New approaches to fast methods for electron correlation in molecules.

MARTIN HEAD-GORDON

Department of Chemistry,
University of California, Berkeley, and,
Chemical Sciences Division,
Lawrence Berkeley National Laboratory
Outline

1. Background and motivation.
   - A tiny bit of history
   - The dominant paradigm & its limitations

2. Fast methods to treat strong and weak correlations.

3. Fast algorithms to implement the new methods

4. Some examples of applications.
“Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry. If mathematical analysis should ever hold a prominent place in chemistry—an aberration which is happily almost impossible— it would occasion a rapid and widespread degeneration of that science.”

Auguste Comte, 1830.
“...in the Schrodinger equation we very nearly have the mathematical foundation for the solution of the whole problem of atomic and molecular structure”

but...

“... the problem of the many bodies contained in the atom and the molecule cannot be completely solved without a great further development in mathematical technique.”

"In conclusion, I would like to emphasize my belief that the era of computing chemists, when hundreds if not thousands of chemists will go to the computing machine instead of the laboratory, for increasingly many facets of chemical information, is already at hand. There is only one obstacle, namely, that someone must pay for the computing time."

Complexity

• Many-body Schrödinger equation is a partial differential equation in $3n$ unknowns— the positions of the electrons.
  • Exact (brute force) solution will scale approximately exponentially with the number of electrons.

• When done in a given basis of 1-electron expansion functions, this is “full configuration interaction” (FCI).
  • Largest FCI’s involve many-body expansions containing billions of terms, for molecules with 2 or 3 of atoms (!).

• Approximations are imperative. Accuracy vs feasibility.
Branches of the family tree

- Wavefunction-based electronic structure theory:
  - Minimize the energy by varying the wavefunction
  - Tremendously complicated unknown function:
    \[ \Psi = \Psi(r_1, r_2, \ldots, r_n) \]
  - Modeling the wavefunction yields “model chemistries”

- Density functional theory
  - The unknown is very simple: \( \rho = \rho(\zeta) \)
  - Hohenberg-Kohn theorem guarantees that: \( E = E\{\rho(\zeta)\} \)
  - True functional is unknown and probably unknowable
  - Modeling the functional gives DFT model chemistries.
Summary of present status of DFT

• The only method in large-scale use for electronic structure studies in condensed matter physics

• The predominant method in use for electronic structure studies of molecules, surfaces, and nanomaterials

• Strengths of present-day DFT:
  – It is the sweet spot between accuracy and feasibility

• Weaknesses of present-day DFT:
  – Not systematically improvable
  – Computational cost is still high relative to empirical methods
  – Inadequate to treat strong and/or non-local correlations
Strong correlations via wavefunctions

- Complete active space (CAS) methods
  - Has all correlations that can be strong (space of bonding and anti-bonding orbitals)
    \[ \Psi = (\hat{T}_1 + \hat{T}_2 + \ldots + \hat{T}_n) \Phi \]
  - Valence-space Schrodinger eq’n -- exponential complexity.
  - Put back remaining correlations via perturbation theory

- Coupled cluster methods
  - Approach the exact wavefunction via a correlation factor
    \[ \Psi = \exp(\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots) \Phi \]
  - Systematic (brute force) approach with rapidly increasing costs \((N^6, N^8, N^{10} \ldots)\) as a function of truncation level
Outline

1. Background and motivation.

2. **Fast methods to treat strong and weak correlations.**
   - Formulating low-scaling methods
   - PP, IP, PP(2) methods

3. Fast algorithms to implement the new methods

4. Some examples of applications.
$N_2$ dissociation (Troy Van Voorhis)

- **Standard coupled cluster Doubles (non-variational)**
- All excitations included
- Only double excitations

How important are quadruples and hextuples?
Double substitutions describe pair correlations

- Correlated fluctuations of 2 electrons
- In general, they are quartic in number.
  - Empty levels
  - Occupied levels

- Correlations decay (somewhat) rapidly with separation.
  - Exponentially between functions of the same electron
  - Algebraically ($R^{-3}$) between the 2 electrons (→dispersion)
Valence space electron correlation

• Divide the correlation problem into two parts:

• “High energy, short wavelength” dynamic correlations:
  • Associated with atomic-like correlations
  • Dynamic correlation is important for quantitative prediction of reaction energies. Also dispersion forces.

