Wave Function Optimization Algorithms in QMCPACK

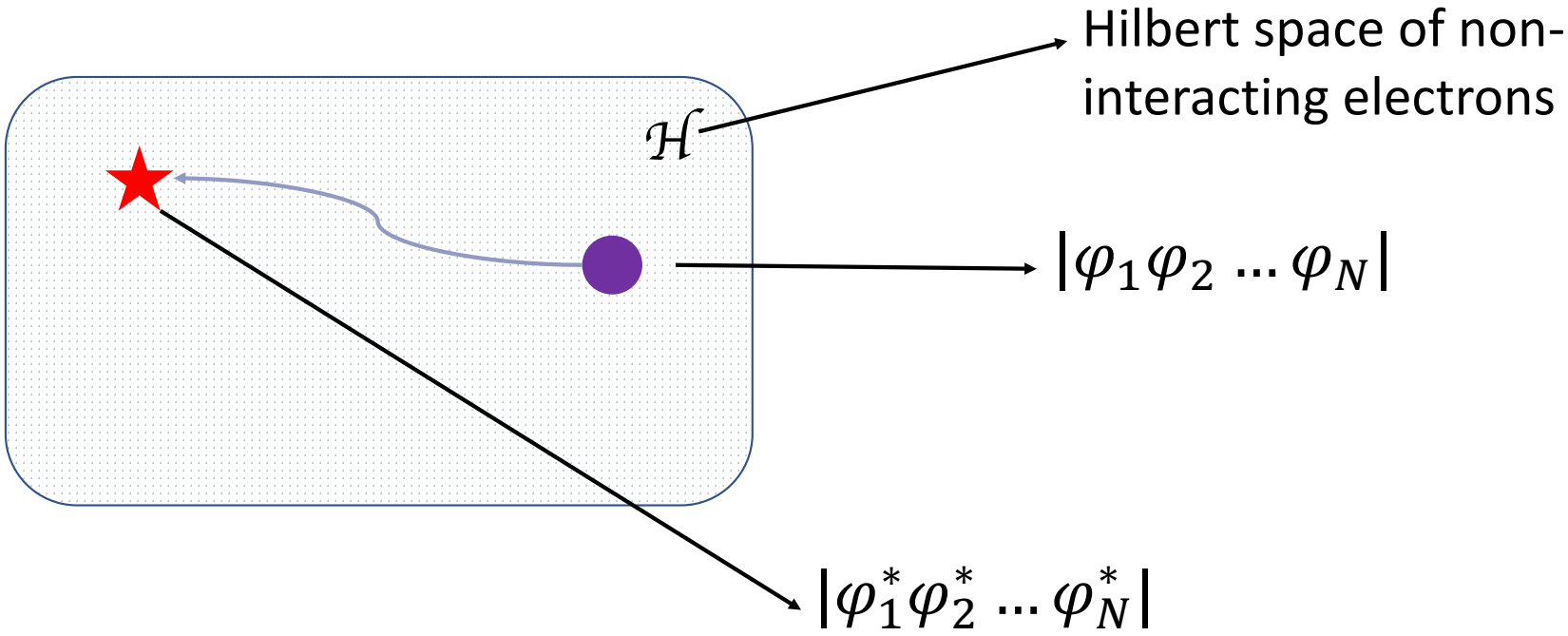
Berkeley
UNIVERSITY OF CALIFORNIA

Luning Zhao

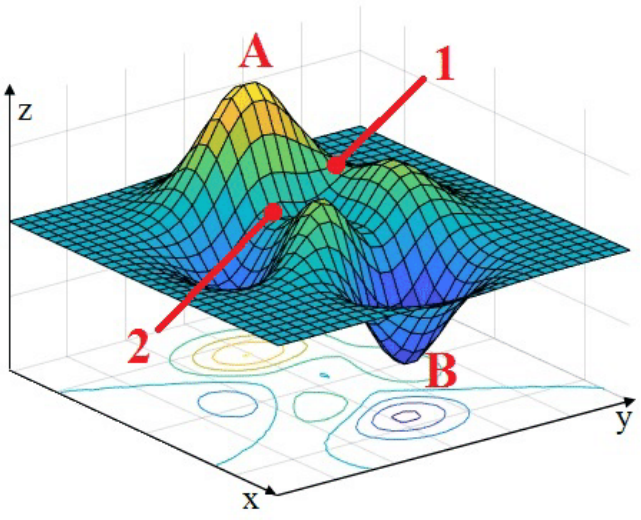
University of California, Berkeley

2019.05.15

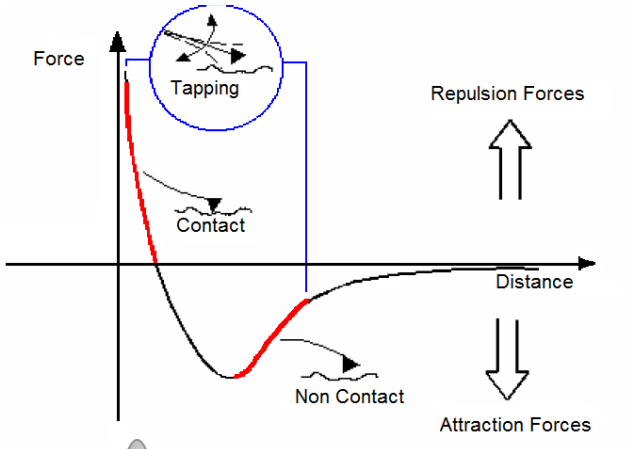
The Motivation of Wave Function Optimization



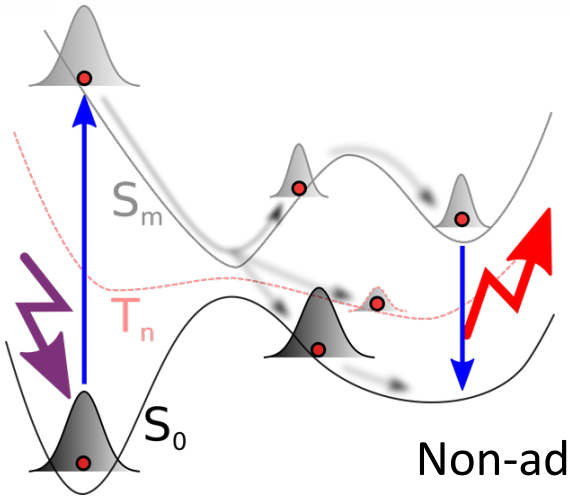
What can we compute with better wave functions?



