

NWChem: Coupled Cluster Method (Tensor Contraction Engine)



Pacific Northwest NATIONAL LABORATORY

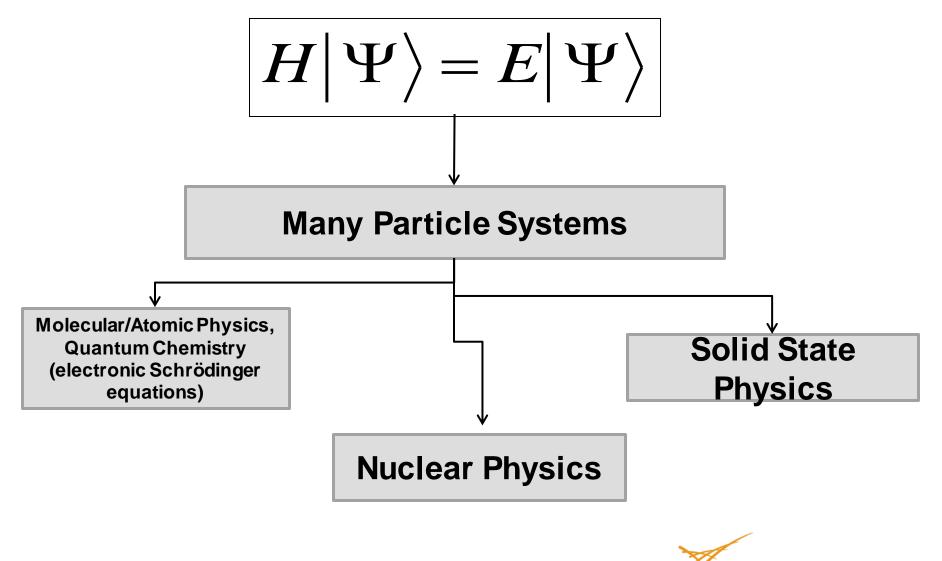


Proudly Operated by Battelle Since 1965

www.emsl.pnl.gov

What we want to solve







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

 C_2 molecule:

12 electrons, 100 orbitals :

#FCI config. $\approx 10^{17}$!!!

Efficient approximations are needed 📈



Approximate wavefunction (WF) methods



- Hartree-Fock method (single determinant)
 - E_{HF} is used to define the correlation energy ΔE

$\Delta E=E-E_{HF}$

In molecules E_{HF} accounts for 99% of total energy but without ΔE making any reliable predictions is impossible

- Correlated methods (going beyond single determinant description)
 - Configuration interaction method (linear parametrization of WF)
 - Perturbative methods (MBPT-n)
 - Coupled Cluster methods
 - and many other approaches





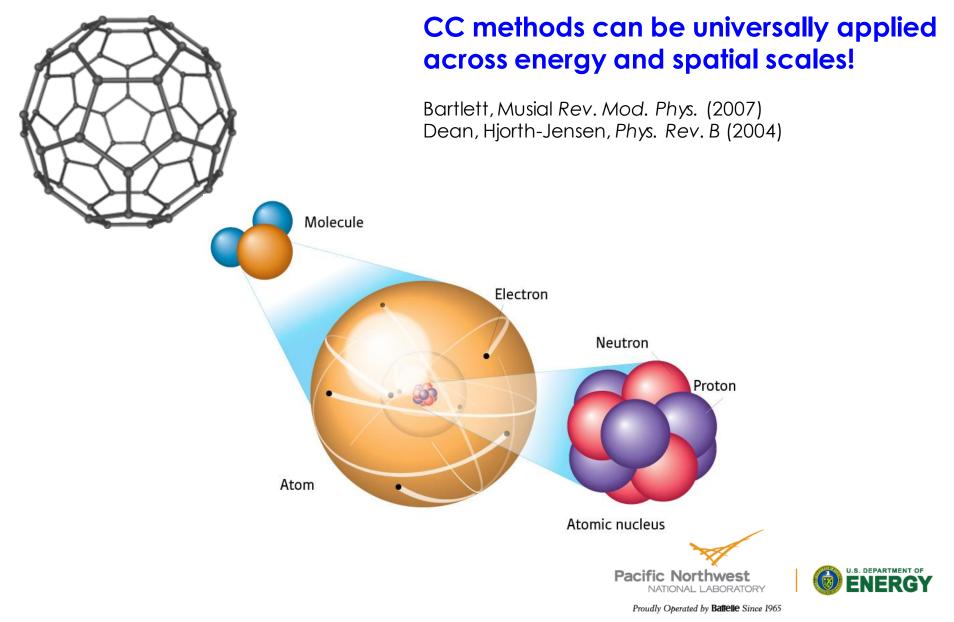
- Correlation effects are important!
- CC is size-extensive theory: can be used to describe dissociation processes.
- Higher-order effects can be approximated by products of lower rank cluster amplitudes.
- Strong ties with many body perturbation theory (MBPT). Effective perturbative methods (CCSD(T)) can be constructed in order to encapsulate higher-order correlation effects.
- Exact energy limit exists full coupled cluster approach.
- Can be applied across energy and spatial scales: from nuclear structure theory to molecular nanosystems





