



**The Abdus Salam
International Centre for Theoretical Physics**



2054-7

Structure and Dynamics of Hydrogen-Bonded Systems

26 - 27 October 2009

**Aqueous Water and Its Liquid-Vapor Surface from First-Principles by a
Car-Parrinello-like Approach**

Thomas D. KUEHNE
*Harvard University
Dept. of Physics, School of Eng.
17 Oxford Street, Cambridge
MA 02138
U.S.A.*

Aqueous Water and Its Liquid-Vapor Surface from First-Principles by a Car-Parrinello-like Approach

Thomas D. Kühne¹, Efthimios Kaxiras¹ and Yousung Jung²

¹Department of Physics and School of Engineering and Applied Sciences
Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA

²Graduate School of EEWS, KAIST, Yuseong-gu, Daejeon 305-701, Korea



Outline

- Briefly review the main ideas our simulation methodology
- Assessing observables that are well established
 - Partial pair-correlation function
 - Velocity autocorrelation function and its power spectrum
- Have a look on observables that are much less established
 - Translational self-diffusion and its convergence wrt system size
 - Determine the shear-viscosity from first-principle
 - Kinetics of hydrogen bonding in liquid water
- First results on the liquid-vapor surface
 - Hydrogen bond and surface definition
 - Hydrogen bonding at the interface
 - Orientational distribution of OH bonds and lone pair orbitals

The Main Conceptual Ideas

T. Kühne, M. Krack, F. Mohamed and M. Parrinello, Phys. Rev. Lett. **98**, 066401 (2007)

- BOMD allows for large timesteps, but incomplete wavefunction minimization leads to more or less pronounced sampling errors
- It is sufficient to approach the BO-surface by design of a CEIMD
- CP-BOMD, to fully bypass the SCF-cycle and iterative minimization by tolerating but correcting for sampling errors
- The DM is solely propagated within a predictor-corrector scheme
- With that a non-self-consistent functional is evaluated
- F_{NSC} and F_{NEF} are analytically solvable using DFPT
- For the short-term integration of the electrons accuracy is crucial
- Therefore the explicit, but non-symplectic ASPC integrator is used
- Non-symplectic integrators generally turn a conservative system into a dissipative one, leading to an exponential decay in energy

Langevin Equation

$$\begin{aligned}
 M_I \ddot{\mathbf{R}}_I &= \underbrace{\mathbf{F}_I^{BO} - M_I \dot{\mathbf{R}}_I (\gamma_D + \gamma_L)} + (\boldsymbol{\Xi}_I^D + \boldsymbol{\Xi}_I^L) \\
 &= \mathbf{F}_I^{PC} - \gamma_L M_I \dot{\mathbf{R}}_I + (\boldsymbol{\Xi}_I^D + \boldsymbol{\Xi}_I^L)
 \end{aligned}$$

- To guarantee an accurate sampling of the Boltzmann distribution, the noise has to obey the fluctuation dissipation theorem:

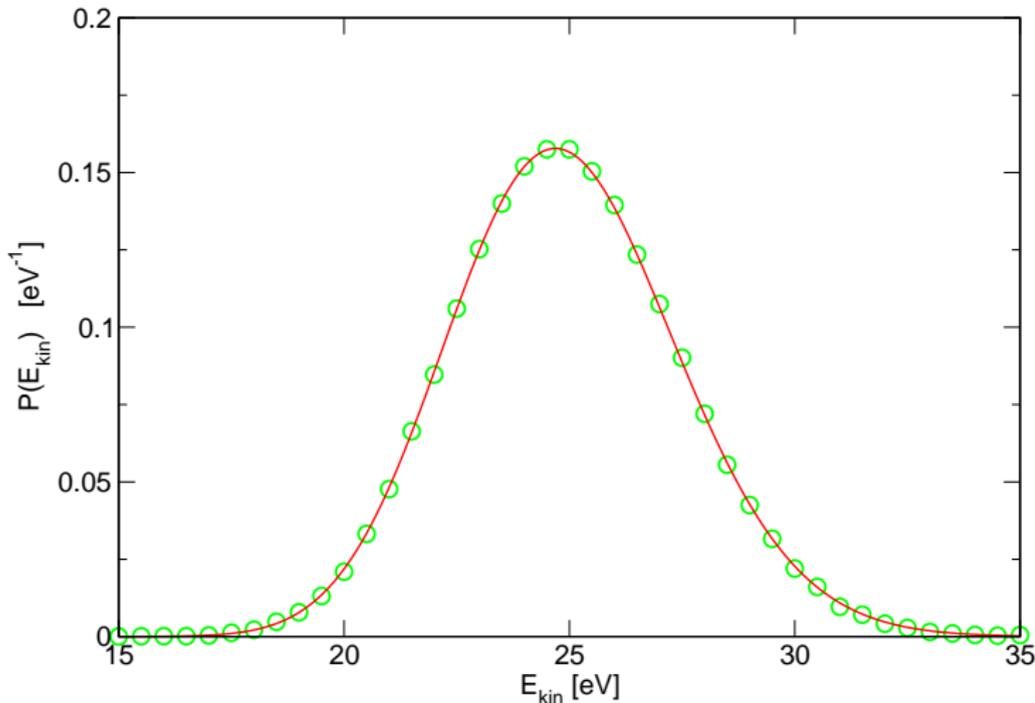
$$\langle \{ \boldsymbol{\Xi}_I^D(0) + \boldsymbol{\Xi}_I^L(0) \} \{ \boldsymbol{\Xi}_I^D(t) + \boldsymbol{\Xi}_I^L(t) \} \rangle = 2(\gamma_D + \gamma_L) M_I k_B T \delta(t)$$

- If the dissipation is indeed exponential decaying, we can do the substitution and still get a rigorously defined canonical ensemble
- γ_D , such that the equipartition theorem $\langle \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 \rangle = \frac{3}{2} k_B T$ holds
- To integrate the equations of motion under Brownian motion, the algorithm of Ricci & Ciccotti is used
- The white noise condition is satisfied by definition

Is our assumption justified?

Is the Kinetic Energy Distribution Maxwell distributed?

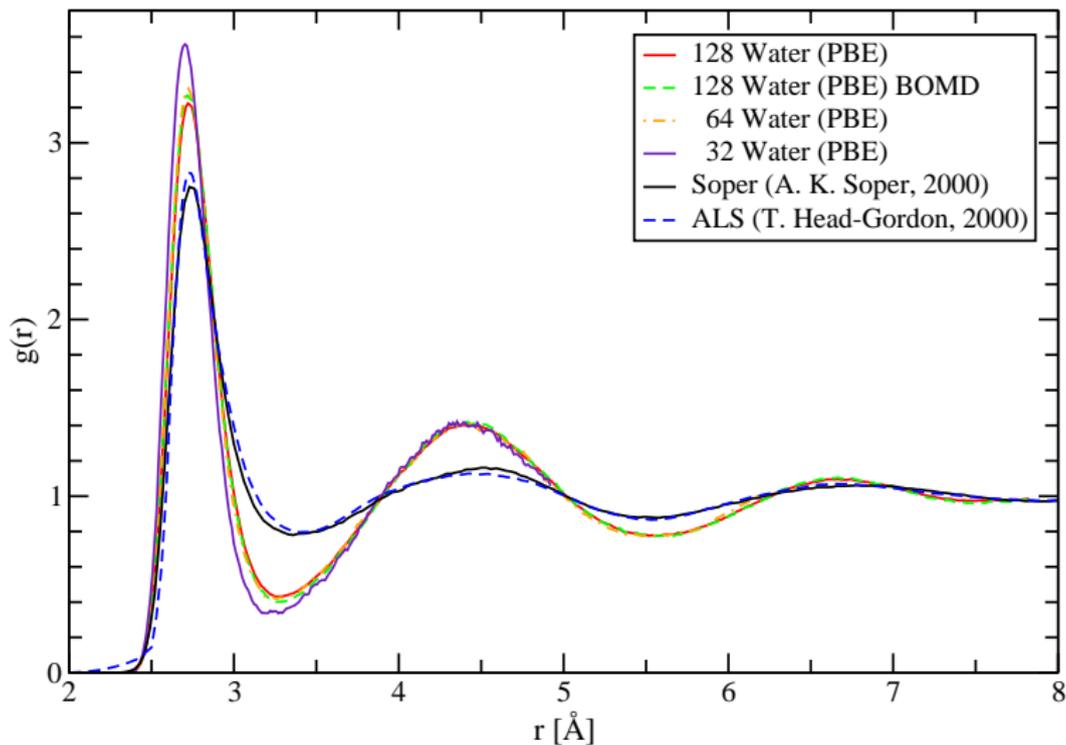
- Si₆₄ (lq)
- 3000 K
- 3D PBC
- LDA
- GTH-PP
- DZVP
- 100 Ry
- t > 1 ns



A correct canonical sampling of the Boltzmann distribution is performed!