Potential Energy Surfaces: Barrier Height, Excitation Energy, Reaction Rate



Forces

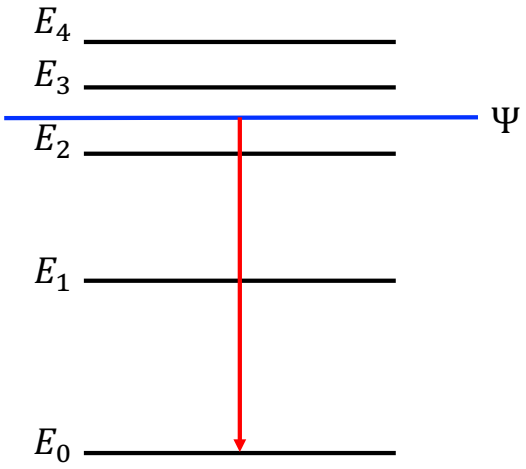


Non-adiabatic Dynamics

Variational Principle For Ground and Excited State

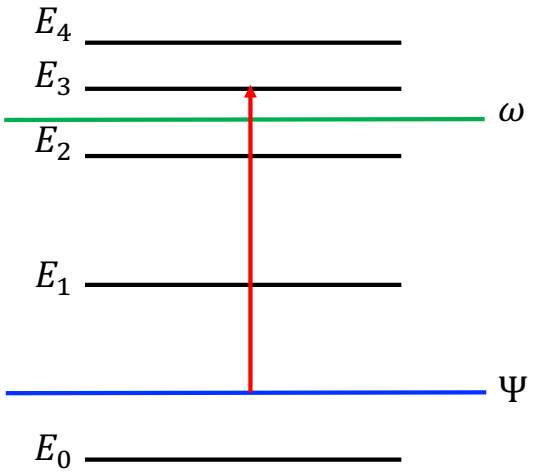
Ground State

$$E[\Psi] = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0$$



Excited State

$$\Omega[\Psi] = \frac{\langle \Psi | \omega - H | \Psi \rangle}{\langle \Psi | (\omega - H)^2 | \Psi \rangle}$$



