

Accuracy of Equilibrium Geometries

Jürgen Gauss



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

General Issues

- geometry optimization routine with analytic gradients
- single calculation provides results without error bars

how to judge effects of approximation in the solution of the Schrödinger equation?

=> sequences of calculations

- to approach exact solution in a systematic manner
- to provide (statistical) error estimates

Approximate Solution of Schrödinger Equation

main issues:

- **basis-set approximation**

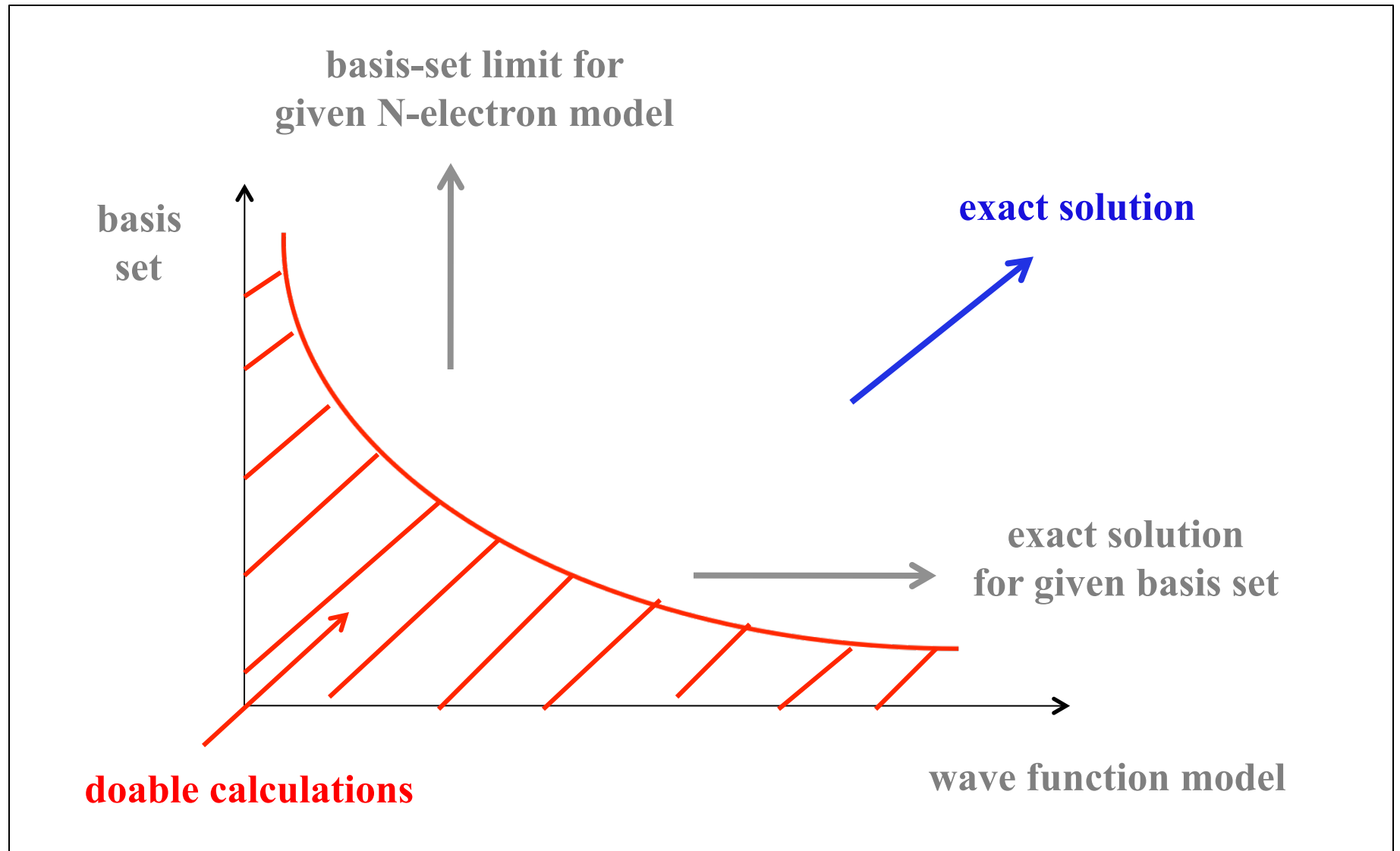
=> one-electron space

- **wave function model**

=> electron correlation

=> N-electron space

Approximate Solution of Schrödinger Equation



Other Issues

- **relativistic effects**

- relevant for heavy-element compounds

- **errors due to Born-Oppenheimer approximation**

- generally considered small and negligible

- **proper comparison with experiment**

- (ro)vibrational corrections

- solvent effects

Systematic Approach to the Solution of the Schrödinger Equation

sequences of calculations based on hierarchies

- hierarchies of basis sets

correlation-consistent basis sets (Dunning)

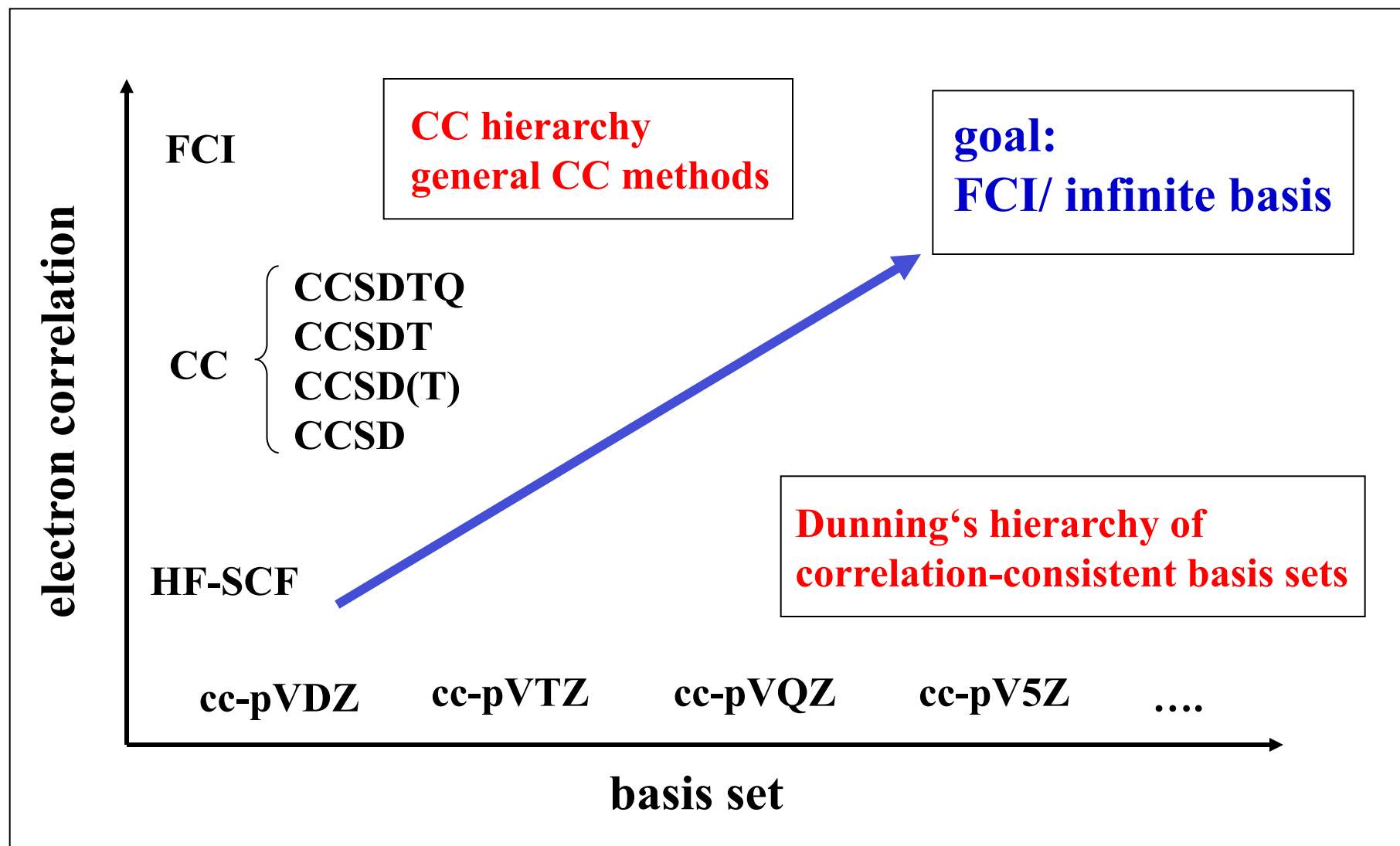
cc-pVXZ, aug-cc-pVXZ, cc-pCVXZ, ... X=D,T,Q,5, ...

