

OpenAtom: Fast, fine grained parallel electronic structure software for materials science, chemistry and physics.

Application Team

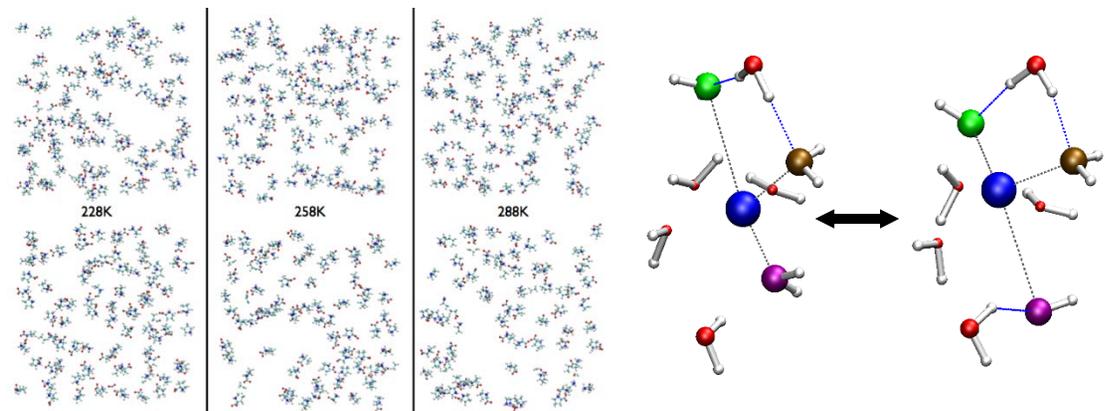
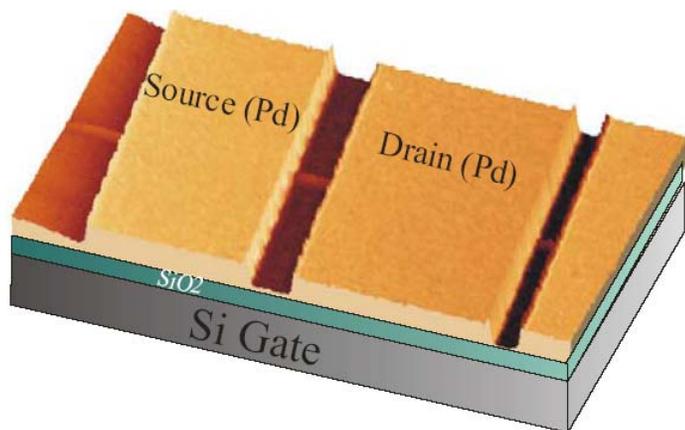
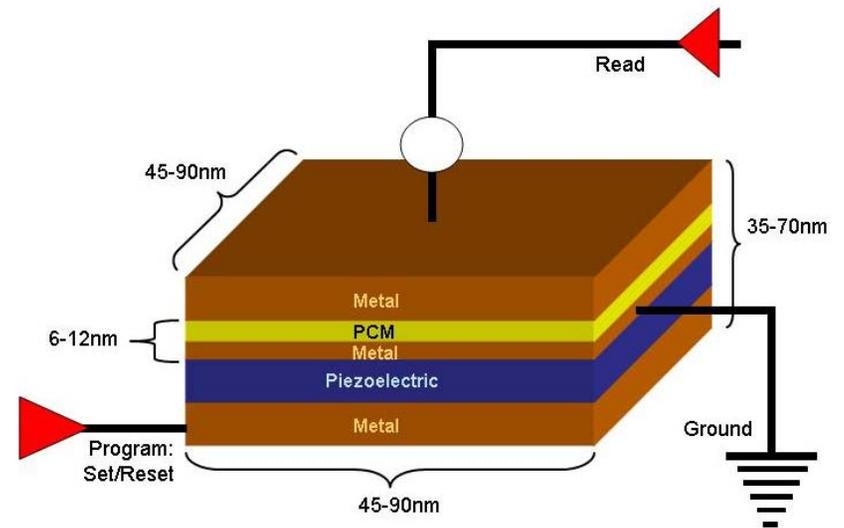
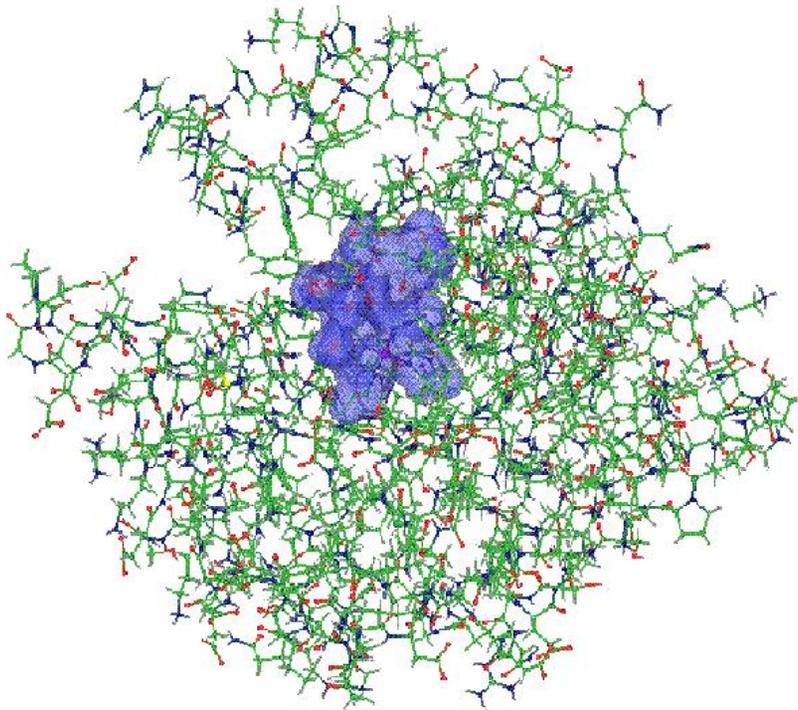
Glenn J. Martyna,	Physical Sciences Division, IBM Research & Edinburgh U.
Sohrab Ismail-Beigi,	Department of Applied Physics, Yale University
Dennis M. Newns,	Physical Sciences Division, IBM Research
Jason Crain,	School of Physics, Edinburgh University
Razvan Nistor,	Department of Chemistry, Columbia University
Ahmed Maarouf,	Egypt NanoTechnology Center.
Marcelo Kuroda,	Department of Physics, Auburn University

Methods/Software Development Team

Glenn J. Martyna,	Physical Sciences Division, IBM Research
Laxmikant Kale,	Computer Science Department, UIUC
Ramkumar Vadali,	Computer Science Department, UIUC
Sameer Kumar,	Computer Science, IBM Research
Eric Bohm,	Computer Science Department, UIUC
Abhinav Bhatele,	Computer Science Department, UIUC
Ramprasad Venkataraman,	Computer Science Department, UIUC
Anshu Arya,	Computer Science Department, UIUC
Nikhil Jain,	Computer Science Department, UIUC
Eric Mikida,	Computer Science Department, UIUC

Funding : NSF, IBM Research, ONRL, ...

Goal : The accurate treatment of complex heterogeneous systems to gain physical insight.



Limitations of *ab initio* MD

- Limited to small systems (100-1000 atoms)*.
- Limited to short time dynamics and/or sampling times.
- Parallel scaling only achieved for
processors \leq # electronic states
until recent efforts by ourselves and others.

Improving this will allow us to sample longer and learn new physics.

*The methodology employed herein scales as $O(N^3)$ with system size due to the orthogonality constraint, only.

Density Functional Theory : DFT

In the Kohn-Sham formulation of density functional theory, the electron density is expanded in a set of orbitals, $\{\psi_i(\mathbf{r})\}$,

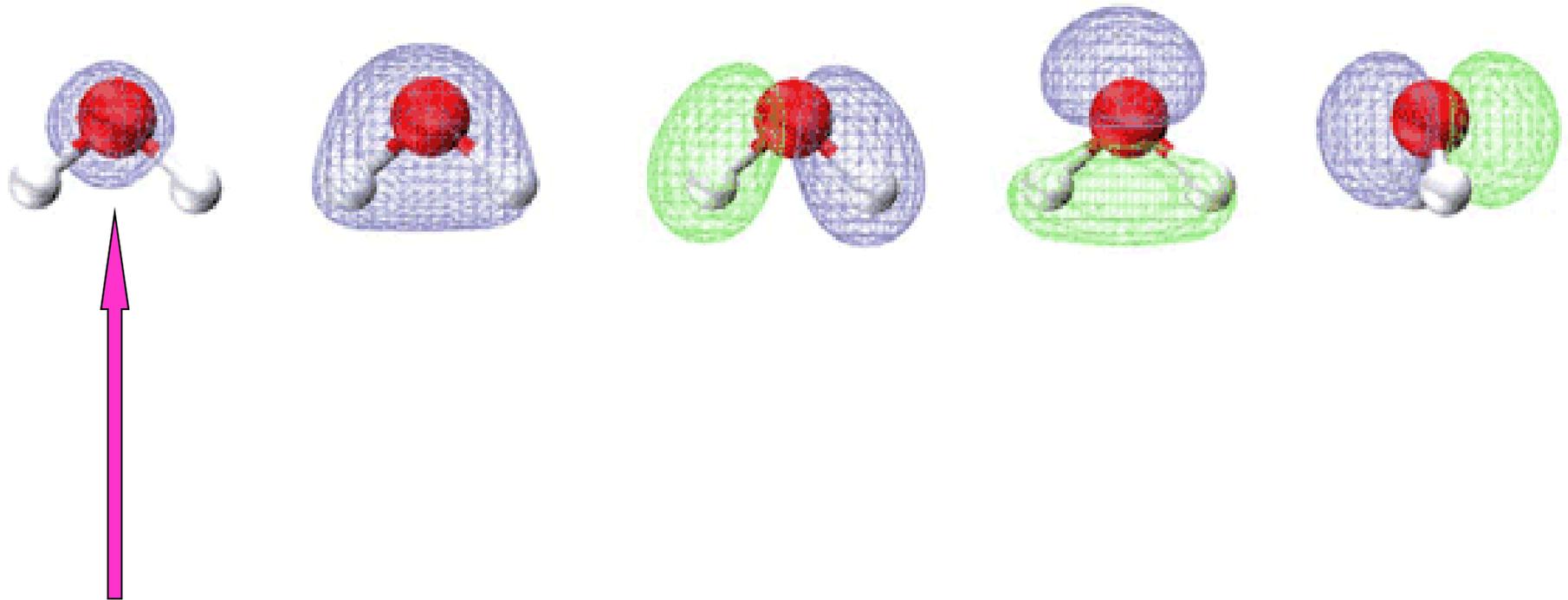
$$n(\mathbf{r}) = \sum_{i=1}^{n_{\text{occ}}} f_i |\psi_i(\mathbf{r})|^2$$

subject to the orthogonality constraint, ($\langle \psi_j | \psi_k \rangle = \delta_{jk}$), where n_{occ} is the number of occupied orbitals and the f_i are the occupation numbers.

The energy functional is given by

$$E[n] = T_s[\{\psi_i\}] + E_H[n] + E_{\text{ext,loc}}[n] + E_{\text{ext,non-loc}}[\{\psi_i\}] + E_{\text{xc}}[n]$$

Electronic states/orbitals of water



Removed by introducing a non-local electron-ion interaction.

Plane Wave Basis Set:

In plane-wave based calculations at the Γ -point, the orbitals and, hence, the density are expanded

$$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{g}} \bar{\psi}_i(\mathbf{g}) \exp(i\mathbf{g} \cdot \mathbf{r})$$
$$n(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{g}} \bar{n}(\mathbf{g}) \exp(i\mathbf{g} \cdot \mathbf{r})$$

Here, $\hat{\mathbf{g}}$ is the vector of integers $\{j, k, m\}$, $\mathbf{g} = 2\pi\hat{\mathbf{g}}\mathbf{h}^{-1}$, $V = \det \mathbf{h}$ is the volume, \mathbf{h} is the matrix whose columns are the Cartesian components of the three vectors describing the parallelepiped enclosing the system and $\mathbf{r} = \mathbf{h}\mathbf{s}$. A plane wave basis can be used to describe

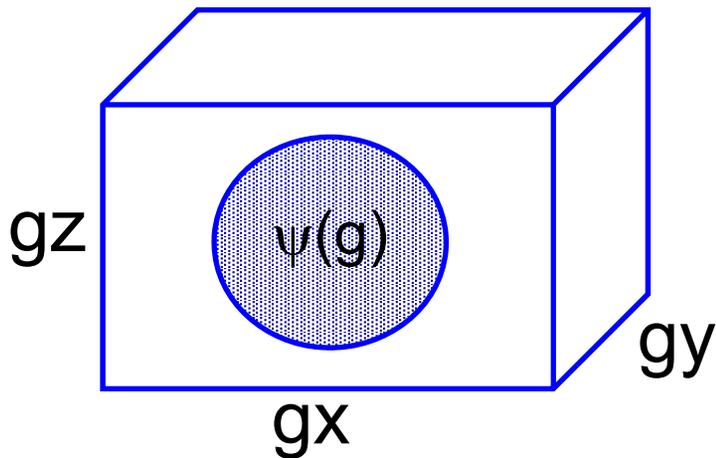
The # of states or orbitals $\sim N$ where N is # of atoms.

The # of pts in \mathbf{g} -space $\sim N$.

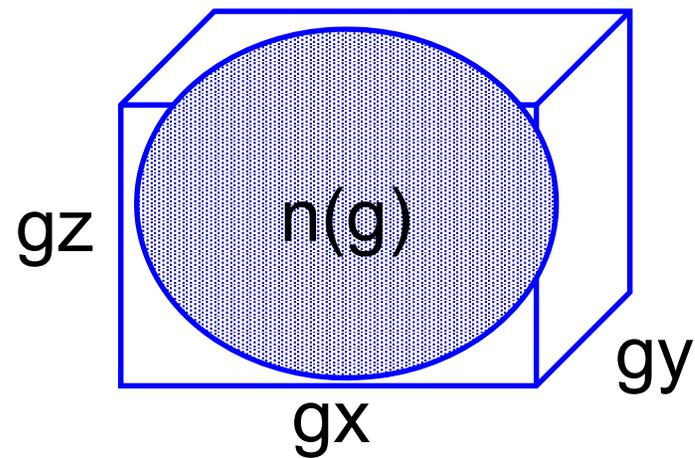
The # of electrons $\sim N$.

Plane Wave Basis Set:

Two Spherical cutoffs in G-space



$\psi(g)$: radius g_{cut}

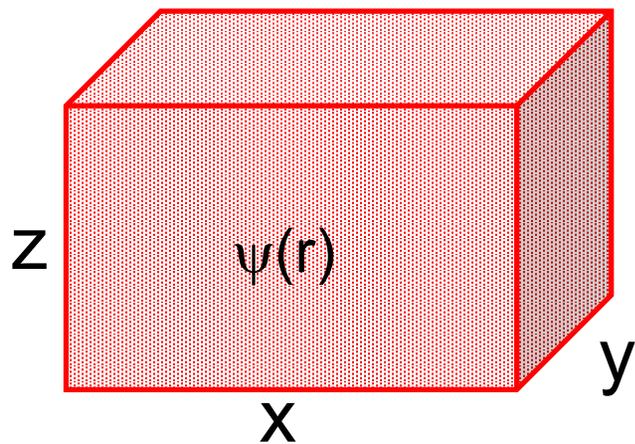


$n(g)$: radius $2g_{\text{cut}}$

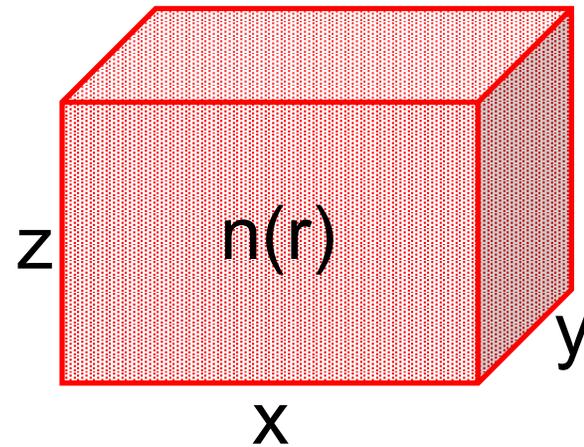
g -space is a discrete regular grid due to finite size of system!!

