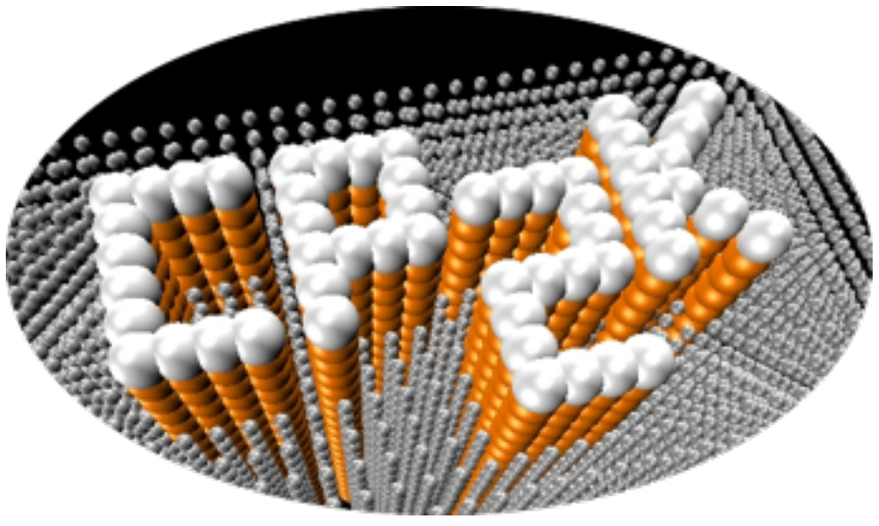
CP2K UK WORKSHOP 2014

27-28 August 2014, London UK

Ab initio Molecular Dynamics

Marcella Iannuzzi

Department of Chemistry, University of Zurich



http://www.cp2k.org

Outline

- MD and statistical mechanics
- Equations of motion
- Numerical integration
- Born Oppenheimer MD
- Car Parrinello MD
- Constraints

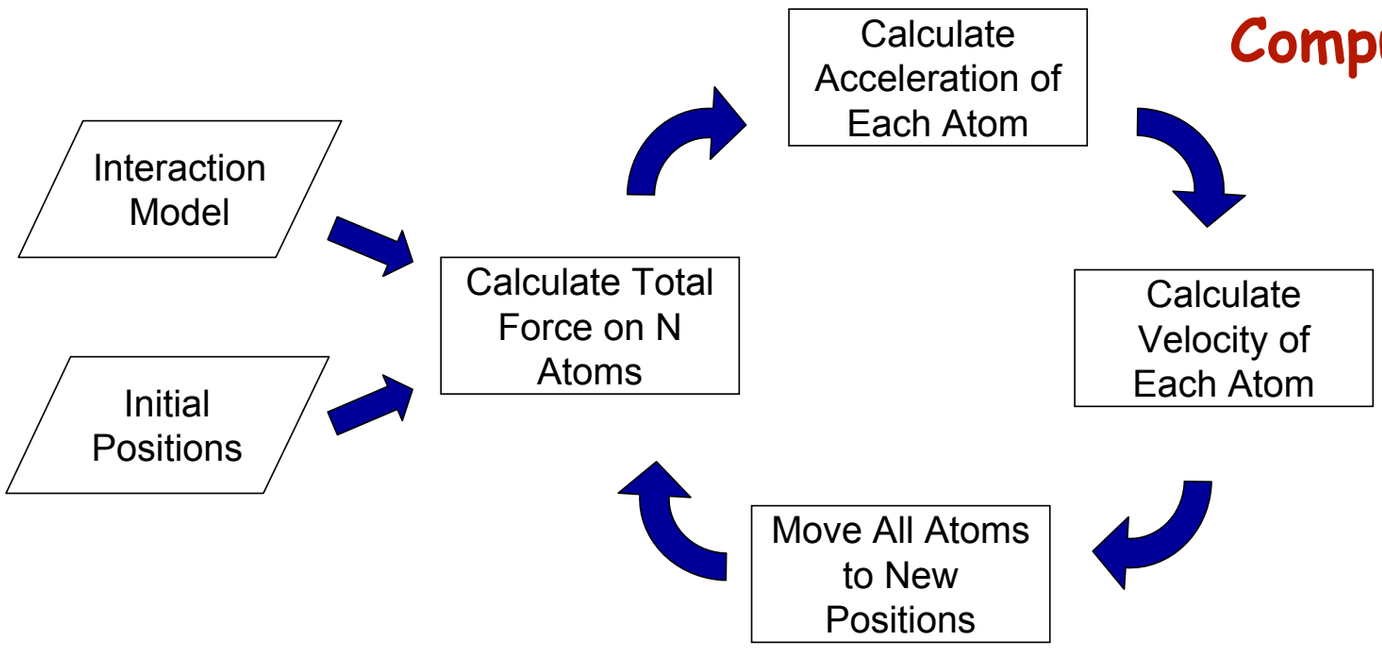
Molecular Dynamics

Given the initial conditions ({RI};{PI}), an interacting potential (H),
and the thermodynamic conditions (T,V,P)

generate deterministic trajectories that sample the phase space according to statistical mechanics

M.P. Allen and D.J. Tildesley, Computer Simulations of Liquids, Claredon Press, Oxford, 1987
D. Frenkel and B. Smit, Understanding Molecular Simulations, Computer Sciences Series, Academic Press, 2002

Computational Experiment



Density of state

In statistical mechanics an observable **macrostate** is defined by thermodynamic parameters

(NVE) MICROCANONICAL ENSEMBLE

its **probability** is given by its degeneracy:

number of **microstates** (configurations) consistent with the **macrostate**

$$\Omega(N,V,E) \propto \int d\mathbf{r}_1 ... \int d\mathbf{r}_N \int d\mathbf{p}_1 ... \int d\mathbf{p}_N \; \delta\left(\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N) - E\right)$$

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

all accessible states are equally probable

A(r,p) ENSAMBLE AVERAGE

$$\langle A \rangle_{\mathrm{NVE}} = \frac{\int d\mathbf{r}_1 ... \int d\mathbf{r}_N \int d\mathbf{p}_1 ... \int d\mathbf{p}_N \; \delta\left(\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N) - E\right) \; A(\mathbf{r}^N,\mathbf{p}^N)}{\int d\mathbf{r}_1 ... \int d\mathbf{r}_N \int d\mathbf{p}_1 ... \int d\mathbf{p}_N \; \delta\left(\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N) - E\right)}$$

Canonical Partition Function

System of particles in equilibrium with a **thermal bath**

(NVT) CANONICAL ENSEMBLE

The Laplace transform of the density of state

$$Q(N,V,T) = \int \exp(-\beta E)\Omega(N,V,E)dE$$

Probability of the macrostate at a given T

Boltzmann's distribution

$$Q(N,V,T) = \frac{1}{N!h^{3N}} \int \exp\left[-\beta\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)\right] d\mathbf{r}^N d\mathbf{p}^N = \frac{1}{\Lambda(\beta)^{3N}N!} Z(N,V,T)$$

one dimensional integral over E replaced by configurational integral
analytic kinetic part is integrated out

$$Z(N,V,T) = \int e^{-\beta\mathcal{U}(\mathbf{r})} d\mathbf{r}$$

configurational partition function

Time Average

Molecular Dynamics: propagation of particle trajectories solving eq. of motion

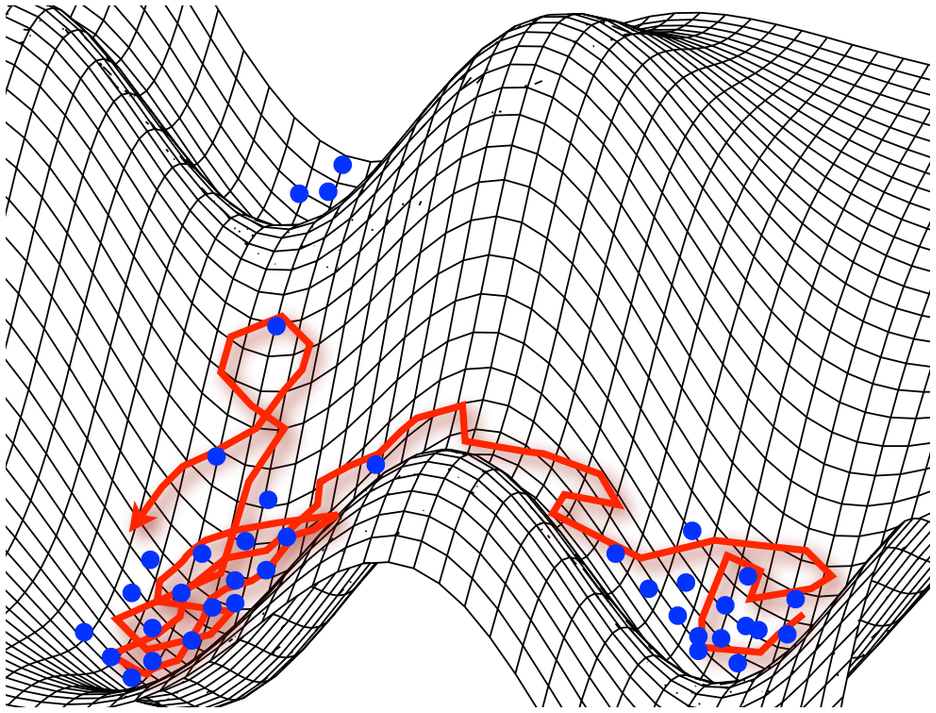
observables from averaging over sufficiently long time

$$\bar{A} = \lim_{t\to\infty} \frac{1}{t} \int_0^t d\tau A(\mathbf{r}^N(\tau), \mathbf{p}^N(\tau))$$

Ergodic Hypothesis: in the limit of infinitely long trajectories, the average does not depend on the initial conditions

$$\bar{A} = \frac{\sum_{\text{init. cond.}} \lim_{t\to\infty} \frac{1}{t} \int_0^t d\tau A_i(\mathbf{r}^N(\tau), \mathbf{p}^N(\tau))}{\Omega(N, V, E)}$$

