Scientific Innovation Through Integration

EMSL

Novel Parallel Algorithms for High-Accuracy Coupled Cluster Calculations

Karol Kowalski

Pacific Northwest National Laboratory

Outline

- Equation of motion coupled-cluster (EOMCC) formalism
  - Short review of methodologies
  - Parallel performance
- Multi-reference CC (MRCC) theories
  - Short review of state-selective methods
  - Challenges
  - New implementations: reference level parallelism
  - Applications

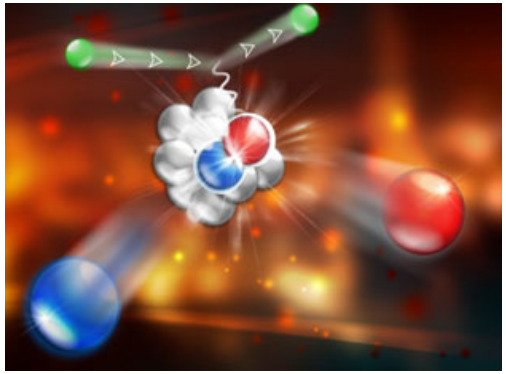
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

What we want to solve

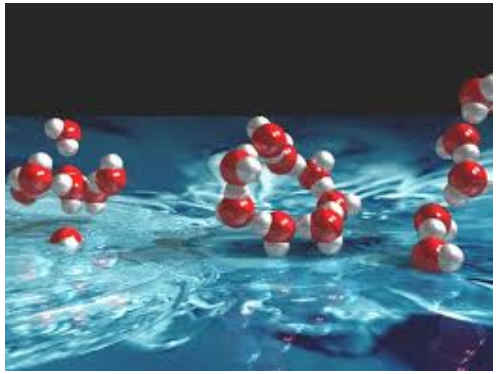
$$\hat{H}\left|\Psi(1,\ldots,N)\right\rangle = E\left|\Psi(1,\ldots,N)\right\rangle$$

Many Particle Systems

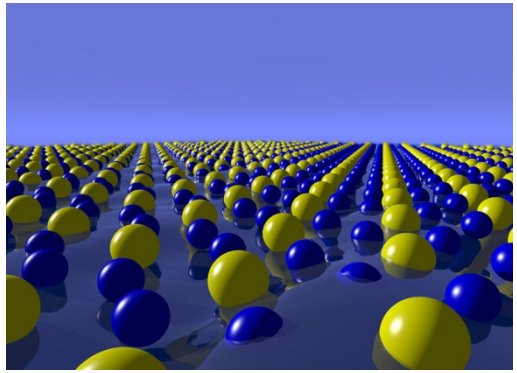
Nuclear Physics



Quantum Chemistry



Solid State Physics



Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Exact solution of Schrödinger equation

Weyl formula (dimensionality of full configuration interaction space) – exact solution of Schrödinger equation

$$f(n,N,S)=\frac{2S+1}{n+1}\binom{n+1}{N/2-S}\binom{n+1}{N/2+S+1}$$

n – total number of orbitals

N – total number of correlated electrons

S – spin of a given electronic state

A simple example: C_2 molecule; *N=12*, *n=100*

$f(n,N,S)\sim 10^{17}$

Equation-of-motion coupled cluster approach

$$\left|\Psi_0\right\rangle = e^T \left|\Phi\right\rangle$$

Coester, Kummel, Cizek,Paldus,Bartlett,...

$$\left|\Psi_K\right\rangle = R_K e^T \left|\Phi\right\rangle$$

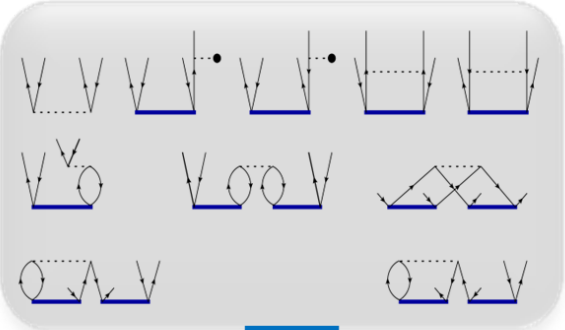
Bartlett, Stanton, ...

Method	Numerical complexity	Applications
EOMCCSD (singles & doubles)	N^6	Excitation energies of singly excited states
CR-EOMCCSD(T) (perturbative triples)	N^7	Excitation energies & potential energy surfaces of more "complicated" states
EOMCCSDT (singles & doubles & triples)	N^8	As CR-EOMCCSD(T) but more accurately

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Tensor Contraction Engine (TCE)

- Symbolic algebra systems for coding complicated tensor expressions: Tensor Contraction Engine (TCE)
- Parallel tools needed: Global Arrays (GA)