Plane Wave Basis Set:

The dense discrete real space mesh.



$$\psi(r) = \text{3D-FFT}\{\psi(g)\}$$

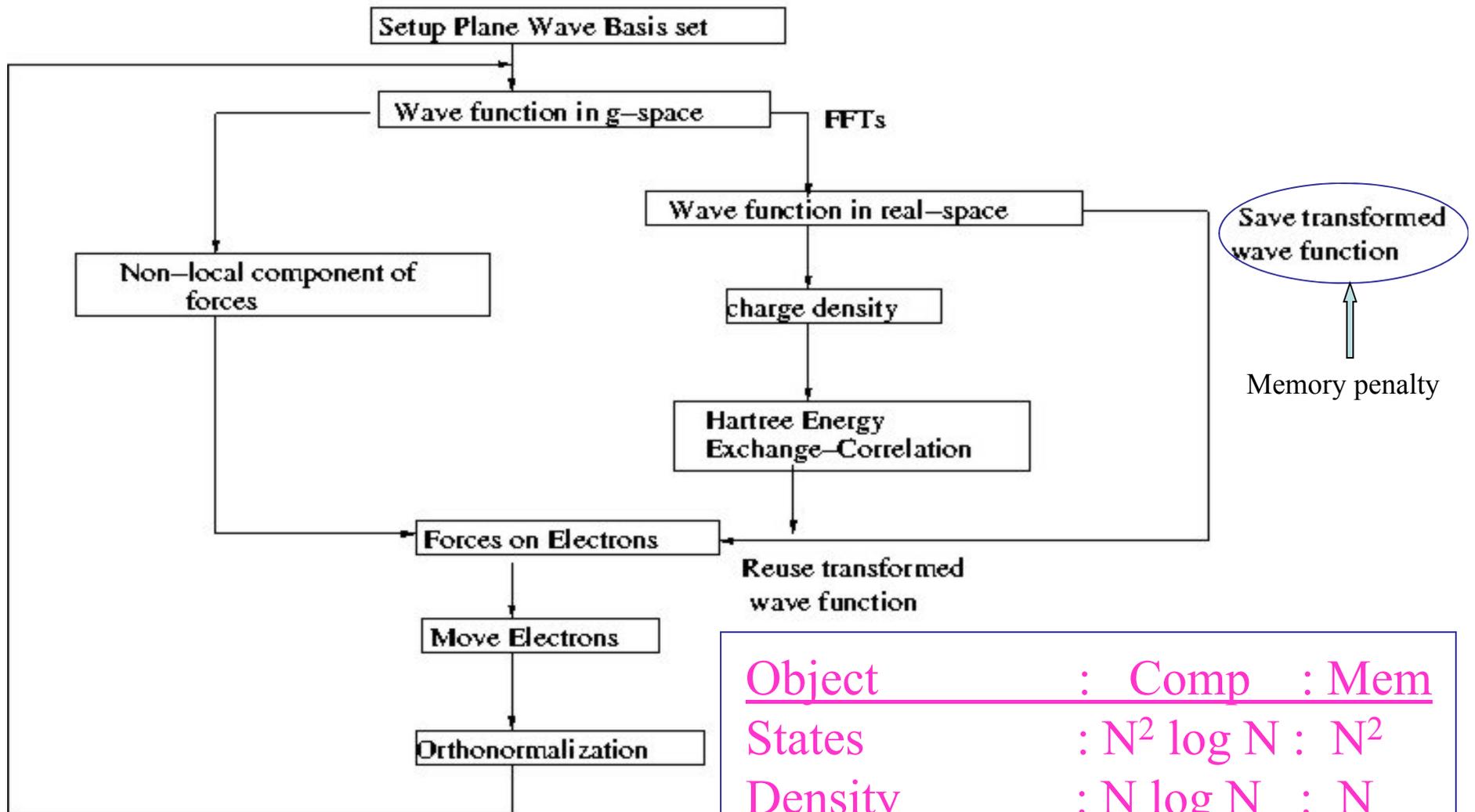


$$n(r) = \sum_k |\psi_k(r)|^2$$

$$n(g) = \text{3D-IFFT}\{n(r)\} \text{ exactly!}$$

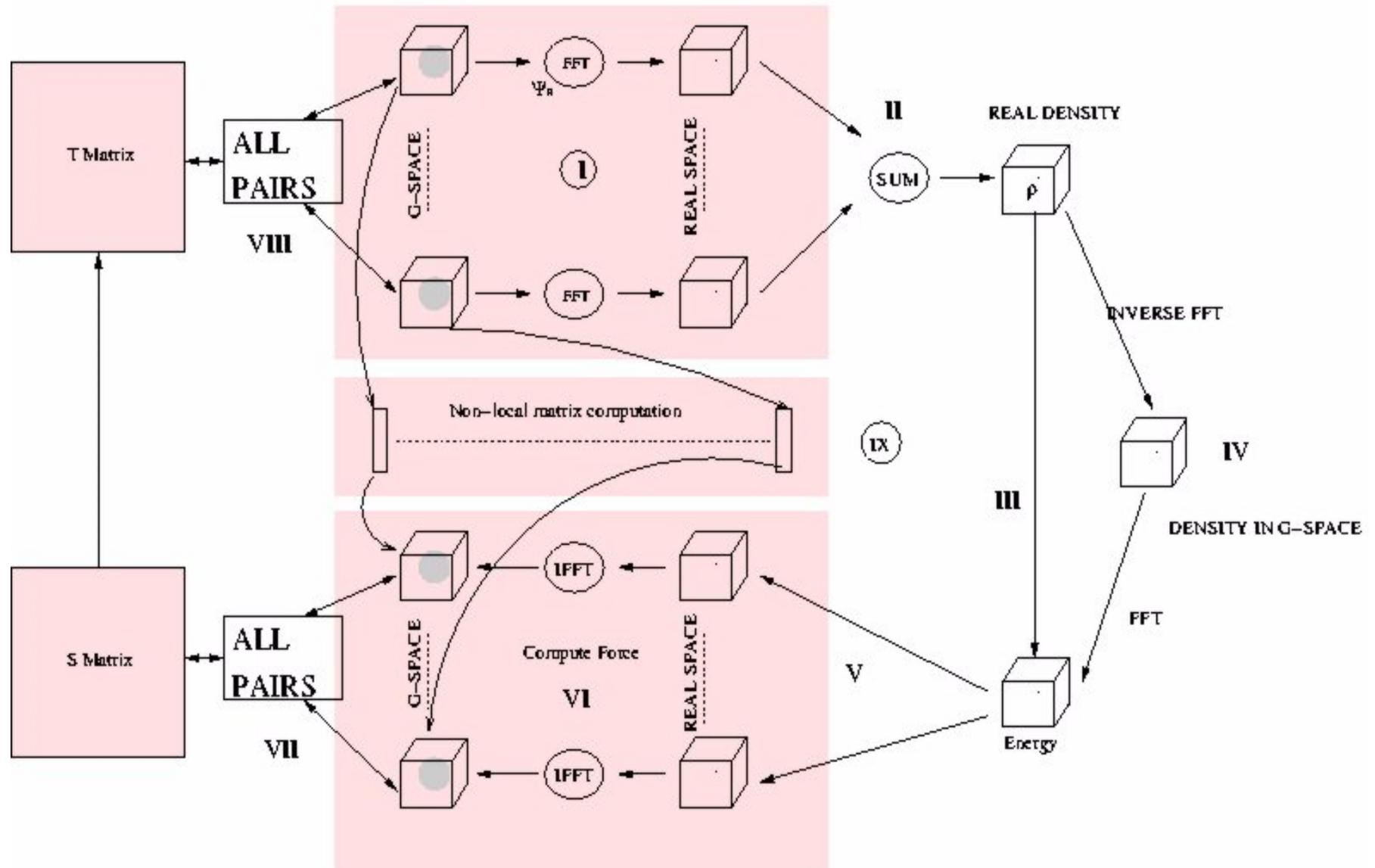
Although r -space is a discrete dense mesh, $n(g)$ is generated exactly!

Simple Flow Chart : Scalar Ops

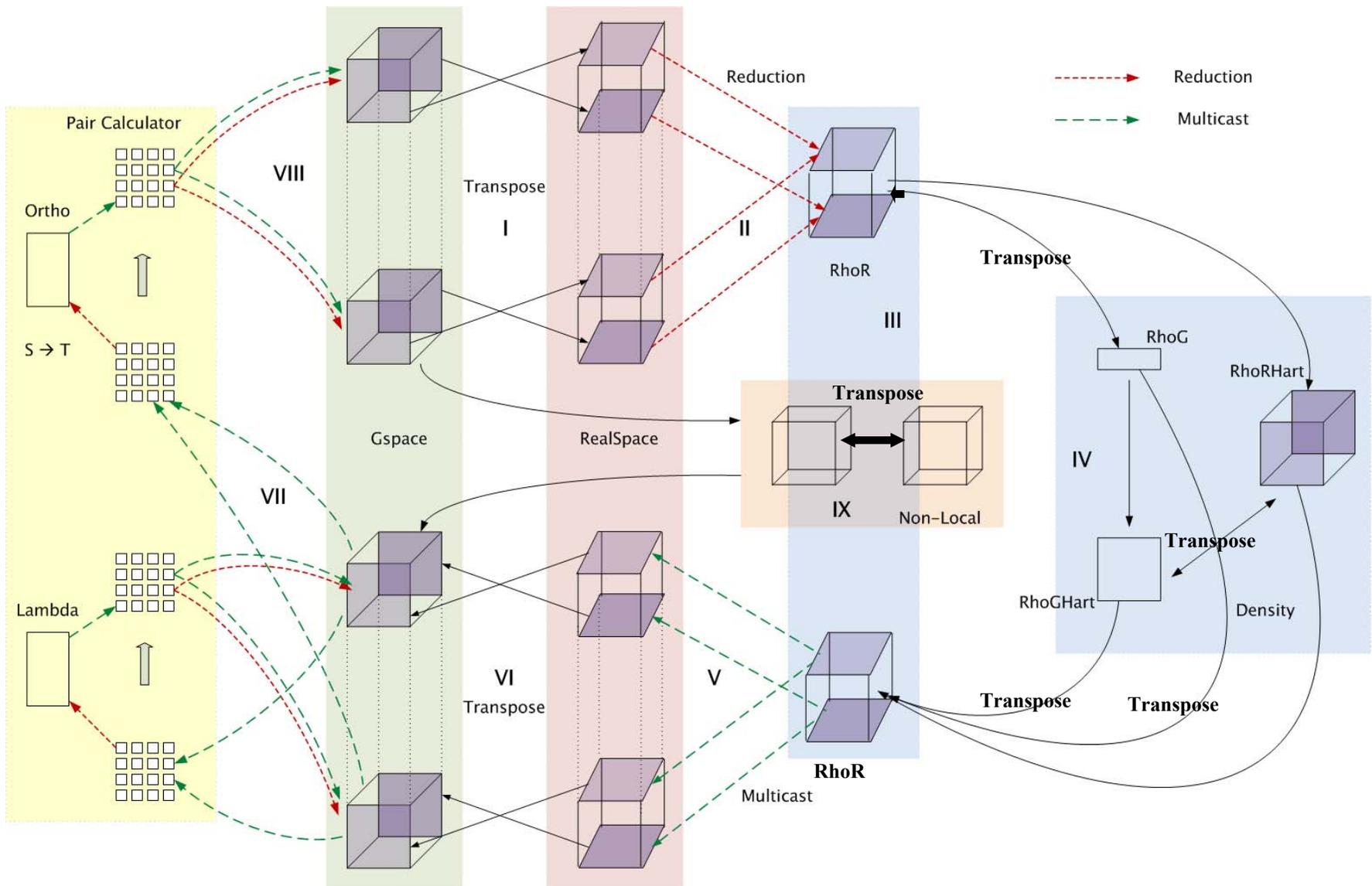


<u>Object</u>	: Comp	: Mem
States	: $N^2 \log N$: N^2
Density	: $N \log N$: N
Orthonormality	: N^3	: $N^{2.33}$

Flow Chart : Data Structures



Parallelization under charm++

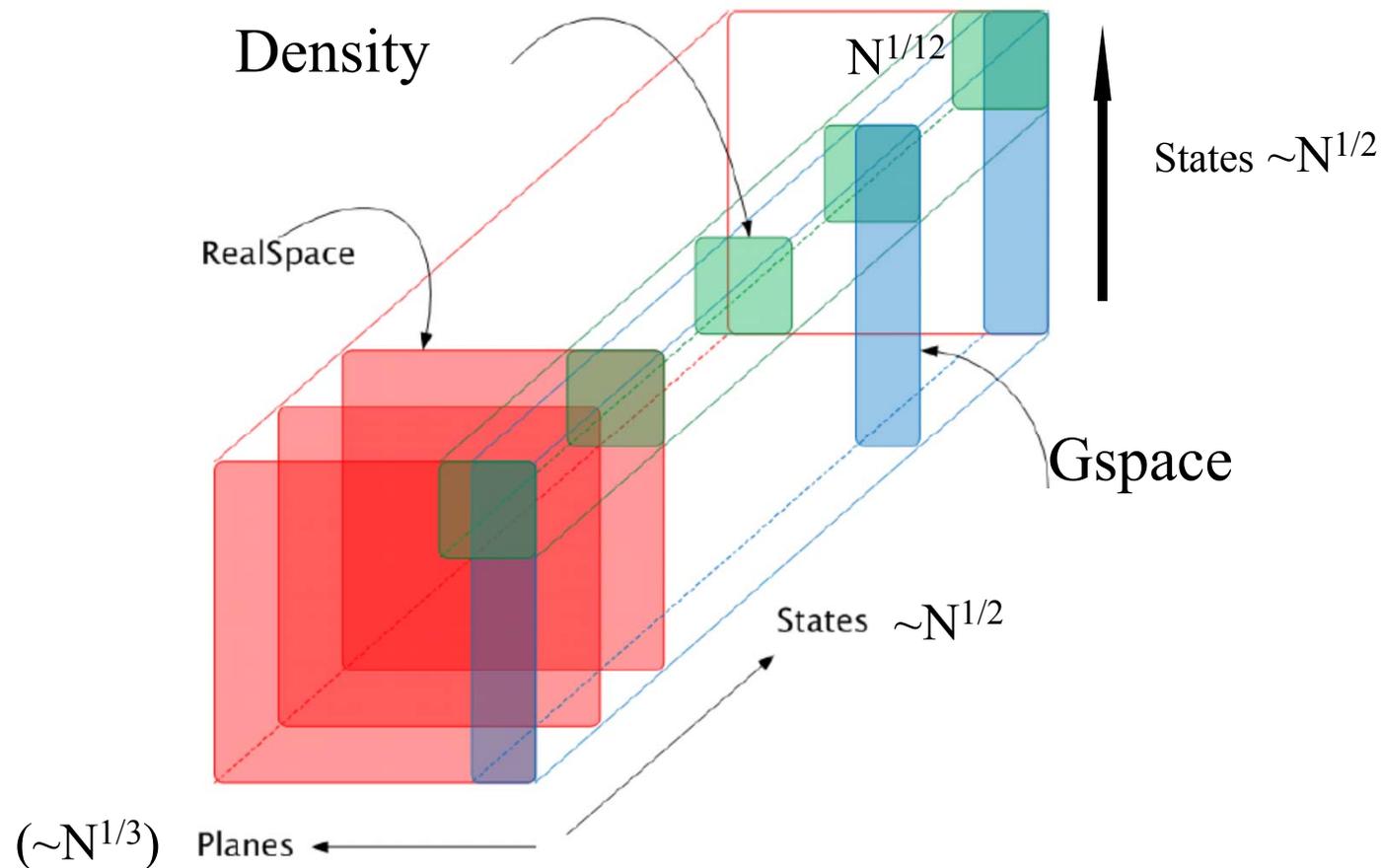


Challenges to scaling:

- Multiple concurrent 3D-FFTs to generate the states in real space require AllToAll communication patterns. Communicate N^2 data pts.
- Reduction of states ($\sim N^2$ data pts) to the density ($\sim N$ data pts) in real space.
- Multicast of the KS potential computed from the density ($\sim N$ pts) back to the states in real space ($\sim N$ copies to make N^2 data).
- Applying the orthogonality constraint requires N^3 operations.
- Mapping the chare arrays/VPs to BG/L processors in a topologically aware fashion.

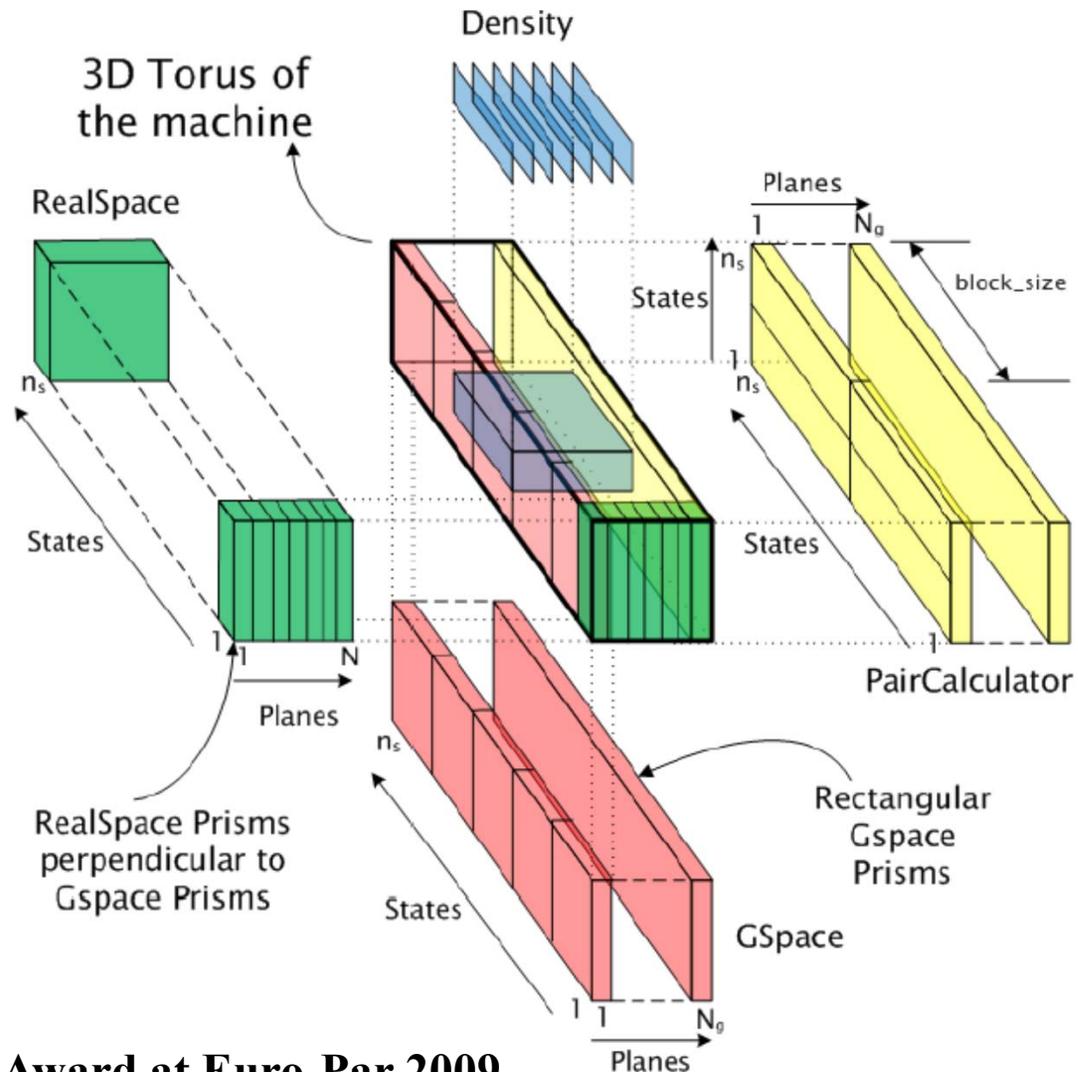
Scaling bottlenecks due to non-local and local electron-ion interactions removed by the introduction of new methods!