Sampling the Phase Space via MD



Ergodic Hypothesis

$$\bar{A} = \langle A \rangle = \frac{\sum A_i}{\Omega} = \lim_{t \to \infty} \frac{1}{t} \int_0^t d\tau \langle A(\tau)$$

- propagation over time
- based on models for molecular-scale interactions
- **Newton's equations** of motions: forces ore needed
- Numerical integration: **Time Step**
- Observations are made based on the evolution

Total Energy

Model potential depending only on the particles' coordinates

no external sources of forces

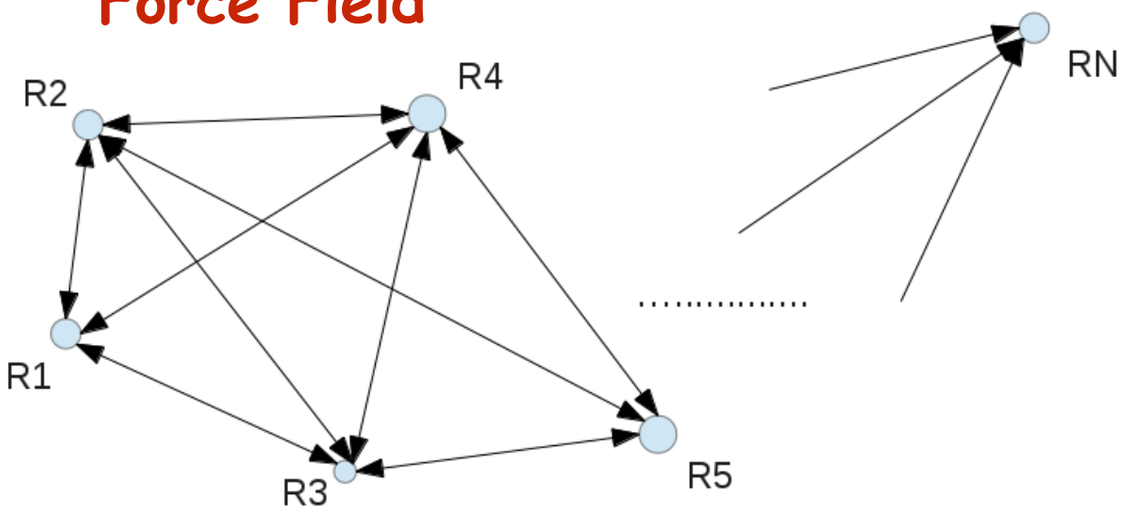
$$\mathcal{H}(\mathbf{R}^N, \mathbf{P}^N) = \sum_{I=1}^{N} \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + U(\mathbf{R}^N)$$

Kinetic **Potential**

N, V, and E are the constants of motion in microcanonical ensemble

$$\frac{d\mathcal{H}}{dt} = \sum_I \left[\frac{\partial \mathcal{H}}{\partial \mathbf{R}_I} \dot{\mathbf{R}}_I + \frac{\partial \mathcal{H}}{\partial \mathbf{P}_I} \dot{\mathbf{P}}_I \right] = 0$$

Force Field



Newton's EOM

$$\mathbf{F}_I = M_I \ddot{\mathbf{R}}_I = -\frac{dU(\mathbf{R}^N)}{d\mathbf{R}_I}$$

Equations of Motion (EOM)

Set of classical particles in a potential

Hamilton EOM

$$\dot{\mathbf{P}}_I = -\frac{\partial \mathcal{H}}{\partial \mathbf{R}_I}$$

$$\dot{\mathbf{R}}_I = -\frac{\partial \mathcal{H}}{\partial \mathbf{P}_I}$$

$$\dot{\mathbf{P}}_I = -\frac{\partial \mathcal{H}}{\partial \mathbf{R}_I} = -\frac{\partial U(\{\mathbf{R}_I\})}{\partial \mathbf{R_I}} = F_I(\{\mathbf{R}_I\})$$

Conservation of energy:

$$\frac{dE}{dt} = \frac{d\mathcal{H}}{dt} = 0$$

Lagrange EOM

$$\mathcal{L}(\{\mathbf{R}_I\}, \{\dot{\mathbf{R}}_I\}) = \sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 - U(\{\mathbf{R}_I\})$$

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{R}}_I} = \frac{\partial \mathcal{L}}{\partial \mathbf{R}_I}$$

Numerical Integration

The system is propagated in the phase space

Discretisation of time

$$t_0 = 0 \ ; t_1 = \Delta t \ ; \ldots \ ; t_N = N \cdot \Delta t \ ; \ldots$$

The fast time scales of the system determine the choice of time step.
A good integrator algorithm:

- Accurate for long time steps: higher order derivatives, more memory storage required
- Minimum number of force calculations
- Long time energy conservation and stability in spite of small perturbations
- Approximation of the true trajectory: Lyapunov instability
- Short time reversibility: invariant for $t \longrightarrow -t$

Source of Errors

- Type of integrator

  time reversible, predictor-corrector, symplectic

- Time Step

  short time accuracy

- Consistency of forces and energy

  e.g. cutoffs leading to non smooth energy surfaces

- Accuracy of forces

  e.g. convergence of iterative optimisation (SCF, constraints)

Velocity Verlet

$$\mathbf{R}(t+\Delta t) = \mathbf{R}(t) + \mathbf{V}(t)\Delta t + \frac{\mathbf{F}(t)}{2M}\Delta t^2 \qquad \mathbf{V}(t+\Delta t) = \mathbf{V}(t) + \frac{\mathbf{F}(t+\Delta t) + \mathbf{F}(t)}{2M}\Delta t$$

- 1 force evaluation, 3 storage vectors
- Positions and velocities available at equal time
- Contains error of order Δt^4
- Time reversible
- Conserves volume in phase space: symplectic
- Long time stability
- Simple adaptation for multiple time steps
- Simple adaptation for constraints

Implementation of Velocity-Verlet

Half Kick -- Drift -- Half Kick

Given **R(:),V(:),F(:)** at time-step itime

Update **V(:)** by half time-step

V(:) := V(:) + dt/(2*M)*F(:)

Then update **R(:)** by the entire time-step

R(:) := R(:) + dt*V(:)

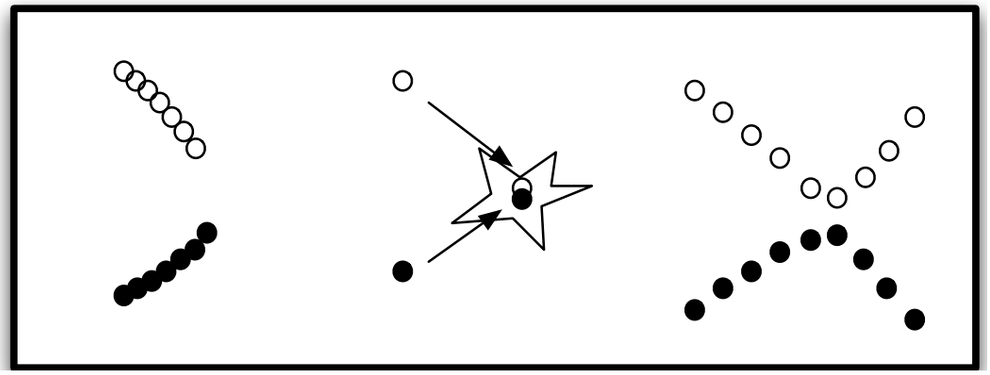
Compute the new **F(:)** by using the updated **R(:)**

Finalize the update of **V(:)** by the second half time-step

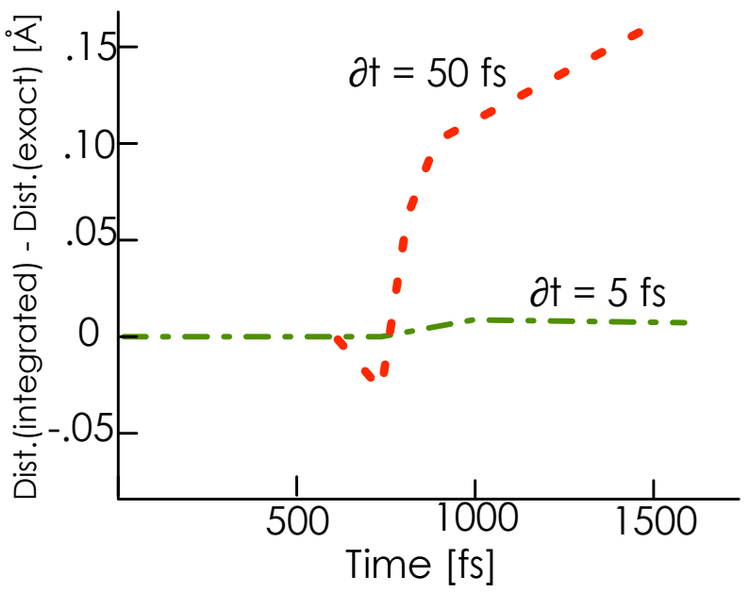
V(:) := V(:) + dt/(2*M)*F(:)

