

# Large-Scale First-Principles Molecular Dynamics Simulations

François Gygi

University of California, Davis

[fgygi@ucdavis.edu](mailto:fgygi@ucdavis.edu)

<http://eslab.ucdavis.edu>

MASP2012 Workshop

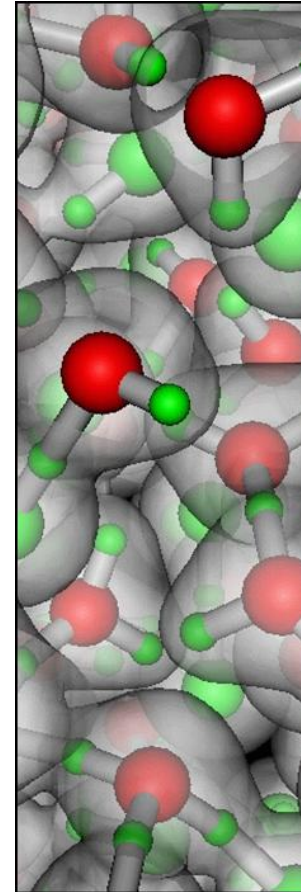
ISSP, Jul 11, 2012

# Acknowledgements

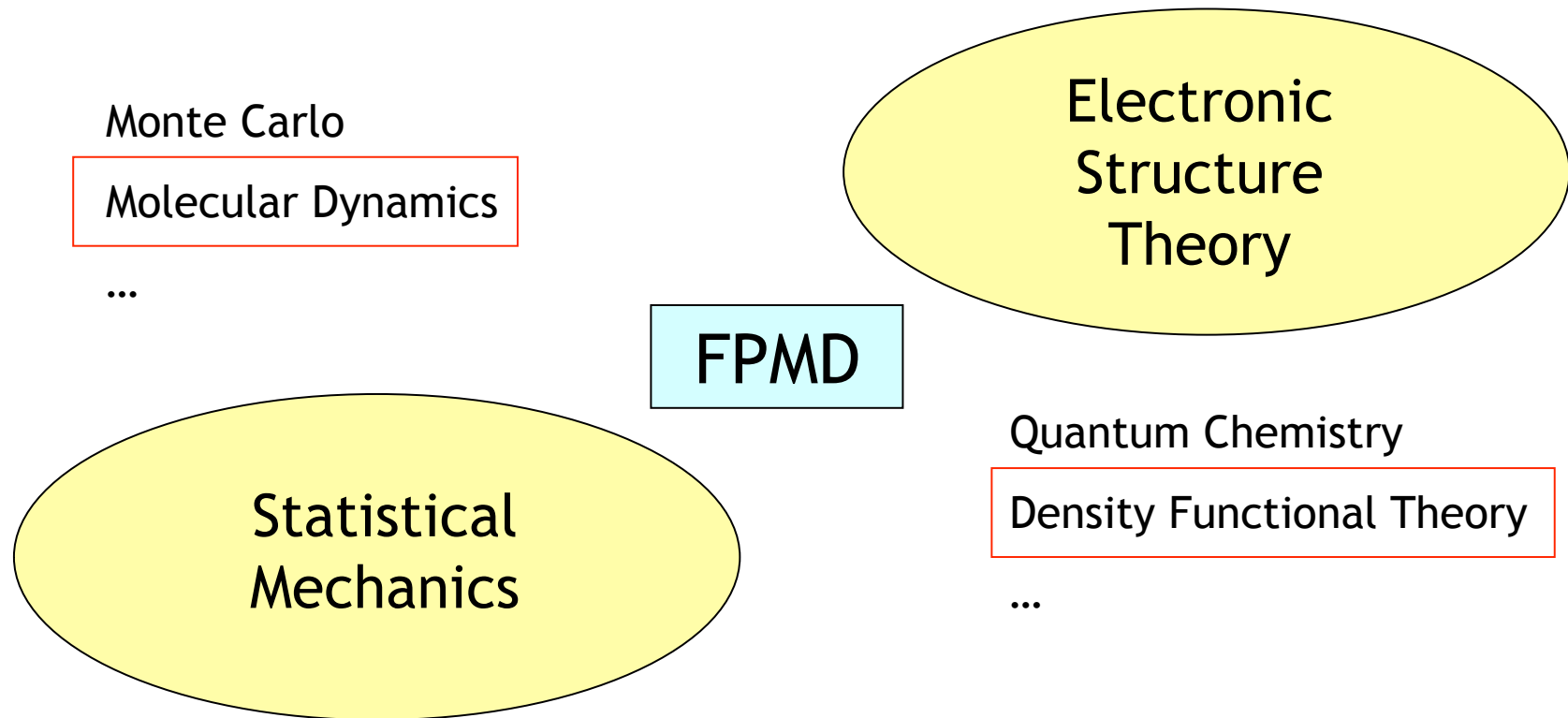
- Giulia Galli (Chemistry, UC Davis)
- Davide Donadio (MPI Mainz)
- Zhaojun Bai (CS, UC Davis)
- \$\$:
  - NSF-OCI PetaApps 0749217
  - DOE SciDAC
- cycles
  - NSF Teragrid (XSEDE)
  - DOE/ANL-ALCF
  - UC ShaRCS
- Ivan Duchemin (post-doc, Appl Sc)
- Gary Yuan (grad st Appl. Sc)
- Jun Wu (grad st, Appl. Sc)
- Yuji Nakatsukasa (grad st, CS)
- William Dawson (grad st, CS)
- Cui Zhang (grad st, Chemistry)

# Outline

- First-Principles Simulations
- Hybrid density functionals
- van der Waals density functionals
- IR spectra in liquids
- Verification and validation of simulation results



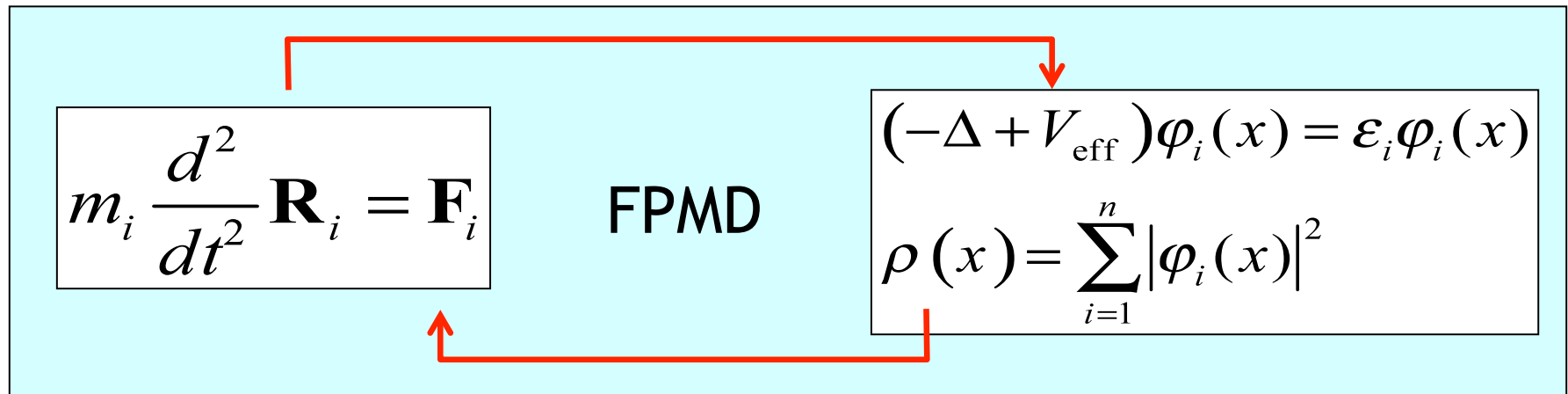
# First-Principles Molecular Dynamics



# First-Principles Molecular Dynamics

Molecular Dynamics

Density Functional Theory



Newton equations

Kohn-Sham equations

# Plane waves and pseudopotentials

- Plane wave basis for electronic wave functions

$$\varphi_j(\mathbf{r}) = \sum_{|\mathbf{q}|^2 < E_{\text{cut}}} c_{\mathbf{q},j} e^{i\mathbf{q}\cdot\mathbf{r}} \quad j = 1, \dots, n$$

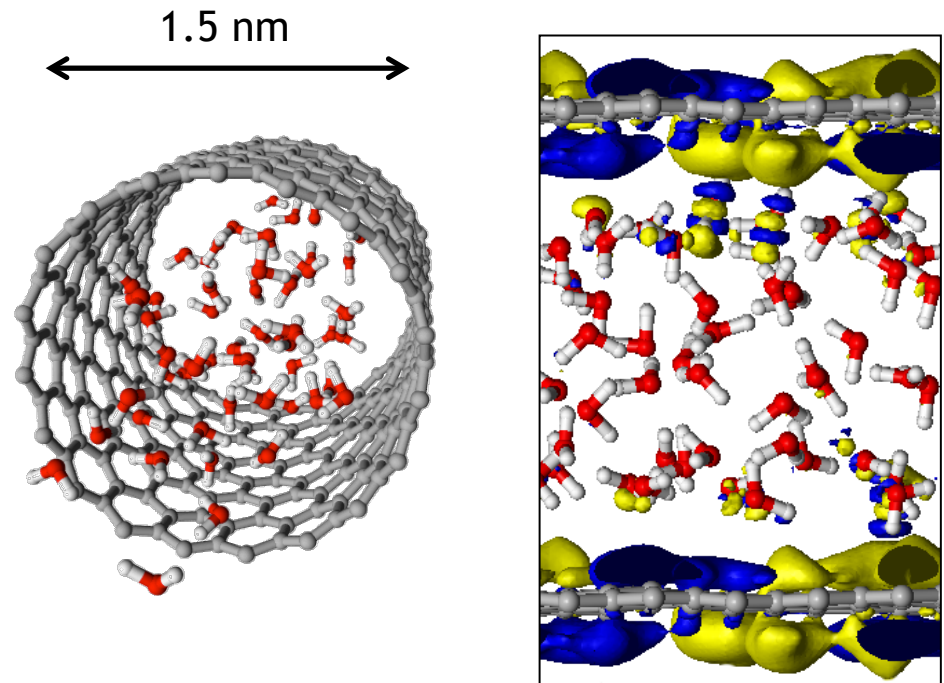
- Solutions represented by an (orthogonal) ( $m \times n$ ) matrix of complex Fourier coefficients

$$Y_{ij} = c_{\mathbf{q}_i,j}$$

- No basis set superposition errors (BSSE)
- Dimensions of  $Y$ :  $10^6 \times 10^4$
- Pseudopotentials: only valence electrons represented

# Water in confined geometries

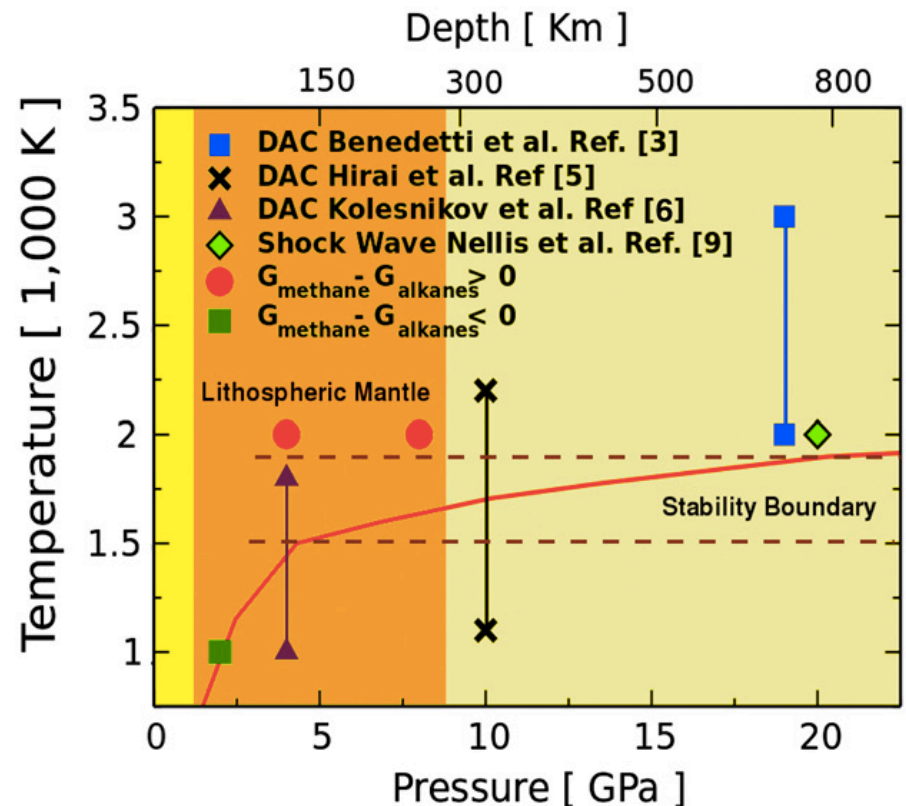
- Simulations of H<sub>2</sub>O confined in carbon nanotubes and between graphene sheets
- Study the effect of the interface on the water hydrogen bond network



G. Cicero, J.C. Grossman, E. Schwegler, F. Gygi, and G. Galli, *JACS* **130**, 1871 (2008).

# Hydrocarbons under pressure

- Investigation of the stability of hydrocarbons at deep Earth pressures and temperatures
- Study the possible creation of hydrocarbons under pressure



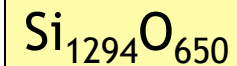
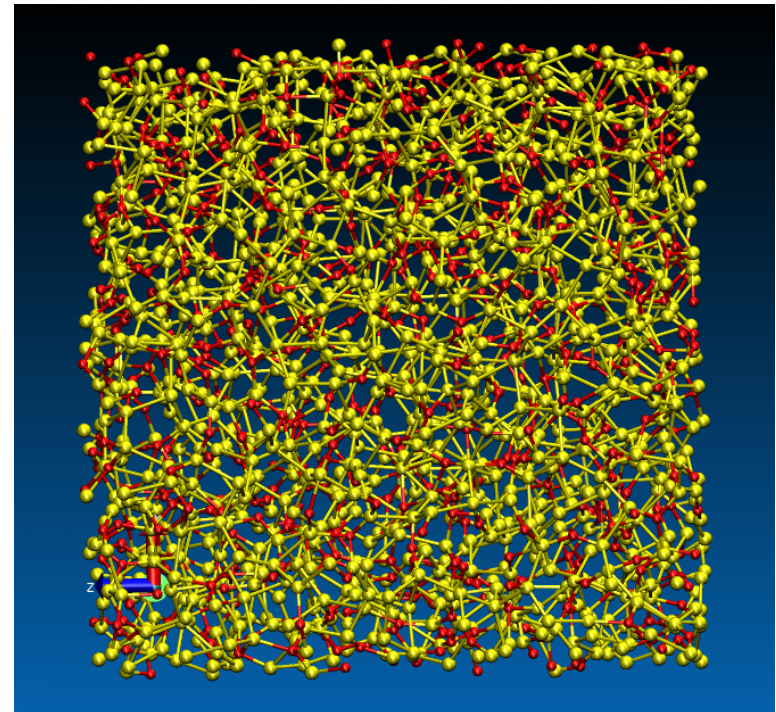
L. Spanu, D. Donadio, D. Hohl, E. Schwegler, and G. Galli,  
*Proc. Nat. Acad. Sci.* **108**, 6843 (2011).

