Scalable Computational Tools for Discovery and Design: Excited State Phenomena in Energy Materials

James R. Chelikowsky
Center for Computational Materials
Institute for Computational Engineering and Sciences
Departments of Physics and Chemical Engineering
University of Texas
Austin, TX 78712

CCM
CENTER FOR COMPUTATIONAL MATERIALS

The Institute for Computational Engineering and Sciences
ICES

ICPS 2014
AUSTIN TEXAS.USA

Principal Investigators

Jim Chelikowsky
University of Texas at Austin

Jeff Neaton
Lawrence Berkeley National Lab
University of California

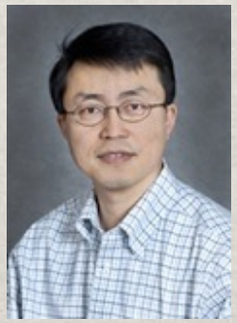
Steve Louie
Lawrence Berkeley National Lab
University of California

Alex Demkov
University of Texas at Austin

Yousef Saad
University of Minnesota

Andrew Canning
Lawrence Berkeley
National Lab

Jack Deslilppe
National Energy Research
Scientific Computing Center



Chao Yang
Lawrence Berkeley
National Lab
FASTMath Institute

The objective of our proposed work is to develop and implement new methods and theories to predict electronic excited state phenomena in energy related materials, e.g., materials for photovoltaics, photocatalysis, and electrical energy storage.

Photovoltaic Materials: Hot Carriers

- Main source of energy loss in solar cells. The dynamics of hot carriers (not in equilibrium) is central to many energy photovoltaic conversion processes
- Characterization of hot carriers has long been a challenge even for the simplest materials
- May be generated by injection or optical excitations

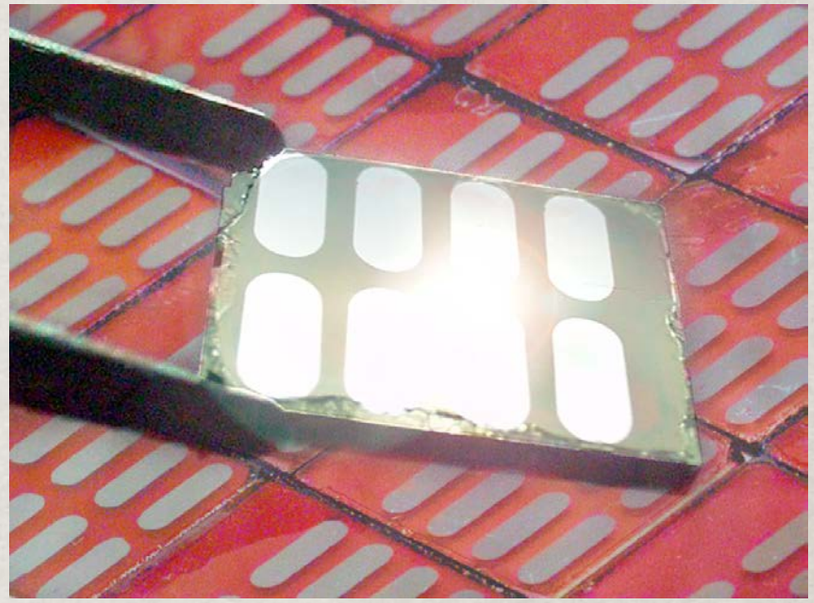
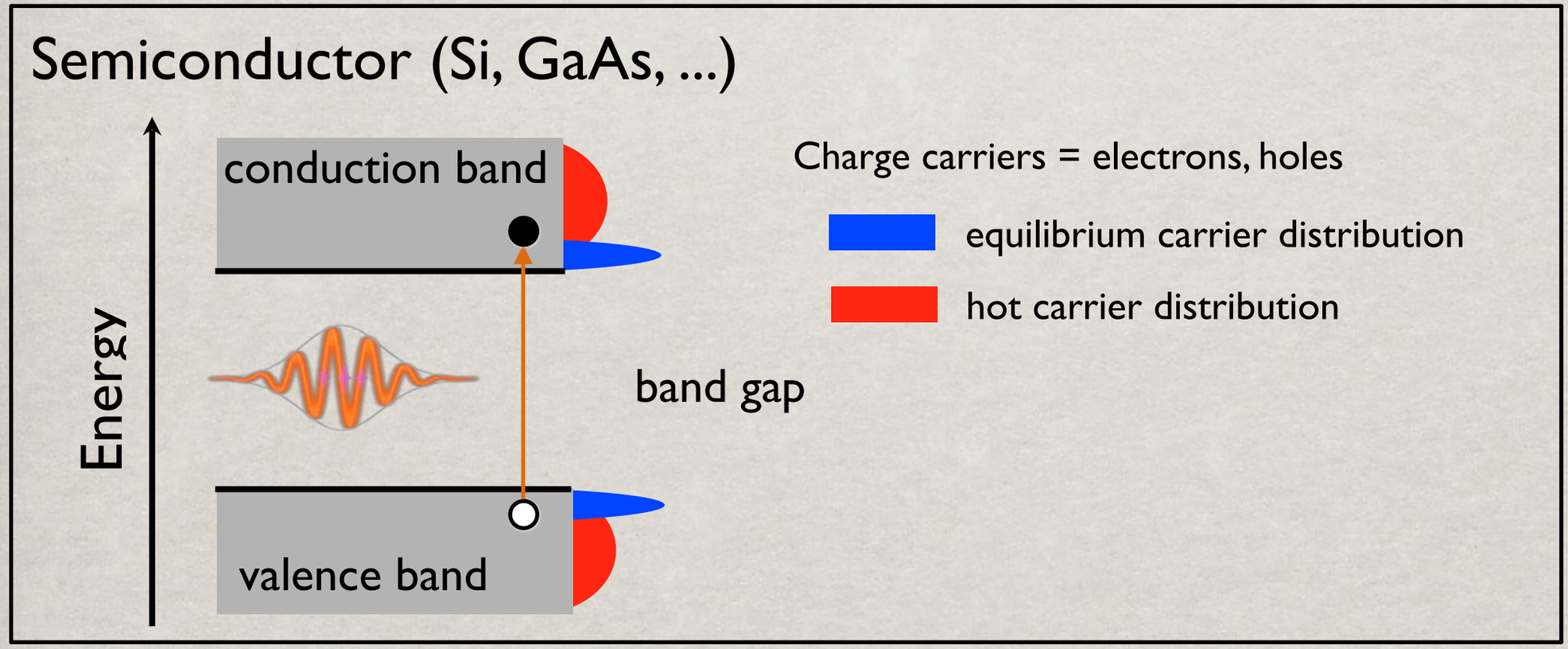


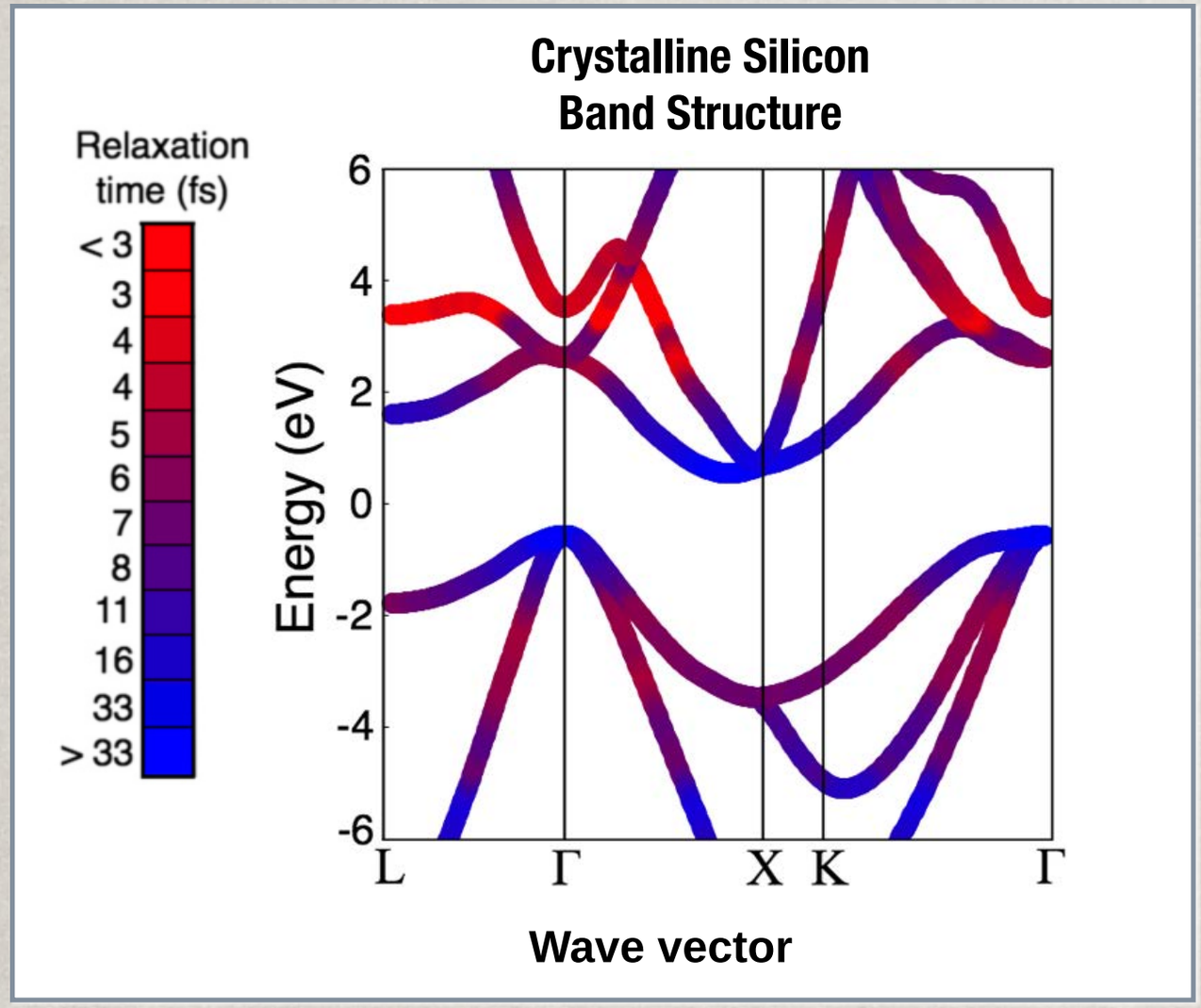
Photo by Roy Kaltschmidt

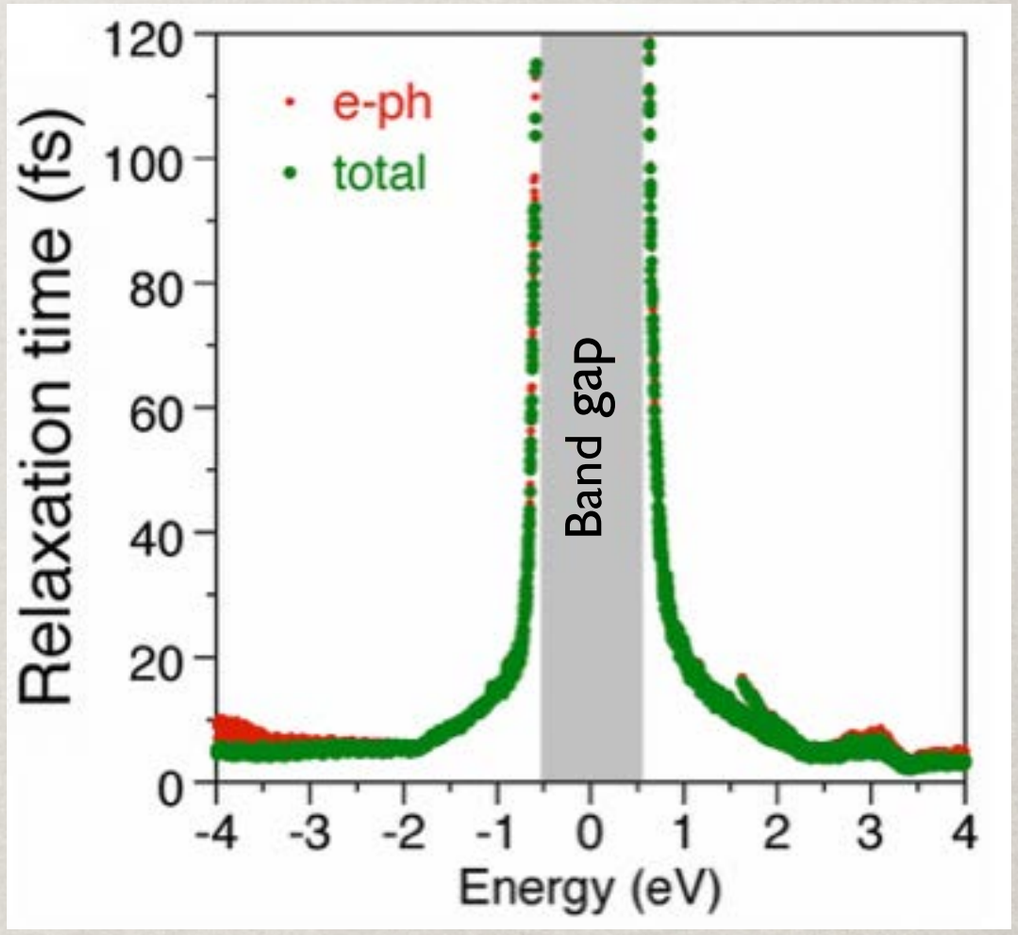


Relaxation time

How long it takes for hot carriers to lose energy

- Fast relaxation away from band edge: ~10 fs
- Slower relaxation near the band edge: >100 fs





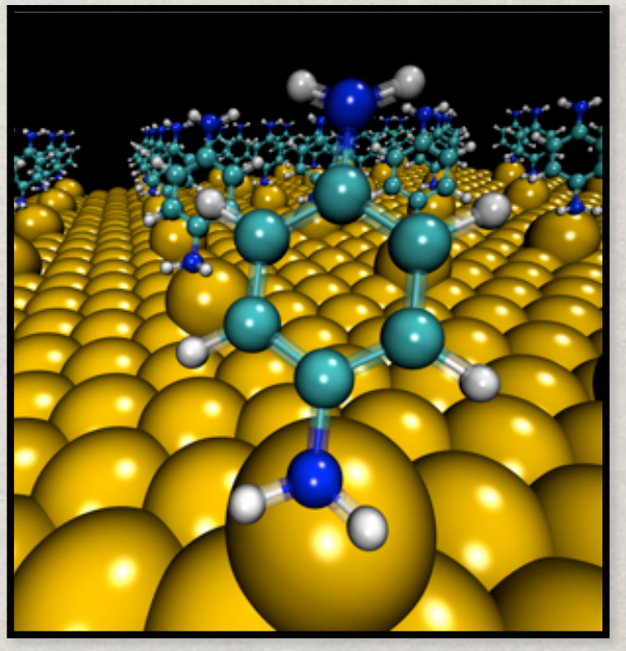
M. Bernardi, D. Vigil-Fowler, J. Lischner, J. B. Neaton and S.G. Louie, Phys. Rev. Lett. 112, 257402 (2014)

COMPUTATIONAL METHODS

Solving an eigenvalue problem for large systems. Determine structural properties and input for response functions and excited state computation.

