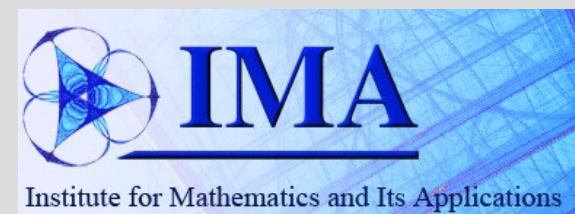
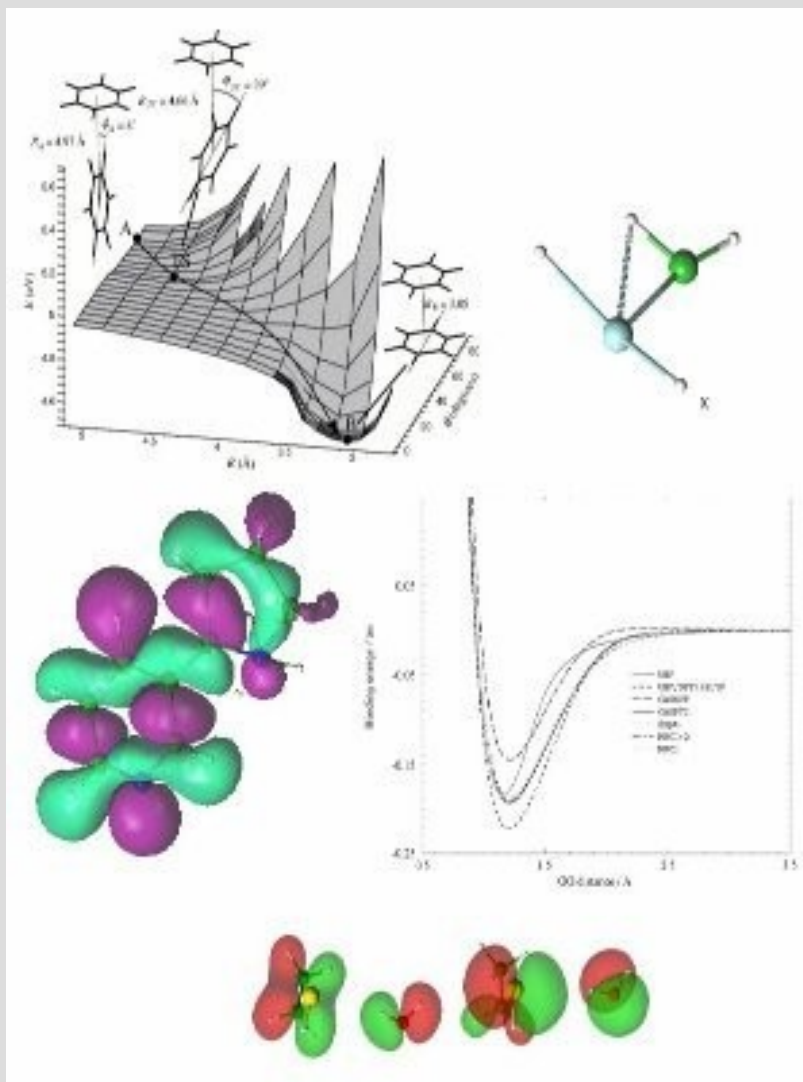




Cholesky Decomposition Techniques in Quantum Chemical Implementations





Outline



- What is MOLCAS?
- A crash course in Cholesky Decomp. (CD)
- LK Exchange
- 1C-CD
- Analytic CD gradients
- aCD on-the-fly RI auxiliary basis set
- AcCD auxiliary basis sets
- Some showcases



What is *MOLCAS*? The hallmark:



a-CASSCF/MS-CASPT2/ANO is and will be our protocol of choice

Typical applications:

- Chemical reactions
- Photo Chemistry
- Heavy Element Chemistry



Chemical reactions:

“Chemiluminescence of 1,2-dioxetane. Mechanism uncovered”