8



# Nanostructures in semiconductors

- Embedded silicon nanoparticle in silicon suboxide ( $\text{SiO}_x$ )
- 1944 atoms, 9076  $e^-$
- Investigation of relaxation of local defect structures, electronic structure, optical properties



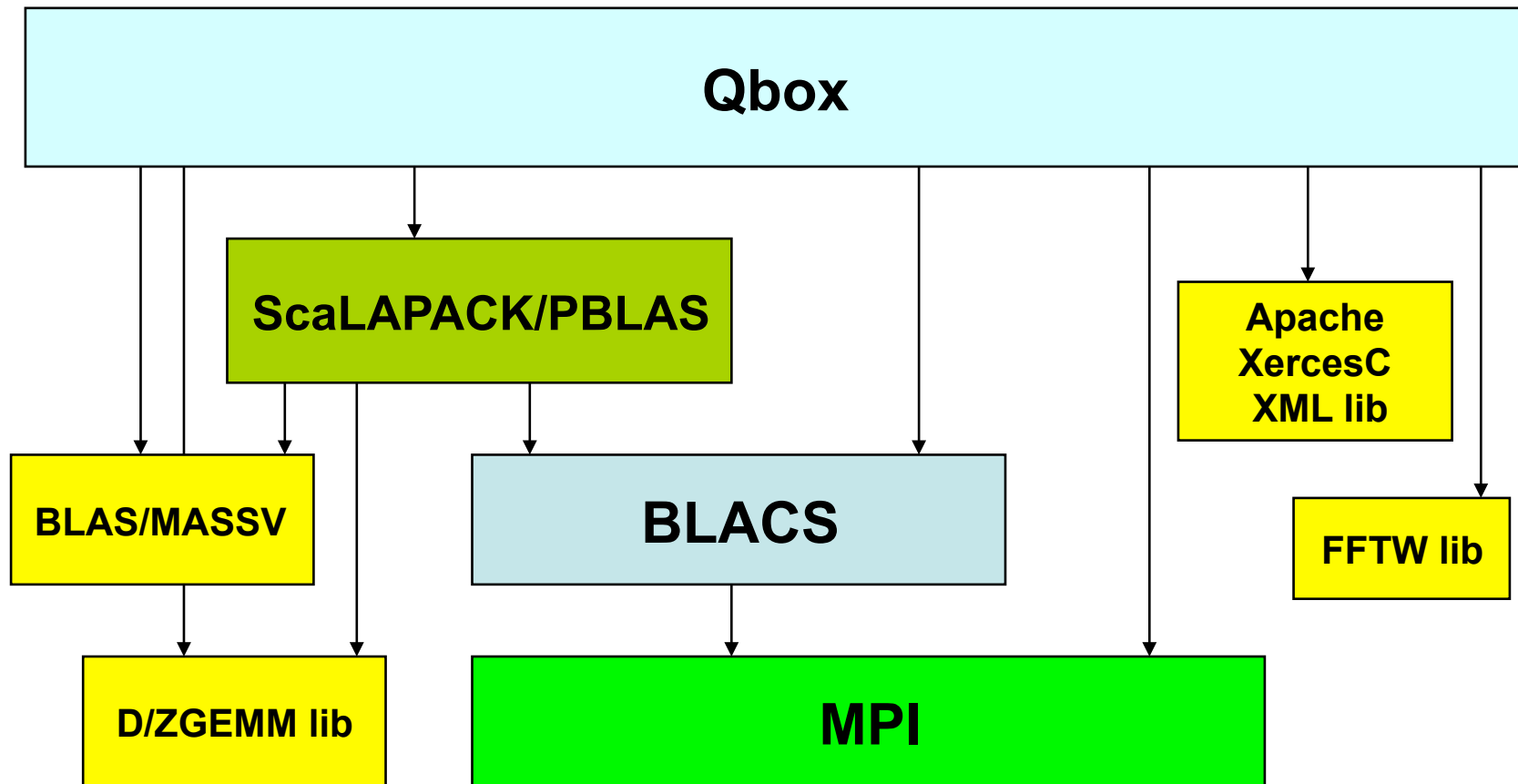
# FPMD simulations: goals and requirements

- Goals
  - simulate larger systems (3-5 nm)<sup>3</sup>
  - use more accurate models (hybrid DFT, vdW DFT)
  - include more features (spectroscopy: IR, Raman)
  - achieve O(N) scaling (with controlled accuracy)
- Needs
  - new scalable algorithms (strong scaling)
    - ANL Intrepid, NICS Kraken, NSF Keeneland, TACC Ranger, NERSC Hopper
  - flexible software infrastructure
    - HPC hardware is a moving target

# Qbox: our scalable implementation of FPMD

- C++/MPI/OpenMP Implementation of FPMD
- Designed for large-scale parallelism
- Built on
  - ScaLAPACK/BLACS library
  - FFTW library
  - BLAS/ATLAS library
  - Apache Xerces-C library
- Client-server capability
- Available under GPL license  
<http://eslab.ucdavis.edu/software/qbox>

# Qbox code structure



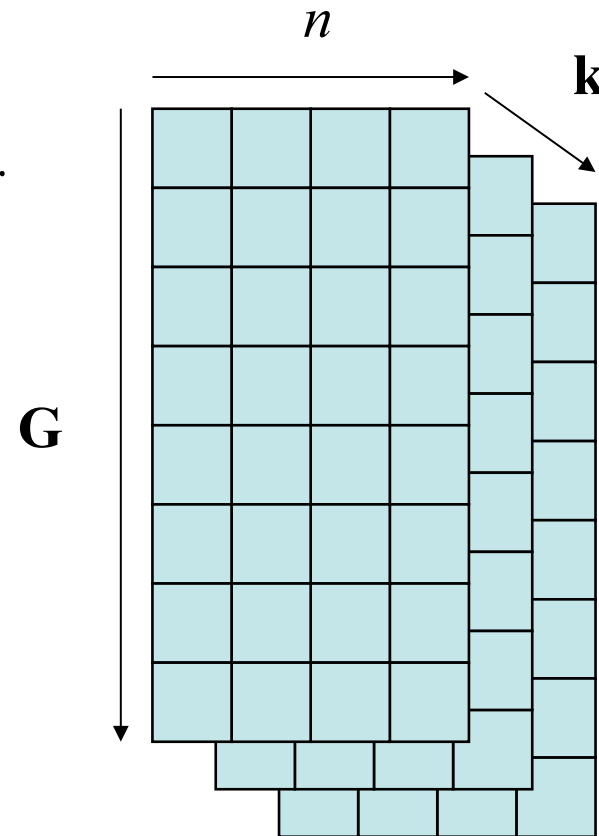
# The cost of FPMD simulations

- The current cost of FPMD is  $O(N^3)$  for  $N$  atoms
  - compare with  $O(N \log N)$  for classical MD
  - Many attempts to reduce to  $O(N)$  or  $O(N \log N)$ 
    - Not feasible without approximations
- Origin of  $O(N^3)$ 
  - orthogonalization of solutions of the KS equations
  - non-local pseudopotentials
- Memory footprint:  $O(N^2)$ 
  - data location is critical

# Data layout

- Distributed plane-wave coefficients

$$\varphi_{n\mathbf{k}}(\mathbf{r}) = \sum_{|\mathbf{k}+\mathbf{G}|^2 < E_{\text{cut}}} c_{n,\mathbf{k}+\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$$

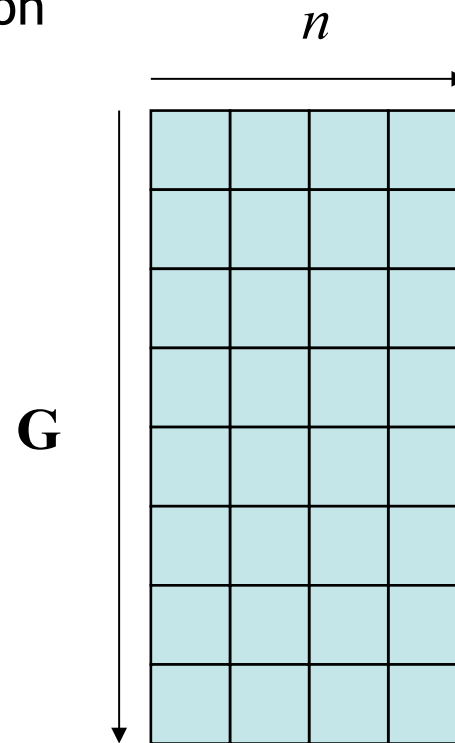


# Data layout

- Single k-point wavefunction:
  - ScaLAPACK matrix block distribution

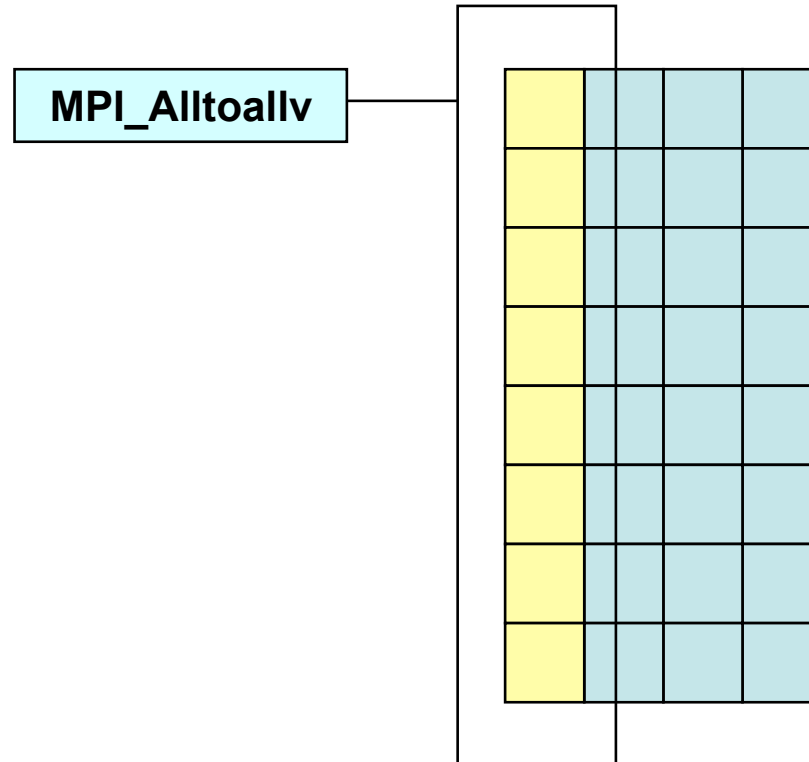
$$\varphi_{n\mathbf{k}}(\mathbf{r}) = \sum_{|\mathbf{k}+\mathbf{G}|^2 < E_{\text{cut}}} c_{n,\mathbf{k}+\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$$

- Dimensions of C:  $10^6 \times 10^4$
- Typical process grid:  $512 \times 16$



# Communication patterns

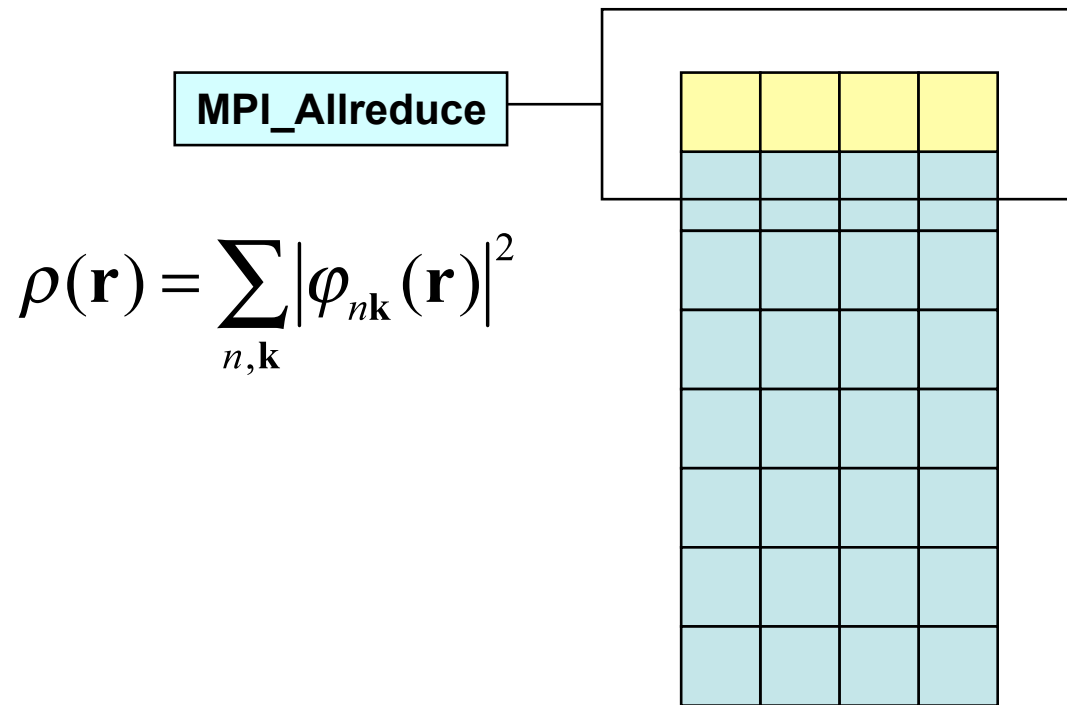
- 3-D Fourier transforms  $\varphi_{n\mathbf{k}}(\mathbf{r}) \leftrightarrow c_{n,\mathbf{k}+\mathbf{G}}$