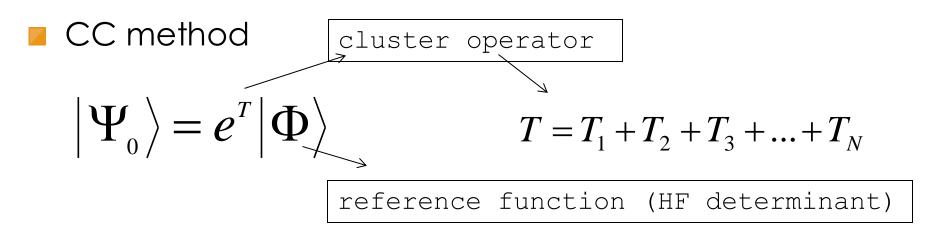
CC methods: across the energy and spatial scales





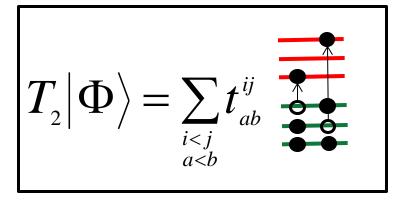
CC method





$$T_{n} = \frac{1}{(n!)^{2}} \sum_{\substack{i_{1}...i_{n} \\ a_{1}...a_{n}}} t_{i_{1}...i_{n}}^{a_{1}...a_{n}} a_{a_{1}}^{+} ... a_{a_{n}}^{+} a_{i_{n}} ... a_{i_{1}}^{-}$$

cluster amplitudes







CI (configuration interaction) and CC methods



Full CI and full CC expansions are equivalent (and this is the only case when CI=CC)

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

$$\begin{split} C_1 &= T_1 \\ C_2 &= T_2 + \frac{1}{2} T_1^{\ 2} \\ C_3 &= T_3 + T_1 T_2 + \frac{1}{6} T_1^{\ 3} \\ C_4 &= T_4 + T_1 T_3 + \frac{1}{2} T_2^{\ 2} + \frac{1}{2} T_1^{\ 2} T_2 + \frac{1}{24} T_1^{\ 4} \end{split}$$

CI amplitudes are calculated from the variational principle while the cluster amplitudes are obtained from projective methods





CC formalism



Working equations:

$$e^{-T} \mid \left\langle He^{T} | \Phi \right\rangle = Ee^{T} | \Phi \rangle$$
$$e^{-T} He^{T} | \Phi \rangle = E | \Phi \rangle$$

$$\left\langle \Phi_{ijk...}^{abc...} \left| (He^T)_c \right| \Phi \right\rangle = 0$$

 $\langle \Phi | (He^T)_c | \Phi \rangle = E$

- CC eqs. are energy independent
- Connected diagrams only: CC theory can properly describe dissociation processes - energy is a sum of energies in the non-interacting system limit



Approximations: CCD,CCSD,CCSDT ...



- CC with doubles (CCD): $T \approx T_2$
- CC with singles and doubles (CCSD): $T \approx T_1 + T_2$
- CC with singles, doubles, and triples (CCSDT): $T \approx T_1 + T_2 + T_3$
- CC with singles, doubles, triples, quadruples (CCSDTQ): $T \approx T_1 + T_2 + T_3 + T_4$



Performance of the CC method



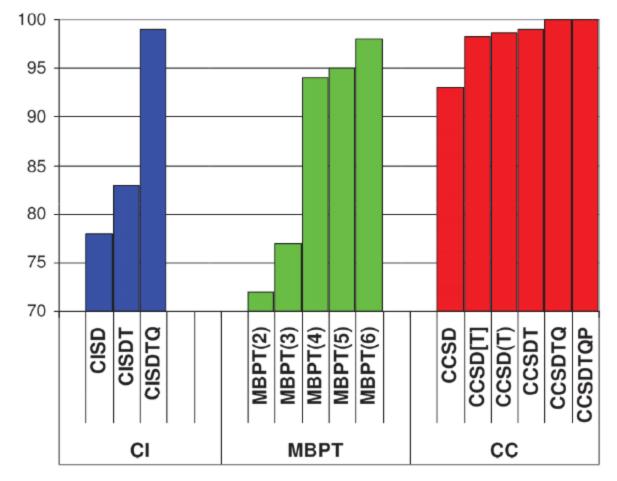


Figure 1. Comparison of CI, MBPT, and CC results with full CI. Results averaged over the series of small molecules: BH, HF, H₂O, SiH₂ and CH₂ in DZP basis set at R_e , $1.5R_e$, $2.0R_e$ and N₂ in cc-pVDZ basis set at R_e and C₂ in cc-pVDZ basis set augmented by diffuse functions at R_e .

R.J. Bartlett Mol. Phys. 108, 2905 (2010).

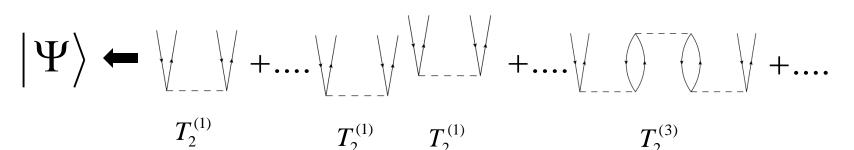


CC and Perturbation Theory (Linked Cluster Theorem)



- Linked Cluster Theorem states:
 - Perturbative expansion for the energy is expressed in terms of closed (having no external lines) connected diagrams only
 - Perturbative expansion for the wavefunction is expressed in terms of linked diagrams (having no disconnected closed part) only

$$\Delta E \leftarrow () + \dots$$



Cluster operator *T* is represented by connected diagrams only





CC and Perturbation Theory

Enable us to categorize the importance of particular cluster amplitudes

$$T_{1} = T_{1}^{(2)} + \dots$$
$$T_{2} = T_{2}^{(1)} + \dots$$
$$T_{3} = T_{3}^{(2)} + \dots$$
$$T_{4} = T_{4}^{(3)} + \dots$$