Targets: Complex molecular systems (organic semiconductors) and nanoscale systems related to energy materials.

Physical approximations: Density functional theory (simplify many body problem) and pseudopotentials (set energy and length scales to valence states).



Target solving the Kohn-Sham problem for large scale systems

THe Kohn-Sham Problem:

$$\left[-\frac{\hbar^2\nabla^2}{2m}+V_{ion}\left(\vec{r}\right)+V_H\left(\vec{r}\right)+V_{xc}\left(\vec{r}\right)\right]\psi_n\left(\vec{r}\right)=E_n\psi_n\left(\vec{r}\right)$$

Charge density: $\rho\left(\vec{r}\right)=e\sum_{occup}\left|\psi_n\left(\vec{r}\right)\right|^2$

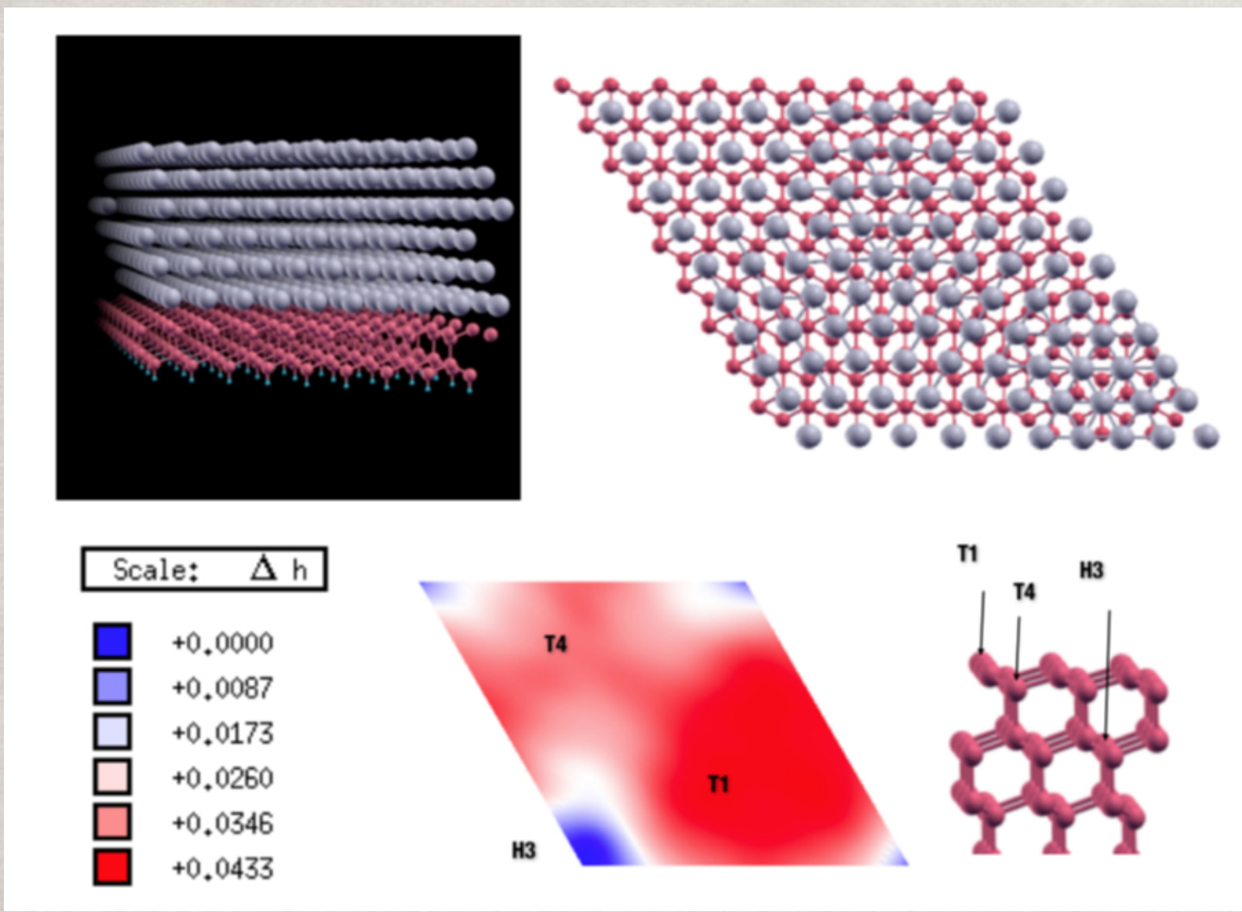
Hartree-Potential: $V_H\left(\vec{r}\right)=e\int\frac{\rho\left(\vec{r}'\right)}{\left|\vec{r}'-\vec{r}\right|}d^3r'$

Effective many body potential: $V_{xc}\left(\vec{r}\right)=V_{xc}\left[\rho\left(\vec{r}\right)\right]$

Electronic energy from a solution to the Kohn-Sham Problem

$$E_{Total} = \frac{1}{2}\sum_{i,j;i\neq j}\frac{Z_i Z_j e^2}{|\vec{R}_i - \vec{R}_j|} + E_{electronic}$$

$$E_{electronic} = \sum_{n,occup} E_n - \frac{1}{2}\int d^3r\, V_H \rho + \int d^3r \left[\varepsilon_{xc} - V_{xc}\right]\rho$$



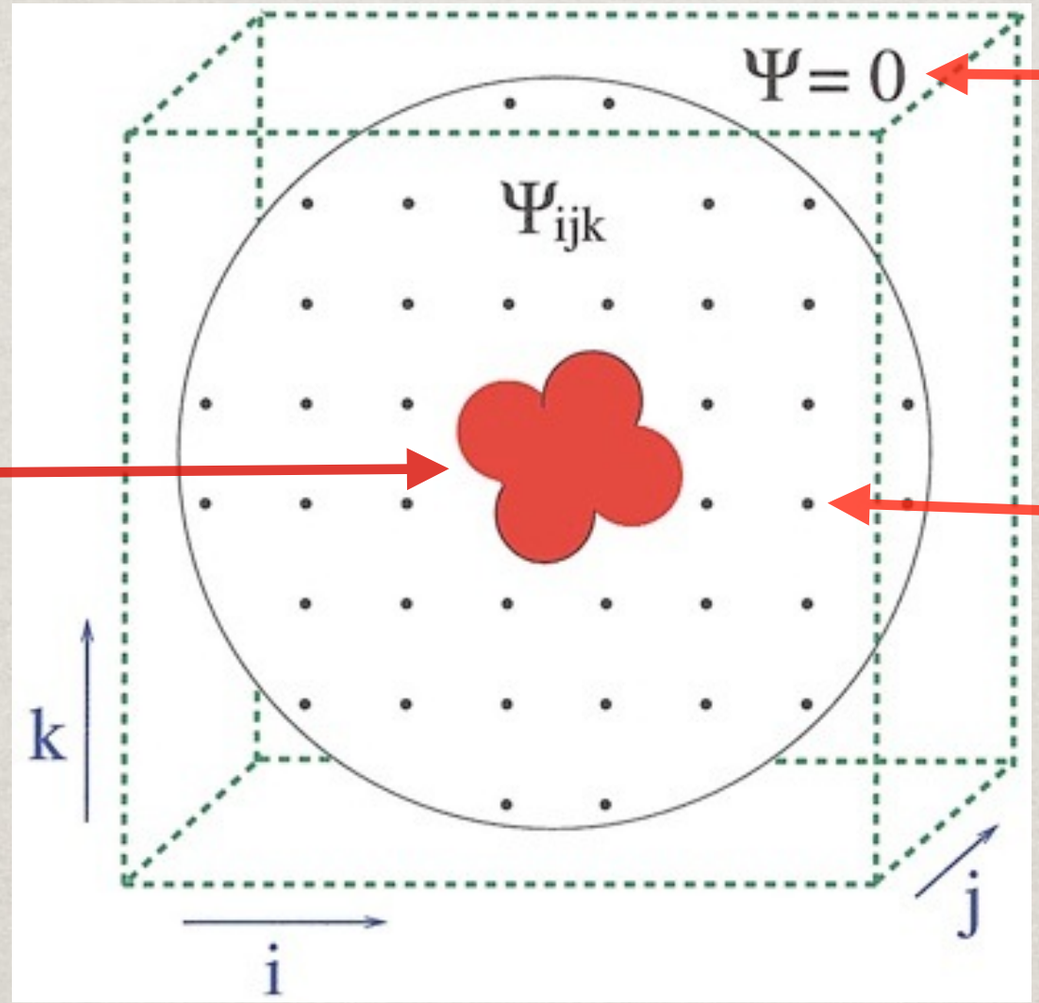
Example: Understanding the structure of a metal-semiconductor interface (Pb on Si). Lattice mismatch has 10 Pb atoms for every 9 Si atoms.

1500 atoms were treated and the structure was fully relaxed.

Algorithms: Discretize the Kohn-Sham problem

Wave function vanishes outside the domain

System of interest (nanocrystal)



Variable of interest is wave function on a grid point

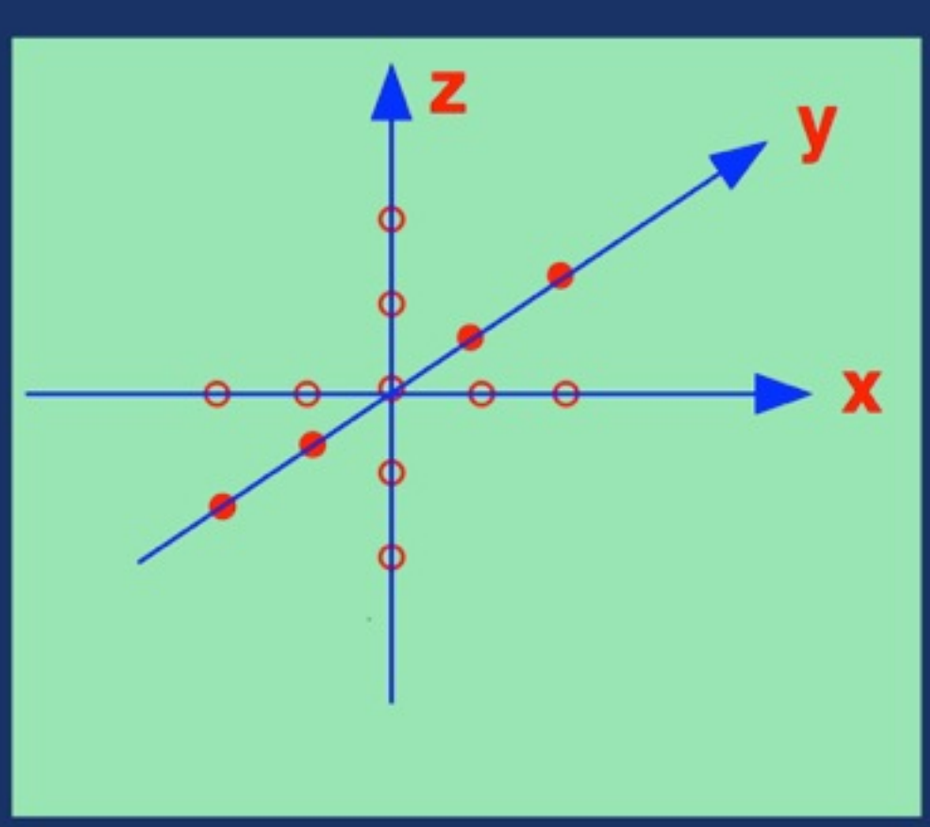
Discretize Kohn-Sham Equation: Solve using high-order finite differencing
Goals: Ease of implementation. Minimize global communications.

