Experiments in Direct Nodal Optimization (Quantum Monte Carlo in the Apuan Alps IV) July 29, 2008



Jonathan L DuBois

Quantum Simulations Group

Physical Sciences

Lawrence Livermore National Laboratory

Overview



- Problems with fermion DMC
 - Getting the nodes right
 - Noncommuting / off-diagonal observables
- PIGS -- yet another way to sample Green's function
 - An application to evaluation of multipole moments
 - Optimizing nodes
 - Released node
- Future directions



Limitations of canonical fixed node DMC



- Nodal surface comes from more approximate method
- Evaluation of nonlocal pseudopotentials introduces a trial function bias in energy
- Observables which don't commute with the Hamiltonian e.g. charge density have a trial function bias



- Better trial functions
 - + Backflow, Pfaffians, Multiple determinants
 - o Scaling with N_e
 - o Effective optimization

- Released node
 - + Faster Computers
 - + Correlated sampling
 - o Naive Scaling $\sim \exp(N_e)$



Physical Sciences 



Starting from Slater-Jastrow ansatz

$$\Psi_{\mathcal{T}}(\mathbf{R}) = e^{J(\mathbf{R})} \mathcal{D}^{\uparrow}(\mathbf{r}_{1}, \mathbf{r}_{2}, ... \mathbf{r}_{N}) \mathcal{D}^{\downarrow}(\mathbf{r}_{1}, \mathbf{r}_{2}, ... \mathbf{r}_{N})$$
$$J(\mathbf{R}) = \sum_{i} f_{1}(\mathbf{r}_{i}) + \sum_{i,j} f_{2}(\mathbf{r}_{i}, \mathbf{r}_{j}) + \sum_{i,j,k} f_{3}(\mathbf{r}_{i}, \mathbf{r}_{j}, \mathbf{r}_{k}) ...$$

parameterizes many body correlations

$$\mathcal{D}(\mathbf{r_1}, \mathbf{r_2}, \dots \mathbf{r_N}) = \begin{vmatrix} \phi_1(\mathbf{r_1}) & \phi_1(\mathbf{r_2}) & \dots \\ \phi_2(\mathbf{r_1}) & \phi_2(\mathbf{r_2}) & \dots \end{vmatrix}$$

determinant of single particle wfns takes care of fermi symmetry

- Multiple determinants
- Many body functions in determinant (backflow)



Backflow



transform coordinates to incorporate "backflow" correlations into single determinant

$$\begin{split} \tilde{\mathbf{r}}_{\mathbf{i}} &\to \mathbf{r}_{\mathbf{i}} + \eta(\mathbf{r}_{1}, \mathbf{r}_{2}, ..., \mathbf{r}_{N}) \\ \tilde{\mathcal{D}}(\mathbf{r}_{1}, \mathbf{r}_{2}, ... \mathbf{r}_{N}) &= \begin{vmatrix} \phi_{1}(\tilde{\mathbf{r}}_{1}) & \phi_{1}(\tilde{\mathbf{r}}_{2}) & ... \\ \phi_{2}(\tilde{\mathbf{r}}_{1}) & \phi_{2}(\tilde{\mathbf{r}}_{2}) & ... \end{vmatrix}$$

the cost

naive scaling is now $\mathcal{O}(N_e^4)$



Optimizing nodes in VMC (benzene dimer example)



Optimization of backflow nodal surface is hard

parameter space is highly nonlinear difficult to separate error in nodal surface from error in correlations at the level of VMC





Example: Optimizing backflow for argon dimer Physical Sciences DMC (HF nodes) Energy (meV) best pair potential -10 -20

separation (Å)

Physical Sciences

Example: Optimizing backflow for argon dimer





Approaches to optimization



- VMC
 - Variance minimization
 - Energy minimization
 - Too much noise from correlations
 - Some other metric (e.g. Lüchow 2007)
- DMC sidewalks
 - Local Energy
 - dE/dτ
 - Expensive
- Some other way -- PIGS
 - Expensive, noisy, probably won't even work -- Lets do it!





Path Integral Ground state Quantum Monte Carlo

$$P(Y) = \psi_T(R_0)\psi_T(R_L) \prod_{\ell=0}^{L-1} G(R_{\ell+1}, R_\ell, \Delta\tau)$$

PIGS projects in imaginary on a single fixed length path

Challenges:

•No Guiding function!

