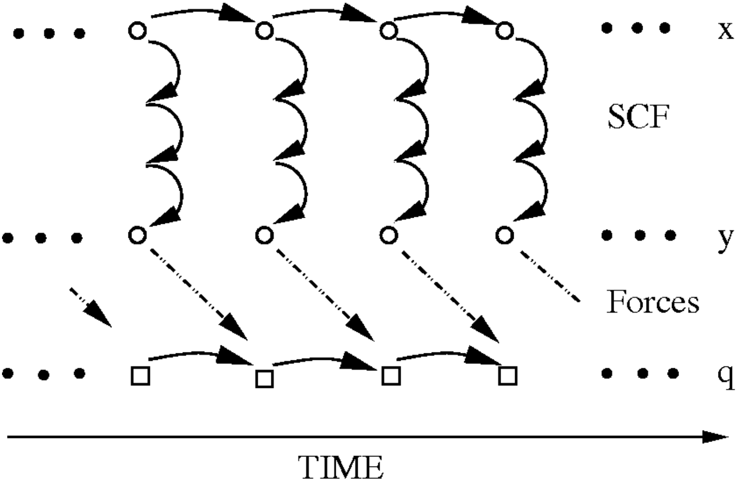
$\mathbf{y} = F(\mathbf{q},\mathbf{x})$	wavefunction optimization
$G(\|\mathbf{x}-\mathbf{y}\|)$	wavefunction retention potential

Equations of motion

$$\begin{aligned} M\ddot{\mathbf{q}} &= -\frac{\partial E}{\partial \mathbf{q}} - \frac{\partial E}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{q}} + k\mu\frac{\partial G}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{q}} \\ \mu\ddot{\mathbf{x}} &= -\frac{\partial E}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{x}} + k\mu\left[\frac{\partial G}{\partial \mathbf{x}} + \frac{\partial G}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{x}}\right] \end{aligned}$$

Extension of Niklasson Lagrangian, PRL **100** 123004 (2008)

Dynamical System



Car–Parrinello Molecular Dynamics

$$\mathbf{y} = \mathbf{x} \quad \Rightarrow \quad G(\|\mathbf{x} - \mathbf{y}\|) = 0$$

Lagrangian

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} M \dot{\mathbf{q}}^2 + \frac{1}{2} \mu \dot{\mathbf{x}}^2 - E(\mathbf{q}, \mathbf{x})$$

Equations of motion

$$M\ddot{\mathbf{q}} = -\frac{\partial E}{\partial \mathbf{q}}$$

$$\mu\ddot{\mathbf{x}} = -\frac{\partial E}{\partial \mathbf{x}}$$

Properties of CPMD

- **Accuracy:** Medium
  Distance from BO surface controlled by mass μ
  Requires renormalization of dynamic quantities (e.g. vibrational spectra)

- **Stability:** Excellent
  All forces can be calculated to machine precision easily

- **Efficiency:** Good
  Efficiency is strongly system dependent (electronic gap)
  Requires many nuclear gradient calculations

Not implemented in CP2K!

BOMD

$$\mathbf{y} = \mathrm{Min}_{\mathbf{x}} E(\mathbf{q}, \mathbf{x}) \quad \text{and} \quad \mu = 0$$

Lagrangian

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} M \dot{\mathbf{q}}^2 + E(\mathbf{q}, \mathbf{y})$$

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} \dot{\mathbf{x}}^2 + kG(\|\mathbf{x} - \mathbf{y}\|)$$

Equations of motion

$$\left.\begin{aligned} M\ddot{\mathbf{q}} &= -\tfrac{\partial E}{\partial \mathbf{q}} \\ \ddot{\mathbf{x}} &= -k\tfrac{\partial G}{\partial \mathbf{x}} \end{aligned}\right\} \text{decoupled equations}$$

BOMD with Incomplete SCF Convergence

$$\mathbf{y} = F(\mathbf{q}, \mathbf{x}) \approx \text{Min}_{\mathbf{x}} E(\mathbf{q}, \mathbf{x}) \quad \text{and} \quad \mu = 0$$

Equations of motion

$$M\ddot{\mathbf{q}} = -\frac{\partial E}{\partial \mathbf{q}} - \frac{\partial E}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{q}}$$

$$\ddot{\mathbf{x}} = -k\frac{\partial G}{\partial \mathbf{x}} - \frac{\partial G}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{x}}$$

SCF Error: Neglect of force terms $\frac{\partial E}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{q}}$ and $\frac{\partial G}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{x}}$.

EOM are coupled through terms neglected!

ASPC Integrator

Integration of electronic DOF (**x**) has to be

- accurate: good wavefunction guess gives improved efficiency
- stable: do not destroy time-reversibility of nuclear trajectory

ASPC: Always Stable Predictor Corrector

- J. Kolafa, J. Comput Chem. **25**: 335–342 (2004)
- ASPC(k): time-reversible to order $2k + 1$

Orthogonality Constraint

Wavefunction extrapolation for non-orthogonal basis sets

$$\mathbf{C}_{\text{init}} = \sum_{j=0}^{K} B_{j+1} \underbrace{\mathbf{C}_{t-j\tau} \mathbf{C}_{t-j\tau}^{T} \mathbf{S}_{t-j\tau}}_{\mathbf{PS}} \mathbf{C}_{t-\tau}$$

Coefficients B_i are given by ASPC algorithm

$$\mathbf{x} = \mathbf{C}_{\text{init}}$$

$$\mathbf{y}[t] = \mathbf{C}_t$$

$\mathbf{S}$ overlap matrix

J. VandeVondele et al. Comp. Phys. Comm. **167**: 103–128 (2005)

Importance of Time-Reversibility

Method	ϵ_{SCF}	Iterations	Drift (Kelvin/ns)
Guess	10^{-06}	14.38	0.4
ASPC(3)	10^{-06}	5.01	0.2
ASPC(3)	10^{-05}	3.02	4.5
Gear(4)	10^{-07}	6.47	5.7
Gear(4)	10^{-06}	5.22	11.8
Gear(4)	10^{-05}	4.60	86.8

