Introduction
Multi-reference methods are designed to treat systems with strong correlations. These calculations are often very expensive in both memory and time. We aim to provide users efficient computational chemistry tools from two different ways: to develop new theoretical methods to reduce the scale of the problem, and to take advantage of the available parallel computing resources.

Implementing GASSCF in NWChem
The MCSCF we have integrated into NWChem is based on General Active Space concept (GASSCF) which is highly flexible. The parallel version has been successfully implemented in NWChem.

Direct CI is the most time consuming part in an MCSCF iteration.

$$\sigma(I_a, I_b) = \sum_{\alpha \beta} \langle \alpha | I_a \alpha \rangle \langle \beta | I_b \beta \rangle \langle \alpha | C_{\alpha \beta} | \beta \rangle,$$

(\(\sigma\) vector)

$$|0\rangle = \sum_{I_a I_b} C(I_a, I_b) |\alpha(I_a) \rangle \langle \beta(I_b) |,$$

(Reference wf)

$$\hat{H} = \sum_{\alpha \beta} \hbar \omega_{\alpha \beta} + \frac{1}{2} \sum_{\alpha \beta} \sum_{i<j} (\hat{c}_i^\dagger \hat{c}_j \hat{c}_j^\dagger \hat{c}_i - \delta_{\alpha \beta} \hat{E}_\alpha),$$

(Hamiltonian)

Two \(\sigma\) contribution is split to 3 parts: 2 \(\alpha\) excitations, 2 \(\beta\) excitations, and mixed excitations.

Orbital optimization uses true 2nd order Newton scheme.

Details on integrating serial LUCIA MCSCF/CI code into NWChem
• Overhaul of allocation to NWChem Memory Allocator infrastructure
• CI vector, \(\sigma\) vector, MO integrals are stored in global arrays that are blocked and spread over processors;
• Multiplet CI and \(\sigma\) vectors are stored to disk in parallel; Future algorithm to store as many in memory as possible
• Parallelization:
  • Distribute block of C to all processors
  • Each processor computes a partial contribution to the \(\sigma\) vector (where code spend 90% of the time)
  • Global reduction of \(\sigma\) vector

The parallel infrastructure of NWChem
Built on the GA toolkits. To tackle the large-scale problem where the memory of a single node are not sufficient, distribution is the only way. And it is desirable to be able to allocate the memory of the whole parallel computer to store large arrays.

Parallelization scaling performance of MCSCF in NWChem
Timing details for most time consuming segments

<table>
<thead>
<tr>
<th>Segment</th>
<th>4 cores</th>
<th>16 cores</th>
<th>32 cores</th>
<th>64 cores</th>
<th>96 cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building o vector</td>
<td>13420</td>
<td>3477</td>
<td>2139</td>
<td>1350</td>
<td>1109</td>
</tr>
<tr>
<td>2-body density matrix</td>
<td>1497</td>
<td>479</td>
<td>306</td>
<td>218</td>
<td>195</td>
</tr>
<tr>
<td>Creation/annihilation mapping</td>
<td>450</td>
<td>114</td>
<td>60</td>
<td>32</td>
<td>24</td>
</tr>
<tr>
<td>Obtain SD blocks</td>
<td>309</td>
<td>43</td>
<td>56</td>
<td>37</td>
<td>44</td>
</tr>
<tr>
<td>Data movement (sigma/CI)</td>
<td>881</td>
<td>991</td>
<td>1218</td>
<td>1180</td>
<td>1157</td>
</tr>
</tbody>
</table>

Conclusion and future work
• We have implemented a parallel version of MCSCF based on active space (GASSCF) in NWChem
• Good scaling for parallel build of \(\sigma\) vector; data movement is the next bottleneck to improve upon
• For the current CASSCF limit (18, 18), each MCSCF iteration can be run in under an 1 hour, comparing to 15+ hours in serial
• Able to tackle bigger science problems not possible before

Next step
• Implement a scalable version of the SplitGAS approach within the NWCHEM GASSCF framework to approximate FCI

Challenging the limit: Free base porphin
Excited states in porphin and metal porphyrines are important for photochemistry and biochemistry. The excited states are difficult to deal with given the large conjugation system of 26 electrons in 24 orbitals. Among them 18 electrons in 20 orbitals are important.

There are 705280360 SDs for the singlet symmetric state.

“PLACEHOLDER, RESULTS AVAILABLE ON FINAL POSTER”

Reference

Acknowledgement
This work is supported by DOE SciDAC program Grant No. DE-SC0008666.