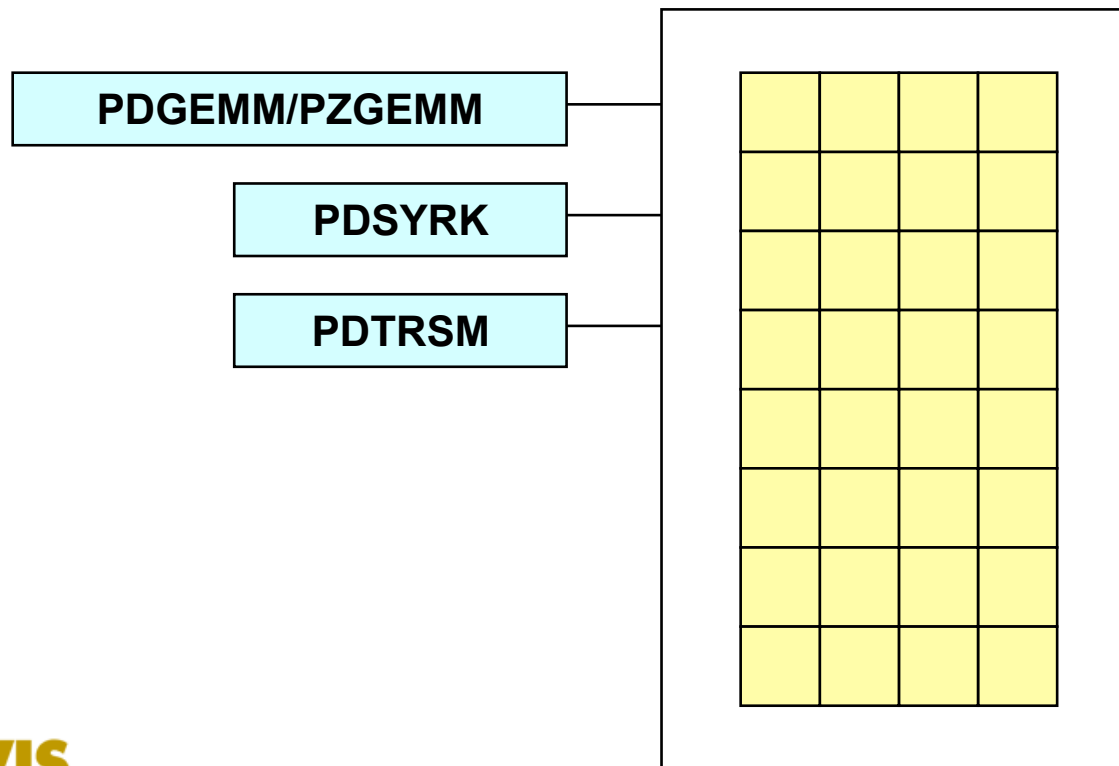
# Communication patterns

- Accumulation of electronic charge density



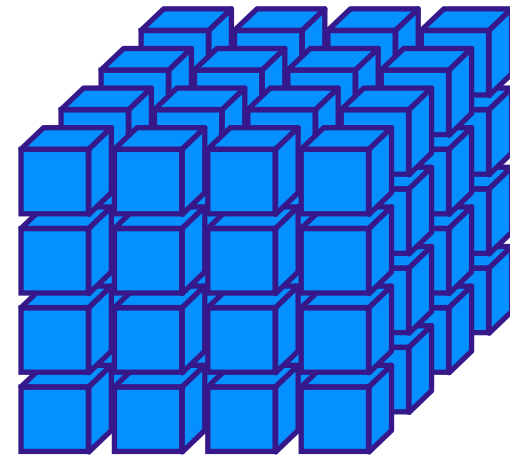
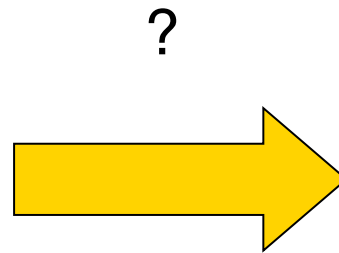
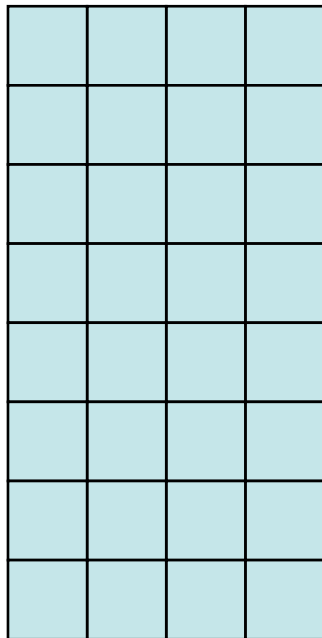
# Communication patterns

- Other operations: (orthogonalization, non-local potential energy, Ritz diagonalization)
  - use the ScaLAPACK linear algebra library



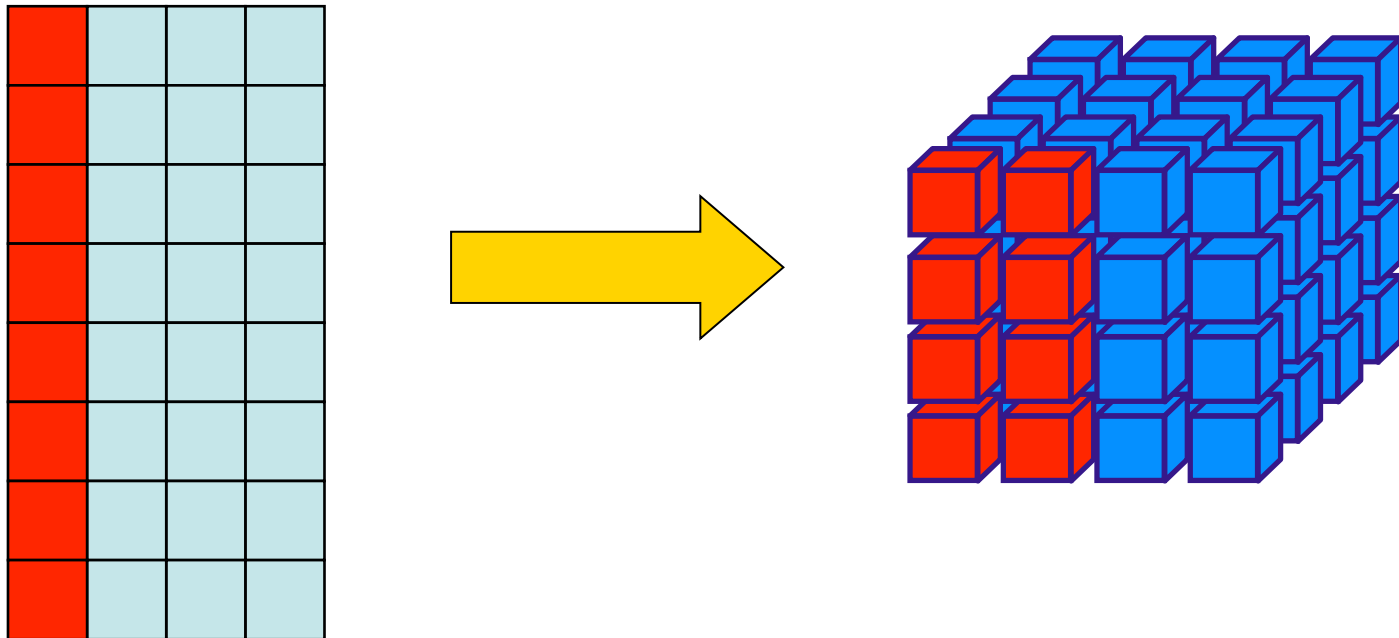
# Mapping tasks to physical nodes

- Mapping a 2-D process grid to a 3-D torus



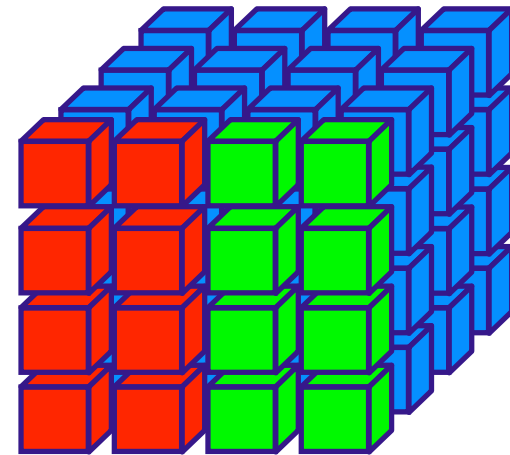
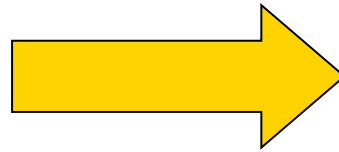
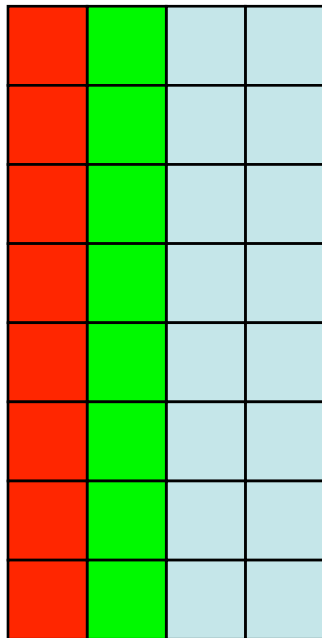
# Mapping tasks to physical nodes

- Mapping a 2-D process grid to a 3-D torus

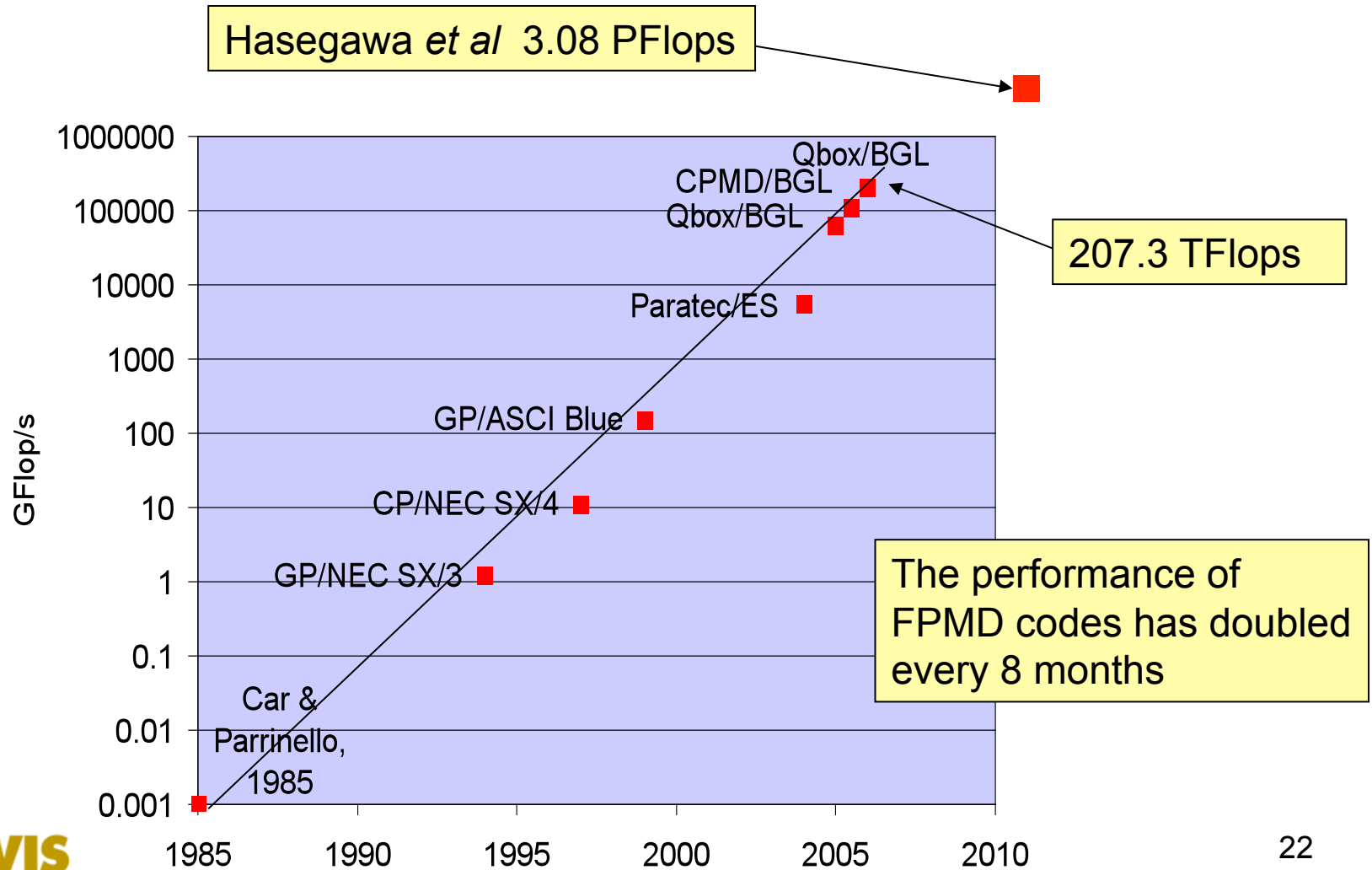


# Mapping tasks to physical nodes

- Mapping a 2-D process grid to a 3-D torus



# History of FPMD performance



# Designing scalable algorithms

- Reevaluate algorithms in terms of their communication patterns
- Example: Jacobi simultaneous diagonalization
  - involves two distinct communication patterns resulting in communication bottlenecks beyond 512 cores
- Using a one-sided Jacobi implementation
  - arithmetic cost is increased
  - simple communication pattern (Eberlein, Park (1990))
  - overall performance improved

# Limits of computational efficiency

- Time scales
  - 1 floating point operation (add/mult): 125 ps
  - fetch a double precision number from DRAM: 4 ns
  - fetch a double precision number from another node: >1  $\mu$ s
- The only viable strategy is to overlap communication and computation
  - reuse data in processor cache
  - use asynchronous communication
- Try to achieve  $t_{compute} > t_{communicate}$



# Limits of computational efficiency

- Fundamental characteristics
  - $F$  = Flop rate (e.g. AMD Opteron: 8 GFlop/s)
  - $B$  = Bandwidth (e.g. Infiniband: 1 GB/s)
  - $S$  = local data Size (e.g. matrix block size: 8 MB)
  - $N(S)$  = number of operations (e.g. matrix multiply:  $N(S) = S^{3/2}$ )

$$t_{\text{compute}} = \frac{N(S)}{F}$$

$$t_{\text{communicate}} = \frac{S}{B}$$

$$\frac{N(S)}{S} > \frac{F}{B}$$

# Limits of computational efficiency

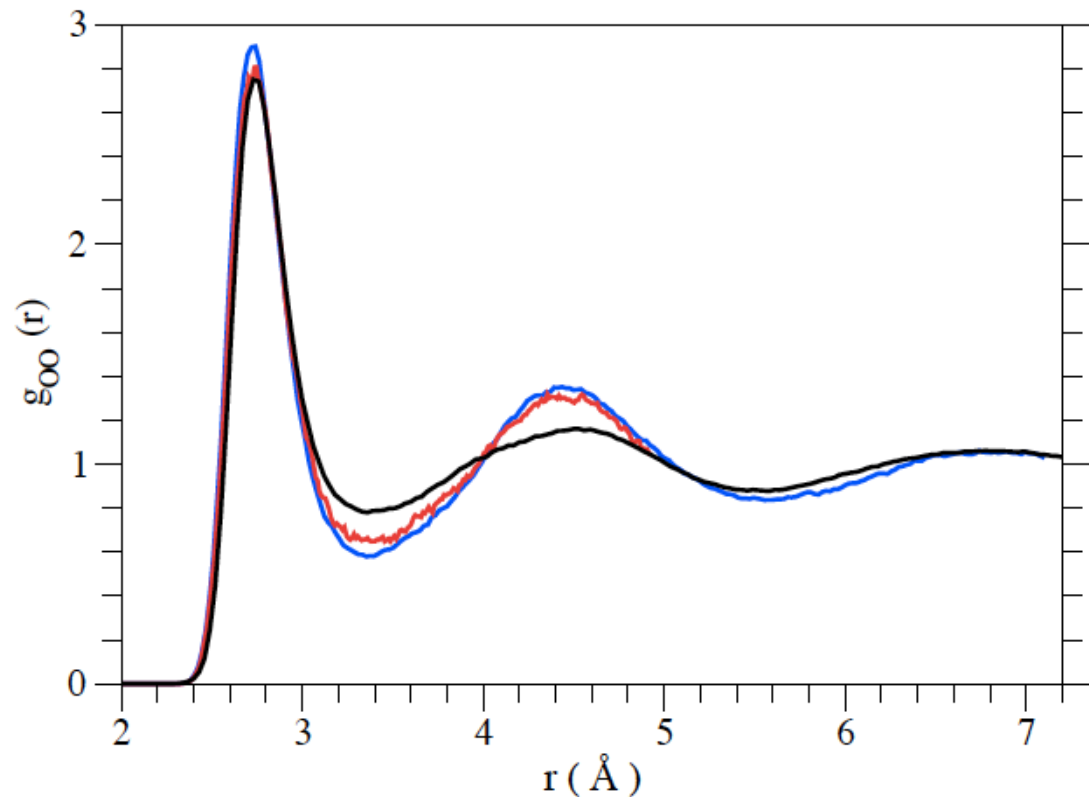
- Trends of new CPU architectures
  - Multicore processors: Flop rate ( $F$ ) increases
  - Bandwidth ( $B$ ) ~ constant

$$\frac{N(S)}{S} > \frac{F}{B}$$

- Only possible for very large problems (large  $S$ )
  - strong scaling is likely unachievable
- $N(S)$  must be superlinear