OCE

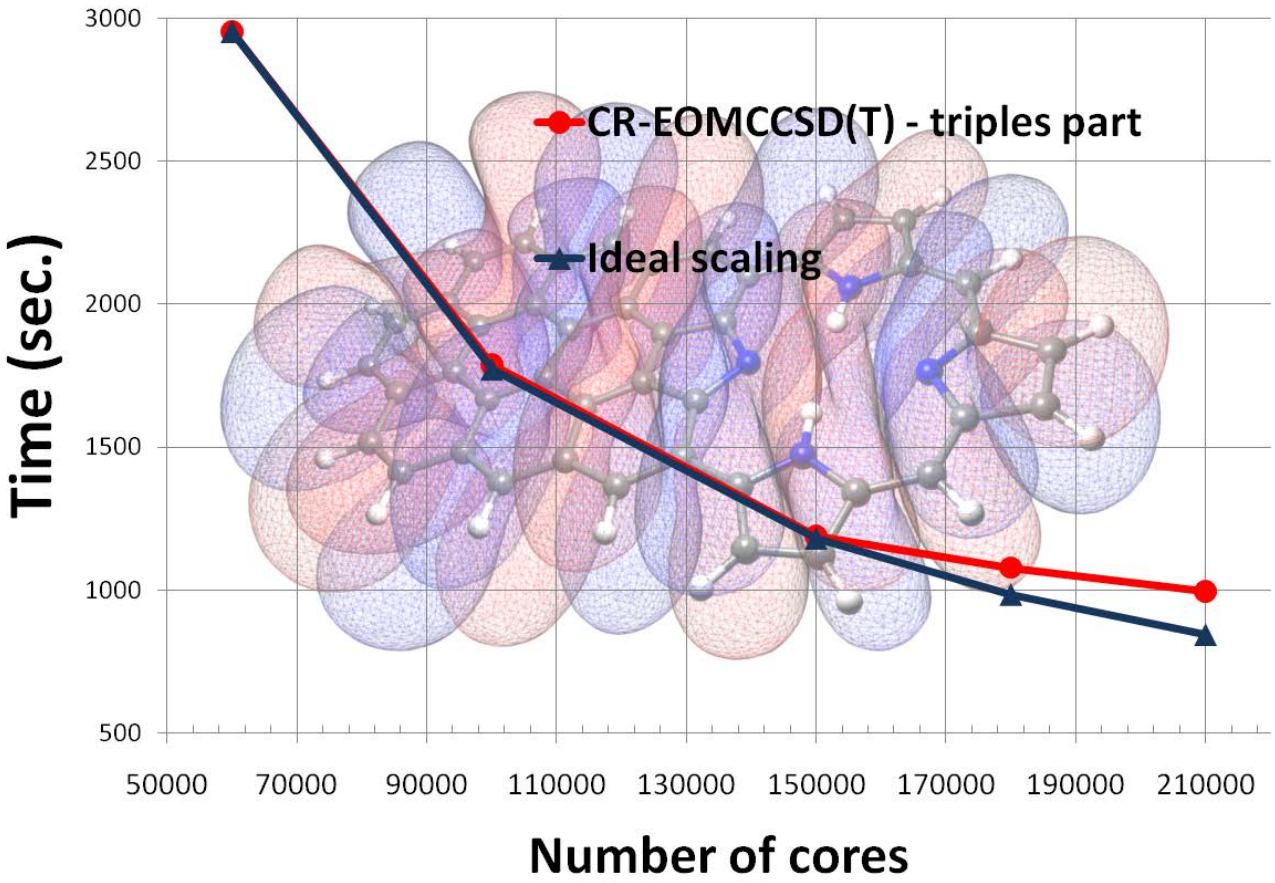
$$+\frac{1}{4}v_{ef}^{mn}t_{ij}^{ef}t_{mn}^{ab}-\frac{1}{2}v_{ef}^{mn}t_{mi}^{ef}t_{nj}^{ab}+$$

TCE

```
next = NXTASK(nprocs, 1)
DO p3b = noab+1,noab+nvab
DO p4b = p3b,noab+nvab
DO h1b = 1,noab
DO h2b = h1b,noab
IF (next.eq.count) THEN
CALL GET_HASH_BLOCK(d_a,dbl_mb(k_a),dim
- 1 + (noab+nvab) * (h1b_1 - 1 + (noab+
+nvab) * (p3b_1 - 1)))))
CALL GET_HASH_BLOCK_I(d_a,dbl_mb(k_a),d
```


Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

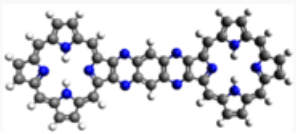
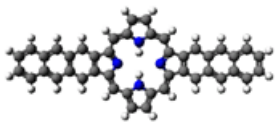
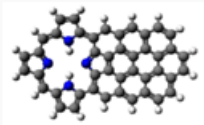
Scalability of the non-iterative EOMCC codes (A.D. 2011)

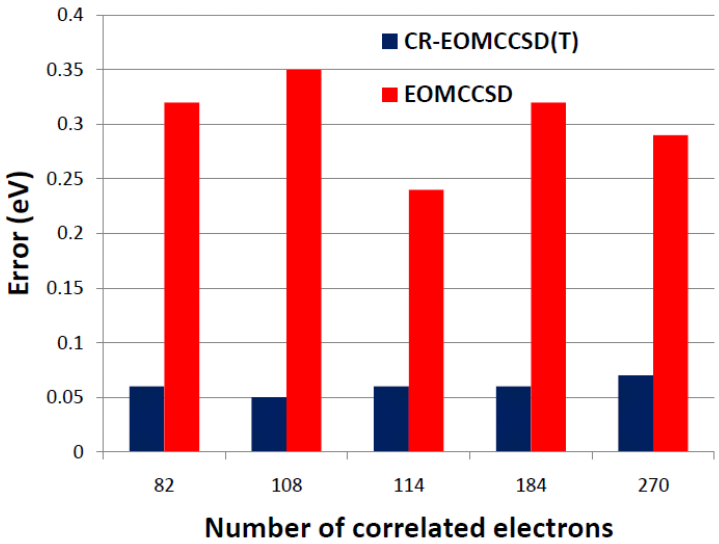


- Scalability of the triples part of the CR-EOMCCSD(T) approach for the FBP-f-coronene system in the AVTZ basis set. Timings were determined from calculations on the Jaguar Cray XT5computer system at NCCS.