Choice of the Time Step

Compromise between **efficiency** and **reliability**



Pairwise potential
∂t = 0.5fs, 50fs, 5fs

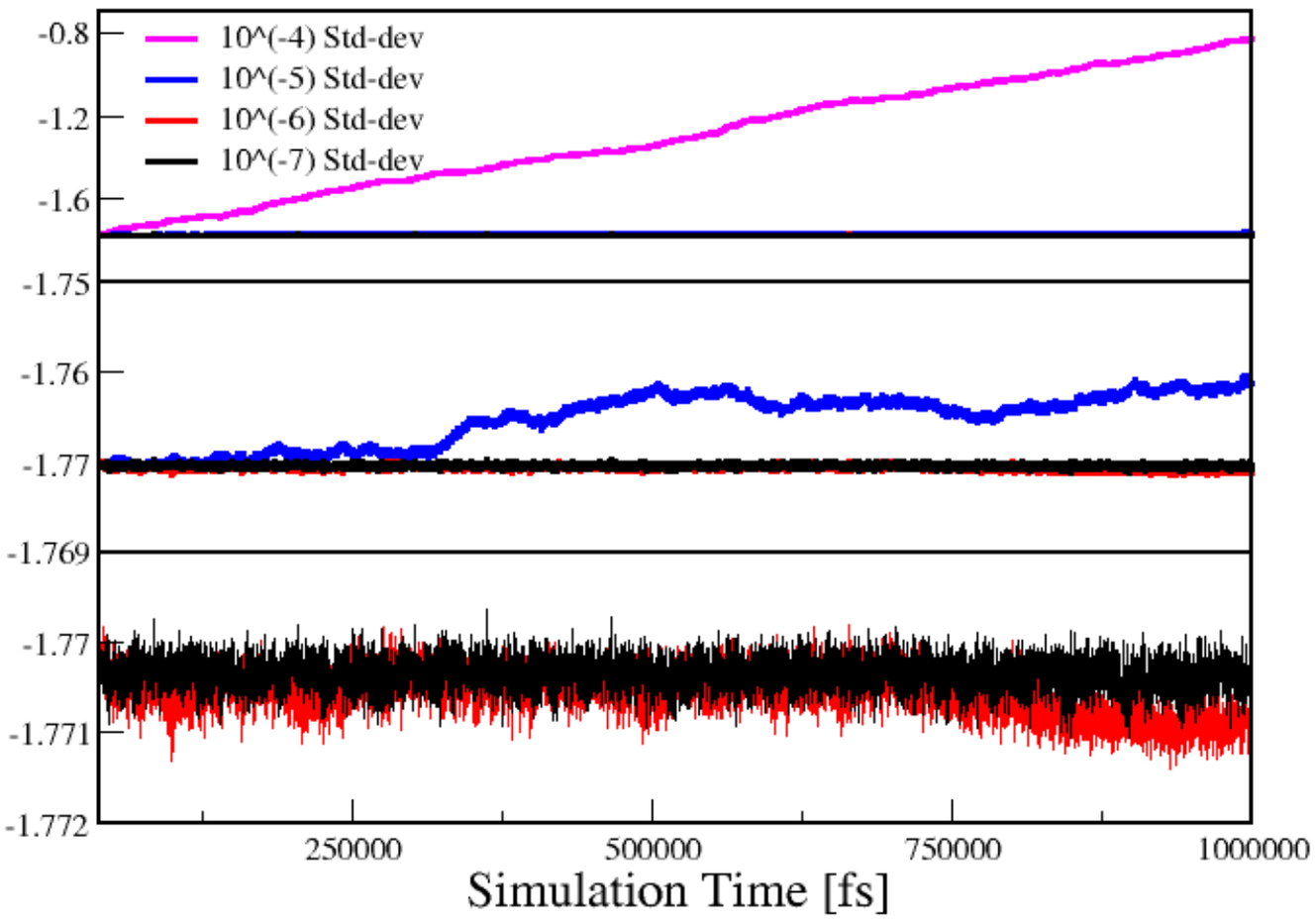


Difference between exact and numerical trajectory

Time-step at least 10 times smaller than the fastest period of motion

Test on Required Accuracy of Forces

Classical FF, 64 H2O at 330 K: TIP3P(flexible), SPME



Stability depends on 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

Temperature

Equipartition of energy over DOF

$$\left\langle \frac{1}{2} M_I V_{I\alpha}^2 \right\rangle = \frac{1}{2} k_{\mathrm{B}} T$$

Number of DOF

$$N_f = 3N - 3$$

Kinetic energy

$$K = N_f \frac{1}{2} k_{\mathrm{B}} T$$

Instantaneous temperature

$$T(t) = \frac{\sum_I \sum_\alpha M_I V_{I\alpha}(t)^2}{N_f k_{\mathrm{B}}}$$

$$T_{\mathrm{NVE}} = \frac{1}{K} \sum_{k=1}^{K} T(t_k)$$

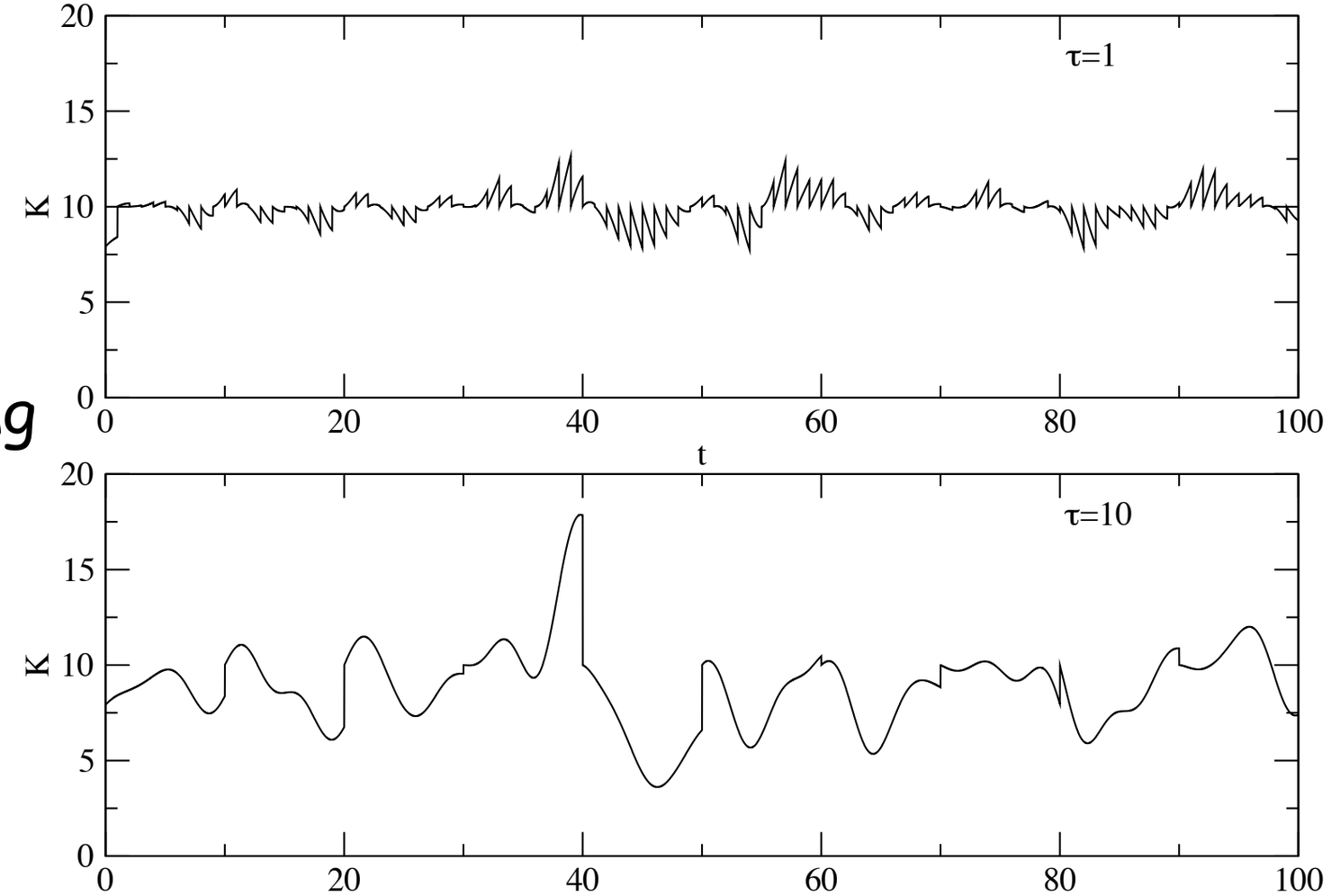
Velocity Rescaling

Keep the temperature at the desired value T_0
multiply all velocities by the same factor

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

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

- Easy
- Tot. energy not conserved
- No correct ensemble sampling
- Used for fast equilibration
- Not time rev.



Extended System

To extend to other ensembles, the Lagrangian equations of motion need to be reformulated: the system moves in a different phase space

impose control of specific thermodynamic variables

Canonical Ensemble

Generate the correct Boltzmann distribution by coupling with a heat bath

$$\mathcal{P}_i = \frac{\exp(-E_i/k_B T)}{\sum_j \exp(-E_j/k_B T)}$$

Andersen Thermostat

- Coupling by stochastic impulsive forces: collision with bath
- Randomly selected particle (local)
- Strength: frequency of collisions
- New velocity drawn from M-B dist.

$$\mathcal{P}(p) = \left(\frac{\beta}{2\pi m}\right)^{3/2} \exp[-p^2\beta/2m] \qquad p_{\text{new}} = \sqrt{m k_B T_{\text{bath}}}\, R$$

- Mixing Newtonian dyn. with stochastic
- Enhanced decay of vel. correlation (no transport)

Nose Thermostat

Heat bath is an integral part of the system:

$$\mathcal{L}_{Nose} = \sum_I \frac{M_I}{2} s^2 \dot{\mathbf{R}}_I^2 - \mathcal{U}(\{\mathbf{R}_I\}) + \frac{Q}{2}\dot{s}^2 - g k_B T \ln s$$

Additional variables s, ds/dt, and coupling strength parameter Q

Conjugate momenta

$$\mathbf{P}_I = \frac{\partial \mathcal{L}_{\text{Nose}}}{\partial \dot{\mathbf{R}}_I} = M_I s^2 \dot{\mathbf{R}}_I \qquad p_s = \frac{\partial \mathcal{L}_{\text{Nose}}}{\partial \dot{s}} = Q\dot{s}$$

EOM

$$\ddot{\mathbf{R}}_I = \frac{\mathbf{F}_I}{M_I} - \dot{s}\dot{\mathbf{R}}_I \qquad \ddot{s} = \frac{1}{Q}\left[\sum_I M_I \dot{\mathbf{R}}_I^2 - g k_B T\right]$$

