EMSL 20YRS

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

U.S. DEPARTMENT OF ENERGY

New computational tools for many-body theories

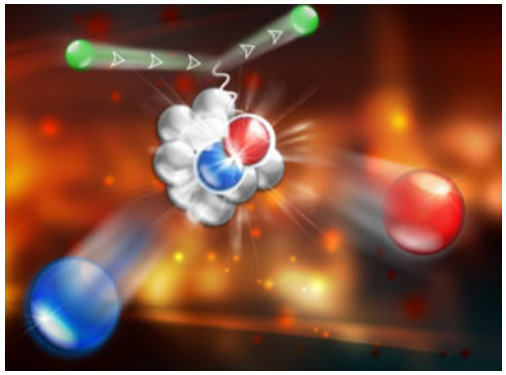
Karol Kowalski & Sriram Krishnamoorthy & Bo Peng

Many-body formulations

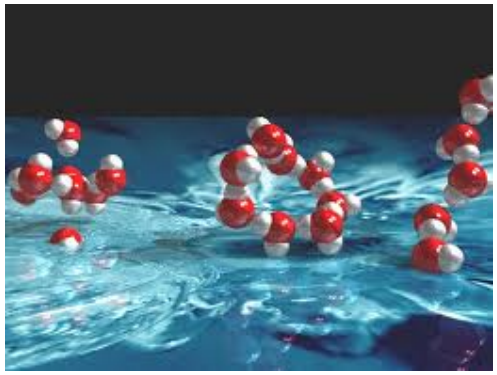
$$\hat{H}\left|\Psi(1,\ldots,N)\right\rangle = E\left|\Psi(1,\ldots,N)\right\rangle$$

Many Particle Systems

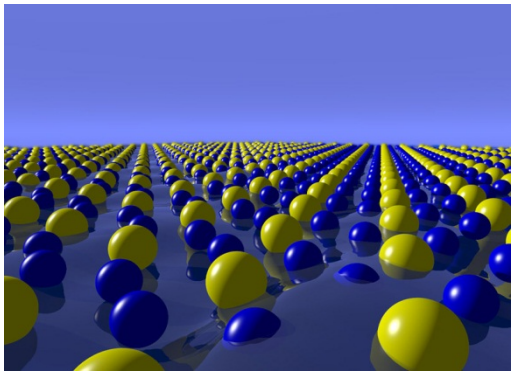
Nuclear Physics



Quantum Chemistry



Solid State Physics



Many-body formulations (cntd.)

- **Remark I**: Certain forms of wave function parametrization are universal across energy/spatial scales in providing accurate energies.
- **Remark II**: All methods that have exact-wave-function-limit are equivalent. It is worthwhile to combine properties of these methods to build more efficient approximations.
- **Remark III:** All many-body formulations should be equally able to "capitalize" on the availability of quantum computers.

Coupled Cluster (CC) formalism – a brief overview

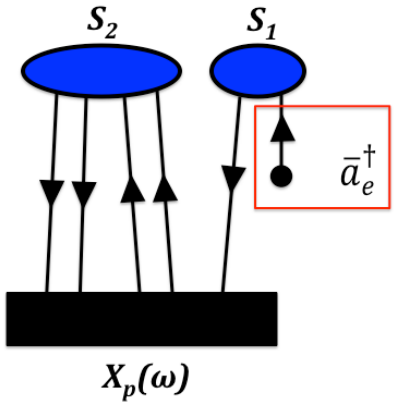
- Exponential wave-function Ansatz

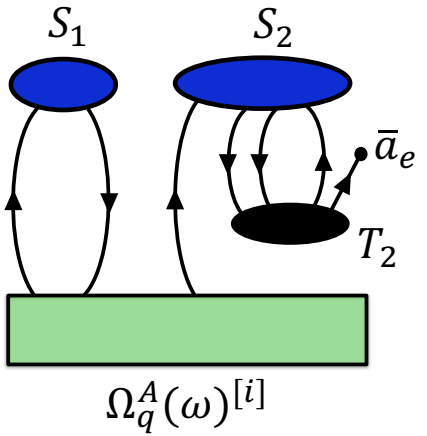
$$|\Psi\rangle = e^{T}|\Phi\rangle$$

- Excited-state extensions (Equation-of-motion CC formalism) $|\Psi_K\rangle = R_K\, e^{T}|\Phi\rangle$

- Multi-reference CC methods: Fock-space/Hilbert-space formulations
  - new algebraic techniques to alleviate intruder state problems
- Green's function CC formulations
  - Analytical ways of calculating CC Green's function/ CC self-energies / **ω**-derivatives
  - Efficient algorithms for representing CC Green's function on entire complex plane
- Things are expensive:
  - CCSD → CCSD(T) → CCSDT → CCSDTQ → …