Topologically aware mapping for CPAIMD



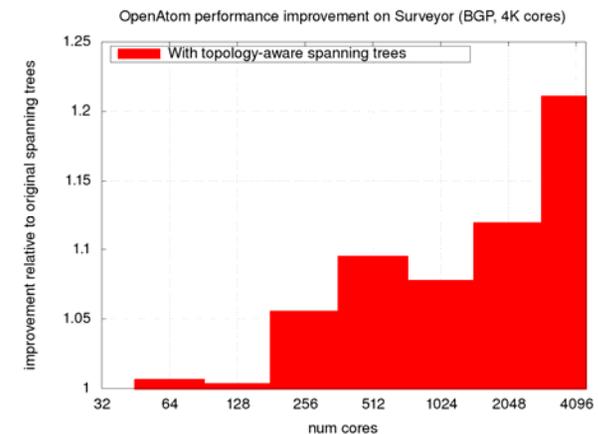
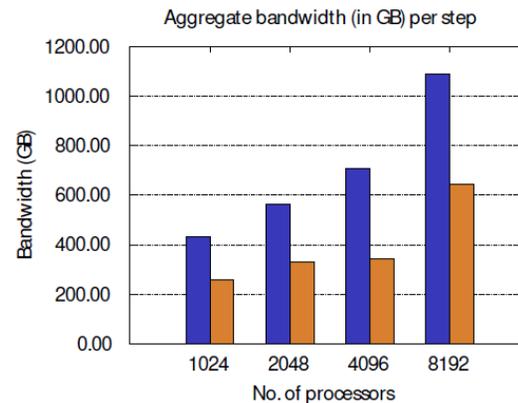
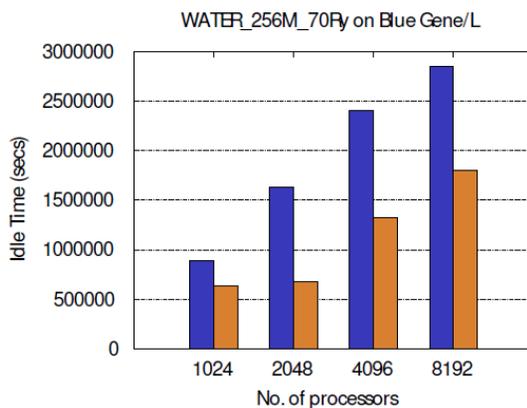
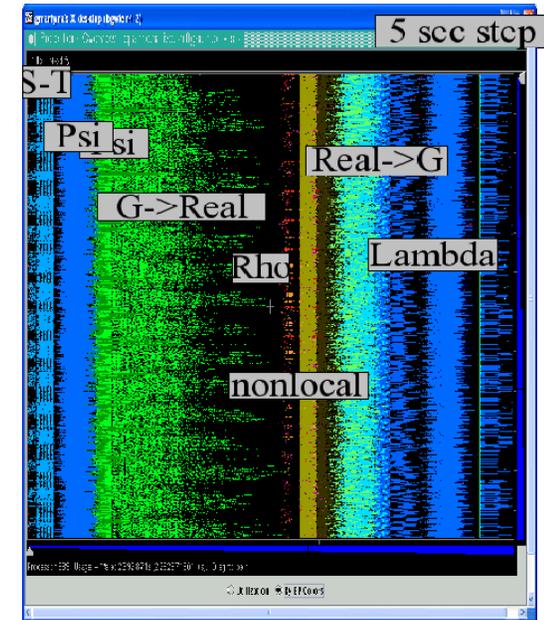
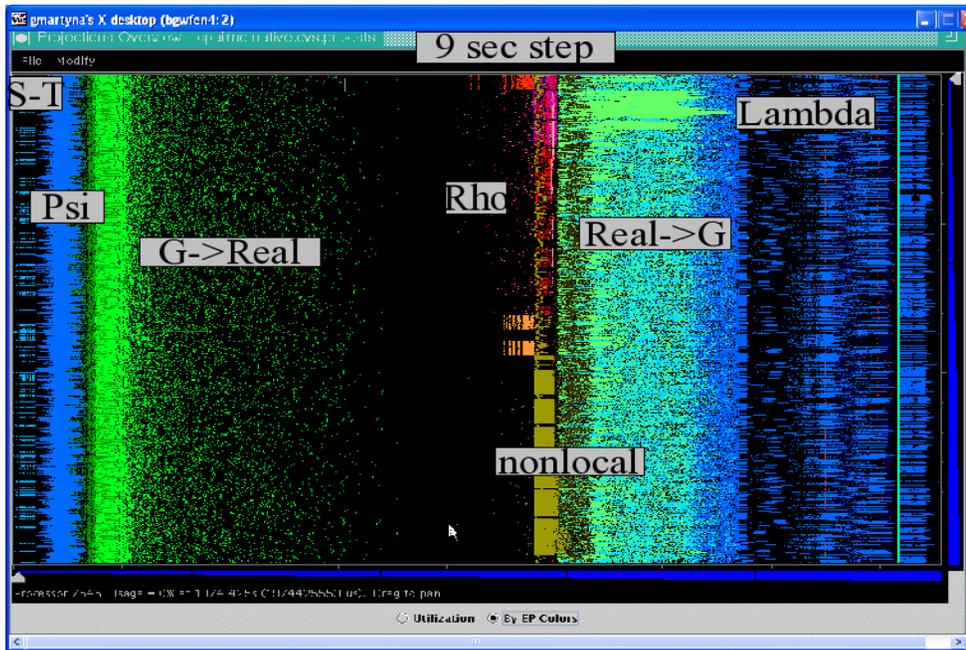
- The states are confined to rectangular prisms cut from the torus to minimize 3D-FFT communication.
- The density placement is optimized to reduced its 3D-FFT communication and the multicast/reduction operations.

Topologically aware mapping for CPAIMD : Details



Distinguished Paper Award at Euro-Par 2009

Improvements wrought by topological aware mapping on the network torus architecture



Density (R) reduction and multicast to State (R) improved.
 State (G) communication to/from orthogonality chares improved.

Parallel scaling of liquid water* as a function of system size on the Blue Gene/L installation at YKT:

CO Mode Native Layer with Optimizations											
Nodes	32	64	128	256	512	1024	2048	4096	8192	16384	20480
Processors	32	64	128	256	512	1024	2048	4096	8192	16384	20480
W8 Time s/step	0.22	0.10	0.082	0.071	0.046	0.026	0.020				
W16 Time s/step	0.73	0.40	0.23	0.15	0.106	0.061	0.041	0.035			
W32 Time s/step	2.71	1.52	0.95	0.44	0.26	0.15	0.11	0.081	0.063		
W64 Time s/step		6.72	3.77	1.88	0.87	0.51	0.31	0.21	0.15		
128 Time s/step					6.9	2.73	1.40	0.91	0.58	0.37	0.3
W256 Time s/step						16.4	8.14	4.83	2.75	1.71	1.54

*Liquid water has 4 states per molecule.

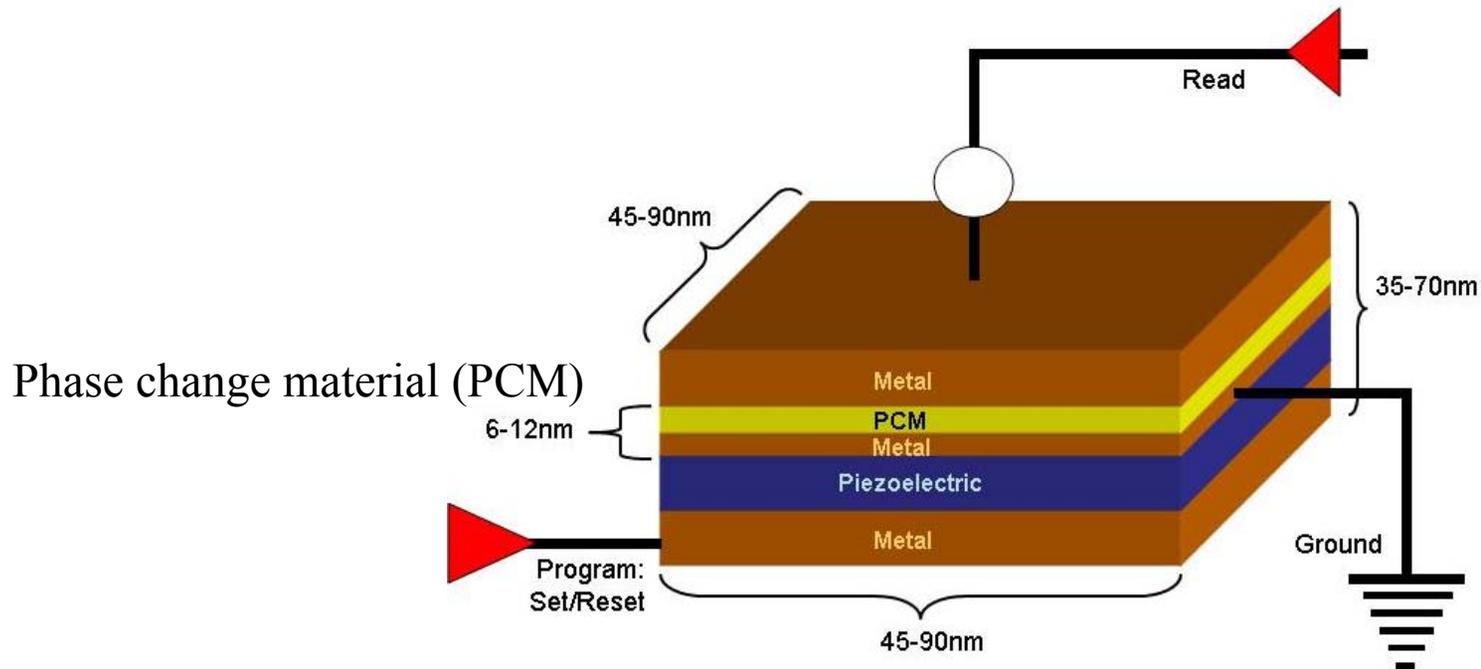
- Weak scaling is observed!
- Strong scaling on processor numbers up to ~60x the number of states!
- IBM J. Res. Dev. (2009).

Software : Summary

- Fine grained parallelization of the Car-Parrinello *ab initio* MD method demonstrated on thousands of processors :
processors >> # electronic states.
- Long time simulations of small systems are now possible on large massively parallel supercomputers.

Application Study if time allows

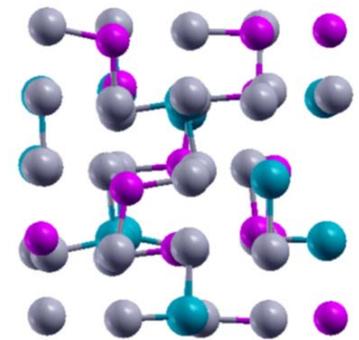
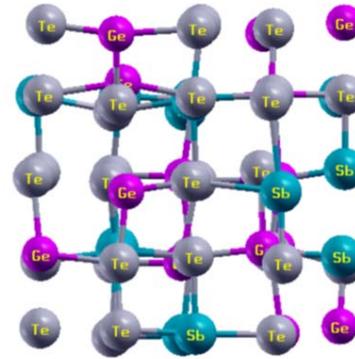
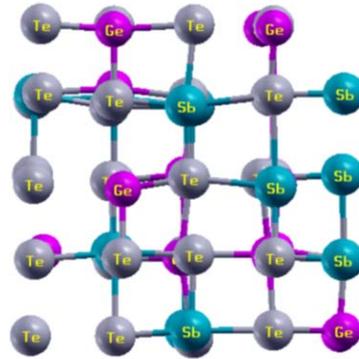
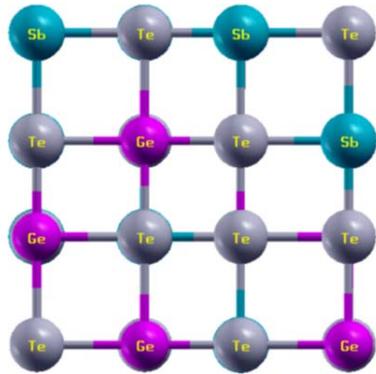
Piezoelectrically driven Phase Change Memory would be fast, cool & scalable:



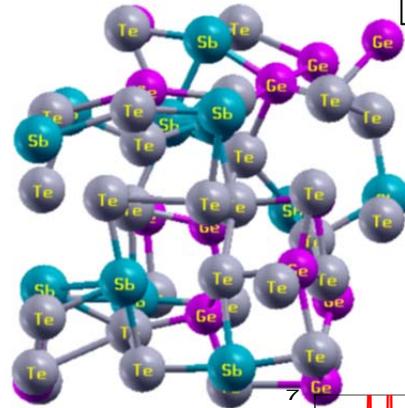
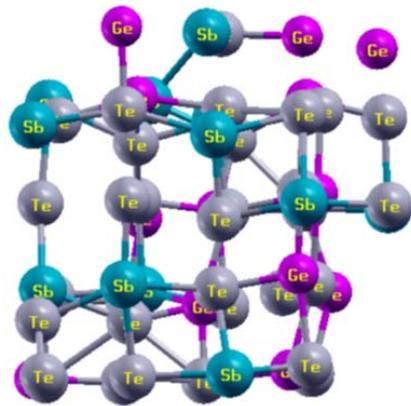
In ON state PCM is in a LOW resistance form → “1”.
In OFF state PCM is in a HIGH resistance form → “0”.

Can we find suitable material that can be switched by pressure using a combined exp/theor approach?

High Resistance State

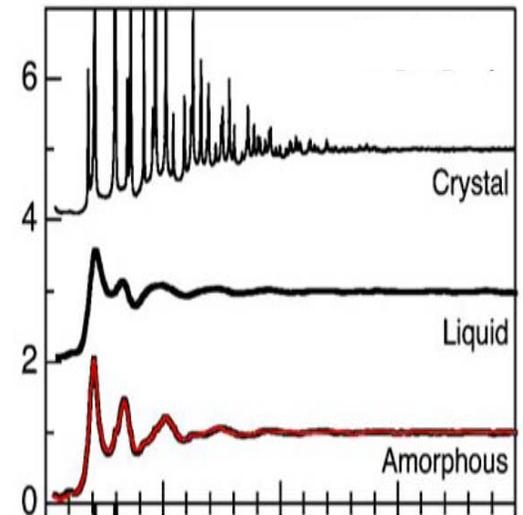
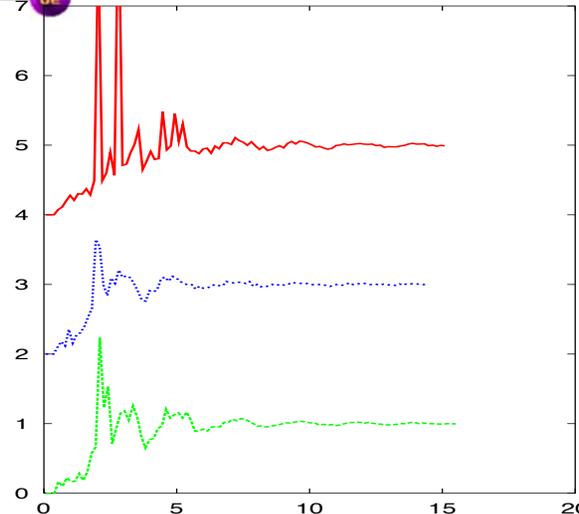


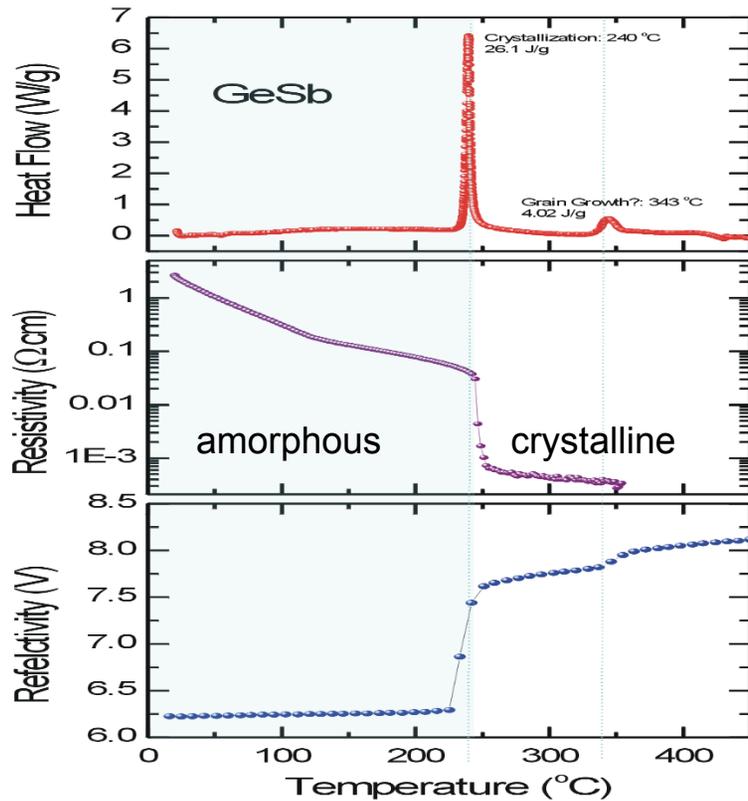
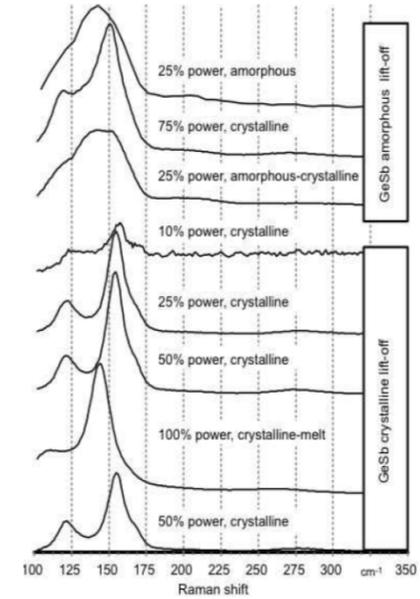
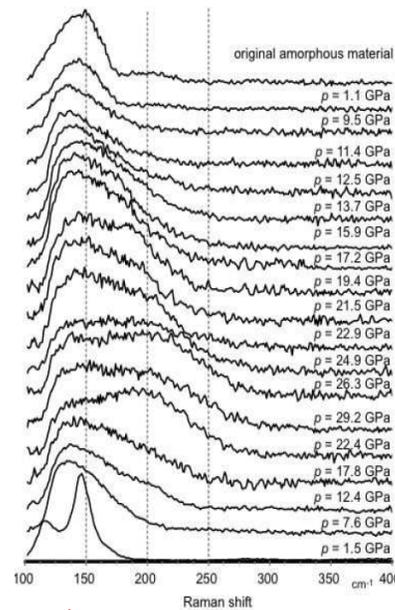
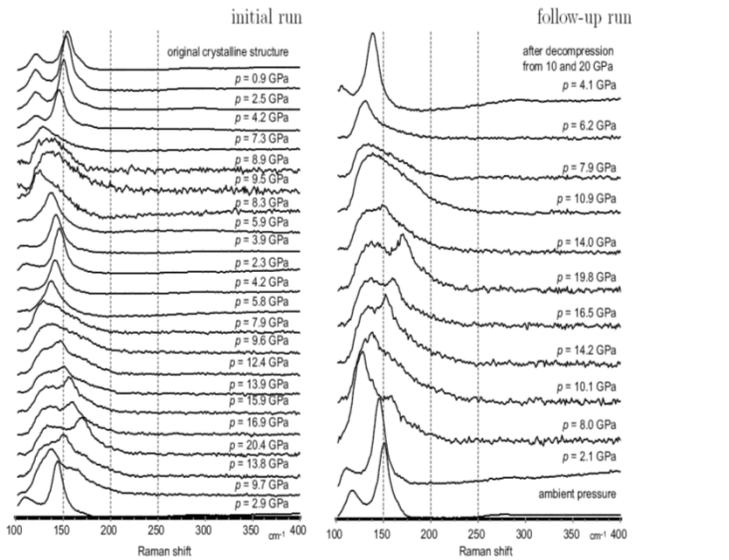
Low Resistance State



$\text{Ge}_2\text{Sb}_2\text{Te}_5$ -undergoes pressure induced “amorphization” both experimentally and theoretically

but the process is not reversible!