Efficiency and Drift

Method	ϵ_{SCF}	Iterations	Drift (Kelvin/ns)
Guess	10^{-06}	14.38	0.4
ASPC(4)	10^{-06}	5.01	0.2
ASPC(4)	10^{-05}	3.02	4.5
ASPC(4)	10^{-04}	1.62	1742.4
ASPC(4)	10^{-02}	1.03	21733.2

Efficiency and Drift

Method	ϵ_{SCF}	Iterations	Drift (Kelvin/ns)
ASPC(4)	10^{-04}	1.62	1742.4
ASPC(5)	10^{-04}	1.63	1094.0
ASPC(6)	10^{-04}	1.79	397.4
ASPC(7)	10^{-04}	1.97	445.8
ASPC(8)	10^{-04}	2.06	24.1

Langevin BOMD

Starting point: BOMD with ASPC(k) extrapolation

Analysis of forces

$$f_{\mathrm{BO}}(\boldsymbol{R}) = \underbrace{f_{\mathrm{HF}}(\boldsymbol{R}) + f_{\mathrm{Pulay}}(\boldsymbol{R})}_{f(\boldsymbol{R})} + f_{\mathrm{nsc}}(\boldsymbol{R})$$

- f_{BO} : correct BO force
- f_{HF} : Hellmann–Feynman force
- f_{Pulay} : Pulay force
- f_{nsc} : non-self consistency error force
- f : approximate BO force

Forces in Approximate BOMD

Approximate f_{nsc} by

$$\tilde{f}_{\mathrm{nsc}} = -\int \mathrm{d}\boldsymbol{r} \left[\left(\frac{\partial V_{\mathrm{xc}}(\rho^{i})}{\partial \rho^{i}} \right) \Delta\rho + V_{\mathrm{H}}(\Delta\rho) \right] \nabla_{I} \rho^{i}$$

with $\Delta\rho = \rho^{o} - \rho^{i}$, ρ^{o} final (output) density, ρ^{i} initial (predicted) density.

Now assume

$$f(\boldsymbol{R}) + \tilde{f}_{\mathrm{nsc}}(\boldsymbol{R}) = f_{\mathrm{BO}}(\boldsymbol{R}) - \gamma_{D} \dot{\boldsymbol{R}}$$

where γ_{D} is a constant friction parameter.

Langevin EOM

$$M\ddot{\boldsymbol{R}} = f_{\mathrm{BO}}(\boldsymbol{R}) - (\gamma_D + \gamma_L)\dot{\boldsymbol{R}} + \Theta$$

with Θ a Gaussian random noise term and

$$\langle \Theta(0)\Theta(t) \rangle = 6(\gamma_D + \gamma_L) M k_B T \delta(t)$$

Given temperature $T = \frac{k_B}{3} \langle M\dot{\boldsymbol{R}}^2 \rangle$ and an arbitrary γ_L this determines γ_D

Langevin BOMD

2nd Generation Car–Parrinello (SGCP)

T. D. Kühne et al., Phys. Rev. Lett. **98** 066441 (2007)

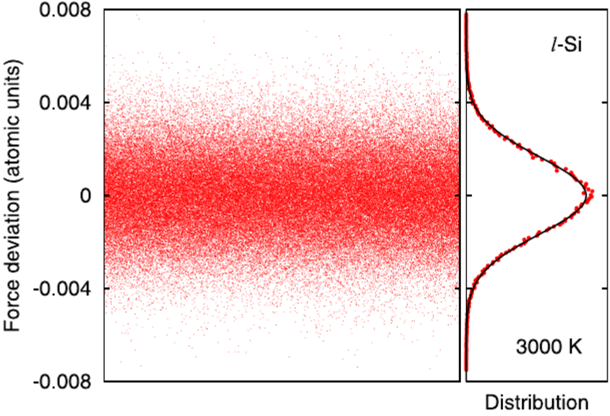
- single SCF step plus force correction needed
  extremely efficient for systems with slow SCF convergence
- γ_D is small: correct statistics and dynamics
- difficult to stabilize in complex systems

SGCP: Examples

T. Musso et al. Eur. Phys. J. B, **91** 148 (2018)

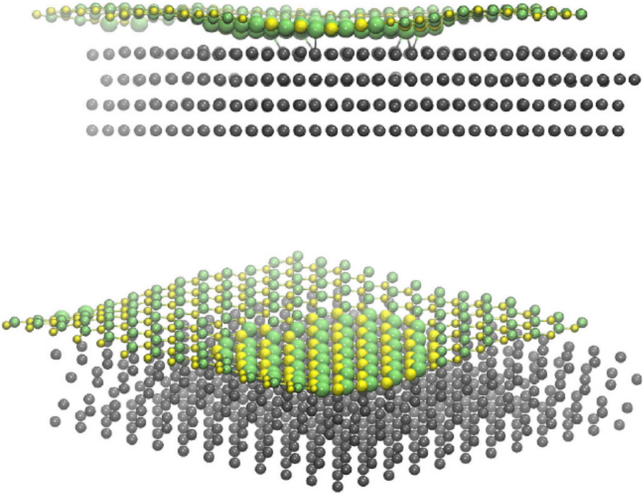
Analysis of Force Errors

Liquid Silicon at 3000 K

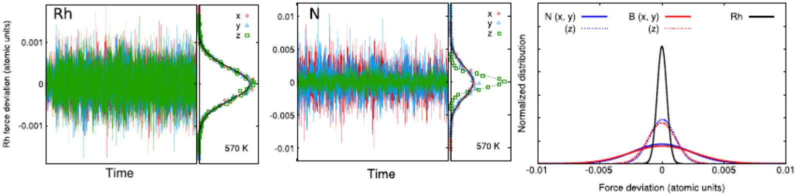


SGCP: Nanomesh

h-BN on Rh(111)