# Structure of DFT water

- PBE  $g_{00}(r)$ 
  - 32-mol ——— (red line)
  - 96 mol ——— (blue line)
- T=408 K
- A shift of ~100K approximately restores the main peak height



# Hybrid density functionals

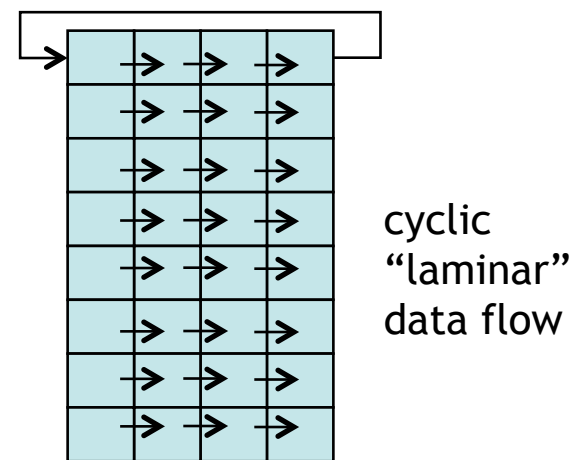
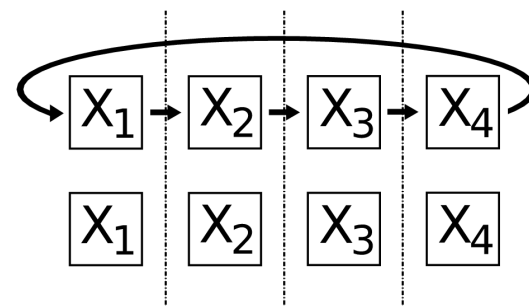
- Conventional density functionals may not be sufficiently accurate to describe weak bonds or optical properties
- *Hybrid density functionals* include a fraction of the Hartree-Fock exchange energy
- The Hartree-Fock exchange energy involves  $N(N+1)/2$  *exchange integrals* (for all e<sup>-</sup> pairs)

$$E_x = -\frac{1}{2} \sum_{i,j}^N \int \frac{\varphi_i^*(r_1)\varphi_i^*(r_2)\varphi_j(r_1)\varphi_j(r_2)}{|r_1 - r_2|} dr_1 dr_2$$

- Cost:  $O(N^3 \log N)$  (with large prefactor)

# Scalability of Hartree-Fock calculations

- The high cost of Hartree-Fock exchange calculations can be mitigated using parallel processes
- Compute exchange integrals for all pairs of wave functions using cyclic permutations of data
- Overlap communication and computation (non-blocking MPI calls)



2-d process grid

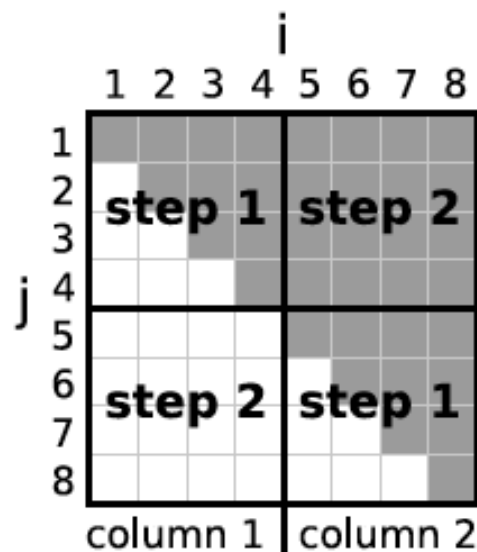
I. Duchemin, F. Gygi. *Comput. Phys. Comm.* **181**, 855 (2010)

E. Bylaska *et al.* *J. Comput. Chem.* **32**, 54 (2011)

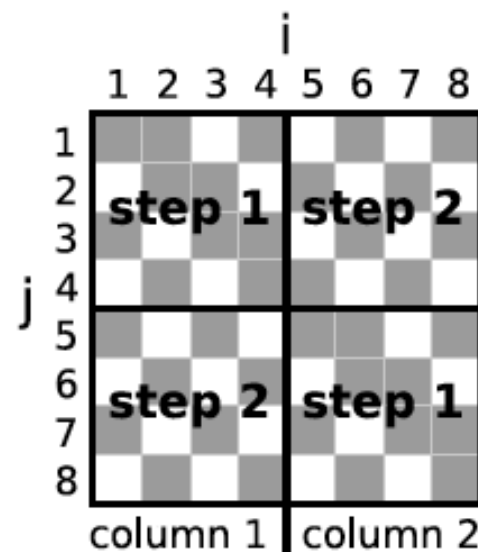
29

# Scalability of Hartree-Fock calculations

- Load-balancing of the computation of exchange integrals



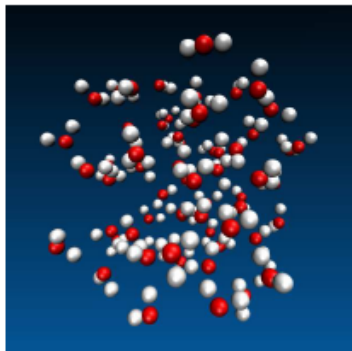
(a)



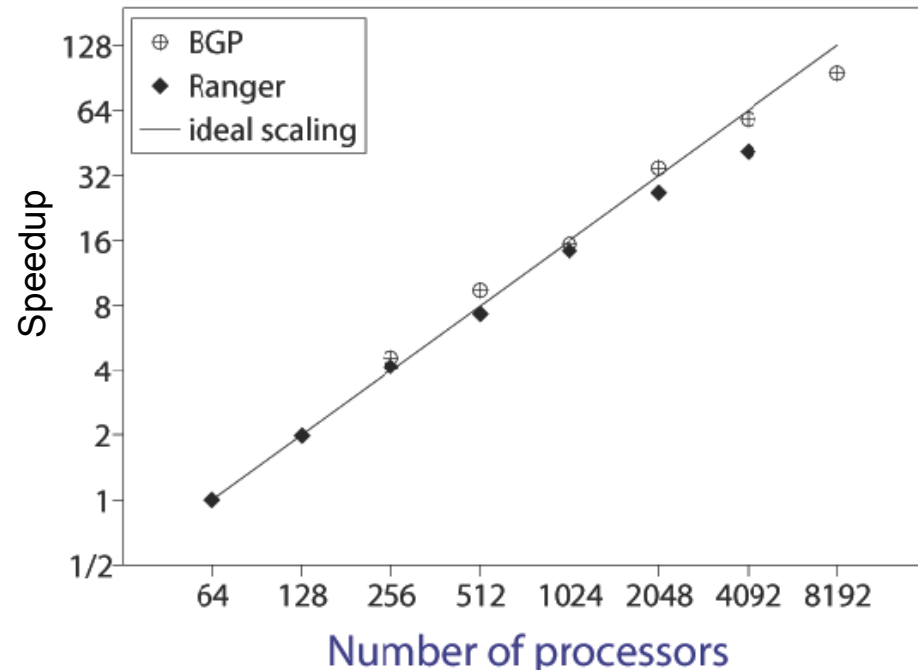
(b)

# Scalable Hartree-Fock and hybrid DFT implementation

- $(\text{H}_2\text{O})_{64}$  PBE0 hybrid DFT



Ranger :  
Sun Constellation Linux Cluster  
4 Quad. Opterons @ 2.3GHz /node  
BGP :  
IBM Blue Gene  
4 PowerPC @ 850MHz / node



# PBE0 water molecule & dimer

- Norm-conserving Hamann pseudopotentials, 85 - 200 Ry pw cutoff
- PBE-PBE0 comparison
- vibrational frequencies
  - dynamical matrix computed by finite differences of forces

C. Zhang, D. Donadio, F. Gygi, G. Galli, JCTC 7, 1443 (2011).

B. Santra, A. Michaelides, M. Scheffler, JCP 131, 124509 (2009).

Xu, X.; Goddard, W. A. J. Phys. Chem. A 108, 2305 (2004).



# PBE0 water molecule & dimer

DFT name	equation
	LDA
SVWN	$1.0E_x(\text{Slater}) + 1.0E_c(\text{VWN})$
	GGA
BLYP	$1.0E_x(\text{Slater}) + 1.0\Delta E_x(\text{B88}) + 1.0E_c(\text{LYP})$
BP86	$1.0E_x(\text{Slater}) + 1.0\Delta E_x(\text{B88}) + 1.0E_c(\text{PZ81,local}) + 1.0\Delta E_c(\text{P86,nonlocal})$
BPW91	$1.0E_x(\text{Slater}) + 1.0\Delta E_x(\text{B88}) + 1.0E_c(\text{PW91})$
PWPW	$1.0E_x(\text{Slater}) + 1.0\Delta E_x(\text{PW91}) + 1.0E_c(\text{PW91})$
mPWPW	$1.0E_x(\text{Slater}) + 1.0\Delta E_x(\text{mPW}) + 1.0E_c(\text{PW91})$
PBEPBE	$1.0E_x(\text{Slater}) + 1.0\Delta E_x(\text{PBE}) + 1.0E_c(\text{PW91,local}) + 1.0\Delta E_c(\text{PBE,nonlocal})$
XLYP	$1.0E_x(\text{Slater}) + 0.722\Delta E_x(\text{B88}) + 0.347\Delta E_x(\text{PW91}) + 1.0E_c(\text{LYP})$
	Hybrid Methods
BH&HLYP	$0.50E_x(\text{HF}) + 0.50E_x(\text{Slater}) + 0.50\Delta E_x(\text{B88}) + 1.0E_c(\text{LYP})$
B3LYP	$0.20E_x(\text{HF}) + 0.80E_x(\text{Slater}) + 0.72\Delta E_x(\text{B88}) + 0.19E_c(\text{VWN}) + 0.81E_c(\text{LYP})$
B3P86	$0.20E_x(\text{HF}) + 0.80E_x(\text{Slater}) + 0.72\Delta E_x(\text{B88}) + 1.0E_c(\text{VWN}) + 0.81\Delta E_c(\text{P86})$
B3PW91	$0.20E_x(\text{HF}) + 0.80E_x(\text{Slater}) + 0.72\Delta E_x(\text{B88}) + 1.0E_c(\text{PW91,local}) + 0.81\Delta E_c(\text{PW91,nonlocal})$
PW1PW	$0.25E_x(\text{HF}) + 0.75E_x(\text{Slater}) + 0.75\Delta E_x(\text{PW91}) + 1.0E_c(\text{PW91})$
mPW1PW	$0.25E_x(\text{HF}) + 0.75E_x(\text{Slater}) + 0.75\Delta E_x(\text{mPW91}) + 1.0E_c(\text{PW91})$
PBE1PBE	$0.25E_x(\text{HF}) + 0.75E_x(\text{Slater}) + 0.75\Delta E_x(\text{PBE}) + 1.0E_c(\text{PW91,local}) + 1.0\Delta E_c(\text{PBE,nonlocal})$

from: Xu, X.; Goddard, W. A. J. Phys. Chem. A 108, 2305 (2004).

# Water monomer (H<sub>2</sub>O)

	Ecut (Ry)	Pseudopotentials	v <sub>1</sub>	v <sub>2</sub>	v <sub>3</sub>	
PBEPBE <sup>1</sup>	-	AE	3702	1601	3804	aug-cc-pVTZ(-f)
PBEPBE <sup>2</sup>	-	AE	3697	1592	3801	aug-cc-pVTZ
PBE <sup>3</sup>	100	TM	3661	1606	3774	
PBE	85	Hamann	3617	1591	3736	
PBE	200	Hamann	3632	1592	3751	
PBE	85	HSCV	3686	1589	3797	
PBE	200	HSCV	3687	1592	3798	
PBE1PBE <sup>1</sup>	-	AE	3862	1643	3965	aug-cc-pVTZ(-f)
PBE0 <sup>4</sup>	-	AE	3856	1633	3962	aug-cc-pV5Z
PBE0	85	Hamann	3772	1633	3892	
PBE0	200	Hamann	3774	1638	3894	
PBE0	85	HSCV	3835	1635	3946	
PBE0	200	HSCV	3828	1639	3939	
Expt. Harm. <sup>5</sup>	-	-	3832	1648	3943	
Expt. Anharm. <sup>5</sup>	-	-	3657	1595	3756	

# PBE0 water dimer

- The dimer binding energy is reduced when using PBE0 (compared to PBE)
- $E_{\text{bind}}(\text{PBE0})=5.12$  kcal/mol
- $E_{\text{bind}}(\text{PBE})=5.24$  kcal/mol
- $E_{\text{bind}}(\text{PBE})-E_{\text{bind}}(\text{PBE0})=0.12$  kcal/mol
  - Santra *et al*: 0.15 kcal/mol
  - Xu *et al*: 0.13 kcal/mol
- Note:
  - BSSE corrections ~ 0.02 - 0.09 kcal/mol (Xu *et al*)
  - BSSE corrections ~ 0.33 - 0.36 kcal/mol (Todorova *et al*)

B. Santra, A. Michaelides, M. Scheffler, JCP 131, 124509 (2009).

Xu, X.; Goddard, W. A. J. Phys. Chem. A 108, 2305 (2004).

T. Todorova, et al, JPC B, 110, 3685 (2006)

# Water dimer

	Ecut (Ry)	Pseudopotentials	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$
PBEPBE <sup>1</sup>	-	AE	3695	1600	3795	3536	1621	3773
PBEPBE <sup>2</sup>	-	AE	3691	1591	3792	3530	1610	3770
PBE	100	TM	3656	1602	3770	3485	1625	3738
PBE	85	Hamann	3611	1590	3727	3443	1612	3699
PBE	200	Hamann	3628	1591	3743	3468	1613	3715
PBE	85	HSCV	3679	1590	3786	3521	1611	3762
PBE	200	HSCV	3681	1592	3788	3522	1612	3764
PBE1PBE <sup>1</sup>	-	AE	3853	1644	3952	3720	1665	3935
PBE0	85	Hamann	3767	1633	3884	3627	1656	3860
PBE0	200	Hamann	3770	1638	3886	3637	1661	3862
PBE0	85	HSCV	3830	1634	3938	3694	1656	3917
PBE0	200	HSCV	3813	1638	3931	3685	1660	3910
Expt. Harm. <sup>11</sup>	-	-	3797	1653	3899	3718	1669	3881
Expt. Anharm. <sup>12</sup>	-	-	3626	1600	3714	3548	1618	3698

# Water dimer

	Ecut (Ry)	Pseudopotentials	$v_7$	$v_8$	$v_9$	$v_{10}$	$v_{11}$	$v_{12}$
PBEPBE <sup>1</sup>	-	AE	651	387	194	180	159	133
PBEPBE <sup>2</sup>	-	AE	646	380	190	158	145	124
PBE	85	Hamann	660	388	209	176	171	137
PBE	200	Hamann	660	388	205	174	170	138
PBE	85	HSCV	650	383	225	187	172	165
PBE	200	HSCV	648	381	211	171	167	146
PBE1PBE <sup>1</sup>	-	AE	654	389	199	164	163	142
PBE0	85	Hamann	661	384	208	172	165	137
PBE0	200	Hamann	668	394	212	194	175	160
PBE0	85	HSCV	655	377	209	167	160	144
PBE0	200	HSCV	665	396	218	202	176	169
Expt. Anharm. <sup>13</sup>	-	-	522	309	173	151	122	116

# PBE0 liquid water ( $D_2O$ )

- PBE0 liquid water simulations
- 32 molecules, norm-conserving PP, 85 Ry
- 17 ps simulations

