

MADN

狂



	<p>Multiresolution Adaptive Numerical Scientific Simulation</p>		S
---	---	--	---

# *Multiresolution Solution of Non-linear Integral Eigen-problems in Electronic Structure*

*George I. Fann<sup>1</sup>, Zhengting Gan<sup>1</sup>, Robert J. Harrison<sup>1,2</sup>*

*<sup>1</sup>Oak Ridge National Laboratory*

*<sup>2</sup>University of Tennessee, Knoxville*

*In collaboration with*

*Gregory Beylkin<sup>4</sup>*

*<sup>4</sup>University of Colorado*

*[harrisonrj@ornl.gov](mailto:harrisonrj@ornl.gov)*

# *The funding*

- This work is funded by the U.S. Department of Energy, the division of Basic Energy Science, Office of Science, under contract DE-AC05-00OR22725 with Oak Ridge National Laboratory. This research was performed in part using
  - the Molecular Science Computing Facility in the Environmental Molecular Sciences Laboratory at the Pacific Northwest National Laboratory under contract DE-AC06-76RLO 1830 with Battelle Memorial Institute,
  - resources of the National Energy Scientific Computing Center which is supported by the Office of Energy Research of the U.S. Department of Energy under contract DE-AC03-76SF0098,
  - and the Center for Computational Sciences at Oak Ridge National Laboratory under contract DE-AC05-00OR22725 .

# References

- The methods in this work are primarily based upon
  - Alpert, Beylkin, Grimes, Vozovoi (J. Comp. Phys., 2002)
  - B. Alpert (SIAM Journal on Mathematical Analysis 24, 246-262, 1993).
  - Beylkin, Coifman, Rokhlin (Communications on Pure and Applied Mathematics, 44, 141-183, 1991.)
  - Beylkin and Mohlenkamp, (Proc. Nat. Acad. 2002)
- The following are useful further reading
  - Daubechies, “*Ten lectures on wavelets*”
  - Walnut, “*An introduction to wavelets*”
  - Meyer, “*Wavelets, algorithms and applications*”
  - Burrus et al, “*Wavelets and Wavelet transforms*”

# *Multiresolution chemistry objectives*

- Complete elimination of the basis error
  - One-electron models (e.g., HF, DFT)
  - Pair models (e.g., MP2, CCSD, ...)
- Correct scaling of cost with system size
- General approach
  - Readily accessible by students and researchers
  - Higher level of composition
  - No two-electron integrals – replaced by fast application of integral operators
- New computational approaches
- *Fast algorithms with guaranteed precision* <sup>5</sup>

# *Molecular electronic Schrödinger equation*

- A 3-N dimensional, non-separable, second-order differential equation

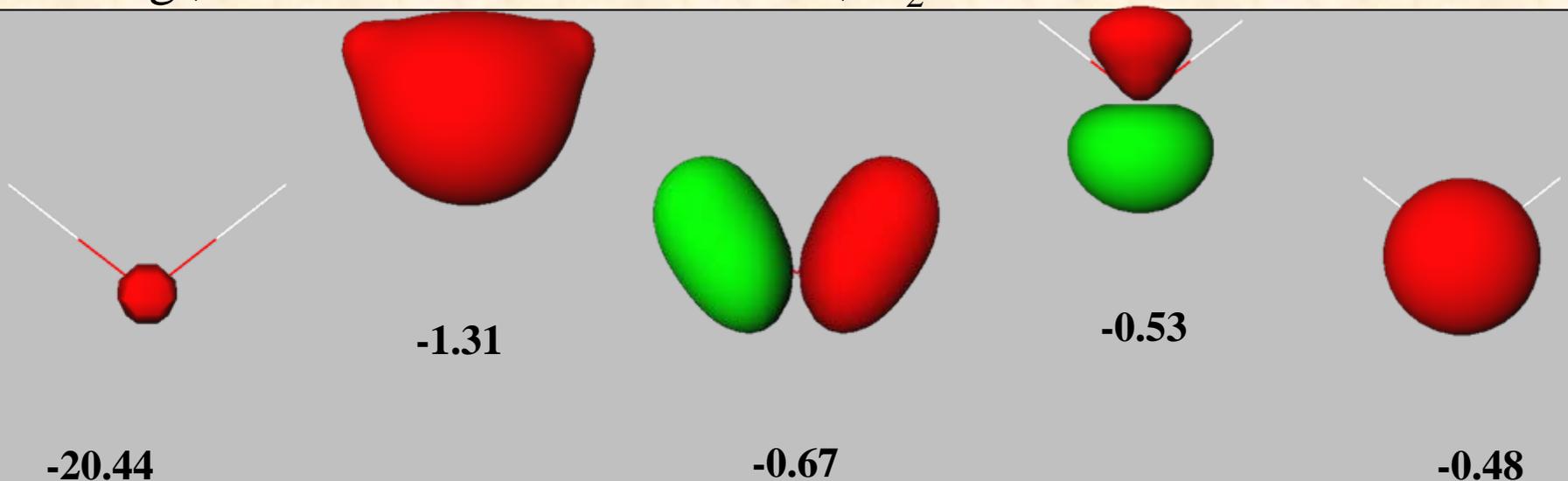
$$H\Psi(r_1, r_2, \dots, r_n) = E\Psi(r_1, r_2, \dots, r_n)$$

$$H = -\frac{1}{2} \sum_{i=1,m} \nabla_i^2 - \sum_{\substack{i=1,n \\ \mu=1,N}} \frac{Z_\mu}{|r_i - r_\mu|} + \sum_{\substack{i=1,n \\ j=1,i-1}} \frac{1}{|r_i - r_j|}$$

- Practical approximations reduce to solving non-linear problems for single-particle (3D) and two-particle (6D) functions

# “Independent” particle models

- Atomic and molecular orbitals
  - Each electron feels the *mean field* of all other electrons (self-consistent field, Hartree-Fock)
  - Replaces linear 3N-D Schrödinger w. non-linear 3-D eigen-problem
  - Provides the structure of the periodic table and the chemical bond
  - Linear combination of atomic orbitals - LCAO
  - E.g., molecular orbitals for water, H<sub>2</sub>O



# *Density functional theory (DFT)*

- Hohenberg-Kohn theorem
  - The energy is a functional of the density (3D)
- Kohn-Sham
  - Practical approach to DFT, parameterizing the density with orbitals (easier treatment of kinetic energy)
  - Very similar computationally to Hartree-Fock, but potentially exact

$$\left( -\frac{1}{2} \nabla^2 + V_{coul}(r; \rho) + V_{xc}(r; \rho) + V_{ext}(r) \right) \phi_i(r) = \varepsilon_i \phi_i(r)$$

$$\rho(r) = \sum_i \phi_i^2(r)$$

# *Integral Formulation*

- Solving the integral equation
  - Eliminates the derivative operator and related “issues”
  - Converges as fixed point iteration *with no preconditioner*

$$\left(-\frac{1}{2}\nabla^2 + V\right)\Psi = E\Psi$$

$$\begin{aligned}\Psi &= -2\left(-\nabla^2 - 2E\right)^{-1} V\Psi \\ &= -2G^*(V\Psi)\end{aligned}$$

$$(G^* f)(r) = \int ds \frac{e^{-k|r-s|}}{4\pi|r-s|} f(s) \quad \text{in 3D ; } k^2 = -2E$$

Such Green's Functions (bound state Helmholtz, Poisson) can be rapidly and accurately applied with a single, sparse matrix vector product.<sup>9</sup>

# *Essential techniques for fast computation*

- Multiresolution  $V_0 \subset V_1 \subset \dots \subset V_n$   
 $V_n = V_0 + (V_1 - V_0) + \dots + (V_n - V_{n-1})$
- Low-separation rank  
 $f(x_1, \dots, x_d) = \sum_{l=1}^M \sigma_l \prod_{i=1}^d f_i^{(l)}(x_i) + O(\varepsilon)$   
 $\|f_i^{(l)}\|_2 = 1 \quad \sigma_l > 0$
- Low-operator rank  
 $A = \sum_{\mu=1}^r \underline{u}_\mu \sigma_\mu \underline{v}_\mu^T + O(\varepsilon)$   
 $\sigma_\mu > 0 \quad \underline{v}_\mu^T \cdot \underline{v}_\lambda = \underline{u}_\mu^T \cdot \underline{u}_\lambda = \delta_{\mu\lambda}$

# *How to “think” multiresolution*

- Consider a ladder of function spaces

$$V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_n$$

- E.g., increasing quality atomic basis sets, or finer resolution grids, ...