- hierarchies of wave function models

HF, MP2, CISD, CCSD, CCSD(T), FCI

systematic increase of accuracy within each hierarchy

Systematic Approach to the Solution of the Schrödinger Equation



Systematic Approach to the Solution of the Schrödinger Equation

sequences of calculations based on hierarchies

- increased computational cost

but

- error estimate based on convergence
- extrapolation to corresponding limits
- statistical error estimates

Statistical Error Estimates

- **single calculation provides no error estimate**

=> accuracy ?

- **statistical analysis**

calculations for (large) test set

analyze performance with respect to experiment

characterization of methods wrt (average) accuracy

facilitates choice of methods in terms of accuracy and cost

Statistical Measures

$$\Delta_i = R_i^{calc} - R_i^{exp}$$

error in individual bond length

$$\bar{\Delta} = \frac{1}{n} \sum_i^n \Delta_i$$

mean error

$$\Delta_{abs} = \frac{1}{n} \sum_i^n |\Delta_i|$$

mean absolute error

$$\Delta_{max} = \max |\Delta_i|$$

maximum error

$$\Delta_{std} = \sqrt{\frac{1}{n-1} \sum_i^n (\Delta_i - \bar{\Delta})^2}$$

standard deviation

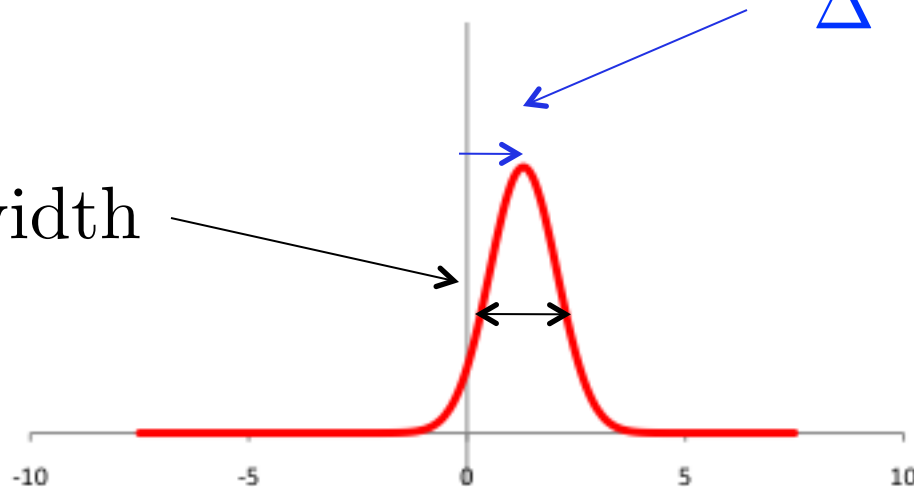
Statistical Measures

normal distribution

$$N \exp \left(-\frac{1}{2} \left(\frac{\Delta - \bar{\Delta}}{\Delta_{std}} \right)^2 \right)$$

$\Delta_{std} \sim$ width

$\bar{\Delta} \sim$ shift



Benchmark Study I

19 closed-shell molecules

CH₂, CH₄, NH₃, H₂O, HF, N₂, C≡O, HC≡N, HN≡C, HC≡CH,
CH₂=CH₂, H₂C=O, HNO, NH=NH, HOF, O=C=O, H₂O₂, O₃, F₂

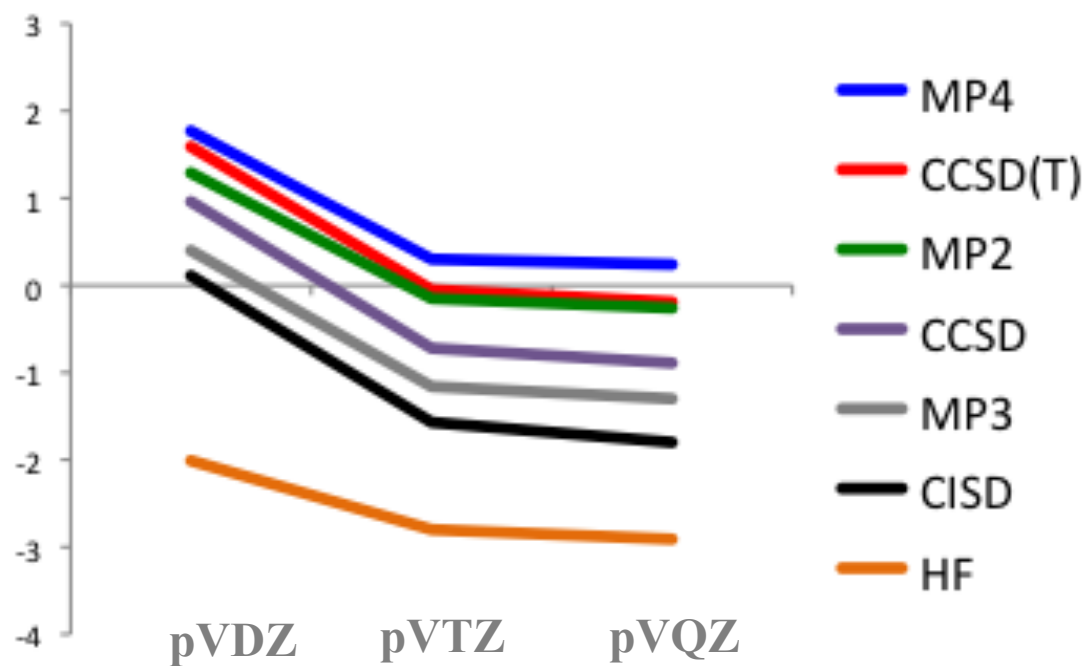
wave function models

HF, CISD, MP2, MP3, MP4, CCSD, CCSD(T)