Time-rescaling (dt' = s dt); P'=P/s
Microcanonical in extended system

Nose-Hoover

Differentiation in real time t' : $\frac{1}{s}\frac{d}{dt'}$

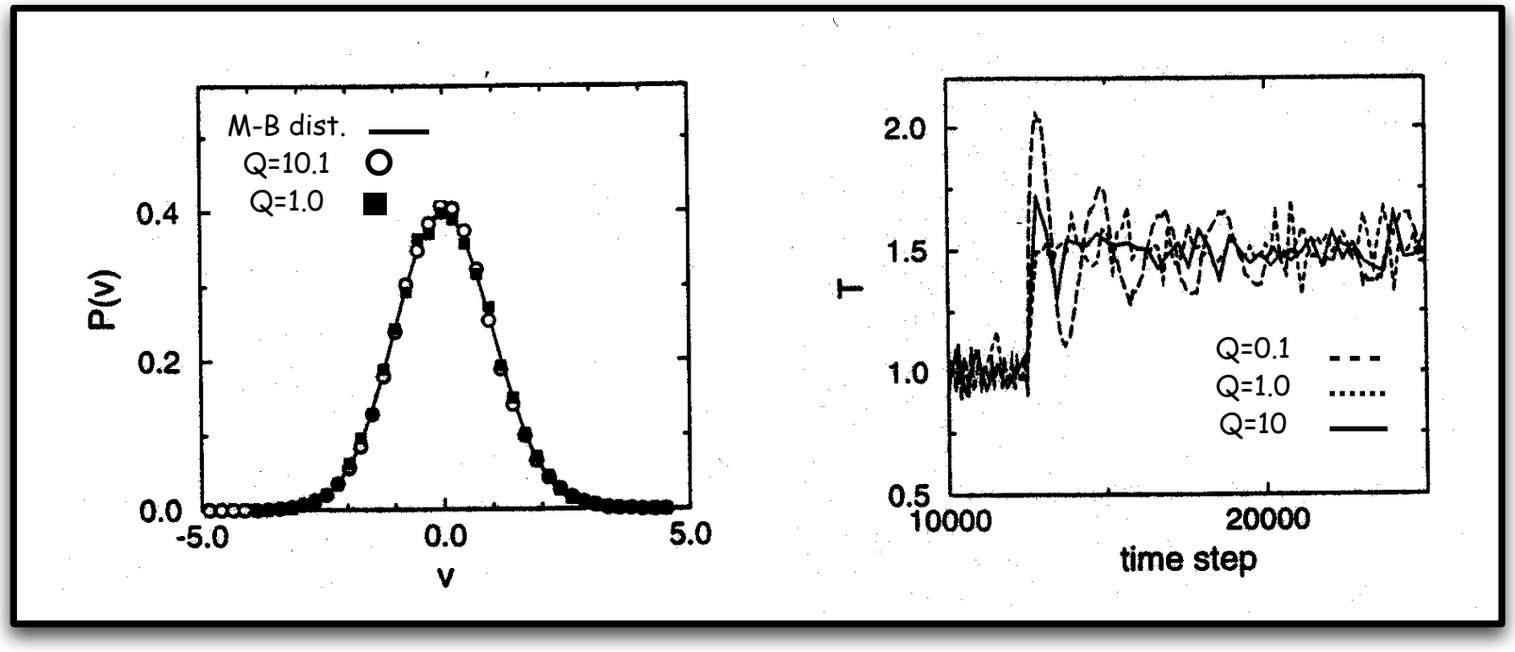
Friction parameter and relaxation time $\xi = \frac{sp_s}{Q}$ $\nu_T = \sqrt{\frac{N_f k_{\rm B} T}{Q}}$

$$\dot{\xi} = \nu_T^2 \left[\frac{\sum_I M_I V_I^2}{N_f k_{\rm B} T} - 1 \right] = \nu_T^2 \left[\frac{\mathcal{T}(t)}{T} - 1 \right]$$

$$\begin{cases} \mathcal{T} < T & \text{smaller friction} \\ \mathcal{T} > T & \text{larger friction} \end{cases}$$

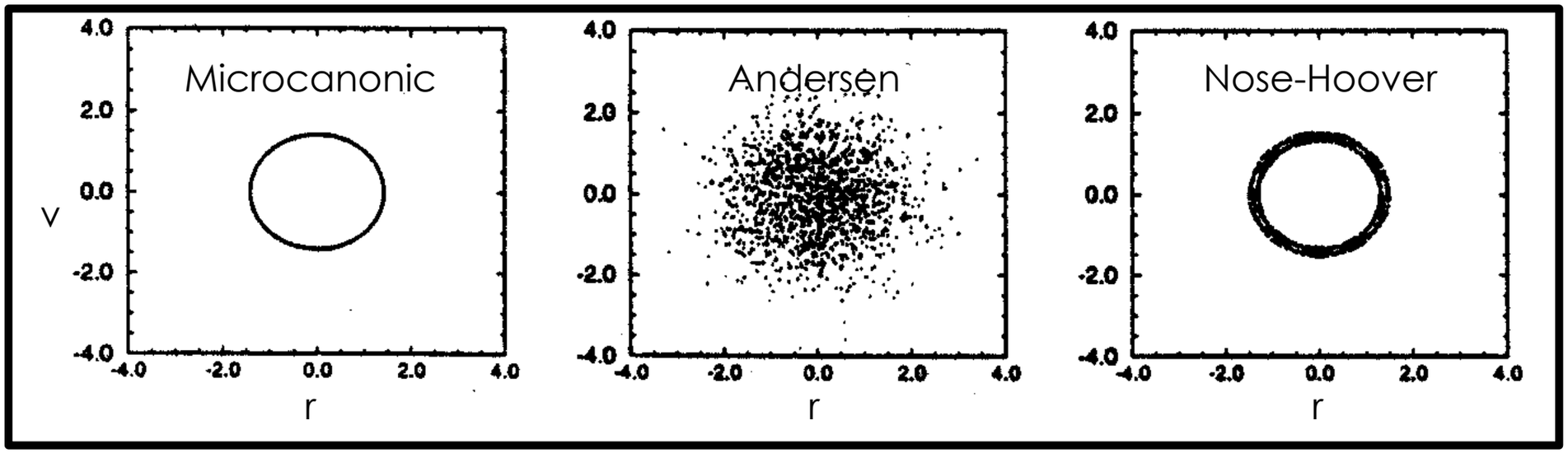
Q large: slow equilibration

Q small: high freq. T

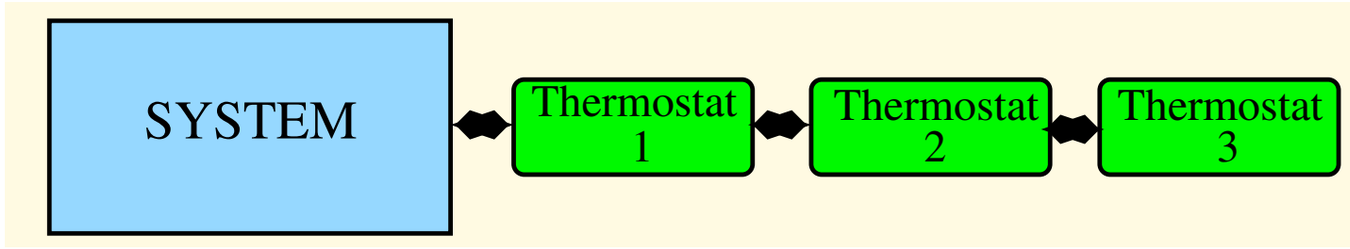


Ergodicity Problems

Simple 1D harmonic oscillator



Ergodicity problems can be solved by implementing a chain of thermostats



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

CP2K MD Input

```
&GLOBAL
   PROJECT pname
   RUN_TYPE md
&END GLOBAL
&MOTION
 &MD
   ENSEMBLE NVT
   STEPS 10
   TIMESTEP [fs] 0.5
   TEMPERATURE [K] 300.0
   &THERMOSTAT
    REGION MASSIVE
    &NOSE
     LENGTH 1
     TIMECON 10
    &END NOSE
  &END THERMOSTAT
&END MD

&PRINT
 &RESTART
   ADD_LAST NUMERIC
   BACKUP_COPIES 1
    &EACH
     MD 5
    &END EACH
&END RESTART
&STRESS
 &EACH
   MD 1
 &END EACH
&END STRESS
&TRAJECTORY
 &EACH
   MD 1
 &END EACH
&END TRAJECTORY
&VELOCITIES
 &EACH
   MD 1
 &END EACH
&END VELOCITIES
 &END PRINT
&END MOTION
```

Born-Oppenheimer Approximation

$$\mathcal{H}(\mathbf{R}^N, \mathbf{x}^{N_{el}})\Psi(\mathbf{R}^N, \mathbf{x}^{N_{el}}) = E\Psi(\mathbf{R}^N, \mathbf{x}^{N_{el}})$$

Separation of time scales

$$\frac{\omega_e}{\omega_n} \sim \sqrt{\frac{m_n}{m_e}} \sim 100$$

- The total wave function is factorised

$$\Psi(\mathbf{R}^N, \mathbf{x}^{N_{el}}) = \Sigma(\mathbf{R}^N)\Phi_{\mathbf{R}}(\mathbf{x}^{N_{el}})$$

- Electronic structure optimised in the potential of fixed ions

$$\mathcal{H}_{el}\Phi_{\mathbf{R}}(\mathbf{x}^{N_{el}}) = U(\mathbf{R}^N)\Phi_{\mathbf{R}}(\mathbf{x}^{N_{el}})$$