    N^6 → N^7 → N^8 → N^{10} → …



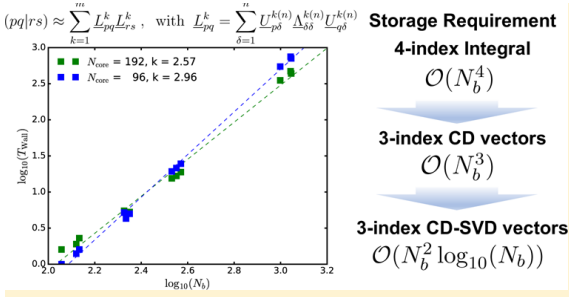


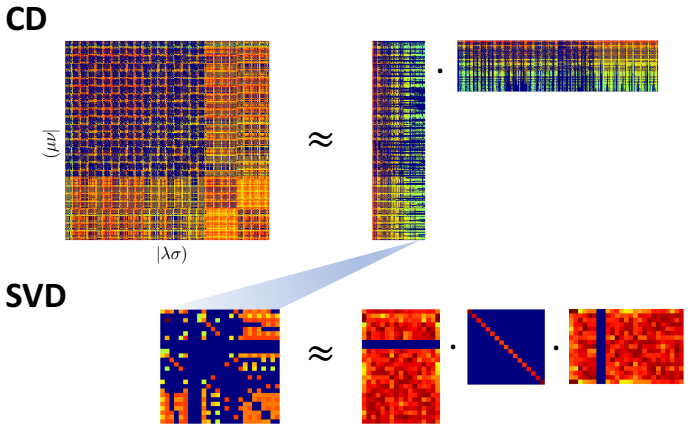
Reduced scaling CC methods in quantum chemistry

- Decomposition of “interaction” tensors – decomposition of 2-electron integrals
  - Possible in quantum chemistry (Coulomb forces) - not sure if applicable to general case for n-n interactions.
- Local methods (O(N) methods)
  - PNO, DLPNO methods
  - Require set of “empirical” threshold
  - Main effort towards engineering theory & parallel implementations

Two-step decomposition: Cholesky decomposition → SVD decomposition

$$(\mu\nu|\lambda\kappa) = \int \mu(1)\nu(1)\frac{1}{r_{12}}\lambda(2)\kappa(2)\,d1\,d2$$





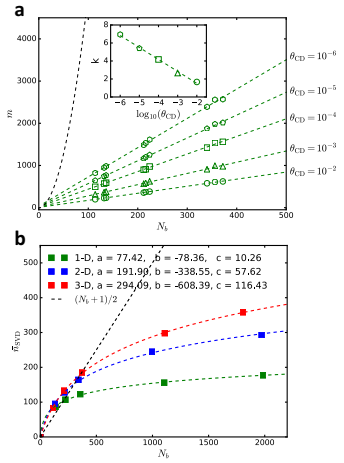
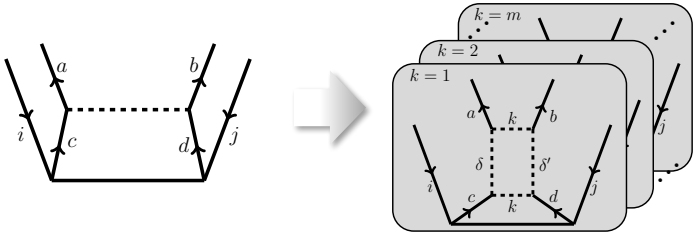


Table 1: Computed transition energy ($^1T_{1u} \leftarrow {}^1A_g$) and dipole polarizabilities (at $\omega = 0.0$ a.u. and 0.0428 a.u.) of C_{60} molecule. The geometry is adopted from Ref. 139, and is subject to D_{2h} symmetry. All 1s electrons were frozen during calculations. Due to the size of the system, we converged all CCSD iterations to 10^{-4}, and set both θ_{CD} and θ_{SVD} to 10^{-4}.

Basis set	N_b	m	$\bar{n}_{SVD}$	$\omega^{EOM-CCSD}_{^1T_{1u}\leftarrow{}^1A_g}$ [a]		$\alpha(\omega)^{LR-CCSD}$		
				CD-SVD	Canonical	ω	CD-SVD	Canonical
ZPolC	1080	2456	367	3.8683 eV	3.8661 eV	0.0 a.u.[b]	555.29 a.u. (82.29 Å^3)	554.71 a.u. (82.20 Å^3)
						0.0428 a.u.[c]	564.85 a.u. (83.70 Å^3)	564.30 a.u. (83.62 Å^3)
aug-cc-pVDZ	1380	4338	418	3.5139 eV	3.5130 eV	0.0 a.u.[b]	559.37 a.u. (82.89 Å^3)	559.44 a.u. (82.90 Å^3)
						0.0428 a.u.[c]	569.08 a.u. (84.33 Å^3)	569.15 a.u. (84.34 Å^3)