basis sets

cc-pVXZ, X=D,T,Q

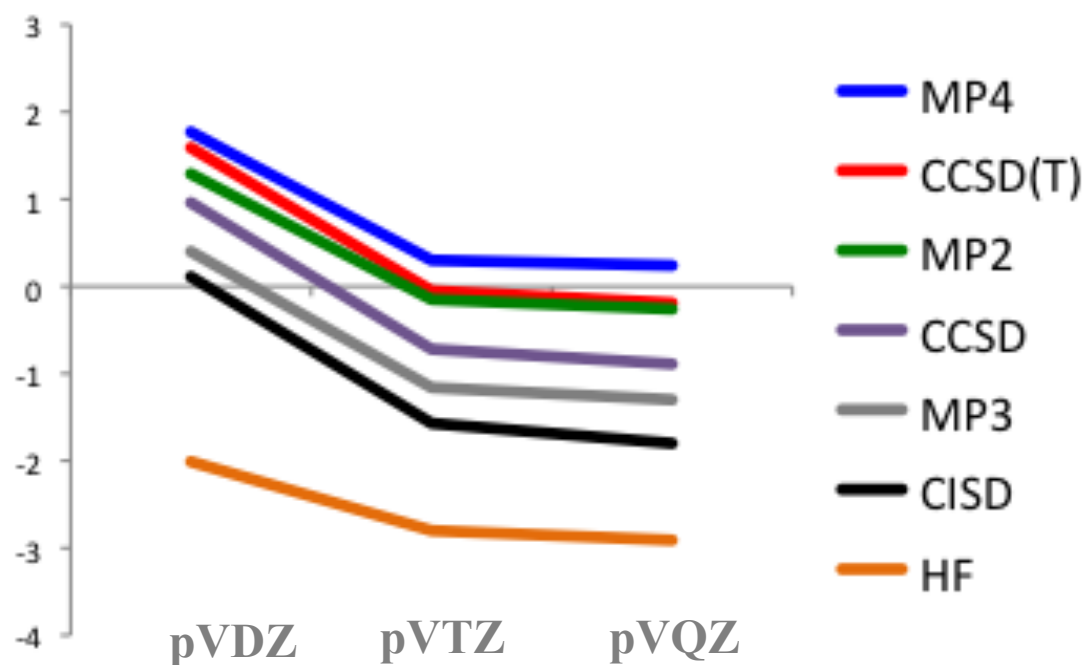
Mean Errors in Bond Lengths



highly systematic behaviour – no crossings !