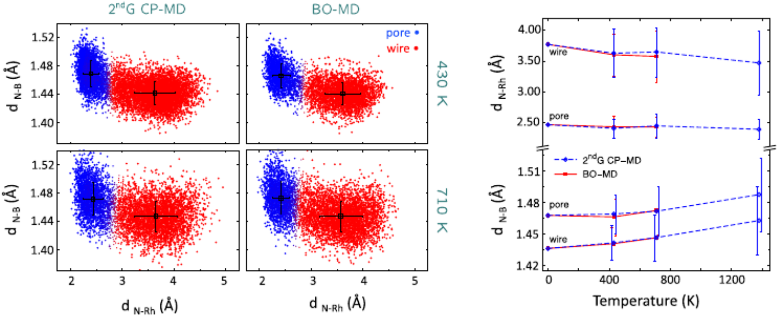
SGCP: Force Errors



	γ_L	γ_D B	γ_D N	γ_D Rh
430 K	0.0010	0.0004	0.0004	0.0002
710 K	0.0010	0.0035	0.0005	0.0020
1380 K	0.0010	0.0007	0.0000	0.0010

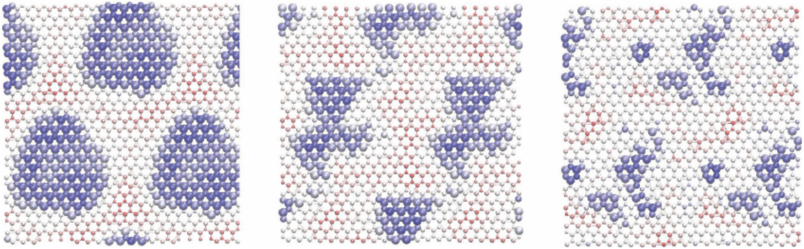
SGCP: h-BN on Rh(111)

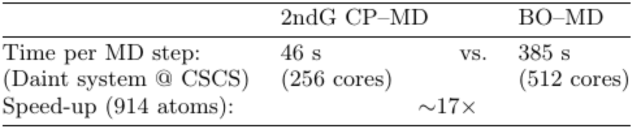
Comparision of MD Results



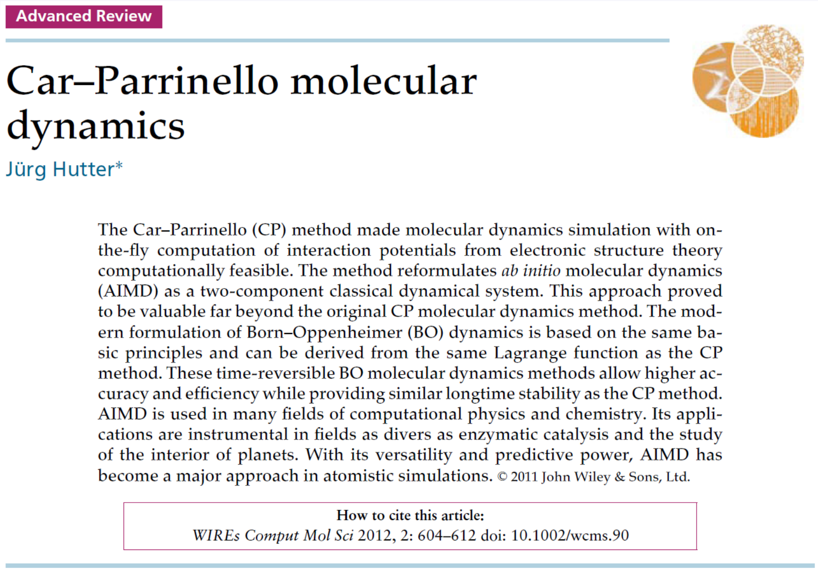
SGCP: h-BN on Rh(111)

Performance





	2ndG CP–MD		BO–MD
Time per MD step:	46 s	vs.	385 s
(Daint system @ CSCS)	(256 cores)		(512 cores)
Speed-up (914 atoms):		∼17×	



Advanced Review

Car–Parrinello molecular dynamics

Jürg Hutter*

The Car–Parrinello (CP) method made molecular dynamics simulation with on-the-fly computation of interaction potentials from electronic structure theory computationally feasible. The method reformulates *ab initio* molecular dynamics (AIMD) as a two-component classical dynamical system. This approach proved to be valuable far beyond the original CP molecular dynamics method. The modern formulation of Born–Oppenheimer (BO) dynamics is based on the same basic principles and can be derived from the same Lagrange function as the CP method. These time-reversible BO molecular dynamics methods allow higher accuracy and efficiency while providing similar longtime stability as the CP method. AIMD is used in many fields of computational physics and chemistry. Its applications are instrumental in fields as divers as enzymatic catalysis and the study of the interior of planets. With its versatility and predictive power, AIMD has become a major approach in atomistic simulations. © 2011 John Wiley & Sons, Ltd.

How to cite this article:
WIREs Comput Mol Sci 2012, 2: 604–612 doi: 10.1002/wcms.90

Micro-Canonical Ensemble

- System variables: **N, V, E**
- Microstate definition: $\Gamma = \{r^N, p^N\}$
- Energy $H(\Gamma) = \sum_i \frac{p_i^2}{2m_i} + U(r^N)$
- All states have the same weight $\delta(H(\Gamma) - E)$
- Partitionfunction

$$\Omega(N, V, E) = \sum_{\Gamma} \delta(H(\Gamma) - E) = \frac{1}{N} \frac{1}{h^{3N}} \int dr^N dp^N \delta(H(\Gamma) - E)$$

Molecular Dynamics

- Fixed number of particles: **N**
- Fixed volume: **V**
- Periodic boundary condition to avoid surface effects
- Total energy

$$E(t) = \sum_i \frac{m_i}{2} \dot{r}_i^2(t) + U(r(t))$$

Constant of Motion: E

$$\begin{aligned}
\frac{dE(t)}{dt} &= \sum_i \frac{m_i}{2} 2\dot{r}_i \ddot{r}_i + \sum_i \frac{dU(r)}{dr_i} \dot{r}_i \\
F_i &= -\frac{dU(r)}{dr_i} = m_i \ddot{r}_i \\
\dot{E} &= \sum_i m_i \dot{r}_i \frac{F_i}{m_i} + \sum_i (-F_i) \dot{r}_i \\
&= \sum_i (\dot{r}_i F_i - \dot{r}_i F_i) \\
&= 0
\end{aligned}$$

The energy is conserved. MD samples correct microstates.