Optimization Method in QMCPACK: The Linear Method

Wave Function with Parameters $\vec{p}$: $\Psi(\vec{p})$

First Order Taylor Expansion around $\vec{p}$:

$$\Psi(\vec{p}+\Delta\vec{p}) = \Psi(\vec{p}) + \sum_{i=1} \frac{\partial \Psi}{\partial p_i} \Delta p_i = \sum_{i=0} \frac{\partial \Psi}{\partial p_i} \Delta p_i$$

Plug $\Psi(\vec{p}+\Delta\vec{p})$ in $E[\Psi]$ or $\Omega[\Psi]$:

$$E[\Psi(\vec{p}+\Delta\vec{p})] = \frac{\sum_{ij} \left\langle \frac{\partial \Psi}{\partial p_i} \middle| H \middle| \frac{\partial \Psi}{\partial p_j} \right\rangle}{\sum_{ij} \left\langle \frac{\partial \Psi}{\partial p_i} \middle| \frac{\partial \Psi}{\partial p_j} \right\rangle} \qquad \Omega[\Psi(\vec{p}+\Delta\vec{p})] = \frac{\sum_{ij} \left\langle \frac{\partial \Psi}{\partial p_i} \middle| \omega - H \middle| \frac{\partial \Psi}{\partial p_j} \right\rangle}{\sum_{ij} \left\langle \frac{\partial \Psi}{\partial p_i} \middle| (\omega - H)^2 \middle| \frac{\partial \Psi}{\partial p_j} \right\rangle}$$

Optimization Method in QMCPACK: The Linear Method

Taking derivatives w.r.t. $\Delta\vec{p}$ and set the derivatives to 0:

$$\sum_j \left\langle \frac{\partial\Psi}{\partial p_i} \middle| H \middle| \frac{\partial\Psi}{\partial p_j} \right\rangle \Delta p_j = E \sum_j \left\langle \frac{\partial\Psi}{\partial p_i} \middle| \frac{\partial\Psi}{\partial p_j} \right\rangle \Delta p_j$$

$$\sum_j \left\langle \frac{\partial\Psi}{\partial p_i} \middle| \omega - H \middle| \frac{\partial\Psi}{\partial p_j} \right\rangle \Delta p_j = \Omega \sum_j \left\langle \frac{\partial\Psi}{\partial p_i} \middle| (\omega - H)^2 \middle| \frac{\partial\Psi}{\partial p_j} \right\rangle \Delta p_j$$

Both equations are in the form of a generalized eigenvalue problem:

$$\tilde{H}\Delta\vec{p} = \lambda\tilde{S}\Delta\vec{p}$$

Optimization Method in QMCPACK: The Linear Method

Perform Inverse Taylor expansion:

$$p_i = p_i + \Delta p_i$$

1. Build $\tilde{H}$ and $\tilde{S}$ matrices.
2. Solve the Generalized Eigenvalue Problem.
3. Update Wave Function Parameters.
4. If not converged, go to step 1.