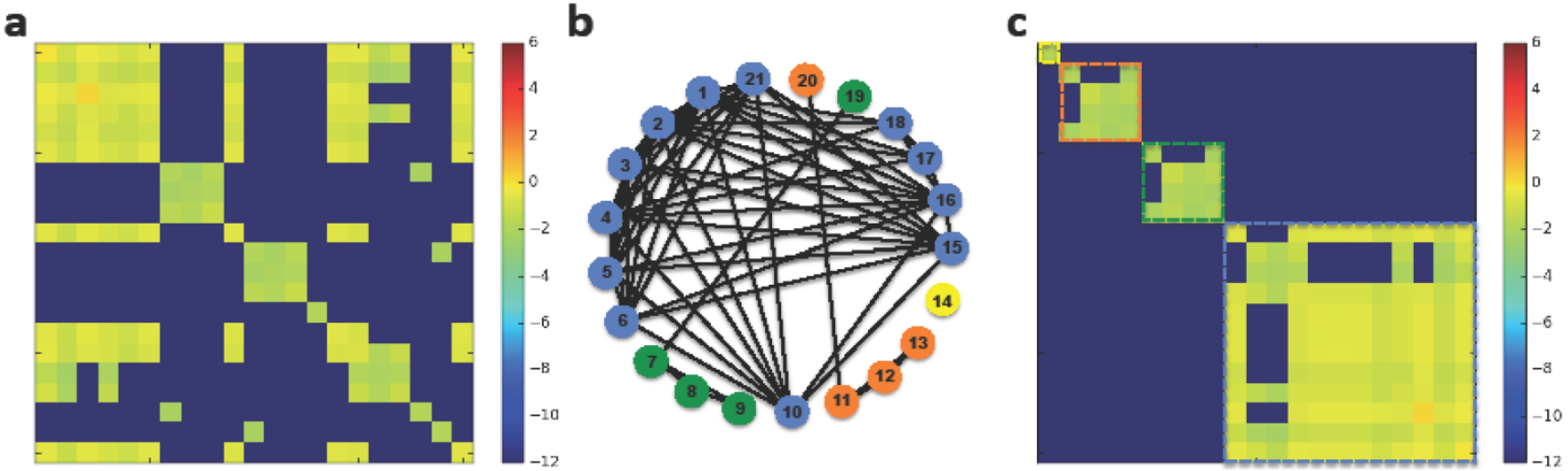
[a] Experimental values are ranging from 3.04 eV to 3.78 eV. See Ref. 132–136.
[b] Experimental value is 76.5 ± 8 Å^3. See Ref. 140.
[c] Experimental value is 79 ± 4 Å^3. See Ref. 141.



Highly Efficient and Scalable Compound Decomposition of Two-Electron Integral Tensor and Its Application in Coupled Cluster Calculations

Bo Peng* and Karol Kowalski*

Reverse Cuthill-McKee (RCM) algorithm for low-rank factorization of electron integral tensors



$$\mathbf{J} = \mathbf{P}_\mathbf{J}\mathbf{B}_\mathbf{J}\mathbf{P}_\mathbf{J}^T \qquad \mathbf{B}_\mathbf{J} = \begin{bmatrix} \mathbf{B}_{\mathbf{J},1} & 0 & \cdots & 0 \\ 0 & \mathbf{B}_{\mathbf{J},2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{B}_{\mathbf{J},n} \end{bmatrix} = \bigoplus_{k=1}^{n} \mathbf{B}_{\mathbf{J},k} \qquad \mathbf{J} = \mathbf{P}_\mathbf{J} \bigoplus_{k=1}^{n} \left(\sum_{m}^{N_k} \mathbf{L}_{\mathbf{J},k}^{m} (\mathbf{L}_{\mathbf{J},k}^{m})^T \right) \mathbf{P}_\mathbf{J}^T,$$

Low-rank factorization of electron integral tensors and its application in electronic structure theory

Bo Peng *, Karol Kowalski **

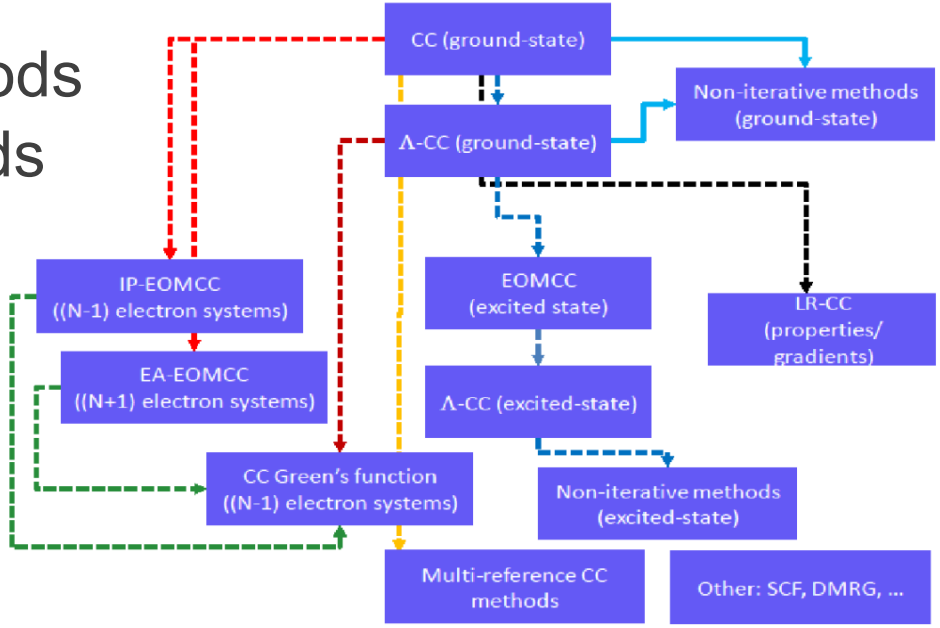
TAMM (Tensor Algebra for Many-body Methods)