• “Low energy, long-wavelength” static correlations
  • From near-degeneracies between orbitals (bond-breaking).
  • Important for highly correlated systems (diradicals)
  • Treat with a limited “perfect pairing” orbital space
  • One correlating orbital, $j^*$, for each occupied orbital, $j$
2 local valence models for pair correlations

(1) Simplest model is perfect pairing
- Only intrapair correlations, only a linear number of variables
  - Hurley, Pople, Lennard-Jones model (1950’s)
  - Popularized by Goddard as “GVB-PP”
  - Coupled cluster version first explored by Cullen (1996)
- Exact for the case of 1 electron pair!

\[ |\Psi_{PP}\rangle = \exp(\hat{T}_{PP}^{2}) |\Phi_{0}\rangle \]
\[ \hat{T}_{PP}^{2} |\Phi_{0}\rangle = \sum_{i}^{\text{occupied}} t_{ii}^{i^{*}i^{*}} |\Phi_{ii}^{i^{*}i^{*}}\rangle \]
2 local valence models for pair correlations

• (2) Next model must be *imperfect pairing*!
  – Includes interpair correlations that don’t transfer electrons.
  – A CC analog of GVB-RCI (but size-consistent!)

\[
|\Psi_{IP}\rangle = \exp\left(\hat{T}_2^{PP} + \hat{T}_2^{\text{interpair}}\right)|\Phi_0\rangle
\]

\[
\hat{T}_2^{\text{interpair}} = \sum_{i<j}^{\text{occupied}} t_{ij}^{*j} \hat{E}_i \hat{E}_j + t_{ij}^{*i} \hat{E}_i \hat{E}_j^*, \quad \text{where} \quad \hat{E}_i^* = \hat{a}_j^\dagger \hat{a}_i + \hat{a}_j^\dagger \hat{a}_i^\dagger
\]
Perturbation analysis: old and new

• Old: many-body perturbation theory.
  • Partition Hamiltonian into mean field + fluctuation parts
  • Expand in powers of the fluctuation potential
  • Lowest order term missing in CCSD is the (T) correction

• New: similarity-transformed perturbation theory.
  • PP is an e’value of a similarity-transformed Hamiltonian in the “strong pair” space
    \[ \Hbar = e^{-T} H e^T = (H e^T)_c \]
  • Partition (“strong pair” and beyond)
  • Expand… leading correction is second order, PP(2):
    \[ E^{(2)} = \langle 0^{(0)} | \Vbar^{(1)} | R^{(1)} \rangle + \langle \Lambda^{(0)} | \Vbar^{(1)} | R^{(1)} \rangle \]
• By construction, the leading correction is 2nd order:

\[ E^{(2)} = \langle 0^{(0)} | \bar{V}^{(1)} | R^{(1)} \rangle + \langle \Lambda^{(0)} | \bar{V}^{(1)} | R^{(1)} \rangle \]

• The first order correction, \( R^{(1)} \), in the q-space,
  (a) 1st term: couples to \( S \oplus D \) of the q-space
    • Contains singles and doubles (doubles are most important)
    • Doubles are the computational bottleneck (5th order)
  (b) 2nd term: couples to \( S \oplus D \oplus T \oplus Q \) of q-space
    • Triples and quadruples correlate excited PP configurations
    • Doubles (and singles) damp the ordinary doubles.
Outline

1. Background and motivation.

2. Fast methods to treat strong and weak correlations.

3. Fast algorithms to implement the new methods

   Making the prefactor small using auxiliary basis sets

4. Some examples of applications.
Efficient algorithms for valence LCC models

- Linear or quadratic number of variables.
- $M^3$ computation. Great improvement over $M^6$ !
- Outer loop is over orbital iterations (until converged).
  - Make $6 \times O$ Coulomb and exchange matrices ($O=\# \text{ of valence pairs}$)... this is the rate-determining step... cubic computation.
  - From these matrices build required 2-electron integrals
  - Solve the amplitude equations (essentially free at present).
  - Form the orbital gradient and update orbitals.