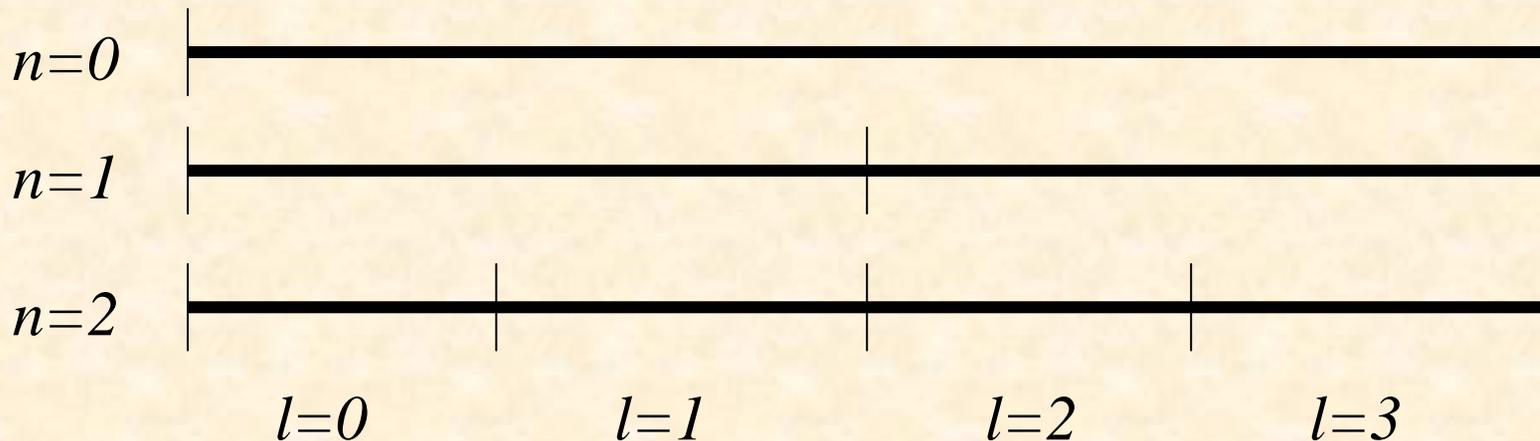
- Telescoping series

$$V_n = V_0 + (V_1 - V_0) + (V_2 - V_1) + \cdots + (V_n - V_{n-1})$$

- Instead of using the most accurate representation, use the difference between successive approximations
- Representation on  $V_0$  small/dense; differences sparse
- Computationally efficient; possible insights

# Scaling Function Basis

- Divide domain into  $2^n$  pieces (level  $n$ )
  - Adaptive sub-division (local refinement)
  - $l^{\text{th}}$  sub-interval  $[l*2^{-n}, (l+1)*2^{-n}]$   $l=0, \dots, n-1$
- In each sub-interval define a polynomial basis
  - First  $k$  Legendre polynomials  $\phi_i(x) = \sqrt{2i+1}P_i(2x-1)$
  - Orthonormal, disjoint support  $\phi_{il}^n(x) = 2^{n/2} \phi_i(2^n x - l)$



# Multiwavelet Basis

- An orthonormal basis to span  $W_{n-1} = V_n - V_{n-1}$

Haar (piecewise constant basis)  $\phi_l^{n-1} = 2^{-1/2} \left( \phi_{2l}^n + \phi_{2l+1}^n \right)$  sum

$\psi_l^{n-1} = 2^{-1/2} \left( \phi_{2l}^n - \phi_{2l+1}^n \right)$  difference

$$\phi_{il}^{n-1}(x) = \sum_{j=0}^{k-1} \left( h_{ij}^{(0)} \phi_{j2l}^n(x) + h_{ij}^{(1)} \phi_{j2l+1}^n(x) \right)$$

$$\psi_{il}^{n-1}(x) = \sum_{j=0}^{k-1} \left( g_{ij}^{(0)} \phi_{j2l}^n(x) + g_{ij}^{(1)} \phi_{j2l+1}^n(x) \right)$$

# *Vanishing moments*

- Critically important property
  - Since  $W_n = V_{n+1} - V_n$  is orthogonal to  $V_n$  the first  $k$  moments of functions in  $W_n$  vanish, i.e.,
$$\int x^j \psi_i(x) dx = 0, \quad j = 0, \dots, k - 1$$
- Compact representation of smooth functions
  - Consider Taylor series ... the first  $k$  terms vanish and smooth implies higher order terms are small
- Compact representation of integral operators
  - E.g.,  $|r-s|^{-1}$  ... interaction decays as  $r^{-2k-1}$
- Derivatives vanish at origin in Fourier space
  - Diminishes effect of singularities at that point

# Truncation Error

- To satisfy the global error condition

$$\|f - f^n\|_2 \leq \varepsilon \|f\|_2$$

- Truncate according to

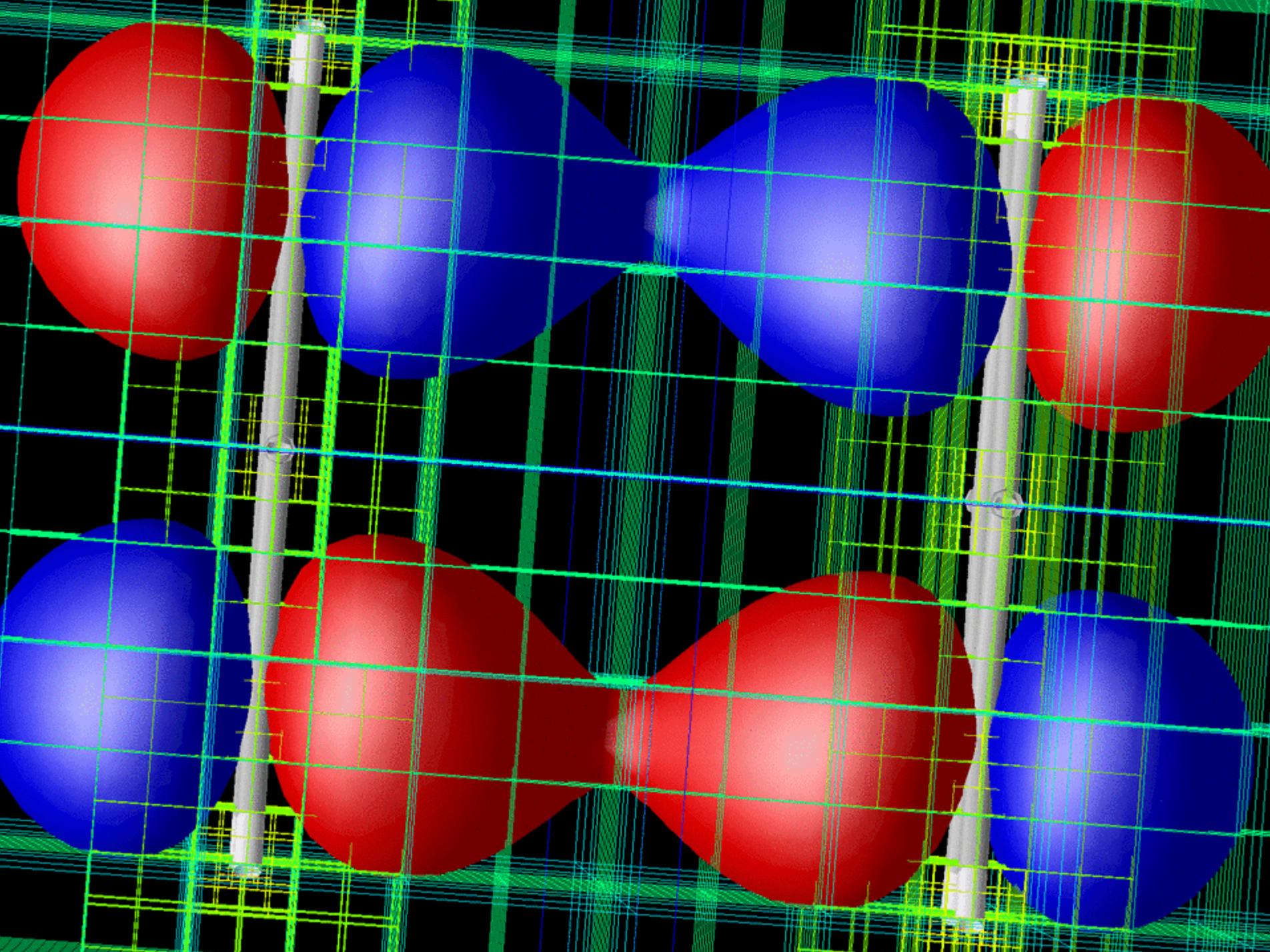
$$\|d_l^n\|_2 \leq 2^{-n/2} \varepsilon \|f\|_2$$

- Above is rather conservative – often use

$$\|d_l^n\|_2 \leq \varepsilon$$

- For accurate function & derivative

$$\|d_l^n\|_2 \leq 2^{-n} \varepsilon$$



# *Iterative solution scheme – v1*

Suppose  $\psi$  is the exact eigenfunction and we seek a correction  $\Delta$  to an approximate eigenvalue  $\varepsilon$

$$\begin{aligned}\psi &= -(T - \varepsilon - \Delta)^{-1} (V\psi) \\ &= -(T - \varepsilon)^{-1} (V\psi) - \Delta(T - \varepsilon)^{-2} (V\psi) + \dots\end{aligned}$$

Left projection with  $V\psi$  yields

$$\Delta = -\langle V\psi | \psi - \bar{\psi} \rangle$$

where

$$\bar{\psi} = -(T - \varepsilon)^{-1} (V\psi)$$

The correction to the wavefunction is then

$$\delta = \frac{\bar{\psi}}{\|\bar{\psi}\|} - \psi$$

Overall linear convergence

Error in eigenvalue is quadratic in error in wavefunction (for linear problem) due to related variational approach.