Real-space Finite Difference Methods

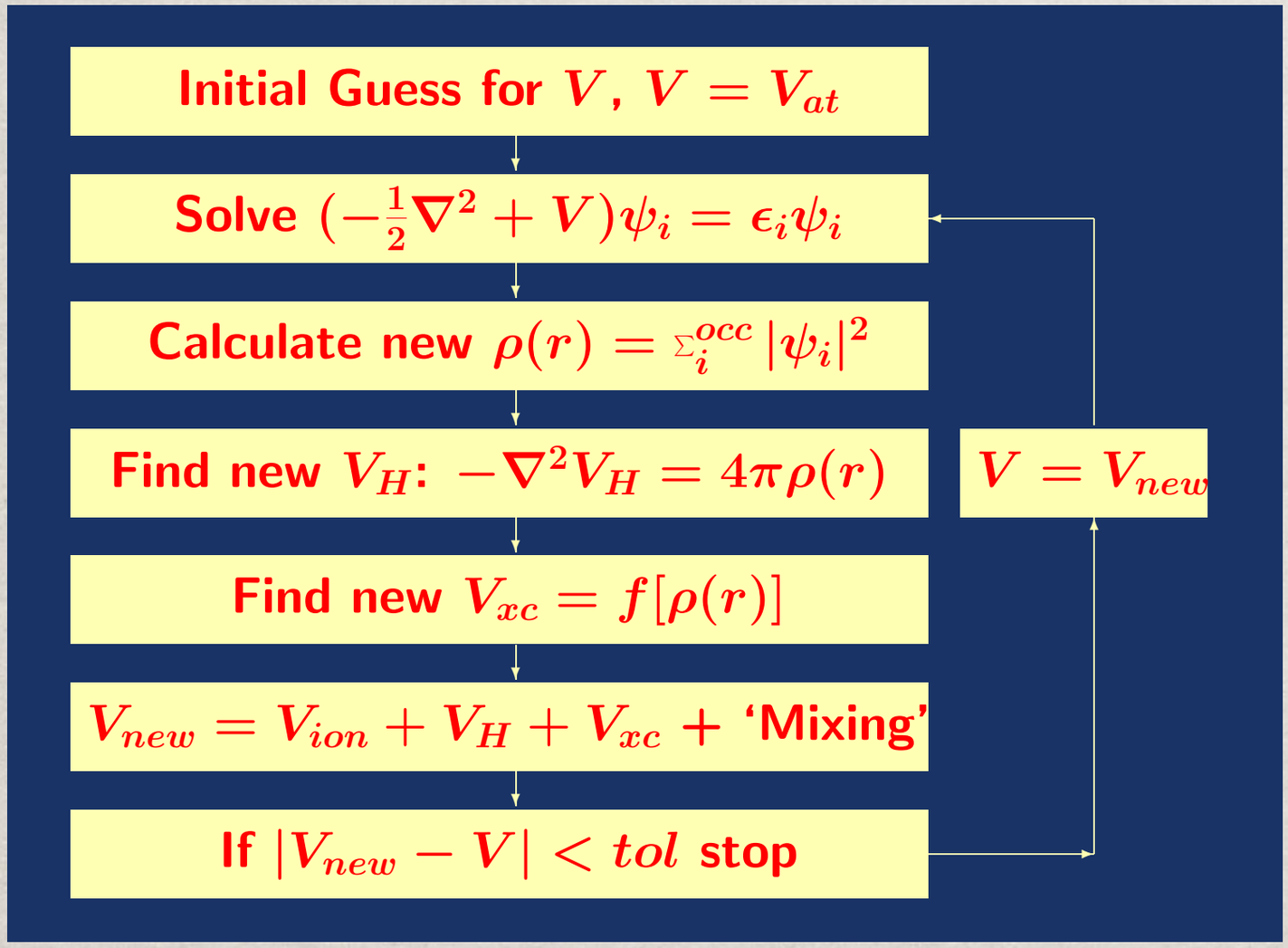
- Use High-Order Finite Difference Methods [Fornberg & Sloan '94]
- Typical Geometry = Cube – regular structure.
- Laplacian matrix need not even be stored.

Order 4 Finite Difference Approximation:

$$\left.\frac{\partial^2 \Psi}{\partial x^2}\right|_{x=x_0} = \frac{1}{h^2} \sum_{m=-M}^{m=M} C_m \Psi(x_0 + mh)$$



The Kohn-Sham problem: Better Eigensolvers



Diagonalization is computationally demanding using "standard" algorithms.

Computational load can be dramatically reduced using subspace filtering.

Filtered Subspace Iteration

Define charge density matrix:

$$P = \Phi^T \Phi \qquad \Phi = \left[\psi_1, \psi_2, \ldots, \psi_{occup} \right]$$

where the diagonal is the charge density.

For any orthonormal matrix U, we can write

$$P = \Phi^T (U^T U) \Phi = (\Phi U)^T (U \Phi)$$

We do not need explicit vectors, we only need to know: $U\Phi$

We can find this using subspace filtering.

Nature of the Filter

Given the diagonalized solution:

Let us consider a polynomial filter, p(H)

$$H = Q^T \Lambda Q \qquad \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N) \qquad Q = [\psi_1, \psi_2, \ldots, \psi_N]$$

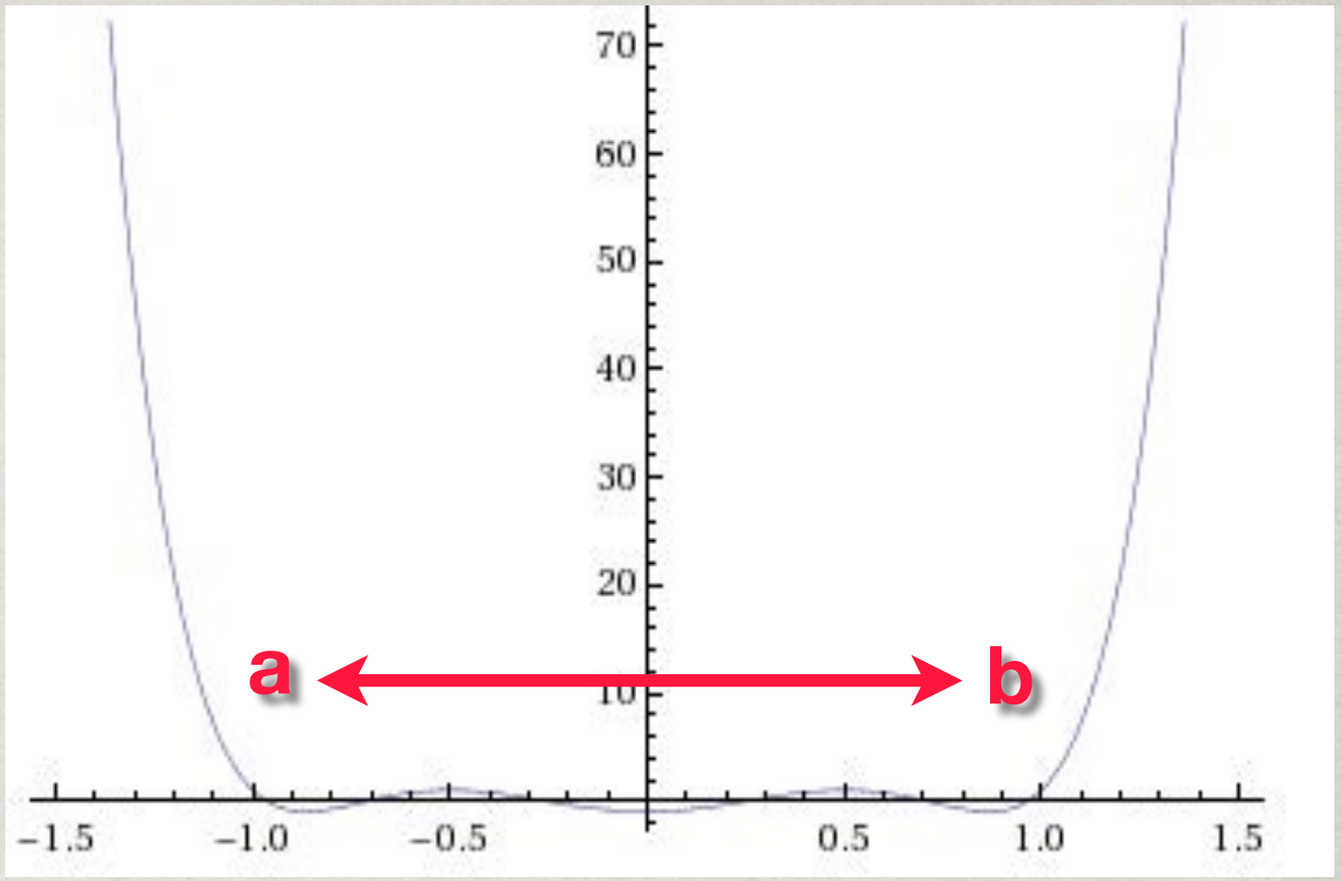
$$P(H) = Q^T P(\Lambda) Q = \sum_{i=1}^{N} P(\lambda_i) \psi_i^T \psi_i$$

$$P(H)v = \sum_{i=1}^{N} P(\lambda_i)(\psi_i^T v)\psi_i$$

Suppose we choose our filter such that it is small for states not of interest, e.g., p≈0, for empty states, then we can approach what we want

$$\hat{\psi}_j = P(H)v_j = \sum_{occup} P(\lambda_i)(\psi_i^T v_j)\psi_i \quad \Rightarrow \quad \hat{\psi} = U\psi$$

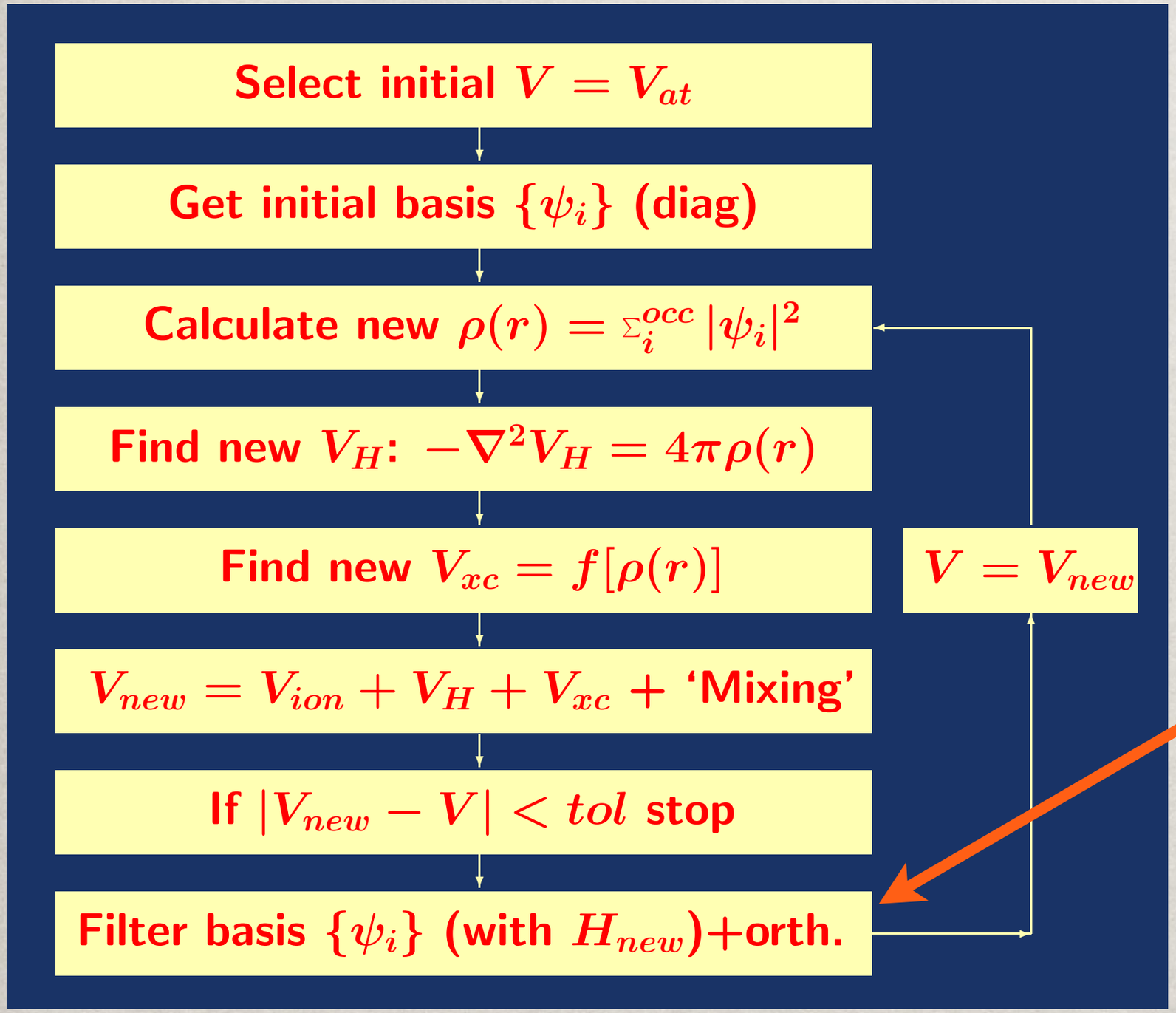
Filtering Operation



E_F E_u

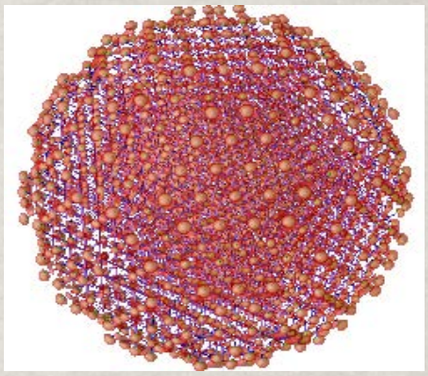
Eigenvalue spectrum

“New” Self-Consistent Loop



Filtering operation replaces diag operation.

“Typical Results”



method	# MV products	# SCF	total_eV /atom	CPU(secs)
CheFSI	124761	11	-77.316873	5946.69
ARPACK	142047	10	-77.316873	62026.37
TRLan	145909	10	-77.316873	26852.84

$Si_{525}H_{276}$, $N = 292584$, $n_{occ} = 1194$, $m = 8$.