Must have very accurate propagator

Advantages:

 No population control
 Non-commuting and off-diagonal observables

Physical Sciences

Use "guided" Green's function

Only move endpoints

PIGS:

All beads can move

Reptation vs PIGS

Reptation:

Use unguided "bare" Green's function

$$P(Y) = \psi_{
u}(\mathbf{R}_0)\psi_{
u}(\mathbf{R}_M)\prod_{\ell}^{M-1}G(\mathbf{R}_{\ell+1},\mathbf{R}_{\ell},\Delta au)$$

 $\psi_{\nu}(\mathbf{R}_0)$







 $\psi_{\nu}(\mathbf{R}_M)$

- Physical Contract Physical Contract Physical Contract Physical Contract Physical PhysicaP
- Could use "exact" pair coulomb density matrix (PIMC)
 - Gets the ee and en cusps exactly right
 - Expensive, tricky to implement
 - Quality is density dependent
 - Doesn't help with arbitrary potentials (e.g. from fixed phase)
- Use a general 4th order propagator

$$\begin{split} G(\mathbf{R}, \mathbf{R}', \Delta \tau) &= \int d\mathbf{R}'' e^{c_1 V(\mathbf{R})} e^{c_2 T(\mathbf{R}, \mathbf{R}'')} e^{c_3 \tilde{V}(\mathbf{R}'')} e^{c_2 T(\mathbf{R}'', \mathbf{R}')} e^{c_1 V(\mathbf{R}')} \\ \tilde{V}(\mathbf{R}) &\equiv [V, [T, V]] \end{split}$$

- Do nothing : charge density builds up at nodes [$ln(d\tau)$ scaling]
- Use image approximation to treat nodal action [~ ($d\tau^2$) scaling]