One Step in LM

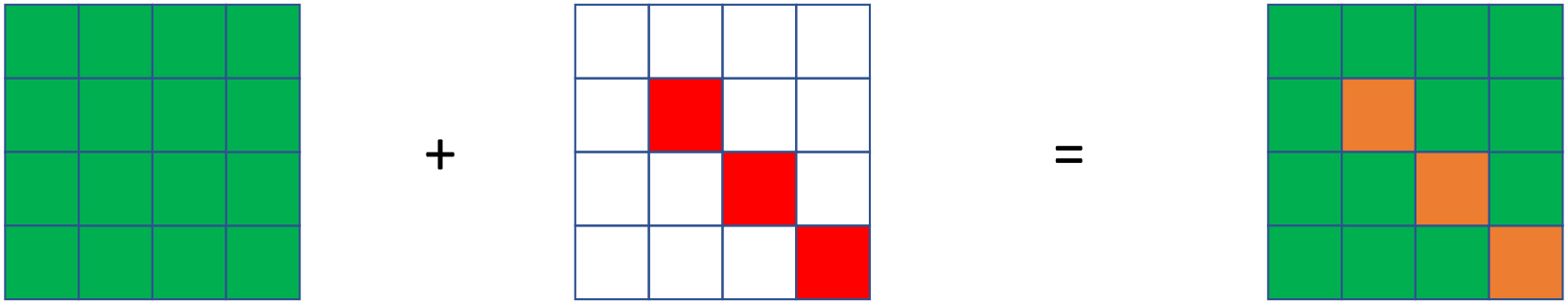
Stabilizing Linear Method: Step Size Control

Modify the Hamiltonian:

$$\widetilde{H} = \widetilde{H} + \alpha A + \beta B$$

Diagonal Shift Matrix A:

$$A_{xy} = \delta_{xy}(1 - \delta_{x0})$$



Stabilizing Linear Method: Step Size Control

Projection Matrix Q: Project the current wave function out

$$Q_{xy} = \delta_{xy} - \delta_{x0}(1 - \delta_{y0})S_{0y}$$

Shift Matrix T: Penalize directions with larger norms

$$T_{xy} = (1 - \delta_{x0}\delta_{y0})[Q^{\dagger}SQ]_{xy}$$

Shift Matrix B: Transform back to the original basis

$$B = \left(Q^{\dagger}\right)^{-1} T Q^{-1}$$

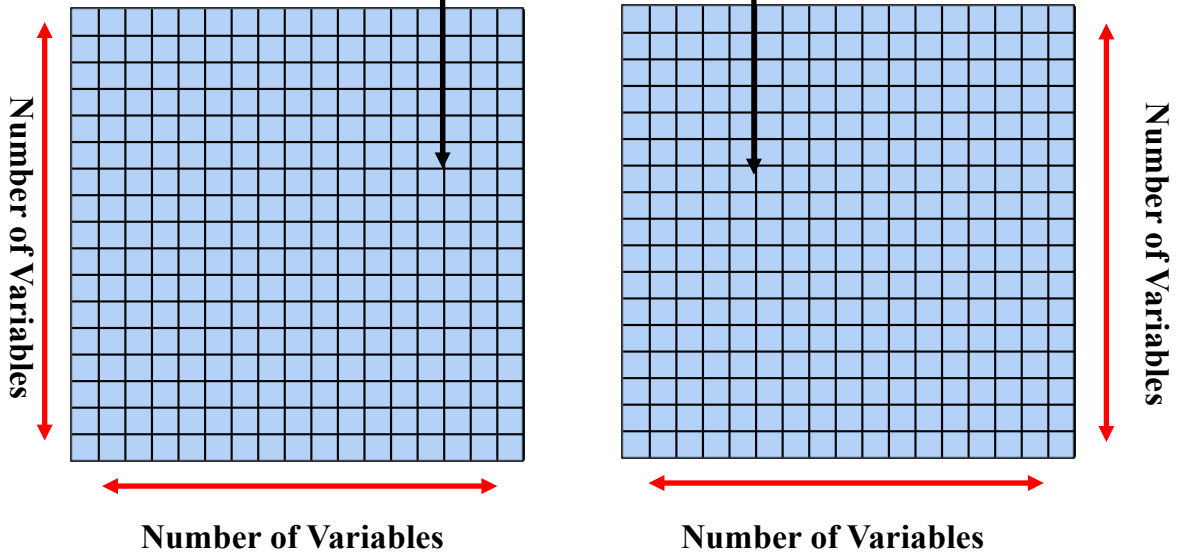
How to Choose the Shift Value?