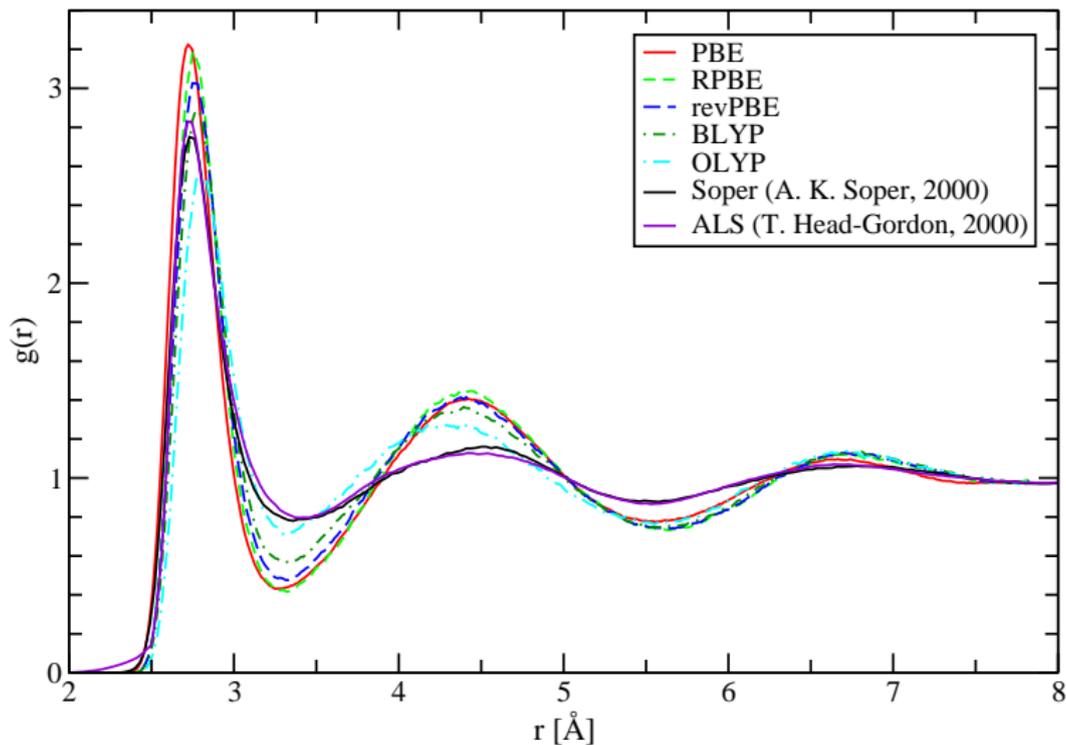
Partial Pair-Correlation Function: Finite Size Effects

T. D. Kühne, M. Krack and M. Parrinello, J. Chem. Theory Comput. **5**, 235 (2009)



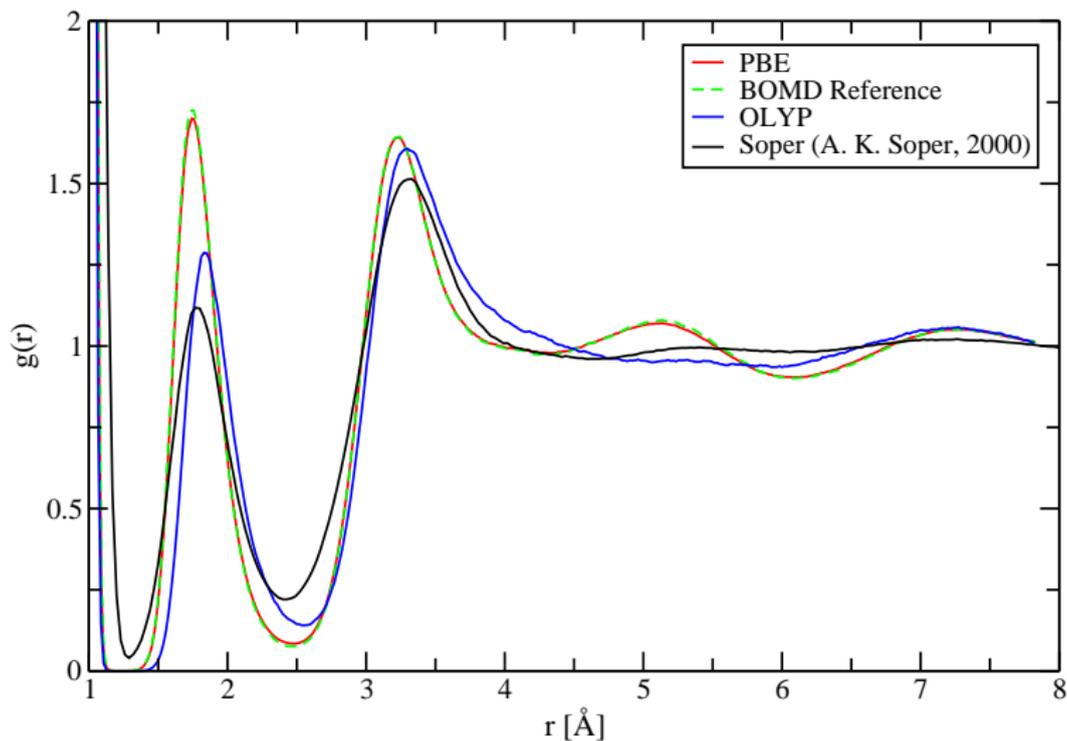
Partial Pair-Correlation Function: XC Functionals

T. D. Kühne, M. Krack and M. Parrinello, J. Chem. Theory Comput. **5**, 235 (2009)



Partial Pair-Correlation Function: Oxygen-Hydrogen

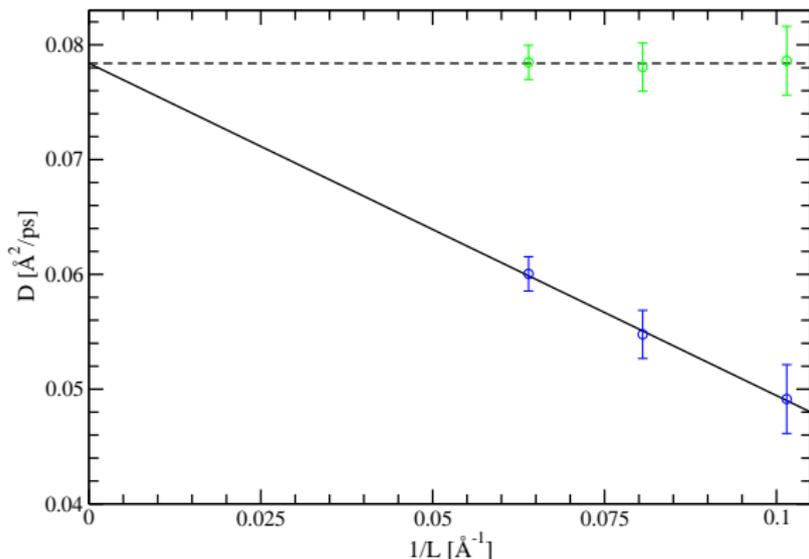
T. D. Kühne, M. Krack and M. Parrinello, J. Chem. Theory Comput. **5**, 235 (2009)