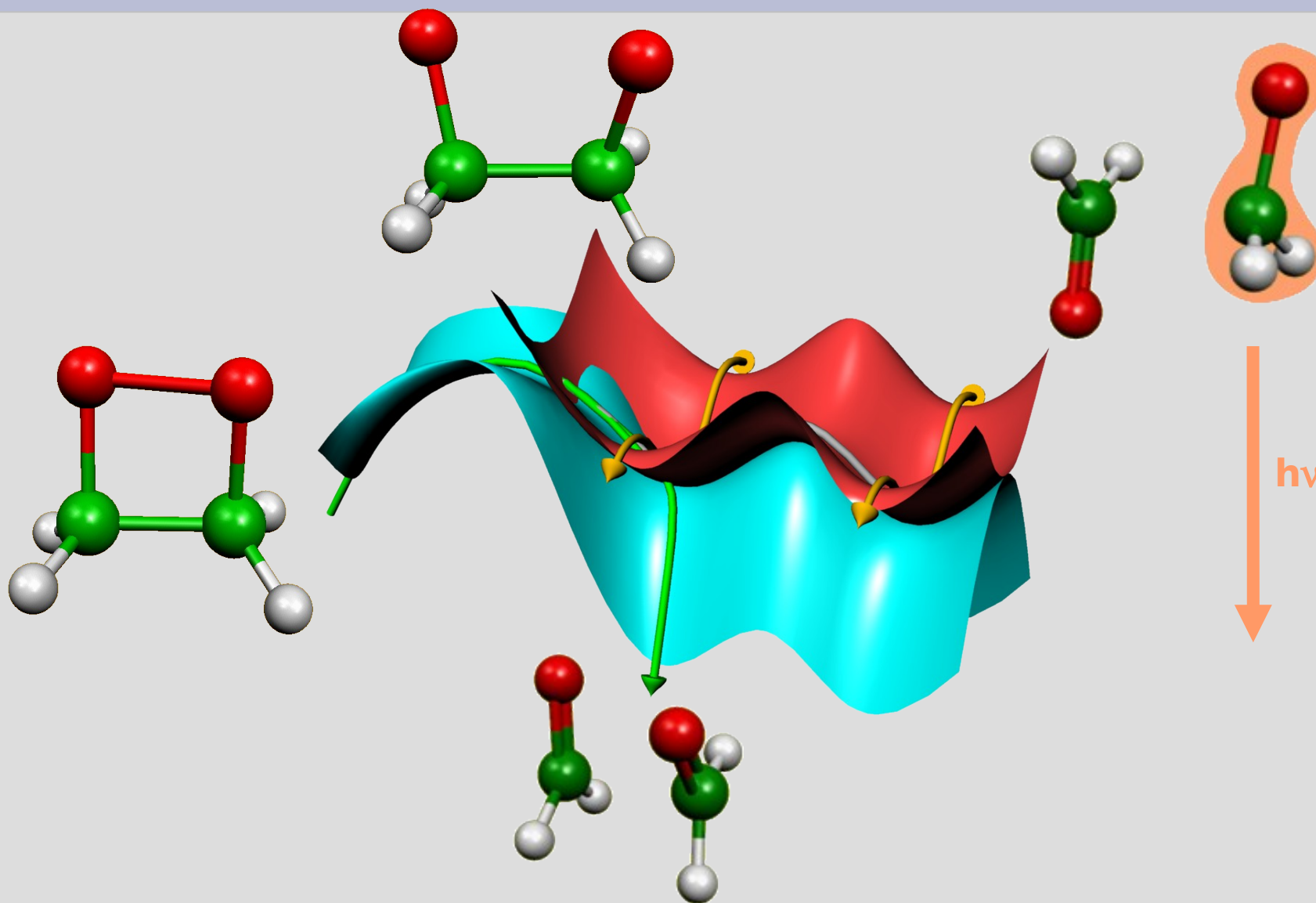
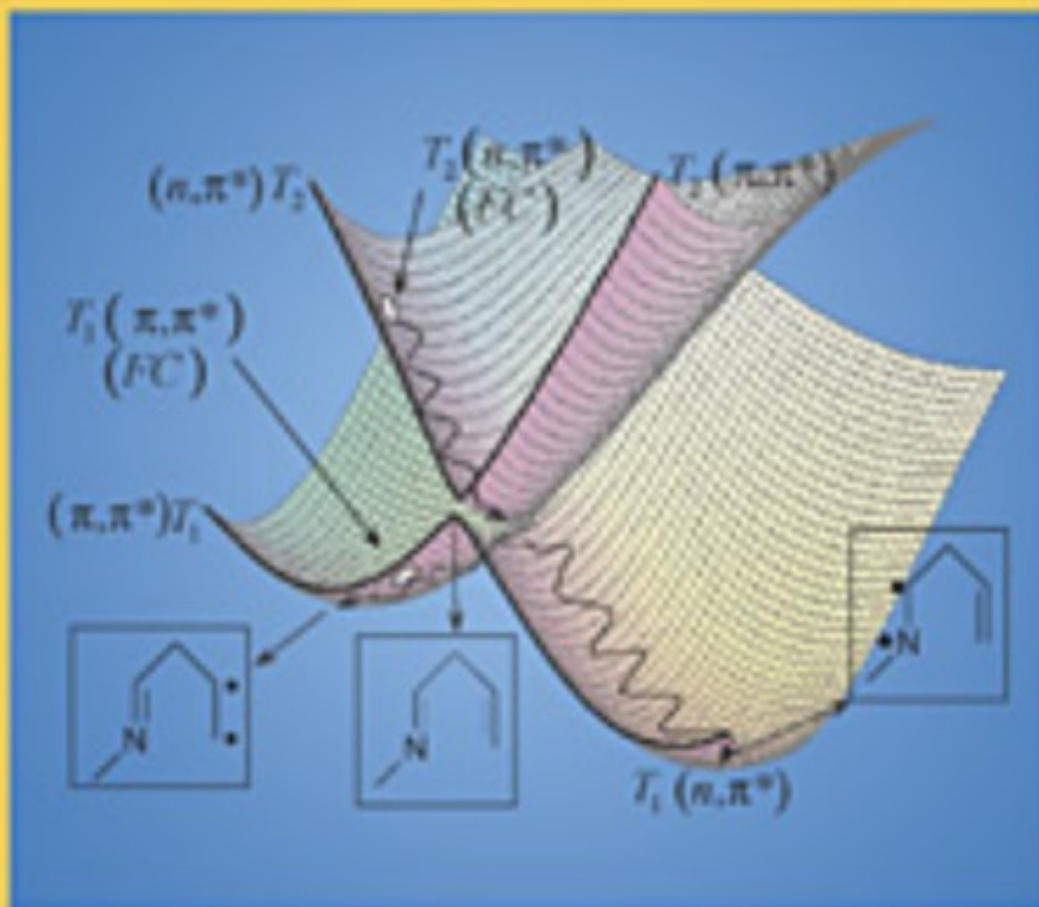




Photo Chemistry:

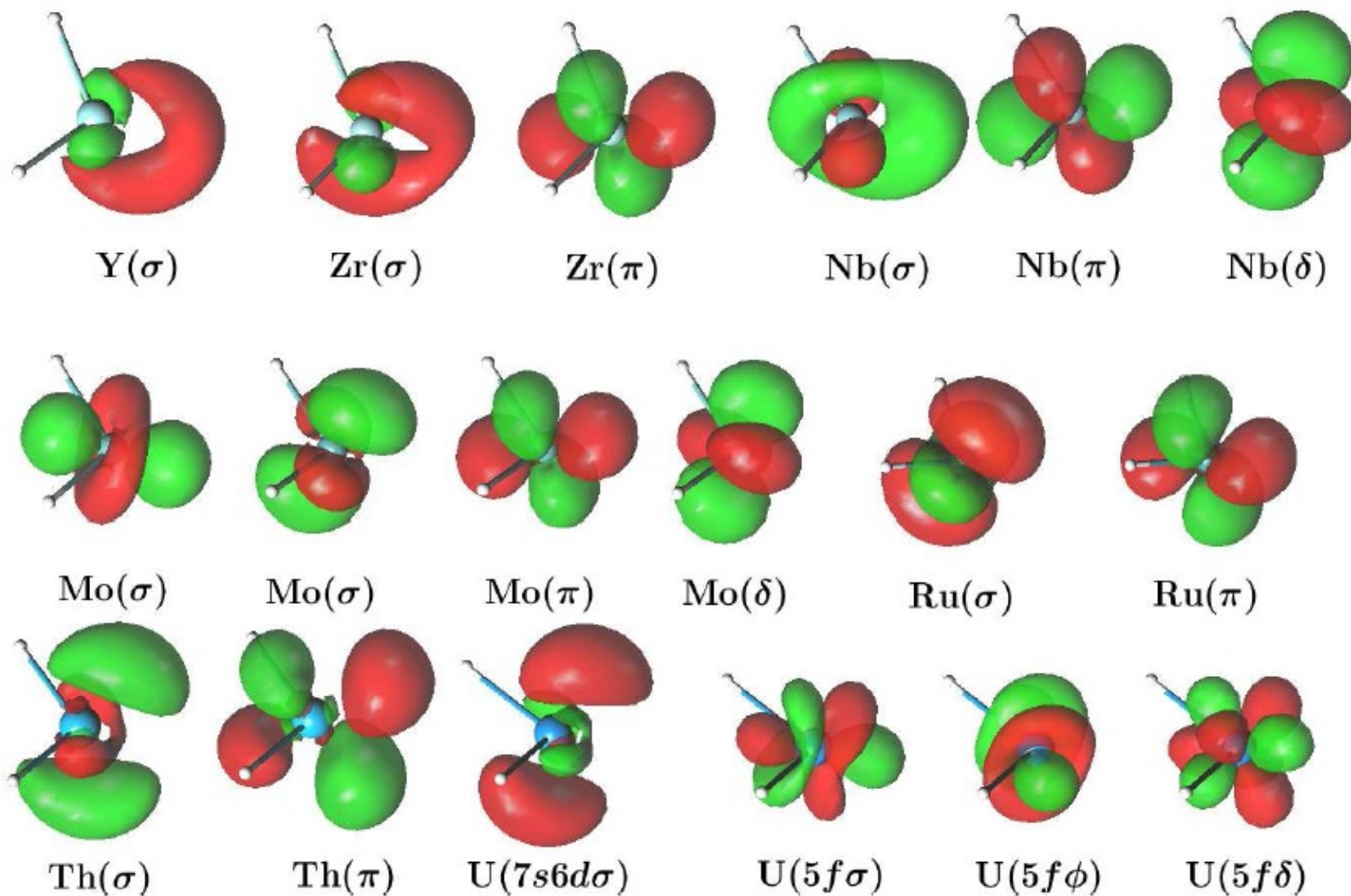
“Intramolecular triplet-triplet energy transfer in oxa- and aza-di-pi-methane photosensitized systems”