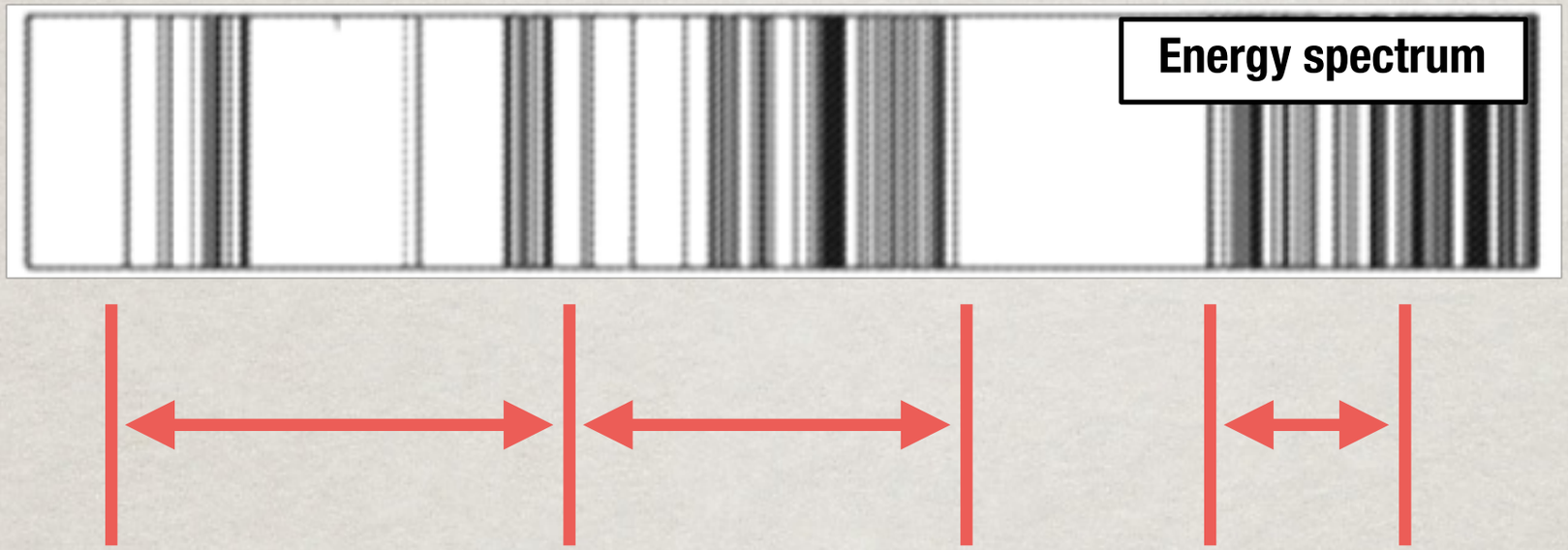
ARPACK: Implicit restart Arnoldi/Lanczos code, one of the best public domain eigenvalue packages, often used for benchmarking. (R. Lehoucq, D. Sorensen, C. Yang).

TRLan: Thick-restart Lanczos. (K. Wu, H. Simon). (improved symmetric eigensolver based on ARPACK).

http://parsec.ices.utexas.edu

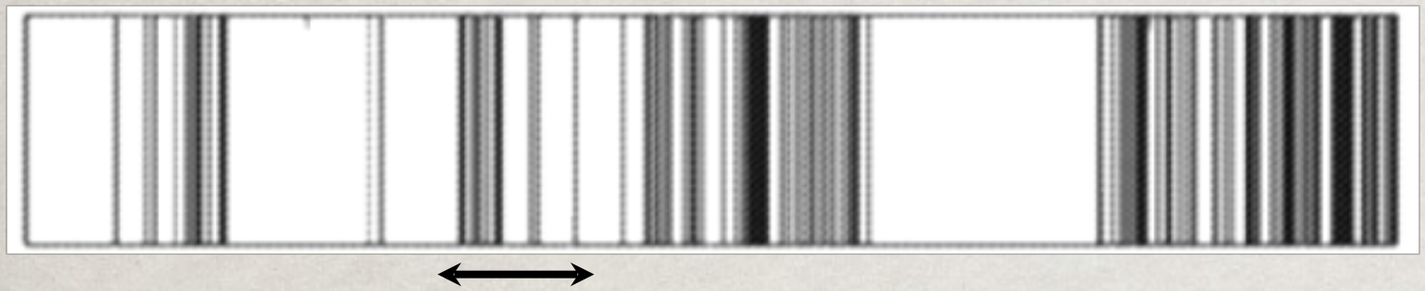
Phys. Rev. E 74, 066704 (2006)

Can we create a “parallel” eigensolver?

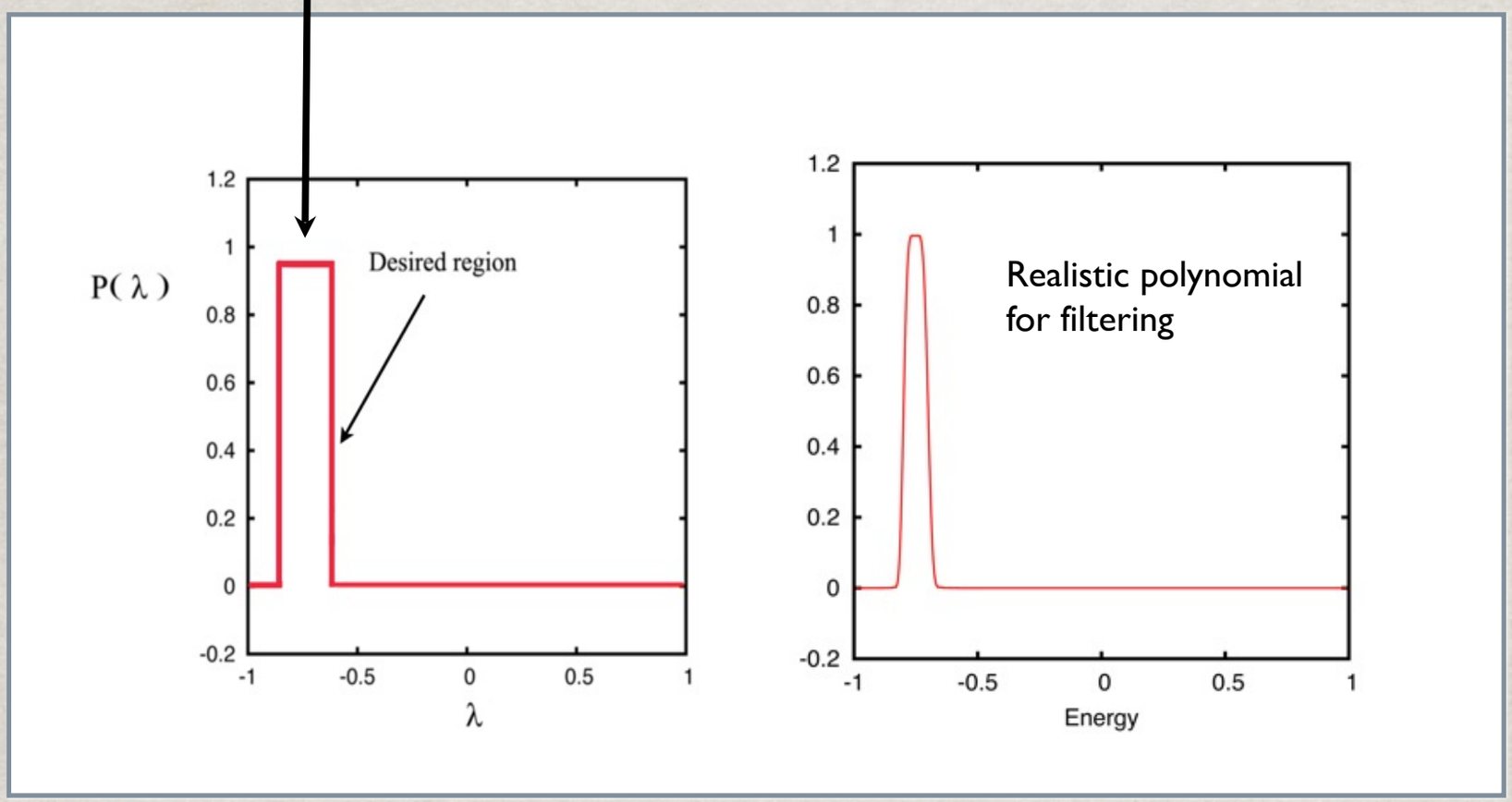


Solve for the energy window (or spectrum slices) independently. Mitigates Ritz step bottleneck— the diagonalization of a dense matrix (set by the number of desired eigenvalues).

Spectrum Slicing for the Kohn-Sham Problem



Create a filtering window for an energy window using polynomials (typically Chebyshev).



G. Schofield, J.R. Chelikowsky and Y. Saad, Comp. Phys. Commun. 183, 497 (2012)

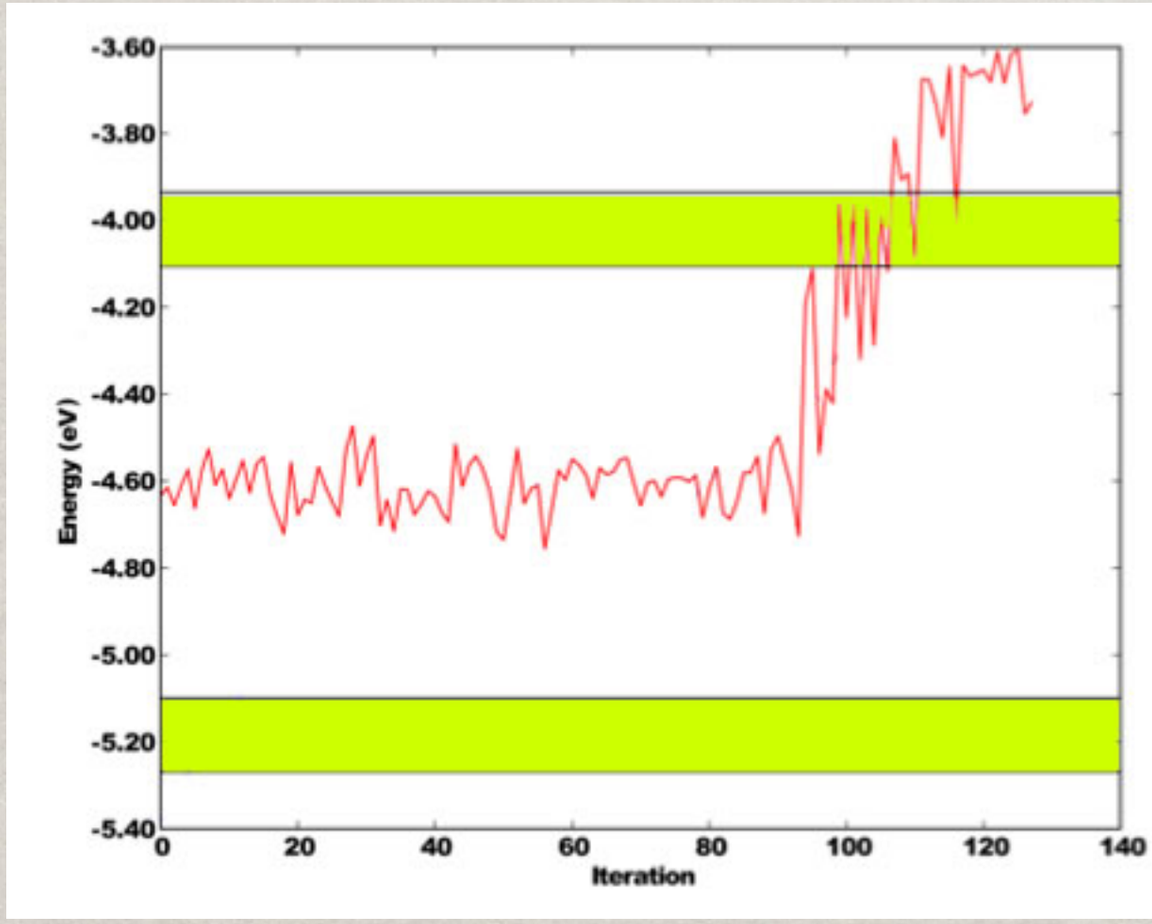
Complications:

Avoid missing eigenvalues

- Overlap energy windows, creating selvage region

Avoid double counting eigenvalues

- Singular value decomposition to eliminate duplicate eigenvalues

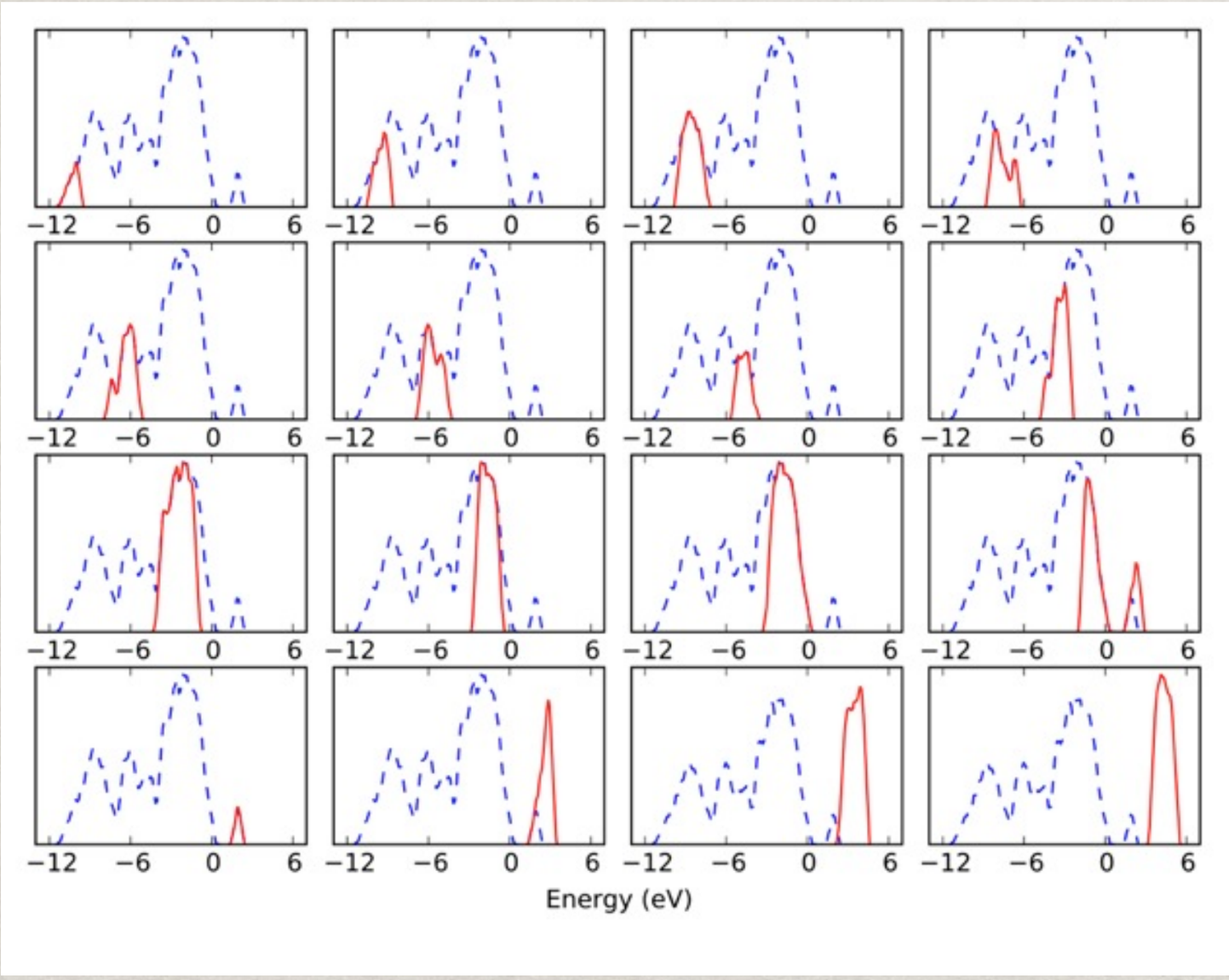


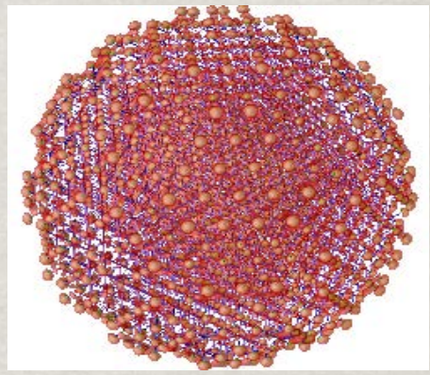
The red line shows the Ritz value of the last converged eigenvalue.