Translational Self-Diffusion for Bulk Water

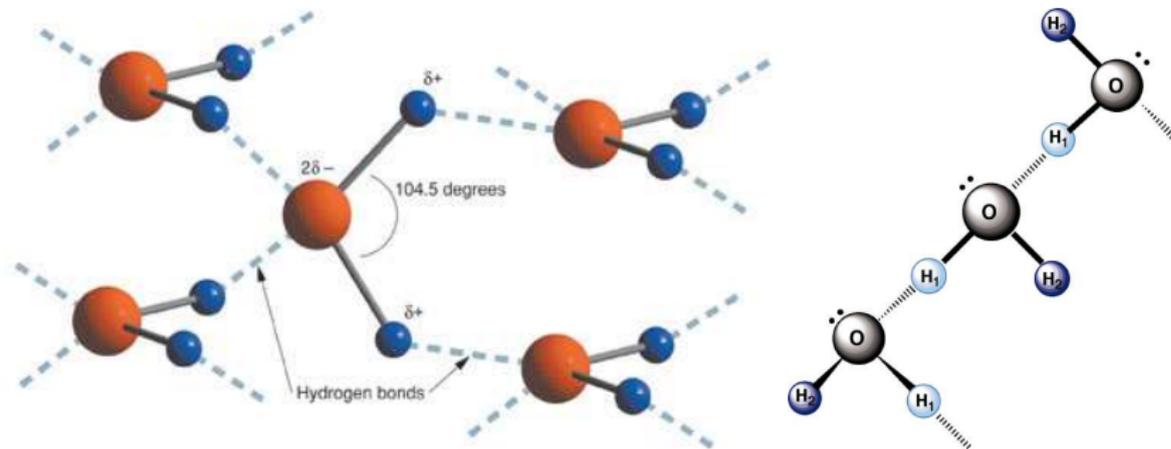
Finite-size effects, extrapolation to infinity and the shear viscosity

- $D_{128}^{\text{BO}} = 0.059(6) \text{ \AA}^2/\text{ps}$
- $D_{128}^{\text{CP}} = 0.060(1) \text{ \AA}^2/\text{ps}$
- $D_{\infty} = D_{\text{PBC}} + \frac{k_B T \zeta}{6\pi\eta L}$
- $D_{\infty}^{\text{CP}} = 0.0789 \text{ \AA}^2/\text{ps}$
- $D_{\text{Exp.}} = 0.2395 \text{ \AA}^2/\text{ps}$
- Shear viscosity:
 $\eta^{\text{CP}} = 21.22 \times 10^{-4} \frac{\text{kg}}{\text{ms}}$
 $\eta_{\text{Exp.}} = 8.92 \times 10^{-4} \frac{\text{kg}}{\text{ms}}$



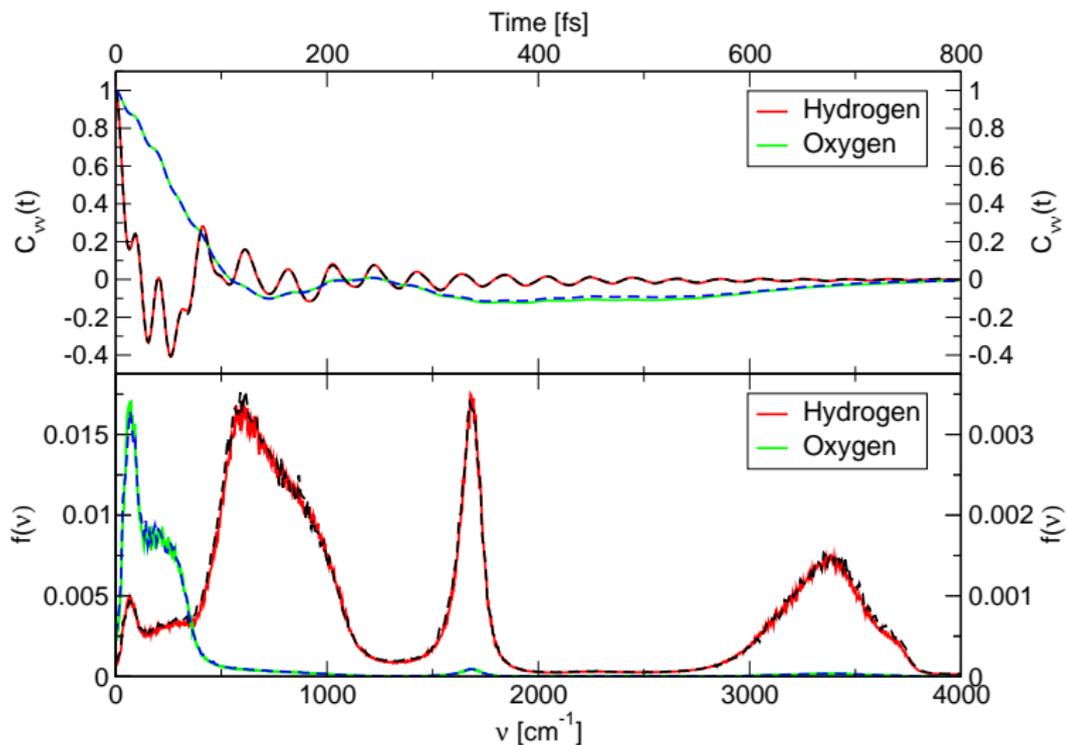
- Diffusion is better reproduced by DFT than generally appreciated
- Translational shear viscosity is much less system size dependent
- System sizes to converge D_{PBC} are essentially out of reach

Hydrogen Bond Network Structure

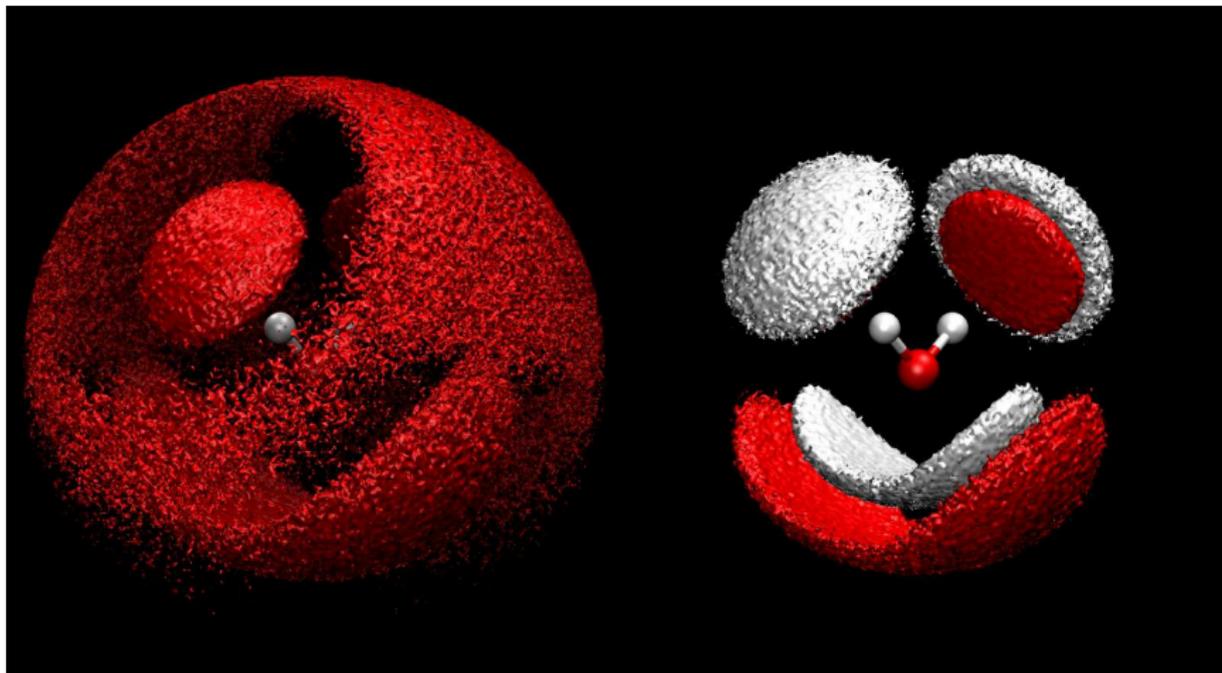


	PBE	RPBE	revPBE	BLYP	OLYP
DD	82.8 %	81.4 %	76.8 %	72.9 %	59.0 %
SD	16.6 %	17.8 %	22.0 %	25.4 %	36.3 %
donated HB's	91.0 %	90.3 %	87.8 %	85.6 %	77.1 %
free HB's	9.0 %	9.7 %	12.2 %	14.4 %	22.9 %
mean HB's	3.642	3.613	3.513	3.423	3.085

Velocity Autocorrelation and Power Spectrum



Spatial Distribution Function



At the DFT level of accuracy liquid water is tetrahedral coordinated!