two converse effects

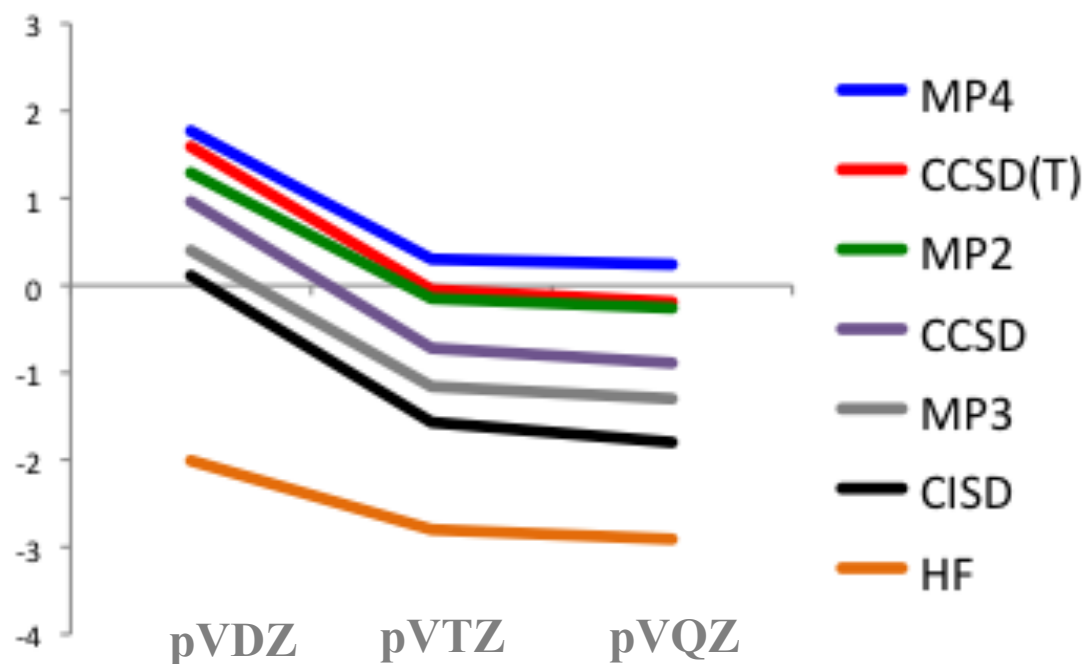
Mean Errors in Bond Lengths



1) bonds shortens with increasing basis

DZ \rightarrow TZ \rightarrow QZ hierarchical decrease

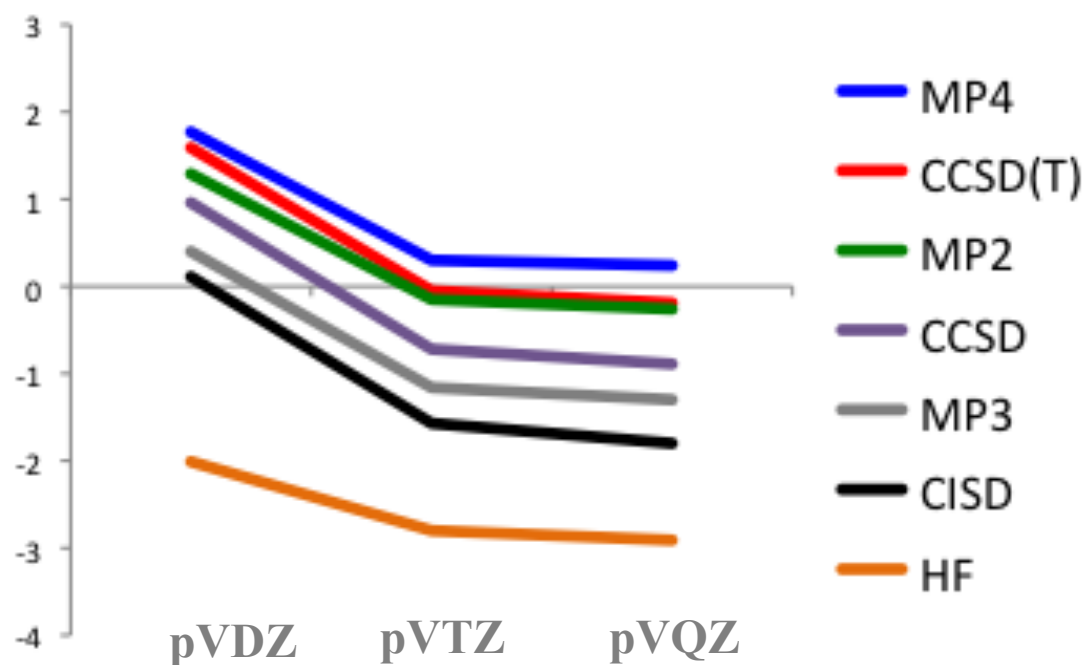
Mean Errors in Bond Lengths



2) bond lengthens with improvement in wave function model



Mean Errors in Bond Lengths



note MP3 < MP2 < MP4 (oscillatory behaviour)

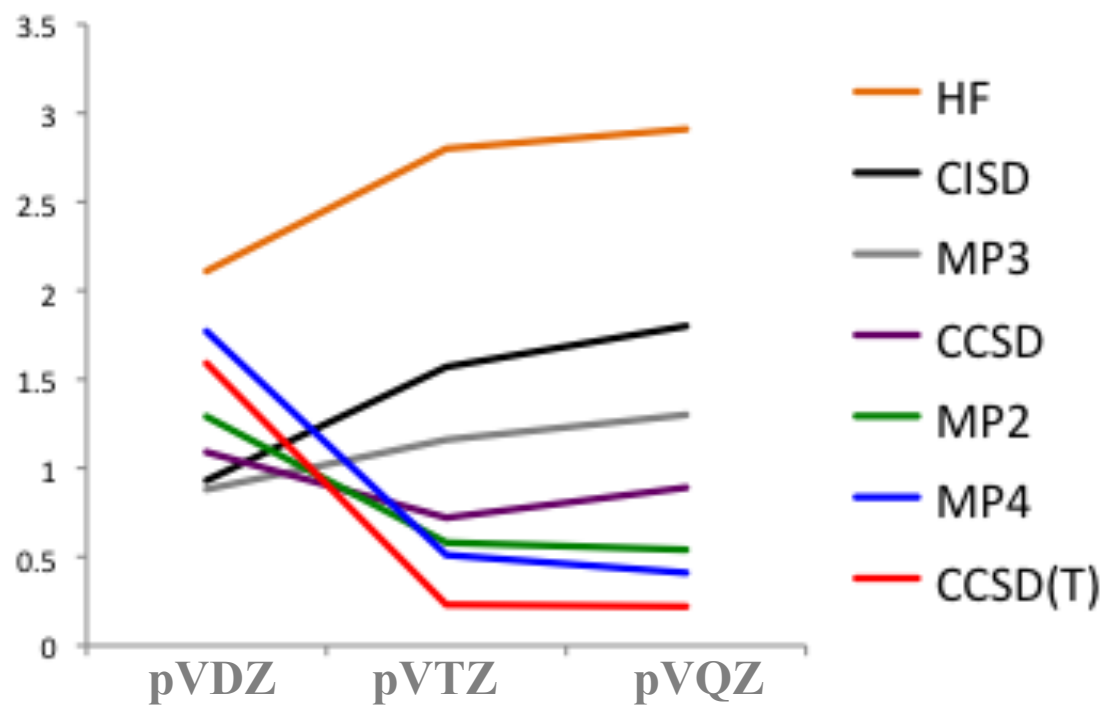
considerable scope for error cancellations: CISD/cc-pVDZ

DZ/TZ Inversion for Correlated Models

	cc-pVDZ	cc-pVTZ
best	CISD, MP3	CCSD(T), MP4
worst	CCSD(T), MP4	CISD, MP3
MP order	MP3<MP2<MP4	MP4<MP2<MP3

DZ sets inadequate for correlated methods

Mean Absolute Error



1) HF

2 - 3 pm

2) CISD, MP3

1 - 2 pm

3) CCSD

≈ 1 pm

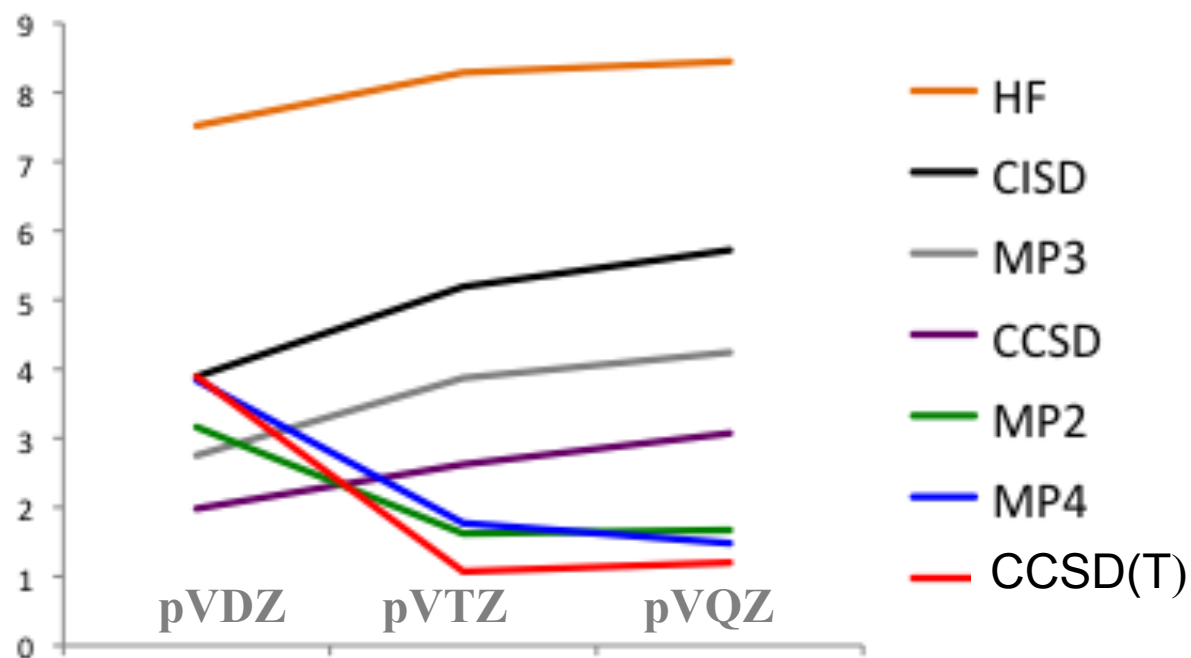
4) MP2, MP4, CCSD(T)