For many-electron systems, we must extract multiple eigenpairs from the Fock operator. Straightforward iteration of the integral equation does not work because all roots will collapse to the lowest root unless the initial guesses are very close to the correct solutions. Two modifications are necessary. First, we use deflation to recast the integral equation for each orbital as a ground-state problem. Let  $P_i$  denote a projector onto the space of the eigenfunctions of lower energy than orbital  $i$ . At convergence, the  $i$ 'th occupied orbital ( $\phi_i$ ) will be the lowest energy solution of

$$(1 - P_i)H(1 - P_i)\phi_i = \epsilon_i\phi_i, \quad (38)$$

which may be rearranged as

$$(H - P_iH(1 - P_i) - (1 - P_i)HP_i + P_iHP_i) \phi_i = \epsilon_i\phi_i. \quad (39)$$

Since  $P_i\phi_i = 0$ , only the first two terms on the left hand side are non-zero. The second term may be included in the potential, thereby incorporating the effect of deflation into iteration of the integral equation. However, we note that if, prior to each iteration, the Hamiltonian or Fock matrix is diagonalized in the space of occupied orbitals, then the second term is also zero and the unmodified integral equation may be used. The second modification is to orthogonalize the updated orbitals in order of increasing energy.

1. destroys pure asymptotic form of eigenfunctions
2. requires additional work to compute differential operator, and
3. requires the integral and differential operators to be consistent to desired precision
4. solves for eigenfunctions which may be delocalized



## Strontium *s* orbitals

# *Iterative Solution v1*

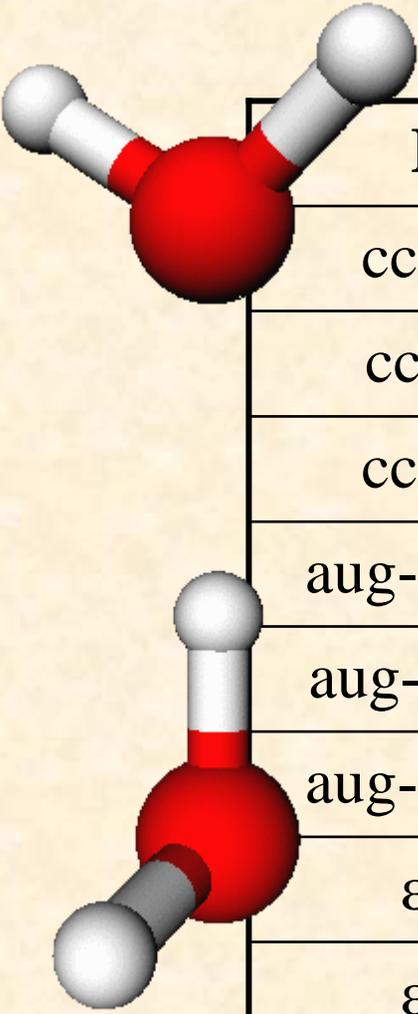
- Form update as described above
- Use this as residual in inexact-Newton (Krylov subspace) non-linear solver
  - Well preconditioned (residual is approximate correction to solution)
- Multiscale solution
  - Solve to low precision at low resolution/threshold
  - Repeat at higher precision with tighter threshold (analogous to h-p refinement)

# *Iterative Solver v1*

- Works well for relatively small systems
  - Typical #iterations 7, 3, 1, 1, ...  
 $(k, \varepsilon) = (5, 10^{-3}), (7, 10^{-5}), (9, 10^{-7}), (11, 10^{-9}), \dots$
- Problems
  - Convergence stalls for large molecules
    - hypothesized to be due to accidental degeneracy leading to poorly resolved rotations
    - inconsistent handling of orthogonality constraints (enforced after making the updated)
  - Inefficient for large molecules (esp. high symmetry) due to delocalized eigenvalues
    - Often only need space spanned by the eigenfunctions

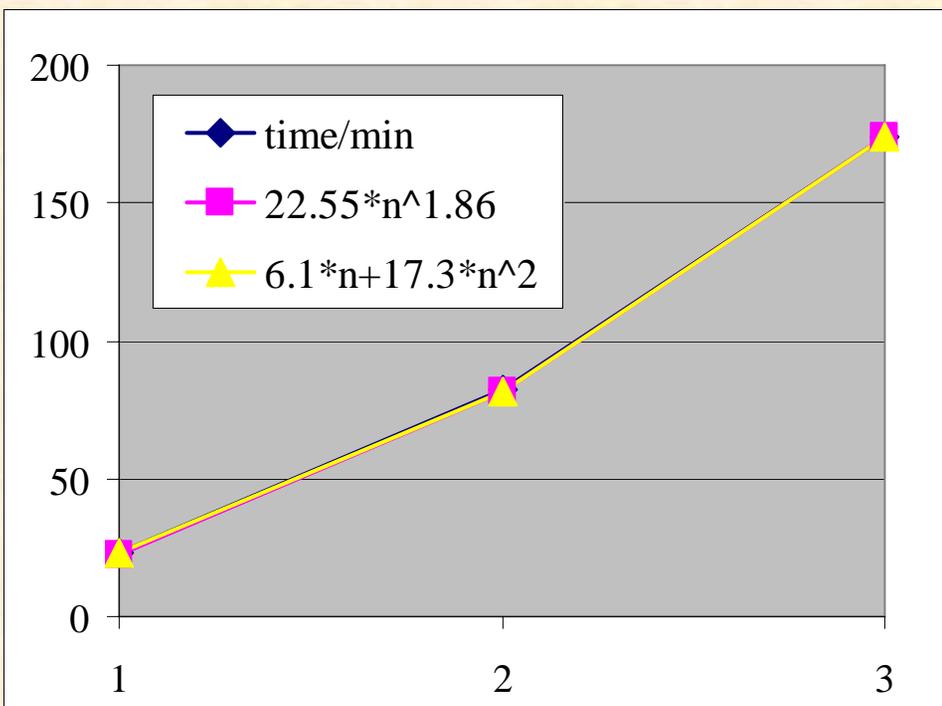
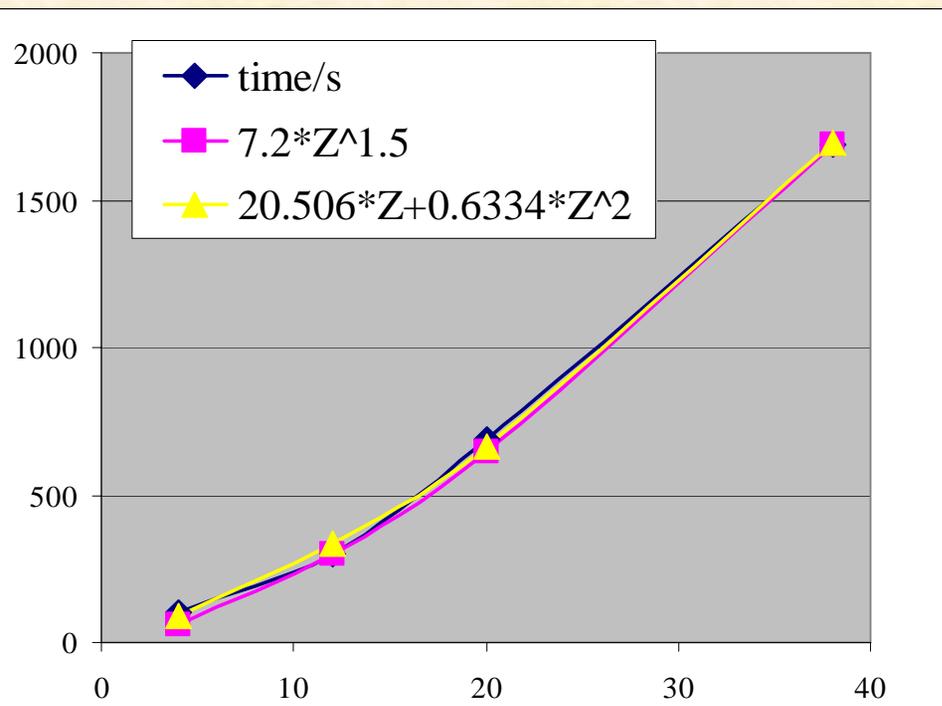
# Water dimer LDA

*aug-cc-pVTZ geometry, kcal/mol.*



<b>Basis</b>	<b>Uncorrected</b>	<b>BSSE</b>	<b>Corrected</b>
cc-pVDZ	-11.733	-3.958	-7.775
cc-pVTZ	-9.464	-1.654	-7.810
cc-pVQZ	-8.708	-0.821	-7.888
aug-cc-pVDZ	-8.187	-0.382	-7.805
aug-cc-pVTZ	-7.992	-0.086	-7.906
aug-cc-pVQZ	-7.995	-0.054	-7.941
$\epsilon=10^{-3}$	-6.483		
$\epsilon=10^{-5}$	-7.932		
$\epsilon=10^{-7}$	-7.943		