Hydrogen Bond Kinetics

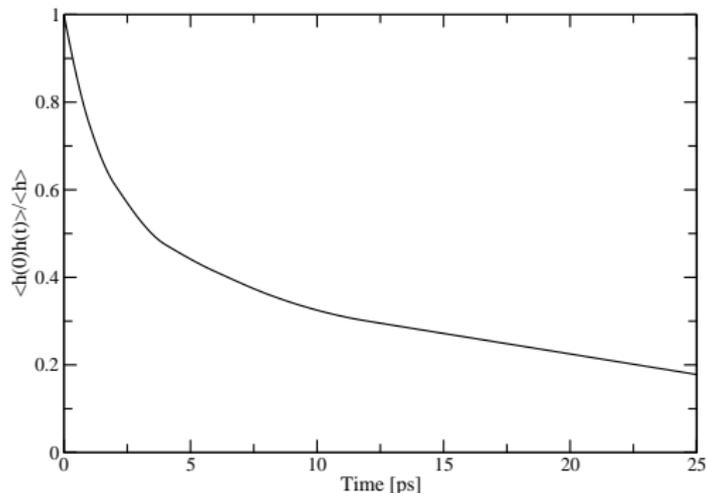
A. Luzar and D. Chandler, *Nature* **379**, 55 (1996) & *Phys. Rev. Lett.* **76**, 928 (1996)

- Reactive flux correlation function:

$$k(t) = -\frac{dc(t)}{dt} = -\frac{\langle (dh/dt)_{t=0} [1 - h(t)] \rangle}{\langle h \rangle} = kc(t) - k'n(t),$$

- $c(t) = \langle h(0)h(t) \rangle / \langle h \rangle$, the HB autocorrelation function
- $n(t) = \langle h(0) [1 - h(t)] H(t) \rangle / \langle h \rangle$, the correlation that two initially bonded water molecules are unbonded but in close proximity
- k and k' are rate constants, determined by least square fitting
- The HB life time $\tau_{\text{HB}} = k^{-1} = 6.98 \text{ ps}$, is 2-3 times larger than with classical force fields (FF)
- This is, at least partially, due to the much slower HB relaxation

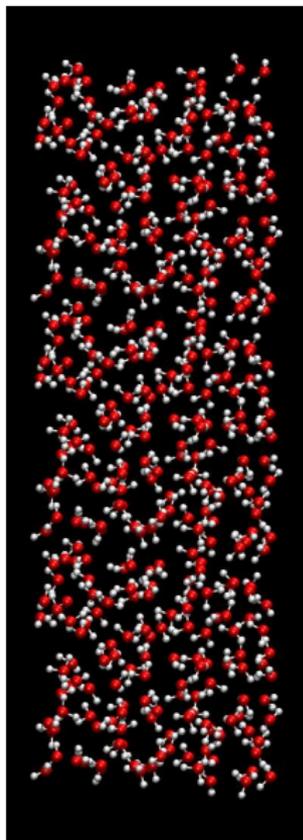
Hydrogen Bond Relaxation



- Contrary to FF $c(t)$ is slowly and non-exponential decaying
- Circumstantial evidences that this might be due to cooperativity
- Substantial dependence on the diffusion and thereby system-size

$$\tau_{rlx} = \frac{\int dt t c(t)}{\int dt c(t)} \simeq 10.25 \text{ ps} \implies \frac{\tau_{rlx}}{\tau_{HB}} = 1.47$$

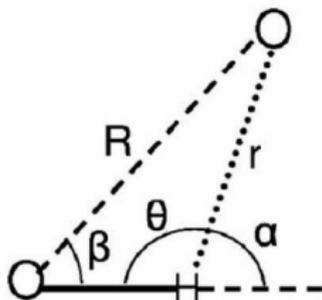
Computational Details



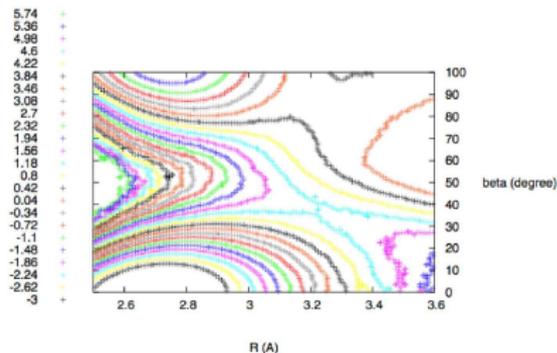
- CP2K/QUICKSTEP (cp2k.berlios.de)
- Orthorhombic box ($15.66 \text{ \AA} \times 15.66 \text{ \AA} \times 55.00 \text{ \AA}$) consisting of 384 light water molecules
- 2D PBC using the decoupling scheme of P. Blöchl
- Temperature very close to 300 K and $\Delta t = 0.5 \text{ fs}$
- Semilocal BLYP GGA XC functional
- Empirical (damped C_6/R^6) vdW correction
- In total more than 300 ps of simulation time
- Molecular optimized TZV2P Gaussian basis set
- BSSE $\approx 0.3 \text{ kcal/mol}$ with a std $\leq 0.01 \text{ kcal/mol}$
- GTH-PP along with a density cutoff of 320 Ry
- $K = 7 \Rightarrow$ time reversibility up to $\mathcal{O}(\Delta t^{12})$
- $\gamma_D = 8.65 \times 10^{-5} \text{ fs}^{-1}$ and $\gamma_L = 1.35 \times 10^{-5} \text{ fs}^{-1}$

Hydrogen Bond Definition via the 2D PMF

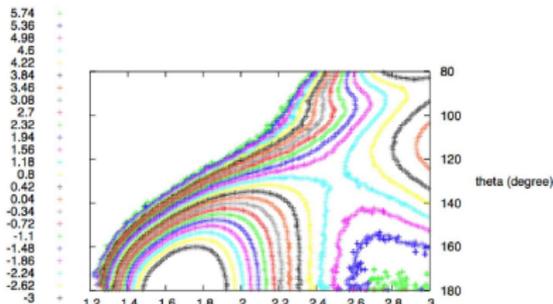
J. L. Skinner et al., J. Chem. Phys. **126**, 204107 (2007)



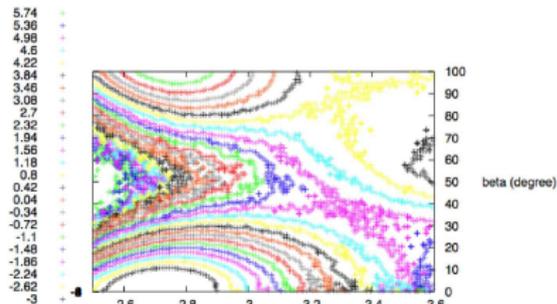
2D PMF $W(R, \beta)$ for bulk region (10-35 Å)



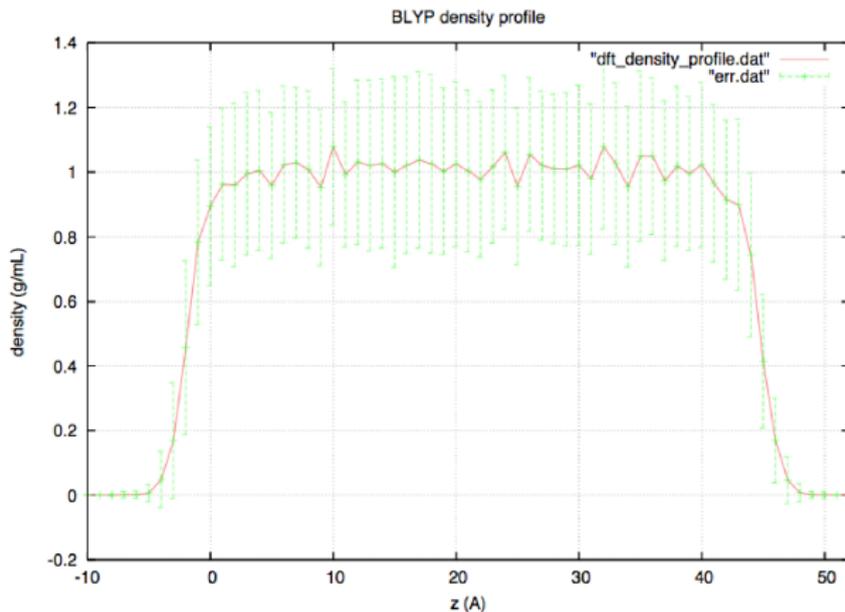
2D PMF $W(r, \theta)$ for bulk region (10-35 Å)