Enable us to express higher-order contributions through lower-order contribution (CCSD(T))

$$T_{3} \approx R_{3}^{(0)} V_{N} T_{2}$$

$$E^{CCSD(T)} = E^{CCSD} + E^{[4]} + E^{[5]}$$





FN

CCSD(T) method



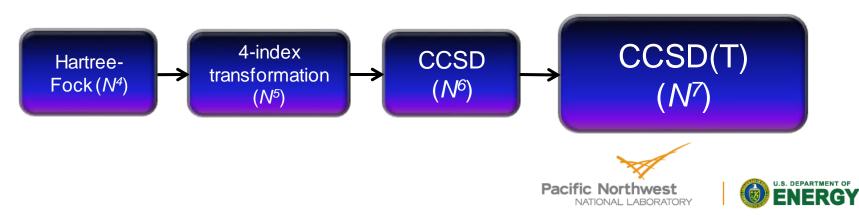
- Driving force of modern computational chemistry (ground-state problems)
- Belongs to the class of non-iterative methods
- Enable to reduce the cost of the inclusion of triple excitations to : required triply excited amplitudes can be generated on-the-fly.
- Storage requirements as in the CCSD approach



Coupled Cluster method



Method	Numerical complexity
CCSD (singles & doubles)	N ⁶
CCSD(T) (perturbative triples)	N ⁷
CCSDT (singles & doubles & triples)	N⁸
CCSDTQ (singles & doubles &triples & quadruples)	N ¹⁰



Diagrammatica



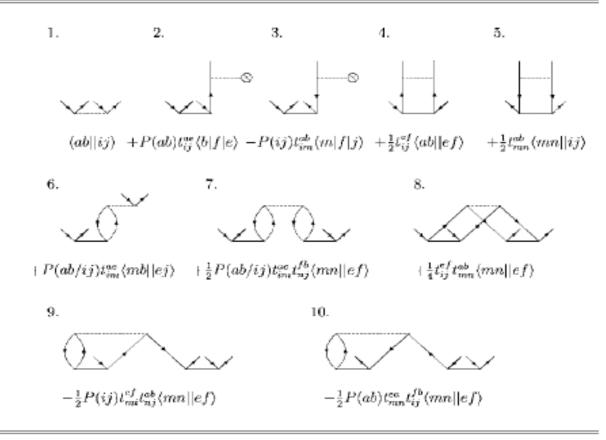


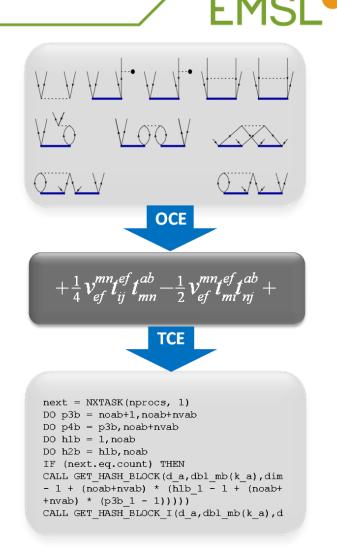
FIG. 10. Diagrammatic representation of the CCD method in the generic spin-orbital form together with the corresponding algebraic expression. Summation over repeated upper and lower indices is assumed. The antisymmetric permutation operator P(pq/rs) is defined as P(pq/rs) = P(pq)(rs) = 1 + (qp)(sr)-(qp)(rs) - (pq)(sr).



rated by Battelle Since 1965

What is Tensor Contraction Engine (TCE)

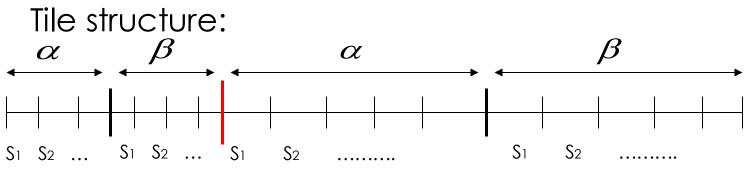
- Symbolic manipulation & program generator
 - Automates the derivation of complex working equations based on a well-defined second quantized many-electron theories
 - Synthesizing efficient parallel computer programs on the basis of these equations.
- Granularity of the parallel CC TCE codes is provided by the so-called tiles, which define the partitioning of the whole spinorbital domain.







What is Tensor Contraction Engine (TCE)



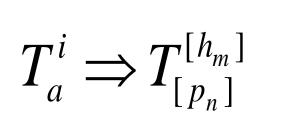
Occupied spinorbitals

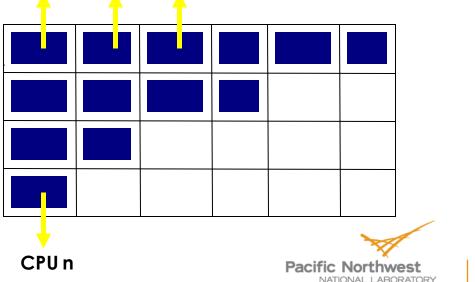
unoccupied spinorbitals

CPU2 CPU3

Tile-induced block structure of the CC tensors:

CPU1





.