Heavy Element Chemistry:

*“Agostic interaction in the
methylidene metal dihydride
complexes H_2MCH_2 ($M = Y, Zr, Nb,$
 $Mo, Ru, Th, or U$)“*





A crash course in CD!



CD technique is trivial and involve no more than elementary vector manipulations.

In an essence CD of 2-electron integrals is a truncated version of a standard Gram-Schmidt orthogonalization in a Coulumbic metric.

Let me demonstrate!



Gram-Schmidt Orthogonalization



The Gram-Schmidt is formulated as

$$|l\rangle = |i\rangle - \sum_{K=1}^{l-1} |K\rangle \langle K|i\rangle$$

Or in matrix form

$$\langle j|l\rangle = \langle j|i\rangle - \sum_{K=1}^{l-1} \langle j|K\rangle \langle K|i\rangle$$

and

$$\langle l|l\rangle = \langle l|i\rangle$$

$$\langle i|l\rangle = \langle i|i\rangle - \sum_{K=1}^{l-1} \langle i|K\rangle \langle K|i\rangle$$



Cont.



Imagine the identity

$$V_{ij} = \sum_{kl} V_{ik} (V^{-1})_{kl} V_{lj}$$

Transform some index to the GS basis

$$V_{ij} = \sum_K V_{iK} (V^{-1/2})_K (V^{-1/2})_K V_{Kj} = \sum_K L_i^K L_j^K$$

Finally expressions

$$L_i^I = (V_{ii} - \sum_{K=1}^{I-1} L_i^K L_i^K)^{-1/2} = L_i^I$$

$$L_j^I = (V_{jj} - \sum_{K=1}^{I-1} L_j^K L_j^K) / L_j^I$$



Truncation - Reduction



The GS procedure lends itself to a single parameter controlled truncation of the GS basis.

A list of all $V_{ii} = V_{ii} - \sum_{k=1}^n V_{ik} V_{ik}$

is stored and updated as we include new GS basis functions. If all remaining V_{ii} are below the threshold then terminate!



DF-RI-CD Unified



In DF-RI-CD the 2-electron integrals are expressed as

$$\langle ij|kl\rangle = \sum_J L_{ij}^J L_{kl}^J$$

$$\langle ij|kl\rangle = V_{ij,kl}$$

For DF/RI we have

$$\langle I|J\rangle = V_{IJ}$$

$$L_{ij}^J = \sum_I \langle ij|I\rangle (V^{-1/2})_{IJ}$$

While for CD we have

$$L_{ij}^J = \langle ij|J\rangle (V_{JJ})^{-1/2}$$

$$L_{JJ} = \left(V_{JJ} - \sum_{K=1}^{J-1} L_{JK}^2 \right)^{1/2}$$
$$L_{IJ} = \left(V_{IJ} - \sum_{K=1}^{J-1} L_{IK} L_{JK} \right) / L_{JJ}$$



Comparison



	DF/RI-version	CD-version
Auxiliary Basis set	External 1-center	Internal (num.) 1- and 2-center
Gradient	Yes	No
Method-dependent?	Yes mostly	No!
Parameter dependent?	Not directly	Yes
Could be exact?	Not automatically	Yes



The LK approach for Exchange



- **CD localization of the occupied orbitals**

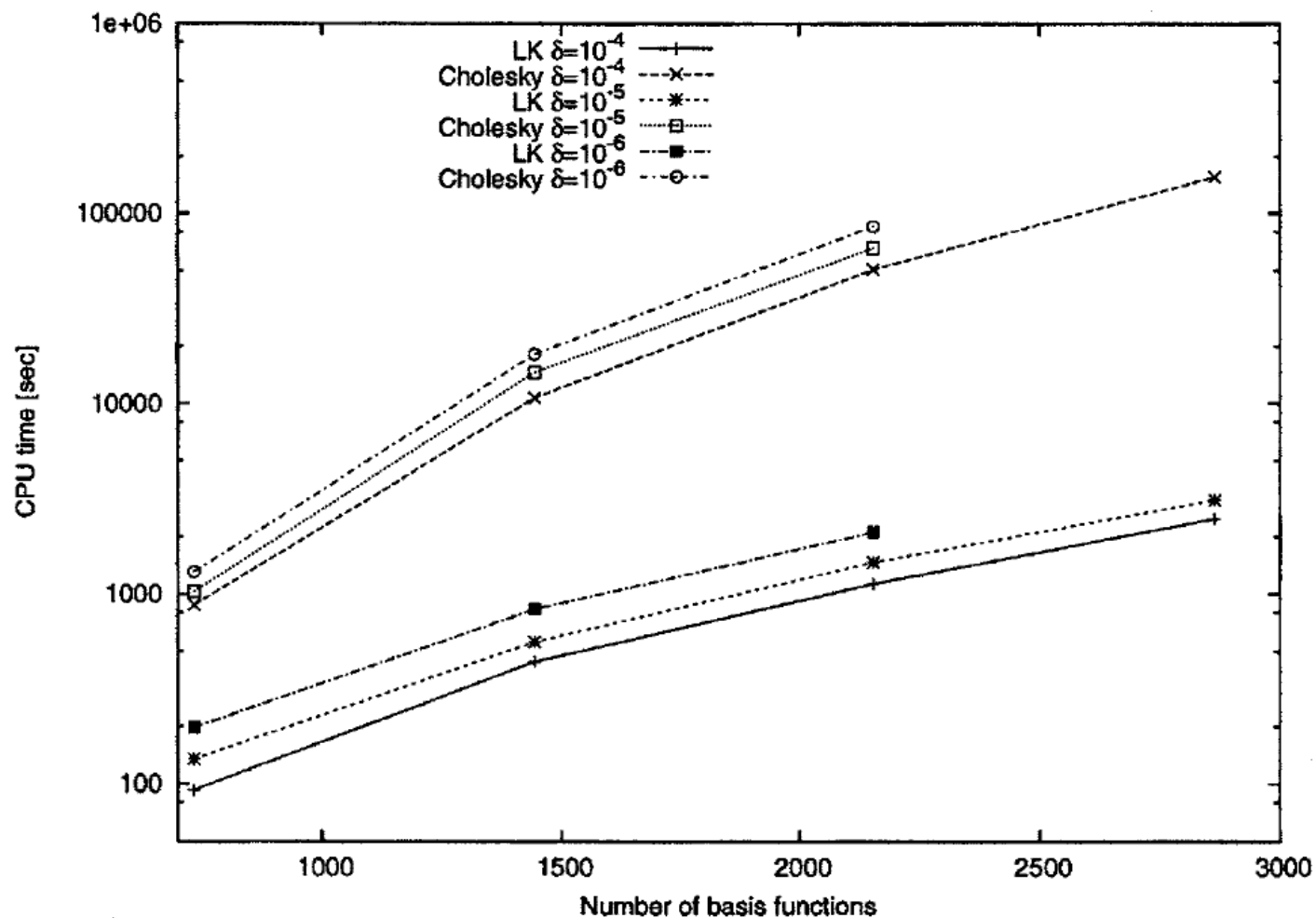
The CD localization scheme is a non-iterative procedure.