Definition of Temperature

Instantaneous temperature

$$T(t) = \frac{2}{3Nk_B}\frac{1}{2}\sum_i m_i \dot{r}_i^2(t)$$

Temperature

$$T = \langle T(t) \rangle$$

Equipartition Theorem

The equipartition theorem is a general formula that relates the temperature of a system with its average energies. For a system described by the Hamiltonian (energy function) $H(r, \dot{r})$

$$\left\langle r_i \frac{\partial H}{\partial r_j} \right\rangle = \left\langle p_i \frac{\partial H}{\partial p_j} \right\rangle = \delta_{ij} k_B T$$

For an ideal gas

$$\langle H \rangle = \langle H_{\mathrm{kin}} \rangle = \frac{3}{2} k_B T$$

Canonical Ensemble

- System variables: **N, V, T**
- Closed system exchanging energy with a heat bath at temperature T.

How to simulate on a computer?

- We can directly manipulate velocities (temperature).
- No extended heat bath needed.
- Molecular dynamics with periodic boundary conditions, but manipulate $\dot{r}_i$ during the integration.

Control of velocities/temperature

- *Differential control:* the temperature is fixed to the predescribed value and no fluctuations occur.
- *Proportional control:* velocities are corrected at each integration step through a coupling constant towards the prescribed temperature. The coupling constant determines the strength of the fluctuations around $\langle T \rangle$.
- *Integral control:* the Hamiltonian is extended and variables are introduced wich reflect the effect of an external system which fix the temperature. Time evolution is derived from the extended Hamiltonian.
- *Stochastic control:* position and velocities are propagated according to modified equations of motion with friction and stochastic forces. The final equations give the correct mean value of the temperature.

NVE

$$\Gamma = \{\mathbf{r}^N, \mathbf{p}^N\} \qquad \text{microstate}$$

$$\mathcal{H}(\Gamma) = \sum_{i=1}^{N} \frac{1}{2m_i}\mathbf{p}_i + \mathcal{U}(\mathbf{r}^N) \qquad \text{const.}$$

All accessible microstates are equally probable

$$W_{\mathrm{NVE}}(\Gamma) = \delta(\mathcal{H}(\Gamma) - E) \qquad \text{statistical weight}$$

Distribution of microstates

$$\Omega_{\mathrm{NVE}} = \sum_{\Gamma} \delta(\mathcal{H}(\Gamma) - E) \approx \frac{1}{N!}\frac{1}{h^{3N}} \int d\mathbf{r}^N \mathbf{p}^N \; \delta(\mathcal{H}(\Gamma) - E)$$

Canonical Ensemble (NVT)

Microstates follow the Boltzmann's distribution

$$W_{\mathrm{NVT}}(\Gamma) = \exp\left[-\frac{\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{k_{\mathrm{B}}T}\right]$$

Canonical partition function: describes the statistical properties

$$\mathcal{Q}_{\mathrm{NVT}} = \sum_{\Gamma} \exp\left[-\frac{\mathcal{H}(\Gamma)}{k_{\mathrm{B}}T}\right] \approx \frac{1}{N!h^{3N}} \int d\mathbf{r}^N d\mathbf{p}^N \; \exp\left[-\frac{\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{k_{\mathrm{B}}T}\right]$$

Probability distribution

$$\mathcal{P}(\Gamma) = \frac{W_{\mathrm{NVT}}(\Gamma)}{\mathcal{Q}_{\mathrm{NVT}}}$$

and the expectation value of the total energy is ($\beta \equiv 1/k_{\mathrm{B}}T$)

$$\langle E \rangle = -\frac{\partial \ln \mathcal{Q}_{\mathrm{NVT}}}{\partial \beta} = \frac{1}{\mathcal{Q}_{\mathrm{NVT}}} \left\{ \sum_{\Gamma} \mathcal{H}(\Gamma) \exp\left[-\frac{\mathcal{H}(\Gamma)}{k_{\mathrm{B}}T}\right] \right\}$$

Velocity Rescaling

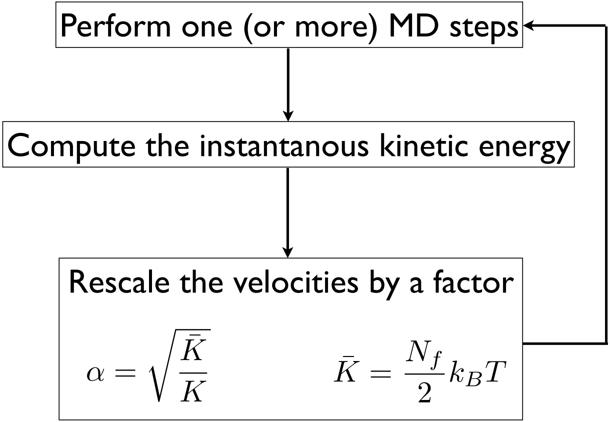
At t and temperature $T(t)$ all velocities are multiplied by λ

$$\Delta T = \frac{1}{2}\sum_{i=1}^{N}\frac{2}{3}\frac{m_i(\lambda v_i)^2}{Nk_\mathrm{B}} - \frac{1}{2}\sum_{i=1}^{N}\frac{2}{3}\frac{m_i v_i^2}{Nk_\mathrm{B}} = (\lambda^2 - 1)T(t)$$

To control the temperature

$$\lambda = \sqrt{T_0/T(t)}$$

- Easy to implement
- Sampling does not correspond to any ensemble
- Not recommended for production MD
- Good for equilibration
- Not time-reversible, not deterministic