CC TCE calculations



- Closed- & open-shell CC calculations with RHF/ROHF/UHF references
- Many-body perturbation theory
- CI methods: CISDT, CISDTQ, ...
- Ground-state methodologies: CCSD, CCSD(T), CCSDT, ...
- Excited-state methods: EOMCCSD,CC2, CR-EOMCCSD(T), EOMCCSDt, EOMCCSDT
- Linear response CC methods for calculating molecular properties: static & dynamic CCSD polarizabilities, static CCSDT polarizabilities, static CCSD hyperpolarizabilities





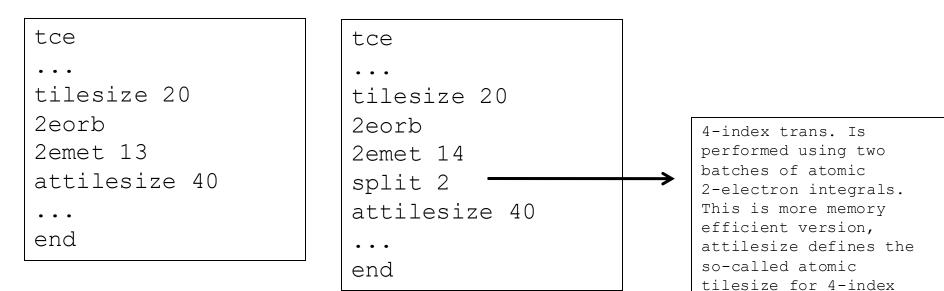
Three types of references can be used in singlereference TCE CC calculations: RHF, ROHF, UHF

scf	scf	scf
thresh 1.0e-10	thresh 1.0e-10	thresh 1.0e-10
tol2e 1.0e-10	tol2e 1.0e-10	tol2e 1.0e-10
singlet	doublet	singlet
maxiter 100	maxiter 100	maxiter 100
rhf	rohf	uhf
end	end	end



How to choose 4-index transformation?

- RHF/ROHF references
 - Default: spinorbital 4-index tarnsformation
 - Alternatives: orbital 4-index transformations



Always: tilesize <= attilesize; available in the GA version only

UHF reference: default spinorbital 4-index transformation will be executed.



trans.



EMS

Local memory management in CC TCE module



- Approaches based on the single and double excitations (CCSD,EOMCCSD,LR-CCSD)
 - ~ (tilesize)⁴
- Perturbative CCSD(T) & CR-EOMCCSD(T) methods
 - 2 * (tilesize)⁶ # choose tilesize wisely
- Iterative CCSDt,CCSDT,EOMCCSDt,EOMCCSDT methods
 - 4 * (tilesize)⁶







Example: CCSD calculation

```
Example: h2o dimer ccsd aug cc pvdz.nw
   scf
   thresh 1.0e-10
   tol2e 1.0e-10
   singlet
   rhf
   end
   tce
   freeze atomic
   ccsd
   maxiter 100
                                  max. number of iterations
   tilesize 15
                            length of the diis cycle
   diis 5
   thresh 1.0d-5
                                    conv. threshold
   2eorb
   2emet 13
                            default value 40
   attilesize 40
   end
                                                                        U.S. DEPARTMENT C
   task tce energy
                                                    Pacific Northwest
                                                      NATIONAL LABORATORY
```

FMS

Examples: open-shell CCSD(T) calculation



Example: cnh2o ccsd t cc pvdz.nw scf thresh 1.0e-10tol2e 1.0e-10 doublet rohf end tce freeze atomic ccsd(t)CCSD(T) calculation will be performed maxiter 100 tilesize 15 diis 5 Level shifting may be helpful in lshift 0.2 converging open-shell CCSD equations thresh 1.0d-5 2eorb 2emet 13 attilesize 40 end





Challenging situations – bond breaking processes: renormalized methods



```
Example: tce_cr_ccsd_t_ozone_pol1.nw
```

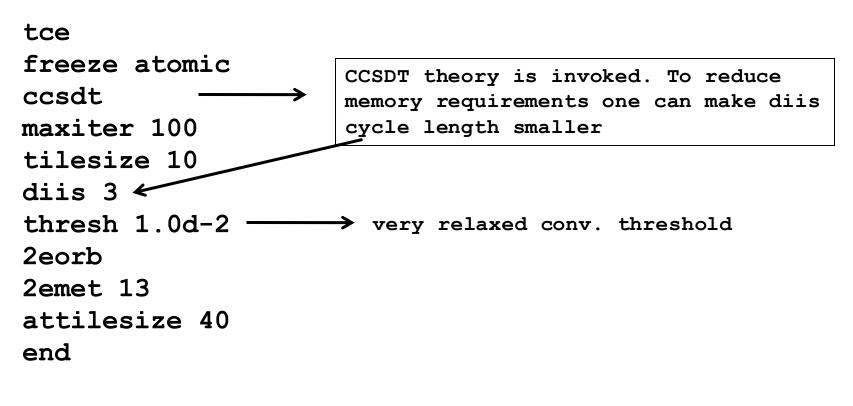
in single-bond breaking/forming processes renormalized methods may provide better description of ground-state potential energy surfaces



Examples: CCSDT calculations



Example: h2o_dimer_ccsdt_cc_pvdz.nw



task tce energy

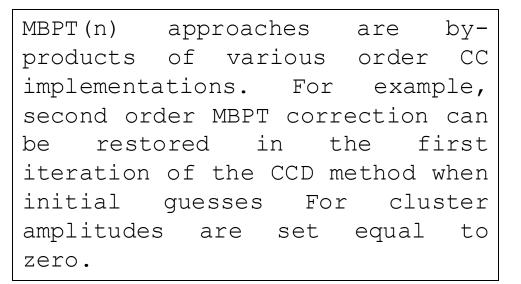


Example: mbpt2_h2o.nw

scf

```
thresh 1.0e-10
tol2e 1.0e-10
singlet
rhf
end
tce
mbpt2
end
```

task tce energy







FMS

Examples: MBPT calculations – towards higher orders



Example: mbpt4sdq_h2o.nw

MBPT(2) < MBPT(3) < MBPT(4, SDQ) < MBPT(4)