K. Kowalski. S. Krishnamoorthy, R.M. Olson, V. Tipparaju, E. Apra, Supercomputing 2011

Functionalization of the porphyrines

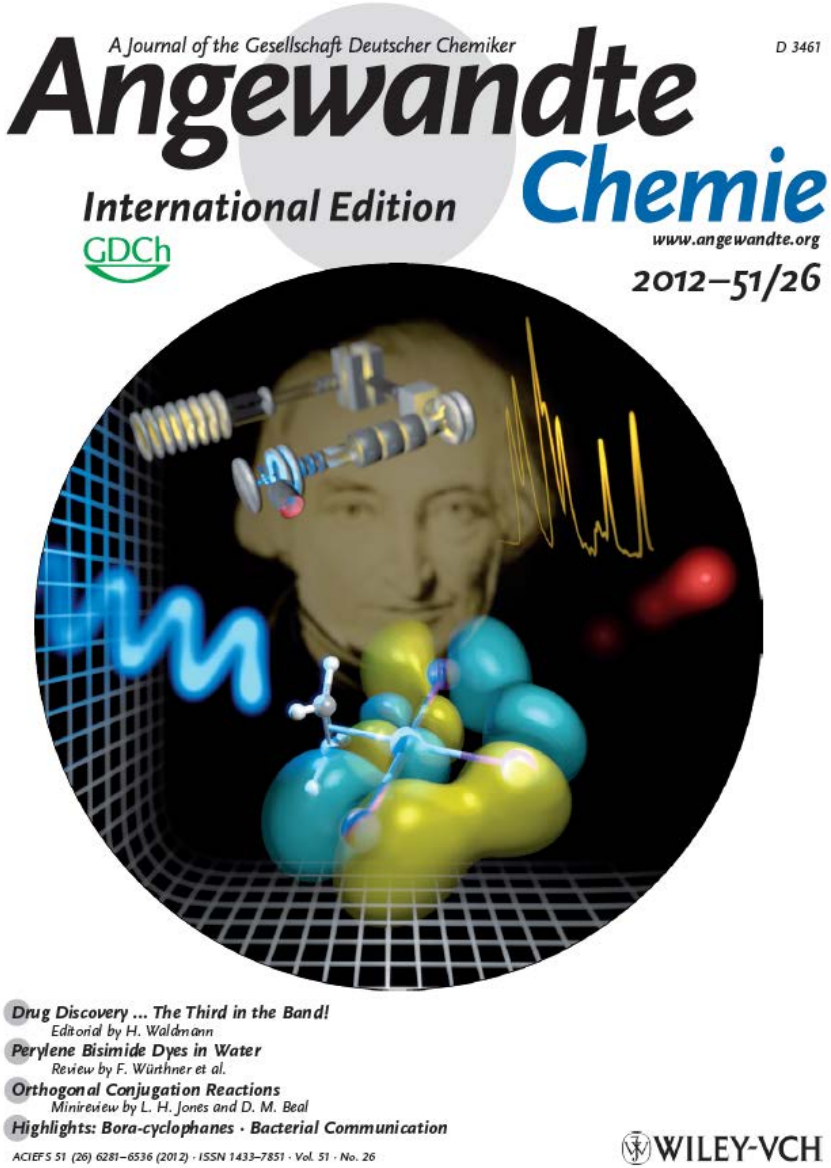
System	Leading excitations	CR-EOMCCSD(T) (eV)
	H→ L, H-1→L+1 H-1→L, H→L+1	2.32 (Expt. 2.27 eV) 1.86 (Expt. 1.91 eV)
	H→L, H-1→L+1, H-2→L+2, H-3→L+3	1.91 (Expt. 1.84 eV)
	H→ L	1.78
	H→ L	1.36

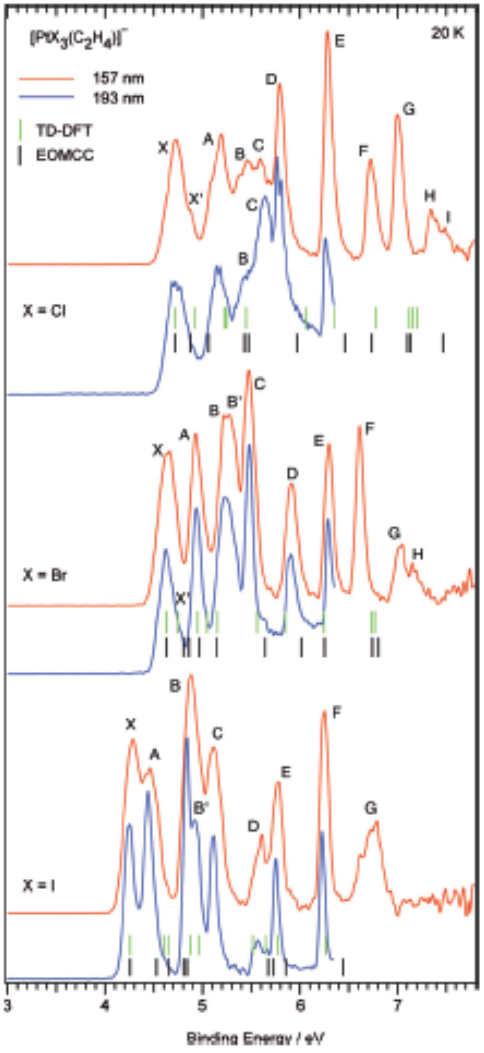


A comparison of calculated and experimental data shows that in order to describe singly excited states in large systems, triple excitations are needed

K. Kowalski, S. Krishnamoorthy, O. Villa, J.R. Hammond, N. Govind, J. Chem. Phys. (2010);
K. Kowalski, R.M. Olson, S. Krishnamoorthy, V. Tipparaju, E. Apra, J. Chem. Theory Comput. (2011))