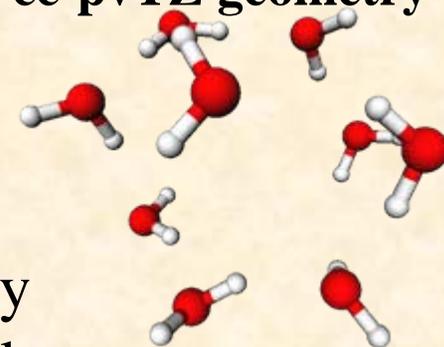
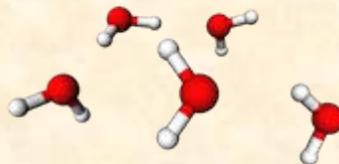
# LDA scaling with Z and system size (energy $\epsilon=10^{-5}$ )



Rare earth atoms  $Z=4,12,20,38$

$(C_6H_6)_n$  MP2 aug-cc-pvTZ geometry

$(H_2O)_n$   $n=5,9 \dots t = O(n^{1.1})$



Stacked benzene – MOs are delocalized by symmetry

Water cluster – MOs are asymptotically localized

(long tail is smooth so is inexpensively treated)

# *Iterative Solver v2*

- Solve for a set of functions to span the space of the lowest  $N$  eigenfunctions
  - C.f., Bernholc, Yang, ...
  - Choose computationally efficient resolution
    - Maximally localized set
    - Strive for overall linear scaling

Return to the variational form of the eigen-problem and introduce the Lagrangian including the orthonormality constraints.

$$L = E - \sum_{i,j=1}^n \lambda_{ij} (\langle i | j \rangle - \delta_{ij})$$

In the linear case,

$$E = \sum_{i=1}^n \langle i | H | i \rangle$$

and in the non-linear case define

$$\frac{\partial E}{\partial \phi_i} = 2H |i\rangle$$

Differentiation w.r.t. the eigenfunctions and multipliers

$$H |i\rangle - \sum_j |j\rangle \lambda_{ji} = 0$$

$$\langle i | j \rangle - \delta_{ij} = 0$$

At this point normally invoke invariance of trace or energy w.r.t. rotations and choose to diagonalize  $\Lambda$

A common approach to generating localized orthonormal orbitals is to solve

$$\min \sum_{i \neq j} \langle i^2 | j^2 \rangle$$

where  $\langle i^2 | j^2 \rangle = \int \phi_i^2 M \phi_j^2 dr^3$  for some positive operator  $M$ .

(Identity - Niessen, square dipole - Boys, Coulomb - Edmiston-Ruedenberg).

This is equivalent to

$$\max \sum_i \langle i^2 | i^2 \rangle$$

A Newton iteration is readily constructed from an orthonormal rotation formed by exponentiating an antisymmetric matrix  $K$

$$W = \sum_i \langle (e^K i)(e^K i) | (e^K i)(e^K i) \rangle$$

$$\left. \frac{\partial W}{\partial k_{ji}} \right|_{K=0} = 4 \langle ji | ii \rangle$$

$$\left. \frac{\partial^2 W}{\partial k_{st} \partial k_{ji}} \right|_{K=0} = 4 \delta_{ti} (\langle js | ii \rangle + 2 \langle ji | si \rangle) + 4 \delta_{st} \delta_{tj} \langle si | ii \rangle$$

The integral equations becomes

$$r_i = \phi_i + (T - \xi_i)^{-1} \left( (V - \xi_i) \phi_i - \sum_j \phi_j \lambda_{ji} \right) = 0$$

$$\langle \phi_i | \phi_j \rangle - \delta_{ij} = 0$$

where  $\Lambda$  must be treated as a sparse matrix and  $\xi_i$  a negative energy probably best chosen as  $\lambda_{ii}$ .

With analogy to the conventional approach of diagonalizing the matrix representation of the Hamiltonian, we first compute the residuals  $r_i$  and then solve in a least squares sense a linearized version of

$$r_i = 0$$

$$\langle \phi_i | \phi_j \rangle - \delta_{ij} = 0$$

fully maintaining the orthonormality constraints.

After (during) determining this update, the orthonormal orbitals are subjected to one (several) Newton iteration of the localization procedure.

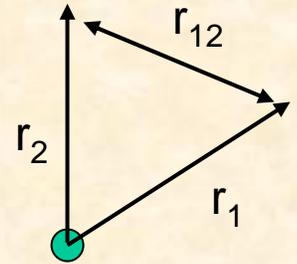
*Iterative  
Solver v2*

# Electron correlation

- All defects in the mean-field model are ascribed to electron correlation
- Consideration of singularities in the Hamiltonian imply that for a two-electron singlet atom (e.g., He)

$$\Psi(r_1, r_2, r_{12}) = 1 + \frac{1}{2} r_{12} + O(r_{12}^2) \quad \text{as } r_{12} \rightarrow 0$$

- Include the inter-electron distance in the wavefunction



- E.g., Hylleraas 1938 wavefunction for He

$$\Psi(r_1, r_2, r_{12}) = e^{-\zeta(r_1+r_2)} (1 + ar_{12} + \dots)$$

- Potentially very accurate, but not systematically improvable, and (until recently) not computationally feasible for many-electron systems

# Conventional approach

- The two-electron wave function is expanded as a product of one-particle functions (orbitals)

$$\Psi(r_1, r_2) = \sum_{ij} c_{ij} \phi_i(r_1) \phi_j(r_2)$$

- Can prove for atoms, that if saturate the atomic basis up to some angular momentum  $L$ , then

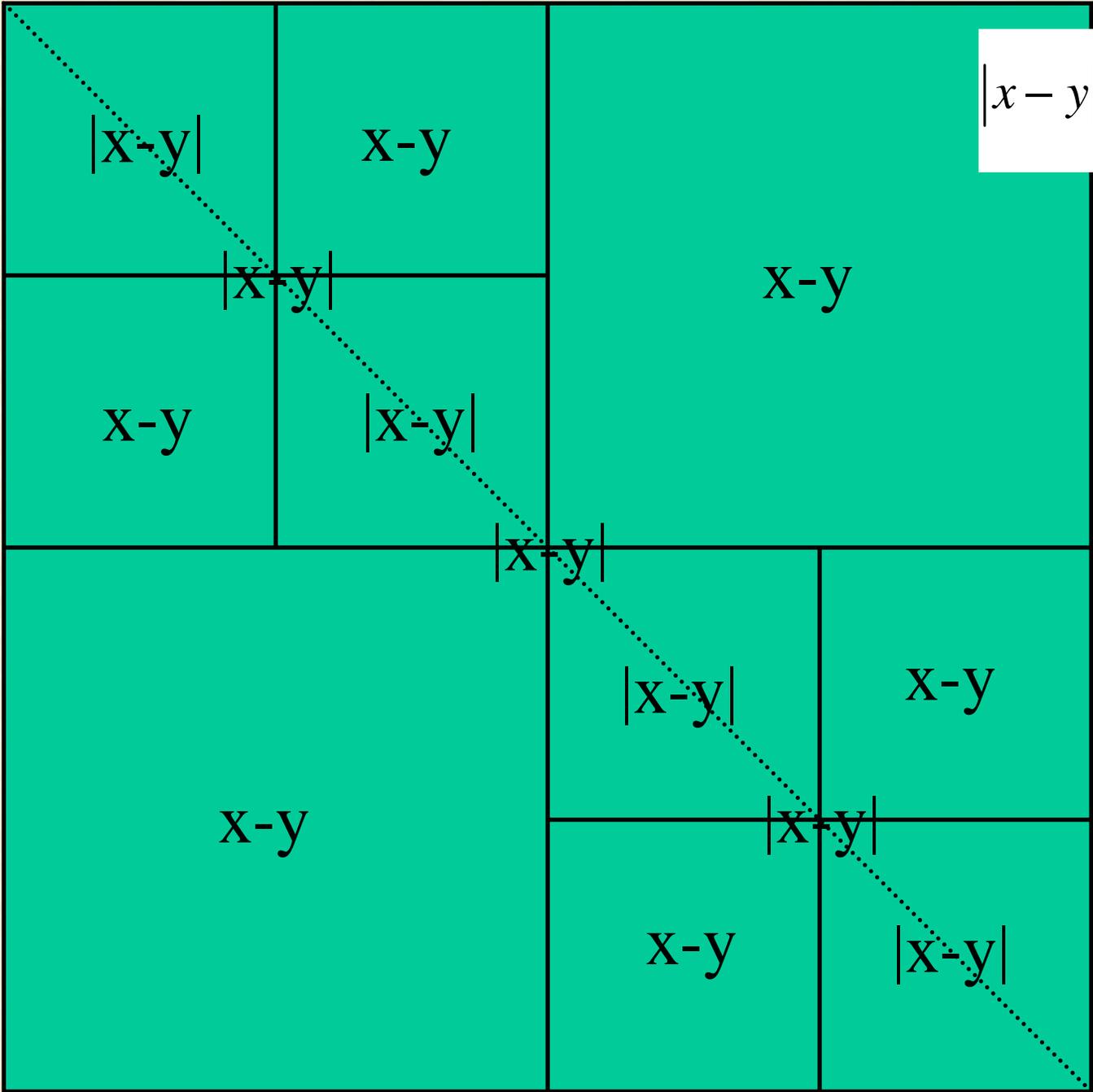
$$\Delta E_{corr} \propto (L + 1)^{-3}$$