- TAMM
  - C++
  - Dense tensor algebra system
    - Extension to sparse tensors (under development)
  - Specification of tensor expression in a domain-specific language
  - Offline analysis & optimization
  - Online optimization (under development)
  - Interoperability with existing TCE implementations

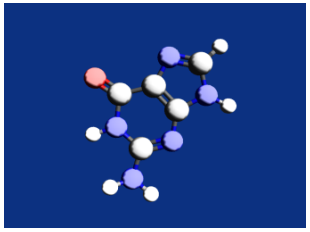
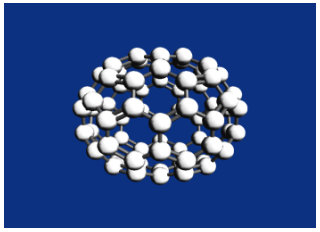
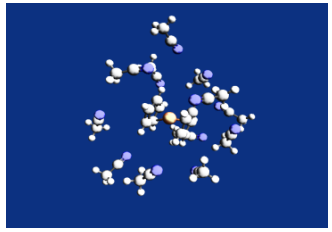
```
t1 {

index h1,h2,h3,h4,h5,h6,h7,h8 = O;
index p1,p2,p3,p4,p5,p6,p7 = V;

array i0[V][O];
array f[N][N]: irrep_f;
array v[N,N][N,N]: irrep_v;
array t_vo[V][O]: irrep_t;
array t_vvoo[V,V][O,O]: irrep_t;
array t1_2_1[O][O];
array t1_2_2_1[O][V];
array t1_3_1[V][V];
array t1_5_1[O][V];
array t1_6_1[O,O][O,V];

t1_1:       i0[p2,h1] += 1 * f[p2,h1];
t1_2_1:     t1_2_1[h7,h1] += 1 * f[h7,h1];
t1_2_2_1:   t1_2_2_1[h7,p3] += 1 * f[h7,p3];
t1_2_2_2:   t1_2_2_1[h7,p3] += -1 * t_vo[p5,h6] * v[h6,h7,p3,p5];
t1_2_2:     t1_2_1[h7,h1] += 1 * t_vo[p3,h1] * t1_2_2_1[h7,p3];
t1_2_3:     t1_2_1[h7,h1] += -1 * t_vo[p4,h5] * v[h5,h7,h1,p4];
t1_2_4:     t1_2_1[h7,h1] += -1/2 * t_vvoo[p3,p4,h1,h5] * v[h5,h7,p3,p4];
t1_2:       i0[p2,h1] += -1 * t_vo[p2,h7] * t1_2_1[h7,h1];
t1_3_1:     t1_3_1[p2,p3] += 1 * f[p2,p3];
t1_3_2:     t1_3_1[p2,p3] += -1 * t_vo[p4,h5] * v[h5,p2,p3,p4];
t1_3:       i0[p2,h1] += 1 * t_vo[p3,h1] * t1_3_1[p2,p3];
t1_4:       i0[p2,h1] += -1 * t_vo[p3,h4] * v[h4,p2,h1,p3];
t1_5_1:     t1_5_1[h8,p7] += 1 * f[h8,p7];
t1_5_2:     t1_5_1[h8,p7] += 1 * t_vo[p5,h6] * v[h6,h8,p5,p7];
t1_5:       i0[p2,h1] += 1 * t_vvoo[p2,p7,h1,h8] * t1_5_1[h8,p7];
t1_6_1:     t1_6_1[h4,h5,h1,p3] += 1 * v[h4,h5,h1,p3];
```

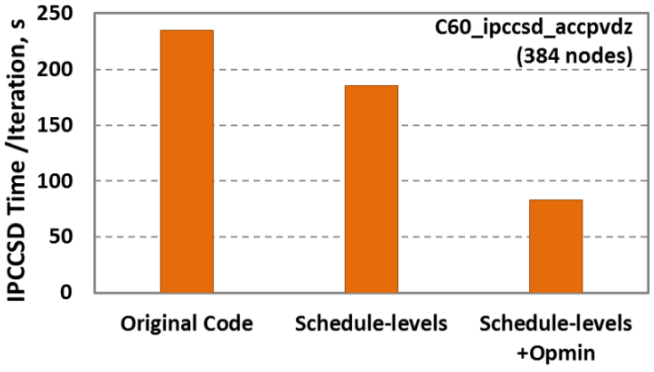
TAMM (Tensor Algebra for Many-body Methods)

- Quick deployment of various many-body methodologies:
  - CC ground-state
  - EOMCC formulation
  - Linear response CC
  - CC Green's function methods
  - Multi-reference CC methods
  - …