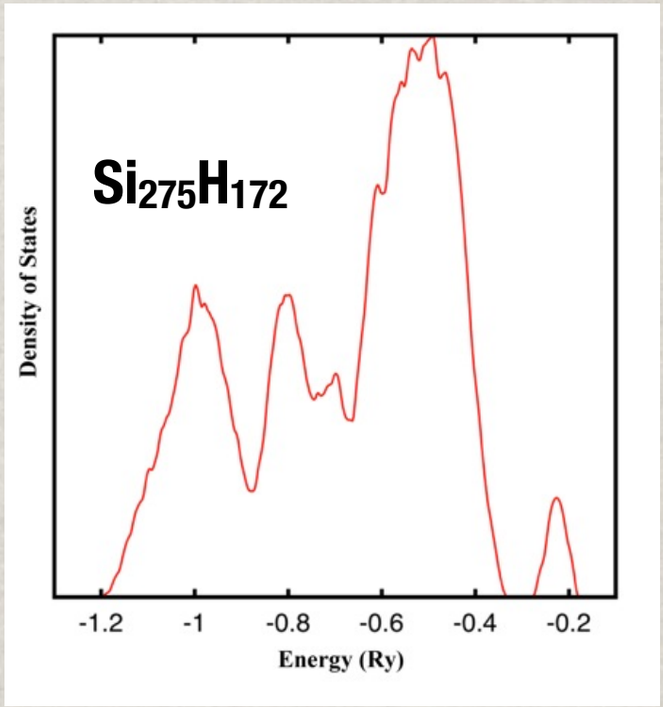
Once the line leaves the energy range of the slice, the method will have found all the eigenvalue of interest.

Proof in principle: Spectrum of Si Nanocrystal

Spectral slicing using filtering allows each energy region to be determined in an independent fashion. Avoids key ottlenecks in the Kohn-Sham solution for large systems.







Spectral slicing using filtering reproduces full eigenvalue spectrum.

Implementation issues

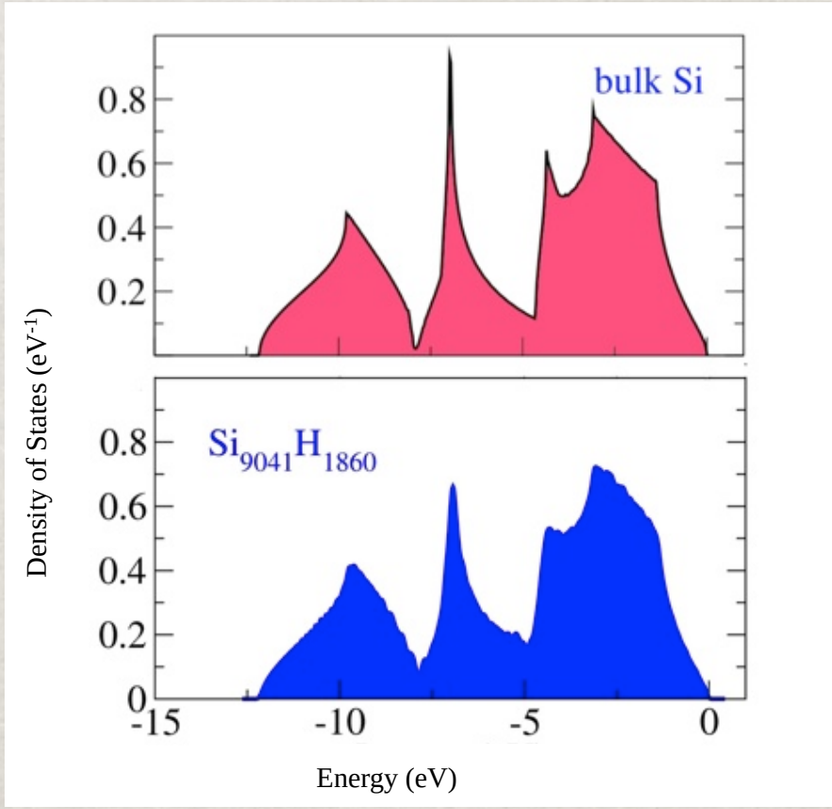
- **Balance windows to contain similar number of eigenvalues**
- **Minimize global communications**
- **Reduce memory load**
- **Expedite mat-vec operations**

Predicting the Spectral Distribution

The density of states (spectral distribution) gives the number of states in a width ΔE where ΔE -> 0

-Provides useful physical information without having to compute all the states

-Several algorithms can be used to determine the DOS and estimate the number of eigenvalues in an interval: kernel polynomial method, approximate Lanczos and continued fractions.



Density of States (Spectral Density) Estimation

$$\phi(\lambda) = trace\left(\delta(H - \lambda I)\right)$$

Kernel polynomial method (KPM): Expand the density of states into polynomials:

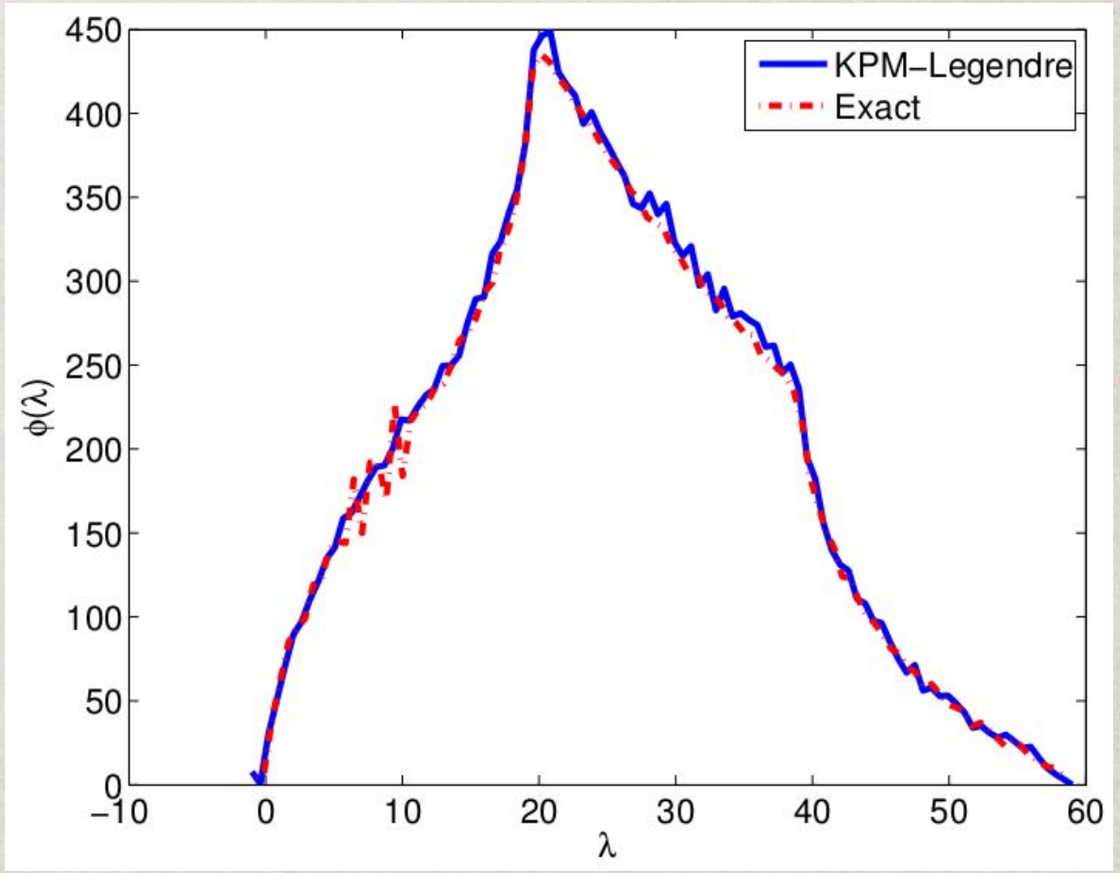
$$\phi(\lambda) \approx \sum_{k=0}^{m} \gamma_k T_k(\lambda)$$

Coefficients γ_k require evaluation of

$$trace(T_k(H))$$

By using stochastic sampling:

$$\text{trace}(H) \approx \sum_i \frac{\psi_i^T H \psi_i}{\psi_i^T \psi_i}$$



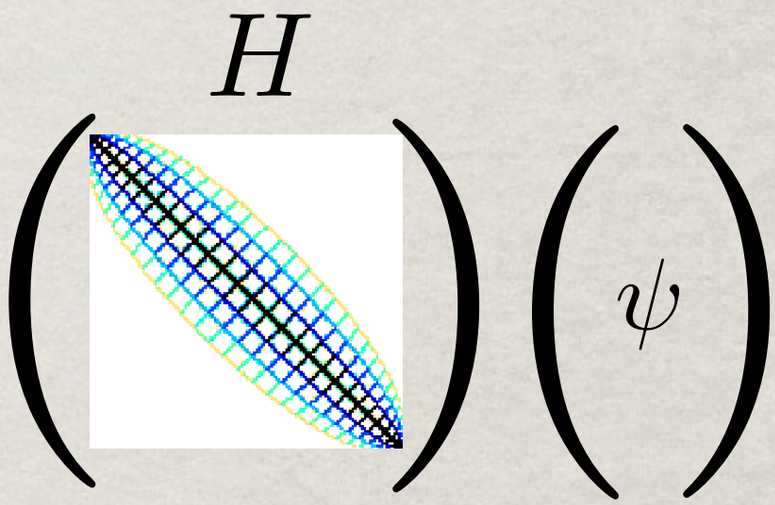
L. Lin, C. Yang and Y. Saad (to be published)

Communication and Cache Management

Sending data over the network must avoid unnecessary overhead.

- -Buffer copies should be eliminated.
- -Arithmetic work needs to happen concurrently with communication.
- -Low overhead remote direct memory access (RDMA) routines should be used.

We need to exploit structure in the Hamiltonian to improve cache utilization, vectorization and minimize indexing.



mat-vec operations dominate the computational load

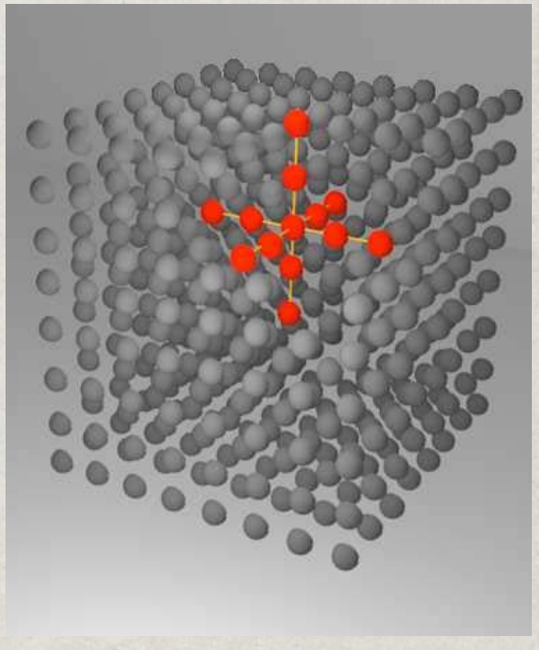
The Laplacian Term

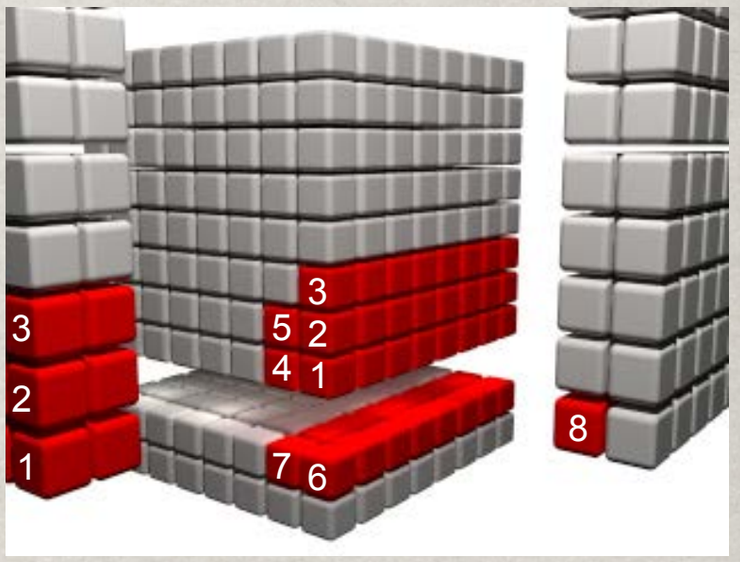
In the new algorithm, the array structures, already present in the problem, are handled as if they were primitive objects.