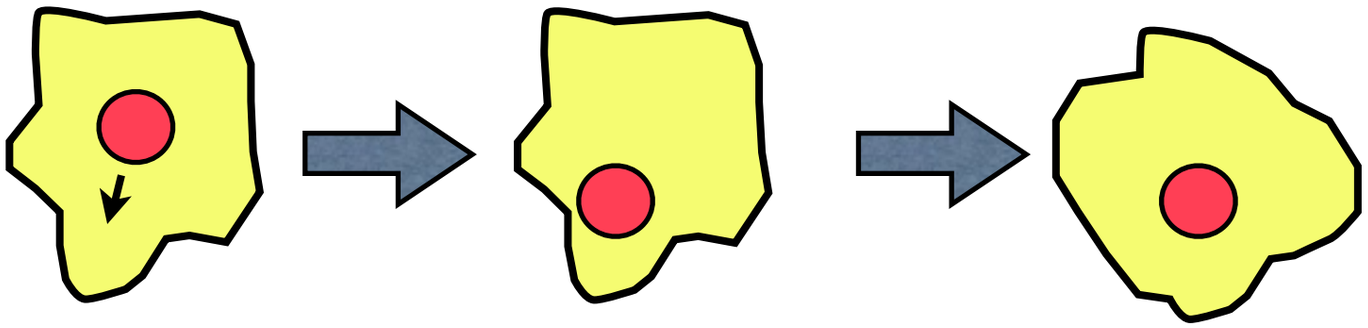
- Electronic and nuclear solutions are decoupled

$$\mathcal{H}_n\Sigma(\mathbf{R}^N) = E\Sigma(\mathbf{R}^N)$$

- Nuclei are in most of the case treated as classical particles

Electronic and nuclear motions are adiabatically separated

Integrator for BOMD



t0 t1 t1 + SCF

Born-Oppenheimer MD
Adiabatic approx.
Semiclassical approx.

$$\mathcal{L}_{\mathrm{BO}}\left(\{\mathbf{R}_I\},\{\dot{\mathbf{R}}_I\}\right)=\sum_{I=1}^{N}\frac{1}{2}M_I\dot{\mathbf{R}}_I^2-\min_{\{\phi_i\}}E_{\mathrm{KS}}\left(\{\phi_i\},\{\mathbf{R}_I\}\right)$$

No electron Dynamics

Classical equations of motion

$$M_I\ddot{\mathbf{R}}_I(t)=-\nabla_I\left[\min_{\{\phi_i\}}E_{\mathrm{KS}}\left(\{\phi_i\},\{\mathbf{R}_I(t)\}\right)\right]$$

Integration step determined by the time scale
of the nuclear dynamics: **~femtoseconds**

$$\mathbf{F}_I=-\left[\langle\Psi_0|\nabla_I\mathcal{H}_{\mathrm{KS}}|\Psi_0\rangle+\langle\nabla_I\Psi_0|\mathcal{H}_{\mathrm{KS}}|\Psi_0\rangle+\langle\Psi_0|\mathcal{H}_{\mathrm{KS}}|\nabla_I\Psi_0\rangle\right]$$

Forces in BO-MD

For exact eigenstates and complete basis sets, the contributions from variations of the wavefunction vanish exactly

$$\mathbf{F}_I^{HFT} = -\langle \Psi_0 | \nabla_I \mathcal{H}_{\mathrm{KS}} | \Psi_0 \rangle$$

Hellman-Feynman

$$\nabla_I \phi_i = \sum_\nu \left(\nabla_I c_{i\nu}\right)\ \varphi_\nu(\mathbf{r};\{\mathbf{R}_I\}) + \sum_\nu c_{i\nu}\ \left(\nabla_I \varphi_\nu(\mathbf{r};\{\mathbf{R}_I\})\right)$$

implicit dependence of the expansion coefficients: not exact self-consistency NSC

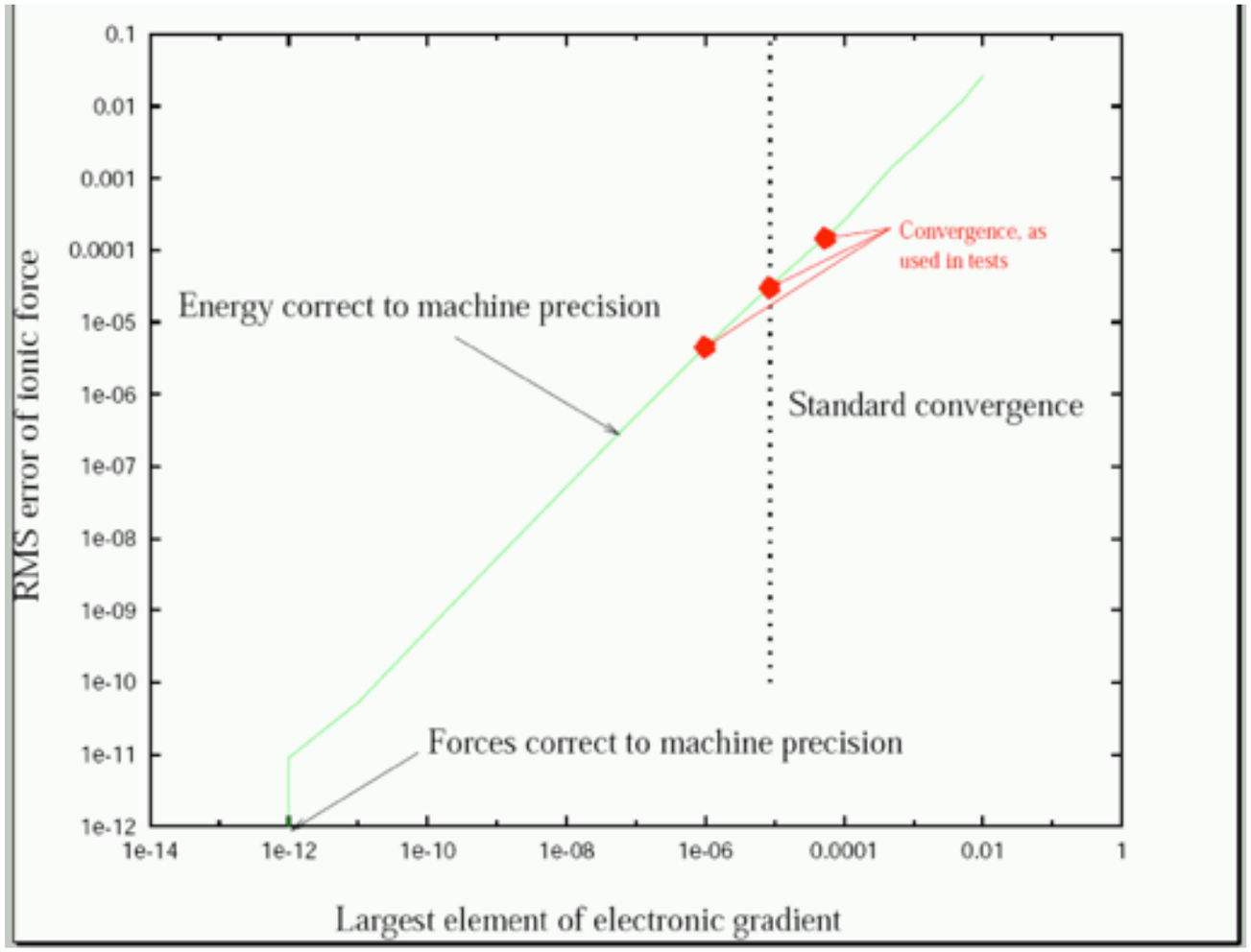
explicit dependence of the basis function IBS

$$\mathbf{F}_I^{\mathrm{NSC}} = -\int d\mathbf{r}\ (\nabla_I n)\left(V^{\mathrm{SCF}} - V^{\mathrm{NSC}}\right)$$

$$\mathbf{F}_I^{\mathrm{IBS}} = -\sum_{i\nu\mu}\left(\left\langle \nabla_I \varphi_\nu \left| H_{\mathrm{e}}^{\mathrm{NSC}} - \epsilon_i \right| \varphi_\mu \right\rangle + \left\langle \varphi_\nu \left| H_{\mathrm{e}}^{\mathrm{NSC}} - \epsilon_i \right| \nabla_I \varphi_\mu \right\rangle\right)$$

P. Bendt et al., PRL, 50, 1684 (1983)
G.P. Srivastava et al., 36, 463 (1987)

Stability in BOMD



64 H2O, 330 K, 1gr/cm^3
TZV2P, PBE, GTH, 280 Ry
0.5fs step

Reference: 1ps, SCF 10^{-10}
Unbiased initial guess

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

Error in Forces $\Longleftrightarrow$ MD Stability

Energy $\min_{\psi} E_{KS}[\{\psi\}]$ error 2nd order in $\delta\psi$

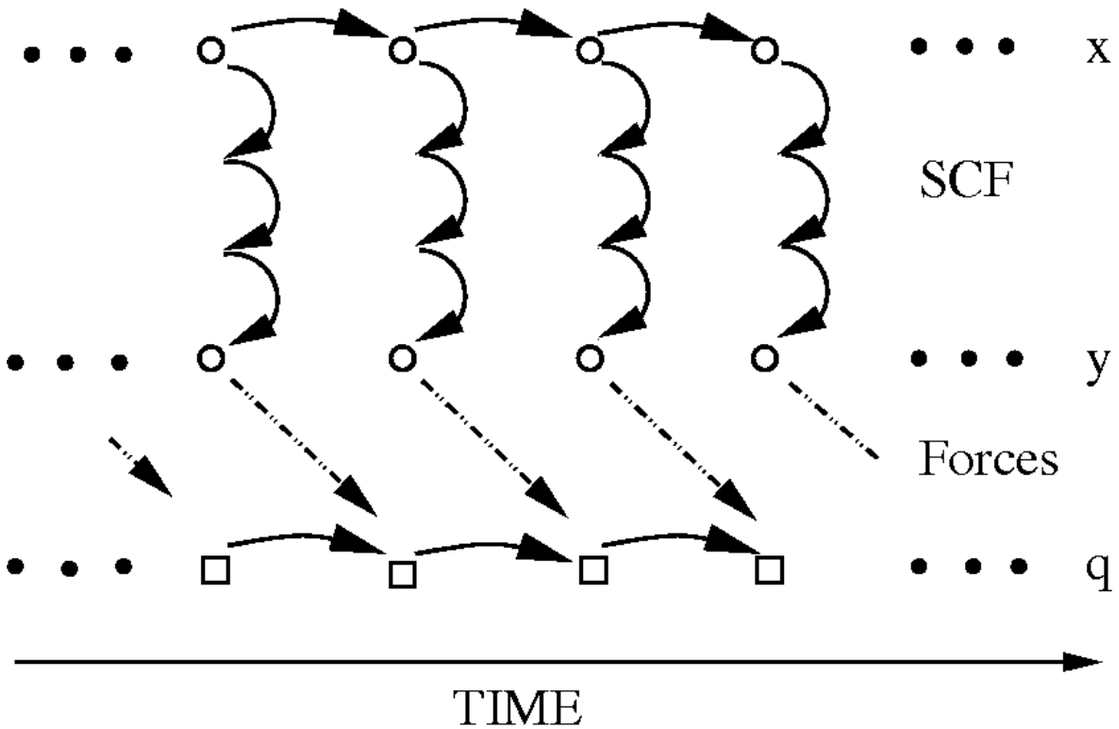
Forces $dE_{KS}[\{\psi\}]/d\mathbf{R}$ error 1st order in $\delta\psi$