2D PMF $W(R, \beta)$ for surface (>42 Å)



Defining the Surface: Density Profile



$$\rho(z) = a \{1 + \tanh [b(z - c)]\} \longrightarrow \text{GDS at } c = 44.5 \text{ \AA} \pm 2.5 \text{ \AA}$$

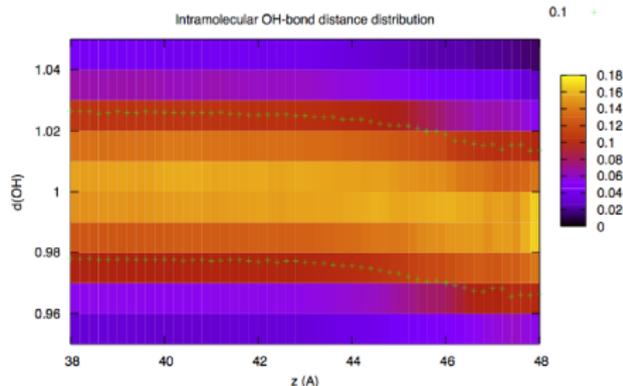
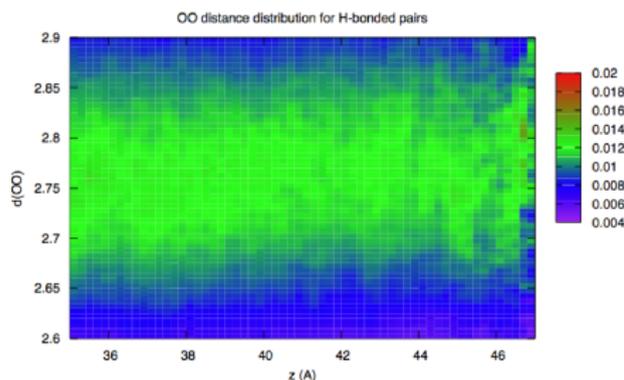
VdW correction systematically improves the density of liquid water

Hydrogen Bonding at the Liquid-Vapor Interface

	2D PMF $W(R, \beta)$		2D PMF $W(r, \theta)$		Kuo & Mundy Science 05		Experiment SFG (Chen)	
	bulk	surf.	bulk	surf.	bulk	surf.	bulk	surf.
DD	84%	49%	84%	48%	40%	14%		
SD	15%	45%	15%	45%	45%	66%		
ND	1%	6%	1%	7%	12%	19%		
free HB	8%	29%	8%	29%	35%	52%	15%	25%
mean HB	3.7	2.9	3.7	2.8	2.5	1.9	3.4	3.0

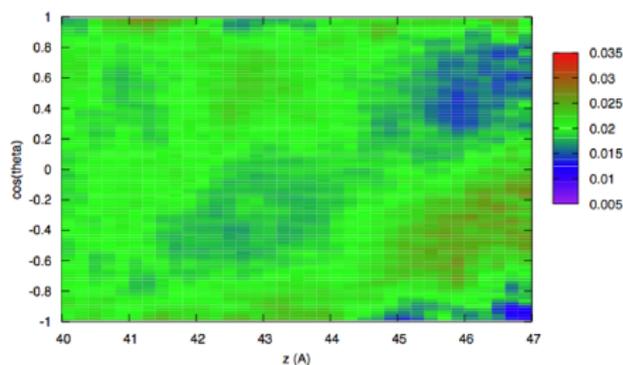
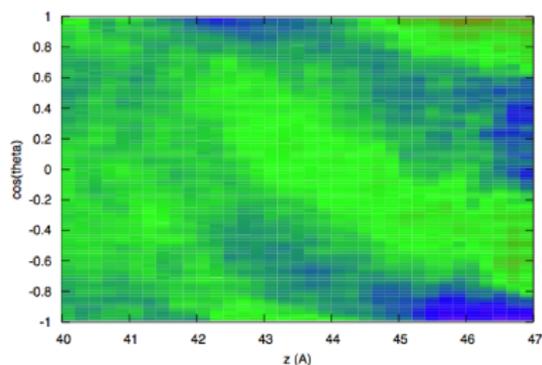
- Very recently Saykally et al. have revisited their own TIY-EXAFS experiments and found that the occurrence of SD and ND are substantially lower than initially reported back in 2001/2002.
- Most of the discrepancy can be explained by the HB definition

Bond Distance Distribution Near the Surface



- The average OO distance increases from 2.809 Å to 2.825 Å
- However, the intramolecular OH bond decreases to 1.007 Å
- Apparently caused by the "missing" HB of the dangling OH bonds
- Also due to the eventual transition to gas phase for $z \geq 47$ Å

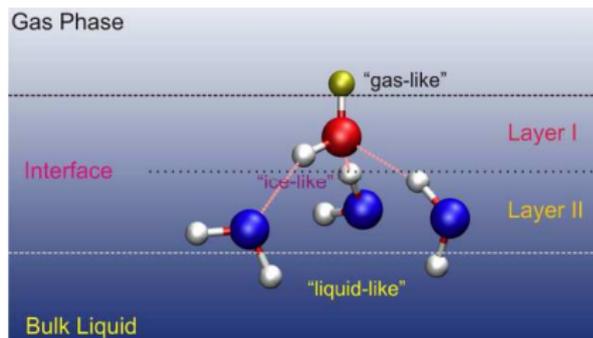
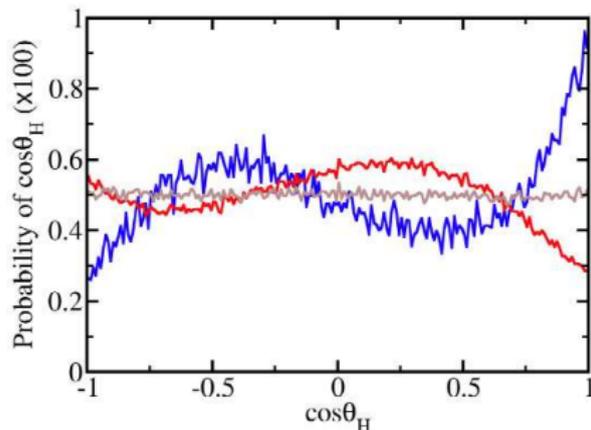
Orientational Distribution Near the Surface: Layer I



- Comparing the distribution of OH-vectors and lone pair orbitals:
- OH bonds prefer to protrude out of the water phase
- However, the lone pair orbitals tend to point into the water phase
- Consistent with the conjecture that the order of Layer I is to maximize the HB and to minimize the exposed partial charge

Orientational Distribution Near the Surface: Layer II

Y. Q. Gao et al., J. Phys. Chem. B **113**, 11672 (2009)



- Only Layer I and II obeys structural order
- The OH vectors of Layer I is peaked at 0° and at around 107° , i.e. tends to donate HB's to Layer II
- In Layer II the opposite is the case - the OH bonds prefer to point into the bulk phase, while providing lone pair orbitals to Layer I

Acknowledgements

- Hagai Eshet (ETH Zürich)
- Rustam Z. Khaliullin (ETH Zürich)
- Matthias Krack (PSI Villingen)
- Michele Parrinello (ETH Zürich)
- CP2K/*Quickstep* Development Team
- The generous allocation of computer time and support from the following institutions and fellows is greatly acknowledged:
 - Harvard University Research Computing Group, Terri Kaltz
 - Swiss National Supercomputing Center, Neil Stringfellow

Thanks for your attention!

