

New computational tools for many-body theories

Karol Kowalski & Sriram Krishnamoorthy & Bo Peng

Scientific Innovation Through Integration • www.emsl.pnnl.gov

Many-body formulations

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

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:
	- \rightarrow CCSD \rightarrow CCSD(T) \rightarrow CCSDT \rightarrow CCSDTQ \rightarrow ...

 N^6 \rightarrow N^7 \rightarrow N^8 \rightarrow N^{10} \rightarrow …

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 \rightarrow SVD decomposition \mathbf{E} lectron Its Application in Coupled Cluster in Coupled Cluster

Table 1. Computed transition energy ($T_{1u} \leftarrow T_{g}$) and upplie polarizabilities (at $\omega = 0.0$ a.u. and 0.0428 a.u.) of C_{60} molecule. The geometry is adopted from $\omega = 0.0$ a.u. and 0.0428 a.u.) or C_{60} molecule. The geometry is adopted from
Ref. 139, and is subject to D_{2h} symmetry. All 1s electrons were frozen during Table 1: Computed transition energy $(^1T_{1u} \leftarrow \ ^1A_g)$ and dipole polarizabilities (at 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} . from ∼100 up to ∼2,000, the observed numerical scaling of our implementation shows [−] () *Nb*

ⁱ ^c ^d ^j i ⁱ ^c ^d ^j b Highly Efficient and Scalable Compound Decomposition of Two-The Lindon Integral Tensor and Its Application in Coupled Cluster
Electron Integral Tensor and Its Application in Coupled Cluster
Calculations *i* Calculations

Bo Peng* and Karol Kowalski* \bullet

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

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)
	- **EXEC** 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];<br>t1_2: t1_2: t1_1: t2_1[h7, h1] += 1 * f[h7,
```

```
t1_21: t1_21[h7,h1] \mathrel{+}= 1 * f[h7,h1];<br>t1_221: t1_221[h7,p3] \mathrel{+}= 1 * f[h7,p3]
```

```
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_22: t1_2 1[h7,h1] t= 1 * t vo[p3,h1] * t1_2 2 1[h7,p3];<br>t1_2 3: t1_2 1[h7,h1] t= -1 * t vo[p4,h5] * v[h5,h7,h1,p4];
```

```
t1_2_3: t1_2_1[h7,h1] \coloneqq -1 * t\_vo[p4,h5] * v[h5,h7,h1,p4];<br>t1_2_4: t1_2_1[h7,h1] \coloneqq -1/2 * t\_vvo[p3,p4,h1,h5] * v[h5,
```

```
t1 2 1[h7,h1] += -1/2 * t vvoo[p3,p4,h1,h5] * v[h5,h7,p3,p4];
```

```
t1_2: i0[p2, h1] \nleftarrows -1 * t_v0[p2, h7] * t1_2_1[h7, h1];<br>
t1 3 1: t1 3 1[p2,p3] i = 1 * f[p2, p3];
```

```
t1_{\_31}: t1_{\_31[p2,p3]} +1 * f[p2,p3];<br>t1_{\_32}: t1_{\_31[p2,p3]} +1 * t t \text{vol}(p4);
```

```
t1\overline{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];<br>t1_4: i0[p2, h1] += -1 * t\_vo[p3, h4] * v[h4, p2, h1, p3]
```

```
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\overline{5_2}: t1\overline{5_1}1[h8,p7] \leftrightarrow 1 * t\_vo[p5,h6] * v[h6,h8,p5,p7];<br>t1\overline{5:} i0[p2,h1] \leftrightarrow 1 * t~vvo[p2,p7,h1,h8] * t1\overline{5_1}1[h8,p5]
```

```
t1_5: i0[p2,h1] += 1 * t_vvoo[p2,p7,h1,h8] * t1_5_1[h8,p7];<br>t1_6 1: t1_6 1[h4,h5,h1,p3] += 1 * v[h4,h5,h1,p3];
```

```
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

 \mathbf{P} …

- ‣ EOMCC formulation
- ‣ Linear response CC
- ‣ CC Green's function methods
- ‣ Multi-reference CC methods

Illustrative examples of TAMM perfromance

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: Scalable Predictive methods for* i
- **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** molecules in various energy regimes. Research using these light sources is vital to answering fundamental riers inhibit their full utilization. The spectra of complex molecular systems from these light sources are entists' ability to simulate these spectra for realistic molecular systems. We will develop theoretical and

Figure 1.1. Our proposed approach (SPEC) will provide the necessary link to interpret spectra obtained at light source

ble us to realize the full capability of DOE's experimental and computational user facilities. when and compatitional does necessarily.

Scalable, open-source simulation tools

Figure 2.1. SPEC consists of an integrated approach combining method development, computer tools and optimization $(Tasks T1)$ open-source libraries.