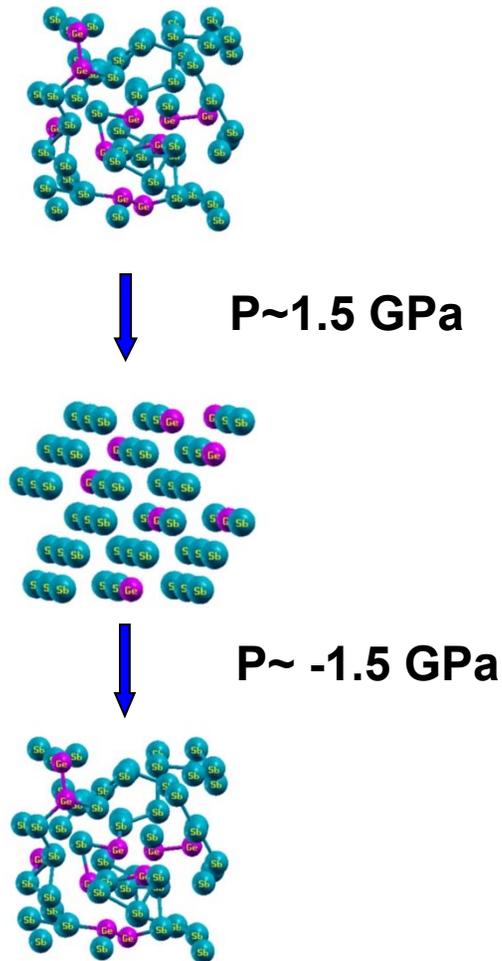


Eutectic GeSb undergoes an amorphous to crystalline transformation under pressure, experimentally!

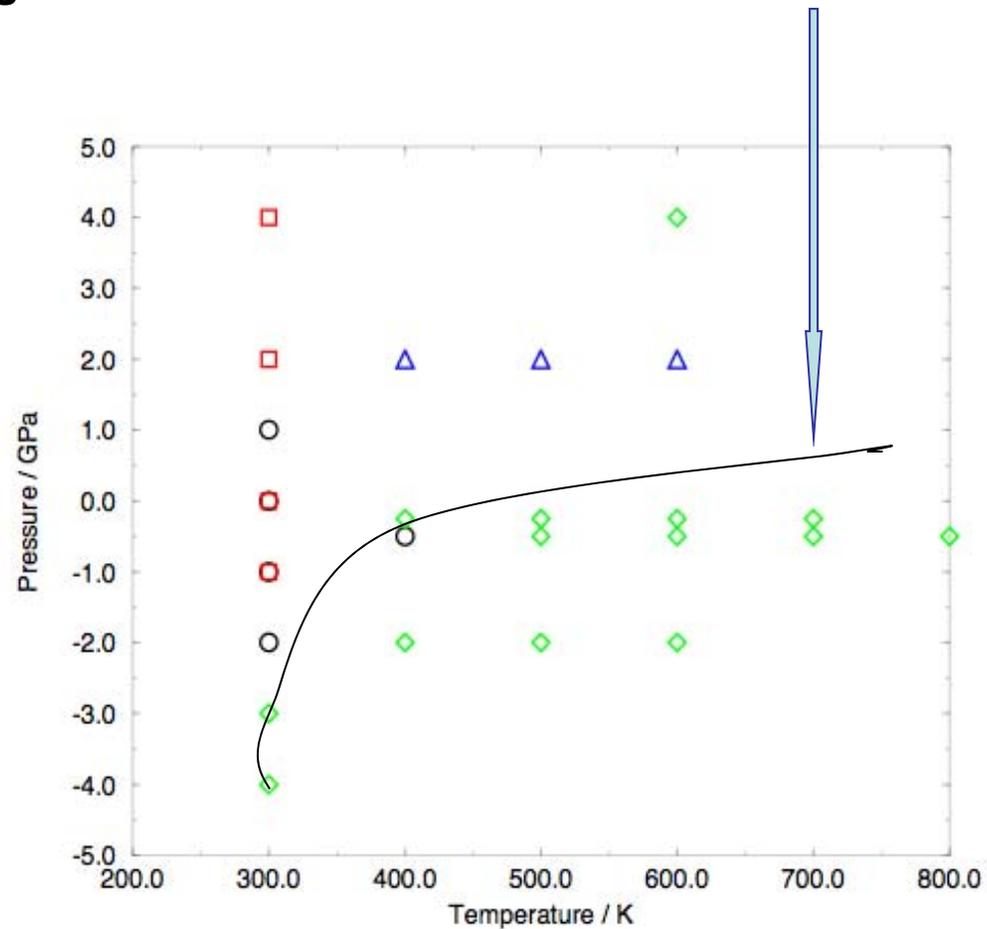
Is the process amenable to reversible switching as in the thermal approach???

Utilize tensile load to approach the spinodal and cause pressure induced amorphization!

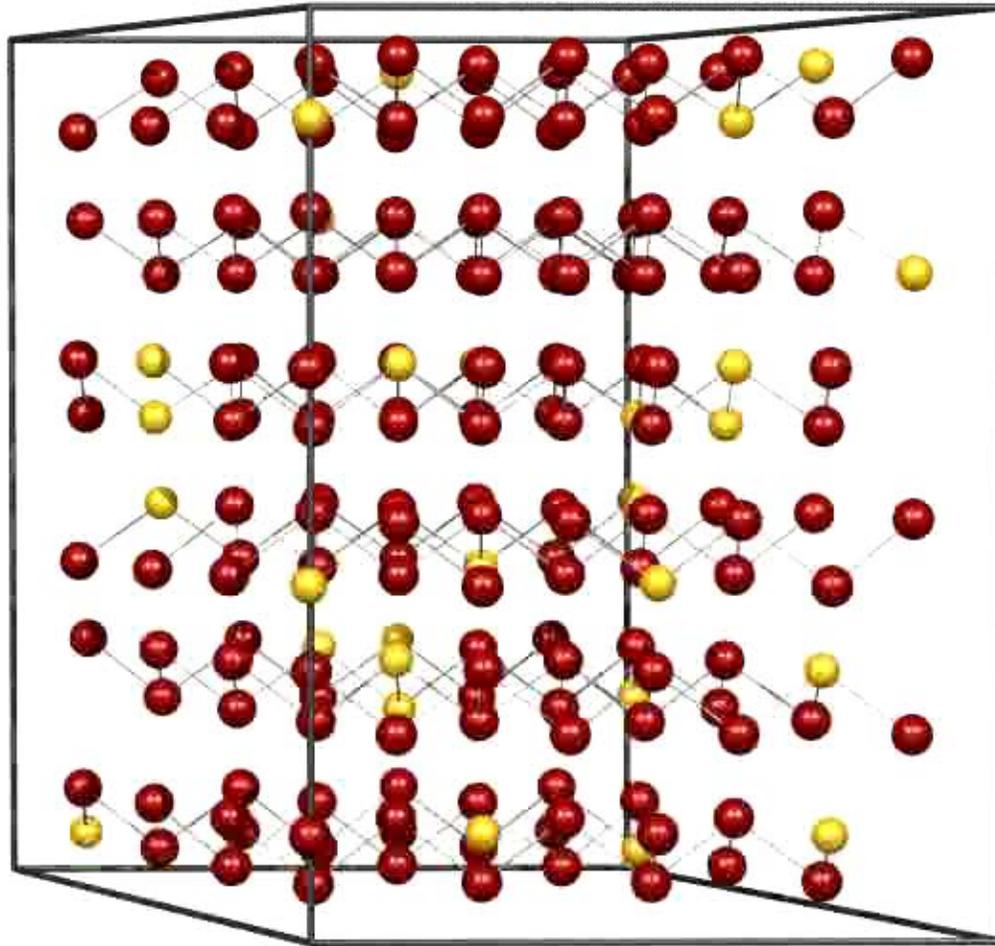
Schematic of a potential device based on pressure switching



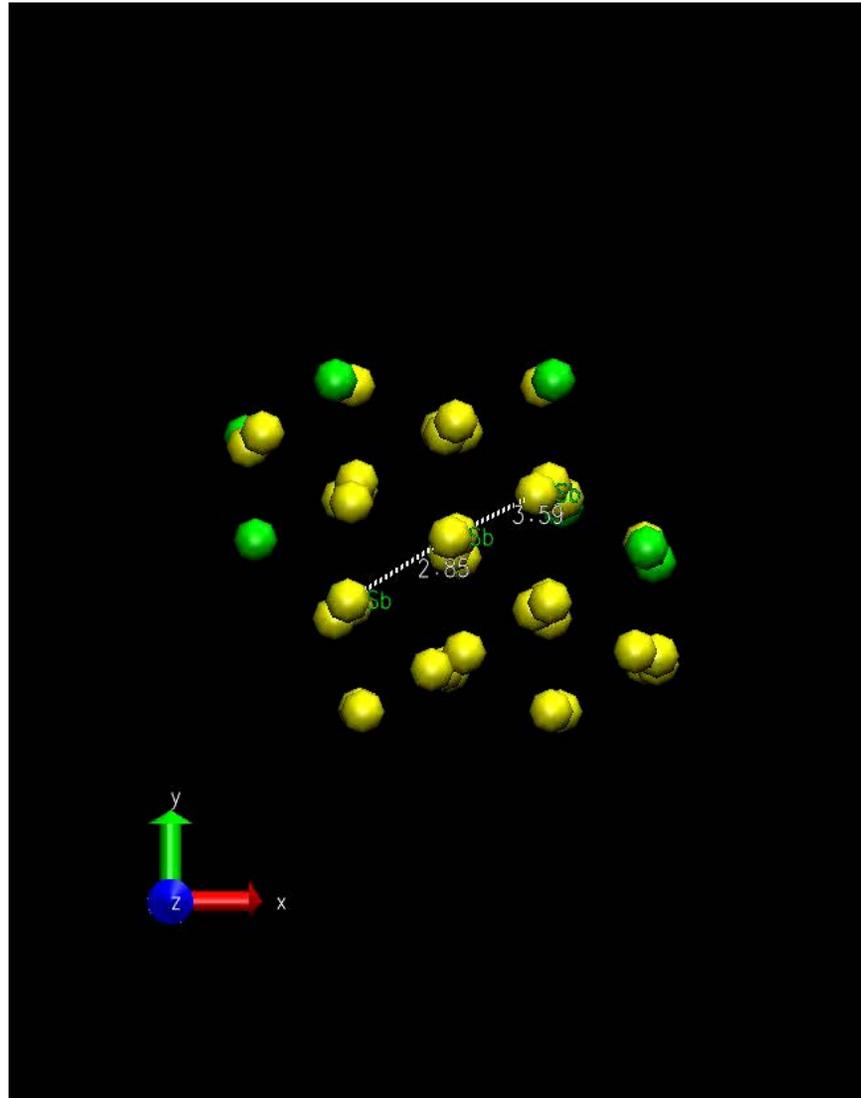
CPAIMD spinodal line!



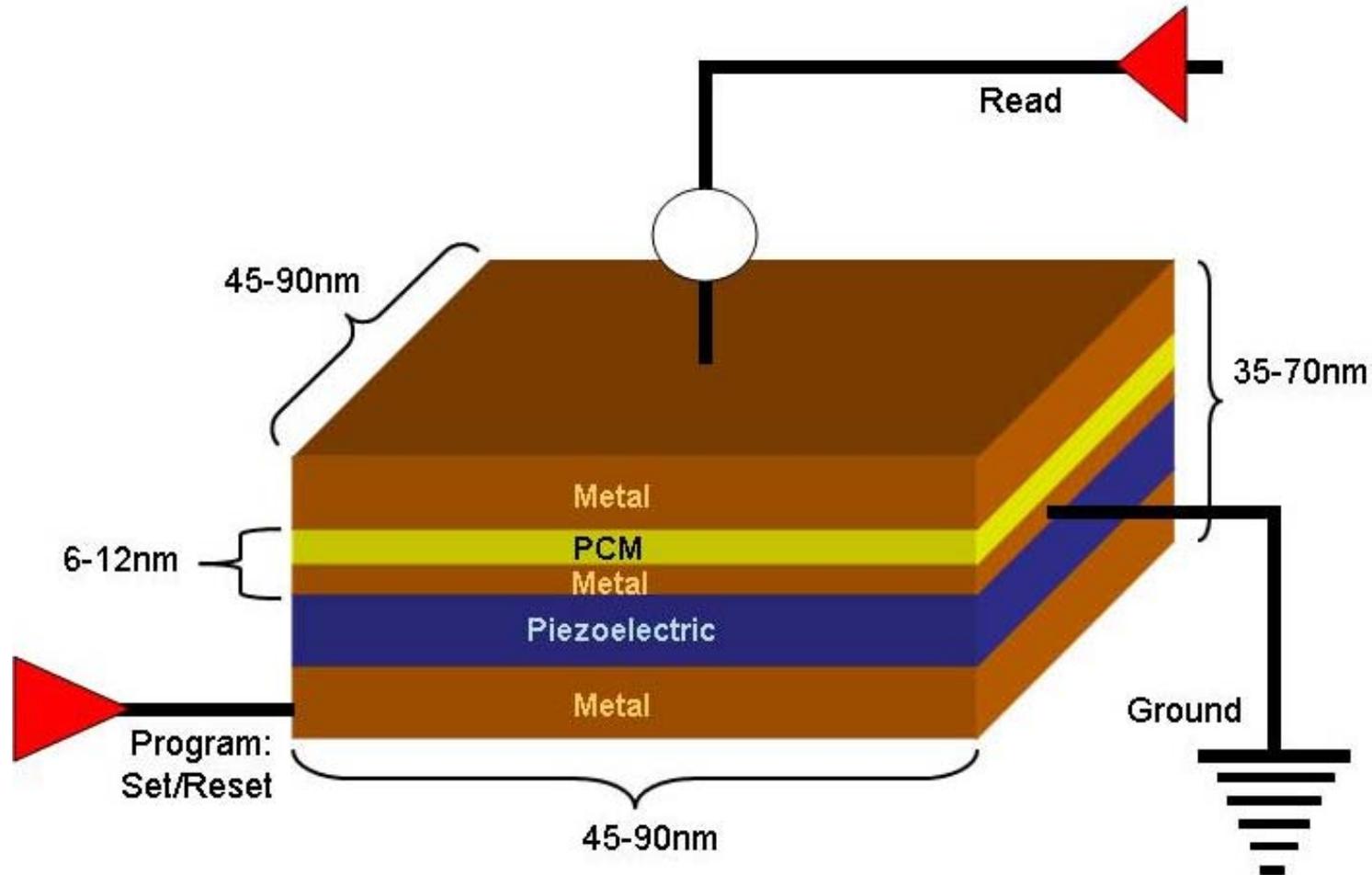
Spinodal decomposition under tensile load



Solid is stable at ambient pressure



IBM's Piezoelectric Memory



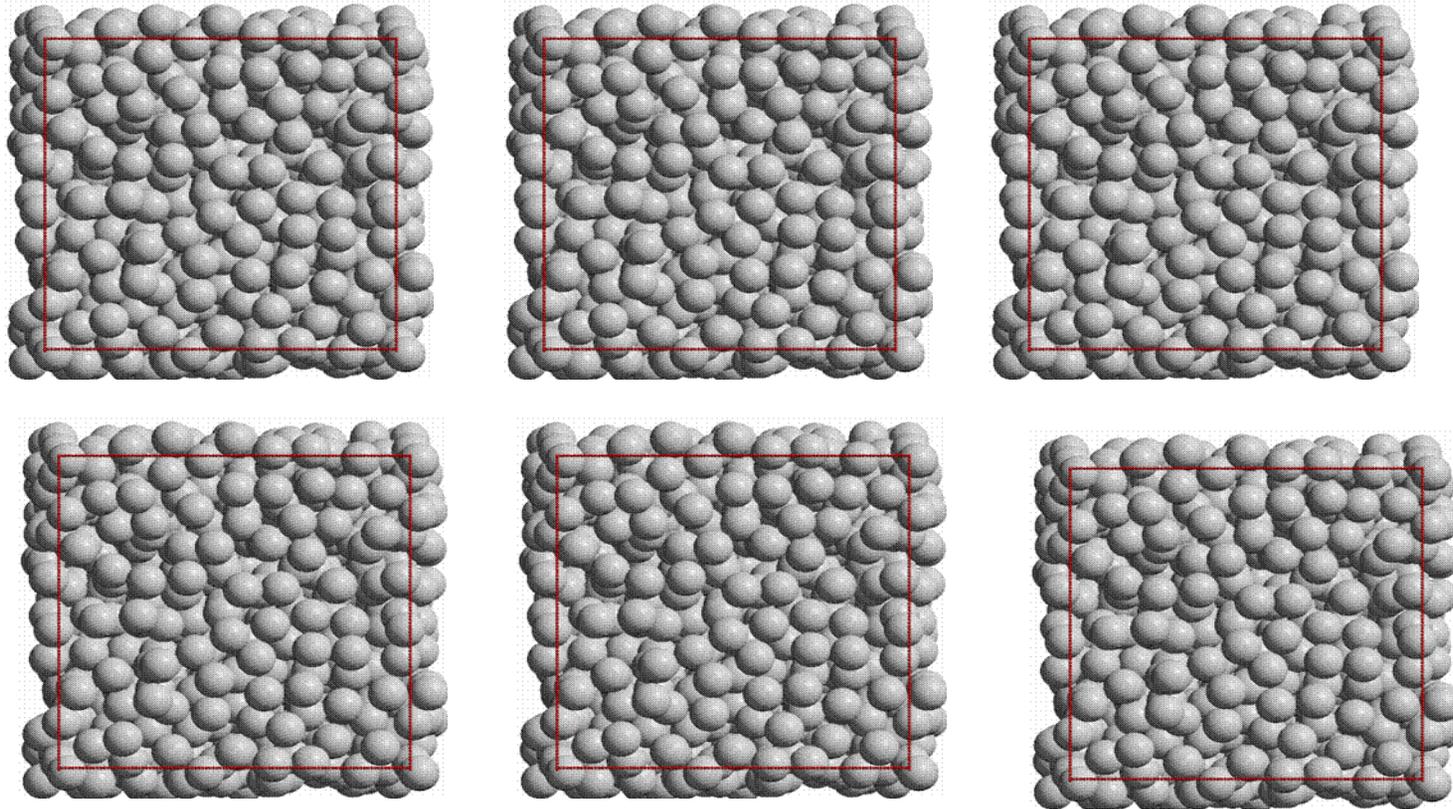
We are investigating other materials and better device designs!
Patent filed. Scientific work has appeared in PNAS.