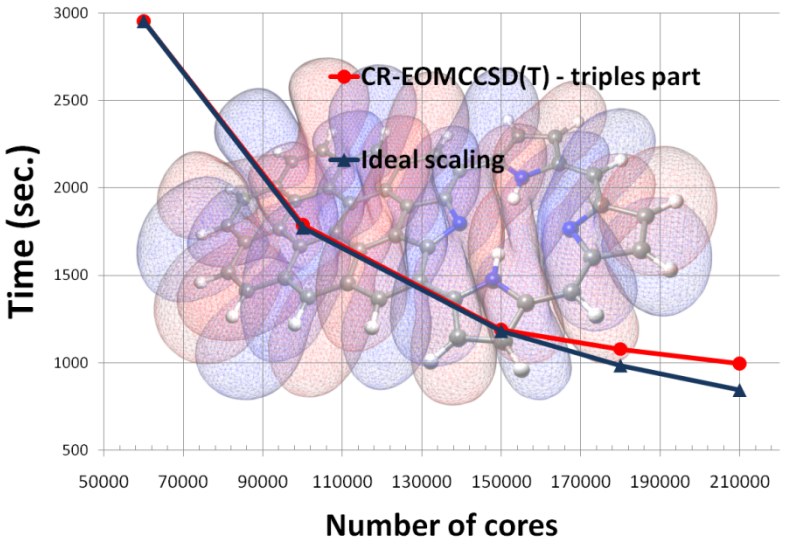
Illustrative examples of TAMM perfromance

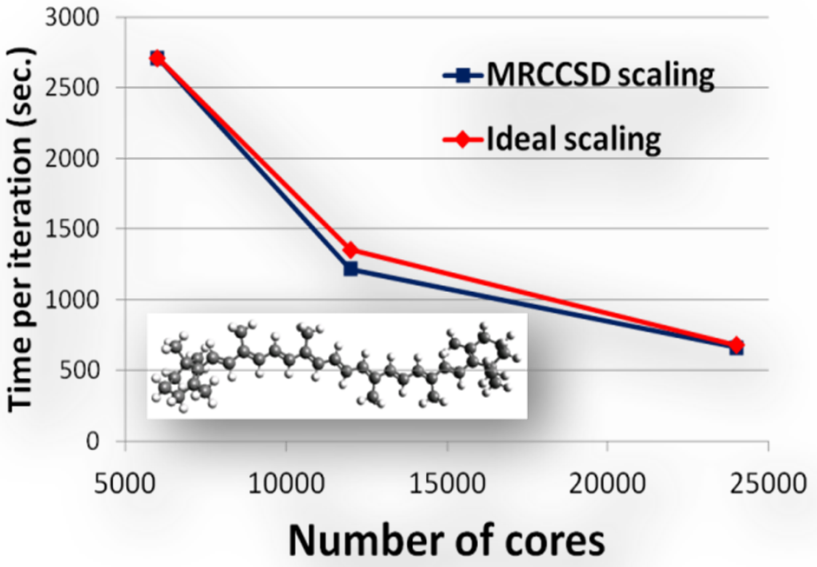





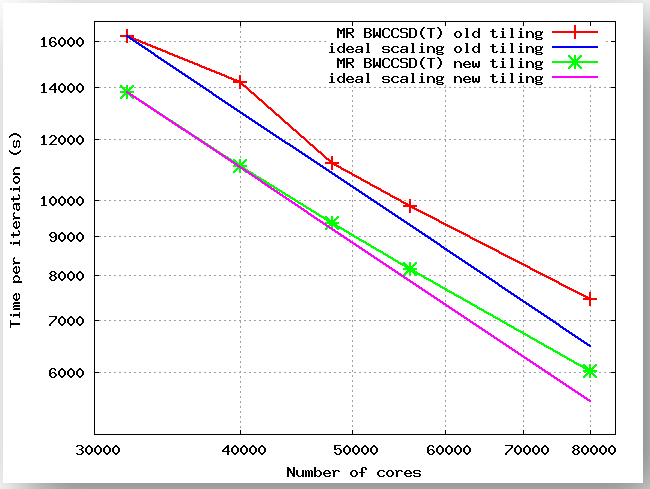


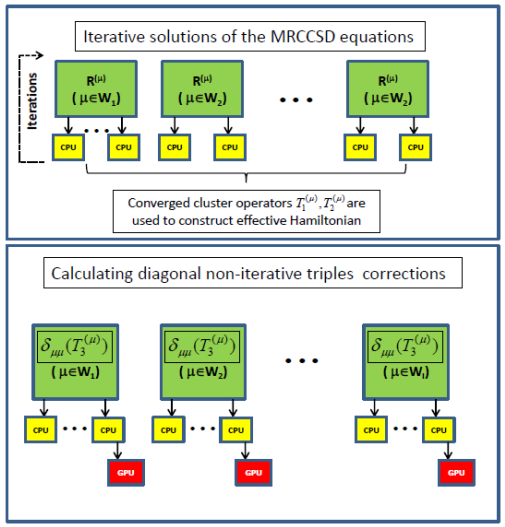
System	Leading amplitudes	IP-EOMCCSD	Koopman's theorem	Expt.
Uracil (aug-cc-pVQZ)	$r^{H} = 0.96$ $r^{H-4} = -0.11$	9.70	10.05	9.4-9.6
Guanine (aug-cc-pVQZ)	$r^{H} = 0.95$	8.21	8.20	8.0-8.3
C_{70} (maug-cc-pVDZ)	$r^{H} = 0.96$	7.48	7.68	7.47±0.02
Ferrocene (cc-pVQZ)	$r^{H-5} = 0.87$ $r^{H-4} = 0.16$ $r^{H-10} = 0.13$	6.77	9.26	6.82
Ferrocene in solution (mixed basis set)	$r^{H-5} = 0.88$	5.98	8.64	

Other examples (1)