Calculating Expectation Values from First Principles

The task is to evaluate expectation values $\langle \mathcal{O} \rangle$ of arbitrary operators $\mathcal{O}(\mathbf{R}_I, \mathbf{P}_I)$ with respect to the configurational Boltzmann distribution

$$\langle \mathcal{O} \rangle = \frac{\int d\mathbf{r} d\mathbf{p} \mathcal{O}(\mathbf{R}_I, \mathbf{P}_I) e^{-\beta E(\mathbf{R}_I, \mathbf{P}_I)}}{\int d\mathbf{r} d\mathbf{p} e^{-\beta E(\mathbf{R}_I, \mathbf{P}_I)}} = \lim_{t \rightarrow \infty} \frac{\sum \mathcal{O}(\mathbf{R}_I(t), \mathbf{P}_I(t)) e^{-\beta E(\mathbf{R}_I(t), \mathbf{P}_I(t))}}{\sum e^{-\beta E(\mathbf{R}_I(t), \mathbf{P}_I(t))}}$$

Opens the door to a variety of relevant static and dynamics properties

- The ions are treated classically (\rightarrow PIMC/PIMD)
- Accuracy of the potential energy surface
- Single particle finite size effects, although PBC helps a lot
- **Statistical error, in particular due to correlation along a trajectory**

Car-Parrinello vs. Born-Oppenheimer Dynamics

Car-Parrinello MD	Born-Oppenheimer MD
<p>WF is propagated, no optimization</p> <p>perfect energy conservation</p> <p>slightly above the BO surface</p> <p>small integration time step</p>	<p>very tight WF optimization</p> <p>reasonable energy conservation</p> <p>on the BO surface</p> <p>big integration time step</p>

- Goal: Unify the best of CPMD & BOMD
 - Efficient, due to solely propagating the electronic density
 - The ions are integrated with at least BOMD time steps
 - Accurate, by approaching BOMD
 - Metals can be treated
 - Ab-initio dynamics method (not competitive in static calculations)

A very brief review of CPMD I

The CP Lagrangian

$$\mathcal{L}_{\text{CP}}(\{\psi_i\}; \mathbf{R}_I, \dot{\mathbf{R}}_I) = \frac{1}{2}\mu \sum_{i=1}^M \langle \dot{\psi}_i | \dot{\psi}_i \rangle + \frac{1}{2} \sum_{I=1}^N M_I \dot{\mathbf{R}}_I^2 - E[\{\psi_i\}; \mathbf{R}_I] + \sum_{i,j} \Lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij})$$



- $\psi_i(\mathbf{r}, t)$ are thought of classical fields
- $E[\{\psi_i\}; \mathbf{R}_I]$ vs. $\min E[\{\psi_i\}; \mathbf{R}_I]$
- Fictitious mass parameter μ
- Ions and e^- are on the same footing

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

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \langle \dot{\psi}_i |} = \frac{\partial \mathcal{L}}{\partial \langle \psi_i |}$$

A very brief review of CPMD II

The Equations of Motion

$$\begin{aligned}
 M_I \ddot{\mathbf{R}}_I &= -\nabla_{\mathbf{R}_I} \left[E[\{\psi_i\}; \mathbf{R}_I] \Big|_{\{\langle \psi_i | \psi_j \rangle = \delta_{ij}\}} \right] \\
 &= -\frac{\partial E}{\partial \mathbf{R}_I} + \sum_{i,j} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \langle \psi_i | \psi_j \rangle \\
 \mu \ddot{\psi}_i(\mathbf{r}, t) &= -\frac{\delta E}{\delta \langle \psi_i |} + \sum_j \Lambda_{ij} |\psi_j\rangle \\
 &= -\hat{H}_e \langle \psi_i | + \sum_j \Lambda_{ij} |\psi_j\rangle
 \end{aligned}$$

- The HF forces are correct, though the theorem doesn't hold
- Under certain conditions the electrons adiabatically follow the ions
- Then the metastable state can be sustained and $\ddot{\psi}_i \simeq 0$
- Energies & Forces are NOT on the BO surface, but are consistent

A very brief review of CPMD II

The Equations of Motion

$$\begin{aligned}
 M_I \ddot{\mathbf{R}}_I &= -\nabla_{\mathbf{R}_I} \left[E[\{\psi_i\}; \mathbf{R}_I] \Big|_{\{\langle \psi_i | \psi_j \rangle = \delta_{ij}\}} \right] \\
 &= -\frac{\partial E}{\partial \mathbf{R}_I} + \sum_{i,j} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \langle \psi_i | \psi_j \rangle \\
 \mu \ddot{\psi}_i(\mathbf{r}, t) &= -\frac{\delta E}{\delta \langle \psi_i |} + \sum_j \Lambda_{ij} |\psi_j\rangle \\
 &= -\hat{H}_e \langle \psi_i | + \sum_j \Lambda_{ij} |\psi_j\rangle
 \end{aligned}$$

- The HF forces are correct, though the theorem doesn't hold
- Under certain conditions the electrons adiabatically follow the ions
- Then the metastable state can be sustained and $\ddot{\psi}_i \simeq 0$
- Energies & Forces are NOT on the BO surface, but are consistent

A very brief review of CPMD III

The fictitious mass parameter μ

$$|\psi_\mu(\mathbf{r}, t) - \psi_0(\mathbf{r}, t)| \leq C\sqrt{\mu}$$

$$\Delta t_{\max} \propto \sqrt{\frac{\mu}{\Delta E_{\text{gap}}}}$$

- μ acts as a continuous slider between speed and accuracy
- Typically, the timestep is $\sim 5 \times -10 \times$ smaller than in BOMD
- Depends on the application if either CPMD or BOMD is to favor
- **Metals are problematic: Finite electron temperature or thermostats**

Desirable to eliminate μ !

The Coupled Electron-Ion Molecular Dynamics

T. Kühne, M. Krack, F. Mohamed and M. Parrinello, Phys. Rev. Lett. **98**, 066401 (2007)

$$2\zeta \stackrel{\text{def}}{=} \gamma \times \sqrt{\frac{\mu|\psi_i\rangle}{\left(\hat{H}_{\text{NSC}}|\psi_i\rangle - \sum_j \Lambda_{ij}|\psi_j\rangle\right)}} = \frac{\gamma}{\omega}$$

$$M_I \ddot{\mathbf{R}}_I = -\frac{\partial E_{\text{NSC}}}{\partial \mathbf{R}_I} + \sum_{i,j} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \langle \psi_i | \psi_j \rangle$$

$$- 2 \sum_i \frac{\partial \langle \psi_i |}{\partial \mathbf{R}_I} \left[\frac{\partial E_{\text{NSC}}}{\partial \langle \psi_i |} - \sum_j \Lambda_{ij} |\psi_j\rangle \right]$$

$$\frac{d^2}{d\tau^2} \psi_i(\mathbf{r}_i, \tau) = -\frac{\delta E_{\text{NSC}}}{\delta \langle \psi_i |} - 2\zeta \frac{d}{d\tau} \psi_i(\mathbf{r}_i, \tau) + \sum_j \Lambda_{ij} |\psi_j\rangle$$

The Always Stable Predictor Corrector (ASPC) Method

Propagating the electrons by a Gear-type predictor-corrector integrator (Kolafa, JCC 04)

- Instead of C the contra-covariant density matrix PS is propagated
- As PS is an approximate projector onto the occupied subspace:

$$C^p(t_n) \cong \sum_{m=1}^K (-1)^{m+1} m \frac{\binom{2K}{K-m}}{\binom{2K-2}{K-1}} \underbrace{C(t_{n-m})C^T(t_{n-m})}_{P(t_{n-m})} \times S(t_{n-m})C(t_{n-1})$$

- The corrector consists of one preconditioned gradient calculation
- For a certain choice of α "always stability" can be proven
- The corrector can be iteratively applied, to control the deviation

$$C(t_n) = \alpha \text{MIN}[C^p(t_n)] + (1 - \alpha)C^p(t_n),$$

$$\text{with } \alpha = \frac{K}{2K - 1}$$

Orbital Transformations as a Corrector

J. VandeVondele and J. Hutter, J. Chem. Phys. **118**, 4365 (2003)

$$C(\mathbf{X}) = C^p(t_n) \cos(U) + \mathbf{X}U^{-1} \sin(U), \text{ with } U = (\mathbf{X}^T \mathbf{S} \mathbf{X})^{1/2}$$

The variable \mathbf{X} has to obey the linear constraint $\mathbf{X}^T \mathbf{S} C^p(t_n) = 0$

$$\frac{\partial(E[C(\mathbf{X})] + \text{Tr}[\mathbf{X}^T \mathbf{S} C^p \Lambda])}{\partial \mathbf{X}} = \frac{\partial E[C(\mathbf{X})]}{\partial C} \frac{\partial C(\mathbf{X})}{\partial \mathbf{X}} + \mathbf{S} C^p \Lambda,$$

$$\text{where } \Lambda = -[(\mathbf{S} C^p)^T \mathbf{S} C^p]^{-1} (\mathbf{S} C^p)^T \frac{\partial E[C(\mathbf{X})]}{\partial \mathbf{X}}$$

- The orthonormalization constraint is fulfilled by construction
- The requirements of a SCF solution and exact eigenfunctions are both fully bypassed
- The non-self-consistent energy functional is solely evaluated:
Neither diagonalization nor iterative minimization is required!