-Handling arrays yields better vectorization.

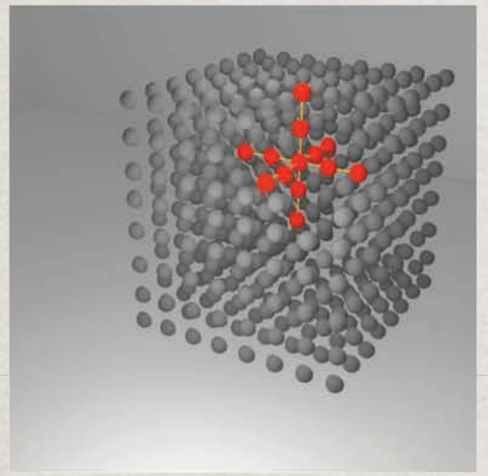
-Once an array is used, it can be reused in other outputs that need that data.

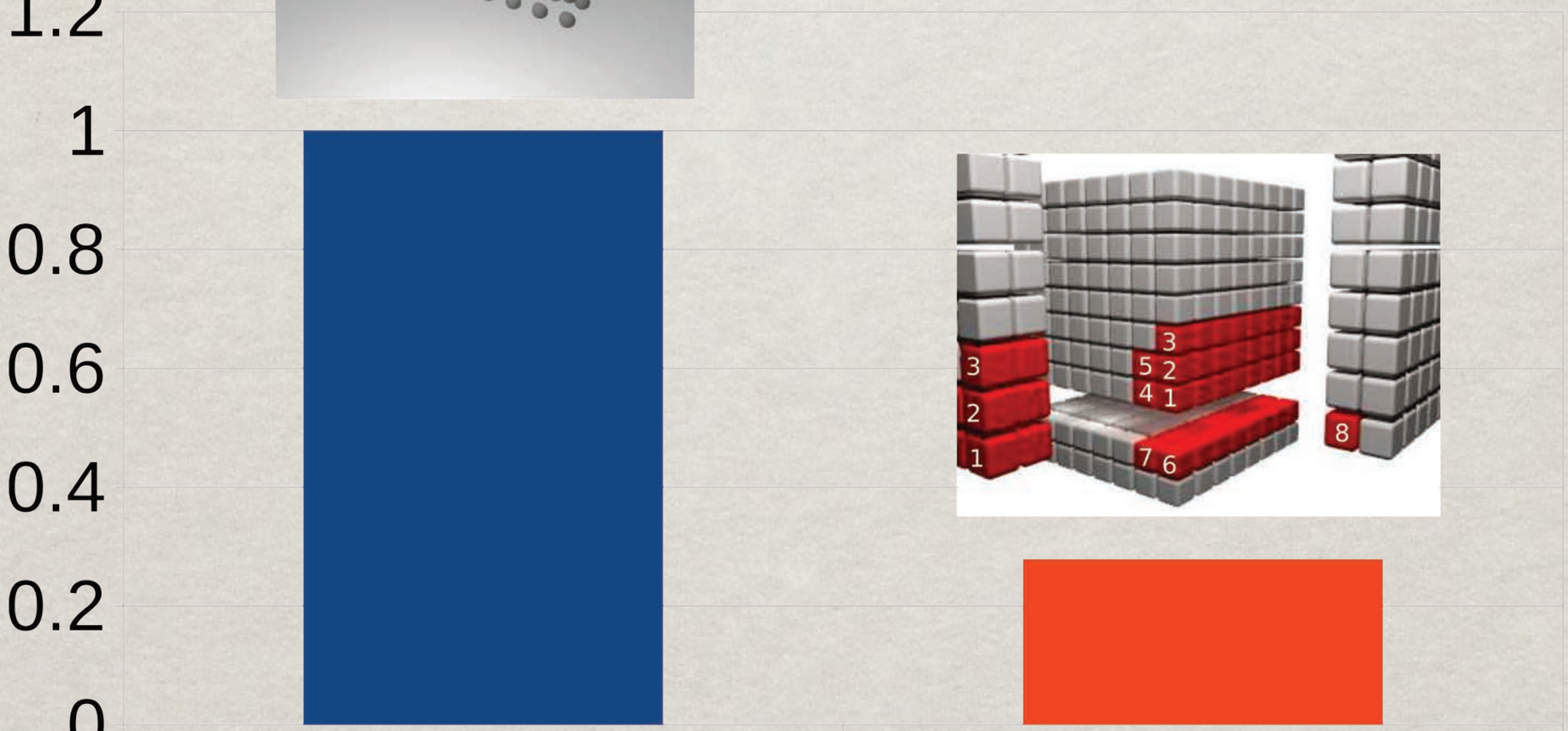
-This gives better cache reuse, leading to more efficient use of memory bandwidth and higher multicore performance.





Example: Speedup comes from vectorization and minimizing indexing



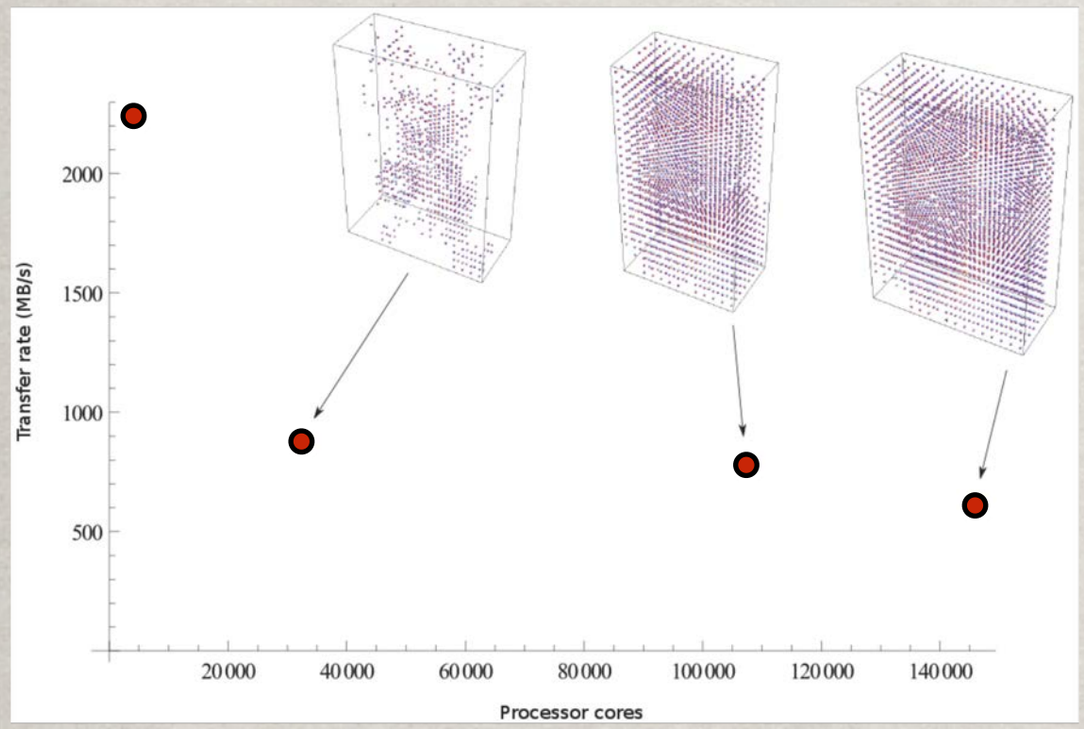


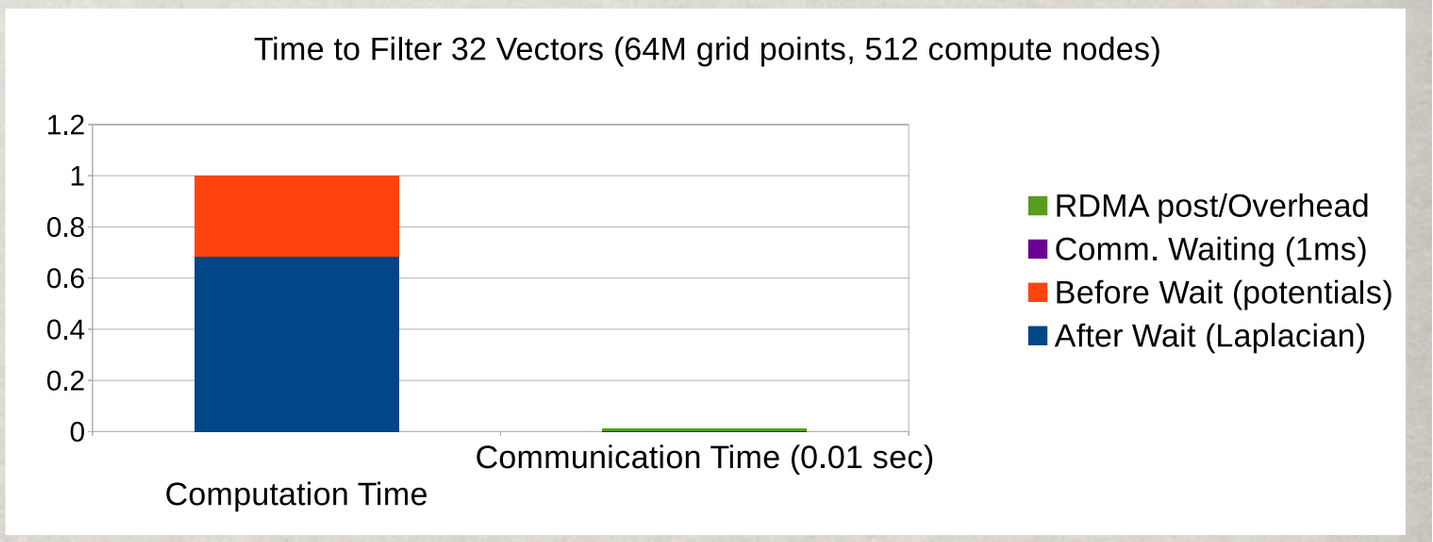
Network Performance

Those blocks of data needed by remote processors are separated and ordered contiguously in memory.

Extra copies into buffers are eliminated in the communication code this way.

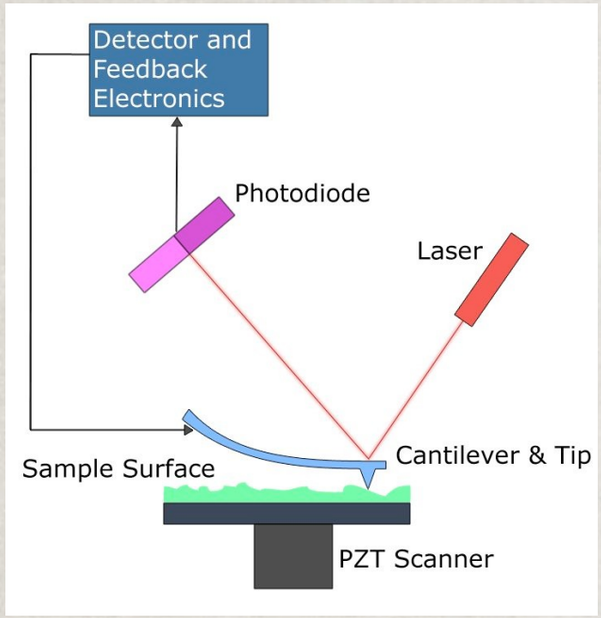
The new code uses low-level remote direct memory access, and overlapping communication.