Generalised Lagrangian

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

$\mathbf{y} = F(\mathbf{q},\mathbf{x})$ wfn optimisation

$G(||\mathbf{x}-\mathbf{y}||)$ wfn retention potential



equations of motion

$$M\ddot{\mathbf{q}} = -\frac{\partial E}{\partial \mathbf{q}} - \frac{\partial E}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{q}} + k\mu\frac{\partial G}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{q}}$$

$$\mu\ddot{\mathbf{x}} = -\frac{\partial E}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{x}} + k\mu\left[\frac{\partial G}{\partial \mathbf{x}} + \frac{\partial G}{\partial \mathbf{y}}\frac{\partial F}{\partial \mathbf{x}}\right]$$

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

Car Parrinello MD

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

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

Extended system approach: add KS orbitals as explicit classical variables

$$\mathcal{L}_{\rm CP} = \underbrace{\sum_I \frac{1}{2}M_I\dot{\mathbf{R}}_I^2 + \sum_i \mu\left\langle\dot{\phi}_i\middle|\dot{\phi}_i\right\rangle}_{\text{kinetic energy}} - \underbrace{E_{\rm KS}[\{\phi_i\},\{\mathbf{R}_I\}]}_{\text{potential energy}} + \underbrace{\sum_{i,j}\Lambda_{ij}\left(\langle\phi_i|\phi_j\rangle - \delta_{ij}\right)}_{\text{constraint}}$$

$$M_I\ddot{\mathbf{R}}_I(t) = -\frac{\partial E_{\rm KS}}{\partial \mathbf{R}_I} + \sum_{ij}\Lambda_{ij}\frac{\partial}{\partial \mathbf{R}_I}\langle\phi_i|\phi_j\rangle$$

$$\mu\ddot{\phi}_i(t) = -\frac{\delta E_{\rm KS}}{\delta\langle\phi_i|} + \sum_j \Lambda_{ij}|\phi_j\rangle$$

Fictitious mass

Constraint forces

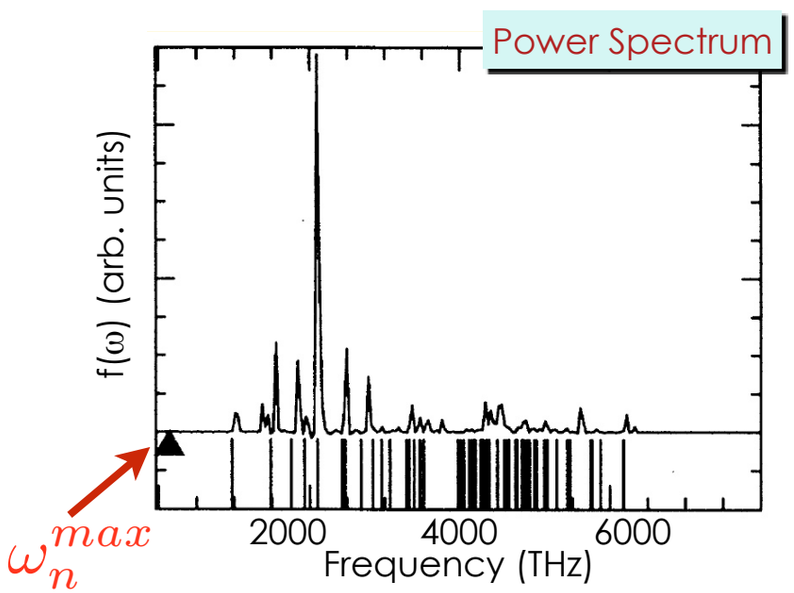
R. Car and M. Parrinello, PRL, 55, 2471 (1985)

Properties of CPMD

- Two systems propagate simultaneously: fictitious dynamics of orbitals
- Stable propagator: exact forces
- Hot nuclei and cold orbitals: el. close to BO surface

$$T_{el} \propto \sum_i \mu \langle \dot{\phi}_i | \dot{\phi}_i \rangle$$

- Decoupled subsystems
- Fictitious dyn. averages out over nuclear time scales



$$\omega_e^{\min} \propto \left(\frac{E_{\text{gap}}}{\mu}\right)^{1/2}$$

$$\delta t^{\max} \propto [\omega_e^{\max}]^{-1} \propto \left(\frac{\mu}{E_{\text{cut}}}\right)^{1/2}$$

μ ~ 500-1000 a.u
δt ~ 0.12-0.24 fs

BOMD with incomplete convergence

BOMD

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

Decoupled Lagrangian dynamics

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

Incomplete convergence

$$\mathbf{y} \approx \mathrm{Min}_{\mathbf{x}} E(\mathbf{q}, \mathbf{x})$$

EOM

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

$$\ddot{\mathbf{x}} = k \left[\frac{\partial G}{\partial \mathbf{x}} + \frac{\partial G}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{x}} \right]$$

coupling via SCF error is neglected

Extrapolation Methods

Integration of electronic DOF has to be

accurate: good wavefunction guess gives improved efficiency

stable: do not destroy time-reversibility of nuclear trajectory

- Unbiased guess

$$\mathbf{C}_{\text{init}} = \mathbf{C}_0$$

- Combinations of previous wavefunctions: unstable

- Extrapolation of the density matrix: **PS** methods, $O(MN^2)$

$$\mathbf{C}(t_n) = \sum_{m=1}^{K} (-1)^{m+1} \begin{bmatrix} K \\ j \end{bmatrix} \mathbf{C}(t_{n-m})\mathbf{C}^{\dagger}(t_{n-m})\mathbf{S}(t_{n-m})\mathbf{C}(t_{n-1})$$

- Always stable predictor corrector (ASPC) based on OT minimisation

Kolafa J., J. Comput. Chem. 25, 335 (2004)

ASPC

Projection onto the occupied subspace

$$\mathbf{C}^p(t_n) = \sum_{m=1}^{K} (-1)^{m+1} m \frac{\binom{2K}{K-m}}{\binom{2K-2}{K-1}} \mathbf{C}(t_{m-m}) \mathbf{C}^\dagger(t_{n-m}) \mathbf{S}(t_{n-m}) \mathbf{C}(t_{n-1})$$

Reversibility $O(\Delta t^{(2K-1)})$

iterate

The corrector step minimises the error and reduces the deviation from ground state

$$\mathbf{C}(t_n) = \omega \mathrm{MIN}[\mathbf{C}^p(t_n)] + (1-\omega)\mathbf{C}^p(t_n) \qquad \omega = \frac{K}{2K-1}$$

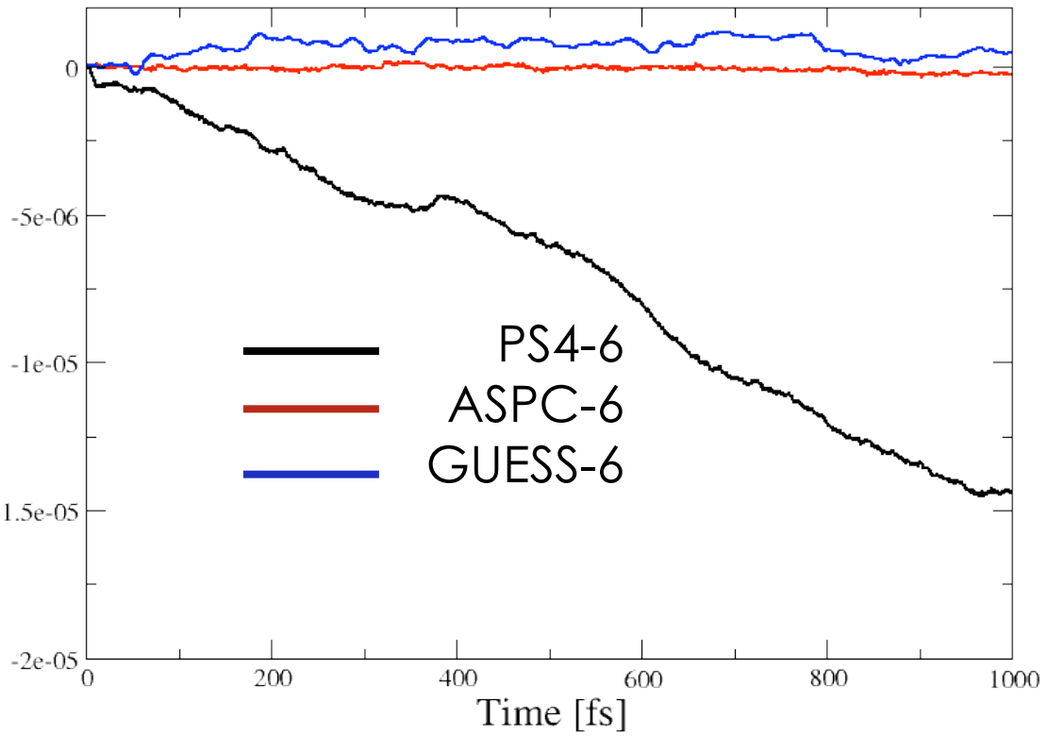
Preconditioned OT minimisation step: large move