- **Error bounded screening**

Reformulation, ERI matrix in AO basis is positive definite and satisfy the Cauchy-Schwarz inequality.



LK Scaling



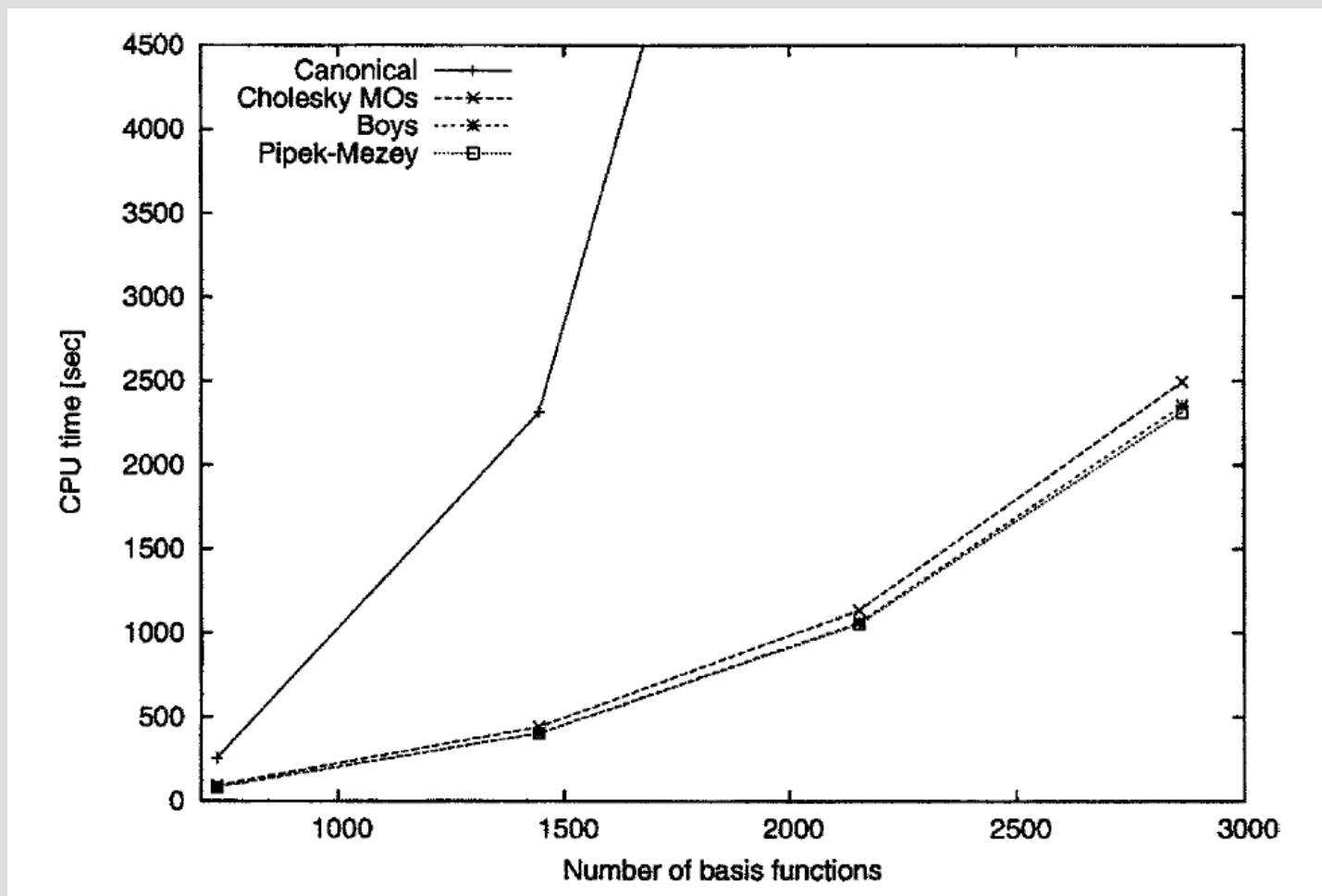
Linear Glycines / cc-pVDZ



Cholesky Localization



**Cholesky localization is not perfect -
do we care?**





1-Center CD



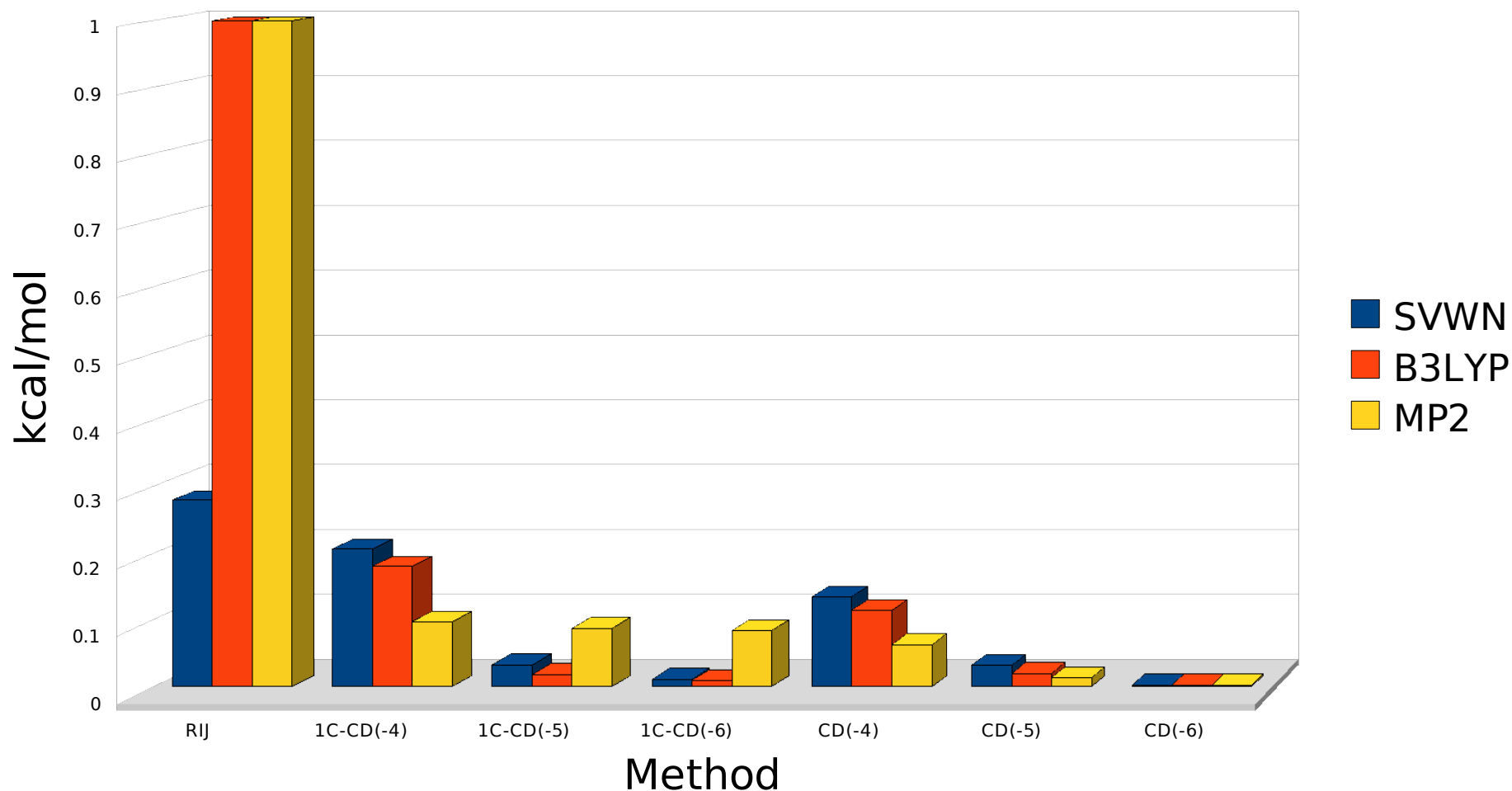
Q: Will a CD procedure which *exclude* all 2-center products as potential auxiliary basis function retain an acceptable accuracy?



1-Center CD vs. full CD



RMS Error / SVP





Observations



- 1-Center approximation *does not* degrade the CD accuracy significantly!
- for 1C-CD the decomposition time is **3-4 times faster than the full CD** with the same threshold.

In the 1C-CD approximation a fixed auxiliary basis set is used, hence we can compute analytic derivatives!



1C-CD gradients: timings



This is now a trivial matter! Use RI (and LK) technology!

Molecule/ Basis set	\mathcal{N}^a	N^b	1C-CD		Integral-direct
			$\delta = 10^{-2}$	$\delta = 10^{-4}$	
<i>cc-pVDZ</i>					
(Gly) ₁	10	95	4.1 (6)	4.6 (7)	8.5 (8)
(Gly) ₁₀	73	734	371 (209)	573 (221)	1257 (666)
(Gly) ₂₀	143	1444	1578 (593)		(2298)
(Gly) ₃₀	213	2154			
<i>cc-pVTZ</i>					
(Gly) ₁	10	220	27.8 (21)	41.3 (23)	95.5 (45)
(Gly) ₁₀	73	1678	2738 (871)	6007 (993)	(4970)

^aNumber of atoms. ^bNumber of basis functions.



1C-CD vs. Conv.: Energies and Bond Distances



TABLE II. 1C-CD DFT/BLYP results at different decomposition thresholds δ , employing the SVP basis set. Accuracy $|\Delta E|$ (in mhartrees) of computed total energies and maximum deviation $|\Delta r|$ (in pm) of bond distances compared to conventional calculations. All calculations have been performed without point group symmetry restrictions.