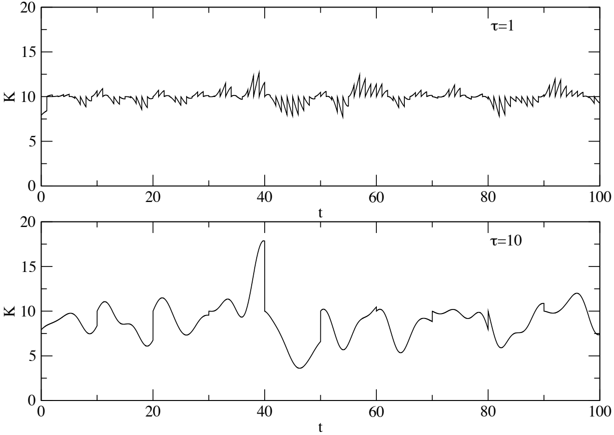
Perform one (or more) MD steps

Compute the instantanous kinetic energy

Rescale the velocities by a factor

$$\alpha = \sqrt{\frac{\bar{K}}{K}} \qquad \bar{K} = \frac{N_f}{2} k_B T$$

Jumps by Rescaling



Berendsen Thermostat

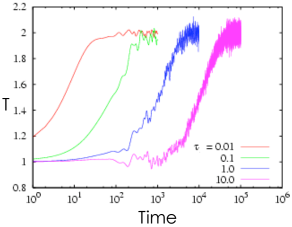
Weak coupling to a heat bath: temperature correction

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_{\text{bath}} - T(t))$$

τ coupling parameter: larger τ weaker coupling
Exponential decay of the system towards the desired temperature

$$\Delta T = \frac{\delta t}{\tau}(T_{\text{bath}} - T(t))$$

$\tau \to \infty$, thermostat is inactive
too small τ, unrealistically low fluctuations
$\tau = \delta t$, standard rescaling
Same problems as velocity rescaling



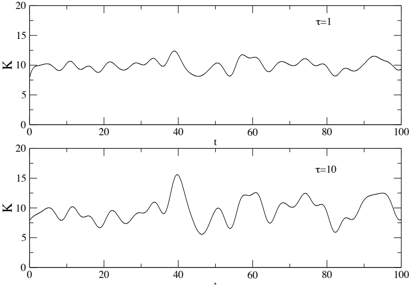
Like a friction

Continuous formulation of the velocity rescaling

$$\dot{\mathbf{p}}_i = \mathbf{f}_i - \zeta \mathbf{p}_i$$

$$(1 - \zeta \delta t) = \left(1 - \frac{\delta t}{\tau} \left(\frac{\sum_i \frac{\mathbf{p}_i^2}{m_i}}{N k_{\mathrm{B}} T} - 1 \right) \right)^{1/2}$$

Friction positive when $\mathcal{K} > N k_{\mathrm{B}} T/2$, $\Rightarrow$ cooling.



Langevin Thermostat

Random forces due to impacts exerted by molecules in the environment: irregular motion and friction. At each step all $\mathbf{v}_i$ are corrected by a random force and a friction term

$$m\ddot{\mathbf{r}}_i = -\frac{\partial \mathcal{U}}{\partial \mathbf{r}_i} - m\Gamma \dot{\mathbf{r}}_i + \mathbf{W}_i(t)$$

Frictional force and random motion are related, i.e. **fluctuation–dissipation theorem** is satisfied

$$\langle \mathbf{W}_i(t), \mathbf{W}_j(t') \rangle = \delta_{ij}\delta(t - t') 6 m \Gamma k_{\mathrm{B}} T$$

Gaussian distribution with infinitely short correlation time (great number of collisions)

Transition probability

$$\frac{\partial \mathcal{P}}{\partial t} = \sum_i \left(-\frac{\mathbf{p}_i}{m_i} \frac{\partial \mathcal{P}}{\partial \mathbf{r}_i} + \frac{\partial \mathcal{U}}{\partial \mathbf{r}_i} \frac{\partial \mathcal{P}}{\partial \mathbf{p}_i} + m_i D \frac{\partial^2 \mathcal{P}}{\partial \mathbf{p}_i^2} + \Gamma \frac{\partial \mathbf{p}_i \mathcal{P}}{\partial \mathbf{p}_i} \right)$$

Stochastic Behavior

The canonical distribution is then recovered as

$$\lim_{t\to\infty} \mathcal{P} = C \exp\left(-\frac{1}{k_\mathrm{B} T}\{\sum_i \frac{m\mathbf{p}_i^2}{2} + \mathcal{U}(\mathbf{r}^N)\}\right)$$

and friction and diffusion are related

$$D = \frac{\Gamma}{m} k_\mathrm{B} T$$

Local thermostat: noise term depends on the particle

- Samples from canonical ensemble and ergodic.
- Allows larger time steps compared to non-stochastic thermostats.
- Momentum transfer is destroyed, i.e., no diffusion coefficients.

Nose Thermostat

Heat bath integral part of the system: s, $\dot{s}$, and Q (coupling)

$$\mathcal{L} = \sum_i \frac{ms^2\dot{\mathbf{r}}_i^2}{2} - \mathcal{U}(\mathbf{r}^N) + \frac{1}{2}Q\dot{s}^2 - gk_\mathrm{B}T\ln s$$

The conjugate momenta in the extended system are

$$\mathbf{p}_i \equiv \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{r}}_i} = m_i s^2\dot{\mathbf{r}}_i \qquad p_s \equiv \frac{\partial\mathcal{L}}{\partial\dot{s}} = Q\dot{s}$$

where the auxiliary variable is a time-scaling parameter

$$dt' = sdt \qquad \mathbf{p}_i' = \mathbf{p}/s \qquad \text{real system momentum}$$

Microcanonical ensemble in the extended system: for $g = 3N + 1$
canonical ensemble for the real system

Nose–Hoover

Using only differentiation in real time t': $\frac{1}{s}\frac{d}{dt'}$

$$\begin{aligned}
\dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} \\
\dot{\mathbf{p}}_i &= -\frac{\partial \mathcal{U}}{\partial \mathbf{r}_i} - \xi \mathbf{p}_i \qquad \xi = s p_s / Q \\
\dot{\xi} &= \frac{1}{Q}\left(\sum_i \frac{\mathbf{p}_i^2}{m_i} - \frac{g}{\beta}\right) \\
\frac{\dot{s}}{s} &= \frac{d \ln s}{dt} = \xi
\end{aligned}$$

$$\mathcal{H}_{\text{Nose}} = \sum_i \frac{\mathbf{p}_i}{2m_i} + \mathcal{U}(\mathbf{r}^N) + \frac{\xi^2 Q}{2} + g\frac{\ln s}{\beta} \qquad \text{conserved}$$