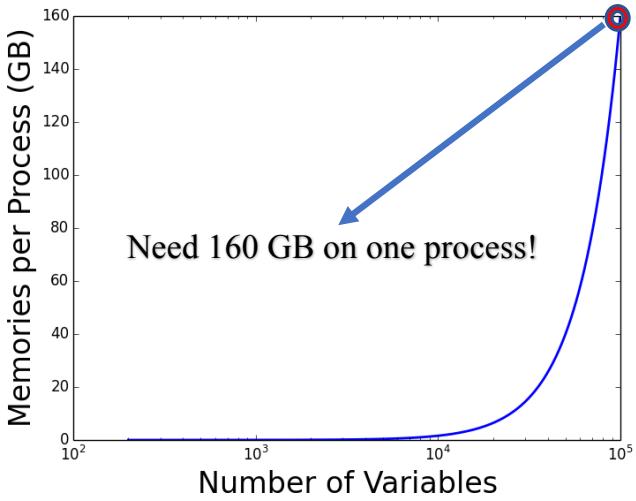
adaptive:

1. Set initial α and β
2. Small shifts: $\alpha/4$, $\beta/4$, Medium Shifts: $\alpha/4$, $\beta/4$, Large Shifts: 4α, 4β
3. Solve the generalized eigenvalue problem for these three sets of shifts.
4. Use correlated sampling to compare the updated wave function based on three shifts
5. Small shifts wins: lower the shifts Medium shifts wins: no change to the shifts Large shifts wins: increase the shifts

Memory Bottleneck of LM

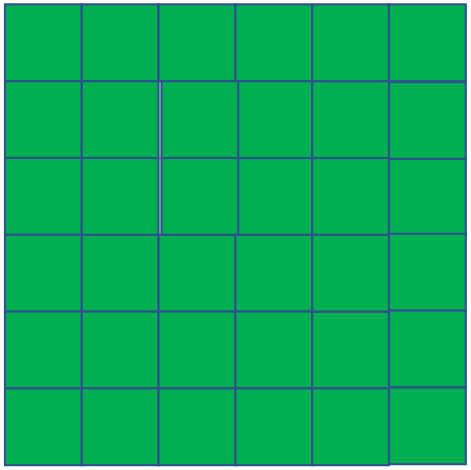
$$\sum_{j=1}^{Nv} \tilde{H}_{ij}\Delta p_j = E \sum_{j=1}^{Nv} \tilde{S}_{ij}\Delta p_j$$