T. Todorova, A.P. Seitsonen, J. Hutter, I.F.W. Kuo, C.J.Mundy, JPC B, 110, 3685 (2006).

M. Guidon, F. Schiffmann, J. Hutter, J. VandeVondele, JCP 128, 214104 (2008)

# Computation of Hartree-Fock exchange in extended systems

- The long range of the Coulomb potential leads to an integrable divergence in Brillouin zone integrals
- Neglecting this singularity leads to significant errors in computed band gaps

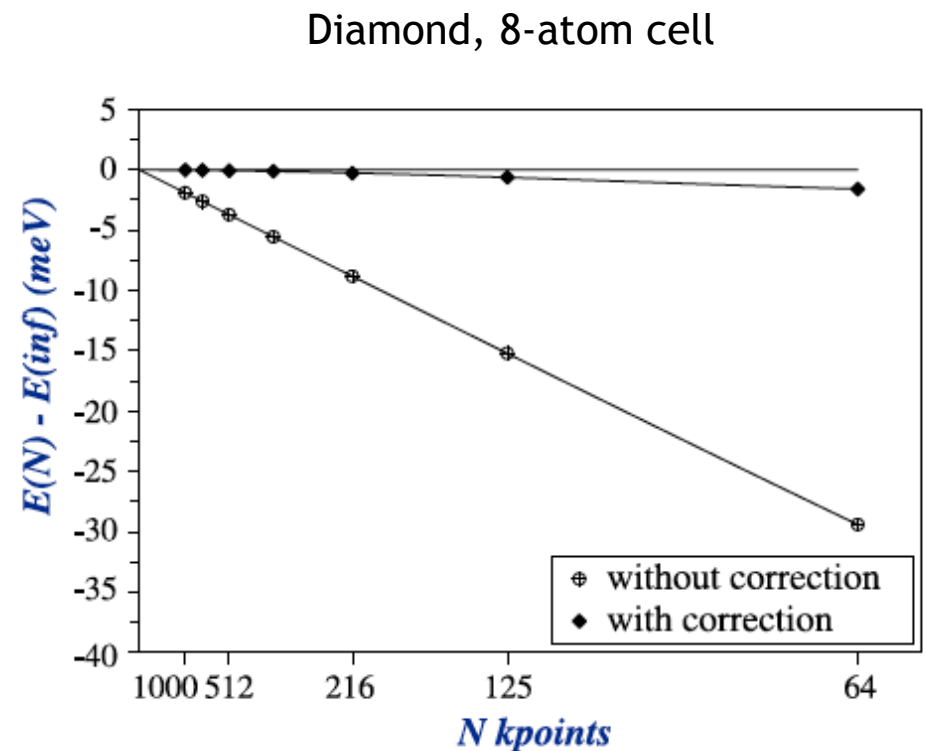
$$E_{ik} = \int_{BZ} \sum_{\mathbf{G}} \sum_j \frac{|\rho_{ijkk'}(\mathbf{G})|^2}{|\mathbf{G} + \mathbf{k} - \mathbf{k}'|^2} d\mathbf{k}'$$

$$\rho_{ijkk'}(\mathbf{q}) = \int u_{jk'}^*(\mathbf{r}) u_{ik}(\mathbf{r}) e^{i\mathbf{q}\mathbf{r}} d\mathbf{r}$$

I. Duchemin, F. Gygi, *Comp. Phys. Comm.* **181**, 855 (2010) 855-860

# Computation of Hartree-Fock exchange in extended systems

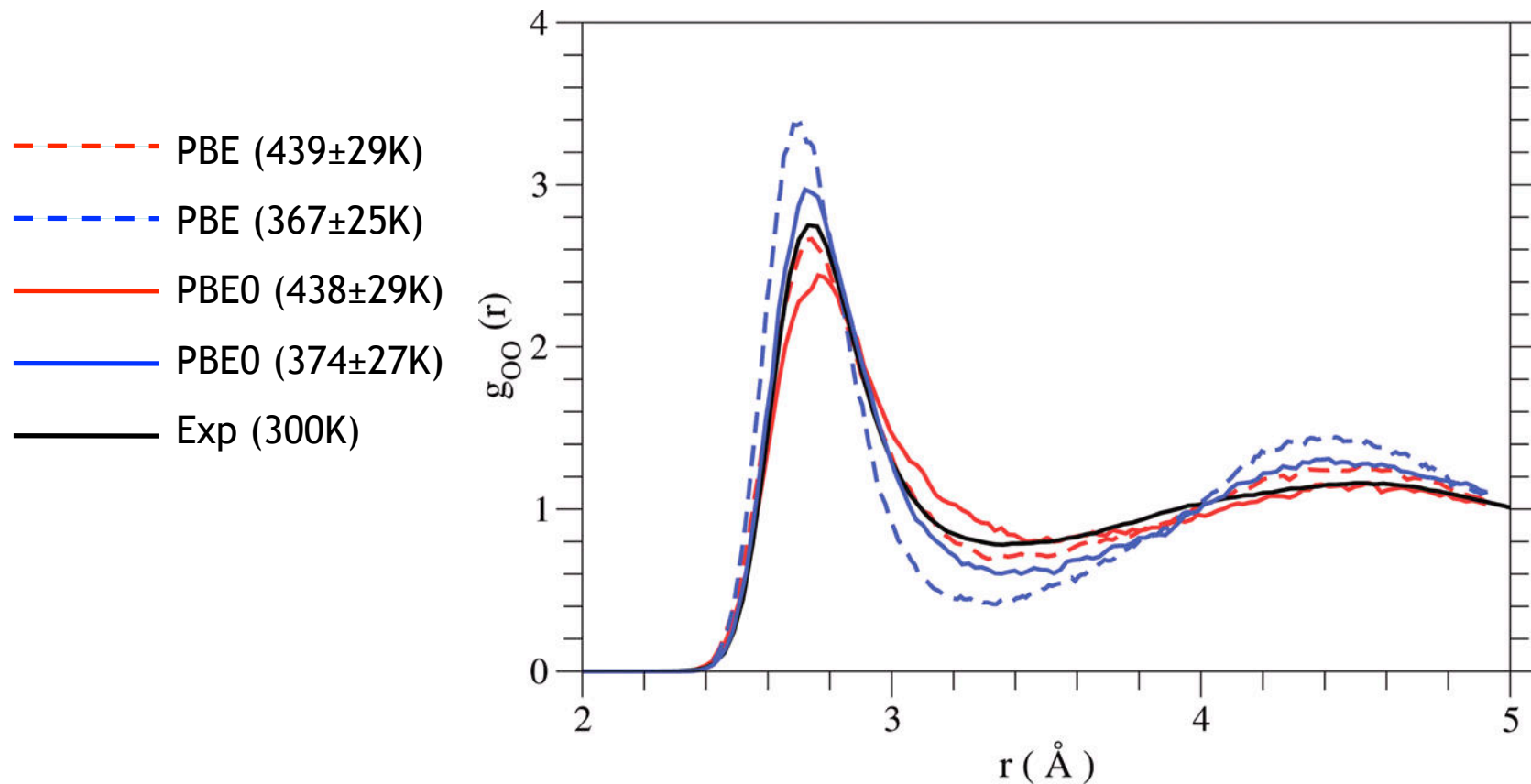
- The divergent term can be accounted for using a simple analytic regularization of the divergence
- Quadratic convergence of Brillouin zone sampling is achieved





# PBE0 water ( $D_2O$ )

- PBE0 water is less structured than PBE water



# Computation of IR spectra in liquids

- Total dipole in the condensed phase is not well defined
- Molecular dipole defined in terms of Maximally Localized Wannier functions (MLWFs)
  - Wannier “centers” are the center of charge of each MLWF
- IR absorption coefficient

$$\alpha(\omega) = \frac{2\pi\omega^2\beta}{3cVn(\omega)} \int_{-\infty}^{\infty} dt e^{-i\omega t} \langle \sum_{ij} \vec{\mu}_i(0) \cdot \vec{\mu}_j(t) \rangle$$

# Maximally Localized Wannier Functions (MLWFs)

- Maximally Localized Wannier Functions (MLWF): linear combinations of electronic orbitals that minimize the spatial spread

$$\sigma_{\hat{x}}^2 = \langle x - \langle x \rangle \rangle^2$$

- MLWFs can be used to compute the electronic polarization in crystals
- Computing MLWFs during a molecular dynamics simulation yields the infrared absorption spectrum

N. Marzari and D. Vanderbilt, Phys. Rev. B56, 12847 (1997)

R. Resta, Phys. Rev. Lett. 80, 1800 (1998)

# Spread Functionals

- Spread of a wavefunction associated with an operator  $\hat{A}$

$$\begin{aligned}\sigma_{\hat{A}}^2(\phi) &= \left\langle \phi \left| \left( \hat{A} - \langle \phi | \hat{A} | \phi \rangle \right)^2 \right| \phi \right\rangle \\ &= \left\langle \phi \left| \hat{A}^2 \right| \phi \right\rangle - \left\langle \phi \left| \hat{A} \right| \phi \right\rangle^2\end{aligned}$$

- Spread of a set of wavefunctions associated with an operator  $\hat{A}$

$$\sigma_{\hat{A}}^2(\{\phi_i\}) = \sum_i \sigma_{\hat{A}}^2(\phi_i)$$

# Spread Functionals

- The spread is *not* invariant under orthogonal transformations

$$\psi_i = \sum_j x_{ij} \phi_j \quad X \in \mathbb{R}^{n \times n} \text{ orthogonal}$$

$$\sigma_{\hat{A}}^2(\{\psi_i\}) \neq \sigma_{\hat{A}}^2(\{\phi_i\})$$

- There exists a matrix  $X$  that minimizes the spread

# Spread Functionals

- Let

$$A, B \in \mathbb{R}^{n \times n} \quad a_{ij} = \langle i | \hat{A} | j \rangle \quad b_{ij} = \langle i | \hat{A}^2 | j \rangle$$

$$\sigma_{\hat{A}}^2(\{\psi_i\}) = \text{tr}(X^T B X) - \sum_{i=1}^n (X^T A X)_{ii}^2$$

- Minimize the spread = maximize  $\sum_{i=1}^n (X^T A X)_{ii}^2$   
= diagonalize  $A$

# Spread Functionals

- Case of multiple operators

operators  $\hat{A}^{(k)}$   $k = 1, \dots, m$

matrices  $A^{(k)}$   $k = 1, \dots, m$

$$\sigma_{\hat{A}}^2(\{\psi_i\}) = \sum_i \sum_k \sigma_{\hat{A}^{(k)}}^2(\psi_i)$$

- Minimize the spread = maximize  $\sum_{i=1}^n \sum_k (X^T A^{(k)} X)_{ii}^2$   
= joint approximate diagonalization of the matrices  $A^{(k)}$

# Spread Functionals

- Example of multiple operators

$$\hat{A}^{(1)} = \hat{X} \quad (\hat{X}\varphi)(x, y, z) \equiv x\varphi(x, y, z)$$

$$\hat{A}^{(2)} = \hat{Y} \quad (\hat{Y}\varphi)(x, y, z) \equiv y\varphi(x, y, z)$$

$$\hat{A}^{(3)} = \hat{Z} \quad (\hat{Z}\varphi)(x, y, z) \equiv z\varphi(x, y, z)$$

- The matrices  $A^{(k)}$  do not necessarily commute, even if the operators  $\hat{A}^{(k)}$  do commute



# Calculation of MLWFs

- In periodic systems

$$\begin{aligned}\hat{A}^{(1)} &= \hat{C}_x \equiv \cos \frac{2\pi}{L_x} \hat{x} & \hat{A}^{(2)} &= \hat{S}_x \equiv \sin \frac{2\pi}{L_x} \hat{x} \\ \hat{A}^{(3)} &= \hat{C}_y \equiv \cos \frac{2\pi}{L_y} \hat{y} & \hat{A}^{(4)} &= \hat{S}_y \equiv \sin \frac{2\pi}{L_y} \hat{y} \\ \hat{A}^{(5)} &= \hat{C}_z \equiv \cos \frac{2\pi}{L_z} \hat{z} & \hat{A}^{(6)} &= \hat{S}_z \equiv \sin \frac{2\pi}{L_z} \hat{z}\end{aligned}$$

F.Gygi, J.L.Fattebert, E.Schwegler, Comput.. Phys.  
Comm. **155**, 1 (2003)

# Calculation of MLWFs

- The spread is minimized by simultaneous diagonalization of the matrices  $C_x, S_x, C_y, S_y, C_z, S_z$
- Positions of the center of mass of the localized solutions (“Wannier centers”)

$$\tau_i = \begin{pmatrix} L_x \frac{\theta_i^x}{2\pi} \\ L_y \frac{\theta_i^y}{2\pi} \\ L_z \frac{\theta_i^z}{2\pi} \end{pmatrix} \quad \theta_i^x = \arctan \frac{(S_x)_{ii}}{(C_x)_{ii}}$$

- Spreads

$$(\sigma_i^2)_x = L_x^2 \left( 1 - (C_x)_{ii}^2 - (S_x)_{ii}^2 \right)$$

# Simultaneous diagonalization

- Jacobi algorithm for simultaneous diagonalization of symmetric matrices

```
repeat
  for each pair  $i, j$ 
    compute Jacobi rotation  $R(i, j)$ 
     $A^{(k)} \leftarrow R^T A^{(k)} R \quad \forall k$ 
until converged
```

A. Bunse-Gerstner, R. Byers, V. Mehrmann, SIAM J. Mat. Anal. Appl. **14**, 927 (1993)

J.F.Cardoso and A. Souloumiac, SIAM J. Mat. Anal. Appl. **17**, 161 (1996).

# Simultaneous diagonalization

$$R(i, j) = \begin{pmatrix} r_{ii} & r_{ij} \\ r_{ji} & r_{jj} \end{pmatrix} = \begin{pmatrix} c & \bar{s} \\ -s & \bar{c} \end{pmatrix}$$

$$c, s \in \mathbb{C}$$

$$|c|^2 + |s|^2 = 1$$

$$c = \sqrt{\frac{x+r}{2r}} \quad s = \frac{y-iz}{\sqrt{2r(x+r)}} \quad r = \sqrt{x^2 + y^2 + z^2}$$

# Simultaneous diagonalization

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ eigenvector of } G \equiv \operatorname{Re} \left( \sum_k h(A^{(k)}) h^H(A^{(k)}) \right)$$

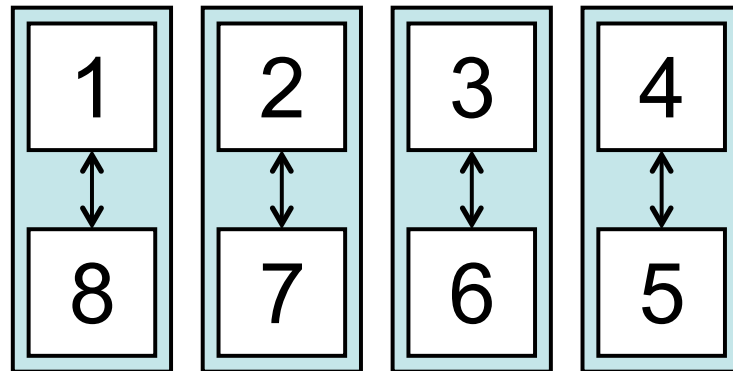
$$h(A) = \begin{pmatrix} a_{ii} - a_{jj} \\ a_{ij} + a_{ji} \\ i(a_{ji} - a_{ij}) \end{pmatrix}$$

J.F.Cardoso and A. Souloumiac, SIAM J. Mat. Anal. Appl. **17**, 161 (1996).