Molecule	$ \Delta E $				$ \Delta r $			
	$\delta=10^{-2}$	10^{-3}	10^{-4}	10^{-5}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
N ₂	3.65	1.48	0.015	0.009	0.0009	0.0298	0.0001	0.0001
C ₂ H ₂	4.35	0.08	0.006	0.003	0.1615	0.0246	0.0107	0.0112
C ₂ H ₆	4.04	1.33	0.021	0.009	0.7536	0.0683	0.0219	0.0276
C ₆ H ₆	7.76	1.54	0.061	0.075	0.0252	0.0297	0.0051	0.0054
CH ₃ OH	3.91	1.32	0.039	0.012	0.1248	0.0067	0.0019	0.0017
HCOCl	4.83	1.25	0.053	0.026	0.1183	0.0148	0.0073	0.0071
SO ₂	5.33	1.38	0.041	0.010	0.0133	0.0099	0.0021	0.0008



The Cholesky auxiliary basis sets!



Q: Given the accuracy of the 1C-CD approach, could it be used to design general DF/RI auxiliary basis sets which are method-free?

Use ***atomic*** CD technique to design the aCD RI basis sets. Plug them into your RI code! aCD/RI

aCD/RI: 1C-CD quality results without the recursive nature of CD



Accuracy of aCD RI basis sets



On-the-fly CD of the atomic two-electron integral matrix
Baker test set of reactions (cc-pVTZ)

Method	Aux. basis	Total Energy	Activation Energy
		Max Error	Max Error
		kcal mol ⁻¹	kcal mol ⁻¹
B3LYP	aCD-3	1.516	0.301
	aCD-4	0.062	0.031
	aCD-5	0.062	0.062
MP2 ^a	aCD-3	1.406	0.293
	aCD-4	0.187	0.062
	aCD-5	0.187	0.047

**aCD/RI is 2-4 times faster than
1C-CD!**



Atomic Compact CD auxiliary basis *acCD*



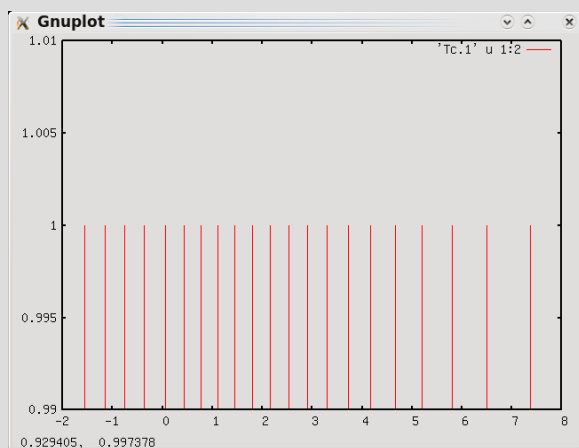
The CD approach does not automatically eliminate the redundancy in the primitive space product space! But we can do that!