Friction Coefficient

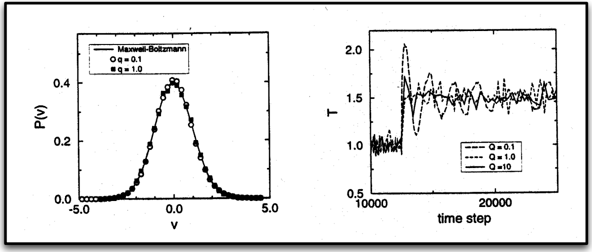
Define the relaxation time as

$$\nu_T \equiv \sqrt{3Nk_{\mathrm{B}}T/Q}$$

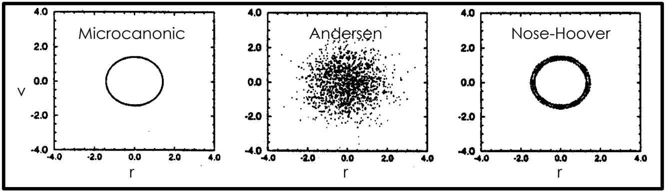
dynamics of friction coefficient by (feed-back mechanism)

$$\dot{\xi} = \nu_T^2 \left[\frac{\sum_i \frac{\mathbf{p}_i^2}{m_i}}{3Nk_{\mathrm{B}}T} - 1 \right] = \nu_T^2 \left[\frac{\mathcal{T}}{T} - 1 \right]$$

Too large Q: canonical distribution after long simulation
Too small Q: high-frequency T oscillations



Ergodicity Problems



NVE: closed loop, constant energy shell
Andersen: points corresponding to a Gaussian velocity distribution
Nose–Hoover: band trajectory determined by the initial conditions

Not sufficiently chaotic to sample all phase-space; trapped in a subspace

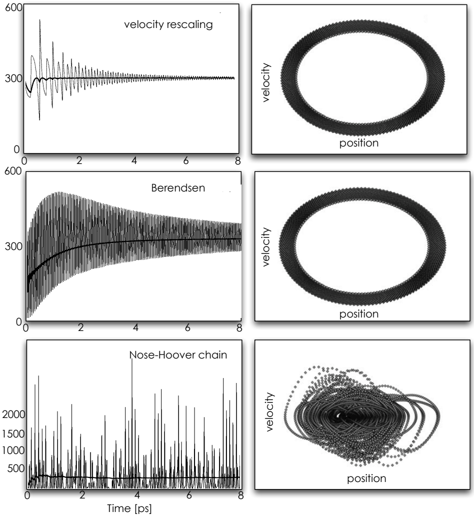
Nose–Hoover Chain

Ergodicity is improved by thermostatting the thermostat variable: M Nose–Hoover thermostats

$$
\begin{aligned}
\dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} \\
\dot{\mathbf{p}}_i &= -\frac{\partial \mathcal{U}}{\partial \mathbf{r}_i} - \xi_1 \mathbf{p}_i \\
\dot{\xi}_1 &= \frac{1}{Q_1}\left(\sum_i \frac{\mathbf{p}_i^2}{m_i}\,\frac{g}{\beta}\right) - \xi_i \xi_2 \\
\dot{\xi}_j &= \frac{1}{Q_j}\left(Q_{j-1}\xi_{j-1}^2 - k_\mathrm{B} T\right) - \xi_j \xi_{j+1} \\
\dot{\xi}_M &= \frac{1}{Q_M}\left(Q_{M-1}\xi_{M-1}^2 - k_\mathrm{B} T\right)
\end{aligned}
$$

$$
\mathcal{H}_\mathrm{NHC} = \sum_i \frac{\mathbf{p}_i}{2m_i} + \mathcal{U}(\mathbf{r}^N) + \sum_j \frac{Q_j \xi_j^2}{2} + g k_\mathrm{B} T s_1 + \sum_{j=2}^{M} k_\mathrm{B} T s_j
$$

Thermostats Comparison



Thermostats Table

	tune	cont.	L/G	correct	ergodic	cons. q.	determ.
Velocity rescaling			G		?		X
Andersen	X		L	X	X		
Berendsen	X	X	G		?		X
Nosé-Hoover	X	X	L/G	X		X	X
Nosé-Hoover chains	X	X	L/G	X	X	X	X
Langevin	X	X	L	X	X	X	
Stochastic velocity rescaling	X	X	L/G	X	X	X	

Other Ensembles

N, V, E	microcanonical	inner energy U
N, V, T	canonical	Helmholtz free energy F
		$F = U - TS$
N, P, H	isobaric-isoenthalpic	enthalpy H
		$H = U + pV$
N, P, T	isobaric-isothermal	Gibbs free energy G
		$G = U - TS + pV = H - TS$

How to simulate an isobaric ensemble?

- We can manipulate the volume of the system.
- This emulates an external piston used to keep the pressure constant.
- Instantaneous pressure $p(t)$ fluctuates, forces on the container
- Averaging $p(t)$ gives the observable internal pressure.