The Harris-Foulkes Functional

$$E_{\text{PC}}[\rho^p] = \text{Tr} [\mathbf{C}^T H[\rho^p] \mathbf{C}] - \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}' \frac{\rho^p(\mathbf{r})\rho^p(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ - \int d\mathbf{r} V_{\text{XC}}[\rho^p]\rho^p + E_{\text{XC}}[\rho^p] + E_{\text{II}},$$

- Linear partial differential equation, since it depends only on ρ^p
- Can be solved non-self-consistently
- Converges to the KS-equation at self-consistency, with $\rho^p \rightarrow \rho^{SC}$
- Not variational, but stationary at the ground state
- Typically a better ground-state estimate than the KS-equation
- The error in the energy is $O((\Delta\rho)^2)$ and 1st order for the forces
- However, as $\Delta\rho \neq 0$ an extra term in the force calculation appears

$$- \int d\mathbf{r} \left\{ \left[\left(\frac{\partial V_{\text{XC}}[\rho^p]}{\partial \rho^p} \right) \Delta\rho + V_{\text{H}}[\Delta\rho] \right] (\nabla_I \rho^p) \right\}$$

The Approximations and Main Conceptual Ideas

T. Kühne, M. Krack, F. Mohamed and M. Parrinello, Phys. Rev. Lett. **98**, 066401 (2007)

- ASPC is not rigorously time-reversible
 - But time-reversible up to $O(\hbar^{2K-2})$ and trivially at exact propagation
- The propagation of the density matrix is not strictly constraint-free
 - Idempotency of the density matrix has to be explicitly enforced, but typically a single purification iteration is sufficient
- C is only an approximate eigenfunction, which leads to a tiny error in the forces, that is negligible but analytically solvable by DFPT
- For the short-term integration of the electrons accuracy is crucial, thus an explicit, though non-symplectic ASPC integrator is used
- Non-symplectic integrators generally turn a conservative system into a dissipative one, leading to an exponential decay in energy

The Modified Langevin Equation

T. Kühne, M. Krack, F. Mohamed and M. Parrinello, Phys. Rev. Lett. **98**, 066401 (2007)

$$\begin{aligned}
 M_I \ddot{\mathbf{R}}_I &= \underbrace{\mathbf{F}_I^{BO} - M_I \dot{\mathbf{R}}_I (\gamma_D + \gamma_L)} + (\boldsymbol{\Xi}_I^D + \boldsymbol{\Xi}_I^L) \\
 &= \mathbf{F}_I^{PC} - \gamma_L M_I \dot{\mathbf{R}}_I + (\boldsymbol{\Xi}_I^D + \boldsymbol{\Xi}_I^L)
 \end{aligned}$$

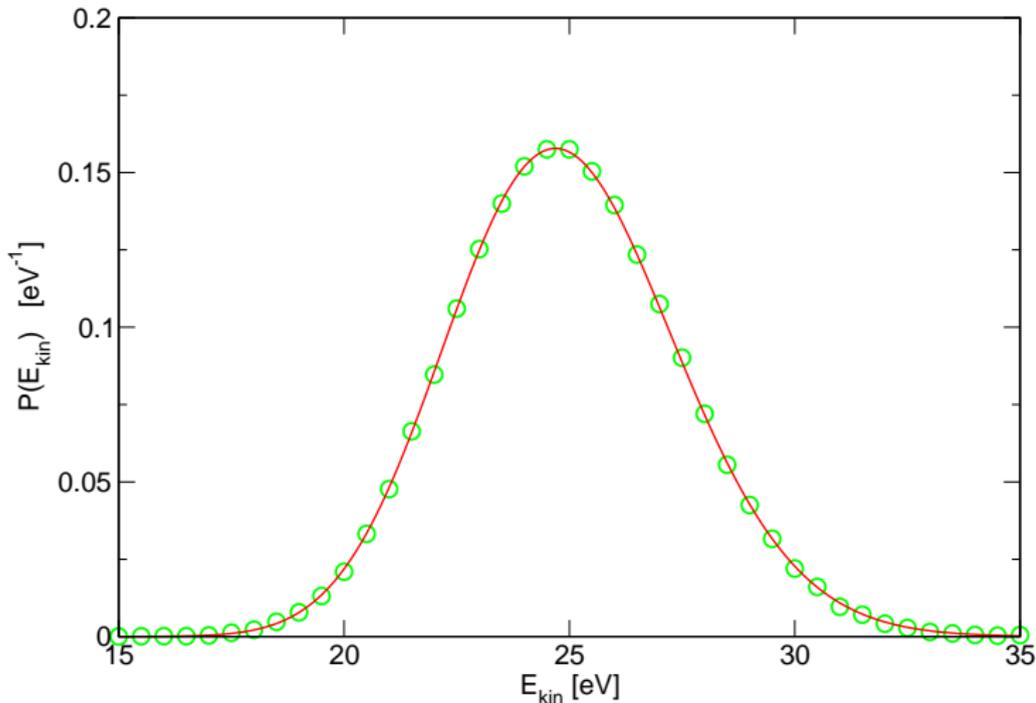
- To guarantee an accurate sampling of the Boltzmann distribution, the noise has to obey the fluctuation dissipation theorem:

$$\langle \{\boldsymbol{\Xi}_I^D(0) + \boldsymbol{\Xi}_I^L(0)\} \{\boldsymbol{\Xi}_I^D(t) + \boldsymbol{\Xi}_I^L(t)\} \rangle = 2(\gamma_D + \gamma_L) M_I k_B T \delta(t)$$
- If the dissipation is indeed exponential decaying, we can do the substitution and still get a rigorously defined canonical ensemble
- γ_D , such that the equipartition theorem $\langle \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 \rangle = \frac{3}{2} k_B T$ holds
- The white noise condition is satisfied by definition

Is our assumption justified?

Is the Kinetic Energy Distribution Maxwellian distributed?

- Si₆₄ (lq)
- 3000 K
- GTH-PP
- DZVP
- 100 Ry
- LDA
- 3D PBC
- t > 1 ns



A correct canonical sampling of the Boltzmann distribution is performed!

The Open-Source Project CP2K

A suite of sampling and force methods for large-scale simulations in condensed phases

- Static calculations
 - Energy, Structure and Transition Path Optimization
 - Property calculation (e.g. Mulliken charges, Wannier functions)
- Sampling Techniques
 - Molecular Dynamics
 - Metadynamics
 - Gibbs-Ensemble Monte Carlo
 - Path Integral Molecular Dynamics
- Energy & Force Methods
 - All-electron calculations (GAPW)
 - QUICKSTEP (GPW)
 - LR-TDDFT (Tamm-Dancoff approximation)
 - Approximate DFT Methods: Kim-Gordon, Harris-Foulkes & DFPT
 - Semi-empirical QC & Tight-Binding Methods (e.g. SCC-DFTB)
 - The classic MM code FIST
 - Embedding Methods (e.g. QM/MM, Orbital-free DFT-embedding)

Decent Sampling: **1 Minute 1 Force!** → Parallel low scaling algorithms

The driving force: QUICKSTEP

The Gaussian and Plane Waves (GPW) formalism (Lippert, Hutter & Parrinello)

- Dual representation of the electronic degrees of freedom
 - The orbitals are expanded in a linear combination of Gaussians
 - The electron density is represented in an auxiliary PW basis set

$$P \rightarrow \rho(\mathbf{R}) \xrightarrow{\text{FFT}} \rho(\mathbf{G}) \rightarrow V_{\text{H}}(\mathbf{G}) = \frac{\rho(\mathbf{G})}{G^2} \xrightarrow{\text{FFT}^{-1}} V_{\text{H}}(\mathbf{R}) \rightarrow \mathbf{V}$$