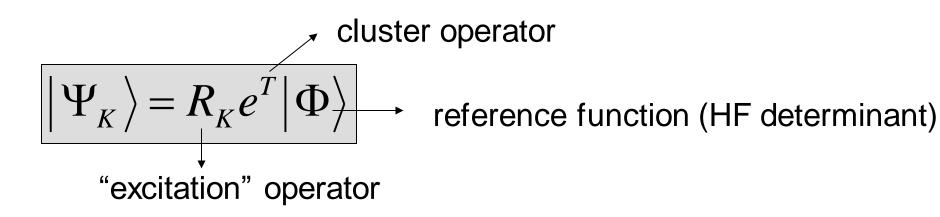
tce mbpt4(sdq) — end

MBPT(3)	+	SE	elec	t€	ed	MBPT(4)
contributions; an approximate						
method	to	inc	lude	Э	the	effect
of	qua	druj	ply			excited
configurations						

task tce energy







$$\overline{H}R_{K}|\Phi\rangle = E_{K}R_{K}|\Phi\rangle$$

$$\overline{H} = e^{-T} H e^{T}$$

(Equation of Motion Coupled Cluster Equations)





EOMCCSD: singly-excited states

$$\Psi_{K}^{EOMCCSD} \rangle = (R_{K,o} + R_{K,1} + R_{K,2})e^{T_{1}+T_{2}} |\Phi\rangle$$

EOMCCSDT: singly and doubly excited states

$$\left|\Psi_{K}^{EOMCCSDT}\right\rangle = (R_{K,o} + R_{K,1} + R_{K,2} + R_{K,3})e^{T_{1}+T_{2}+T_{3}}\left|\Phi\right\rangle$$

Perturbative methods: EOMCCSD(T) formulations





EOMCCSD < CR-EOMCCSD(T) < EOMCCSDT < EOMCCSDTQ

Method	Numerical complexity	
EOMCCSD (singles & doubles)	N ⁶	Excitation energies of singly Excited states
CR-EOMCCSD(T) (perturbative triples)	N7	
EOMCCSDT (singles & doubles & triples)	N ⁸	Excited-state potential energy surfaces, doubly excited state
EOMCCSDTQ (singles & doubles & triples & quadruples)	N ¹⁰	



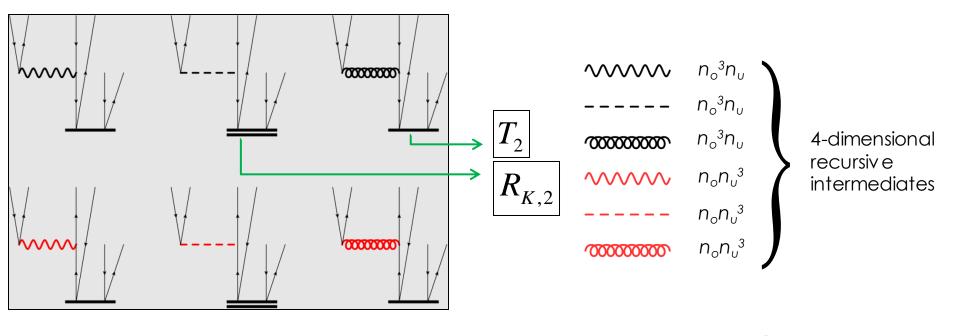


Non-iterative EOMCC methods: CR-EOMCCSD(T) approach



$$\omega_{K}^{CR-EOMCCSD(T)} = \omega_{K}^{EOMCCSD} + \delta_{K}(T) \leftarrow \delta_{K}(T) = \frac{\langle \Psi_{K} | Q_{3}M_{K,3} | \Phi \rangle}{\langle \Psi_{K} | e^{T_{1}+T_{2}}(R_{K,0} + R_{K,1} + R_{K,2}) | \Phi \rangle}$$

Kowalski and Piecuch "A new class of non-iterative correction for the excited states: completely renormalized EOMCCSD(T) approach" J. Chem. Phys. **120**, 1715 (2004)







Reduced cost CR-EOMCCSD(T) methods

- Active-space CR-EOMCCSD(T) approach
 - Active space can be defined by orbitals with orbital energies falling into a predefined energy interval $[\alpha, \beta]$
 - Significant reduction of the numerical cost associated with calculating (T) correction: $N^7 => N_{act}{}^5N^2$: Applicable to large molecular systems