- Compute a normal aCD basis set.
- Do a complementary CD elimination in the primitive product space. Keep essential products as exponents of the acCD basis
- Do a least-square fit to the aCD basis set to get the contract coefficients of the acCD basis.

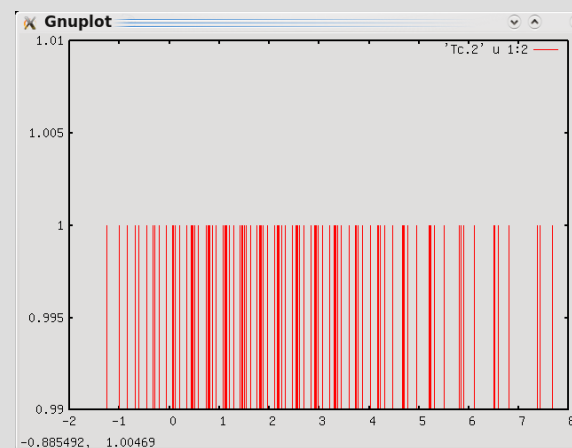


Tc ANO-RCC auxiliary basis set

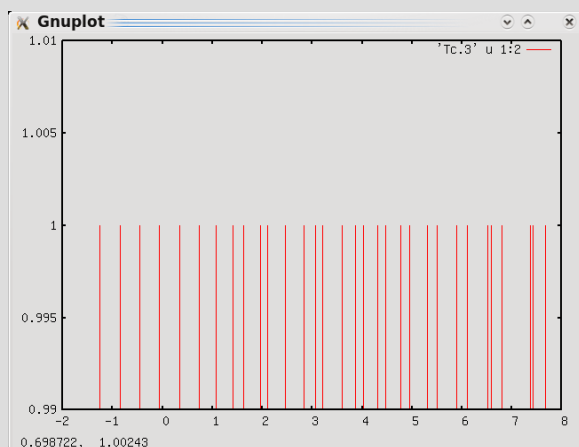
$\tau=1.0E-4$



Tc-ANO: 21 s-functns



Tc-aCD: 231 s-functns



Tc-acCD: 32 s-functns

Total # of products

conv: 23871
aCD: 17272
acCD: 3946



The CD-RI hierarchy in MOLCAS



- CD($\tau=0$) - Conventional
- CD(τ)
- 1C-CD(τ)
- aCD(τ)/RI
- acCD(τ)/RI
- External aux. bfn/RI



CD developments: material published so far



- “Analytic derivatives for the Cholesky representation of the two-electron integrals” - **CD gradients**
- “Unbiased auxiliary basis sets for accurate two-electron integral approximations” - **CD-RI auxiliary basis sets**
- “Cholesky decomposition-based multiconfiguration second-order perturbation theory (CD-CASPT2): Application to the spin-state energetics of Co-III(diiminato) (NPh) - **CD-CASPT2**
- “ Accurate ab initio density fitting for multiconfigurational self-consistent field methods” - **CD-CASSCF**
- “Quartic scaling evaluation of canonical scaled opposite spin second-order Moller-Plesset correlation energy using Cholesky decompositions” - **CD-MP2**
- “Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals” - **CD-HF**



CD-CASPT2 Example: Relative energies of spin- states of Ferrous complex

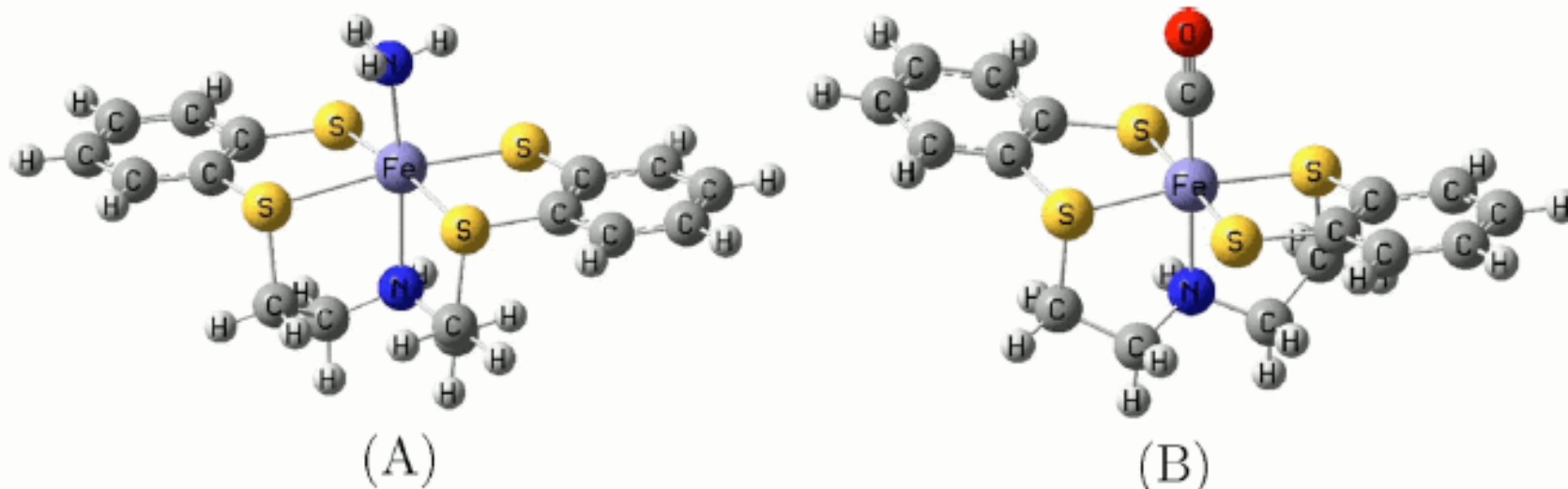


FIG. 1. (Color online) Examples of $[\text{Fe}(\text{L})(\text{NHS}_4)]$ complexes. (A) The *cis*- NH_3 complex. (B) the *trans*-CO complex.

K. Pierloot et al.

-CD(-6)-CASPT2/CASSCF(14-in-16)/ANO

-810 bfn (no symmetry) / 964 bfn (C_2 symmetry)



CD-CASPT2 example



Cholesky decomposition-based multiconfiguration second-order perturbation theory (CD-CASPT2): Application to the spin-state energetics of Co-III(diiminato)(NPh)

Aquilante et al.