L	d	f	g	h
$(L+1)^{-3}$	0.04	0.016	0.008	0.0046

- Correlation consistent basis sets (Dunning) are currently the best choice – cost is  $O(\varepsilon^{-4})$
- Explicitly correlated wave functions yields  $O(\varepsilon^{-2.4})$
- Fully numerical promises  $O(\log \varepsilon^{-1})$

$x \rightarrow$

$y \downarrow$



$$|x - y| = \sum_{\mu=1}^r f_{\mu}(x) g_{\mu}(y)$$

$r =$  separation rank

In 3D, ideally must be one box removed from the diagonal

Diagonal box has full rank

Boxes touching diagonal (face, edge, or corner) have increasingly low rank

Away from diagonal  $r = O(-\log \epsilon)$

# *Multiresolution solver of two-electron Schrödinger equation*

- Wavefunction in 6-D multiresolution representation
- Solve integral equation
  - The 6D GF nominally has 12 indices! Separated representation of operator accurate and efficient
- Partly or fully use SVD to represent 6-D tensor coefficient sets
  - Blocks separated from the diagonal have low rank (1 or 2, the full rank being  $k^3$ )
  - Directly analogous to linear CI expansion but not global
- Can compute directly in this form, but other refinements make it much more practical

# Summary

- Multiresolution in multi-wavelet bases
  - Made practical in three and higher dimensions by separated representations
- Integral form of non-linear eigen-problem
  - Well preconditioned
  - Original formulation not consistent thru first order with poor treatment of accidental degeneracies
  - New formulation
    - consistent (?)
    - well resolved (must still treat ambiguity in localization)
    - hopefully much more efficient ( $O(N)$  ?)

# Separated form for integral operators

$$T * f = \int ds K(r - s) f(s)$$

- Approach in current prototype code
  - Represent the kernel over a finite range as a sum of Gaussians

$$r_{ii',jj',kk'}^{nl} = \sum_i \omega_i X_{ii'}^{nl_x} Y_{ii'}^{nl_y} Z_{ii'}^{nl_z} + O(\varepsilon) \quad \text{in 3D}$$

$$K(r) = \sum_i \omega_i e^{-t_i r^2} + O(\varepsilon)$$

- Only need compute 1D transition matrices (X,Y,Z)
- SVD the 1-D operators (low rank away from singularity)
- Apply most efficient choice of low/full rank 1-D operator
- Even better algorithms not yet implemented

# Accurate Quadratures

$$\begin{aligned}\frac{e^{-\mu r}}{r} &= \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-x^2 t^2 - \mu^2 / 4t^2} dt \\ &= \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2 e^{2s} - \mu^2 e^{-2s} / 4 + s} ds\end{aligned}$$

- Trapezoidal quadrature
  - Geometric precision for periodic functions with sufficient smoothness.

The kernel for  $x=1e-4, 1e-3, 1e-2, 1e-, 1e0$ .  
The curve for  $x=1e-4$  is the rightmost

Automatically generated representations of  $\exp(-30r)/r$  accurate to  $1e-10$ ,  $1e-8$ ,  $1e-6$ ,  $1e-4$ ,  $1e-2$  (relative error) for  $r$  in  $[1e-8, 1]$  (92, 74, 57, 39 and 21 terms, respectively).

Low-energy *scattering* states also possible (but stronger dependence on range)

Periodic systems (cubic subgroups) straightforward.

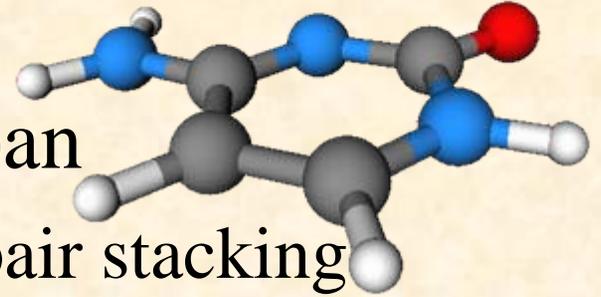
# *Current Capabilities*

- Open/closed shell Hartree-Fock and DFT
  - Wide range of GGAs, hybrid ( $O(N)$  HF exchange), and asymptotically corrected functionals
  - Energies and analytic derivatives
  - Full TDDFT and RPA for excitation energies \*\*\*\*\*
  - Abelian point groups
  - Parallel execution on shared memory computers
  - Interfaces to NWChem and GAMESS-US
- Working prototypes for computing in 6D
  - Direct solution of the pair equation for polyatomic systems (initial target is basis-set limit MP2)



## *Prototype code in use by ...*

- Hideo Sekino at Toyohashi, Japan
  - Benchmark calculations of base pair stacking
  - Electron transport in molecules
- Schaefer and Allen U. Georgia, Athens, USA
  - Benchmark HF energies and structures
- Tennant, U. Sheffield, England
  - New solvation models



# *High-level composition using functions and operators*

- Conventional quant. chem. uses explicitly indexed sparse arrays of matrix elements
  - Complex, tedious and error prone

- Python classes for Function and Operator

- in 1,2,3,6 and general dimensions  $H\varphi = -\frac{1}{2}\nabla^2\varphi + V\varphi$

- wide range of operations

`Hpsi = -0.5*DelSq*psi + V*psi`

`J = Coulomb.apply(rho)`

$$J(r) = G^* \rho$$

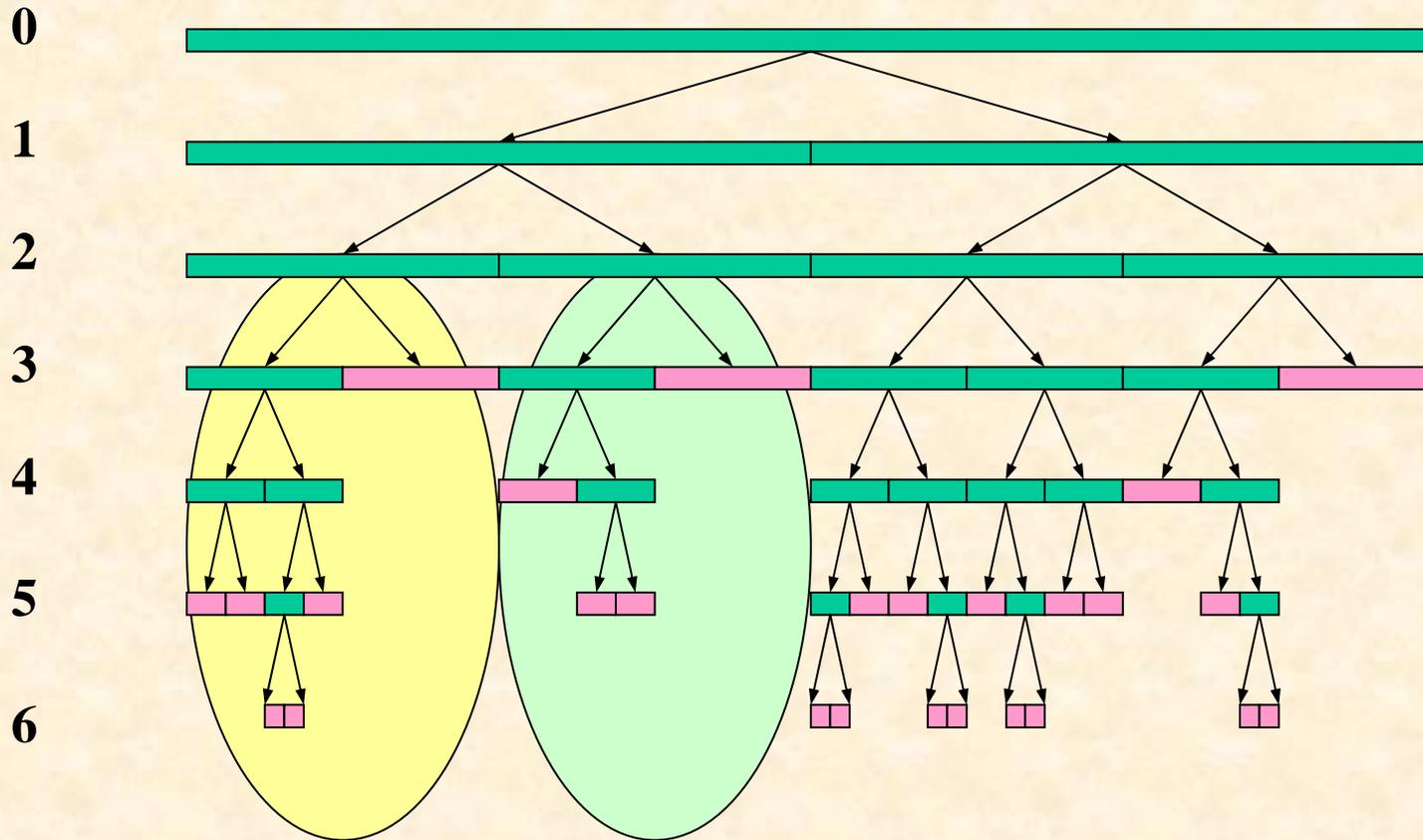
$$= \int \frac{\rho(s)}{|r-s|} ds$$

- All with guaranteed speed and precision

# *New solver being developed*

- Working with localized orbitals
  - $O(1)$  application of operators to one orbital
  - $O(N)$  computation of Coulomb potential (already)
  - $O(N)$  computation of Fock-like matrices
    - As a result of localized orbitals
  - More robust convergence
- Near total rewrite in C++
  - Two-levels of parallelism targeting massively parallel computer using multi-processor nodes
  - In anticipation of highly-threaded processors

# 1-D Example Sub-Tree Parallelism



Both sub-trees can be done in parallel.