K-points, Path Integrals and Parallel Tempering

Instance parallelization

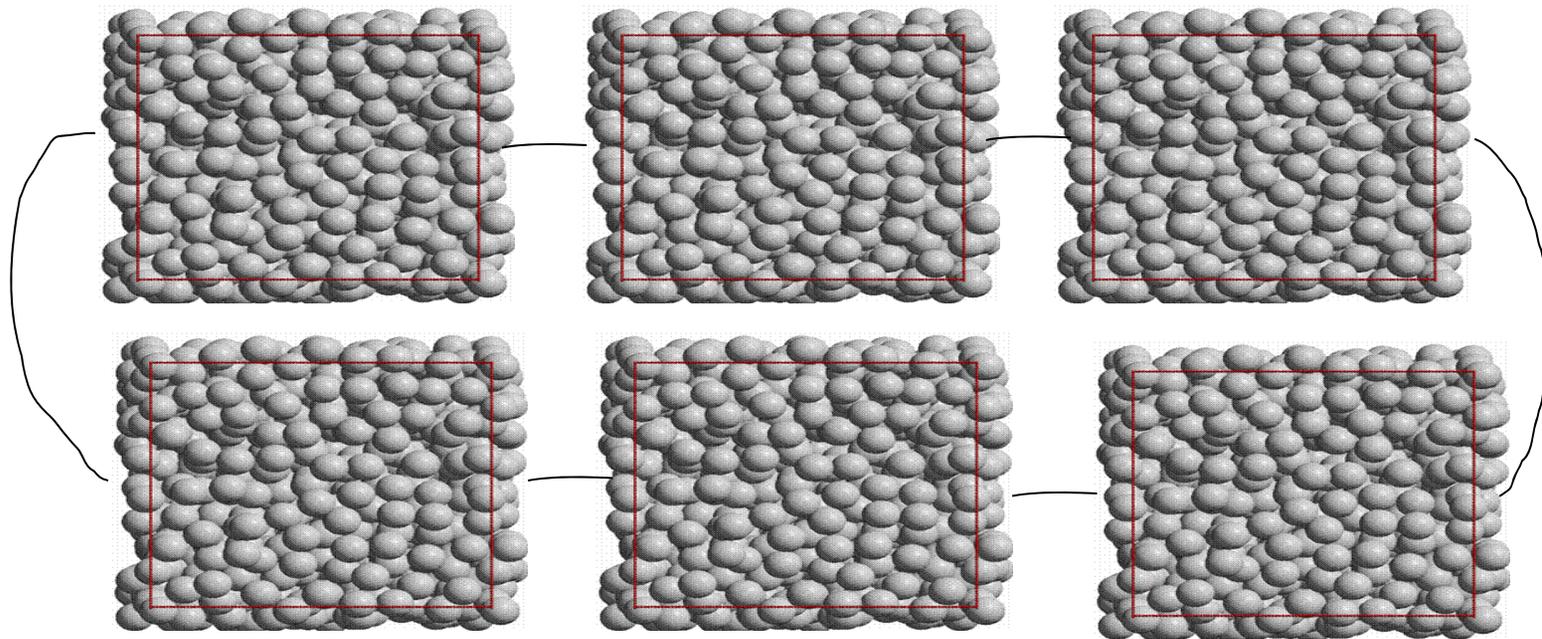
- Many simulation types require fairly uncoupled instances of existing chare arrays.
- Simulation types in this class include:
 - 1) Path Integral MD (PIMD) for nuclear quantum effects.
 - 2) k-point sampling for metallic systems.
 - 3) Spin DFT for magnetic systems.
 - 4) Replica exchange for improved atomic phase space sampling.
- A full combination of all 4 simulation is both physical and interesting

Replica Exchange : M classical subsystems each at a different temperature acting indepently

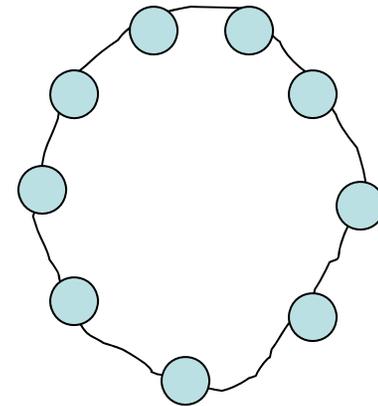


Replica exchange uber index active for all chares.
Nearest neighbor communication required to exchange temperatures
and energies

PIMD : P classical subsystems connect by harmonic bonds



Classical particle

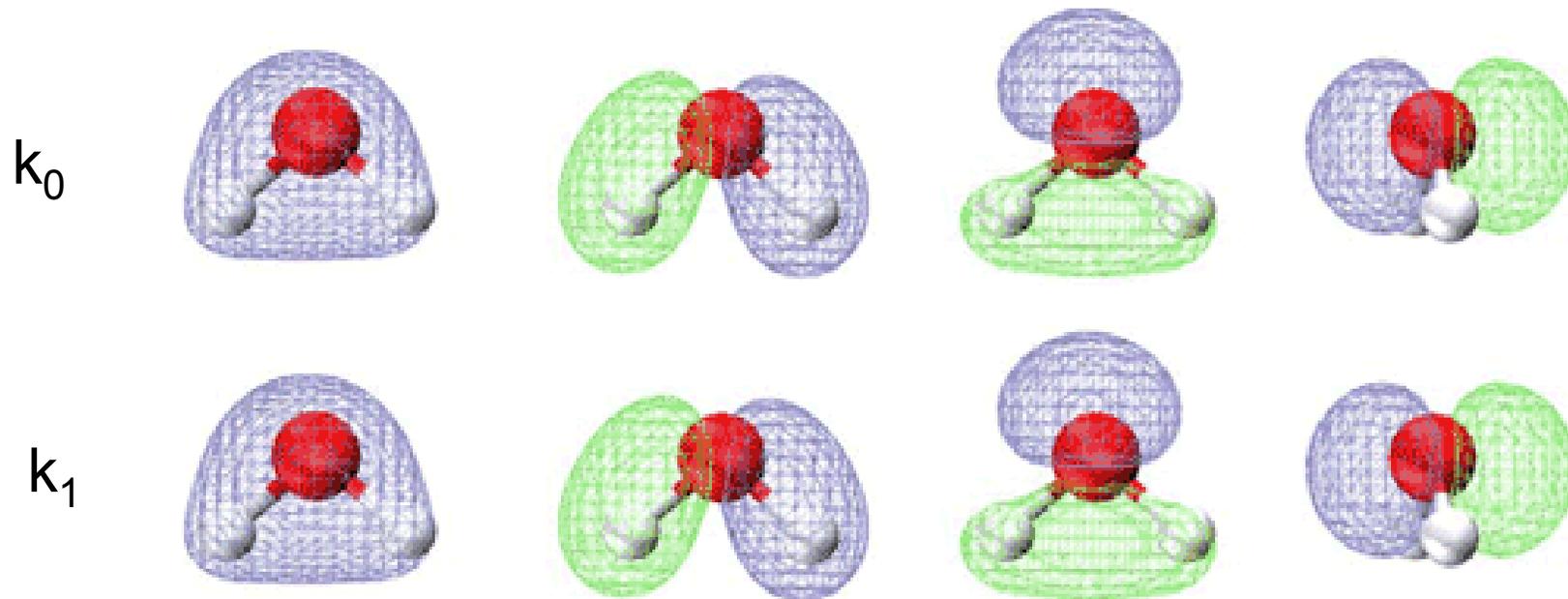


Quantum particle

PIMD uber index active for all chares.

Uber communication required to compute harmonic interactions

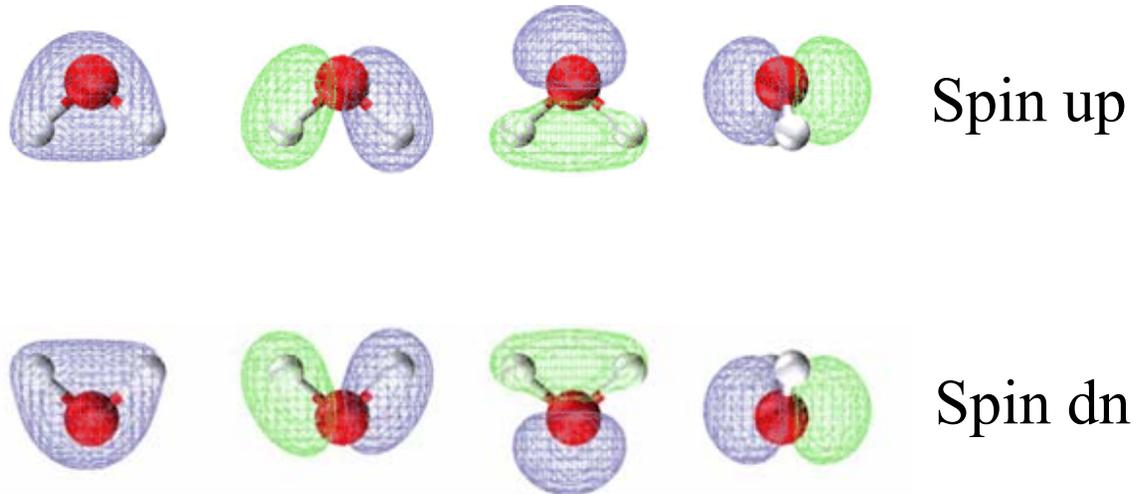
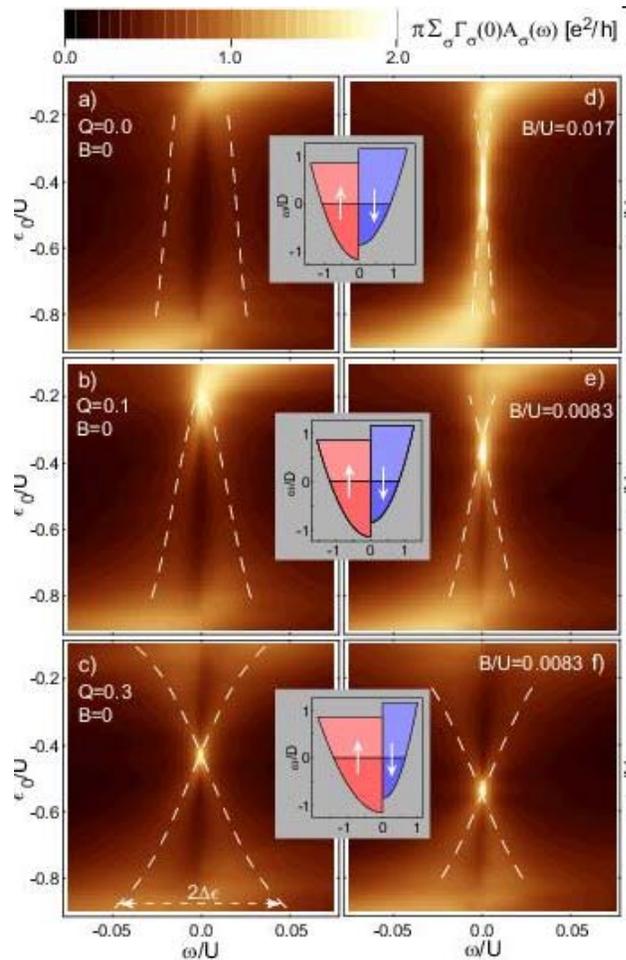
K-points : N-states are replicated
and given a different phase.



Atoms are assumed to be part of a periodic structure and are shared between the k-points (crystal momenta).

The k-point uber index is not active for atoms and electron density.
Uber reduction communication require to form the e-density and atom forces.

Spin DFT : States and electron density are given a spin-up and spin-down index.



The spin index is not active for atoms.

Under reduction communication require to form the atom forces

``Uber'' charm++ indices

- Chare arrays in OpenAtom now possess 4 uber ``instance'' indices.
- Appropriate section reductions and broadcasts across the ``Ubers'' have been enabled.
- All physics routines are working.

Describing excited electrons:

**what, why, how,
and**

what it has to do with charm++

Sohrab Ismail-Beigi

Applied Physics, Physics, Materials Science
Yale University

Density Functional Theory

For the ground-state of an interacting electron system we solve a Schrodinger-like equation for electrons

$$\left[-\frac{\hbar^2 \nabla^2}{2m_e} + V_{ion}(r) + \phi(r) + V_{xc}(r) \right] \psi_j(r) = \epsilon_j \psi_j(r)$$

Density Functional Theory

For the ground-state of an interacting electron system we solve a Schrodinger-like equation for electrons

$$\left[-\frac{\hbar^2 \nabla^2}{2m_e} + V_{ion}(r) + \phi(r) + V_{xc}(r) \right] \psi_j(r) = \epsilon_j \psi_j(r)$$

Approximations needed for $V_{xc}(r)$: LDA, GGA, etc.

Density Functional Theory

For the ground-state of an interacting electron system we solve a Schrodinger-like equation for electrons

$$\left[-\frac{\hbar^2 \nabla^2}{2m_e} + V_{ion}(r) + \phi(r) + V_{xc}(r) \right] \psi_j(r) = \epsilon_j \psi_j(r)$$

Approximations needed for $V_{xc}(r)$: LDA, GGA, etc.

Tempting: use these electron energies ϵ_j
to describe processes where
electrons change energy
(absorb light, current flow, etc.)

DFT: problems with excitations

Energy gaps (eV)

Material	LDA	Expt. [1]
Diamond	3.9	5.48
Si	0.5	1.17
LiCl	6.0	9.4

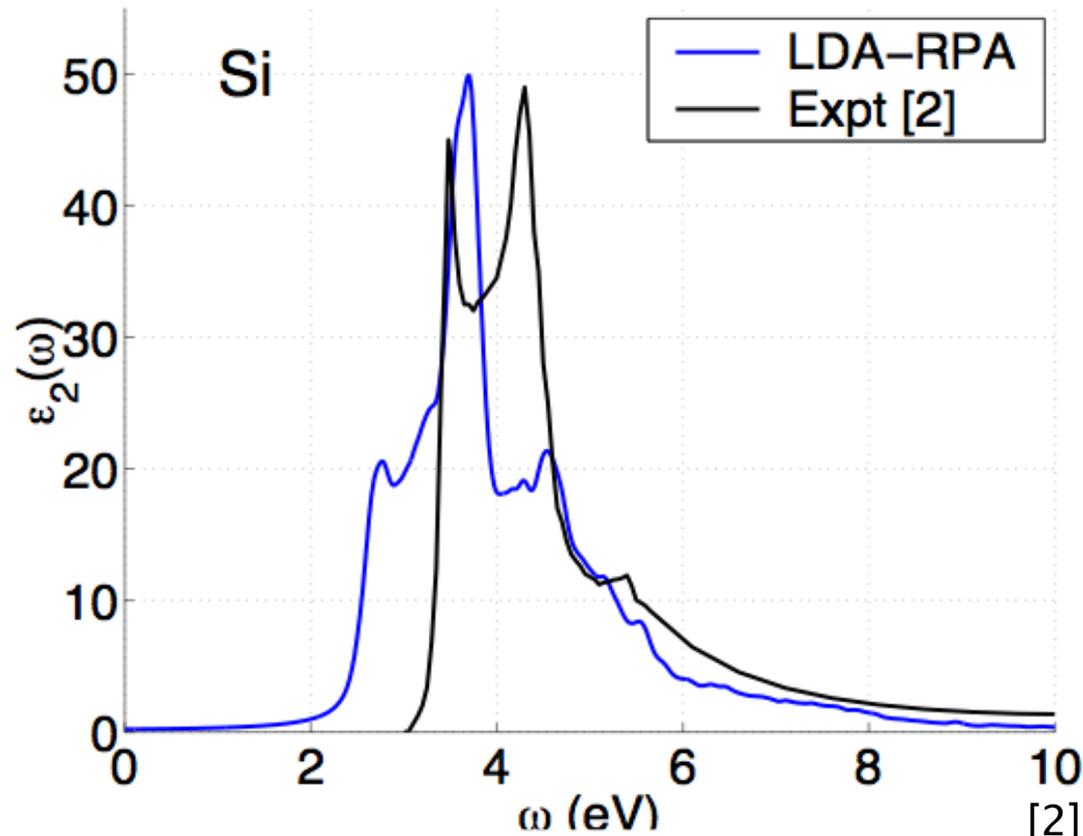
[1] Landolt-Bornstien, vol. III; Baldini & Bosacchi, Phys. Stat. Solidi (1970).

DFT: problems with excitations

Energy gaps (eV)

Material	LDA	Expt. [1]
Diamond	3.9	5.48
Si	0.5	1.17
LiCl	6.0	9.4

[1] Landolt-Bornstien, vol. III; Baldini & Bosacchi, Phys. Stat. Solidi (1970).