Science impact: synergy between experiment, theory, and HPC



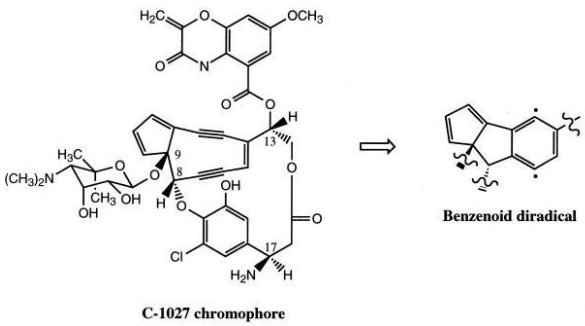


Pacific Northwest NATIONAL LABORATORY

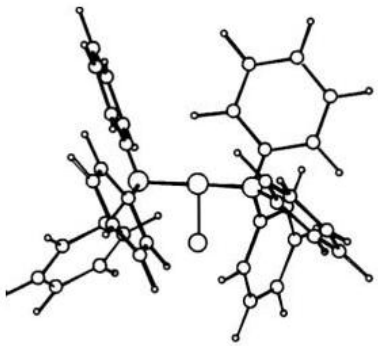
U.S. DEPARTMENT OF ENERGY

Proudly Operated by Battelle Since 1965

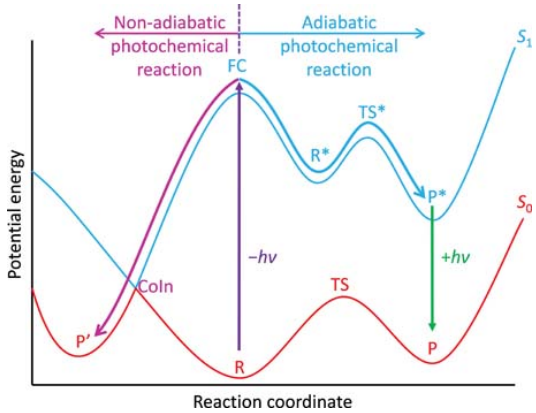
MRCC theory: new approaches and implementations



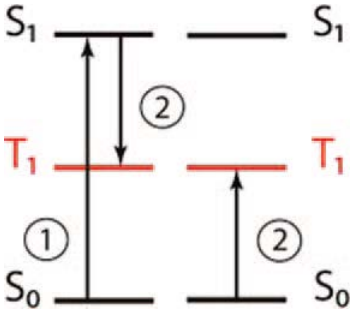
W. Liu, B. Shen, Antimicrob Agents Chemother. 2000, **44**: 382–392.



$Rh(PPh_3)_2Cl$, F. Maseras, A. Lledos, Catalysis by Metal Complexes, 2002, **25**, pp. 6



B. Lasorne, G.A. Worth, M.A. Robb, Comp. Mol. Sci. 2011, **1**, 460-475.



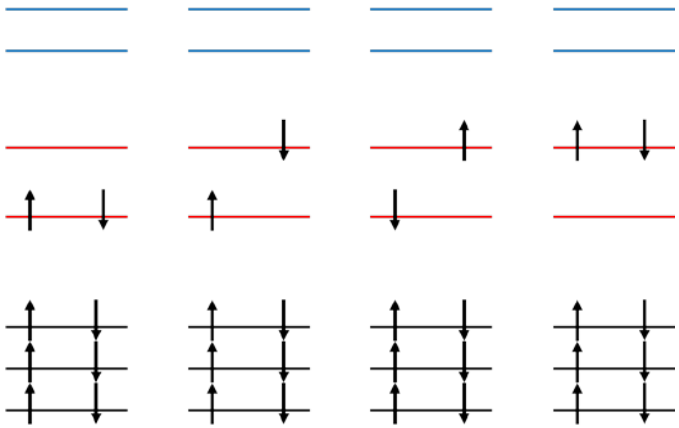
M.B. Smith, J. Michl, Chem. Rev. **110**, 6891 (2010).

MRCC theory in a nutshell

$$\left|\Phi\right\rangle \longrightarrow M_0 = ls\left\{\left|\Phi_\mu\right\rangle\right\}_{\mu=1}^{M}$$

Reference function

Model space



Schematic representation of the complete model space corresponding to two active electrons distributed over two active orbitals (red lines). Only determinants with M_S=0are included in the model space.

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

MRCC theory in a nutshell

$$\left|\Psi\right\rangle = \sum_{\mu=1}^{M} c_\mu e^{T^{(\mu)}} \left|\Phi_\mu\right\rangle$$