Kolafa J., J. Comput. Chem. 25, 335 (2004),
VandeVondele et al, JCP, 118, 4365 (2003)

Efficiency and Drift

64 H2O, 330 K, 1gr/cm^3

Method	ϵ_{SCF}	Iterations	Drift (μHartree/ns)
Guess	10^{-06}	14.38	253
PS4	10^{-10}	14.95	–
PS4	10^{-08}	8.05	-195
PS4	10^{-07}	6.47	-3441
PS4	10^{-06}	5.22	-7186
PS4	10^{-05}	4.60	52771
ASPC	10^{-06}	5.01	-115
ASPC	10^{-05}	3.02	-2758
ASPC	10^{-04}	1.62	-1059843
ASPC	10^{-02}	1.03	-13219651



Gear not time reversible

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

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

DFT section with ASPC

```
&DFT
  BASIS_SET_FILE_NAME ../BASIS_SET
  POTENTIAL_FILE_NAME ../GTH_POTENTIAL
  &MGRID
    CUTOFF 300
  &END MGRID
  &QS
    EPS_DEFAULT 1.0E-12
    EXTRAPOLATION ASPC
    EXTRAPOLATION_ORDER 4
  &END QS
  &SCF
    EPS_SCF 1.0E-5
    SCF_GUESS ATOMIC
    &OT ON
      MINIMIZER DIIS
    &END OT
  &END SCF
  &XC
    &XC_FUNCTIONAL BLYP
    &END XC_FUNCTIONAL
  &END XC
&END DFT
```

```
&SUBSYS
  &CELL
    ABC 8.0 8.0 8.0
  &END CELL
  &COORD
  O   0.000000    0.000000   -0.065587
  H   0.000000   -0.757136    0.520545
  H   0.000000    0.757136    0.520545
  &END COORD
  &KIND H
    BASIS_SET DZVP-GTH-BLYP
    POTENTIAL GTH-PADE-q1
  &END KIND
  &KIND O
    BASIS_SET DZVP-GTH-BLYP
    POTENTIAL GTH-PADE-q6
  &END KIND
&END SUBSYS
```

Forces in Approximated BOMD

exact $$\mathbf{F}_{\mathrm{BO}}(\mathbf{R}) = \mathbf{F}_{\mathrm{HF}}(\mathbf{R}) + \mathbf{F}_{\mathrm{Pulay}}(\mathbf{R}) + \mathbf{F}_{\mathrm{nsc}}(\mathbf{R})$$

approximated $$\tilde{\mathbf{F}}(\mathbf{R}) = \mathbf{F}_{\mathrm{HF}}(\mathbf{R}) + \mathbf{F}_{\mathrm{Pulay}}(\mathbf{R})$$

Now assume $$\tilde{\mathbf{F}}(\mathbf{R}) + \mathbf{F}_{\mathrm{nsc}}(\mathbf{R}) = \mathbf{F}_{\mathrm{BO}}(\mathbf{R}) - \gamma_D \dot{\mathbf{R}}$$ friction

Langevin dynamics to correct the error (dissipative drift)

$$M_I \ddot{\mathbf{R}}_I = \mathbf{F}_I^{\mathrm{BO}} - (\gamma_D + \gamma_L)\dot{\mathbf{R}}_I + \Xi_I^D + \Xi_I^L$$

Gaussian random noise guarantees accurate Boltzmann sampling

fluctuation dissipation theorem

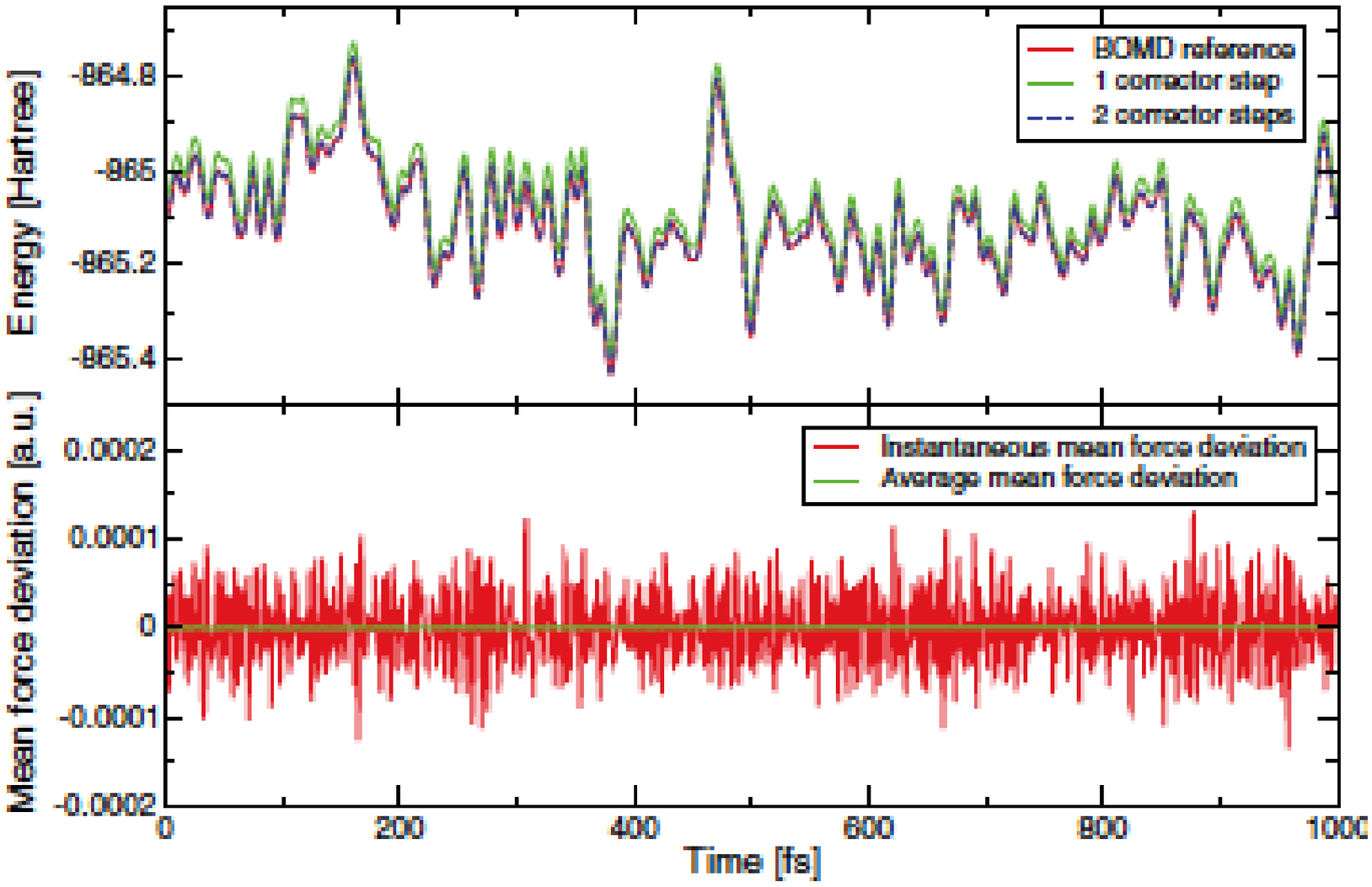
$$\langle (\Xi_I^D(0) + \Xi_I^L(0))(\Xi_I^D(t) + \Xi_I^L(t)) \rangle = 6(\gamma_D + \gamma_L) M_I k_{\mathrm{B}} T \delta t$$

given $$\left\langle \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 \right\rangle = \frac{3}{2} k_{\mathrm{B}} T$$ this determines the friction