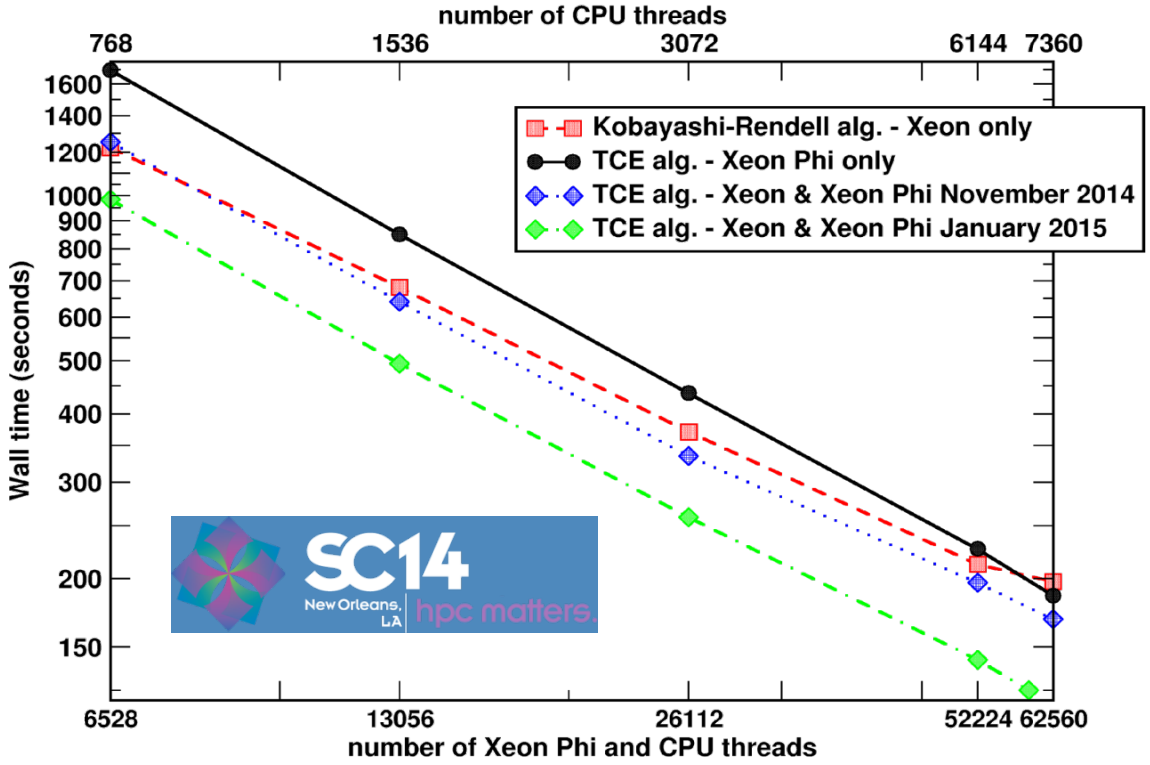








Other examples (2)



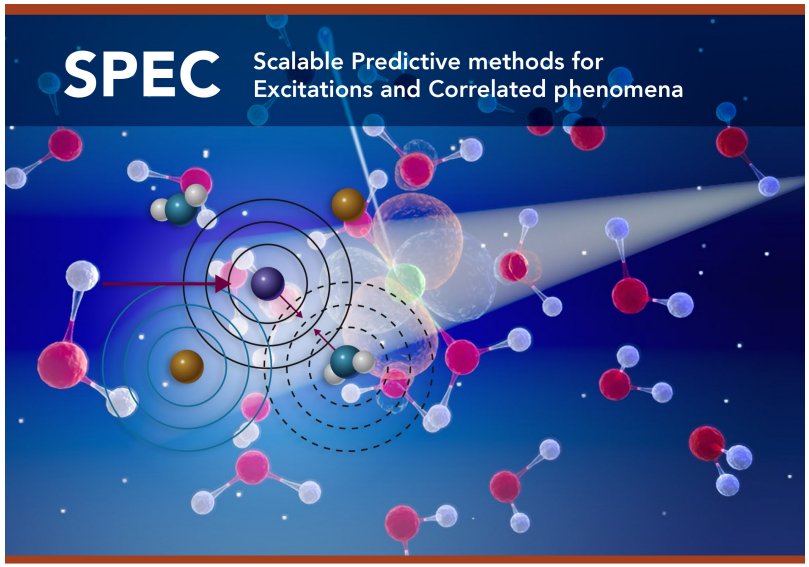
PNNL's science challenges that require many-body simulations

- Active sites in proteins
  - Effect of the environment in the active site function
- Enzymatic reactions
- Metabolic processes/metabolic cycles (effect of biotic and abiotic factors/stresses)
- Catalysis/photo-catalysis (low-pressure/low-temperature)
- Catalytic conversion of biomass to biofuels
- Electron transport in biological systems (long-range processes, respiration processes, DNA damage & DNA repair mechanisms)
- Interactions of biological systems with light