In 3-D nodes split into 8 children ... in 6-D there are 64 children

# *Next Generation Languages/Mechanisms*

- X10 (IBM) (funded under DARPA HPCS)
  - Derivative of Java
  - Futures/Async(Dynamic Scheduling)
  - Var@Place (notion of locality)
- Chapel (Cray) (funded under DARPA HPCS)
  - Roots in ZPL, MTA stuff, and others
  - Futures
  - Domains (locality and scheduling)
- In both cases still waiting on a more complete/up-to-date language spec

# *Futures*

- Parallel language construct that simplifies parallel execution with a dependency upon result
  - Store an unevaluated expression in a variable of type Future
  - May be executed by another thread as resources permit
  - Reference to the result forces it to be evaluated if it has not yet been done so.
- Implementation as templated C++ class with thread pool (similar interface to Java)

# *Recursive composition with Futures*

## Sequential

```
void Function::_reconstruct(OctTreeT *t) {  
    Tensor<double> ss = unfilter(t->data());  
  
    FORIJK(OctTreeT *child = t->child(i,j,k);  
        if (child) {  
            child->data()...;  
            _reconstruct(child);  
        }  
        else {  
            t->insert_child(...);  
        }  
    });  
};
```

## Multithreaded

```
void Function::_reconstruct(OctTreeT *t) {  
    Tensor<double> ss = unfilter(t->data());  
  
    Future<void,Function,OctTreeT*> fut[2][2][2];  
  
    FORIJK(OctTreeT *child = t->child(i,j,k);  
        if (child) {  
            child->data()...;  
            fut[i][j][k].start(&Function::_reconstruct, this, child);  
        }  
        else {  
            t->insert_child(...);  
        }  
    });  
  
    FORIJK( fut[i][j][k].force(); );  
};
```

# Analytic removal of cusp(s)

- Two-electron Hamiltonian
 
$$H = -\frac{1}{2}\Delta_1 - \frac{1}{2}\Delta_2 + V(r_1, r_2) + \frac{1}{|r_1 - r_2|}$$
- Two-electron wavefunction
 
$$\Psi(r_1, r_2) = e^{u(|r_1 - r_2|)} \phi(r_1, r_2)$$
- Transformed problem
 
$$e^{-u} H e^u \phi = -\frac{1}{2}\Delta_1 \phi - \frac{1}{2}\Delta_2 \phi - u' \frac{r_1 - r_2}{|r_1 - r_2|} \cdot (\nabla_1 \phi - \nabla_2 \phi) + \left( V - ((u')^2 + u'') + \frac{1}{|r_1 - r_2|} (1 - 2u') \right) \phi$$
- Choose  $u$  to eliminate singularity at  $r_{12}=0$
- Quantum Monte Carlo calculations
  - Best to eliminate all  $r_1=0, r_2=0, r_{12}=0, r_1=r_2=r_{12}=0$

# Smoothed potential and wave function

- Similarity-transformed Hamiltonian with correlation factor (cf. transcorrelated Hamiltonian)
  - The effective wavefunction  $\Phi$  as well as the transformed Hamiltonian is smoothed at  $r_{12}=0$

$$\bar{H} = e^{-u(r_{12})} \hat{H} e^{u(r_{12})}$$

Correlation factor:  $u(r_{12})$

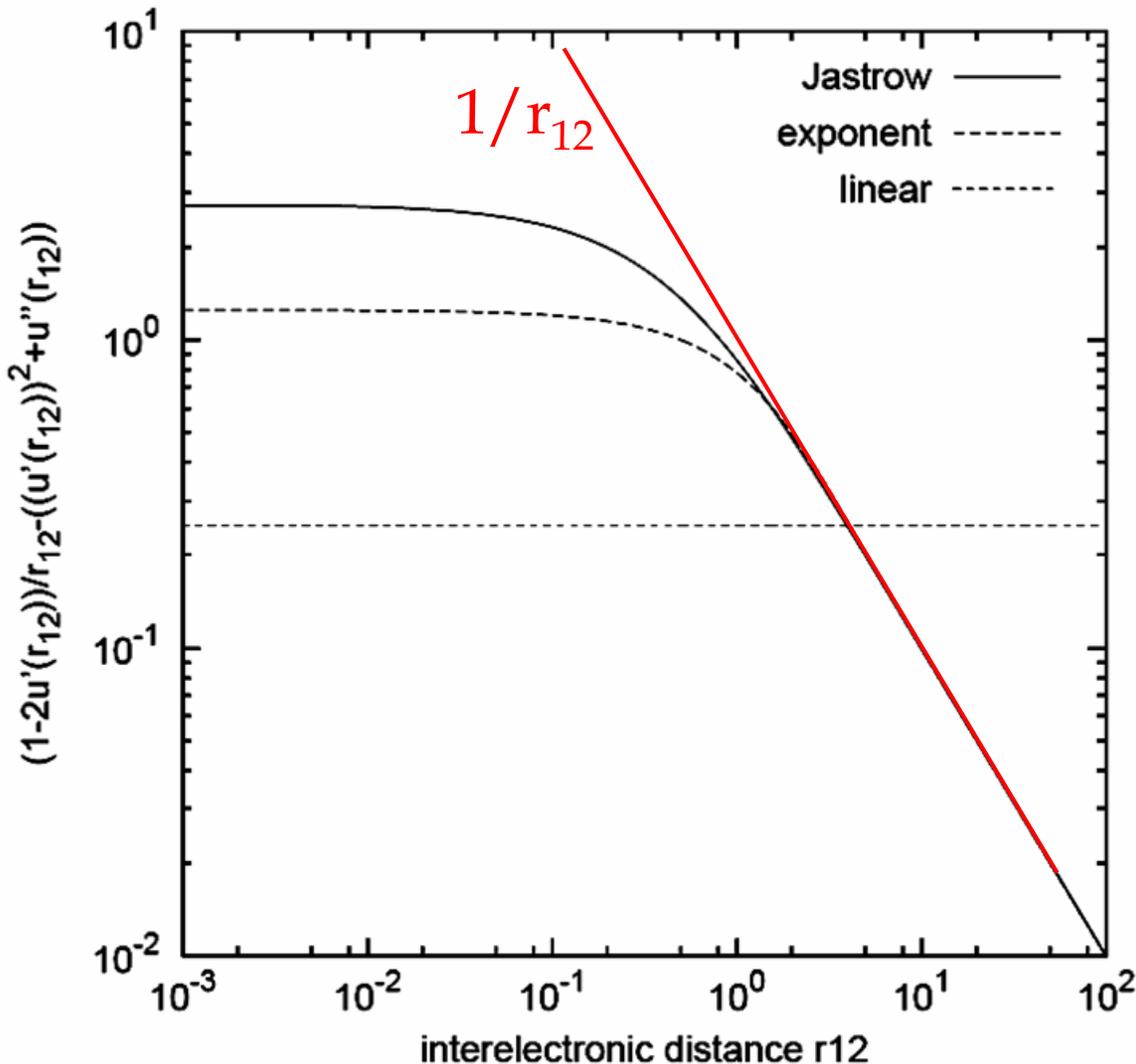
$$\Psi(r_1, r_2) = e^{u(r_{12})} \Phi(r_1, r_2)$$

- Electron-electron repulsion is smoothed

$$\frac{1}{r_{12}} \rightarrow \frac{1}{r_{12}} (1 - 2u'(r_{12})) - u'(r_{12})^2 - u''(r_{12}) - u'(r_{12}) \frac{\mathbf{r}_1 - \mathbf{r}_2}{r_{12}} \cdot (\nabla_1 - \nabla_2)$$

# Smoothed interelectron potential

- Correlation factor



- Jastrow

$$u(r_{12}) = \frac{ar_{12}}{1 + br_{12}}$$

- Exponential

$$u(r_{12}) = a(1 - e^{-br_{12}})$$

- Linear

$$u(r_{12}) = \frac{1}{2}r_{12}$$

# *Solve for the correlation correction to the HF wavefunction*

- The smoothed wavefunction is separated into Hartree-Fock wavefunction and its perturbation.
  - The perturbative wavefunction is numerically smaller than Hartree-Fock wavefunction.

$$\Phi(r_1, r_2) = \Phi^{HF}(r_1, r_2) + \underline{\delta\Phi(r_1, r_2)}$$

$$\left\{ \begin{array}{l} \|\Phi^{HF}(r_1, r_2)\|_2 = 1.0 \\ \|\underline{\delta\Phi(r_1, r_2)}\|_2 \approx 0.1 \quad \text{For He atom} \end{array} \right.$$