T. Kuehne et al, PRL , 98(6), 066401 (2007)

Validation

liquid silica, 24 SiO_2 at 3500 K
Time step: Δt = 1 fs
γ_D 10^{-4} fs^{-1} , K=4

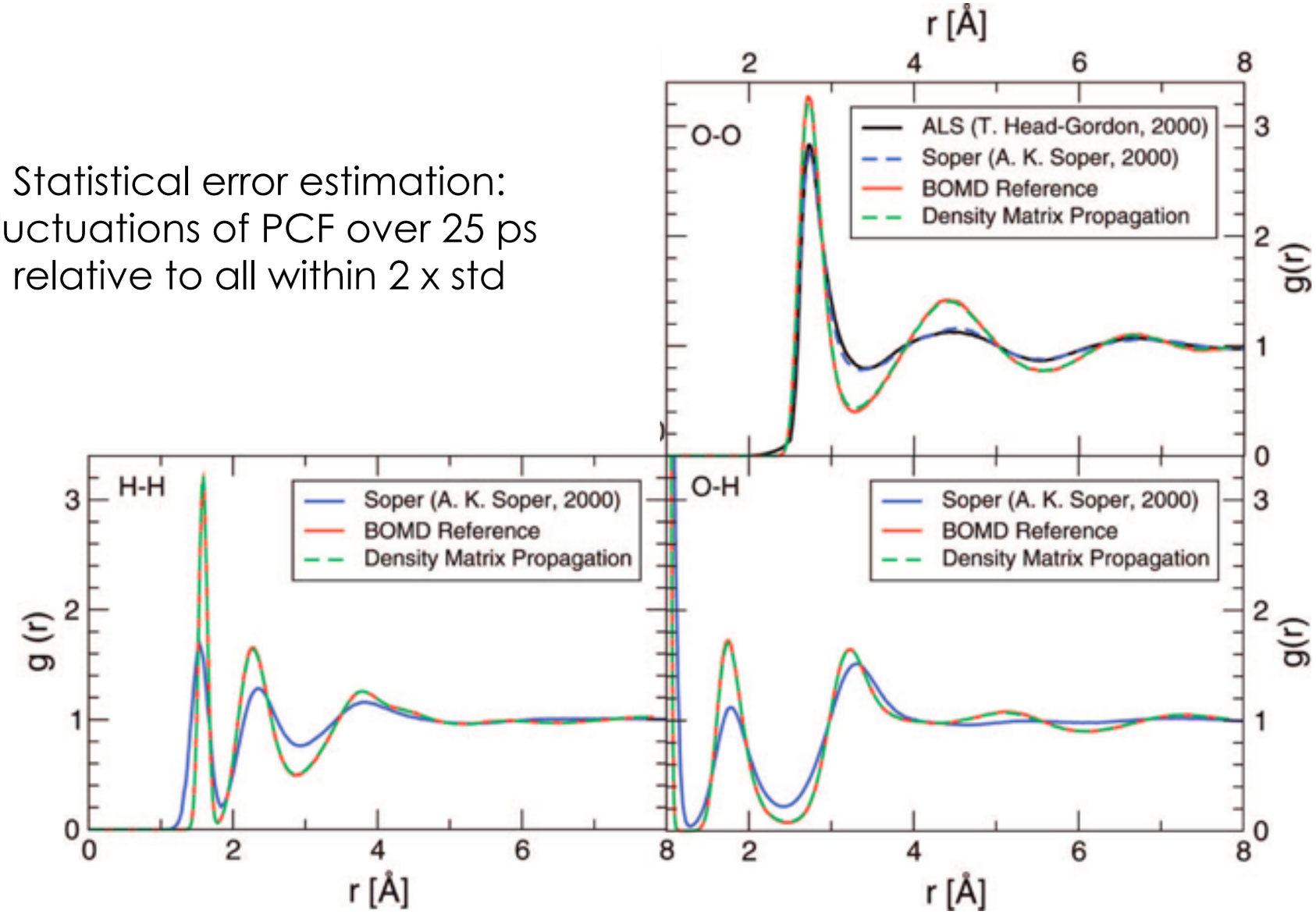


Bonds are swiftly broken and formed
Worst case scenario for P propagation, as the electronic density is rapidly varying

Liquid Water

PBE, TZV2P, 320 Ry
300 K, Δt = 0.5 fs, **25+250 ps** trajectories
γ_D 8.65 10^{-5} fs^{-1} , K=7 => 1 PC step, deviation 10^{-5} au

Statistical error estimation:
fluctuations of PCF over 25 ps
relative to all within 2 x std



T. Kuehne et al, J. Chem Theory Compt , 5(2), 235 (2008)

Geometrical Constraints

Implicit functions of the degrees of freedom of the system

$$\sigma(\{\mathbf{R}_I\}, \mathbf{h}, \Psi) = 0 \qquad \dot{\sigma}(\{\mathbf{R}_I\}, \mathbf{h}, \Psi) = 0$$

- To freeze fast degrees of freedom and increase the time step: e.g., intra-molecular bonds by distance constraints
- To explore only a sub-region of the conformational space
- As reaction coordinates : constrained dynamics and thermodynamic integration
- To prevent specific events or reactions

Lagrange formulation for simple constraints, functions of $\mathbf{R}_I$

$$\mathcal{L}'(\{\mathbf{R}_I\}, \{\mathbf{P}_I\}) = \mathcal{L}(\{\mathbf{R}_I\}, \{\mathbf{P}_I\}) - \sum_\alpha \lambda_\alpha \sigma(\{\mathbf{R}_I\})$$

The Lagrange multipliers ensure that positions and velocities satisfy the constraints

Shake-Rattle algorithm

Modified velocity Verlet scheme by additional constraint forces

- First update of velocities (first half step) and positions

$$V_I' = V_I(t) + \frac{\delta t}{2M_I} F_I(t) \qquad R_I' = R_I(t) + \delta t V_I'$$

- Positions' correction by **constraint forces**

$$R_I(t+\delta t) = R_I' + \frac{\delta t^2}{2M_I} g_I^{(p)}(t)$$

- Calculation of the new forces $F_I(t+\delta t)$

- Update of velocity (second half step)

- Correction by the constraint forces

$$V_I(t+\delta t) = V_I' + \frac{\delta t}{2M_I}[F_I(t+\delta t) + g_I^{(v)}(t+\delta t)]$$

Constraint Forces

$$g_I^{(p)}(t) = -\sum_\alpha \lambda_\alpha^{(p)} \frac{\partial \sigma_\alpha(\{R_I\})}{\partial R_I}$$

$$e_\alpha(\{\lambda_\gamma\}) = -\sum_\beta \mathbf{J}^{-1}_{\alpha\beta}\, \sigma_\beta(\{\lambda_\gamma\}) \qquad \mathbf{J}_{\alpha\beta} = \frac{\partial \sigma_\alpha(\{\lambda_\gamma\})}{\partial \lambda_\beta}$$

Set of non-linear equations **solved iteratively**

$$g_I^{(v)}(t) = -\sum_\alpha \lambda_\alpha^{(v)} \frac{\partial \sigma_\alpha(\{R_I\})}{\partial R_I}$$

$$\sum_I \frac{\partial \sigma_\alpha}{\partial \mathbf{R}_I} \mathbf{V}_I = \sum_I \frac{\partial \sigma_\alpha}{\partial \mathbf{R}_I} \cdot \mathbf{V}_I' + \sum_\beta \left(\sum_I \frac{\delta t^2}{2M_I} \frac{\partial \sigma_\alpha}{\partial \mathbf{R}_I} \frac{\partial \sigma_\beta}{\partial \mathbf{R}_I} \right) \lambda_\beta^v = 0$$

H.C. Andersen, J. Comp. Phys., **52**, 24 (1983)

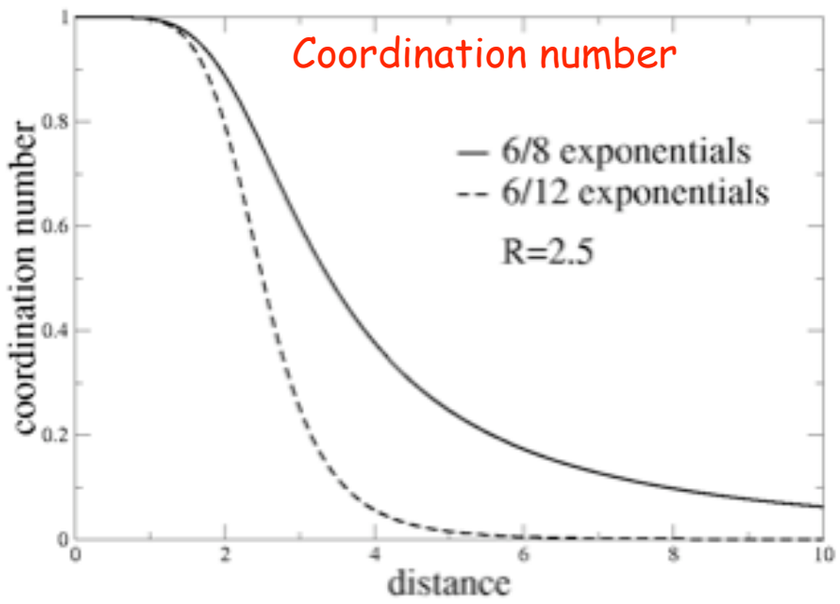
Some simple collective variables

Derivable function of the degrees of freedom to which a given value can be assigned

- Distance $|\mathbf{R}_I - \mathbf{R}_J|$
- Angle $\theta(\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_k)$
- Dihedral $\Theta(\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_k, \mathbf{R}_L)$
- Difference of distances $|\mathbf{R}_I - \mathbf{R}_J| - |\mathbf{R}_J - \mathbf{R}_K|$

- Generalised coordination number

$$C_{L_1 L_2} = \frac{1}{N_{L_1}} \sum_{j=1}^{N_{L_1}} \left\{ \sum_{i=1}^{N_{L_2}} \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^n}{1 - \left(\frac{r_{ij}}{r_0}\right)^m} \right\}$$



- Generalised displacement

$$D^{[klm]}_{L_1 L_2} = \frac{1}{N_{L_1}} \sum_{i \in L_1} \mathbf{d}_i \cdot \hat{\mathbf{v}}_{[klm]} - \frac{1}{N_{L_2}} \sum_{j \in L_2} \mathbf{d}_j \cdot \hat{\mathbf{v}}_{[klm]}$$

CP2K input for Collective variables

In SUBSYS add one section per CV

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

```
&COLVAR
 &DISTANCE_FUNCTION
   ATOMS 4 6 6 1
   COEFFICIENT -1.00000
   # distance 1 = ( 4 - 6 )
   # distance 2 = ( 6 - 1 )
 &END DISTANCE_FUNCTION
&END COLVAR
```

```
&COLVAR
 &DISTANCE
   AXIS  X
   ATOMS 1 4
 &END DISTANCE
&END COLVAR
```

```
&COLVAR
 &RMSD
   &FRAME
    COORD_FILE_NAME planar.xyz
   &END
   &FRAME
    COORD_FILE_NAME  cage.xyz
   &END
   SUBSET_TYPE LIST
   ATOMS 1 3 5 6 8 9
   ALIGN_FRAMES  T
 &END RMSD
&END
```

Constraints and Restraints

In MOTION add one section per constraint

```
&CONSTRAINT
  &COLLECTIVE
    COLVAR 1
    INTERMOLECULAR
    TARGET       5.
    TARGET_GROWTH   1.1
    TARGET_LIMIT    10.
  &END COLLECTIVE
&END CONSTRAINT
```

```
&COLLECTIVE
  TARGET [deg] 0.0
  MOLECULE 1
  COLVAR 1
  &RESTRAINT
    K   [kcalmol] 4.90
  &END RESTRAINT
&END COLLECTIVE
```