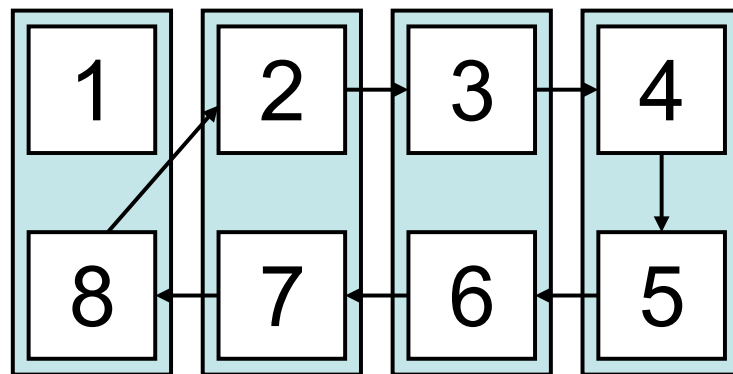
# Simultaneous diagonalization algorithm

- The cost of simultaneous diagonalization is  $O(N^3)$ 
  - a scalable algorithm is desirable
- Conventional implementations scale up to  $N_{CPU} \sim N/2$
- scalable algorithm based on the one-sided Jacobi algorithm (P.J. Eberlein SIAM J. Alg. Disc. Meth. (1987))

# Conventional Parallel Jacobi algorithm



Each processor holds a pair of columns



G. H. Golub and C. F. Van Loan, Matrix Computations, 3<sup>rd</sup> ed., (1996)

Limited to  $n/2$  processors for  $n$  columns

# One-sided Jacobi algorithms

Conventional Jacobi

```
repeat  
 $A \leftarrow R^T AR$   
 $U \leftarrow UR$   
until converged
```

One-sided Jacobi

```
repeat  
 $A \leftarrow AR$   
 $U \leftarrow UR$   
until converged
```

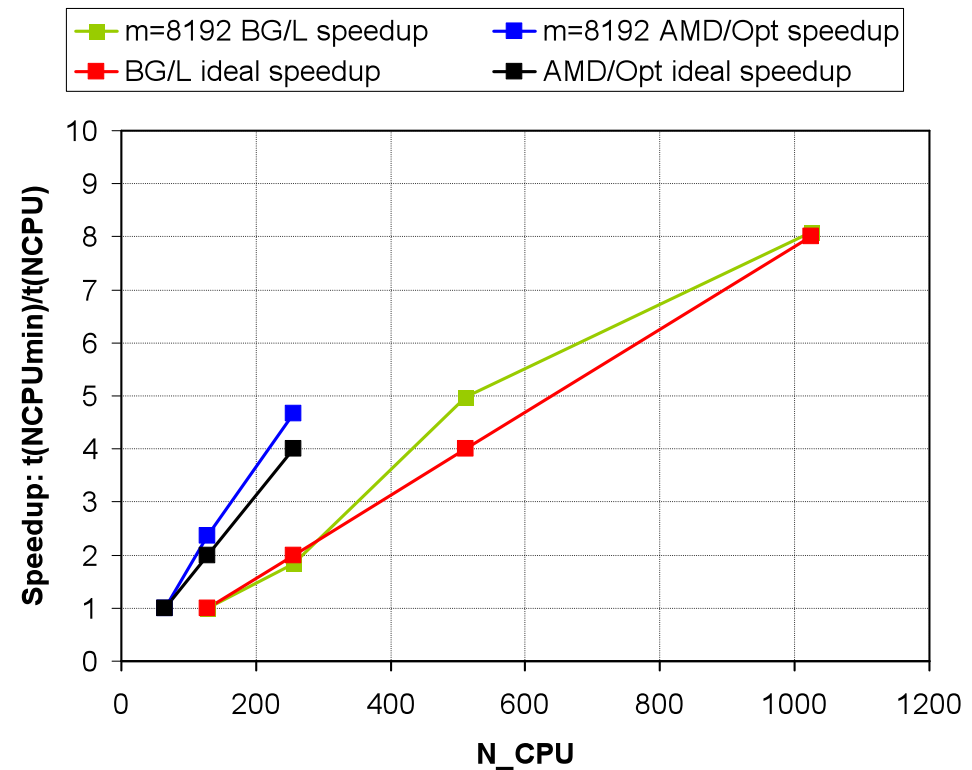
P.J. Eberlein, H. Park, J. Parallel and Distrib. Comp. 8 (1990), 358–366.

The one-sided Jacobi algorithm reduces the amount of global communication



# Scalability of simultaneous diagonalization

- One-sided Jacobi simultaneous diagonalization algorithm:



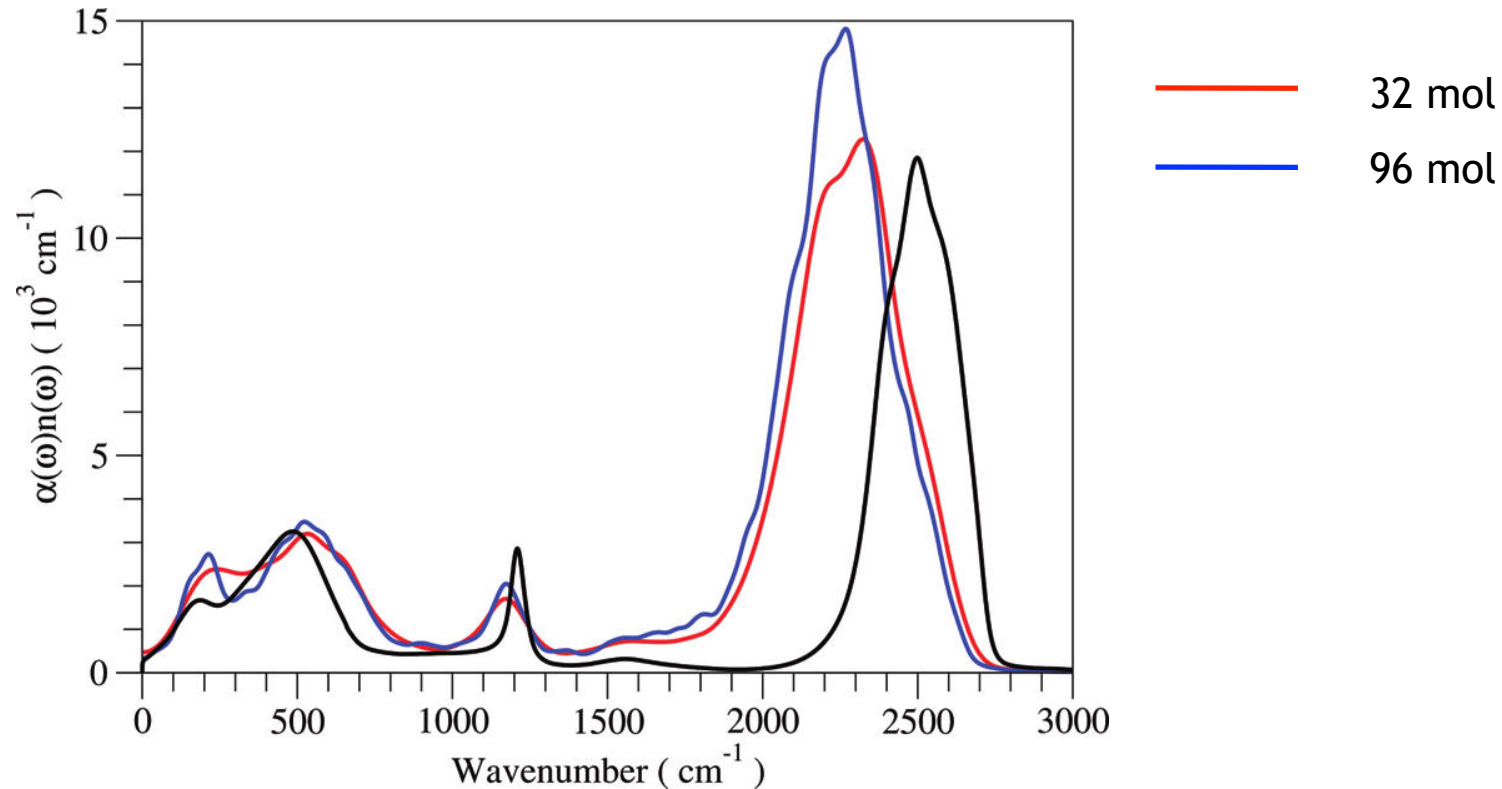
# Examples of MLWF calculations

- One-sided Jacobi simultaneous diagonalization algorithm included in Qbox
- Performance data: time to compute MLWFs starting from extended Kohn-Sham eigenstates (worst case)
  - 256 AMD 2.0 GHz cores

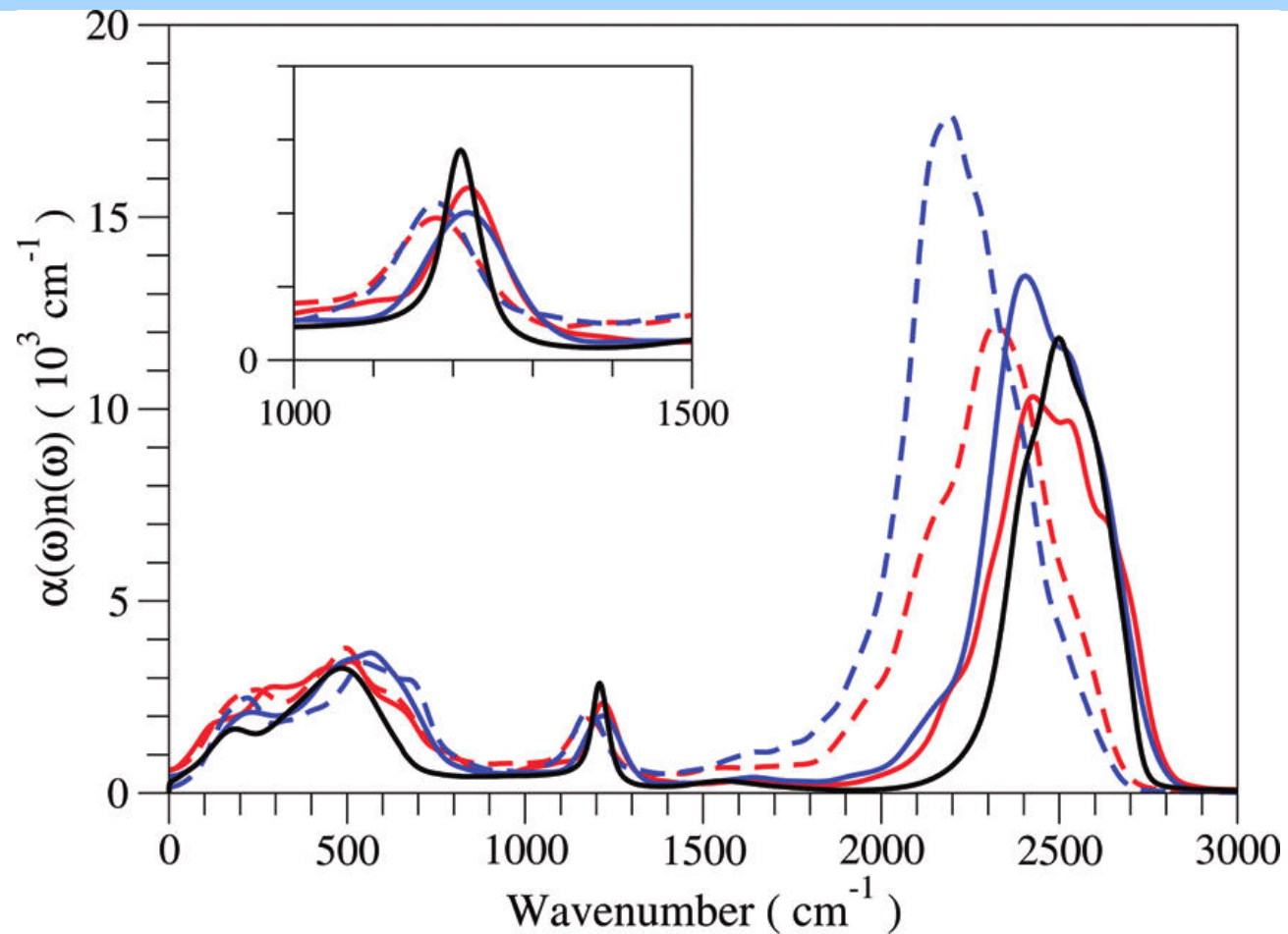
System	$N_{el}$	$E_{cut}$ (Ry)	MLWF time (s)
(HfSiO <sub>4</sub> ) <sub>32</sub>	1216	70	38
(H <sub>2</sub> O) <sub>512</sub>	4096	72	440

# Water IR spectra (D<sub>2</sub>O)

- PBE IR spectrum from MLWFs at T=407 K



# Water IR spectra (D<sub>2</sub>O)



PBE0: 438 K (solid red), 374 K (solid blue)

PBE: 439 K (dash red), 367 K (dash blue).

# PBE0 water vs PBE water

- PBE0 yields a less structured liquid
- PBE0 yields smaller dipole moments
- PBE0 IR spectrum in better agreement with experiment than PBE

# Van der Waals density functionals

- Dion-Rydberg-Schroeder-Langreth-Lundqvist (DRSLL)
- Lee-Murray-Kong-Lundqvist-Langreth (LMKLL)
- Klimes-Bowler-Michaelides (optB88)
- Wang-Roman-Perez-Soler-Artacho-Fernandez-Serra (DRSLLPBE)

# Dion *et al.* (2004) (DRSLL)

$$E_{xc}^{vdW} = E_x^{GGA} + E_c^{LDA} + E_c^{nl}$$

$$E_c^{nl} = \iint d\mathbf{r} d\mathbf{r}' \rho(\mathbf{r})\rho(\mathbf{r}')\Phi(\mathbf{r}, \mathbf{r}'),$$

$$\Phi(\mathbf{r}, \mathbf{r}') = \phi(d, d').$$

$$\phi(d, d') = \phi(q |\mathbf{r} - \mathbf{r}'|, q' |\mathbf{r} - \mathbf{r}'|)$$

$$q = q_0(r)$$

$$q' = q_0(r')$$

$$q_0(\mathbf{r}) = -\frac{4\pi}{3}\varepsilon_{xc}^{LDA}(\rho(\mathbf{r})) - \frac{Z_{ab}}{9}s^2(\mathbf{r})k_F(\mathbf{r}).$$

# Reduction of the cost to $O(N)$ (Soler)

$$\begin{aligned}\phi(d, d') &= \phi(q |\mathbf{r} - \mathbf{r}'|, q' |\mathbf{r} - \mathbf{r}'|) \\ &\simeq \sum_{\alpha, \beta=0}^{N_q-1} \phi(q_\alpha |\mathbf{r} - \mathbf{r}'|, q_\beta |\mathbf{r} - \mathbf{r}'|) p_\alpha(q) p_\beta(q'),\end{aligned}$$

$$\theta_\alpha(\mathbf{r}) = \rho(\mathbf{r}) p_\alpha(q_0(\mathbf{r}))$$



# Reduction of the cost to $O(N)$ (Soler)

$$E_c^{\text{nl}} = \frac{1}{2} \sum_{\alpha, \beta=0}^{N_q-1} \iint d\mathbf{r} d\mathbf{r}' \theta_\alpha(\mathbf{r}) \theta_\beta(\mathbf{r}') \phi_{\alpha\beta}(|\mathbf{r} - \mathbf{r}'|),$$

$$E_c^{\text{nl}} = \frac{\Omega}{2} \sum_{\alpha, \beta=0}^{N_q-1} \sum_{\mathbf{G}} \theta_\alpha^*(\mathbf{G}) \theta_\beta(\mathbf{G}) \phi_{\alpha\beta}(|\mathbf{G}|),$$

$$\theta_\alpha(\mathbf{G}) = \int d\mathbf{r} \theta_\alpha(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}},$$

$$\phi_{\alpha\beta}(k) = \frac{4\pi}{k} \int_0^\infty r \phi_{\alpha\beta}(r) \sin kr dr.$$