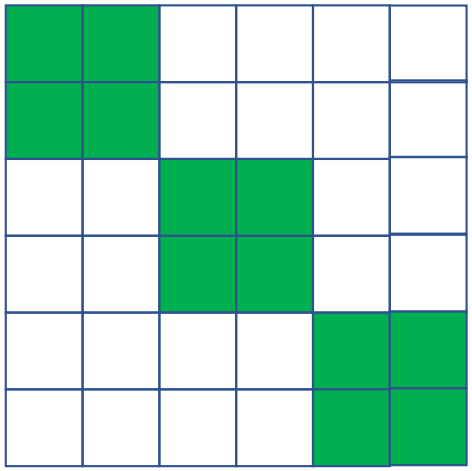




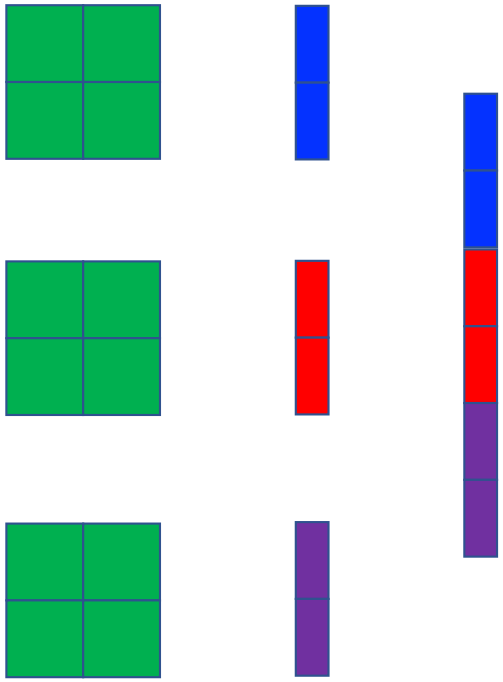
A few GB of memory

Blocked LM: A Crude Idea

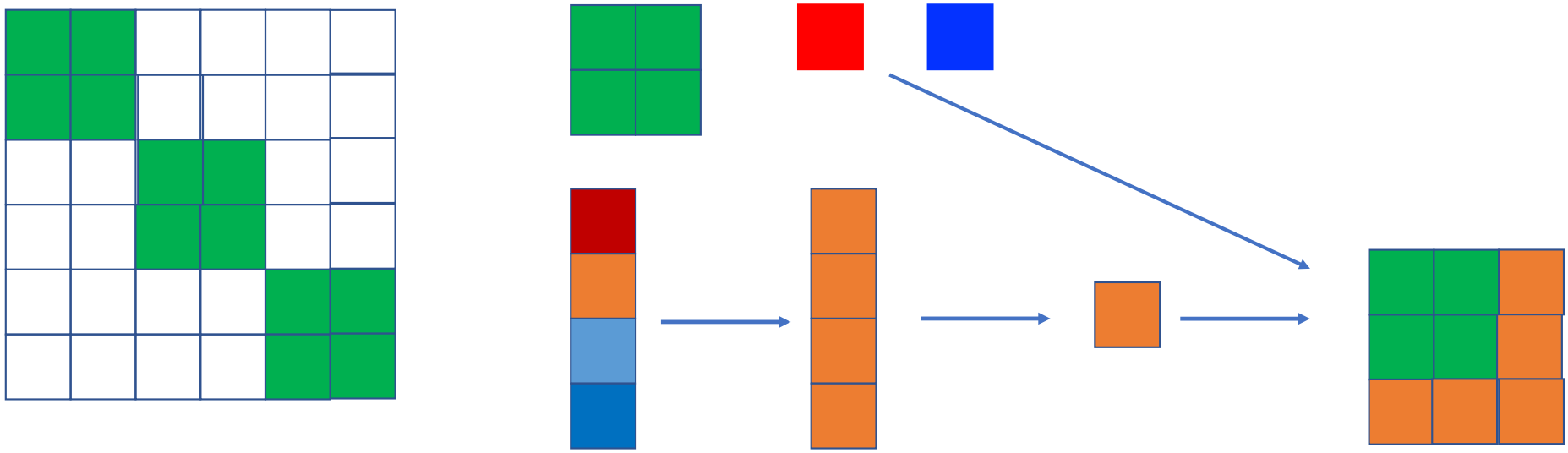




Divide Variables into Blocks

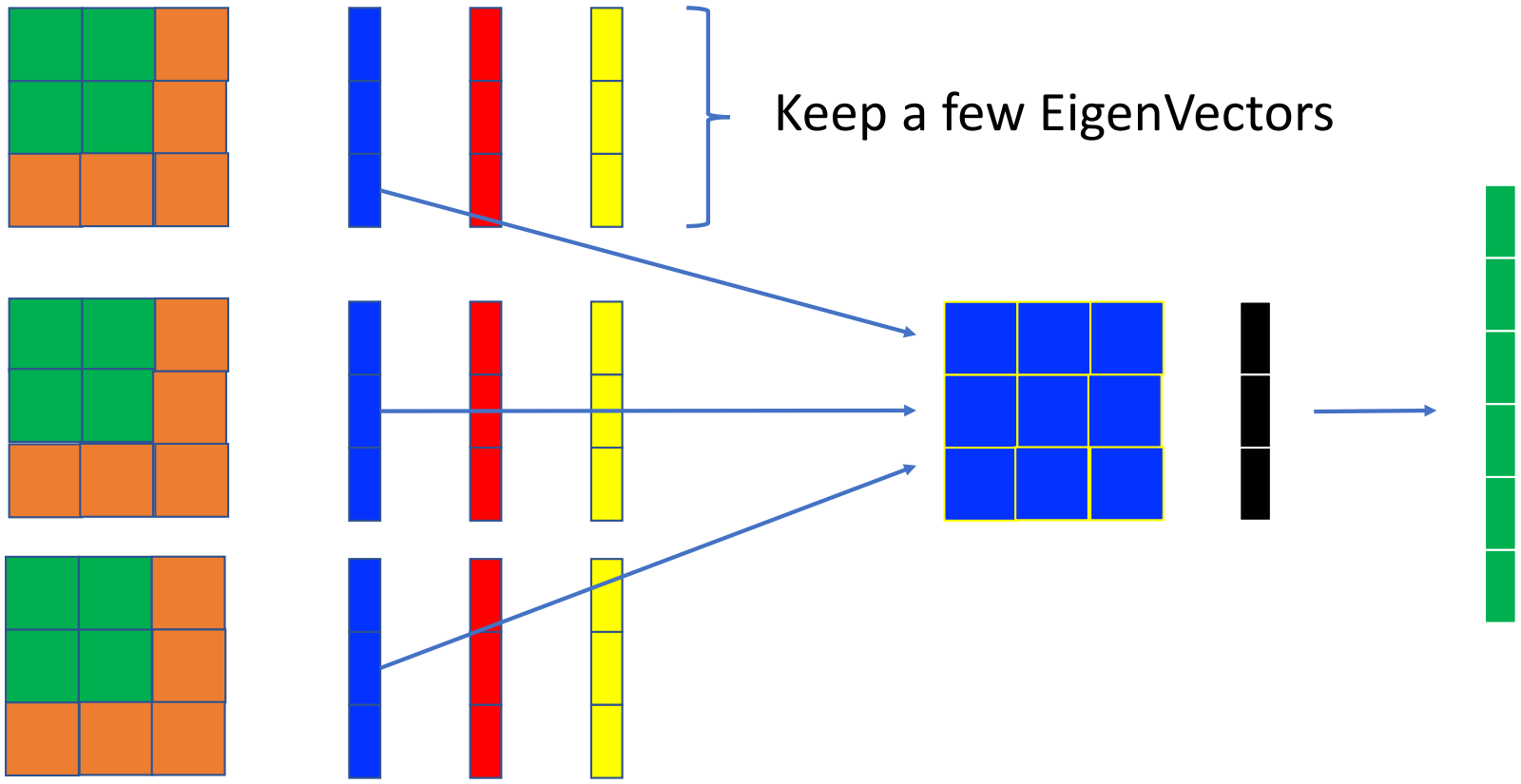


Blocked LM: Account for Inter-Block Coupling



Old Update Directions

Blocked LM: Account for Inter-Block Coupling



How to use LM/BLM in QMCPACK

```
<loop max="10">
    <qmc method="linear" move="pbyp" checkpoint="-1" gpu="no">
      <parameter name="MinMethod"> adaptive </parameter>
      <parameter name="targetExcited"> yes </parameter>
      <parameter name="omega"> -17.779  </parameter>
       …….
    </qmc>
</loop>
```

How many steps you want to run

Specify linear method

Parameters of LM in QMCPACK

1. targetExcited: yes/no: whether to do target function/energy minimization
2. omega: real number: Energy shift used to target different excited states
3. shift_i: 1.0: value of the identity shift
4. shift_s: 1.0: value of the overlap shift
5. max_param_change: 0.3: the maximum allowed parameter update
6. chase_lowest: yes/no: whether to chase the lowest eigenvector in iterative solver
7. chase_closest: yes/no: whether to chase the eigenvector that is closest to the initial guess