# *Preliminary results for He atom*

	Variational E	$\Delta E$	residual
<b>HF</b>	<b>-2.861 61</b>		
<b>Iter. 0</b>	<b>-2.871 08</b>		<b>0.414 73</b>
<b>1</b>	<b>-2.894 92</b>	<b>-0.023 84</b>	<b>0.017 28</b>
<b>2</b>	<b>-2.900 43</b>	<b>-0.005 51</b>	<b>0.007 94</b>
<b>3</b>	<b>-2.902 18</b>	<b>-0.001 75</b>	<b>0.003 84</b>
<b>4</b>	<b>-2.902 88</b>	<b>-0.000 70</b>	<b>0.002 02</b>
<b>5</b>	<b>-2.903 20</b>	<b>-0.000 32</b>	<b>0.001 25</b>
<b>6</b>	<b>-2.903 39</b>	<b>-0.000 20</b>	<b>0.000 91</b>
<b>...</b>	<b>...</b>	<b>...</b>	<b>...</b>
<b>12</b>	<b>-2.903 73</b>	<b>-0.000 04</b>	<b>0.000 36</b>
<b>13</b>	<b>-2.903 73</b>	<b>+0.000 004</b>	<b>0.000 32</b>
<b>14</b>	<b>-2.903 77</b>	<b>-0.000 04</b>	<b>0.000 28</b>

## Computational details:

- 5-th order multiwavelets
- Wavelet threshold:  $2 \times 10^{-5}$
- SVD threshold:  $2 \times 10^{-6}$
- Exponential correlation factor

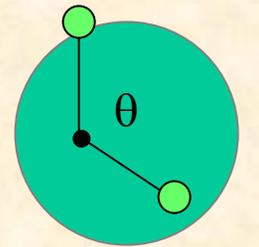
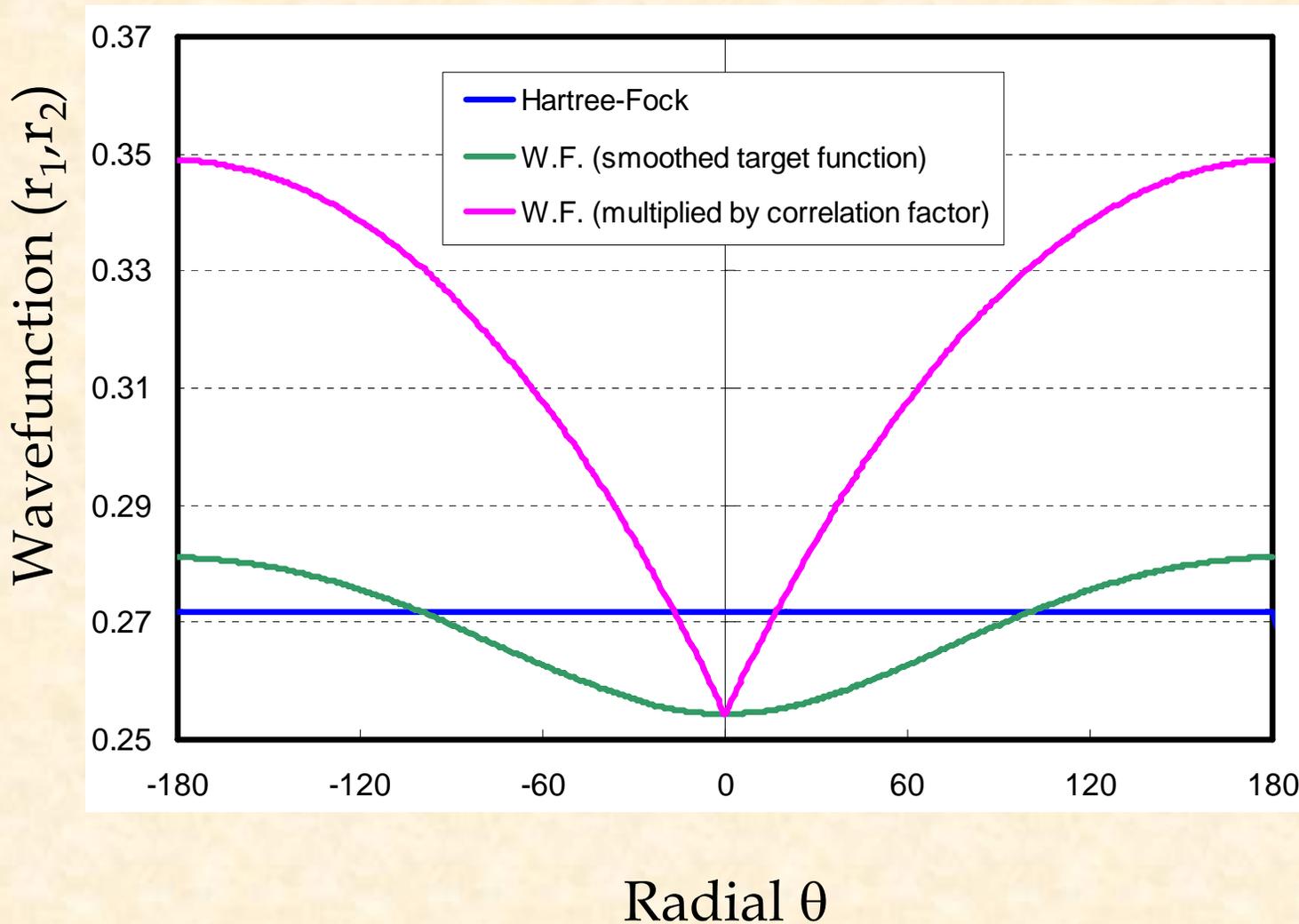
## Perturbative wavefunction:

- Maximum refinement:  $n=4$
- Memory: 132M in full SVD form

<b>exact</b>	<b>-2.903 74 (E(HF)=-2.861 68)</b>
<b>Hylleraas (6 terms)</b>	<b>-2.903 24</b>
<b>Löwdin and Redei</b>	<b>-2.895 4</b>
<b>cc-pV6Z</b>	<b>-2.903 48 (FCI) (E(HF)= -2.861 67)</b>

-Energy is variational  
(small non-variational is  
just truncation err)

# *Coulomb hole (He)*



$$r_1 = r_2 = 0.5a_0$$

# Summary

- Multiresolution analysis provides a general framework for computational chemistry
  - Accurate and efficient with high-level composition
  - Multiwavelets provide high-order convergence and readily accommodate singularities/boundary conditions
  - General framework readily accessible to researchers
  - Real impact will be application to many-body models
- Separated form for operators and functions
  - Critical for efficient computation in higher dimension
- Precision is guaranteed
  - Excited states, non-linear response, ...

# “Time-dependent” DFT

- Linear response to frequency dependent perturbation

– Eigenvalues yield excitation energies

Transition density  
for one excited state

$$(1 - \hat{\rho}^0) \left[ \left( \hat{F}^0 - \varepsilon_i^0 \right) x_i(r) + \frac{\partial V_{coul}}{\partial \rho} * \left( \sum_j^{occ} x_j(r) \phi_j^\dagger(r') \right) \phi_i(r) \right] = \omega x_i(r)$$


– Actually solve corresponding integral equation

# *Low Separation Rank Representation*

$$f(x_1, \dots, x_d) = \sum_{l=1}^M \sigma_l \prod_{i=1}^d f_i^{(l)}(x_i) + O(\varepsilon)$$

$$\|f_i^{(l)}\|_2 = 1 \quad \sigma_l > 0$$

- Different from low operator rank
  - Identity has full operator rank, but unit separation rank
- Beylkin & Mohlenkamp Proc. Nat. Acad. 2002
  - Many functions and operators have low sep. rank
  - E.g., Poisson GF; the many-electron Schrödinger Hamiltonian
- We are combining adaptive multires. & separated representations to compute in 6D

*Putting it all together –  
A path to  $O(N)$  exact MP2*

- HF provably  $O(N)$  to arbitrary finite precision
  - Based upon the density matrix (Geodecker, Beylkin and Coult)
  - Orbital-based schemes more expensive without enforcing localization (e.g., Bernholc)
- Need an MP2 scheme based upon density matrices

# *Characteristics of H and P*

- Hamiltonian
  - Globally and locally has low separation rank
  - Globally has full operator rank
  - Locally off-diagonal blocks have low op-rank
- Density matrix has low separation & operator rank
  - Full density matrix has operator rank  $N_{elec}$
  - Local blocks of the density matrix have a much lower operator rank independent of  $N_{elec}$ 
    - Due to localization and orbitals being locally smooth
  - E.g., water monomer 2.6, dimer 2.5, trimer 2.3 average off-diagonal block rank.