0.2 - 0.5 pm

deteriorate with X

improve with X

Maximum Absolute Error



worst case errors:

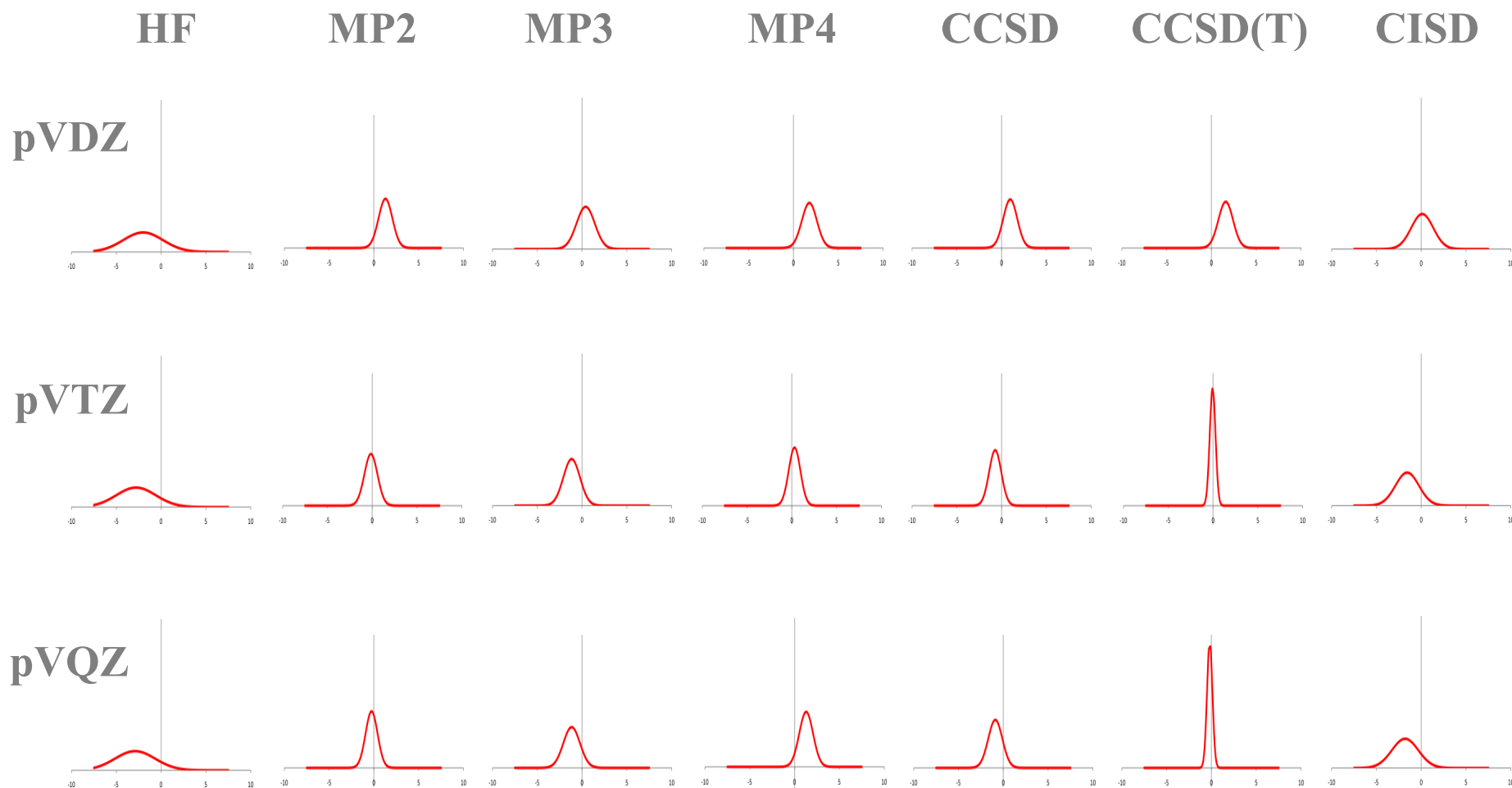
$$\Delta_{max} = \max |\Delta_i| \sim 3\bar{\Delta}_{abs}$$

for CCSD(T):

$$\Delta_{max} \sim 6\bar{\Delta}_{abs}$$

uncertainty in experimental results

Normal Distributions



spectacular performance of CCSD(T) !

Problems

- **no uniform accuracy in experimental equilibrium geometries**
=> issue for highly-correlated methods
- **test set contained single-reference and multi-reference cases**
=> exclude ozone (O_3)
- **core correlation should be treated with core-polarized basis sets**
=> use cc-pCVXZ sets
- **importance of higher excitations**
=> CCSDT, CCSDTQ,
- **restriction to first-row elements**
=> different test set

Experimental Geometries

r_e	equilibrium geometry
r_g	vibr. averaged distances
r_α, r_z	vibr. averaged geometry
r_0	from B_0 by ignoring vibrational effects
r_s	substitution structure

bond distances
for methane

	CH ₄	CD ₄
r_e	108.59	108.59
r_0	109.40	109.19
r_g	110.68	110.27
r_z	109.91	109.55

differences of 1- 2 pm

Equilibrium Geometries

- **minimum on Born-Oppenheimer potential energy surface**

usually/routinely determined in quantum chemistry

- **experimental determination**

$$B_0 = B_e - \frac{1}{2} \sum_r \alpha_r^B + \dots$$

vibrationally averaged rotational constant
from experiment

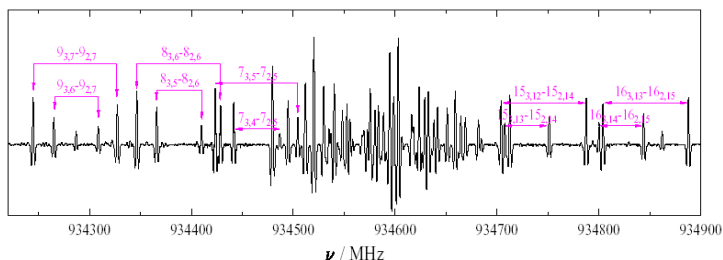
equilibrium value

$B_e = \frac{h}{8\pi^2 c I_e} \longrightarrow I_e = \mu R_e^2$
(diatomic molecule)