Jeziorski-Monkhorst Ansatz

$$H^{eff} c = Ec$$

$$H^{eff} = [\langle \Phi_\nu | (He^{T^{(\mu)}})_C | \Phi_\mu \rangle]$$

State-Universal MRCC methods

- Energies for M states
- Bloch Equation

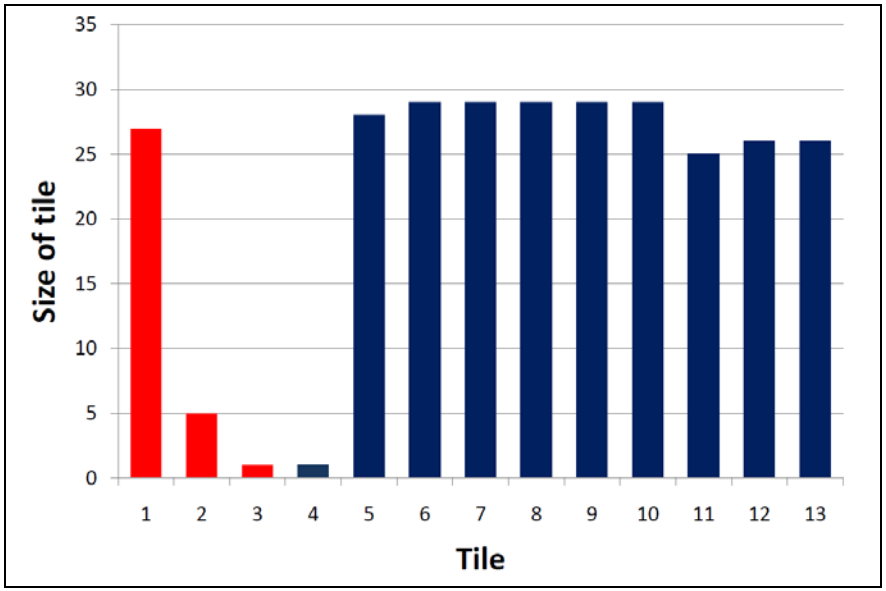
State-Specific MRCC methods

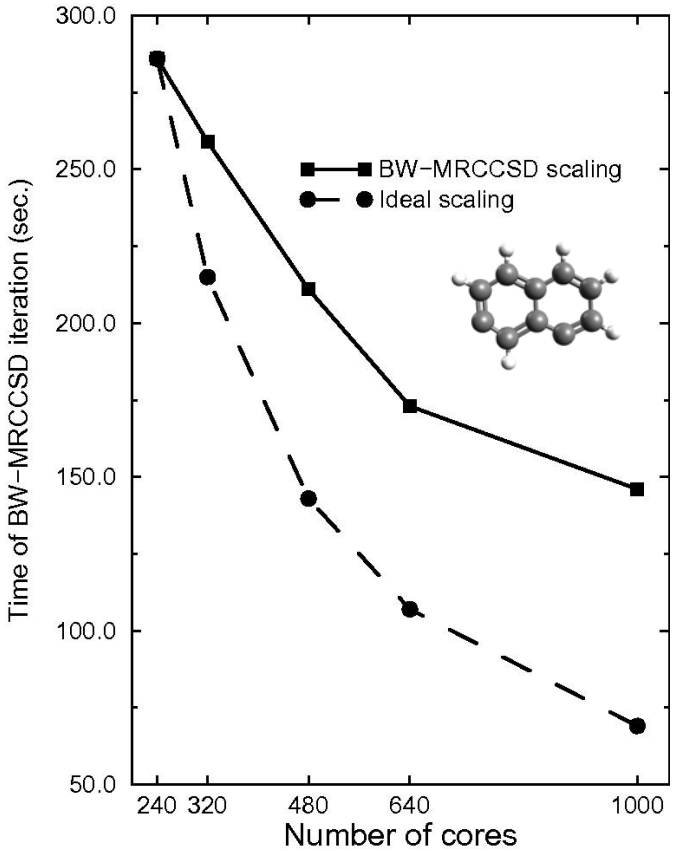
- Target one state at time (less prone to intruder states)
- Sufficiency conditions: BW-MRCC, Mk-MRCC, …

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Towards parallel MRCC – blazing the trail

- Problems with load balancing caused by active-orbital logic

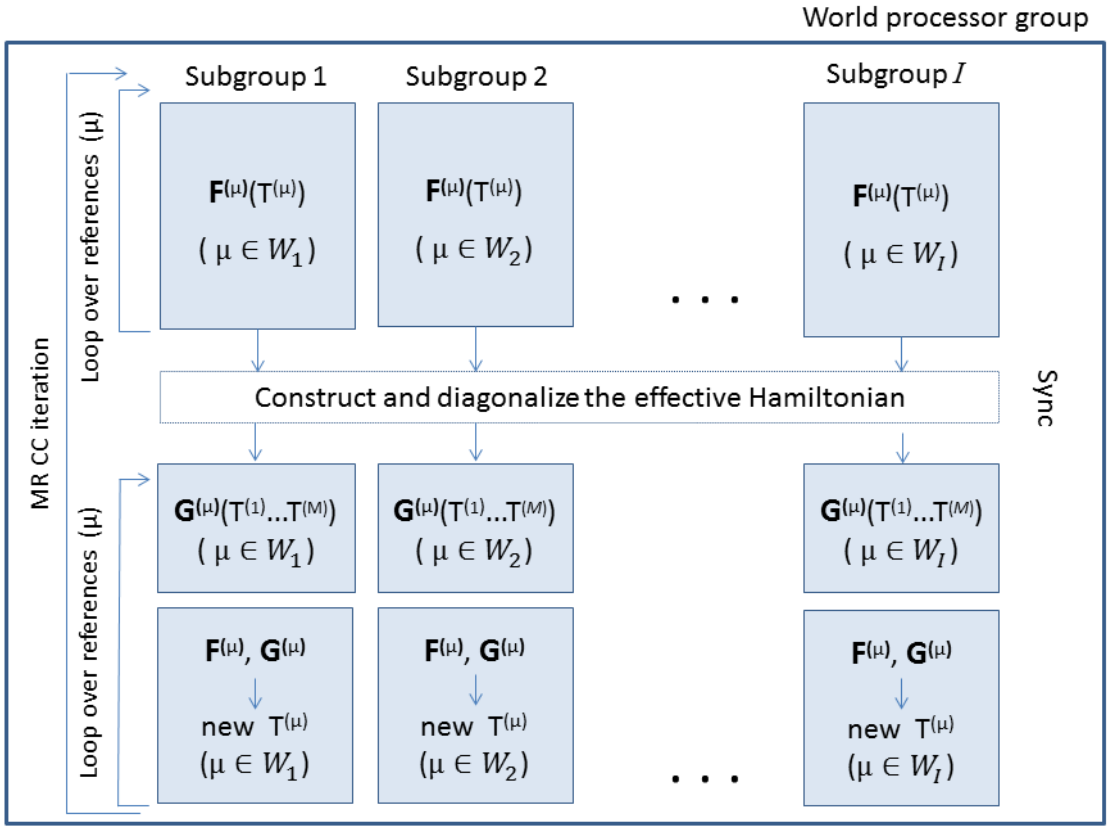




Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Processor groups (PGs) and reference level parallelism

$$R^{(\mu)} = F^{(\mu)}(T^{(\mu)}) + G^{(\mu)}(T^{(1)},\ldots,T^{(\mu)},\ldots,T^{(M)}) = 0 \qquad \forall_{\mu=1,\ldots,M}$$