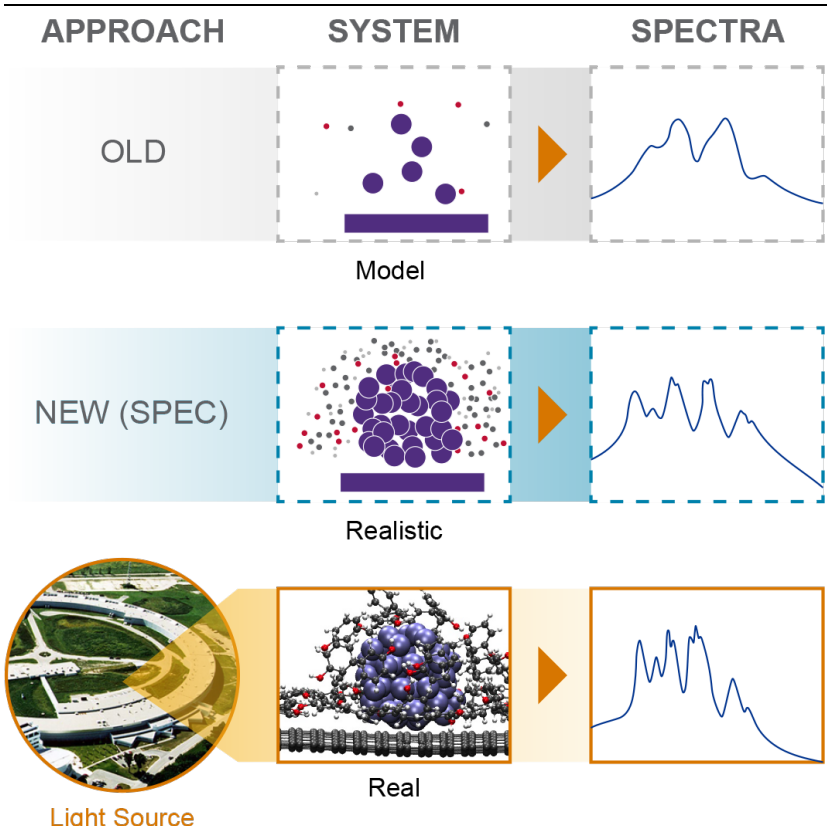
To describe these processes a detailed characterization of energetics & dynamics is needed: reliable methods are needed (cutting corners is not always a good idea)

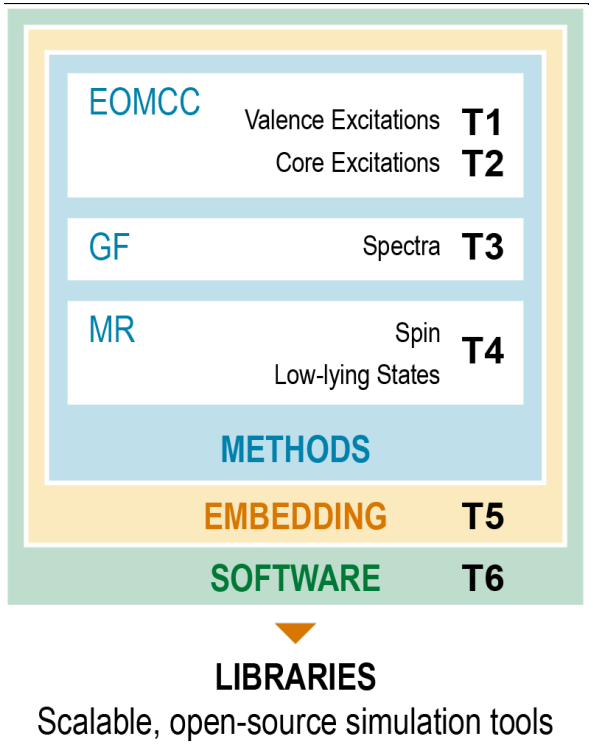
Ongoing projects targeting exa-scale (conventional) computing

- NWChem
  - Gaussian basis DFT / plane-wave DFT / high-accuracy methods / multi-scale methods
- NWChemEx (ex-scale extension of NWChem)
  - Ground-state CC formulations (canonical/reduced-scaling)
    - DLPNO-CC formulations for ground-state calculations
- SPEC
  - Equation-of-motion CC formalisms
  - Relativistic approaches
  - Hierarchical Green's function approaches:
    - GW /Bethe-Salpeter/ Cumulant theory
    - GF(n) perturbative approached
    - CC Green's function formulation
    - Self-energy embedding
  - MRCC methods



SPEC: Scalable Predictive methods for Excitations and Correlated Phenomena





Hybrid computing

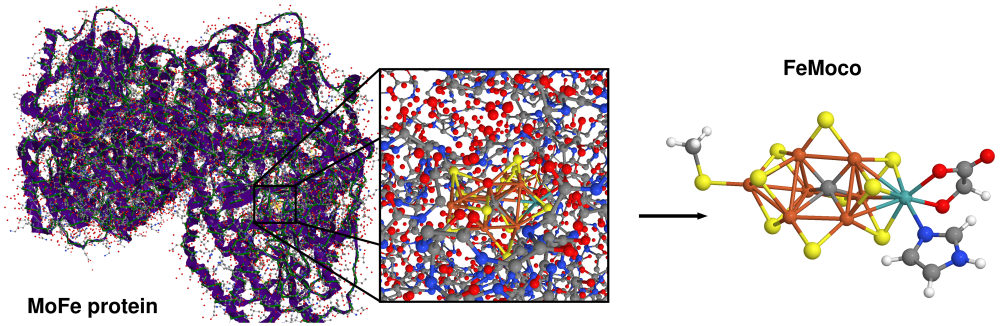
- Having 100-200 qubits, what can we do?:
  - **A:** Perform quickly CC/unitary-CC calculations for 50-100 (or more) orbitals?
  - **B:** Perform quickly FCI calculations for 50-100 orbitals?
  - **C:** Perform FCI calculations for 50-100 orbitals in "no time" and integrate it with conventional CC calculations for 1,000-2,000- or more orbitals

Hybrid computing

Perhaps accuracy of many-body models utilizing quantum computing can be further amplified by integrating them with conventional computational models at exa-scale?