Physical Sciences

~~~~~

## Efficiency compared to DMC (not great)

- Even with nodal action and 4th order propagator
  - PIGS time step ~  $5 \times 10^{-4}$
  - DMC time step ~ 10<sup>-2</sup>
- Surprisingly, PIGS is only ~factor of 20 slower than DMC in total CPU time vs error bar.
- Nodal action appears to be the limiter with  $d\tau^2$  scaling
- Tests with hydrogenic atoms show 4th order propagator is 4th order even for en cusp -- correct g(r)





#### **Application to noncommuting observables**



Multipole Moments for some first row dimers

$$Q_{2} = \frac{ZR^{2}}{2} - \sum_{i} r_{i}^{2} P_{2}(\cos \theta_{i})$$

$$= \frac{ZR^{2}}{2} + \frac{1}{2} \sum_{i} [-3z_{i}^{2} + r_{i}^{2}],$$

$$Q_{4} = \frac{ZR^{4}}{8} - \sum_{i} r_{i}^{4} P_{4}(\cos \theta_{i})$$

$$= \frac{ZR^{4}}{8} + \frac{1}{8} \sum_{i} [-35z_{i}^{4} + 30r_{i}^{2}z_{i}^{2} - 3r_{i}^{4}]$$

|                       | Experiment          | VMC      | PIGS     |
|-----------------------|---------------------|----------|----------|
| _i <sub>2</sub>       | Q 2                 | 7.79(4)  | 8.4(6)   |
|                       | Q 4                 | 72(1)    | 56(7)    |
| <b>D</b> <sub>2</sub> | Q <sub>2</sub> 0.25 | -0.63(5) | 0.38(3)  |
|                       | Q <sub>4</sub>      | 3.8(3)   | 4.8(9)   |
| $N_2$                 | $Q_2$ 1.09 ± 0.07   | -1.53(3) | 0.68(5)  |
|                       | $Q_4 - 8.0 \pm 2.4$ | -8.0(3)  | -10.6(4) |



#### A direct optimization scheme







Maximize Volume / surface area within nodal pocket





- 1. Generate a set of paths  $\{Y_1, Y_2, ..., Y_M\}$  given initial boundary conditions (nodal surface)
- 2. Evaluate average distance to nearest boundary at center of path  $<|R_c^i R_0^i(\alpha)|>$  and  $<\nabla_{\alpha}|R_c^i R_0^i(\alpha)|>$
- 3. Use favorite optimization scheme to generate new parameters  $\alpha$ '
- 4. Repeat 2.-3. with  $\alpha$ ' on original {Y<sub>1</sub>,Y<sub>2</sub>,...,Y<sub>M</sub>} reweight  $<|R_c^i - R_0^i(\alpha')|>$  and  $<\nabla_\alpha |R_c^i - R_0^i(\alpha')|>$  by change in nodal action  $\Pi_l G_{node}(R_l,R_{l+1};\alpha')/G_{node}(R_l,R_{l+1};\alpha)$



Test: Optimizing V/S for single particle in 2D box

Physical OMA

Boundary defined by 4 b-splines : volume conserved



10<sup>6</sup> paths, 16 iterations



Estimate distance to node |R-R<sub>0</sub>| assuming

 $\Psi_{\nu}(\mathbf{R}) = \mathbf{a} \cdot \mathbf{R} + b$ 

 $\mathbf{a} = 
abla f(\mathbf{R})$ 

$$|\mathbf{R} - \mathbf{R}_0| = \left|\mathbf{R} - rac{
abla \Psi_
u(\mathbf{R}) \cdot \mathbf{R} - \Psi_
u(\mathbf{R})}{
abla \Psi_
u(\mathbf{R})}
ight|$$

- This is a poor approximation so results for <|R<sub>c</sub><sup>i</sup> R<sub>0</sub><sup>i</sup>(α)|> must be filtered -- typically 80% of configurations are lost
- Exploring steepest descent to find nearest root.





- Trial function based on Umrigar multideterminant wfn
- Optimize determinant coefficients and orbitals

|                 | Original Umrigar<br>DMC Energy | After node<br>Optimization | Experiment |
|-----------------|--------------------------------|----------------------------|------------|
| Li <sub>2</sub> | -14.9938                       | -14.9941(1)                | -14.9954   |
| N <sub>2</sub>  | -109.505                       | -109.515(3)                | -109.5423  |
| O <sub>2</sub>  | -150.277                       | -150.248(5)                | -150.3268  |

 Obtained modest improvements over Umrigar's published optimized parameters for these wfns (except Oxygen)

Work in progress... more sophisticated wfns / larger systems



#### **Released node with PIGS**





Release nodes for fixed number of time slices at path endpoints

Evaluate local energy at path endpoints. Paths on **–** side contribute with a negative weight.







**Physical Sciences** 

24

#### Released node Li<sub>2</sub>

![](_page_24_Picture_1.jpeg)

![](_page_24_Figure_2.jpeg)

### **Released node C<sub>2</sub>**

![](_page_25_Picture_1.jpeg)

![](_page_25_Figure_2.jpeg)

#### **Released node C<sub>2</sub>**

![](_page_26_Picture_1.jpeg)