Auxiliary basis expansions

- Early contributions by Whitten, Dunlap, Baerends, Almlof, etc
  - Auxiliary basis or “resolution of the identity” (RI) or “density fitting” (DF)

- Popularized by Ahlrichs and co-workers for DFT, MP2.
  - Demonstrated efficiency
  - Developed standardized auxiliary basis sets: 3-4 times the AO basis size.

- Replace 4-center integrals by (inexact) expansions:
  \[
  \langle ia|jb \rangle \approx \sum_{KL} C^K_{ia} \langle K|L \rangle C^L_{jb} = \sum_{K} B^K_{ia} B^K_{jb}
  \]

- Coefficients C minimize the Coulomb deviation of the fit:
  \[
  C^K_{ia} = \sum_{L} \langle ia|L \rangle \langle L|K \rangle^{-1}
  \]
  \[
  B^K_{ia} = \sum_{L} \langle ia|L \rangle \langle L|K \rangle^{-1/2}
  \]
Auxiliary basis algorithms for valence LCC models

1*. Form \((L | M)^{-1/2}\) \((X^3)\)

2*. Form \((\mu \nu | M)\) \((\text{NFP } X)\)

3*. Contract: \(B_{\mu \nu}^L = \sum_M (\mu \nu | M)(M | L)^{-1/2}\) \((\text{NFP } X^2)\)

4. Contract: \(B_{\mu[i,i^*]}^L = \sum_{\mu} B_{\mu \nu}^L C_{\mu[i,i^*]}\) \((\text{NFP } X \circ)\)

5. Contract: \(B_{[ii,i^* i^*]}^L = \sum_{\nu} B_{[ii,i^*]}^L C_{\nu[i,i^*]}\) \((X \cdot N \circ)\)

6. Contract: \(K_{\mu \nu}^{[ii,i^* i^*]} = \sum_L B_{\mu [ii^*]}^L B_{(i^* i^*)}^L\) \((X \cdot N^2 \circ)\)

7. Contract: \(J_{\mu \nu}^{[ii,i^* i^*]} = \sum_L B_{\mu \nu}^L B_{[ii,i^*]}^L\) \((\text{NFP } X \circ)\)

Alex Sodt, Greg Beran, MHG (to be published)
Timings as a function of system size (Alex Sodt)

cc-pVDZ basis
linear alkanes

2 Ghz IBM 970fx
(Apple Xserve)
Use of auxiliary basis expansion reduces the scaling by one power of system size in the “small molecule” regime.

Additionally it permits formulation as matrix multiplies.
  • Much higher efficiency.

In practice, speedups approach an order of magnitude!

<table>
<thead>
<tr>
<th>Chain length</th>
<th>RI-PP CPU (s)</th>
<th>RI-IP CPU (s)</th>
<th>PP CPU (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8.0</td>
<td>8.8</td>
<td>25.0</td>
</tr>
<tr>
<td>4</td>
<td>54.1</td>
<td>63.2</td>
<td>281.1</td>
</tr>
<tr>
<td>6</td>
<td>170.4</td>
<td>198.7</td>
<td>1120.1</td>
</tr>
<tr>
<td>8</td>
<td>321.1</td>
<td>402.6</td>
<td>3061.8</td>
</tr>
</tbody>
</table>
Overall timings for the new algorithms

- 2 GHz PowerPC 970fx (Apple Xserve)
- Alex Sodt, Greg Beran, MHG (to be published)
Outline

1. Background and motivation.

2. Fast methods to treat strong correlations.


4. Some examples of applications.
   - The first indefinitely stable singlet diradical?
   - The longest C-C bond?
What is a singlet diradical?

• Like aromaticity, it is not defined by a single quantity

• Experimental perspective:
  – Low singlet-triplet gap; high reactivity
  – Low-lying electronic excitations (e.g. infra-red)

• Wavefunction perspective: presence of strong correlations
  – Small energy gap between occupied and empty energy levels
  – Instead of the HOMO having 2 electrons and the LUMO 0, we approach having 1 electron in each… antiferromagnetically coupled.