# vdW implementation

$$q q' \phi(q |\mathbf{r} - \mathbf{r}'|, q' |\mathbf{r} - \mathbf{r}'|)$$

$$\simeq \sum_{\alpha, \beta=0}^{N_q-1} q_\alpha q_\beta \phi(q_\alpha |\mathbf{r} - \mathbf{r}'|, q_\beta |\mathbf{r} - \mathbf{r}'|) p_\alpha(q) p_\beta(q').$$

$$q_0(\mathbf{r}) = -\frac{4\pi}{3} \varepsilon_{xc}^{\text{LDA}}(\rho(\mathbf{r})) - \frac{Z_{ab}}{9} s^2(\mathbf{r}) k_F(\mathbf{r}),$$

$$\phi(q |\mathbf{r} - \mathbf{r}'|, q' |\mathbf{r} - \mathbf{r}'|)$$

$$\simeq \sum_{\alpha, \beta=0}^{N_q-1} \phi(q_\alpha |\mathbf{r} - \mathbf{r}'|, q_\beta |\mathbf{r} - \mathbf{r}'|) \frac{q_\alpha p_\alpha(q)}{q} \frac{q_\beta p_\beta(q')}{q'}.$$

# vdW implementation

$$E_c^{\text{nl}} = \frac{1}{2} \sum_{\alpha\beta} \iint d\mathbf{r} d\mathbf{r}' \eta_\alpha(\mathbf{r}) \eta_\beta(\mathbf{r}') \phi_{\alpha\beta}(|\mathbf{r} - \mathbf{r}'|)$$

$$= \frac{\Omega}{2} \sum_{\mathbf{G}} \sum_{\alpha\beta} \eta_\alpha^*(\mathbf{G}) \eta_\beta(\mathbf{G}) \phi_{\alpha\beta}(\mathbf{G}),$$

$$\eta_\alpha(\mathbf{r}) = \frac{q_\alpha \rho(\mathbf{r}) p_\alpha(q_0(\mathbf{r}))}{q_0(\mathbf{r})},$$

# vdW implementation

$$\begin{aligned}\sigma_{ij}^{\text{nl-c}} &= -\frac{1}{\Omega} \frac{\partial E_c^{\text{nl}}}{\partial u_{ij}} = -\frac{\delta_{ij}}{\Omega} \int \left[ \varepsilon_c^{\text{nl}}(\rho(\mathbf{r})) - v_{\text{nl-c}}^{(1)}(\mathbf{r}) \right] \rho(\mathbf{r}) d\mathbf{r} \\ &\quad - \frac{v_{\text{nl-c}}^{(2)}}{\Omega} \int \left[ \delta_{ij} |\nabla \rho|^2 + (\nabla \rho)_i (\nabla \rho)_j \right] d\mathbf{r} \\ &\quad - \frac{1}{2\Omega} \sum_{\alpha, \beta=0}^{N_q-1} \int \int d\mathbf{r} d\mathbf{r}' \eta_\alpha(\mathbf{r}) \eta_\beta(\mathbf{r}') \frac{d\phi_{\alpha\beta}(|\mathbf{r} - \mathbf{r}'|)}{du_{ij}},\end{aligned}$$

# vdW implementation

$$v_{\text{nl-c}}^{(1)} = \sum_{\alpha=0}^{N_q-1} u_{\alpha}(\mathbf{r}) \frac{\delta \eta_{\alpha}(\mathbf{r})}{\delta \rho},$$

$$v_{\text{nl-c}}^{(2)} = -\frac{1}{|\nabla \rho(\mathbf{r})|} \sum_{\alpha=0}^{N_q-1} u_{\alpha}(\mathbf{r}) \frac{\delta \eta_{\alpha}(\mathbf{r})}{\delta |\nabla \rho|},$$

# vdW implementation

$$\begin{aligned} u_{\alpha}(\mathbf{r}) &= \sum_{\beta=0}^{N_q-1} \int d\mathbf{r}' \eta_{\beta}(\mathbf{r}') \phi_{\alpha\beta}(|\mathbf{r} - \mathbf{r}'|) \\ &= \sum_{\mathbf{G}} \left( \sum_{\beta=0}^{N_q-1} \eta_{\beta}(\mathbf{G}) \phi_{\alpha\beta}(\mathbf{G}) \right) e^{i\mathbf{G}\cdot\mathbf{r}} \end{aligned}$$

# vdW implementation

$$\begin{aligned}\sigma_{ij}^{\text{nl-c}} &= -\frac{\delta_{ij}}{\Omega} \int \left[ \varepsilon_c^{\text{nl}}(\rho(\mathbf{r})) - v_{\text{nl-c}}^{(1)}(\mathbf{r}) \right] \rho(\mathbf{r}) d\mathbf{r} \\ &\quad - \frac{v_{\text{nl-c}}^{(2)}}{\Omega} \int \left[ \delta_{ij} |\nabla \rho|^2 + (\nabla \rho)_i (\nabla \rho)_j \right] d\mathbf{r} \\ &\quad + \frac{1}{2} \sum_{\mathbf{G}} \sum_{\alpha, \beta=0}^{N_q-1} \eta_{\alpha}^*(\mathbf{G}) \eta_{\beta}(\mathbf{G}) \frac{d\phi_{\alpha\beta}(|\mathbf{G}|)}{d|\mathbf{G}|} \frac{G_i G_j}{|\mathbf{G}|}, \\ \frac{\partial |\mathbf{G}|}{\partial u_{ij}} &= -\frac{G_i G_j}{|\mathbf{G}|}.\end{aligned}$$

# vdW implementation

$$\phi_{\alpha\beta}(|\mathbf{G}|) = \frac{1}{q_{\alpha}^3} F_{\alpha\beta} \left( \frac{|\mathbf{G}|}{q_{\alpha}} \right).$$

$$F_{\alpha\beta}(k) = \frac{4\pi}{k} \int_0^{\infty} u \phi \left( u, \frac{q_{\beta}}{q_{\alpha}} u \right) \sin(ku) du,$$



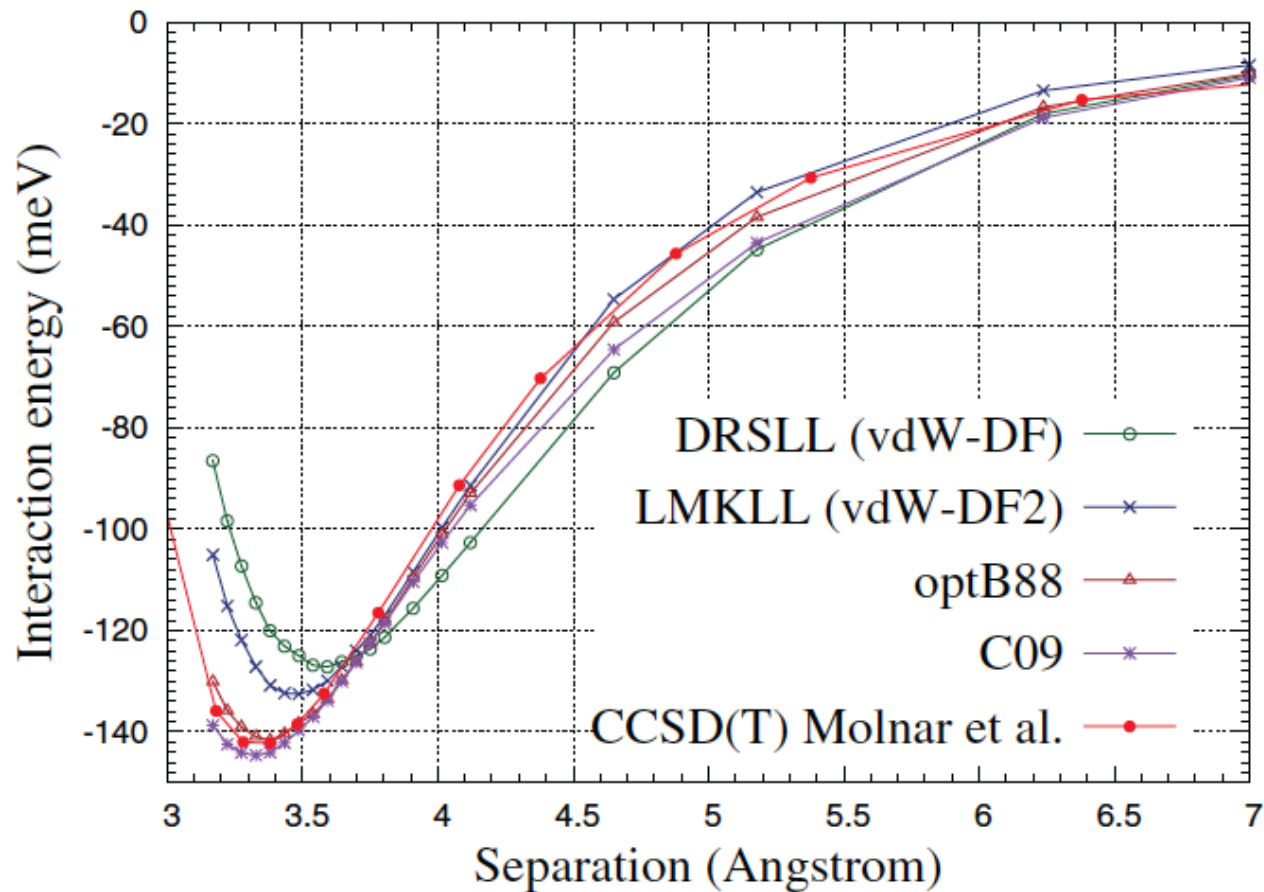
# Water-benzene complex

Functional		$d_{\min}$ (Å)	$E_{\text{bind}}$ (meV)
DRSLL	This work	3.59	127
	Ref. 5	3.58	124
	Ref. 21	3.55	120
LMKLL	This work	3.48	133
	Ref. 5	3.44	129
optB88	This work	3.37	142
	Ref. 21	3.40	135

Ref 5: Lee *et al.* (2010)

Ref 21: J. Klimes (private comm.)

# Water-benzene complex



# Pbca benzene

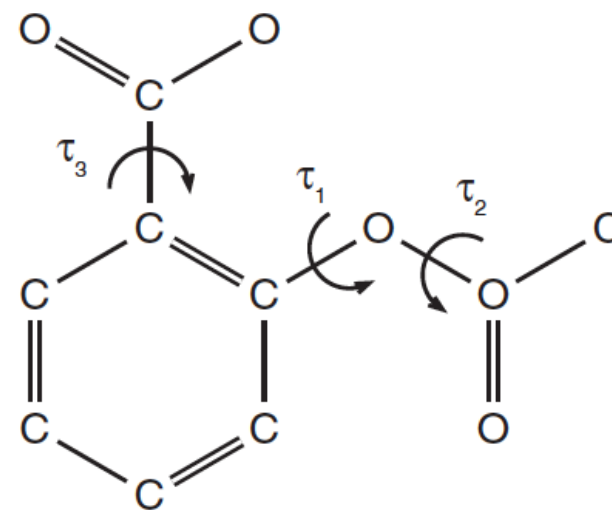
	$E_{\text{coh}}$ (kJ/mol)	$a$ (Å)	$b$ (Å)	$c$ (Å)	Volume (Å <sup>3</sup> )	$B_0$ (GPa)
LDA	52.1	7.063	9.040	6.369	406.7	13.9
PBE	9.3	7.947	9.975	7.232	573.3	4.2
DRSLL	57.5	7.669	9.725	6.926	516.6	10.3
LMKLL	56.3	7.459	9.477	6.763	478.1	12.6
optB88	67.9	7.375	9.407	6.529	453.0	14.5
C09	69.2	7.183	9.352	6.541	439.3	14.1
DRSLLPBE	85.6	7.371	9.418	6.655	462.0	30.1
Expt.	52.3 <sup>a</sup>	7.351 <sup>b</sup>	9.364 <sup>b</sup>	6.695 <sup>b</sup>	460.8 <sup>b</sup>	8 <sup>a</sup>

J. Wu and F. Gygi, JCP **136**, 224107 (2012)

E.J.Meijer and M. Sprik, JCP 105, 8684 (1996).

# Aspirin

- Crystal structure involves both hydrogen bonds and vdW interactions
- Monoclinic structure  $P2_1/c$
- New polymorph discovered in 2005



C. Ouvrard and S. L. Price, *Cryst. Growth Des.* **4**, 1119 (2004).

P. Vishweshwar, J. A. McMahon, M. Oliveira, M. L. Peterson, and M. J.

Zaworotko, *J. Am. Chem. Soc.* **127**, 16802 (2005).

# Aspirin structure ( $P2_1/c$ )

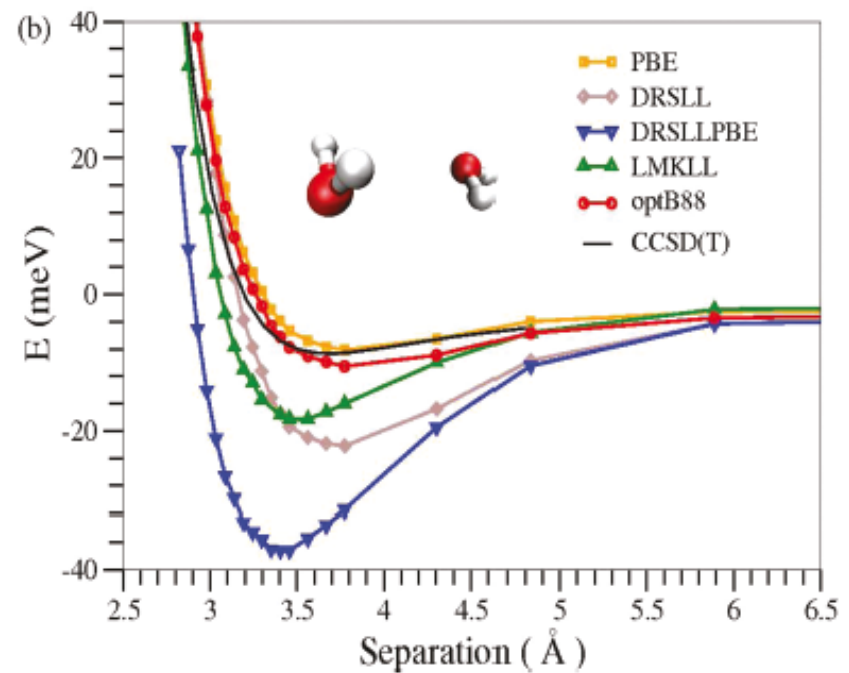
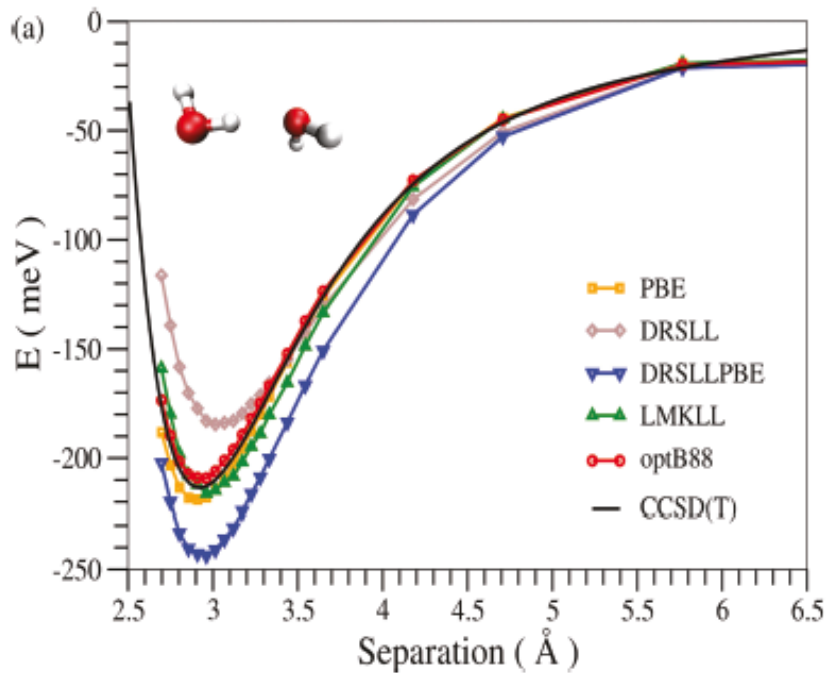
	Form	$E_{\text{coh}}$ (kJ/mol)	$a$ (Å)	$b$ (Å)	$c$ (Å)	Volume (Å <sup>3</sup> )	$\beta$ (deg)
PBE	I	49.1	12.020	6.509	11.970	935.30	90.1
	II	48.6	13.765	6.174	12.442	954.46	115.5
optB88	I	154.5	11.168	6.671	11.180	823.50	98.6
	II	155.3	11.782	6.504	11.314	817.69	109.4
Expt.	I		11.233	6.544	11.231	821.22	95.9
	II		12.095	6.491	11.323	827.1	111.5