- CASSCF/CASPT2
- ANO-RCC-VTZP
- 869 bfn

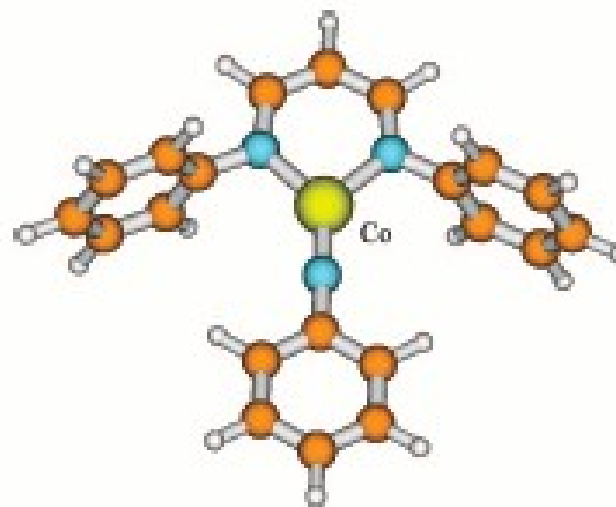


Figure 1. Model of the Co-imido complex used in the calculations.



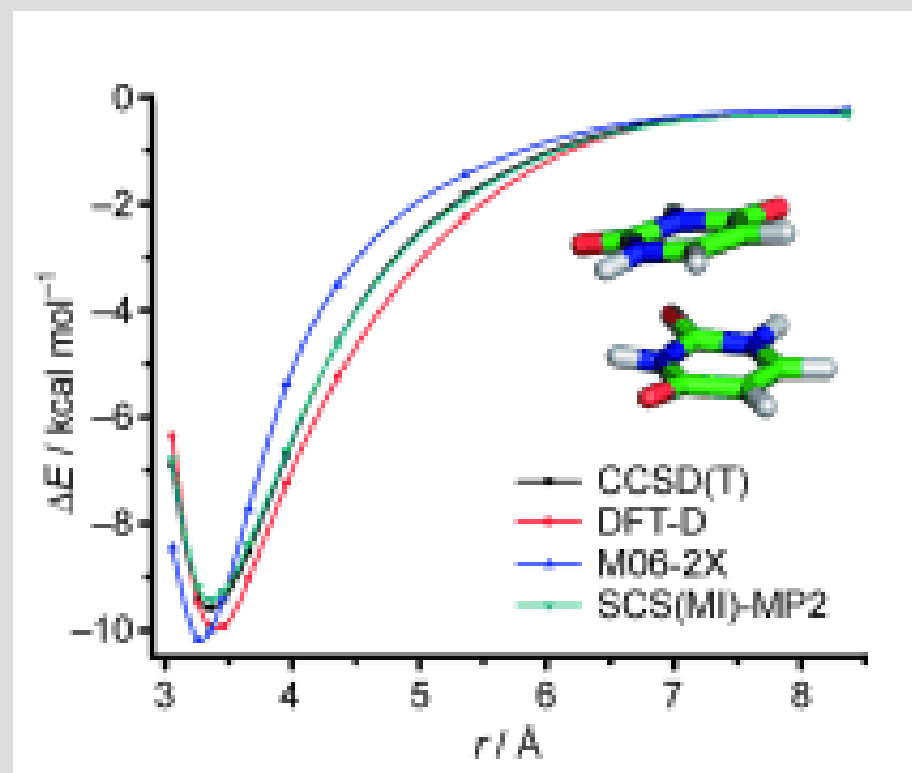
CD-CCSD(T)



Highly Accurate CCSD(T) and DFT-SAPT Stabilization Energies of H-Bonded and Stacked Structures of the Uracil Dimer

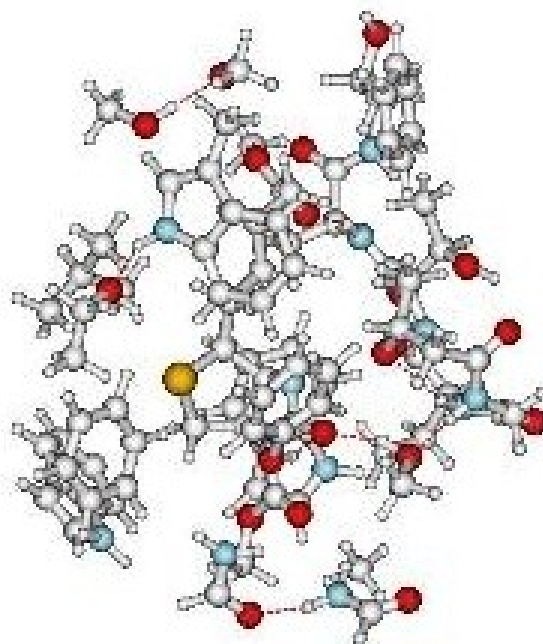
Pitonak et al., CPC

- MP2 - 1648 bfn
- CCSD(T) - 920 bfn





CD-MP2 Aquilante et al.



Model for Biotin@Avidin: 257 atoms, 924 e⁻

6-31G* : 2091 bsf

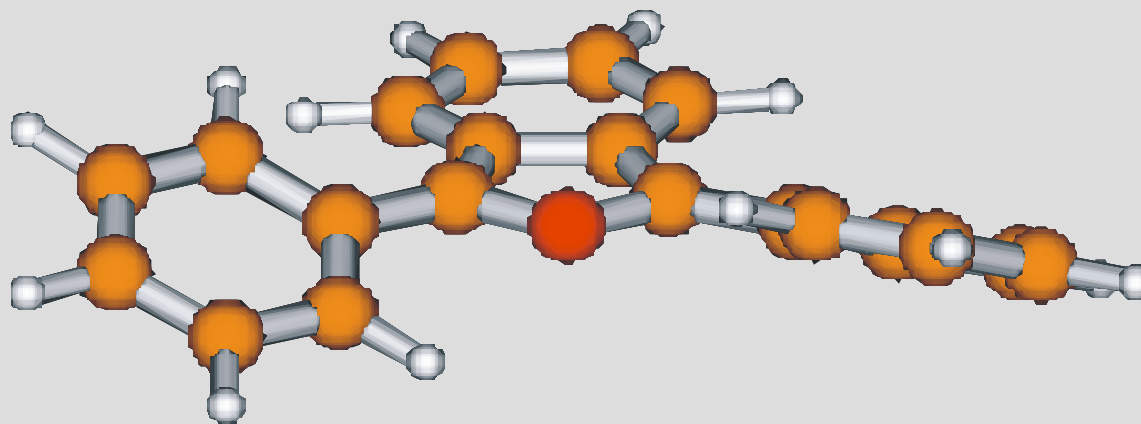
Cholesky-MP2: 44 hr on 4 CPUs

0.1 mH accuracy in the total energy ($\delta = 10^{-4}$)



1,3-DIPHENYLISOBENZOFURAN: Photovoltaic material with singlet fission

Zdenek Havlas, Andrew Schwerin, and
Jozef Michl



CAS(16el/14orb; 7a,7b)

ANO-L(C,O: 4s3p2d1f, H: 3s2p1d, Ryd(8,8,8)/[1,1,1])