• Measure diradical character by $100 \times n(\text{LUMO})\%$
  – $n(\text{LUMO})$ is the occupation number of the (nominal) LUMO
A new stable singlet diradical (?)


- Planar BPBP ring

- Stable at room temp.

- B-B distance of 2.6 Å

- No ESR signal.
Compare against Si(100) surface

- Use Si₉H₁₂ cluster to model a dimer on Si(100).
- 48 valence electrons (24 pairs).
- We optimized the structure to test whether the dimer buckles (it does not).
- HOMO: 1.68 electrons  
  LUMO: 0.32 electrons
How diradicaloid is this molecule? (Yousung Jung)

- Valence active space coupled cluster calculations
  - 94 active electrons (47 pairs), 6-31G* basis

- HOMO: 1.83 electrons
- LUMO: 0.17 electrons

- Stability comes from reduced (17%) diradical character
Phenalenyl dimer: a strong $\pi$ stacking complex

$\text{C}_{13}\text{H}_9^\cdot$: a very stable radical in solution and in the solid state.

$\text{C}_{13}\text{H}_9^\cdot$ dimerizes to form a stable $\pi$ stacking complex

A crystal structure of the dimer of the tri-\text{t-Bu} derivative has been obtained experimentally.
DFT potential surface (B3LYP/6-31G*)

Relative Energy (kcal/mol) vs. Separation (Å)
Phenalenyl dimer potential energy surface

![Graph showing the relative energy vs. R (Å) for Phenalenyl dimer with PP(1), CP-MRMP2, and MRMP2 curves.](image-url)
**Character of the chemical bond**

- Diradicaloid character:
  - LUMO occupation number (ON) at R=3.1 Å: 0.25 electrons
  - i.e. 25% diradicaloid character.

<table>
<thead>
<tr>
<th>R (Å)</th>
<th>ON(LUMO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6</td>
<td>0.15</td>
</tr>
<tr>
<td>3.1</td>
<td>0.25</td>
</tr>
<tr>
<td>3.3</td>
<td>0.32</td>
</tr>
<tr>
<td>5.0</td>
<td>0.95</td>
</tr>
</tbody>
</table>

- Dispersion-assisted 12-center 2-electron bond:
  - Covalency alone cannot overcome repulsions to give net binding.
  - This weak and diradicaloid chemical bond is assisted by dispersion.
  - This interplay between weak covalency and dispersion makes the bond in the phenalenyl dimer distinctive relative to weaker π complexes.
Summary of present status

- Efficient new algorithms developed for treating strong correlations
  - Perfect and imperfect pairing models
  - Combining them with auxiliary basis (RI) methods makes them faster!
  - Scaling with system size is like mean-field theory!!

- Remaining weak correlations treated by perturbation theory
  - Use similarity-transformed perturbation theory from the PP reference
  - Have completed an efficient auxiliary basis algorithm
  - Scaling with system size is like MP2 (pert. theory from mean field)!

- Promising application areas include characterizing diradicaloids
  - Computation can give insight that is very hard to gain experimentally
  - Two examples were discussed. Many more to come!
Ongoing work and future issues

• The chemistry of PP, IP, PP(2)
  – Investigate extent of applicability and chemical limitations
  – Breakdowns may occur for some types of radicals
  – New ideas always emerge from studies of the worst cases!

• Calculations on molecules in the 100+ atom regime
  – Need to exploit “natural sparsity” akin to linear scaling DFT
  – We are currently investigating this problem.
  – Will yield quadratic-scaling (asymptotic linear scaling)
(Very Grateful) Acknowledgements.

<table>
<thead>
<tr>
<th>Name</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greg Beran</td>
<td>PP(2) theory and implementation</td>
</tr>
<tr>
<td>Alex Sodt</td>
<td>Fast algorithms for PP and IP</td>
</tr>
<tr>
<td>Yousung Jung</td>
<td>Auxiliary basis; chemical applications</td>
</tr>
</tbody>
</table>

$: \quad$ DOE SciDAC program

DOE Basic Energy Sciences