![](_page_26_Figure_2.jpeg)

27

#### **Future Directions**

![](_page_27_Picture_1.jpeg)

- Extend optimization tests to
  - larger systems
  - More sophisticated trial functions
- Evaluation of single particle density matrix and other off-diagonal observables
- Correlated path released node

![](_page_27_Picture_7.jpeg)

#### **Correlated paths**

![](_page_28_Picture_1.jpeg)

![](_page_28_Figure_2.jpeg)

## Off diagonal observables (for fermionic systems)

![](_page_29_Picture_1.jpeg)

- E.g. Single particle density matrix
- Exchange-correlation hole.

![](_page_29_Figure_4.jpeg)

![](_page_30_Picture_0.jpeg)

![](_page_30_Picture_1.jpeg)

#### **Other adventures with PIGS: Correlated Sampling**

![](_page_31_Picture_1.jpeg)

Correlated path sampling  

$$P(Y) = \psi_{\nu}^{a}(\mathbf{R}_{0})\psi_{\nu}^{b}(\mathbf{R}_{M}) \prod_{\ell=0}^{M/2-1} G^{a}(\mathbf{R}_{\ell+1}, \mathbf{R}_{\ell}, \Delta\tau)G^{b}(\mathbf{R}_{M/2+\ell+1}, \mathbf{R}_{M/2+\ell}, \Delta\tau)$$

$$E_{a} - E_{b} = \frac{\langle \phi_{a} | \hat{H}_{a} - \hat{H}_{b} | \phi_{b} \rangle}{\langle \phi_{a} | \phi_{b} \rangle} = \frac{\langle \phi_{a} | \hat{V}_{a} - \hat{V}_{b} | \phi_{b} \rangle}{\langle \phi_{a} | \phi_{b} \rangle}$$

 $E_a - E_b \approx \frac{1}{L} \sum_{k=0}^{L} V_a(R_{M/2}^{(k)}) - V_b(R_{M/2}^{(k)})$ 

#### **Acknowledgments**

![](_page_32_Picture_1.jpeg)

Randolph Q Hood(LLNL)Eric Schwegler(LLNL)

![](_page_32_Picture_3.jpeg)

![](_page_32_Picture_4.jpeg)

**Physical Sciences** 

CASINO

The Cambridge quantum Monte Carlo code

![](_page_32_Picture_8.jpeg)

![](_page_32_Picture_9.jpeg)

![](_page_32_Picture_10.jpeg)

![](_page_33_Picture_1.jpeg)

## Test I :: Noble atom solid (FCC Argon)

- Continuum system
  - QMC 'always' Gamma point so large calculation required to minimize finite size errors.
- Well characterized experimentally

## Test II :: Benzene dimers

- Chemically 'non-trivial' system
- Well studied with high level Q-Chem methods

![](_page_33_Picture_9.jpeg)

![](_page_34_Picture_1.jpeg)

Goal is to treat very large scale (hundreds of atoms) which imposes constraints on our approach

- Use Pseudopotentials
- Use a simple Slater-Jastrow variational ansatz
  - Orbitals from DFT
  - Minimal number of determinants
  - Backflow if necessary

Do errors in the nodal surface have a strong influence on relatively weak dispersive intermolecular interactions?

## Validating our approach Example I :: FCC Solid Argon

![](_page_35_Picture_1.jpeg)

128 argon atom supercell(1024 electrons)

•Single determinant

No backflow

![](_page_35_Figure_5.jpeg)

## Validating our approach Example I :: FCC Solid Argon

![](_page_36_Picture_1.jpeg)

#### Putting the QMC result into context...

![](_page_36_Figure_3.jpeg)

![](_page_36_Picture_5.jpeg)

#### **Example II :: Benzene dimers**

![](_page_37_Picture_1.jpeg)

T-shaped

![](_page_37_Figure_3.jpeg)

|         | LDA +<br>DMC | LDA+BF+<br>DMC | HF+MP2+<br>DMC | CCSD(T) |
|---------|--------------|----------------|----------------|---------|
| d (Å)   | 4.9          | 4.9            | 4.9            | 5.0     |
| E (meV) | 110          | 116            | 130            | 119     |

CCSD(T) results : *J. Phys. Chem. A* 110, 10656 (2006).

HF-MP2-DMC results : J. Chem. Phys. 123, 184106 (2005).

#### Parallel displaced

![](_page_37_Picture_9.jpeg)

|         | LDA +<br>DMC | LDA+BF+<br>DMC | HF+MP2+<br>DMC | CCSD(T) |
|---------|--------------|----------------|----------------|---------|
| d (Å)   | 3.8          | 3.8            | 3.41           | 3.9     |
| E (meV) | 130          | 128            | 155            | 121     |

![](_page_37_Picture_12.jpeg)

#### Example II :: Benzene 'pancake' dimer

Physical Contract Sciences

![](_page_38_Figure_2.jpeg)

Using different DFT functionals to generate the single particle orbital inputs to QMC allows for rough estimation of systematic error.

![](_page_39_Picture_1.jpeg)

- For 'reasonable' separations, fixed node error appears to cancel for purely dispersive interactions