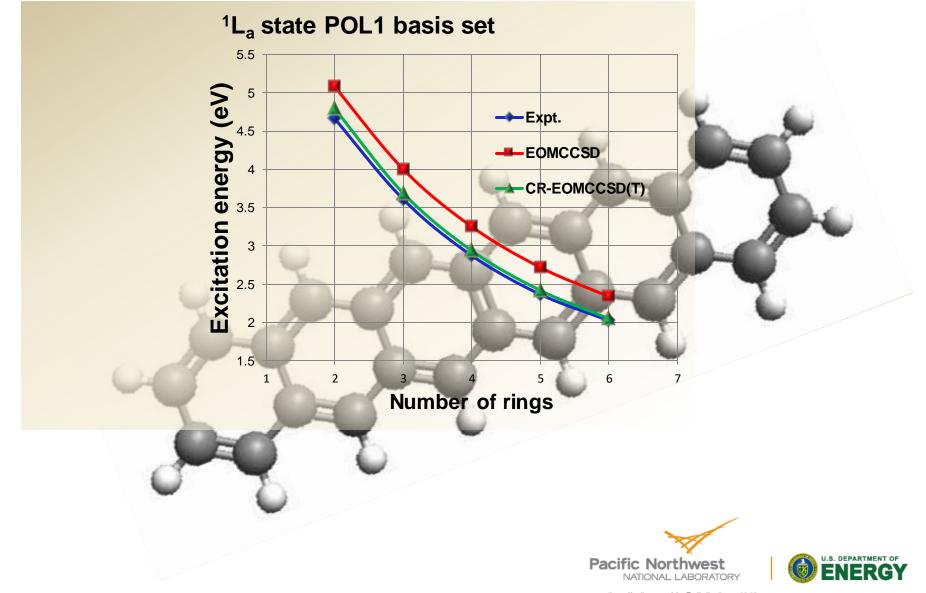
Double excitations are allowed only within the active space

$$\begin{array}{c|c} & & & \\ \hline T_1, t_2 \\ \hline R_{K,1}, r_{K,2} \end{array}$$

$$\begin{aligned} & \text{Active-space CR-EOMCCSD(T) correction} \\ & \delta_{K}(t) = \frac{\left\langle \Psi_{K} \left| q_{3} M_{K,3} \right| \Phi \right\rangle}{\left\langle \Psi_{K} \left| e^{T_{1}+t_{2}} \left(R_{K,0} + R_{K,1} + r_{K,2} \right) \right| \Phi \right\rangle} \\ & \left[q_{3} = \sum_{\substack{i < J < K \\ A < B < c}} \left| \Phi_{iJK}^{ABc} \right\rangle \right\rangle \\ \end{aligned}$$



Illustrative examples of large-scale excited-state calculations – components of light harvesting systems EMSL



Excited-state calculations: EOMCCSD

							-	
Freeze le .		EOM-CCSD right-hand	d side iter	ations				
Example: h2o dimer eomccs	d aug og pydz n	Residuum	Omega / h	artree	Omega / eV	Cpu	Wall	
		····						
		 Iteration 29 using	48 trial	vectors				
tce		0.0000082390224	0.2870037		7.80977			
		0.0000084487979	0.3499939	129169	9.52382	2.0	2.3	
freeze atomic	;	Ttorationa convora						
ccsd	-	Iterations converged largest EOMCCSD amplitudes: R1 and R2						
		Singles contributions						
tilesize 20		11a' (alpha)		ι ₁ ,		571088259	Э	
		11a' (alpha)		(alpha)		193546713		
diis 5		13a' (alpha) 14a' (alpha)		(alpha)		324980230 L54368698		
thresh 1.0d-5		14a' (alpha) 15a' (alpha)		(alpha) (alpha)		59219332 [°]		
		16a' (alpha)		(alpha)		331210023		
2eorb		16a' (alpha)	9a'	(alpha)	-0.33	310076628	3	
2emet 13		18a' (alpha)		(alpha)	0.14	1971579	5	
		Doubles contributio	ons					
nroots 2>	number of roots							
eomsol 1 →	"old" eigensolvers	(default option)	- reauires	5				
end	more memory but	· · · ·						
end	1							
	doubly excited sto	lies						
-								

task tce energy

Pacific Northwest NATIONAL LABORATORY



FMS

Excited-state calculations: EOMCCSD



```
Example:
h2o dimer eomccsd aug cc pvdz eomsol2.nw
tce
freeze atomic
ccsd
tilesize 20
                               EOMCCSD
                                             solver
                                                       with
                       new
diis 10
                       improved memory management -
thresh 1.0d-5
2eorb
                       should be used for singly excited
2emet 13
                       states only; initial starts taken from
nroots 1
                       the CIS calculations
eomsol 2
symmetry
                states of a' symmetry will be calculated
targetsym a'
end
```

task tce energy





Excited-state calculations: EOMCCSD



tce freeze atomic ccsd tilesize 20 diis 10 thresh 1.0d-5 2eorb 2emet 13 nroots 2 eomsol 2 symmetry targetsym a' end

Sometimes it is easier to converge n-root in the presence of the n+1 root. We can apply separate conv. criteria for n roots of interest (thresheom) and auxiliary (n+1-st) root (threshl).

set tce:thresheom 1.0d-4
set tce:threshl 1.0d-3

task tce energy



37

Excited-state calculation: EOMCCSDT



Example: tce_h2o_eomccsdt_cc-pvdz.nw

- # CCSDT/EOMCCSDT methods are much more expensive than
- # the CCSD/EOMCCSD formalisms

tce

freeze core atomic

ccsdt

dipole ------>

```
thresh 1.0d-6
```

nroots 1

end

task tce energy

calculates excited-state dipole moments and transition moments



Excited-state calculations: active-space EOMCCSDT methods (EOMCCSDt)