# Aspirin internal angles

	Form	$\tau_1$ (deg)	$\tau_2$ (deg)	$\tau_3$ (deg)
PBE	I	89.4	1.3	1.4
	II	97.6	0.7	1.9
	Monomer	81.7	2.9	2.9
optB88	I	81.7	2.1	1.1
	II	84.9	2.7	1.1
	Monomer	78.6	3.7	3.5
Expt.	I	81.9	3.6	1.2
	II	84.0	5.4	3.0

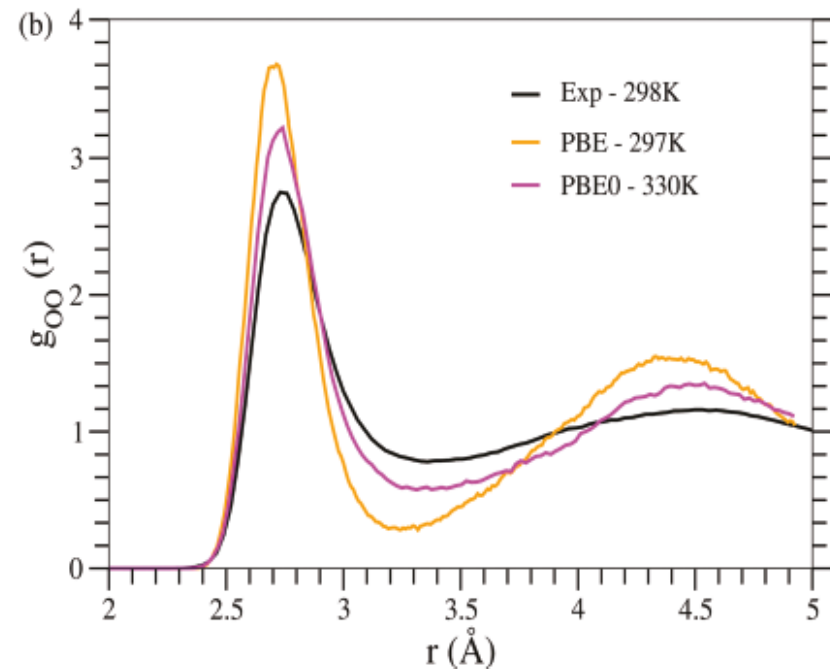
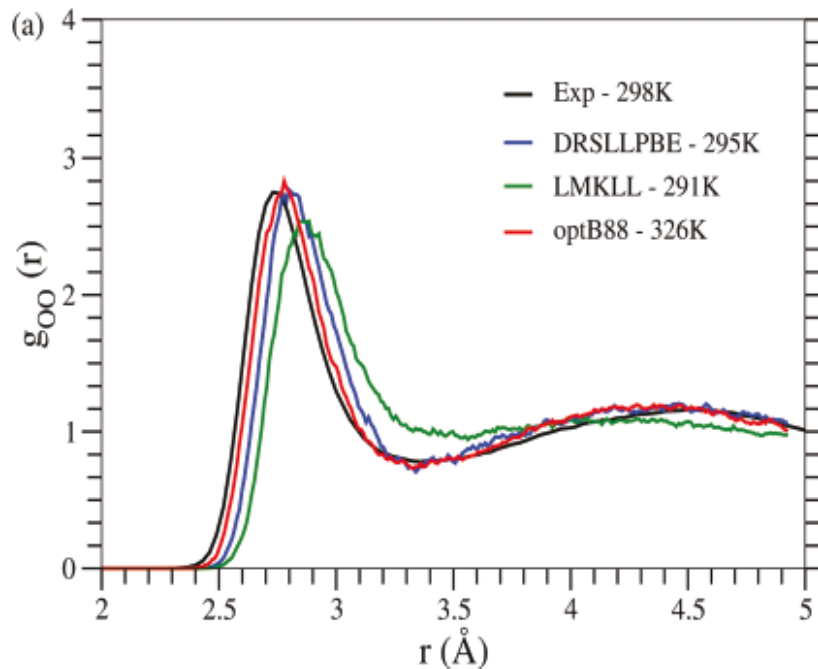
# vdW water dimer

## Water dimer: Comparison of vdW functionals



# vdW liquid water

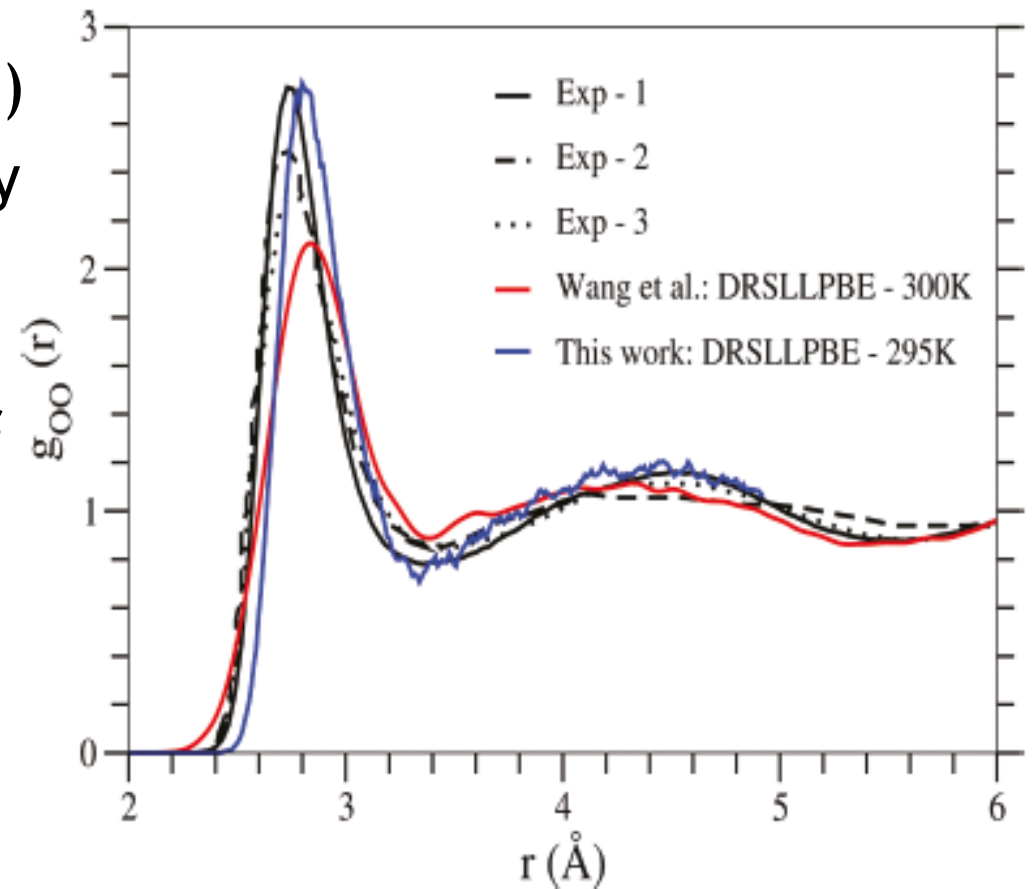
- vdW  $g_{OO}(r)$  improved w.r.t. PBE
- vdW effect differs from PBE0 effect



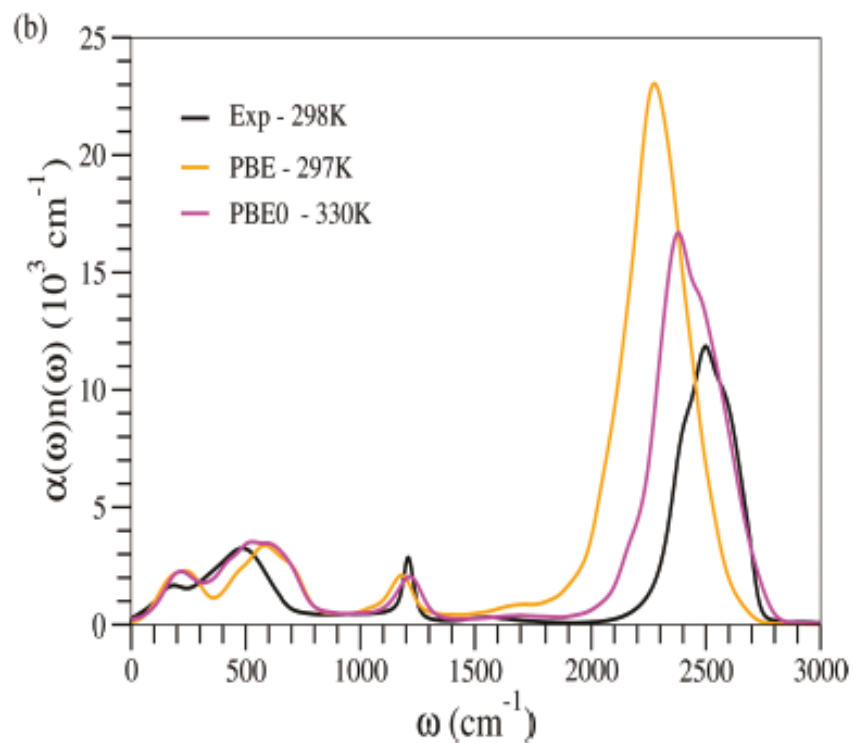
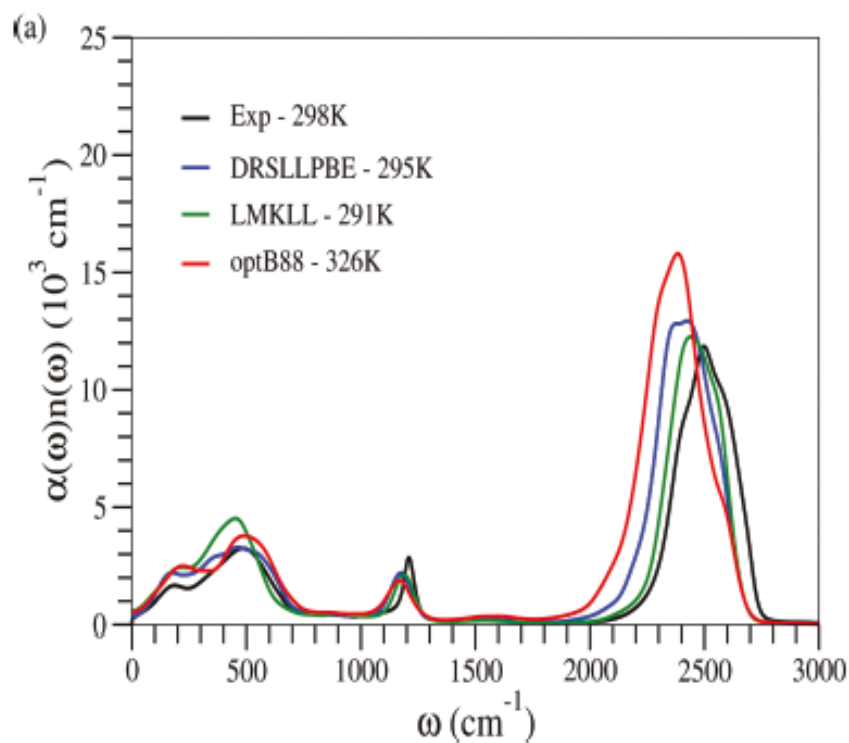


# DRSLLPBE: a modified vdW DF

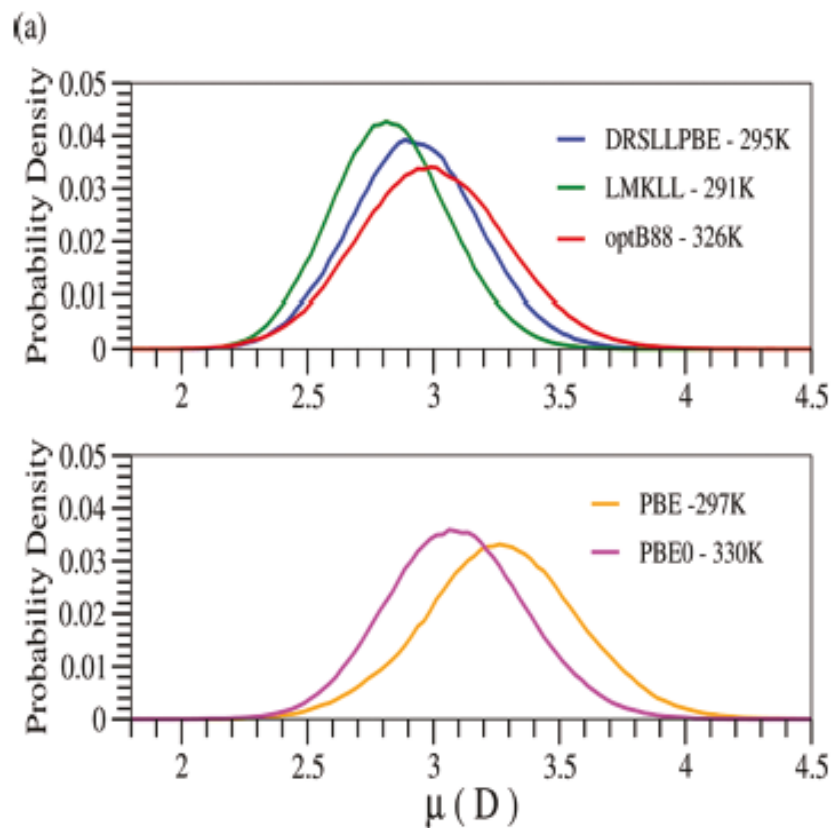
- Wang *et al* (DRSLLPBE)
  - revPBE replaced by PBE in DRSLL
- Note: basis set differences (SIESTA vs PW)



# vdW vibrational spectra



# vdW liquid water: dipole moment



# Summary

- Large-scale plane-wave pseudopotential FPMD simulations provide an accurate model of structure and electronic properties, with controlled accuracy
- Hybrid and vdW functionals improve the description of liquid water over PBE
- New software tools for verification and validation of simulations are needed