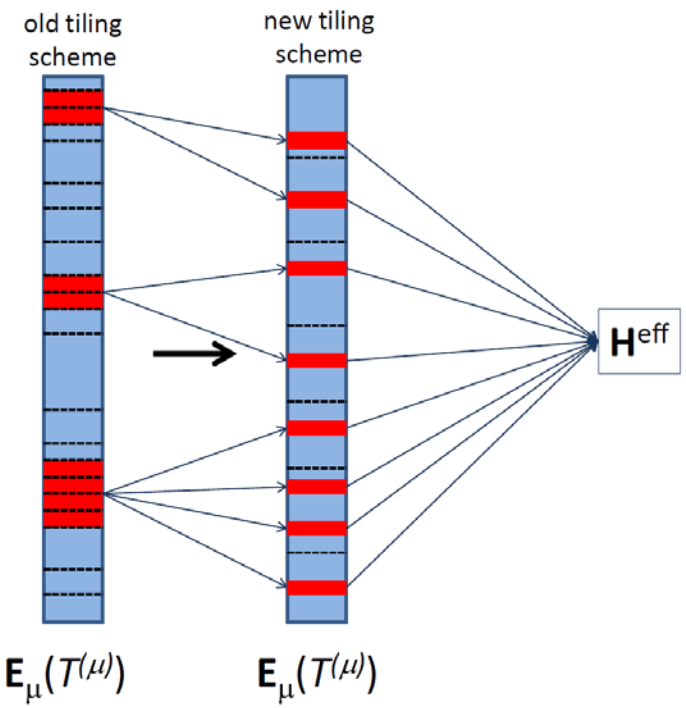


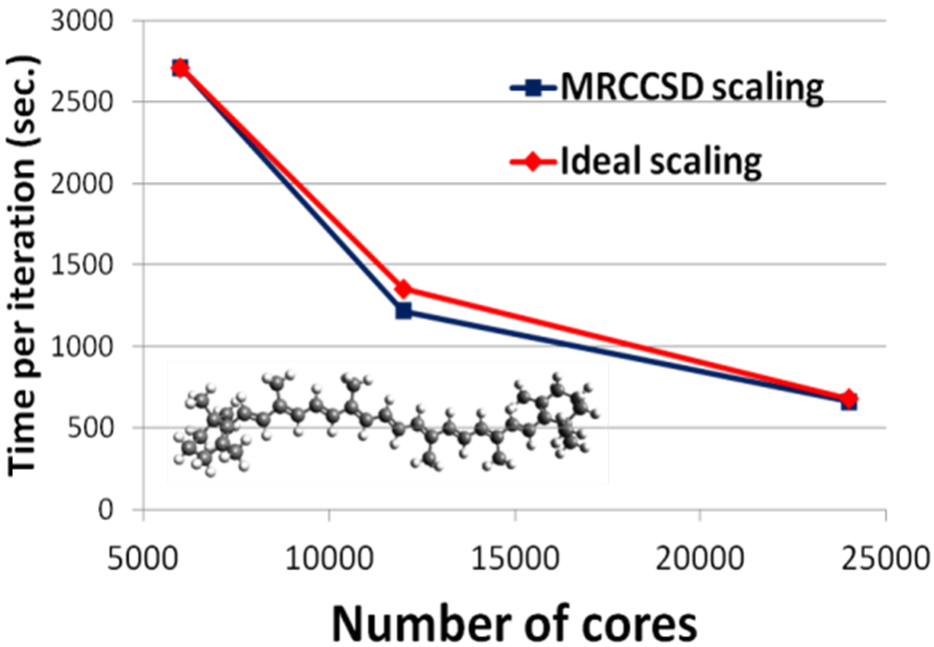
- The reference level parallelism can be applied in:
  - Solving coupled reference-dependent MRCC iterative equations
  - Build efficient parallel schemes for non-iterative MRCC methods

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Processor groups (PGs) and reference level parallelism

- Scalability of the BW-MRCCSD methods for β-carotene in 6-31G basis set (~470 basis set functions); (4,4) complete model space model space (20 reference functions) was used





JCTC Journal of Chemical Theory and Computation — Article
pubs.acs.org/JCTC

Parallel Implementation of Multireference Coupled-Cluster Theories Based on the Reference-Level Parallelism

Jiří Brabec,[†] Jiří Pittner,[†] Hubertus J. J. van Dam,[‡] Edoardo Aprà,[‡] and Karol Kowalski*[,‡]

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Processor groups and reference level parallelism

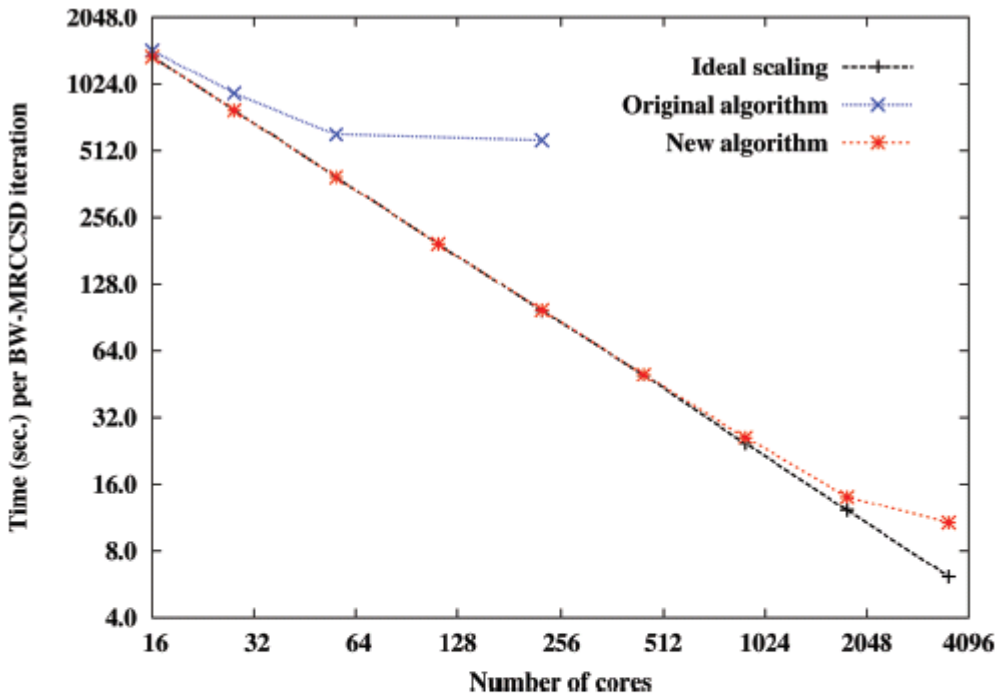


Figure 3. The scalability of the BW-MRCCSD/CMS(6,6) (112 reference functions; $M = 112$) calculations for the N_2 molecule in the aug-cc-pVTZ basis set (two lowest occupied and two highest unoccupied orbitals were kept frozen in the calculations). The calculations were performed using 16, 28, 56, 112, 224, 448, 896, 1792, and 3584 cores using the following processor groups: $PG(7,16,1)$, $PG(4,28,1)$, $PG(2,56,1)$, $PG(1,112,1)$, $PG(1,112,2)$, $PG(1,112,4)$, $PG(1,112,8)$, $PG(1,112,16)$, and $PG(1,112,32)$, respectively (see text for details).