- Nonlocal pseudopotentials give error from locality approximation of up to ~100 meV -- difficult (but not impossible) to resolve at the VMC level
- A good variational ansatz (backflow) can still be problematic to optimize and does not strongly influence Van der walls binding energies
- -- Iterative DMC-VMC required

![](_page_39_Picture_6.jpeg)

## Some preliminary results for H2 on aromatic carbon

![](_page_40_Picture_1.jpeg)

![](_page_40_Figure_2.jpeg)

Can we say anything about scaling and non-additivity of vdW on these adsorbents?

![](_page_40_Picture_4.jpeg)

#### H2 on benzene

Physical Contract Physical Sciences

![](_page_41_Figure_2.jpeg)

#### H2 on coronene

300QMC LDA 200100 $E \;(meV)$  $\bigcirc$ 0 -100 -20022.53 3.54.555.546 a (Angstrom)

Single h2 binding energy is ~200 ± 12 meV

![](_page_42_Picture_5.jpeg)

## H2 on planar Graphene (1/3 filling )

![](_page_43_Picture_1.jpeg)

![](_page_43_Figure_2.jpeg)

...Work in progress... multiple determinants required

![](_page_43_Picture_5.jpeg)

#### Scaling of H2 binding on planar aromatic carbon

Physical OMA

![](_page_44_Figure_2.jpeg)

![](_page_45_Picture_1.jpeg)

Simple picture of competing electron correlation effects

![](_page_45_Figure_3.jpeg)

ideal substrate binding

![](_page_45_Picture_5.jpeg)

![](_page_45_Figure_6.jpeg)

ideal H2 - H2 binding

![](_page_45_Picture_8.jpeg)

![](_page_46_Figure_1.jpeg)

![](_page_46_Figure_2.jpeg)

![](_page_47_Picture_1.jpeg)

- Carbon based materials offer many different possibilities for tuning binding energetics of H2 e.g.
  - Curvature
  - Damage
  - Doping
  - Decorating
  - Charging

![](_page_47_Picture_8.jpeg)

## In progress / future directions (methodology)

- Improved 'nodal-centric' variational metrics
- Path based variational optimization (inline DMC)
- Relax some of our constraints
  - E.g. use pfaffians for variational ansatz

![](_page_48_Picture_5.jpeg)

**Physical Sciences** 

![](_page_49_Picture_1.jpeg)

In each iteration, an ensemble of many-body coordinates  ${f R}=\{{f r}_1,{f r}_2,\ldots,{f r}_N\}$  are drawn from the probability distribution  $|\psi_T(R)\phi_0(R)|$ 

![](_page_49_Figure_3.jpeg)

Never "see" the full wavefunction only

$$\{\mathbf{R}_1,\mathbf{R}_2,\mathbf{R}_3,\ldots,\mathbf{R}_M\}$$

Collect statistics on observables of interest

$$\langle \psi_T | \hat{\mathcal{O}} | \phi_0 \rangle \approx \frac{1}{M} \sum_i^M \hat{\mathcal{O}}(\mathbf{R}_i)$$

![](_page_49_Picture_8.jpeg)

![](_page_50_Picture_1.jpeg)

Ground state of full many-body Schrödinger equation

$$\hat{H} = -\frac{1}{2} \sum_{i} \nabla_i^2 + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i,\alpha} \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{r}_\alpha|} + \sum_{\alpha < \beta} \frac{Z_\alpha Z_\beta}{|\mathbf{r}_\alpha - \mathbf{r}_\beta|}$$

By starting with a guess  $\Psi_T(\mathbf{R})$  where  $\mathbf{R} = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$ 

And projecting out the exact ground state

$$e^{-\tau \hat{H}} \Psi_T(\mathbf{R}) = e^{-\tau \hat{H}} \sum_i c_i \phi_i(\mathbf{R})$$
$$= \sum_i c_i e^{-\tau E_i} \phi_i(\mathbf{R})$$
$$\lim_{\tau \to \infty} \propto \phi_0(\mathbf{R})$$

![](_page_50_Picture_7.jpeg)

#### **Overview**

![](_page_51_Picture_1.jpeg)

- What is Quantum Monte Carlo
- Why is it good for Van der Waals interactions
- Limitations of the method
- Our approach for large systems
- Example applications (solid argon, benzene dimer)
- Current results for H2 on carbon adsorbants

![](_page_51_Picture_8.jpeg)

#### **QMC Benchmarks for Van der Waals interactions**

Quantitative determination of Van der Waals interactions from first principles is an active area of research

For complex / large systems where experimental results are difficult to obtain, high quality benchmarks are required

Fixed node Diffusion Monte Carlo requires minimal uncontrolled approximations :: (total QMC energy is a variational upper bound and depends on the accuracy of nodal surface)

![](_page_52_Picture_4.jpeg)

![](_page_52_Picture_5.jpeg)