α_r^B **vib.-rot. constants**
from experiment (in principle)

vibrational corrections

Determination of Equilibrium Geometries



exp. rotational spectra



**rotational constants (B_0)
for various isotopologues**

analysis

r_0, r_s, \dots structures

computed
vibrational
corrections

rotational constants B_e

r_e structure

Vibration-Rotation Interaction Constants

second-order vibrational perturbation theory (VPT2) yields

$$\alpha_r^B = -\frac{2B^2}{\omega_r} \left[\sum_{\xi} \frac{3(a_r^{(b\xi)})^2}{4I_{\xi}} + \sum_s (\zeta_{r,s}^{(b)})^2 \frac{(3\omega_r^2 + \omega_s^2)}{\omega_r^2 - \omega_s^2} + \pi \left(\frac{c}{h} \right)^{1/2} \sum_s \phi_{rrs} a_s^{(bb)} \left(\frac{\omega_r}{\omega_s^{3/2}} \right) \right]$$

ω_r harmonic frequency of mode r

I_{ξ} ξ -th principal component of inertia tensor

$a_r^{(ab)}$ derivative of the inertia tensor wrt normal coordinate Q_r

$\zeta_{r,s}^{(b)}$ Coriolis matrices

ϕ_{rst} cubic force constant

accessible via quantum-chemical calculations

Errors in Equilibrium Geometries

• error in experimental B_0 \Rightarrow negligible

• error in least-squares fit \Rightarrow negligible

• error in computed α_r^B

$$\delta\alpha_r^B \approx 0.001 \text{ cm}^{-1} \quad \frac{\delta R_e}{R_e} = \frac{1}{4} \sum_r \frac{\delta\alpha_r^B}{B_0}$$

\Rightarrow 0.06 pm accuracy

• error in B_0 expansion \Rightarrow negligible

accuracy of empirical R_e geometries better than 0.1 pm

Accuracy of Empirical Geometries

	$r_e(\text{exp})$	$r_e(\text{emp})$	$r_e(\text{best theory})$
HF	91.680(8)	91.69	91.71
CO	112.832	112.84	112.84
N ₂	109.768(5)	109.77	109.75
H ₂ O	95.72	95.75	95.78
NH ₃	101.1(6)	101.16	101.09

CCSD(T)/cc-pVQZ vibrational corrections

Pawłowski et al., *J.Chem. Phys.* **116**, 6482 (2003)

Heckert et al., *J. Chem. Phys.* **125**, 044108 (2006)

Benchmark Study II

19 closed-shell molecules

CH₂, CH₄, NH₃, H₂O, HF, N₂, C≡O, HC≡N, HN≡C, HC≡CH,
CH₂=CH₂, H₂C=O, HNO, NH=NH, HOF, O=C=O, H₂O₂, H₂, F₂

wave function models

HF, MP2, CCSD, CCSD(T)

basis sets

cc-pCVXZ, X=D,T,Q

empirical equilibrium geometries (except for H₂O₂)

CCSD(T) Calculations of Bond Lengths: Statistics

	old	new (with H ₂ O ₂)			new (without H ₂ O ₂)		
	Q	D	T	Q	D	T	Q
$\bar{\Delta}$	0.19	1.61	0.17	0.04	1.67	0.20	0.00
$\bar{\Delta}_{abs}$	0.22	1.61	0.22	0.09	1.67	0.21	0.06
Δ_{std}	0.30	0.78	0.18	0.16	0.77	0.14	0.08
Δ_{max}	1.20	4.42	0.49	0.59	4.42	0.49	0.21

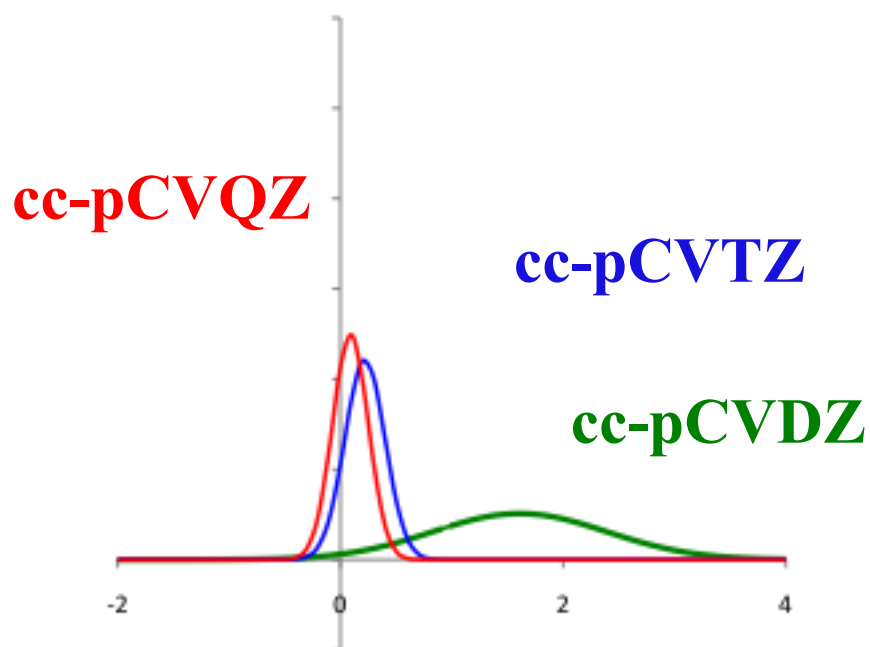
problem case H₂O₂:

experimental R_{OH} ~ 0.5 pm too long

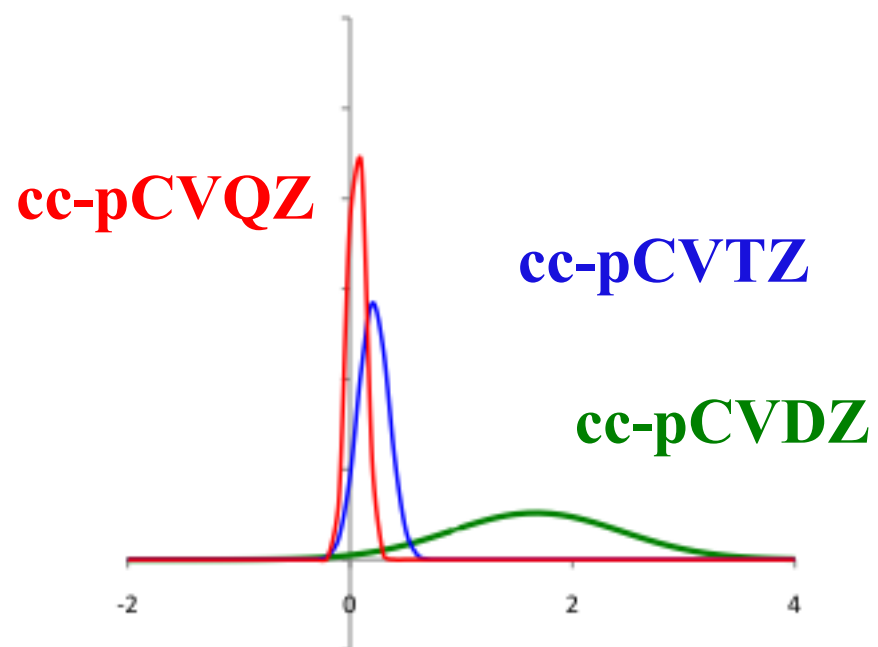
experimental R_{OO} ~ 0.6 pm too long

Errors in Equilibrium Geometries

CCSD(T) with H_2O_2

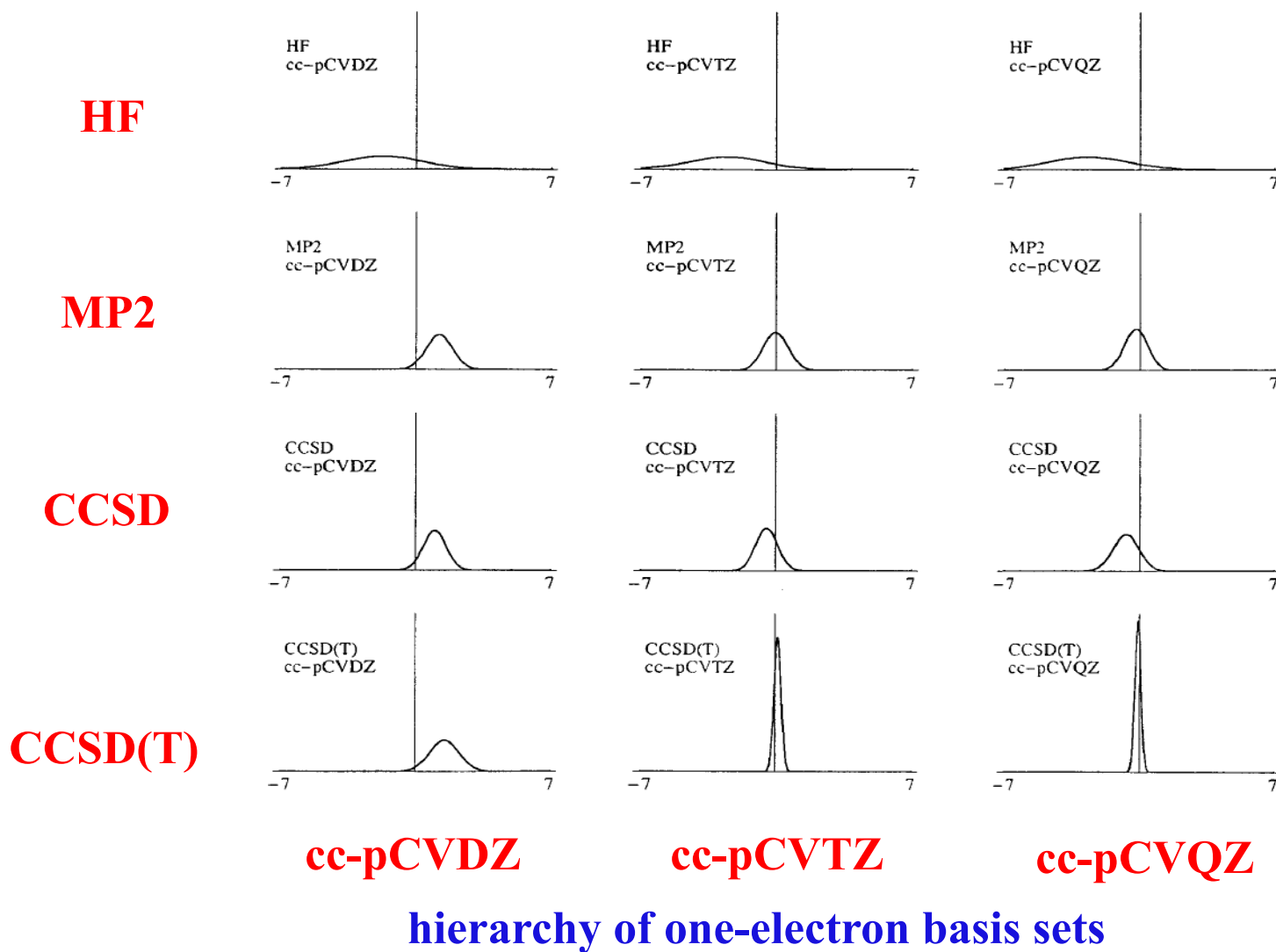


CCSD(T) without H_2O_2

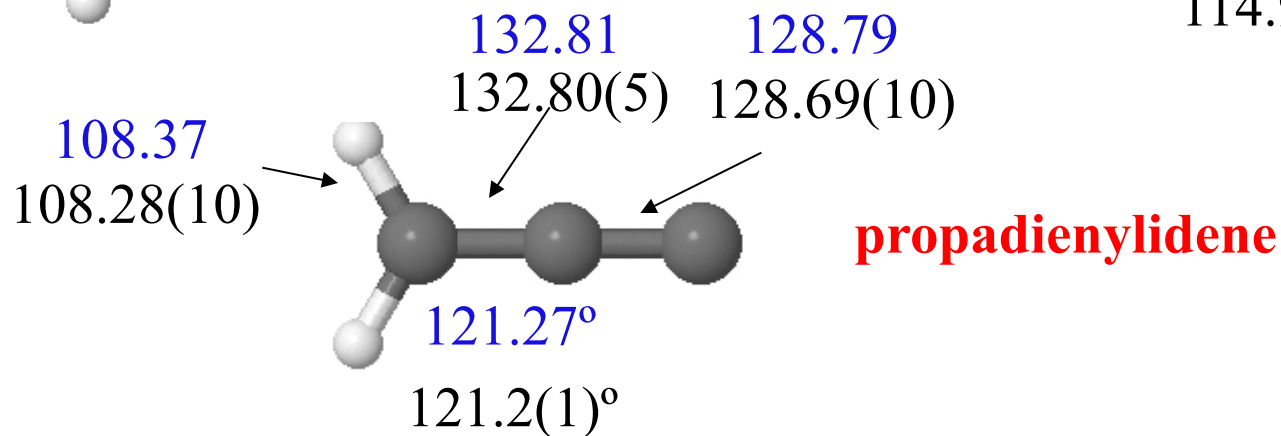
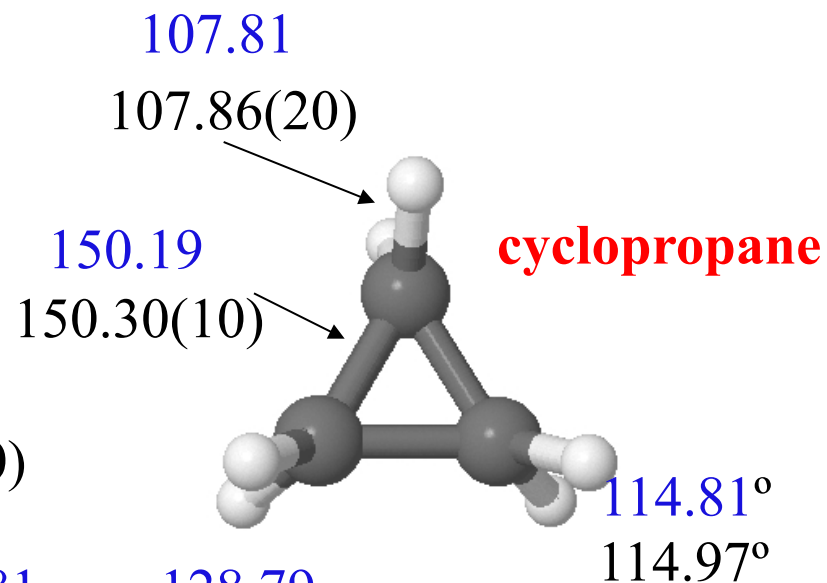
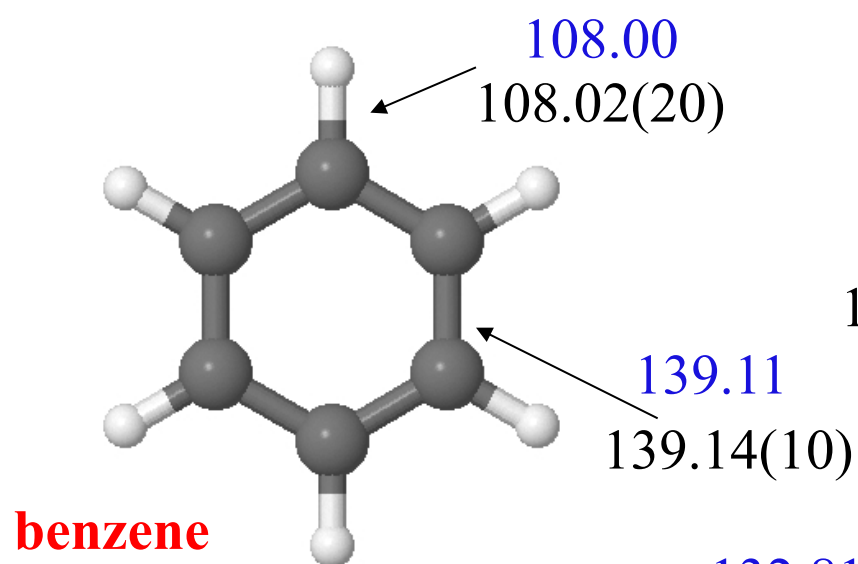


Systematic Solution to the Schrödinger Equation

hierarchy of N-electron models



Examples