JCTC Journal of Chemical Theory and Computation Article
pubs.acs.org/JCTC

Parallel Implementation of Multireference Coupled-Cluster Theories Based on the Reference-Level Parallelism

Jiří Brabec,[†] Jiří Pittner,[†] Hubertus J. J. van Dam,[‡] Edoardo Aprà,[‡] and Karol Kowalski*[,‡]

Pacific Northwest NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY
Proudly Operated by Battelle Since 1965

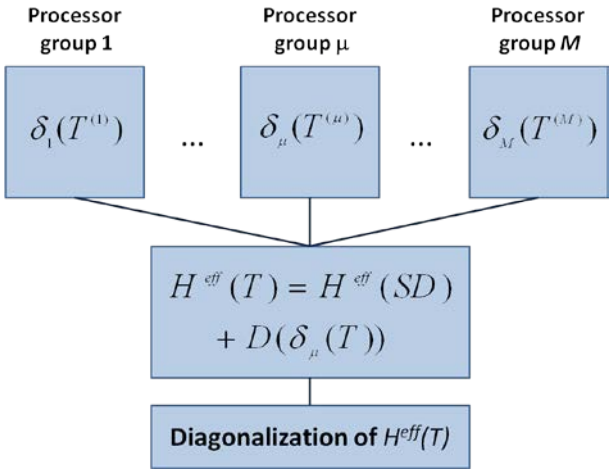
When triple excitations are needed: MRCCSD(T)

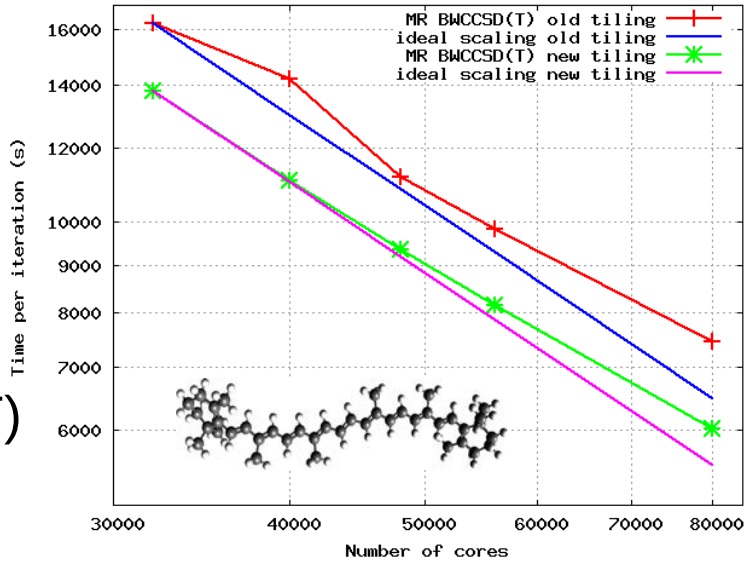
- Improve the quality of the MRCCSD approaches
- Counteract the intruder-state problem

$$H^{eff}_{\mu\mu}(T) = H^{eff}_{\mu\mu}(SD) + \delta_\mu(T^{(\mu)})$$

Numerical complexity $\sim M \times N^7$

Scalability $\sim M \times$ *(scalability of the CCSD(T) approach)*





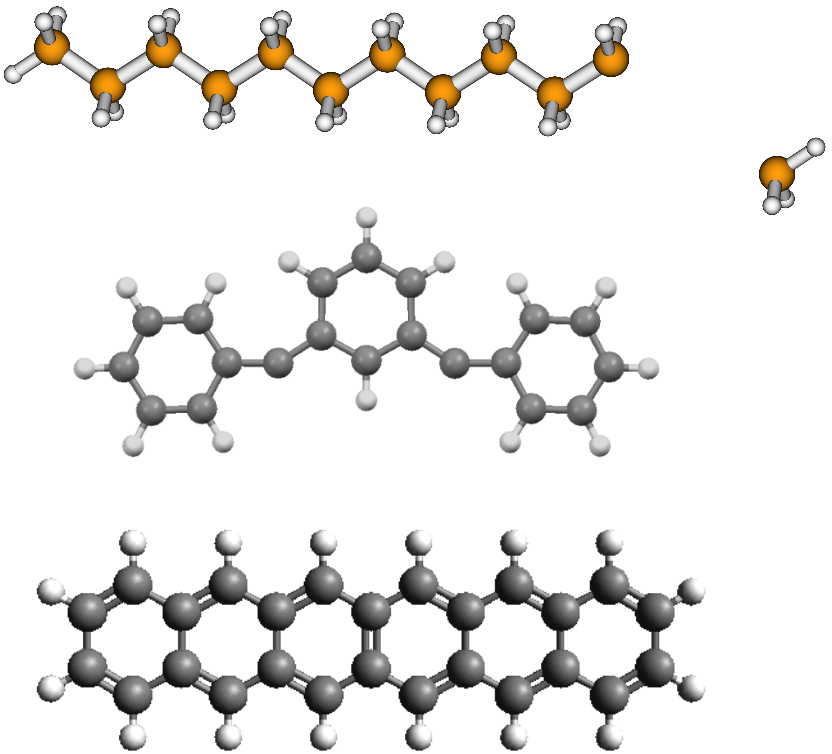
GPU implementations of the MRCCSD(T) approaches are currently tested

K. Bhaskaran-Nair, J. Brabec, J. Pittner, H.J.J. van Dam, E. Apra, K. Kowalski, JCP (accepted).

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Recent MRCC applications

- The MRCC capabilities can be applied to real science cases
  - Bond breaking
  - Open-shell singlet states in organic magnetic materials
  - Low-high spin splittings
  - Excited states: singlet fission (multielectron excitations)