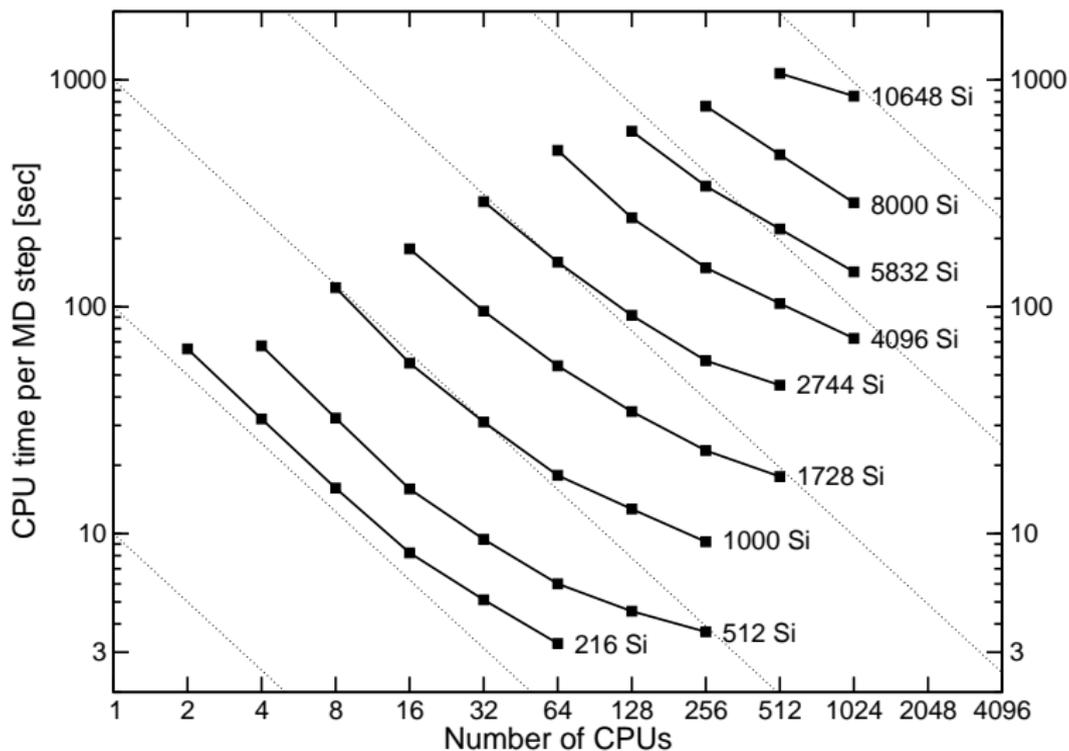
$\underbrace{\hspace{15em}}_{O(n \log n)}$

- This way the best of both worlds is combined with benefit!
 - \Rightarrow The Kohn-Sham matrix built up scales as $O(n \log n)$
 - \Rightarrow The density matrix update is in principle also linear scaling
- Gaussian Augmented Plane Waves (GAPW)
 - "GPW with P. Blochl's PAW density representation"
 - Efficient and accurate all-electron molecular dynamics (M. Krack)

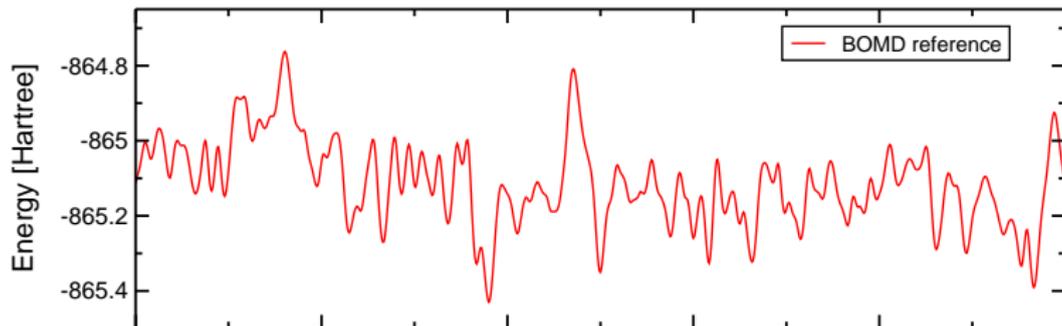
<http://cp2k.berlios.de/>

Scalability and Performance

Once again liquid Silicon in a cubic box with a converged basis set and density cutoff

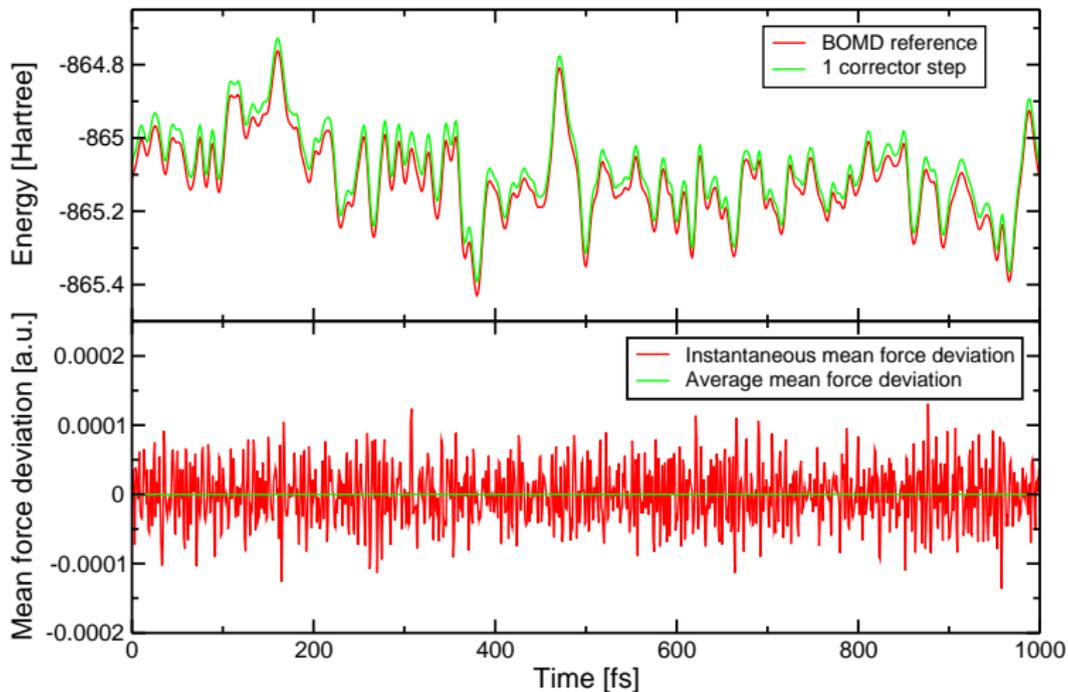


Deviation from the Born-Oppenheimer Surface I

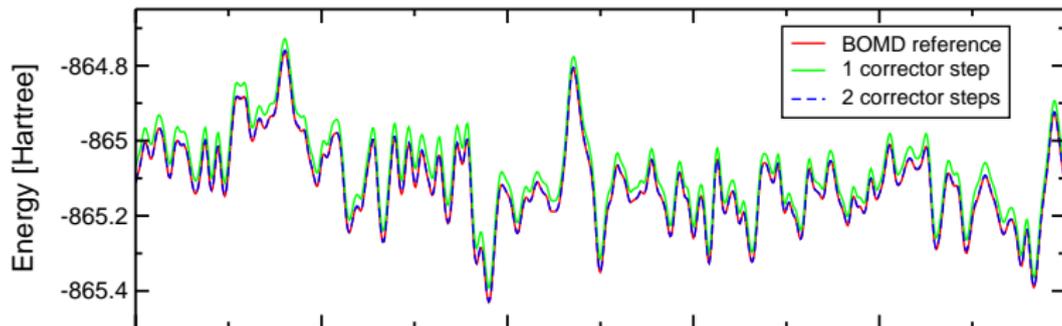


- 24 SiO₂ (lq) at 3500 K
- A DZVP basis set, 280 Ry density cutoff and $t = 1.0$ fs are used
- Constantly bonds are swiftly broken and formed
- Worst case scenario, as the density matrix is rapidly varying

Deviation from the Born-Oppenheimer Surface II



Deviation from the Born-Oppenheimer Surface III



- The energy shift is reasonable small and importantly constant
- Energy differences are very well reproduced
- The electrons are solely propagated: No SCF cycle at all
- The deviation from the BO surface is systematically controllable