Example: tce active ccsdt be3.nw # EOMCCSDt uses selected set of triply excited amplitudes # - it makes it less expensive than the full EOMCCSDT approach ! tce freeze atomic ccsdta tilesize 12 thresh 1.0d-4 active oa 3 Definition of the active space: active occupied alpha spinorbitals active ob 3 active occupied beta spinorbitals active virtual alpha spinorbitals active va 9 Aactive virtual beta spinorbitals active vb 9 t3a lvl 2 nroots 1 Only t_{iIK}^{ABc} amplitudes included symmetry targetsym al end

task tce energy

NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Pacific Northwest



Excited-state calculations: CR-EOMCCSD(T)



Example: tce_cr_eom_t_ozone_p	oll.nw				
<pre># Excitation energie # EOMCCSD < CR-EOMCC</pre>	-				
tce freeze atomic					
2eorb 2emet 13	RHF reference is employed, orbital from of 2-electron integrals can be used				
creomsd(t) —— tilesize 15	CR-EOMCCSD(T) calculation is composed of several steps:				
thresh 1.0d-4 nroots 1 symmetry	 (1) CCSD calculation (2) EOMCCSD calculation (3) calculation of the CR-EOMCCSD(T) 				
targetsym b2 end	non-iterative correction				

task tce energy





Excited-state calculations: active-space CR-EOMCCSD(T) method

tce freeze atomic 2eorb 2emet 13 creom(t)ac	
oact 21 uact 99 tilesize 15	Number of active occupied/unoccupied orbitals. Alternative the active orbitals can be defined by "energy window"
thresh 1.0d-4 nroots 1 symmetry targetsym a'	emin_act -0.5 emax_act 1.0
end	

task tce energy



FM^S

Linear response CC calculations



Property calculations with the CC method: The cluster operator T(t) is expanded in order of timedependent perturbation

$$T(t) = T^{(0)} + T^{(1)}(t) + T^{(2)}(t) + \dots$$

Dipole polarizabilities can be calculated as a linear response function

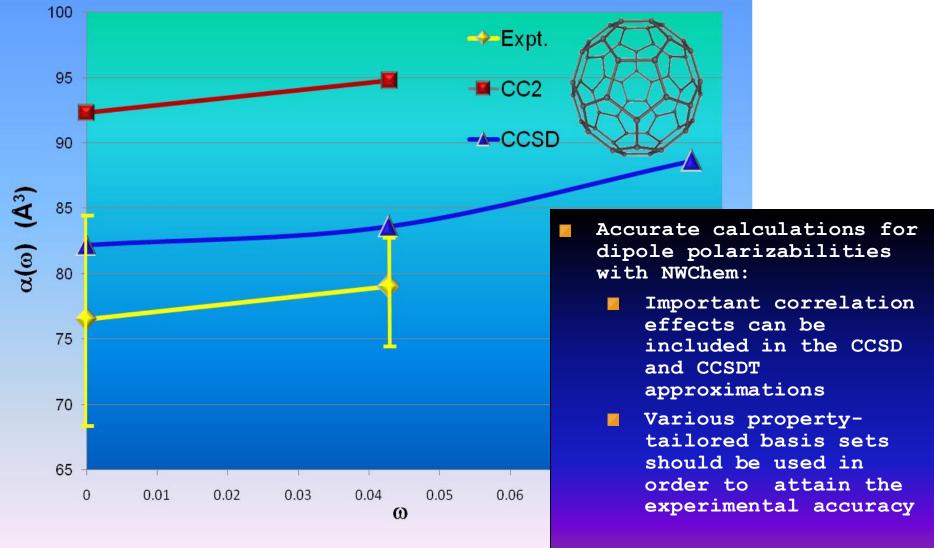
$$\alpha_{ij}(\omega) = - <<\mu_i; \mu_j >>_{\omega}$$

CC eq. $\rightarrow \Lambda$ -CC eq. \rightarrow first-order LR-CC eq.



Linear response CC calculations

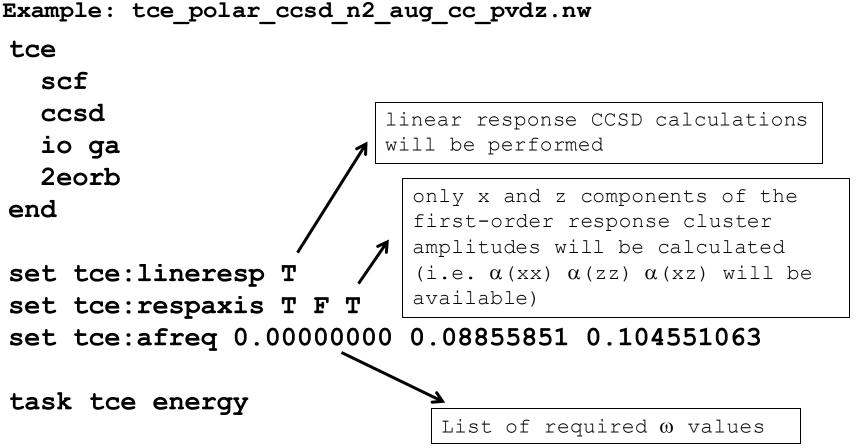






Property calculations: CCSD dipole polarizabilities







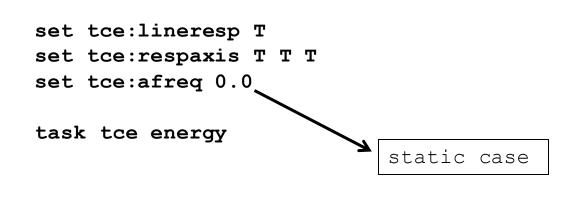
Property calculations: CCSD dipole polarizabilities – benzene