Pressure: Virial Theorem

From equipartition theorem

$$\langle H_{\mathrm{kin}} \rangle = \left\langle \sum_i \frac{p_i^2}{2m_i} \right\rangle = \frac{1}{2} \sum_i \left\langle r_i \frac{\partial U(r)}{\partial r_i} \right\rangle$$

The total potential $U(r)$ is coming from in internal potential $V_{int}(r)$ and and a wall potential $W(r)$.

$$\sum_i \left\langle r_i \frac{\partial W(r)}{\partial r_i} \right\rangle = p \int df \cdot r = p \int dV(divr) = 3pV$$

The virial theorem is then

$$3pV = 2\langle H_{\mathrm{kin}} \rangle - \sum_i \left\langle r_i \frac{\partial V(r)}{\partial r_i} \right\rangle$$

Isobaric-Isoenthalpic Ensemble

The system exchange work with an external surface: Ω is a variable
Cubic box

$$\mathbf{h} = \begin{pmatrix} \Omega^{\frac{1}{3}} & 0 & 0 \\ 0 & \Omega^{\frac{1}{3}} & 0 \\ 0 & 0 & \Omega^{\frac{1}{3}} \end{pmatrix}$$

Scaled coordinates

$$\mathbf{r}_i = \Omega^{1/3}\mathbf{s}_i$$

Lagrangian in terms of scaled coordinates

$$\mathcal{L} = \frac{1}{2}\sum_i m_i \Omega^{\frac{2}{3}} \dot{\mathbf{s}}_i^2 - \mathcal{U}(\mathbf{s}^N, \Omega)$$

External applied pressure P

$$-P\Delta\Omega$$

work on the system to balance P: the internal pressure fluctuates

Barostat

Correct the internal pressure by scaling the voluem, i.e. the inter-particle distances

The system is coupled to a barostat

Isobaric-Isoenthalpic ensemble NPH

$$\mathcal{P}(\mathbf{r}^N, \mathbf{p}^N, \Omega) \propto \delta(C - \mathcal{K}(\mathbf{p}^N) - \mathcal{U}(\mathbf{r}^N, \Omega) - \Pi\Omega)$$

where Π is the internal pressure
Isobaric-Isothermal ensemble NPT

$$\mathcal{P}(\mathbf{r}^N, \mathbf{p}^N, \Omega) \propto \exp\left(-\frac{\mathcal{K}(\mathbf{p}^N) + \mathcal{U}(\mathbf{r}^N, \Omega) + \Pi\Omega}{k_\mathrm{B} T}\right)$$

Berendsen Barostat

Deviation from desired pressure

$$\frac{dP(t)}{dt} = \frac{P_0 - P(t)}{\tau_P} \qquad \tau_P\text{relaxation time}$$

Rescale volume by η the volume, $\eta^{1/3}$ coordinates

$$\eta(t) = 1 - \frac{\delta t}{\tau_P}\gamma(P_0 - P(t))$$

γ is the compressibility.

Asotropic or anisotropic, τ_P strength of coupling,
not correct ensemble distribution

Andersen Barostat

Associate a kinetic energy to the variable volume

$$\mathcal{L} = \frac{1}{2}\sum_i m_i \Omega^{\frac{2}{3}} \dot{\mathbf{s}}_i^2 + \frac{1}{2} W \dot{\Omega}^2 - \mathcal{U}(\mathbf{s}^N, \Omega) - P\Omega$$

The coupling mimics the action of a piston of mass W (strength of coupling)
Equations of motion

$$\ddot{\mathbf{s}}_i = -\frac{1}{m_i \Omega^{\frac{2}{3}}} \frac{\partial U}{\partial \mathbf{s}_i} - \frac{2}{3} \frac{\dot{\Omega}}{\Omega} \mathbf{s}_i$$

$$W\ddot{\Omega} = \frac{1}{3\Omega}\left(\Omega^{\frac{2}{3}} \sum_i m_i \dot{\mathbf{s}}_i^2 - \sum_i \mathbf{s}_i \frac{\partial \mathcal{U}}{\partial \mathbf{s}_i} \right) - P = P_{\text{int}} - P$$

Feedback loop to adjust the simulation volume: internal pressure fluctuates around the external pressure

Correct Distribution

Conserved Hamiltonian of the $6N + 2$ dimensional system

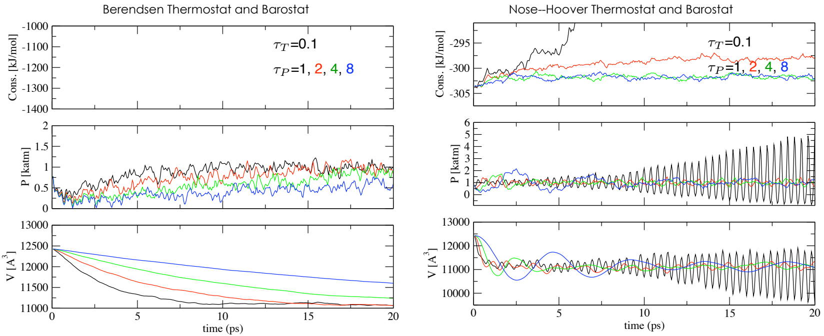
$$\mathcal{H} = \frac{1}{2} \sum_i \frac{\boldsymbol{\pi}_i^2}{m_i \Omega^{\frac{2}{3}}} + \frac{\Pi^2}{2W} + \mathcal{U}(\mathbf{s}^N, \Omega) + P\Omega$$

low W will result in rapid box size oscillations
large W slow adjustmen of the volume
Partition function

$$\Delta(N, P, \mathcal{H}) = \int d\Omega \int d\Pi \int d\boldsymbol{\pi}^N \int \mathbf{s}^N \delta \left(H + \frac{\Pi^2}{2W} - \mathcal{H} \right)$$

trajectories consistent with the NPH ensemble
The combination with one of the constant temperature methods allows the simulation in the NPT ensemble.