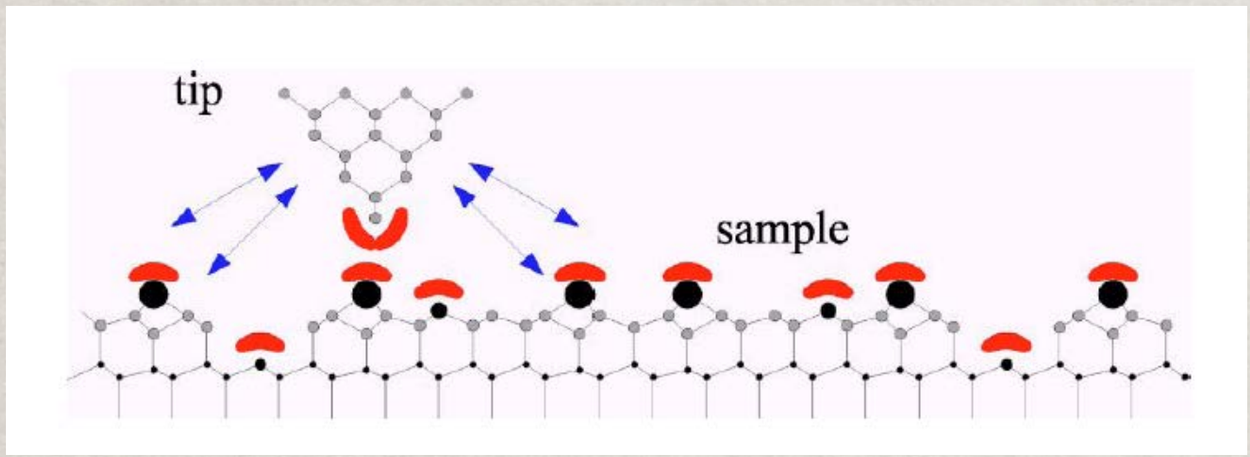


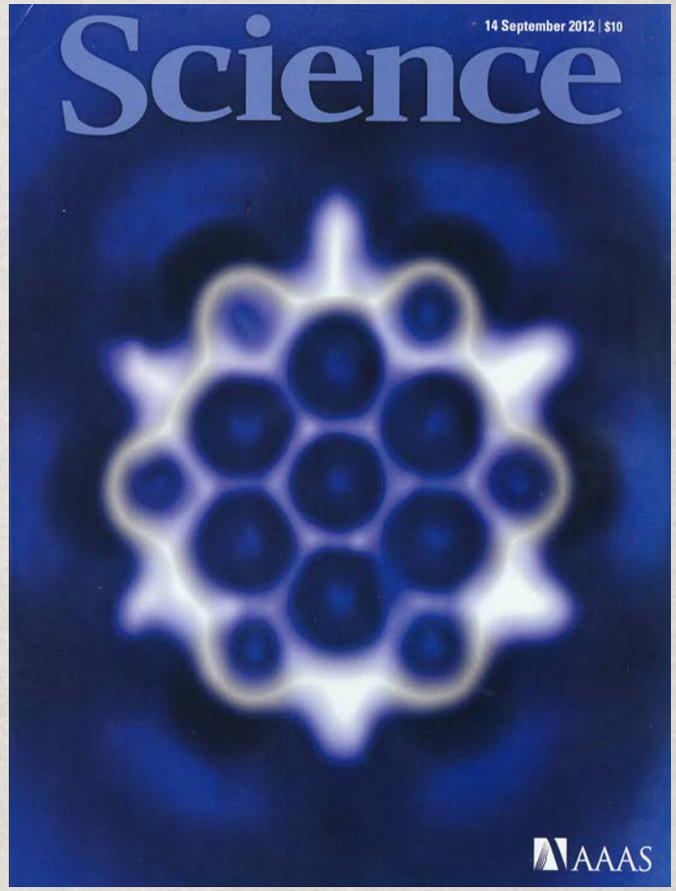


Example of an Application:
Simulating Atomic Force Microscopy Images

Atomic Force Microscopy Images

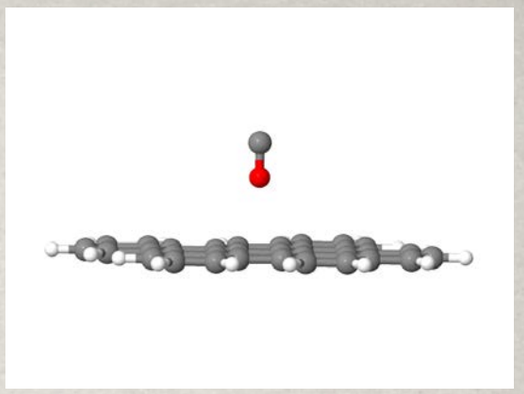


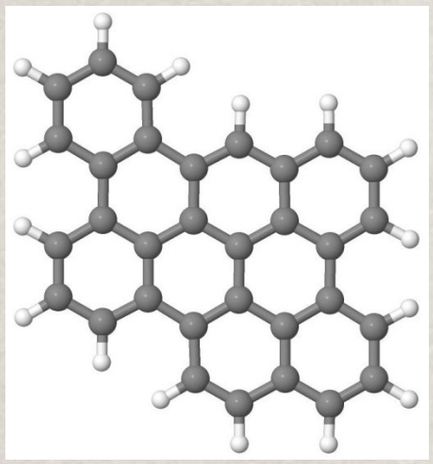




Simulation of AFM image requires calculating the force on the tip. The number of calculated geometries can easily approach ~10^4. Modeling the tip and sample can involve hundreds of atoms.

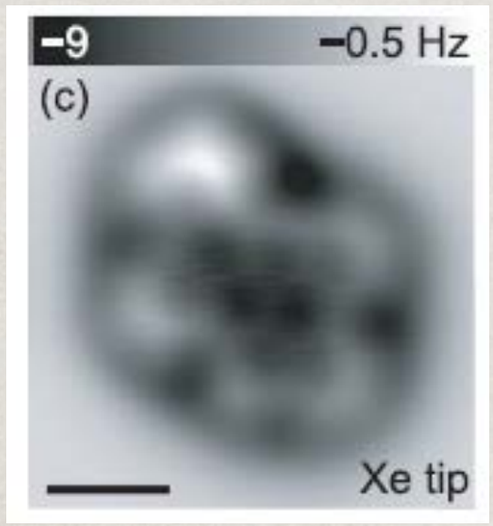
Simulate the images of a functionalized tip with a CO molecule.

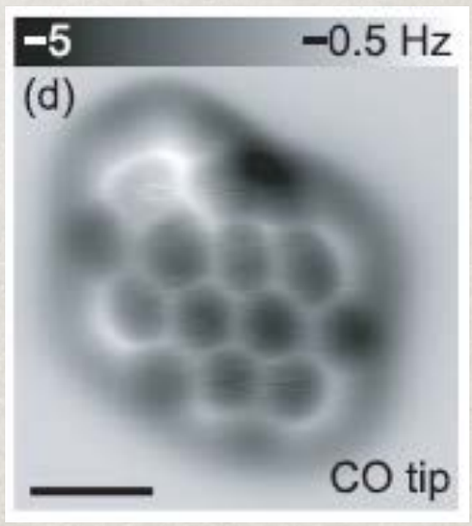




dibenzo(cd,n)naphtho-
(3,2,1,8-pqra)perylene

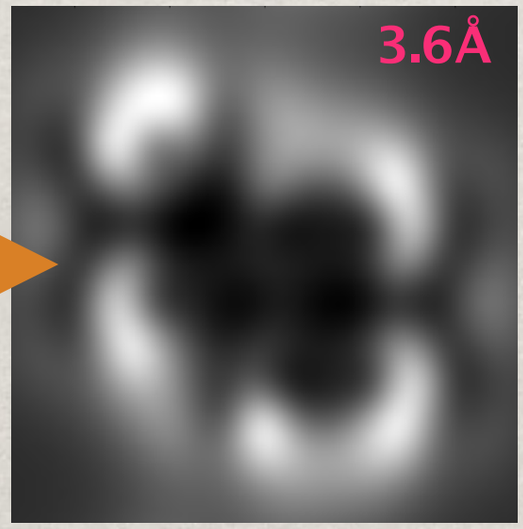
Experimental AFM images with Xe and CO tip

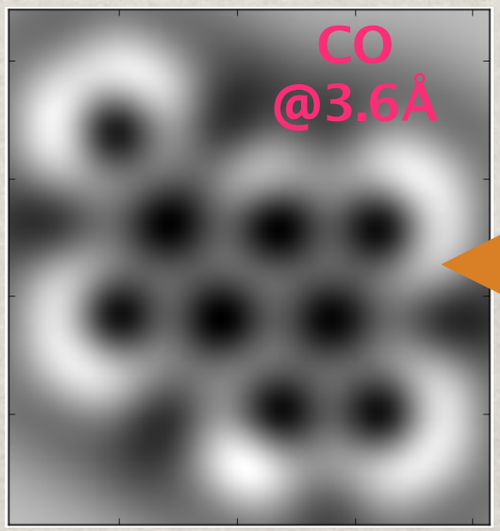




Mohn, et al. Appl. Phys. Lett. (2013)

Simulated AFM images with chemically inert tip





Simulated AFM images with CO tip

Posters for “Excited State” Applications and Methods

Excited State Computations: GW Method

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{ion}} + V_{\text{H}} + \Sigma(E_n^{\text{QP}})\right]\psi_n^{\text{QP}} = E_n^{\text{QP}}\psi_n^{\text{QP}}$$

The quasiparticle energy corresponds to the energy create to particle-like excitation in a system, e.g., adding a particle to or removing a particle from a system of N interacting particles.

$$\Sigma(\mathbf{r},\mathbf{r}',\omega) = i\int \frac{dE'}{2\pi} e^{-i\delta\omega'} G(\mathbf{r},\mathbf{r}';\omega-\omega')W(\mathbf{r},\mathbf{r}';\omega')$$

∑ is the non-local, energy-dependent, non-Hermitian, self-energy operator; its exact form is unknown, but progress can be made by approximating it within many-body perturbation theory, e.g., as the first term of an expansion in the screened Coulomb interaction.

Theoretical developments and applications in excited-state calculations

James R. Chelikowsky, Jack R. Deslippe, Felipe H. da Jornada, Steven G. Louie, Jeffrey B. Neaton, Johannes Lischner, Fang Liu, Diana Y. Qiu, Jaime S. de Sousa, Sahar Sharifzadeh, Derek Vigil-Fowler, Chao Yang

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY | Office of Science

Nature of Plasmon Satellites in Photoemission Spectroscopy

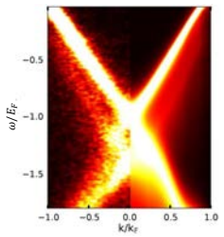
Motivation

- Plasmons and their interactions with electrons and holes are of increasing interest for technological applications, such as plasmonics and catalysis.
- **Goal**: Understand the nature of plasmon satellites that appear in the angle-resolved photoemission spectroscopy (ARPES) measurements on doped graphene and bulk Si.

Doped Graphene

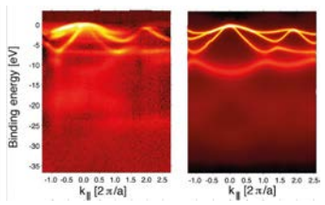
- Previous GW calculations based on simplified models of electron systems, such as the linear-bands model for graphene and the electron gas, predicted the existence of a composite particle, the *plasmaron,* consisting of a hole and a plasmon.
- ARPES experiments on doped graphene showing significant plasmon satellites appeared to support this theory.
- *Ab initio* treatment of substrate and the GW plus cumulant expansion (GW+C) of the interacting green's function give ARPES spectra in good agreement with experiment for all momenta measured.
- No plasmaron appears in spectra, only weak plasmon satellites.
- Agreement of previous theoretical work with experiment was due to cancellation of errors in use of GW approximation and overestimation the screening from the substrate.

Doped Graphene ARPES



Bulk Si

- First measurement and calculation of ARPES spectra of bulk Si, including QP and satellite peaks.
- GW+C again gives good agreement with experiment and gives no plasmaron.



Conclusions

- While sufficient for calculating QP properties, the GW approximation is insufficient for satellite properties, for which you need GW+C.
- The plasmaron is an artifact of the GW approximation and does not exist in nature.
- Common model dielectric functions overestimate substrate screening.

Future Works

- Understand electron-plasmon coupling on surfaces for catalytic applications.
- Include extrinsic contribution to GW+C ARPES spectra to improve agreement of QP and satellite peak weights with experiment.

Collaborators

Theory: Derek Vigil-Fowler, Johannes Lischner, and Steven G. Louie

Experiment: Gunnar K. Palsson and Charles S. Fadley

Accurate GW calculations for closed-shell molecules.

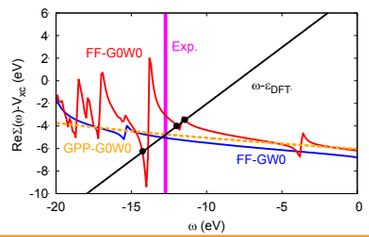
Motivation

- Knowledge of quasiparticle energies of molecules is important for interpreting spectroscopy experiments and the design of novel devices for energy technology, such as Graetzel cells or OLEDs.
- **Goal**: Understanding the effects of self-consistency, full-frequency approaches and generalized plasmon-pole models on quasiparticle energies in closed-shell molecules.

Pitfalls in full-frequency G0W0 calculations

- In molecules, the frequency-dependent electron self energy exhibits poles which describe shake-up excitations consisting of a quasiparticle and an electron-hole pair.
- For molecules with a small HOMO-LUMO gap, the self-energy poles can occur at unphysically low energies inducing large errors in the computed quasiparticle energies.
- Replacing the DFT Kohn-Sham energies by quasiparticle energies within a self-consistent GW0 approach pushes the self-energy poles to higher energies and results in good agreement with experiment.

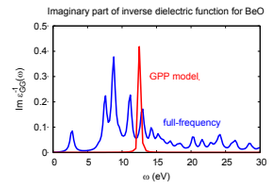
LUMO of ozone



G0W0 calculations with a generalized plasmon-pole model

- The generalized plasmon-pole model replaces the complicated dielectric response of a molecule by a single excitations which must have a high energy to fulfill the f-sum rule.
- As a result, the self-energy poles are shifted to higher energies and accurate quasiparticle energies are obtained.

Imaginary part of inverse dielectric function for BeO



Conclusions

- For molecules with small HOMO-LUMO gaps, full-frequency G0W0 calculations can result in large errors for quasiparticle energies because of unphysical self-energy poles.
- Self-consistent GW0 calculations give good agreement with experiment for quasiparticle energies.
- G0W0 calculations employing a generalized plasmon pole model provide accurate quasiparticle energies at a significantly lower numerical cost than self-consistent full-frequency GW0 calculations.

Future Works

- Understand the effects of full-frequency approaches, self-consistency and the generalized plasmon-pole model on quasiparticle energies in other systems, such as oxide materials.

Collaborators

Theory: Johannes Lischner, Sahar Sharifzadeh, Jack R. Deslippe, Jeffrey R. Neaton and Steven G. Louie.

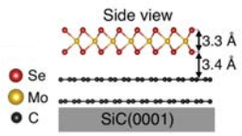
Effect of Substrate Screening on the Quasiparticle and Optical Properties of $MoSe_2$

Motivation

- Transition metal dichalcogenides, such as MoS2 and MoSe2, are interesting candidate materials for electronic and photovoltaic applications.
- **Goal**: Understand the role of substrate used in experiments on the quasiparticle and excitonic properties of MoSe2.

In-Plane Substrate Approximation

- Experimental collaboration: both scanning tunneling spectroscopy and photoluminescense measurements were performed on the same sample.