Hybrid computing

- Quantum FCI calculations for nitrogen fixation by the enzyme nitrogenase

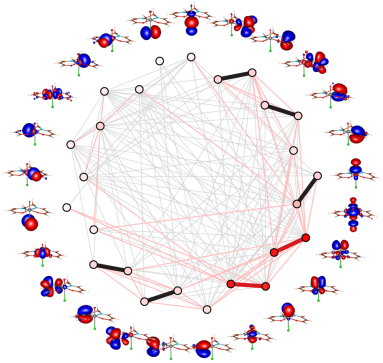


Elucidating reaction mechanisms on quantum computers

Markus Reiher[a,1], Nathan Wiebe[b,1], Krysta M. Svore[b], Dave Wecker[b], and Matthias Troyer[b,c,2]

Fig. 1. (*Left*) X-ray crystal structure 4WES (21) of the nitrogenase MoFe protein from *Clostridium pasteurianum* taken from the protein database (the backbone is colored in green, and hydrogen atoms are not shown), (*Middle*) the close protein environment of the FeMoco, and (*Right*) the structural model of FeMoco considered in this work (C, gray; O, red; H, white; S, yellow; N, blue; Fe, brown; and Mo, cyan).

- Externally corrected/Tailored CC formalisms (DMRG → CCSD)



T_3, T_4 *from cluster analysis of DMRG wave funciton*

T_1 and T_2 amplitudes (CCSD) formalism iterated in the presence of T_3 and T_4 excitations

THE JOURNAL OF PHYSICAL CHEMISTRY Letters — Letter — pubs.acs.org/JPCL

Coupled Cluster Method with Single and Double Excitations Tailored by Matrix Product State Wave Functions

Libor Veis,[*,†,⊥] Andrej Antalík,[†,⊥] Jiří Brabec,[†] Frank Neese,[‡] Örs Legeza,[*,§] and Jiří Pittner[*,†]

Integrating Full Configuration Interaction (FCI) method with single reference CC approaches

- Externally corrected/Tailored CC methods
  - CCSD amplitudes (T_1 & T_2) when iterated in the presence of exact T_3 & T_4 operators provide exact electronic energies

$$f_{SD}(T_1, T_2, T_3, T_4) = 0$$

FCI – extracting T_3 & T_4 through cluster analysis of the FCI wave-function :

$$e^T|\Phi\rangle = \ln(1 + C)\,|\Phi\rangle$$

$$T_1 = C_1$$

$$T_2 = C_2 - \tfrac{1}{2}C_1^2$$

$$T_3 = C_3 - C_1C_2 + \tfrac{1}{3}C_1^3$$

…

Sources for determining T_3 & T_4 operators:

- MCSCF
- DMRG

Methods of Moments of CC equations (renormalized CC formulations)

- Trial wave-function is needed to calculate renormalized corrections

$$E = E_{CCSD} + \frac{\langle \Psi | M_{CCSD} | \Phi \rangle}{\langle \Psi | e^{T_{CCSD}} | \Phi \rangle}$$

QFCI wave function: ground- and excited-state formalisms (when coupled with MM EOMCC formulations)

"CI-CC compromise" - sources for the trial wave functions:

- CI (approximate)
- CC (approximate)
- perturbative (approximate)

Multi-reference formulations for strongly correlated systems

- MR-MBPT methods
  - CASPT2 → diagonalization of electronic Hamiltonian in the model spaces (40 active orbitals)
    - Moving to 50-100 active orbitals would significantly extend the applicability range of CASPT2 methods
- MR-CC methods
  - Diagonalization of effective Hamiltonian in large model space

What can be simulated with hybrid formulations?

- Active sites/Enzymatic reactions/Catalysis:
  - Externally corrected/Tailored CC methods & renormalized CC for systems described by 1,000-3,000 orbitals (400-1200 correlated electrons)
  - Multi-reference methods (CASPT2/MRCC) – models spaces – 100 orbitals
- Photo-catalysis
  - Very accurate EOMCC methods for systems composed of 400-1200 correlated electrons
  - Accurate descriptions of excited-state potential energy surfaces (for multiple electronic states)
    - Assuming that quantum FCI diagonalization algorithm can “lock” multiple eigenvalues/eigenvectors

Acknowledgement

- Collaborators: Sriram Krishnamoorthy's group, Bo Peng, NWChem group, GA group
- BER – NWChem
- ASCR – NWChemEx project
- BES – SPEC project
- Intel Parallel Computing Center
- PNNL LDRD program

For more information, visit EMSL's website

www.emsl.pnnl.gov