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

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

 $\frac{1}{2}$ split-localized orbitals and the mutual information (M $\frac{1}{2}$ and 312) for 1 and 312 for 12 and 312 for 12 and 312 for 13 and 312 for

Quantum Chemical Methods for Mechanistic Studies A_{th} the heart of any chemical process is its mechanism, the elucida-

of dinitrogen-coordination \mathcal{F} . The dinitrogenecometer \mathcal{F} intermediate and spin states in the feasible and spin states and spin states are feasible and spin states are f α assessed with relative energy. Especially, the relative energy. Especially, α kineticmodeling poses tightlimits on the activation energies entering the argument of exponentials in rate expressions. $\mathbb{F}_{\mathbb{F}_{q}}$ nitrogenese, and \mathbb{F}_{q} $\overline{}$ $\overline{\mathbf{a}}$ into the proton polynomial polyno

mutual information is constant red lines correspond to the strongest correlations (order of magnitude 1), followed by black (10−11−1), followed by black (10−11−1), followed by black (10−11−1), followed by black (10−11−1),

). One-site entropy values are represented by a color gradient of the respective dot, red by a color gradient of the largest value and white α

chemically active species

§ Quantum FCI calculations for nitrogen fixation by the enzyme nitrogenase

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). Fig. 1. (Left) X-ray crystal structure 4WES (21) of the nitrogenase MoFe protein from *Clostridium pasteurianum* taken from the protein database (the

 $DMRG \rightarrow CCS$ Externally corrected/Tailored CC formalisms (DMRG \rightarrow CCSD) fast a quantum computer would have to be to be to perform such cal-**Externally corrected** Tailored CC formali

> all affects and feasibility analysis (because electrostatic \mathcal{M} embedding will not change the number of orbitals considered for the wave function construction). We carried out (full) molecu-

> energy to every molecular structure. The accurate calculation of the pixotal challenge, here considered by $\mathcal{L}_{\mathcal{A}}$ computing. Characteristic molecular structures are optimized to provide local minimum structures indicating stable intermedi-

> > \sim

 \mathcal{A} states of oxo-Mathematic contract basis \mathcal{A}

d \mathcal{T}_2 amplitudes (**requiring exorbitant resources.** formalism iterated e presence of $\overline{T_3}$ excitations, very accurate exp that quantum simulation will be able to provide answers to prob- T_1 and T_2 amplitudes (CCSD) $\qquad \qquad \mid$ This enzyme accomplishes the remarkable transformation of \mathcal{F}_1 dinitrogen into two ammonia molecules under ambient conditions. Whereas the industrial Haber–Bosch catalyst requires high temperature and pressure and is therefore energy-intensive, the active site of Mo-dependent nitrogenase, the iron molybdenum in the presence of T_3 and T_4 the MoFe protein of nitrogenase (Fig. 1, *Left*) and the FeMoco buried in this protein (Fig. 1, *Middle*). Despite the importance of this process for fertilizer production that makes nitrogen

understanding and prediction of complex reaction \mathcal{L}

into account the substantial overhead of quantum error correc-

Chemistry
.e*tters*

Our work addresses the question of compelling killer applications for although quantum computers. Although quantum computers. a strong candidate, the lack of details of α details of α details of α Letter pubs.acs.org/JPCL

Significance

to emerge from existing research and further gives confidence

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

 $\text{Libor Veis,}^{\ast\circ\uparrow,\perp}\text{ Andrej Antalik,}^{\uparrow,\perp}\text{ Jiří Brabec,}^{\uparrow}\text{ Frank Neese,}^{\ddagger}\text{ Örs Legeza,}^{\ast\circ\circ}\text{ and Jiří Pittner}^{\ast\uparrow}$

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

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

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

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

\n
$$
T_{1} = C_{1}
$$

\n
$$
T_{2} = C_{2} - \frac{1}{2} C_{1}^{2}
$$

\n
$$
T_{3} = C_{3} - C_{1} C_{2} + \frac{1}{3} C_{1}^{3}
$$

\n...

Sources for determining T_3 & *T4* operators:

MCSCF

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

• DMRG

Methods of Moments of CC equations (renormalized CC formulations)

§ Trial wave-function is needed to calculate renormalized corrections

Multi-reference formulations for strongly correlated systems

§ MR-MBPT methods

- \triangleright CASPT2 \rightarrow 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

EMSL'S IMPACT

- Biofuels
- Catalysis
- Energy Storage
- Solar Power
- > Explore EMSL's Impact

Become an EMSL User

EMSL is known for its cross-cutting diversity of instruments and expertise available under one roof. Scientists and scientific teams can accelerate new discoveries through a no-cost collaboration with EMSL.

> What can EMSL do for you? > Get Started

Kathryn Mireles, EMSL User Washington State University

EMSL's 2014 Call for Proposals is now open

A new study shows that scientists can

www.emsl.pnnl.gov