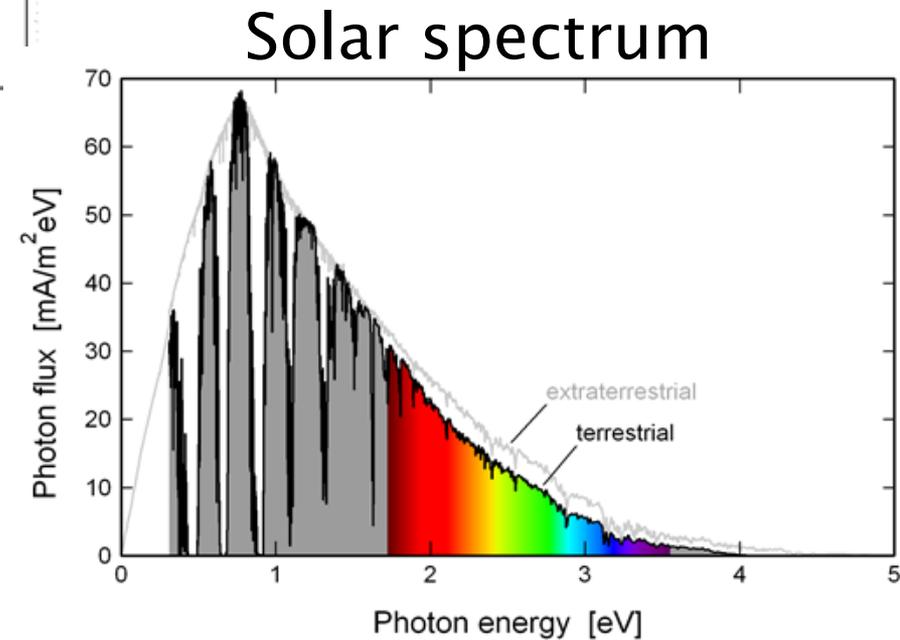
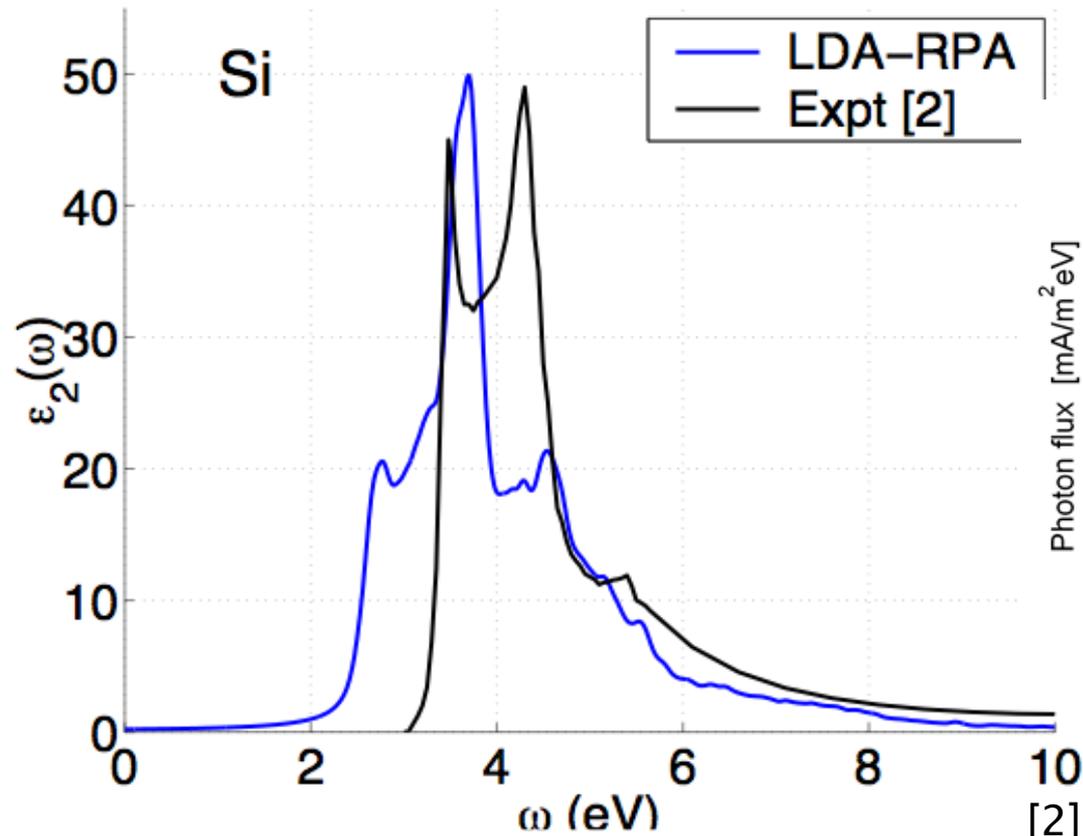
[2] Aspnes & Studna, Phys. Rev. B (1983)

DFT: problems with excitations

Energy gaps (eV)

Material	LDA	Expt. [1]
Diamond	3.9	5.48
Si	0.5	1.17
LiCl	6.0	9.4

[1] Landolt-Bornstien, vol. III; Baldini & Bosacchi, Phys. Stat. Solidi (1970).



[2] Aspnes & Studna, Phys. Rev. B (1983)

Green's functions successes

Energy gaps (eV)

Material	DFT-LDA	GW*	Expt.
Diamond	3.9	5.6	5.48
Si	0.5	1.3	1.17
LiCl	6.0	9.1	9.4

Green's functions successes

Energy gaps (eV)

Material	DFT-LDA	GW*	Expt.
Diamond	3.9	5.6	5.48
Si	0.5	1.3	1.17
LiCl	6.0	9.1	9.4

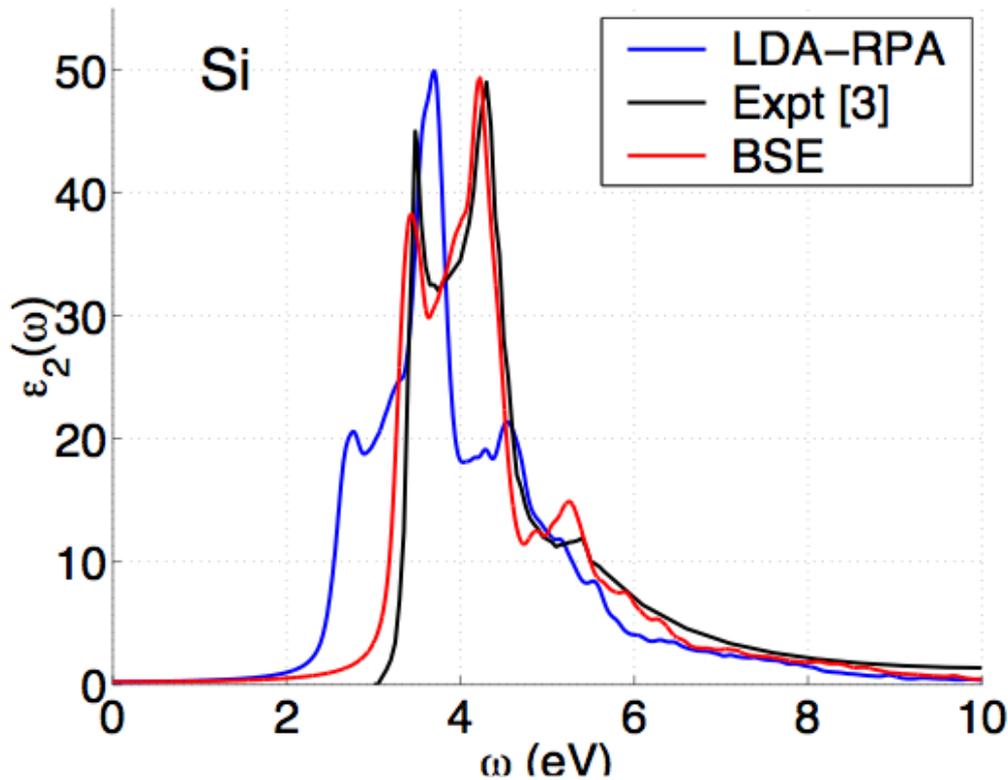
* Hybertsen & Louie, Phys. Rev. B (1986)

Green's functions successes

Energy gaps (eV)

Material	DFT-LDA	GW*	Expt.
Diamond	3.9	5.6	5.48
Si	0.5	1.3	1.17
LiCl	6.0	9.1	9.4

* Hybertsen & Louie, Phys. Rev. B (1986)

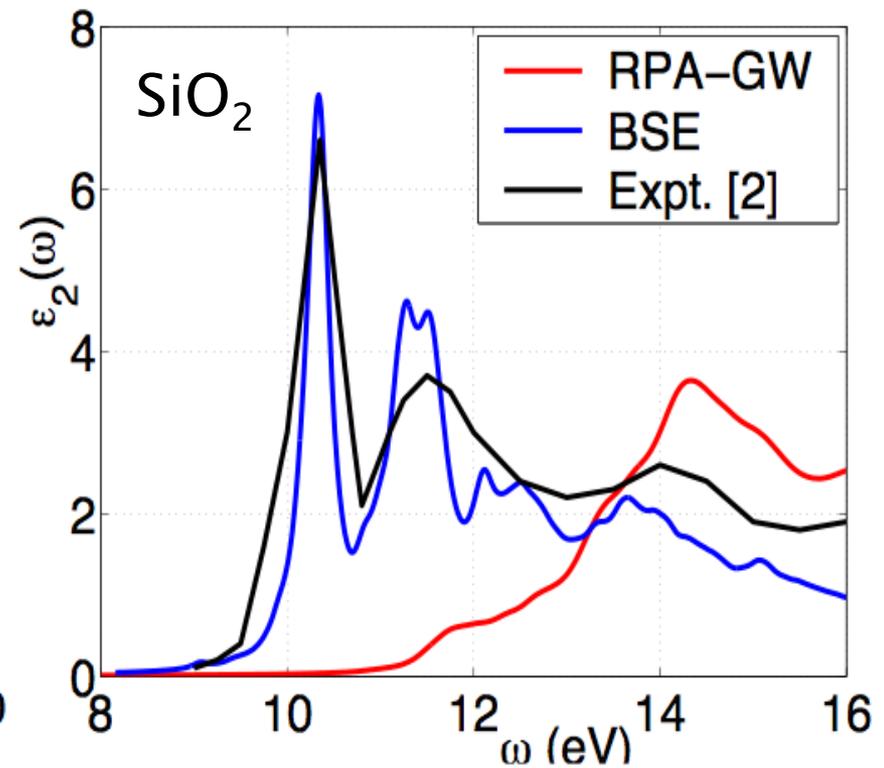
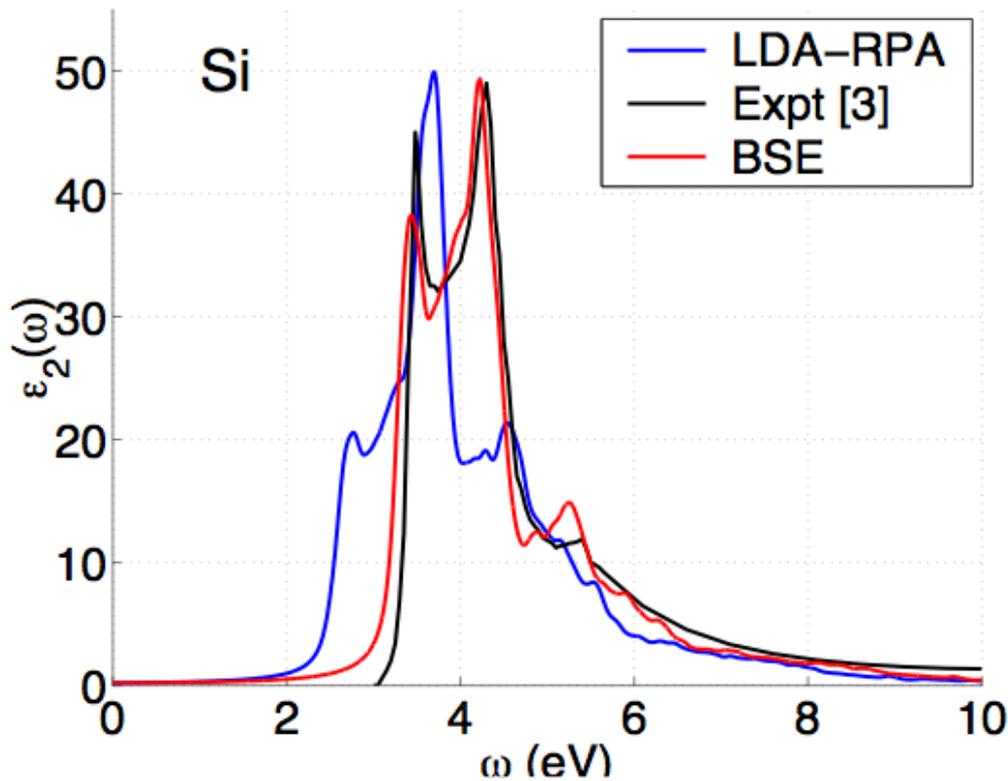


Green's functions successes

Energy gaps (eV)

Material	DFT-LDA	GW*	Expt.
Diamond	3.9	5.6	5.48
Si	0.5	1.3	1.17
LiCl	6.0	9.1	9.4

* Hybertsen & Louie, Phys. Rev. B (1986)



GW-BSE: what is it about?

DFT is a ground-state theory for electrons

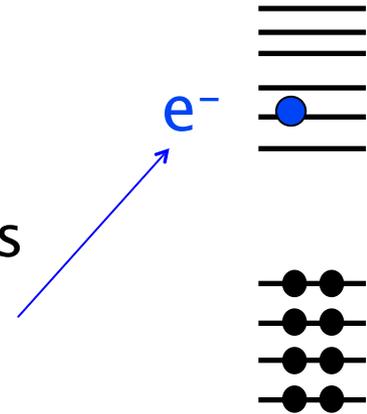
But many processes involve exciting electrons:

GW-BSE: what is it about?

DFT is a ground-state theory for electrons

But many processes involve exciting electrons:

- Transport of electrons in a material or across an interface: dynamically adding an **electron**



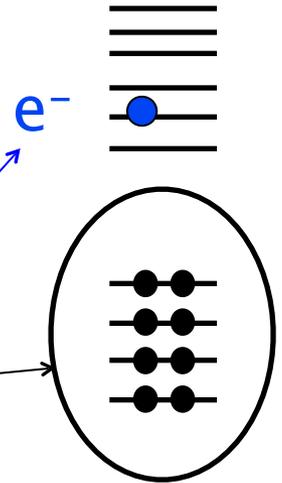
GW-BSE: what is it about?

DFT is a ground-state theory for electrons

But many processes involve exciting electrons:

- Transport of electrons in a material or across an interface: dynamically adding an **electron**

→ The other electrons respond to this and modify energy of **added electron**

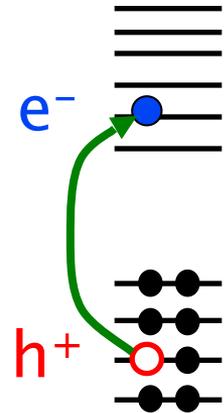


GW-BSE: what is it about?

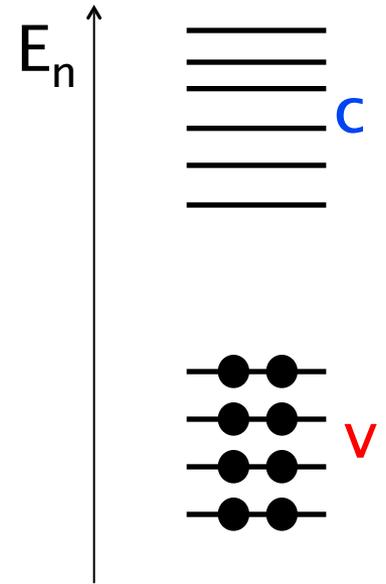
DFT is a ground-state theory for electrons

But many processes involve exciting electrons:

- Transport of electrons
- Excited electrons: optical absorption promotes **electron** to higher energy



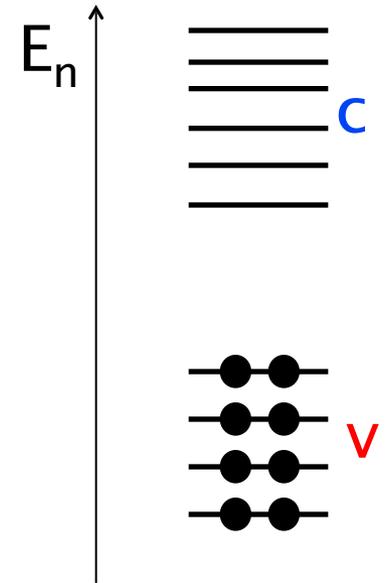
Optical excitations



Optical excitations

Single-particle view

- Photon absorbed
- one e^- kicked into an empty state



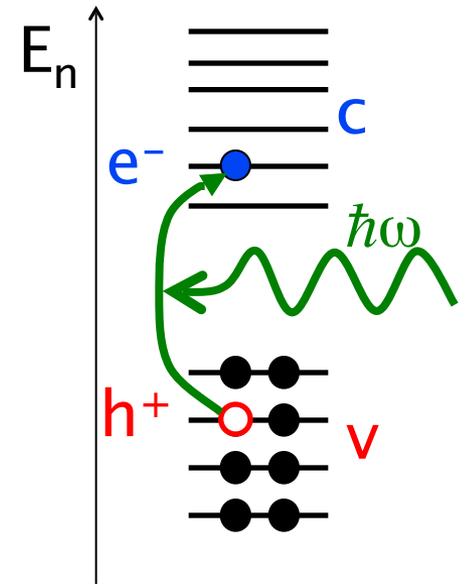
Problem:

- e^- & h^+ are charged & interact
- their motion must be correlated

Optical excitations

Single-particle view

- Photon absorbed
- one e^- kicked into an empty state



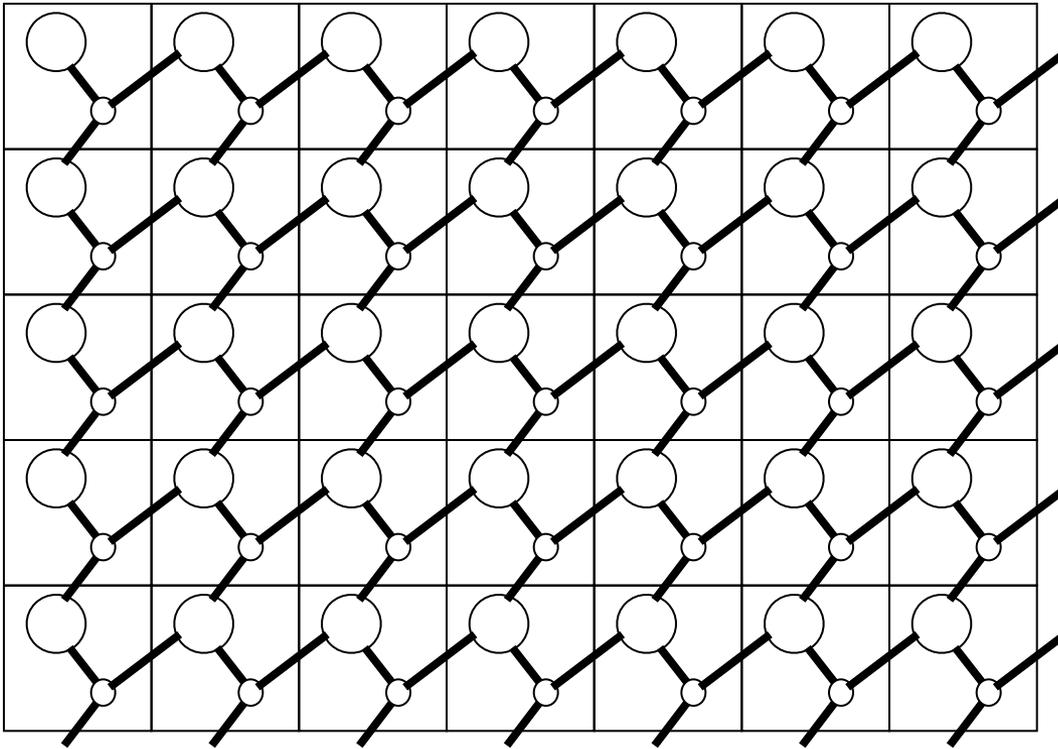
Problem:

- e^- & h^+ are charged & interact
- their motion must be correlated

Optical excitations: excitons

Exciton: correlated e^-h^+ pair excitation

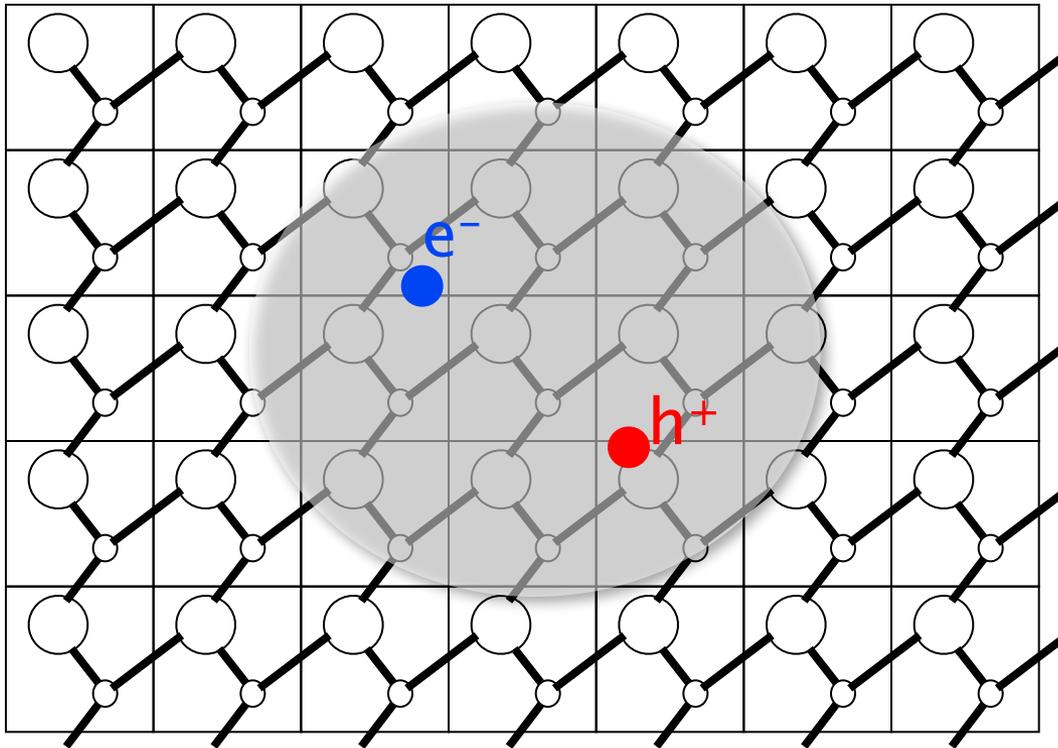
Low-energy (bound) excitons: hydrogenic picture



Optical excitations: excitons

Exciton: correlated e^-h^+ pair excitation

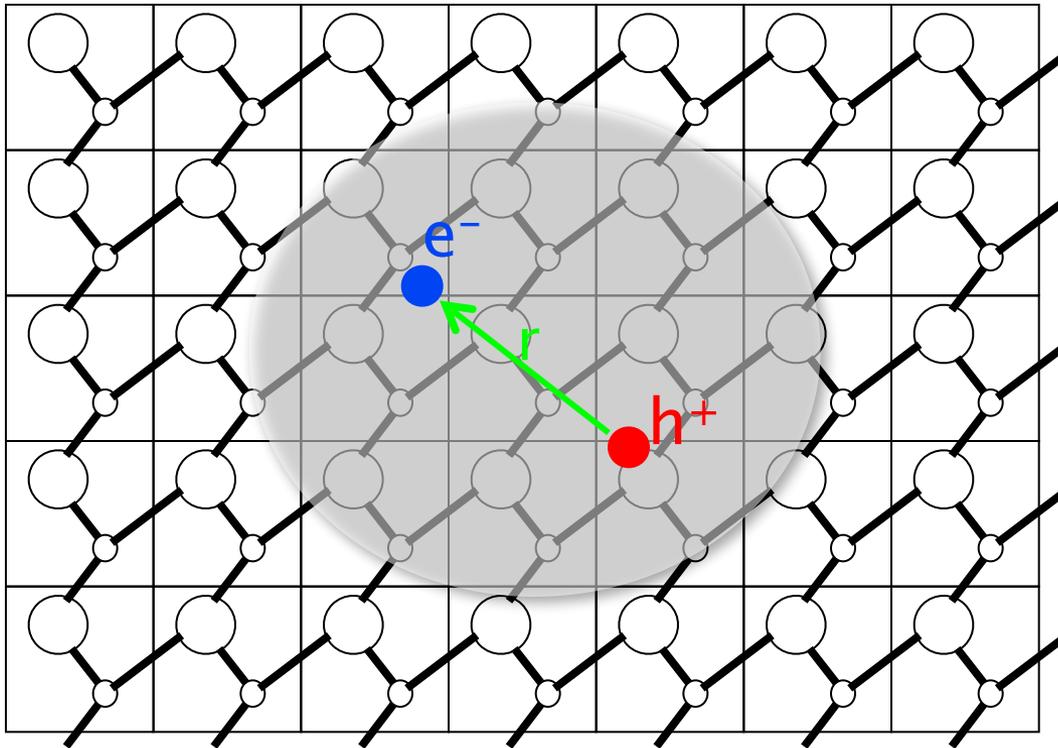
Low-energy (bound) excitons: hydrogenic picture



Optical excitations: excitons

Exciton: correlated e^-h^+ pair excitation

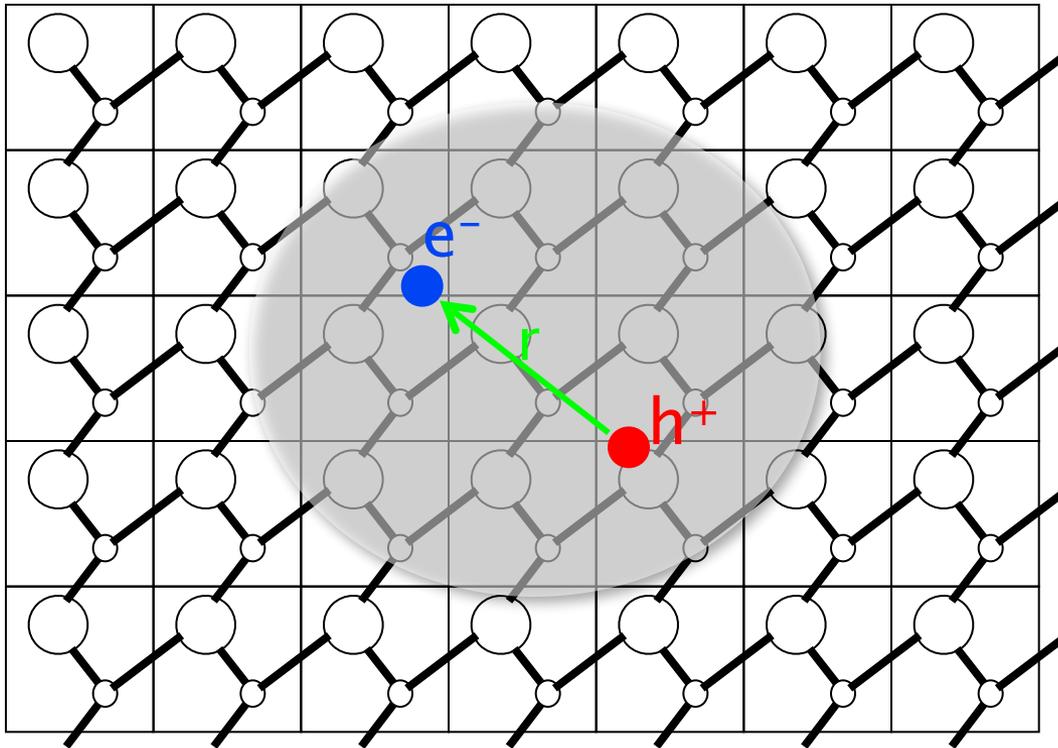
Low-energy (bound) excitons: hydrogenic picture



Optical excitations: excitons

Exciton: correlated e^-h^+ pair excitation

Low-energy (bound) excitons: hydrogenic picture



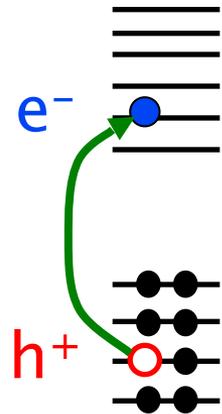
Material	r (Å)
InP	220
Si	64
SiO	4

GW-BSE: what is it about?

DFT is a ground-state theory for electrons

But many processes involve exciting electrons:

- Transport of electrons
 - Excited electrons: optical absorption promotes **electron** to higher energy
- The missing electron (**hole**) has **+** charge, attracts **electron**:
modifies excitation energy and absorption strength



GW-BSE: what is it about?

DFT is a ground-state theory for electrons

But many processes involve exciting electrons:

- Transport of electrons, electron energy levels
- Excited electrons

Each/both critical in many materials problems, e.g.

- Photovoltaics
- Photochemistry
- “Ordinary” chemistry involving electron transfer

GW-BSE: what is it for?

DFT is a ground-state theory for electrons

But many processes involve exciting electrons:

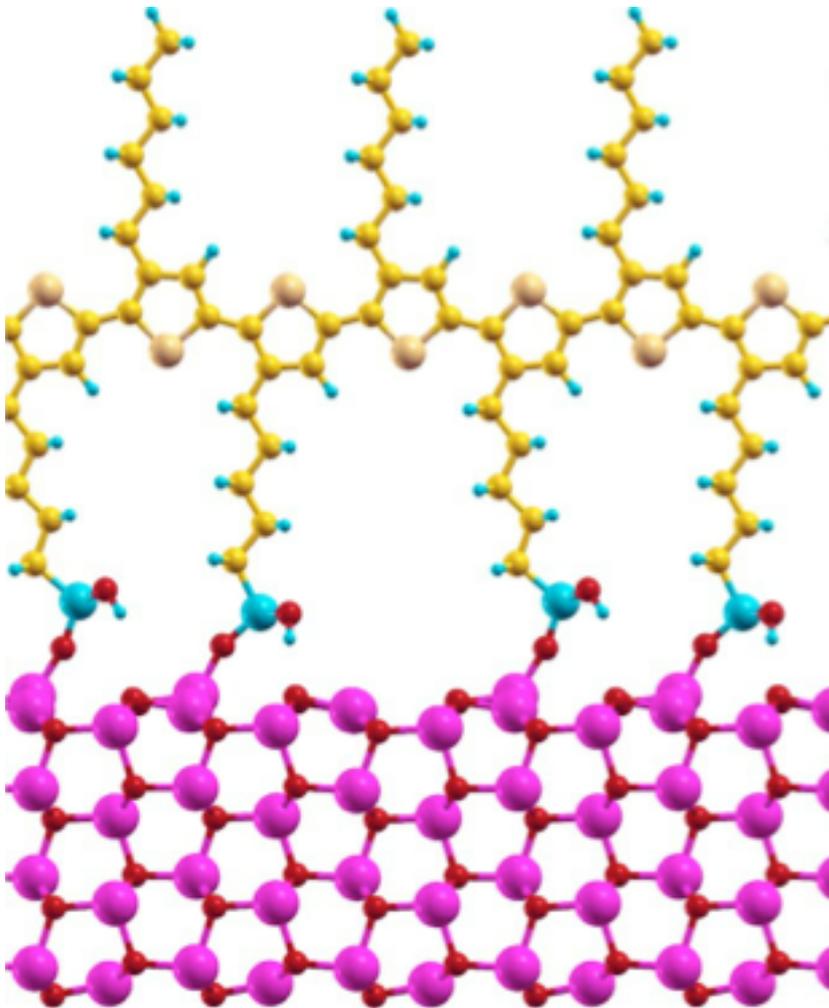
- Transport of electrons, electron energy levels
- Excited electrons

DFT --- in principle and in practice --- does a poor job of describing both

- GW : describe added electron energies
including response of other electrons
- BSE (Bethe-Salpeter Equation): describe optical processes
including electron-hole interaction and GW energies

A system I'd love to do GW-BSE on...

P3HT polymer



Zinc oxide nanowire

But with available
GW-BSE methods

it would take
“forever”

i.e. use up all my
supercomputer
allocation time

GW-BSE is expensive

Scaling with number of atoms N

- DFT : N^3
- GW : N^4
- BSE : N^6

GW-BSE is expensive

Scaling with number of atoms N

- DFT : N^3
- GW : N^4
- BSE : N^6

But in practice the GW is the killer

e.g. a system with 50–75 atoms (GaN)

- DFT : 1 cpu x hours
- GW : 91 cpu x hours
- BSE : 2 cpu x hours

GW-BSE is expensive

Scaling with number of atoms N

- DFT : N^3
- GW : N^4
- BSE : N^6

But in practice the GW is the killer

e.g. a system with 50–75 atoms (GaN)

- DFT : 1 cpu x hours
- GW : 91 cpu x hours
- BSE : 2 cpu x hours

Hence, our first focus is on GW

Once that is scaling well, we will attack the BSE

What's in the GW?

Key element : compute response of electrons to perturbation

$$P(r, r') = \frac{\partial n(r)}{\partial V(r')} = \sum_i^{\text{filled}} \sum_j^{\text{empty}} \frac{\psi_i(r)\psi_j(r)\psi_i(r')\psi_j(r')}{\epsilon_i - \epsilon_j}$$

What's in the GW?

Key element : compute response of electrons to perturbation

$$P(r, r') = \frac{\partial n(r)}{\partial V(r')} = \sum_i^{\text{filled}} \sum_j^{\text{empty}} \frac{\psi_i(r)\psi_j(r)\psi_i(r')\psi_j(r')}{\epsilon_i - \epsilon_j}$$

$P(r, r')$ = Response of electron density $n(r)$ at position r
to change of potential $V(r')$ at position r'

What's in the GW?

Key element : compute response of electrons to perturbation

$$P(r, r') = \frac{\partial n(r)}{\partial V(r')} = \sum_i^{\text{filled}} \sum_j^{\text{empty}} \frac{\psi_i(r)\psi_j(r)\psi_i(r')\psi_j(r')}{\epsilon_i - \epsilon_j}$$

$P(r, r')$ = Response of electron density $n(r)$ at position r
to change of potential $V(r')$ at position r'

Challenges

1. Many FFTs to get wave functions  $\psi_i(r)$ functions
2. Large outer product to form P
3. Dense r grid : $P(r, r')$ is huge in memory
4. Sum over j is very large

What's in the GW?

Key element : compute response of electrons to perturbation

$$P(r, r') = \frac{\partial n(r)}{\partial V(r')} = \sum_i^{\text{filled}} \sum_j^{\text{empty}} \frac{\psi_i(r)\psi_j(r)\psi_i(r')\psi_j(r')}{\epsilon_i - \epsilon_j}$$

$P(r, r')$ = Response of electron density $n(r)$ at position r
to change of potential $V(r')$ at position r'

Challenges

1. Many FFTs to get wave functions  $\psi_i(r)$ functions
2. Large outer product to form P
3. Dense r grid : $P(r, r')$ is huge in memory
4. Sum over j is very large

1 & 2 : Efficient parallel FFTs and linear algebra

3 : Effective memory parallelization

4 : replace explicit j sum by implicit inversion

Summary

GW-BSE is promising as it contains the right physics

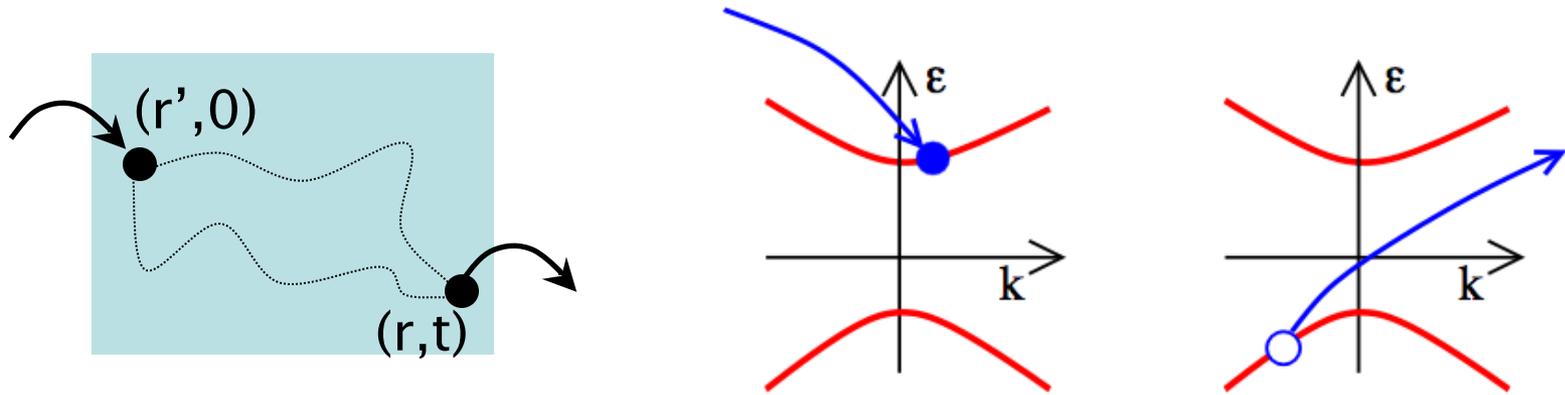
Very expensive : computation and memory

Plan to implement high performance version in
OpenAtom for the community (SI2-SSI NSF grant)

Two sets of challenges

- How to best parallelize existing GW-BSE algorithms?
Will rely on Charm++ to deliver high performance
Coding, maintenance, migration to other computers
much easier for user
- Need to improve GW-BSE algorithms to use the computers
more effectively (theoretical physicist/chemist's job)

One particle Green's function



$$G_1(r, r', \omega) = \sum_j \frac{\psi_j(r) \psi_j(r')^*}{\omega - \epsilon_j}$$

Dyson Equation:

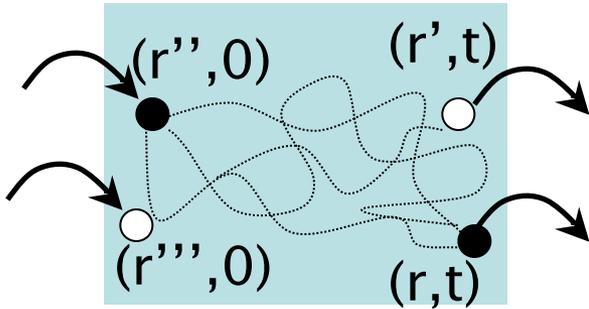
$$\left[-\frac{\nabla^2}{2} + V_{ion}(r) + V_H(r) + \Sigma(r, r', \epsilon_j) \right] \psi_j(r) = \epsilon_j \psi_j(r)$$

DFT:

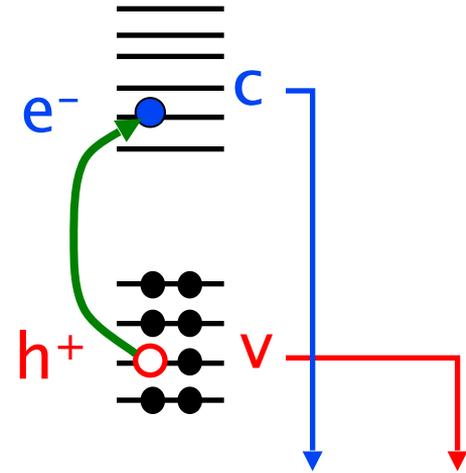
$$\Sigma \approx iG_1W \quad , \quad W = \epsilon^{-1}(\omega) * v_c \quad (RPA)$$

$$\left[-\frac{\nabla^2}{2} + V_{ion}(r) + V_H(r) + V_{xc}(r) \right] \psi_j(r) = \epsilon_j \psi_j(r)$$