Experiment with barostat and thermostat


Berendsen Thermostat and Barostat

Nose--Hoover Thermostat and Barostat

Multiple Time Step Method (r-RESPA)

M.E. Tuckerman et al. JCP, **97** 1990 (1992)

Equation of motion

$$\dot{r} = \frac{p}{m} \qquad \dot{p} = F_{\mathrm{fast}} + F_{\mathrm{slow}}$$

Liouville operator

$$iL = \left(\frac{p}{m}\frac{\partial}{\partial r} + F_{\mathrm{fast}}\frac{\partial}{\partial p} \right) + F_{\mathrm{slow}}\frac{\partial}{\partial p} + iL_{\mathrm{ref}} + iL_{\mathrm{slow}}$$

Integrator

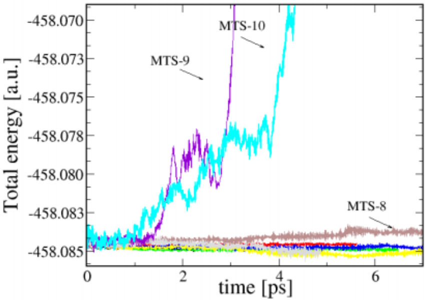
$$\begin{aligned} x(\delta t) &= \exp(iL\delta t)x(0) \\ &= \exp(iL_{\mathrm{slow}}\delta t/2)\left[\exp(iL_{\mathrm{ref}}\delta t)\right]^{n}\exp(iL_{\mathrm{slow}}\delta t/2) + \mathcal{O}(\delta t^{3}) \end{aligned}$$

MTS in CP2K

- Two FORCE_EVAL sections defining methods
- Example DFT with hybrid functionals

$$F_{\mathrm{ref}} = F_{\mathrm{GGA}}$$

$$F_{\mathrm{slow}} = F_{\mathrm{hybrid}}$$

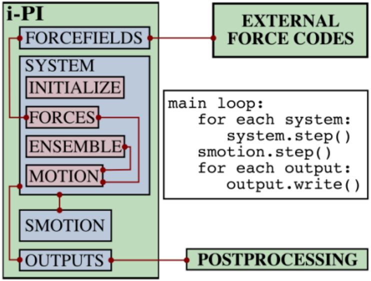


PLUMED: MD Driver and Enhanced Sampling

- Definition of many Collective Variables (CV)
- Analysis features for MD and MetaDynamics
- Bias methods for enhanced sampling MD
- Logarithmic Mean Force Dynamics Method
- Experiment Directed Simulation Methods

i-Pi: External MD Driver

Kapil et al., Comp. Phys. Comm. 236, 214-223 (2018)



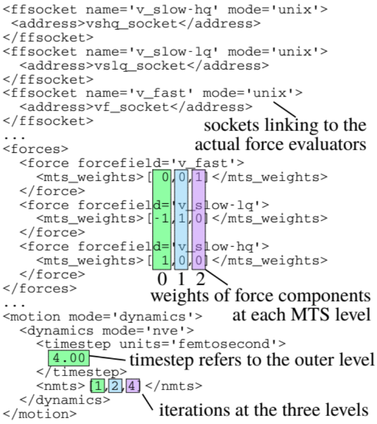
```
main loop:
    for each system:
        system.step()
    smotion.step()
    for each output:
        output.write()
```

Driver software communicating with force engines (e.g. CP2K) through socket interfaces

i-Pi: Features

- MD and Path-Integral MD (PIMD) for NVE, NVT, NPT
- Ring Polymer Contraction (RPC) MD and Centroid MD
- Generalized Langevin Equation (GLE) thermostats
- PI+GLE and PIGLET
- Ring Polymer Instantons
- Thermodynamic Integration, Geomerty optimization, Saddle point search
- Harmonic vibrations
- Multiple Time Step algorithms
- Metadynamics (interface to PLUMED)
- Replica Exchange MD
- Second Generation CP like integration

i-PI: Multiple Time Steps



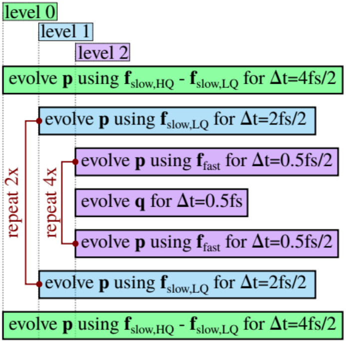
```
<ffsocket name='v_slow-hq' mode='unix'>
 <address>vshq_socket</address>
</ffsocket>
<ffsocket name='v_slow-lq' mode='unix'>
  <address>vslq_socket</address>
</ffsocket>
<ffsocket name='v_fast' mode='unix'>
  <address>vf_socket</address>
</ffsocket>
...
<forces>
  <force forcefield='v_fast'>
    <mts_weights>[ 0, 0, 1] </mts_weights>
  </force>
  <force forcefield='v_slow-lq'>
    <mts_weights>[-1, 1, 0] </mts_weights>
  </force>
  <force forcefield='v_slow-hq'>
    <mts_weights>[ 1, 0, 0] </mts_weights>
  </force>
</forces>
...
<motion mode='dynamics'>
  <dynamics mode='nve'>
    <timestep units='femtosecond'>
     4.00
    </timestep>
    <nmts>[1,2,4]</nmts>
  </dynamics>
</motion>
```

sockets linking to the actual force evaluators

weights of force components at each MTS level

timestep refers to the outer level

iterations at the three levels



level 0

level 1

level 2

evolve **p** using $\mathbf{f}_{\mathrm{slow,HQ}} - \mathbf{f}_{\mathrm{slow,LQ}}$ for Δt=4fs/2

repeat 2x

evolve **p** using $\mathbf{f}_{\mathrm{slow,LQ}}$ for Δt=2fs/2

repeat 4x

evolve **p** using $\mathbf{f}_{\mathrm{fast}}$ for Δt=0.5fs/2

evolve **q** for Δt=0.5fs

evolve **p** using $\mathbf{f}_{\mathrm{fast}}$ for Δt=0.5fs/2

evolve **p** using $\mathbf{f}_{\mathrm{slow,LQ}}$ for Δt=2fs/2

evolve **p** using $\mathbf{f}_{\mathrm{slow,HQ}} - \mathbf{f}_{\mathrm{slow,LQ}}$ for Δt=4fs/2

Summary: BOMD in CP2K

- Born–Oppenheimer MD with ASPC(3) is default in CP2K
- SCF convergence criteria depends on system
  $10^{-5} - 10^{-6}$ is a reasonable starting guess
- Best used together with OT and the
  FULL_SINGLE_INVERSE preconditioner, for large system
  an iterative update of the preconditioner should be used
  PRECOND_SOLVER INVERSE_UPDATE
- Langevin dynamics is an option
  needs some special care, analysis of forces