Parameters of BLM in QMCPACK

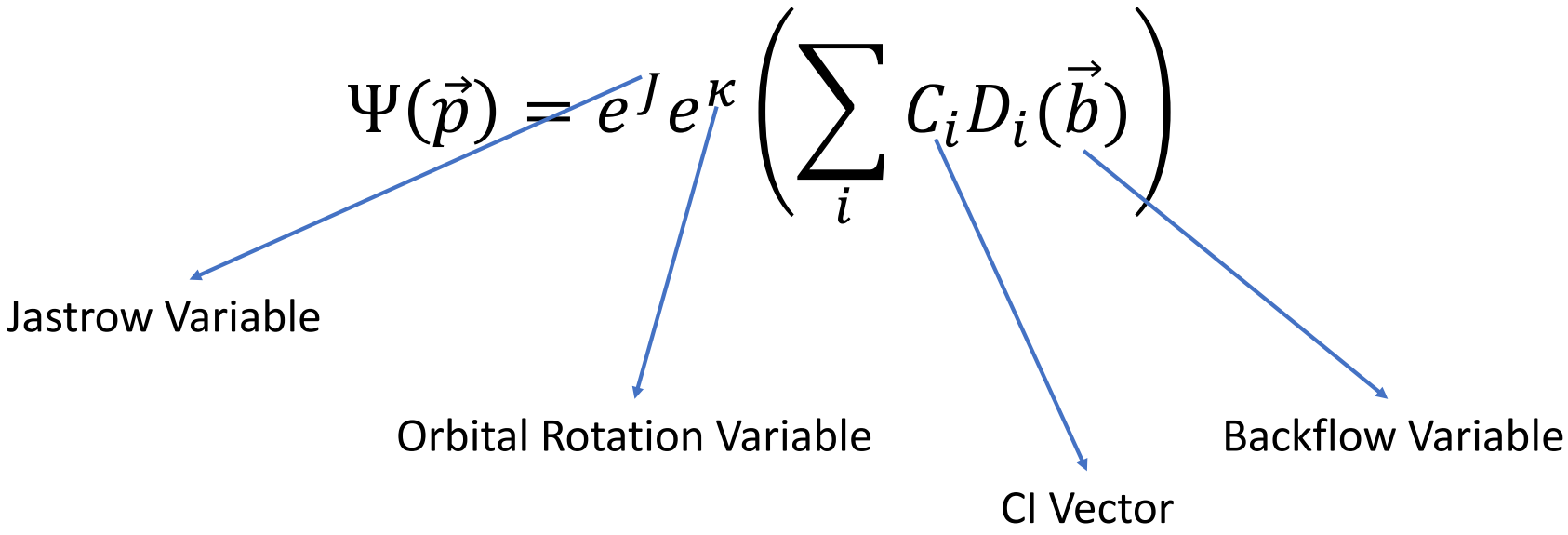
1. block_lm: yes/no: whether to switch to BLM
2. nblocks: integer: number of blocks in BLM
3. nkept: integer: number of eigenvectors to keep per block in BLM
4. nolds: integer: number of old update vectors used in BLM

Rule of Thumb: BLM should be turned on when the number of parameters > 5000.

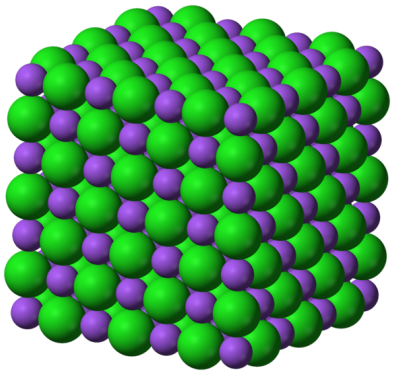
Usually a few hundred blocks and a handful of nkept and nolds are sufficient to balance memory and accuracy.

Type of Parameters that LM/BLM can optimize

Everything!



Upcoming Optimization Features in QMCPACK



Complex Wave Function Parameters: Solid State

$$\Psi_G^2(R) = |\Psi(R)|^2 + \varepsilon \sum_i \Psi_i^2(R)$$

$$\Psi_G^2(R) = |\Psi(R)|^2 + \varepsilon |\nabla^2 \Psi(R)|$$

Guiding Function other than $\Psi(R)$

Examples in the Hands on Session

1. Energy Optimization of all-electron H_4 with Jastrow Multi-Slater Wave Function and Orbital Optimization.
2. Excited State Optimization of H_2O with Jastrow Multi-Slater Wave Function and Orbital Optimization.
3. Blocked LM Optimization of C_2.

Acknowledgement

- Eric Neuscamman
- Graduate students and postdocs in Neuscamman group
- Collaborators in CPSFM

QMCPACK

CPSFM