Two particle Green's function



$$G_2(\omega) = \sum_S \frac{\Phi_S \cdot \Phi_S^*}{\omega - \Omega_S}$$



Exciton amplitude:

Bethe-Salpeter Equation: $HA^S = \Omega_S A^S$
(BSE)

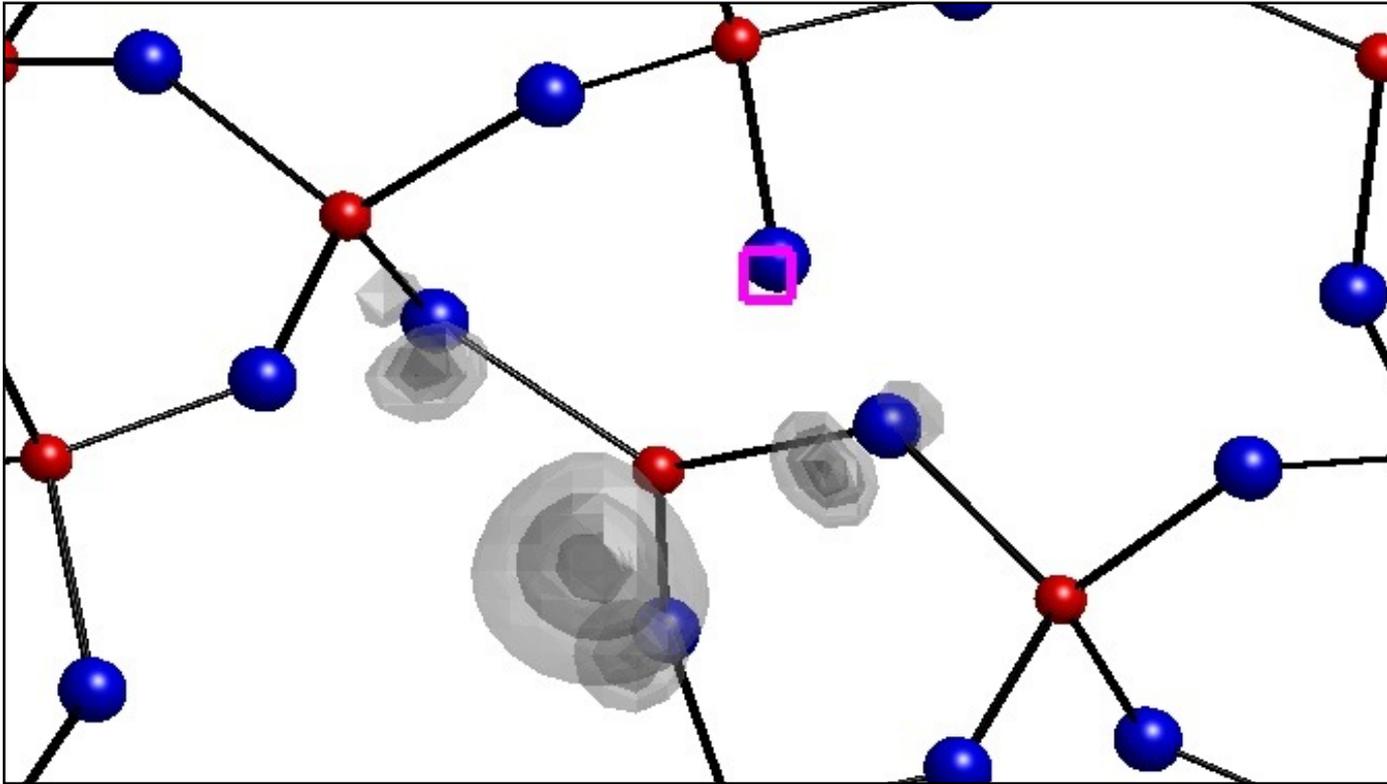
$$H = \epsilon_c - \epsilon_v + K_{int}$$

$$\langle K_{int} \rangle = - \int dr \int dr' |\Phi_S(r, r')|^2 W(r, r') \quad \text{attractive (screened direct)}$$

$$+ \int dr \int dr' \Phi_S^*(r, r) \Phi_S(r', r') v_c(r - r') \quad \text{repulsive (exchange)}$$

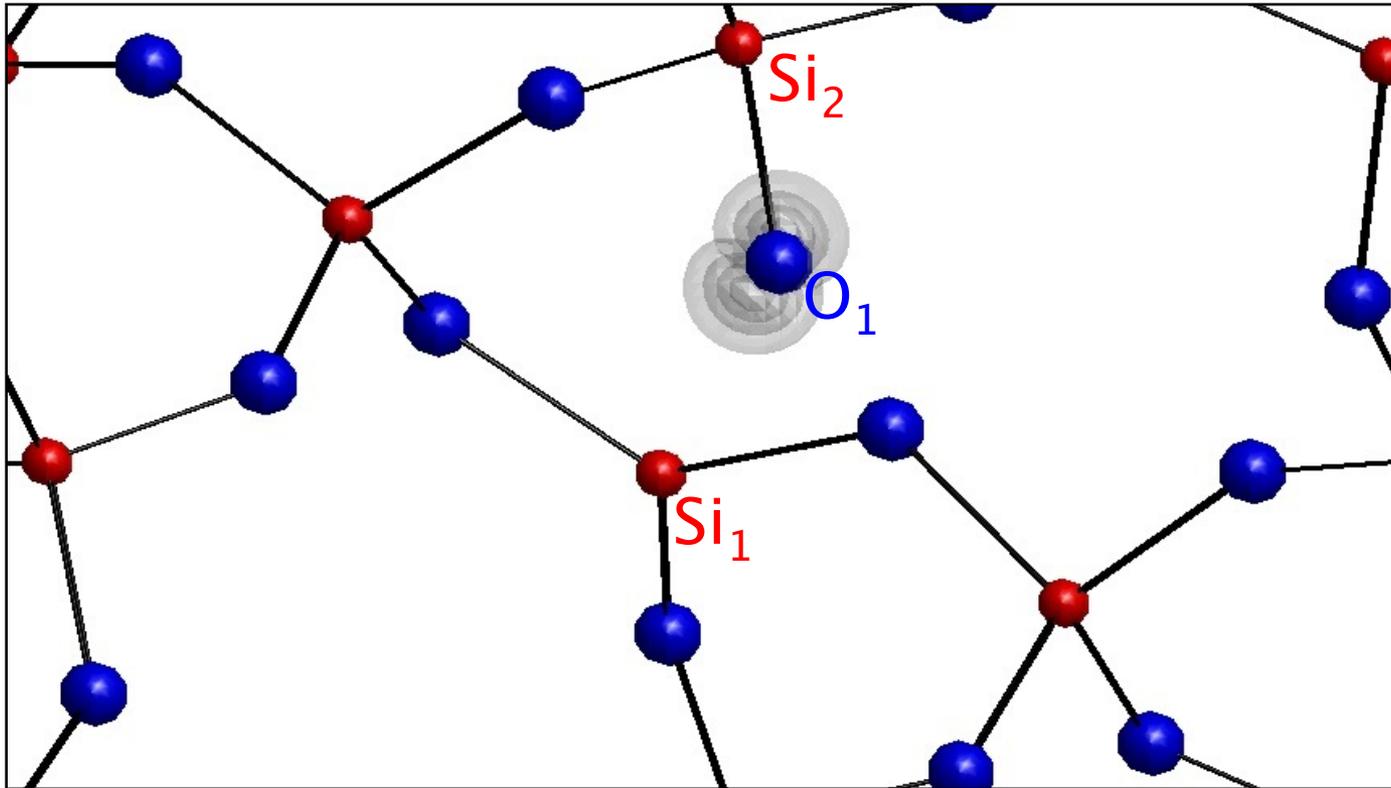
STE geometry

Prob : 20,40,60,80% max



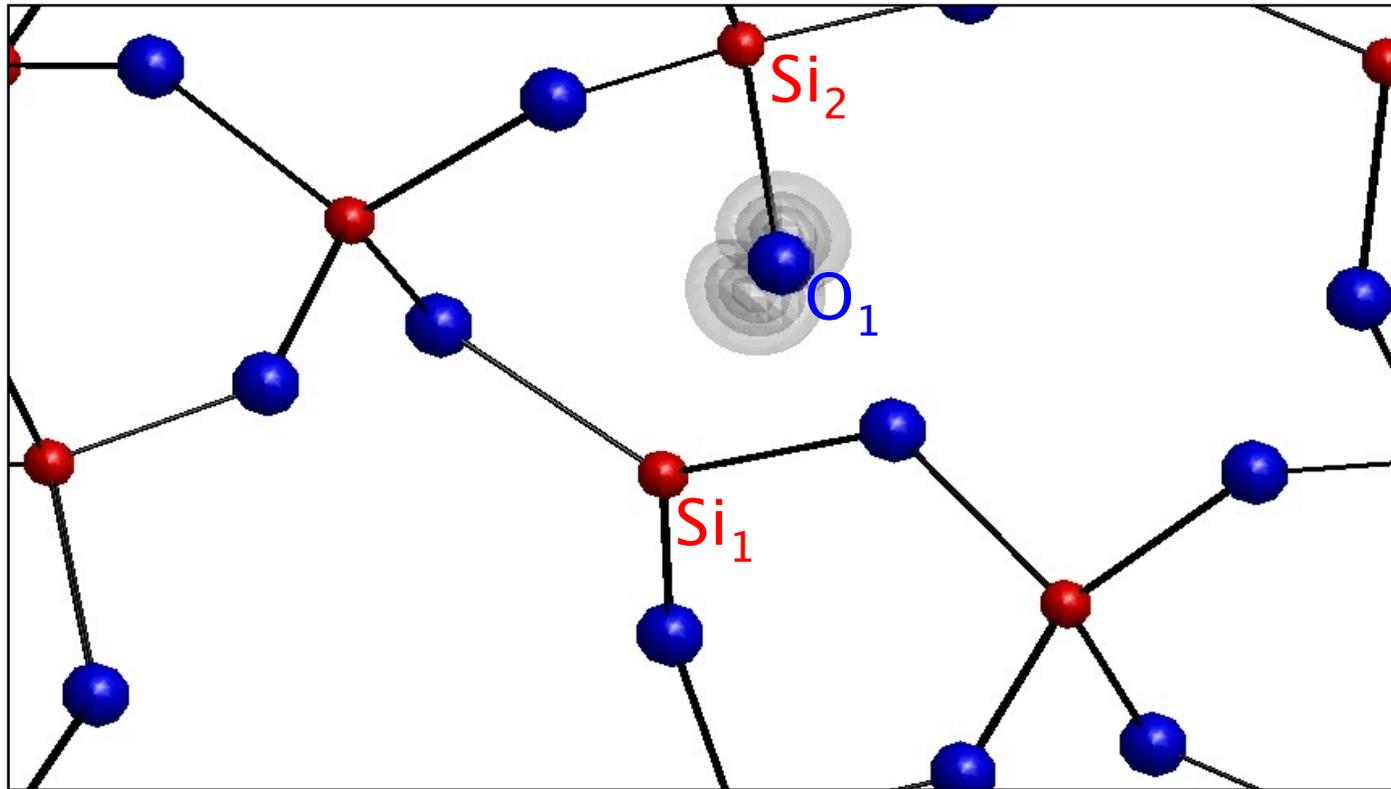
STE geometry

Prob : 20,40,60,80% max



STE geometry

Prob : 20,40,60,80% max



Bond (Å)	Bulk	STE
Si	1.60	1.97 (+23%)
Si	1.60	1.68 (+5%)
Si	1.60	1.66 (+4%)

Angles	Bulk	STE
O	109	≈ 85
O	109	≈ 120

Exciton self-trapping

Defects → localized states: exciton can get trapped

Interesting case: self-trapping

- If exciton in ideal crystal can lower its energy by localizing
 - defect forms spontaneously
 - traps exciton

Exciton self-trapping

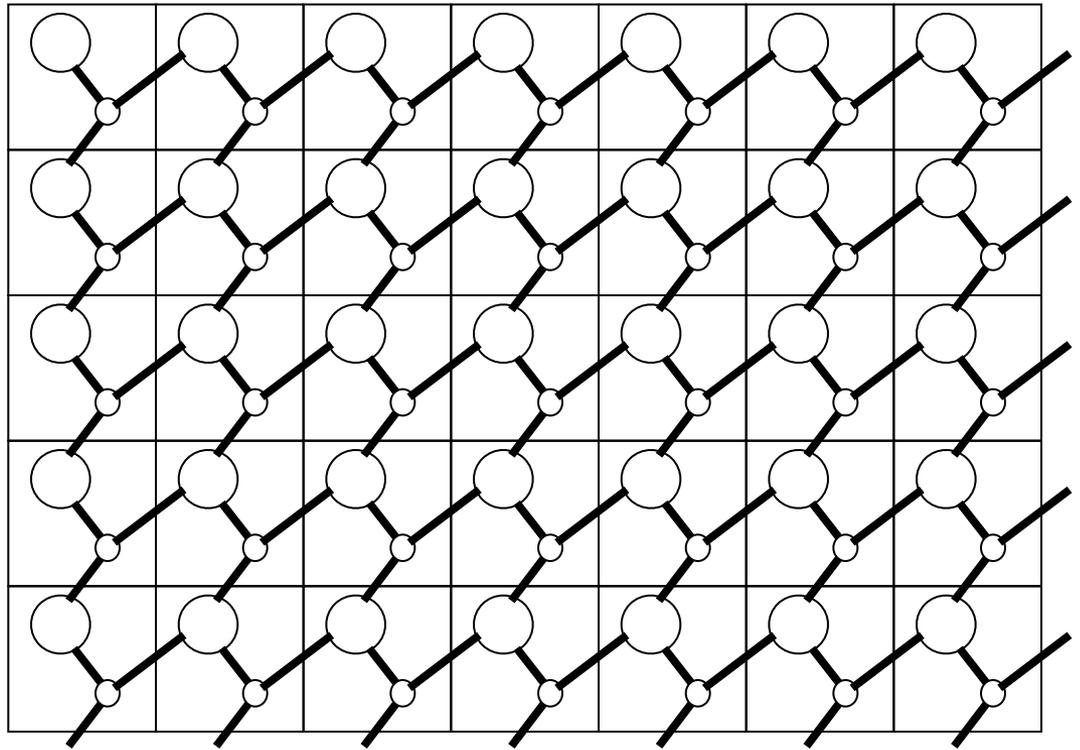
Defects → localized states: exciton can get trapped

Interesting case: self-trapping

- If exciton in ideal crystal can lower its energy by localizing

→ defect forms spontaneously

→ traps exciton



Exciton self-trapping

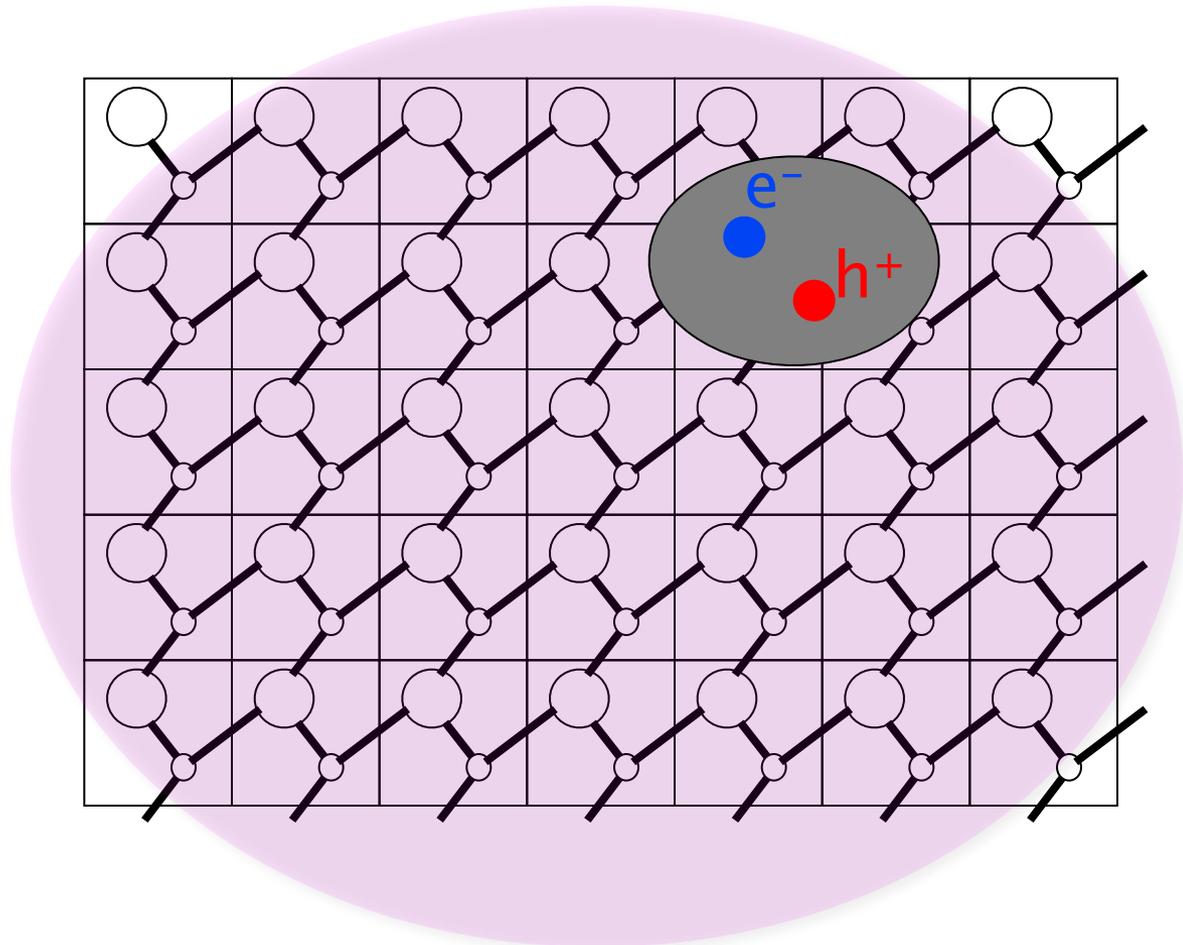
Defects → localized states: exciton can get trapped

Interesting case: self-trapping

- If exciton in ideal crystal can lower its energy by localizing

→ defect forms spontaneously

→ traps exciton



Exciton self-trapping

Defects → localized states: exciton can get trapped

Interesting case: self-trapping

- If exciton in ideal crystal can lower its energy by localizing

→ defect forms spontaneously

→ traps exciton