# *Shotglass (not yet the holy grail)*

- Global separation of the Hamiltonian
- Sign function (purification) iteration to produce the density matrix

$$S_0 = F / \|F\|_2$$

$$S_{k+1} = (3S_k - S_k^3) / 2, \quad k = 0, 1, \dots$$

- Python classes (+ Fortran)
- Working, but not practical
  - Cost of maintaining LS form is too high
  - LS rank  $r$  is  $O(N)$  ... rank reduction costs  $O(r^3)$

# *Efficient representation of the density matrix and Green's Function*

- Currently exploring these representations
  - Global low separation rank
  - Local low separation rank + NS/S forms
  - Local low operator rank + NS/S forms ←

# The Resolvent has low separation rank

- Already known – Almlöf Laplace factorization

$$\left(\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j\right)^{-1} = \int_0^{\infty} e^{-\left(\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j\right)t} dt$$

$\varepsilon_i$  = Fock matrix  
eigenvalue

$i, j$  = occupied states

$a, b$  = unoccupied states

$$\approx \sum_{\mu=1}^M \omega_{\mu} e^{-\left(\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j\right)t_{\mu}}$$

$$E_2 = -\frac{1}{4} \sum_{ijab} \frac{\langle ij || ab \rangle^2}{\left(\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j\right)}$$

$$\approx -\frac{1}{4} \sum_{ijab\mu} \langle i_{\mu} j_{\mu} || a_{\mu} b_{\mu} \rangle^2$$

$$|i_{\mu}\rangle = \omega_{\mu}^{1/8} e^{-\left(\varepsilon_f - \varepsilon_i\right)t_{\mu}} |i\rangle \quad |a_{\mu}\rangle = \omega_{\mu}^{1/8} e^{-\left(\varepsilon_a - \varepsilon_f\right)t_{\mu}} |a\rangle$$

# Density matrix form of MP2

$$\begin{aligned}
 E_2 &= -\frac{1}{4} \sum_{ijab\mu} (i_\mu a_\mu | j_\mu b_\mu) \left( 2(i_\mu a_\mu | j_\mu b_\mu) - (i_\mu b_\mu | j_\mu a_\mu) \right) \\
 &= -\frac{1}{4} \sum_{\mu} \int 2\rho_{\mu}(r_1, r_3) \rho_{\mu}(r_2, r_4) q_{\mu}(r_1, r_3) q_{\mu}(r_2, r_4) r_{12}^{-1} r_{34}^{-1} dr_1 dr_2 dr_3 dr_4 \\
 &\quad + \text{exchange} \\
 &= -\frac{1}{4} \sum_{\mu} \int 2V_{\mu}(r_2, r_3) V_{\mu}(r_2, r_3) dr_2 dr_3 + \text{exchange}
 \end{aligned}$$

$$\rho_{\mu}(r_1, r_2) = \omega_{\mu}^{1/4} e^{(H_0(1) - \varepsilon_f)t} e^{(H_0(2) - \varepsilon_f)t} \rho(r_1, r_2)$$

$$q = 1 - \rho$$

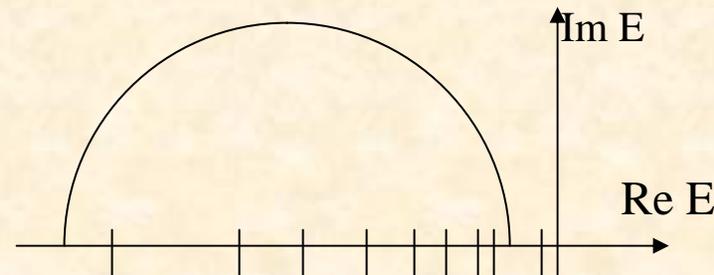
$$V_{\mu}(r_2, r_3) = \int \rho_{\mu}(r_1, r_3) q_{\mu}(r_1, r_3) r_{12}^{-1} dr_1$$

# *The KS Green's Function*

- Purification loses precision for all-electron potential, has limited opportunity for parallelism, and requires full precision throughout
- GF approach is self-correcting, efficient (iterative solution of well-conditioned equations), and highly parallel

$$G(r, r'; E) = G_0(r, r'; E) + \int G_0(r, r''; E) V(r'') G(r'', r'; E) dr''$$

$$\rho(r, r') = \frac{-1}{2\pi} \text{Im} \int_{-\infty}^{-\mu} G(r, r'; E) dE$$



# Summary

- Multiresolution analysis provides a general framework for computational chemistry
  - Accurate and efficient with high-level composition
  - Multiwavelets provide high-order convergence and readily accommodate singularities and boundary conditions
  - Real impact will be application to many-body models for which we must compute in 6D
- Separated form for operators and functions
  - Critical for efficient computation

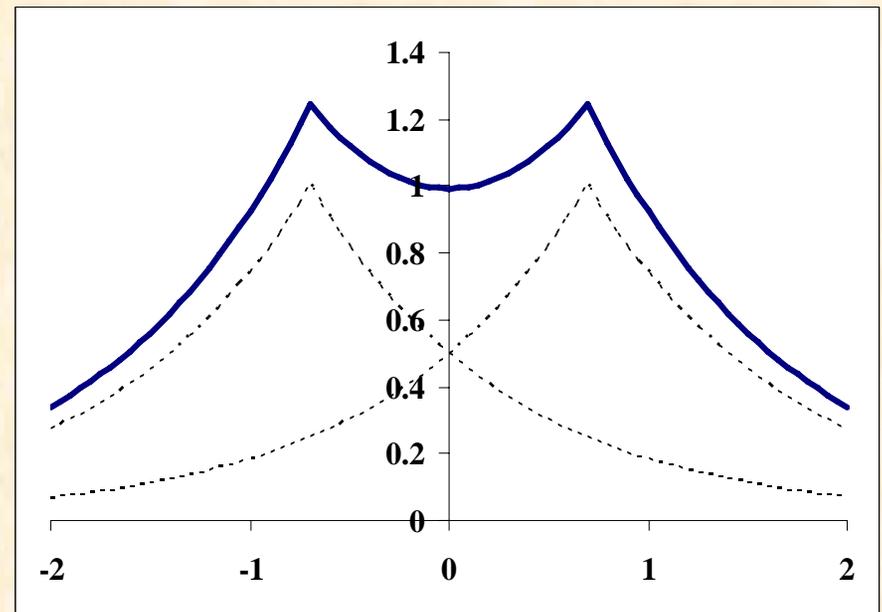
# Linear Combination of Atomic Orbitals (LCAO)

- Molecules are composed of (weakly) perturbed atoms
  - Use finite set of atomic wave functions as the basis
  - Hydrogen-like wave functions are exponentials
- E.g., hydrogen molecule ( $\text{H}_2$ )

$$1s(r) = e^{-|r|}$$

$$\phi(r) = e^{-|r-a|} + e^{-|r-b|}$$

- Smooth function of molecular geometry
- MOs: cusp at nucleus with exponential decay



# LCAO

- A fantastic success, but ...
- Basis functions have extended support
  - causes great inefficiency in high accuracy calculations
  - origin of non-physical density matrix
- Basis set superposition error (BSSE)
  - incomplete basis on each center leads to over-binding as atoms are brought together
- Linear dependence problems
  - accurate calculations require balanced approach to a complete basis on every atom
- Must extrapolate to complete basis limit
  - unsatisfactory and not feasible for large systems

The Kohn-Sham orbital LDA equations (equations 7.2.7-9 and 7.4.3 in [1]) result from minimization of the LDA energy functional

$$E[\rho] = 2 \sum_{i=1}^N \int dr \phi_i(r) \left( -\frac{1}{2} \nabla^2 + V_{coul}(r) + V_{ext}(r) \right) \phi_i(r) + E_{xc}[\rho] + E_{nuc}, \quad (15)$$

with respect to variation of the occupied orbitals ( $\phi_i(r)$ ,  $i = 1, \dots, N$ ) which define the electron density  $\rho$ ,

$$\rho(r) = 2 \sum_{i=1}^N |\phi_i(r)|^2. \quad (16)$$

The negative Laplacian is the kinetic energy operator. In this paper, the external potential,  $V_{ext}$ , includes only the attraction of the electrons to the nuclei,

$$V_{ext}(r) = - \sum_{\alpha} \frac{Z_{\alpha}}{|r - r_{\alpha}|}. \quad (17)$$

The Coulomb potential,  $V_{coul}(r)$ , describes the repulsion between electrons,

$$V_{coul}(r) = \int dr' \frac{\rho(r')}{|r - r'|}. \quad (18)$$

In (15)  $E_{xc}[\rho]$  is the exchange correlation energy, the exact form of which is unknown. The repulsion energy of the nuclei,

$$E_{nuc} = \sum_{\alpha < \beta} \frac{Z_{\alpha} Z_{\beta}}{|r_{\alpha} - r_{\beta}|}, \quad (19)$$

does not depend upon the density, but must be included to obtain the dependence of the total energy upon the nuclear coordinates.

The resulting equations

$$\left[ -\frac{1}{2} \nabla^2 + V(r) \right] \phi_i(r) = \epsilon_i \phi_i(r) \quad (20)$$

$$V(r) = V_{ext}(r) + V_{coul}(r) + V_{xc}^{LDA}(r) \quad (21)$$

define the occupied orbitals as the lowest  $N$  eigenfunctions of the Kohn-Sham operator (also casually referred to as the Fock operator since the Hartree and Hartree-Fock equations are very similar [1]) which implicitly depends upon the orbitals through the density.

# Scaling Function Basis - III