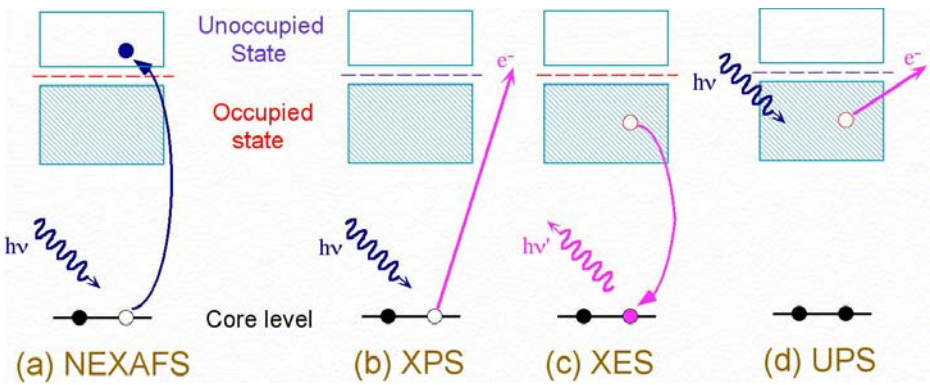
J. Brabec, K. Bhaskaran Nair, K. Kowalski, J. Pittner, H.J.J. van Dam, Chem. Phys. Lett. DOI:10.1016/j.cplett.2012.05.064 (in press)

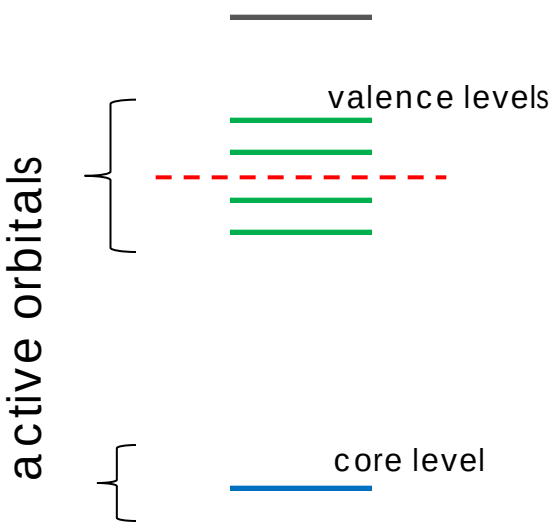
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

MRCC methods for core-level spectroscopies

- Core-level excited states:
  - How to correlate both ground- and core-level states
  - MRCC offers a possibility of balanced description of correlation effects by proper choice of the model space:

$$M_0 = M_0(g) \cup M_0(c)$$





System/Excitation	Basis set	MR-BWCCSD	Expt.	Error
H_2O: O(1s) → (σ* a_1)	cc-pVTZ	534.3	534.0	0.3
H_2O: O(1s) → (σ* b_1)	cc-pVTZ	536.2	535.9	0.3
CO: O(1s) → (π* b_1)	cc-pVTZ	534.2	534.2	0.0
CO: C(1s) → (π* b_1)	cc-pVTZ	287.3	287.4	-0.1
SiH_4: Si(1s) → (σ*)	cc-pVTZ	1846.1	1842.5	3.6

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Planned CC developments

- Development of new programming models for the CC theories: New TCE is needed to properly address algorithmic challenges posed by new generation of first-principle methods
  - Heterogeneous computer architectures
  - Very large model spaces in the MRCC methods
  - Calculation of analytical gradients for high-level EOMCC approaches
- Going across spatial scales: using currently existing CC as drivers for integrated approaches
  - Local coupled cluster methods (adaptive formulations)
  - Dynamic Mean Field Theories for material science
- Building the synergy between theory, high-performance computing, and experiment
  - Highly-accurate excited state simulations: dynamical processes involving valence excited states, core-level spectroscopy

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

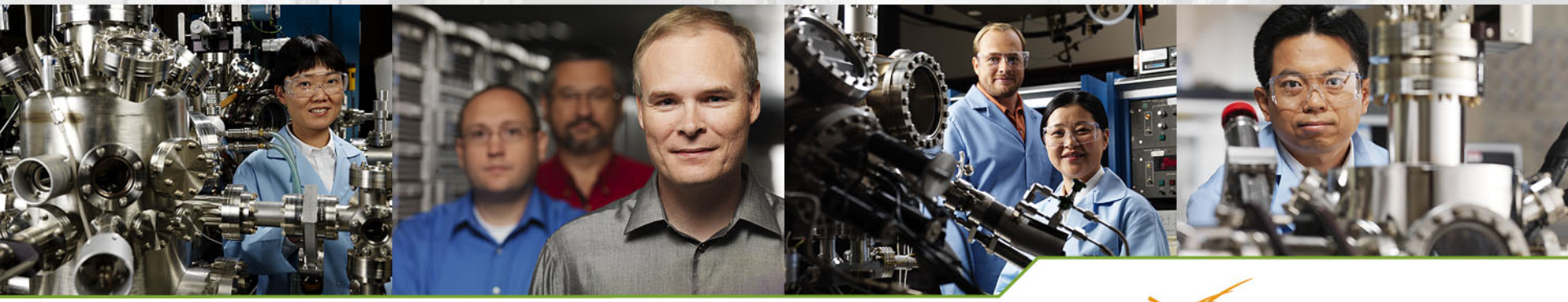
Acknowledgment

- Jirka Brabec, Jiri Pittner, , Huub van Dam (development of the iterative MRCC codes)
- Kiran Bhaskaran-Nair (development of the MRCCSD(T) methods, GPU implementations of the MRCCSD(T) methods)
- Sriram Krishnamoorthy (new algorithms for iterative CC methods)
- Oreste Villa, Sriram Krishnamoorthy (GPU)
- Niri Govind, Ken Lopata (TDDFT,excited states)
- Eduardo Apra (scalability tests on Jaguar)

LDRD project – eXtreme Scale Computing Initiative
NWChem (EMSL/BER)

EMSL

THANK YOU



www.emsl.pnl.gov

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

U.S. DEPARTMENT OF ENERGY