Cholesky (Thrs= 1.0d-5)

35 atoms ($O_1C_{20}H_{14}$)

835 orbitals (419a, 416b)

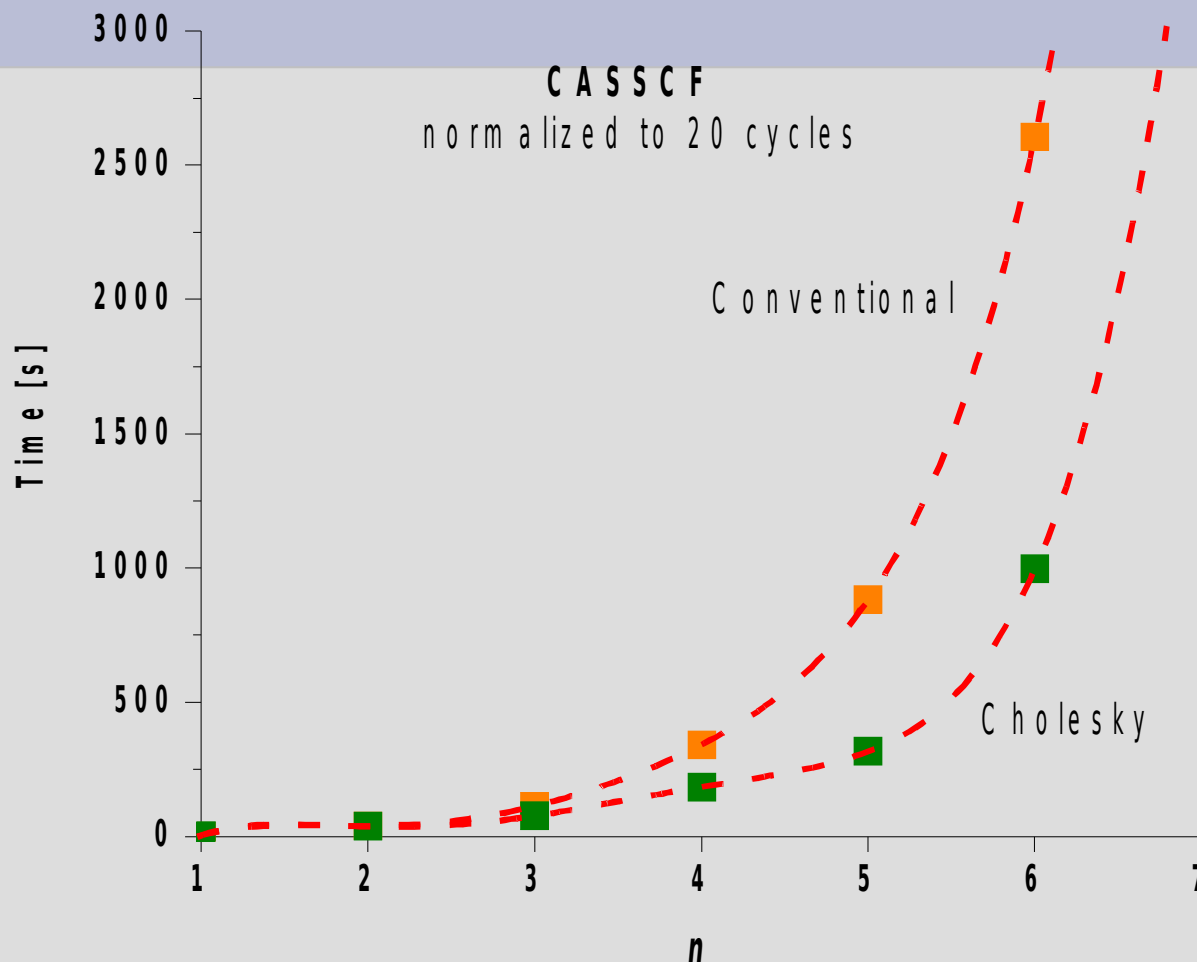
Speed up: 48 h - > 15 h



CHOLESKY DECOMPOSITION IN MOLCAS 7.0 (Z. Havlas et al.)



(Ethylene)_n, stacked with distance 3Å



- Basis: cc-pVDZ
- Active space: $n \times (2,2)$, π orbitals only
- Processors: AMD Athlon 64 X2 Dual Core Processor 4800+
- CPU speed: 2.4 GHz



Interested in more examples?



File Edit View History Bookmarks Tools Help

Back Forward Reload Stop Home Print http://www.teokem.lu.se/~roland/show_case/covers.html per siegbahn

Most Visited ISI MOLCAS Google OpenMP Anslag&stip Journals Tinker QM/MM Rotary The Vega Science Tr... Home pages LinkedIn S f

Google Maps <http://ww...vers.html>

The Show Case Room



This web page is dedicated to scientific paper in which MOLCAS has been of significance. This we hope will inspire old and new MOLCAS users to see what is possible with the MOLCAS software. This web page contains four sections:

1. a dynamically updated ISI list of all papers which quotes the official "MOLCAS paper" ([MOLCAS: a program package for computational chemistry](#)).
2. a list of MOLCAS hallmark application papers (if you have a paper you think should be on the list email roland.lindh@teokem.lu.se, please!).
3. a list of MOLCAS methods papers, and

[MOLCAS reference papers since September 2004:search the ISI data base!](#)

MOLCAS hallmark application papers ([see also the news stand below](#)):

1. **"Molecular Basis of DNA Photodimerization: Intrinsic Production of Cyclobutane Cytosine Dimers"**,
D. Roca-Sanjuán, G. Olaso-González, I. González-Ramírez, L. Serrano-Andrés, and M. Merchán,
J. Am. Chem. Soc., In press (2008).
2. **"Cholesky decomposition based multiconfigurational second order perturbation theory (CD-CASPT2): Application to the spin state energetics of Co(III)(diiminato)(NPh)"**,
F. Aquilante, P.-Å. Malmqvist, T.B. Pedersen, A. Ghosh, B.O. Roos,
J. Chem. Theor. Comp., 4, 694-702 (2008). [[DOI](#)]
3. **"A Three-State Model for the Photophysics of Guanine"**,
L. Serrano-Andrés, M. Merchán, A.C. Borin,



Furture work!



- HF 1C-CD and RI gradients
- CASSCF 1C-CD and RI gradients
- MP2 1C-CD gradients
- Localized and linear scaling RI and CD
- Numerical problems with accurate RI/CD
- CD in the N-particle space



Summary



- DF, RI and CD are interrelated
- LK Exchange
- Gradients for CD
- **1C-CD** approximation is **equivalent in performance and accuracy to DF/RI**
- **1C-CD** approach can be used to derive “**method-free**” **aCD RI auxiliary basis sets**
- Much smaller **acCD** auxiliary basis sets can be derived from **aCD** basis sets without any further loss of accuracy.