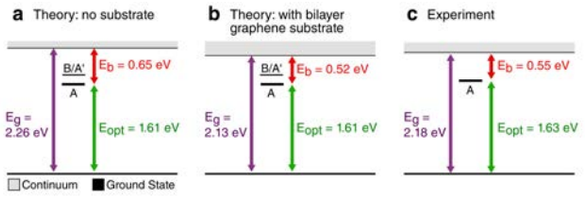


- Challenge: calculate quasiparticle band structure and exciton binding energy of monolayer $MoSe_2$ including the screening from the substrate. The MoSe2 monolayer is only commensurable with the substrate on a 3x3 supercell.
- We developed an *ab initio* method to calculate the screening from the substrate without having to construct a large supercell for the material + substrate.
- Idea: we fully consider the perpendicular component of the screening, but neglect in-plane local fields:

$$\tilde{\chi}^{0,\text{bilayer}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \chi^{0,\text{bilayer}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})\,\delta_{\mathbf{G}_x\mathbf{G}'_x}\delta_{\mathbf{G}_y\mathbf{G}'_y}$$

$$\Delta\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \left[1 - v\left(\chi^{0,\text{MoSe2}}(\mathbf{q}) + \tilde{\chi}^{0,\text{bilayer}}(\mathbf{q})\right)\right]^{-1}_{\mathbf{G}\mathbf{G}'} - \left[1 - v\,\chi^{0,\text{MoSe2}}(\mathbf{q})\right]^{-1}$$

Effect of the Substrate on Monolayer $MoSe_2$



- Good agreement with experiment when substrate screening is included.

Conclusions

- The In-plane Substrate Approximation is an efficient *ab inito* technique to include the substrate screening.
- For MoSe2 on bilayer graphene substrate, the substrate closes the quasiparticle gap by 130 meV, and decreases the exciton binding energy by a similar amount.

Future Works

- Calculate the effect of metallic screening on $MoSe_2$.
- Apply the method for molecular systems.

Collaborators

Theory: Felipe H. da Jornada, Diana Y. Qiu and Steve G. Louie

Experiment: Feng Wang and Mike F. Crommie groups

Efficiency and Accuracy of PW-PP GW Calculations

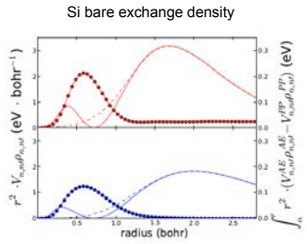
Motivation

- Efficient methods for calculating the QP properties of large systems are needed to adequately treat systems of current interest, including materials with defects and complex materials with large unit cells.
- **Goal**: To understand the effect of the plane-wave pseudopotential (PW-PP) approximation on the accuracy of GW calculations.

Atomic

- Contributions to bare exchange coming from states that have a small (large) separation between outer and inner wavefunction peaks are smaller (larger) for PW-PP GW calculations relative to AE GW calculation.

Si bare exchange density



- Atomic Si, Ga, As, and Ar show small differences in bare exchange between all-electron (AE) and PW-PP GW calculations (~1%).
- Trends are robust across all atoms studied, which vary in localization of electrons and presence of d-states, indicating the trends are robust across different system types.

Bulk

- Wavefunctions in the bulk have larger inner-outer peak separation than in the atomic case due to bonding. The bare exchange is larger for PW-PP GW calculations, as expected from the atomic case.
- The valence band has a larger difference between PW-PP and AE GW calculations than the conduction band, leading to an opening of the gap for PW-PP GW calculations, on the order of 0.1-0.2 eV.
- A cheap correction scheme is to perform a PW-PP calculation for the dynamical self energy and get bare exchange from AE calculation.

Conclusions

- The error due to the PW-PP approximation is small, on the order 0.1-0.2 eV for many standard semiconductors.
- For systems with deep cores, this error can be reduced to ~0.05 eV by calculating only the bare exchange in the AE formalism, with only a minimal loss in computational efficiency. This could be very useful for accurate calculations on large systems.

Future Works

- Investigate efficacy of PAW formalism for shallow-core systems, where computational efficiency is drastically reduced with PW-PP GW.
- Design improved pseudopotentials including accurate exchange integrals as constraint during generation.

Collaborators

Theory: Derek Vigil-Fowler, Brad D. Malone and Steven G. Louie

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, under Contract No. DE-AC02-05CH11231.

MSD Materials Sciences Division

New algorithms and improvements to the BerkeleyGW excited state code

James R. Chelikowsky, Jack R. Deslippe, Felipe H. da Jornada, Steven G. Louie, Jeffrey B. Neaton, Johannes Lischner, Fang Liu, Diana Y. Qiu, Jaime S. de Sousa, Sahar Sharifzadeh, Derek Vigil-Fowler, Chao Yang

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY | Office of Science

Dimension reduction techniques for the GW self energy approximation

Motivation

The GW approximation to the self-energy required in the Green's function approach to model single electron excitation requires the frequency dependent screening dielectric matrix $\epsilon(\omega) = I - V\chi_0(\omega)$ to be computed efficiently. The standard approach requires computing the irreducible polarizability operator $\chi_0(\omega)$, which in turn requires all eigenpairs of a single particle Hamiltonian to be computed. We would like to eliminate the need to compute the unoccupied eigenpairs by constructing $V\chi_0(\omega)$ directly without forming $\chi_0(\omega)$ explicitly. However, this approach requires solving many linear equations. We show that dimension reduction techniques can be used to reduce the number of equations to be solved.

It follows from eigenvector perturbation analysis that $V\chi_0(\omega)$ can be obtained by solving the Sternheimer equation

$$(H - \varepsilon_v - (\pm\omega \pm i\eta))\Delta\Psi^{\pm}_{v,\omega} = -P^{\perp}_{occ}\,(\mathrm{Diag}(\psi_v)V)$$

$$(V\chi_0)(\omega) = 2\left[\sum_v \mathrm{Diag}(\psi^*_v)\left(\Delta\Psi^{+}_{v,\omega} + \Delta\Psi^{-}_{v,\omega}\right)\right]$$

- $P^{\perp}_{occ} = I - \sum_v \psi_v\psi^*_v$
- For each frequency, solve $n \times n \times n_v$ equations to get a $n \times n$ dielectric matrix.

Low-rank structure of $v\chi_0$

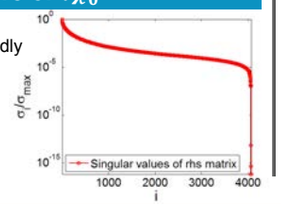
Wison, Lu, Gygi, Galli (2009)

- Eigenvalues of $V\chi_0$ decrease to 0 rapidly
- Low-rank approximation: $V\chi_0 \approx U\Lambda U^*$

$V\chi_0$ low rank

⇒ $\Delta\Psi$ low rank?

⇒ $-P^{\perp}_{occ}\mathrm{Diag}(\psi_v)V$ low rank?



Constructing low-rank approximation to ϵ and Σ

- $-P^{\perp}_{occ}\,(\mathrm{Diag}(\psi_v)v) \approx U_v\Sigma_v V^*_v$ (Frequency independent!)
- Solve $(H - \varepsilon_v + \omega)\Delta\Psi_v = U_v$ ($n \times n_s$ matrix, $n_s \ll n$)
- $\epsilon(\omega) \approx I - \sum_v Diag(\psi_v)\,\Delta\Psi_v\Sigma_v V^*_v$
- Further rank reduction $\epsilon(\omega) = I - X(\omega)CY^*$
- $\epsilon^{-1}(\omega) = I + X(\omega)C^{-1}Y^*$ (Inverting an $n_s \times n_s$ matrix C!)
- $\langle\psi_n|\Sigma_C(\omega)|\psi_n\rangle = tr(Z^*_Y H^{-1}(\omega)Z_X(\omega)C^{-1})$
  $Z_X(\omega) = \mathrm{Diag}(\psi_n)X(\omega), \quad Z_Y = \mathrm{Diag}(\psi_n)Y$

Numerical results for a methane molecule

Accuracy in $\epsilon^{-1}(0)$

Truncation Level	Rel Err($\epsilon^{-1}(0)$) Our approach	Rel Err($\epsilon^{-1}(0)$) Galli's approach
2205	0.16E-15	0.80E-14
2000	0.44E-07	0.48E-07
1000	0.18E-05	0.10E-05
200	0.11E-03	0.22E-03
100	0.44E-03	0.11E-02
50	0.15E-02	0.38E-02
20	0.52E-02	0.14E-01

Accuracy in $\Sigma(\omega)$

Observation

For small molecules, a 90% truncation of the singular values/vectors of $v\chi_0$ yields sufficiently accurate approximation of the self energy and quasi-particle energy. Consequently both computational time and storage requirement can be reduced significantly by the low-rank approximation.

Improved Interpolation Schemes

Motivation

- Excitons are correlated electron-hole pairs which can be predicted from the solutions of the Bethe-Salpeter equation (BSE).
- Because of their correlated nature, it's necessary to use interpolation schemes to solve the BSE.

Interpolation of the Bethe-Salpeter Equation

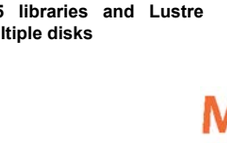
- In BerkeleyGW, we calculate the Kernel of the BSE on a coarse grid and interpolate them using the projection between wave functions the **coarse** and **fine** grids.

$$\langle vc\mathbf{k}_{\mathrm{fi}}|K|v'c'\mathbf{k}'_{\mathrm{fi}}\rangle = \sum_{i_1,n_2,n_3,n_4} C^{\mathbf{k}_{\mathrm{co}}}_{c,n_1} C^{*\mathbf{k}_{\mathrm{co}}}_{v,n_2} C^{*\mathbf{k}'_{\mathrm{co}}}_{c',n_3} C^{\mathbf{k}'_{\mathrm{co}}}_{v',n_4} \langle n_2 n_1 \mathbf{k}_{\mathrm{co}}|K|n_4 n_3 \mathbf{k}'_{\mathrm{co}}\rangle$$

- We implemented space decomposition algorithms and caching techniques to speedup the interpolation: over **10^4 speedup** for graphene.

Before: $\mathcal{O}(N^2_{\mathrm{co}}N^3_{\mathrm{fi}})$

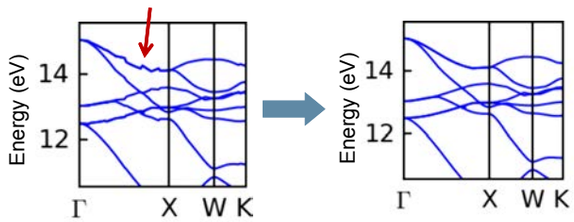
After: $\mathcal{O}(N_{\mathrm{co}}N_{\mathrm{fi}})$



Delaunay Tessellation

- In order to obtain smooth interpolated band structures, we implemented an algorithm that first tessellates the k-points based on Delaunay triangulation. This removes interpolation discontinuities.

Delaunay tessellation of an inhomogeneous k-mesh.



Old interpolation algorithm | New algorithm

Ongoing and Future Works

- Improve the interpolation of the dielectric matrix, which is ill-behaved for systems with reduced dimensionality.
- Support interpolation of kernels generated without the Tamm-Dancoff approximation (partially implemented)

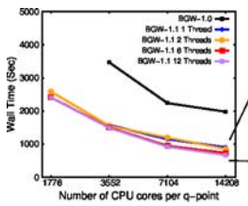
Collaborators **Developers:** Felipe H. da Jornada, Jack R. Deslippe and Steven G. Louie

Node Level Parallelism in BerkeleyGW 1.1

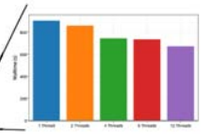
Motivation

- Improve the scalability of GW implementations to massively parallel DOE machines – utilizing both inter- and intra-node parallelism

On Node Parallelism

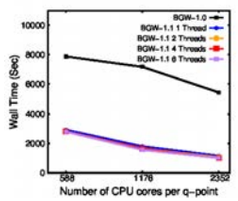


Epsilon Performance



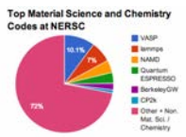
- **2X – 10X performance improvement throughout Package**
- **Hybrid OpenMP/MPI model added to support current and next-generation DOE machines.**
- **New efficient algorithms to reduce complexity. DFT orbital requirements reduced 5x.**

Sigma Performance

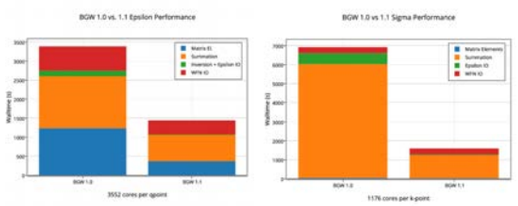


Impact

- **BerkeleyGW was 2.2% of the entire NERSC workload in the first 8 months of 2013. Compared to 1.2% in 2012.**
- **Users Study: photovoltaics, Interpretation of DOE Light-Source photoemission spectra, LED, Electronic, transport, and optical properties of novel**



BerkeleyGW 1.1 vs 1.0



- Substantial speedups throughout package from IO (next panel) OpenMP and vectorization.

Collaborators **Developers:** Jack R. Deslippe, Felipe H. da Jornada, Derek Vigil-Fowler, Fang Liu, Chao Yang and Steven G Louie

Improvements in BerkeleyGW

Motivation

- For large systems run at scale on DOE supercomputers, IO has become a major bottleneck in BerkeleyGW Performance.
- We overcome this bottleneck with parallel-IO using DOE supported HDF5 libraries.

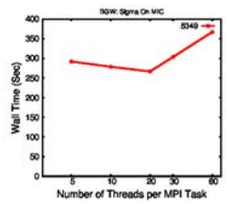
Parallel IO

BGW 1.0 IO Signature

BGW 1.1 IO Signature

- **IO improved from 90 MB/sec to > 2 GB/sec**
- **Average write size increased by orders of magnitude**
- **Utilize DOE supported parallel HDF5 libraries and Lustre filesystems features: striping across multiple disks**

Support for Many-Core Architectures



Rev 4770: Initial Code

Rev 4896: Refactor code to have loops targeting MPI, OpenMP, SIMD

Rev 5338: OpenMP Pragmas added

Rev 5349: Vectorization Ensured

Ongoing and Future Works

- Parallel IO for more files formats (wavefunctions)
- Support of GPUs
- Collaboration with Intel engineers to ensure vectorization, optimization

Collaborators **Developers:** Jack R. Deslippe, Felipe H. da Jornada, Derek Vigil-Fowler, Fang Liu, David A. Strubbe, Chao Yang and Steven G. Louie

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, under Contract No. DE-AC02-05CH11231.

MSD Materials Sciences Division

The End