Example: tce_polar_ccsd_benzene.nw

scf

singlet							
rhf			nse polarizab 2000000 / au	ility/au			
thresh 1.0e-10		ength =	Inf / nm				
end	atomic units (bohr^3))	angstroms	^ <u>3</u>	
tce		Х	Ŷ	Ζ	X	Y	Z
scf	 X	68.4305909	0.0000000	0.0000000	10.1403696	0.0000000	0.0000000
freeze core ccsd	Y Z	0.0000000 0.0000000	68.4364388 0.0000000	0.0000000 20.1980487	0.000000 0.000000	10.1412361 0.0000000	0.0000000 2.9930427
io ga	Eigs = Isotrop Anisotr		68.4364388 52.3550261 48.2354664		7 10.140369 	7.14123 7.75821 7.14776	61
2eorb end			40.2004004	• 	I 	/.14//00	JZ







Multireference CC (MRCC) theory



Polycarbene (n=2). Structure of the model space representation of the lowest singlet state (c_{μ}^{2} s). (4,4) model space and 6-31G basis set was used (208 functions).

Occupation/Method	12R-BWCCSD a p.	12R-MkCCSD	
2200	0.000	0.000	1
lpha lpha eta eta	0.654	0.533	Ser YY Ser
$\alpha\beta\alpha\beta + \alpha\beta\beta\alpha$	0.240	0.437	
2020 + 0220 + 2002 + 0202	0.106	0.030	
0022	0.000	0.000	

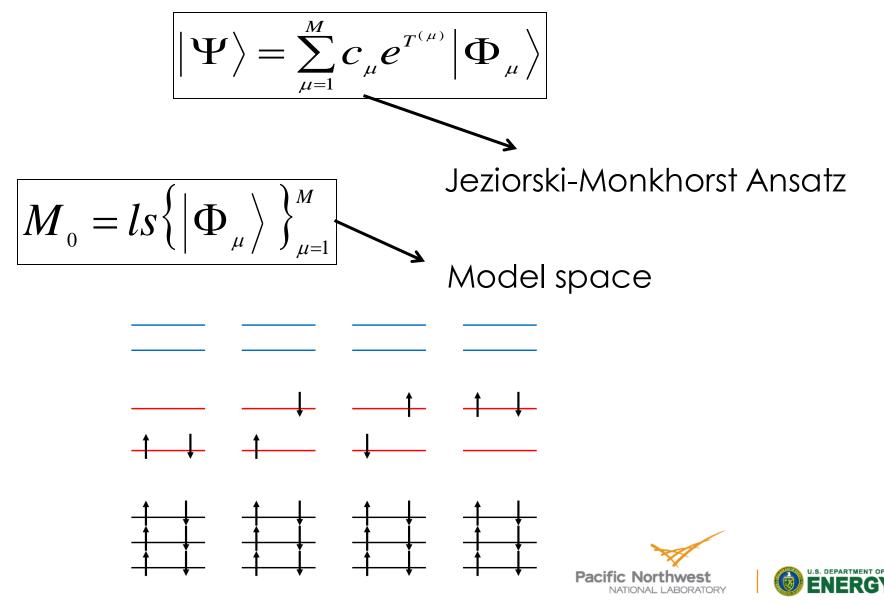
12R-BW-MRCCSD-ap	12R-Mk-MRCCSD	
-766.792581		Closed-shell singlet
-766.884663	-766.883870	Open-shell singlet





MRCC theory



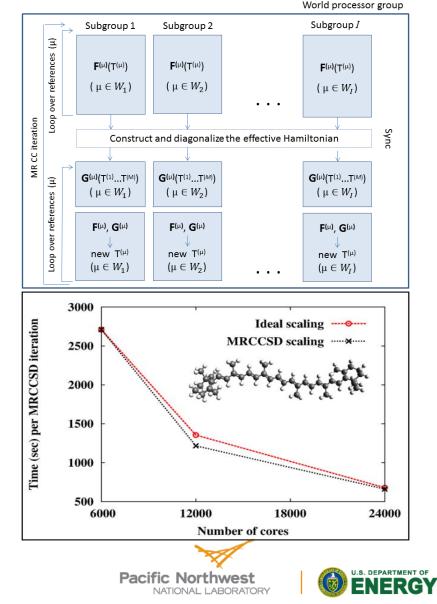


MRCC implementations in NWChem

EMSL

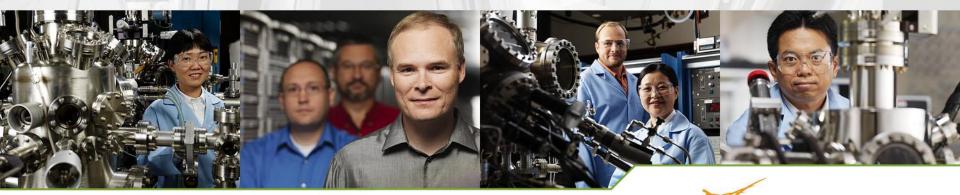
 Multireference CC methods in NWChem (next release)

- Strongly correlated excited states
- Implemented MRCC approaches
 - Brillouin-Wigner MRCCSD
 - Mukherjee Mk-MRCCSD approach
 - State-Universal MRCCSD (under testing)
 - Perturbative triples corrections MRCCSD(T)
- Novel paralellization strategies based on the processor groups
- Demonstrated scalability of MRCCSD across 24,000 cores





Questions?



www.**emsl**.pnl.gov