CCSD(T)/cc-p(C)VQZ calculations

empirical r_e geometry

Benchmark Study III

31 molecules containing second-row elements

HCl, H₂S, PH₃, SiH₂, SiH₄, P₂, PN, Cl₂, ClF, ClOH, HCP,
ClNO, Cl₂O, CCl₂, BCl, BCl₃, HSiCl, ClBS, CS, CS₂, OCS,
HBS, FBS, S₂O, SO₂, SO₃, SiS, H₂CS, SiO, SiF₂, SiCl₂

wave function models

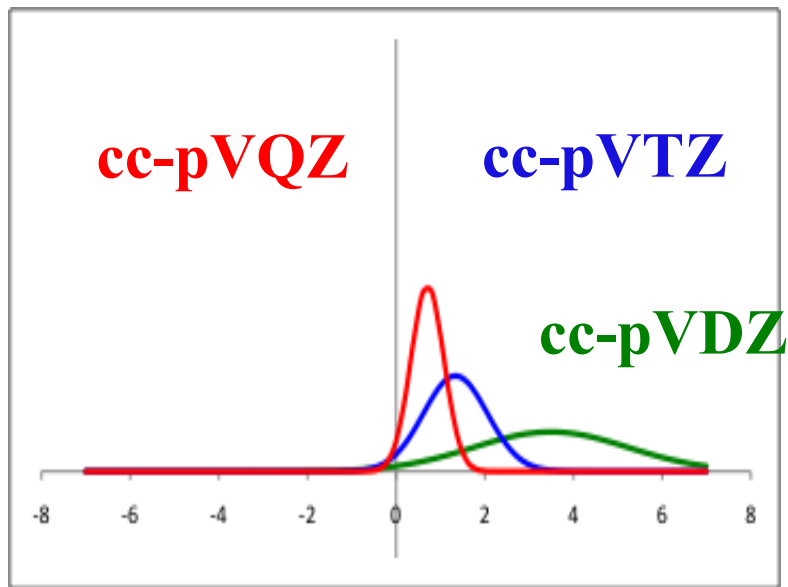
HF, MP2, CCSD, CCSD(T)

basis sets

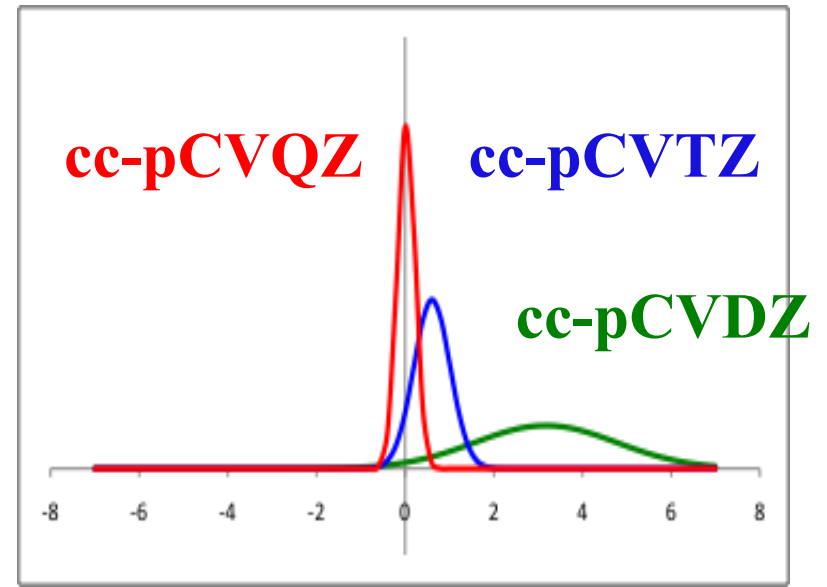
cc-p(C)VXZ, X=D,T,Q

Core-Correlation Effects

frozen core

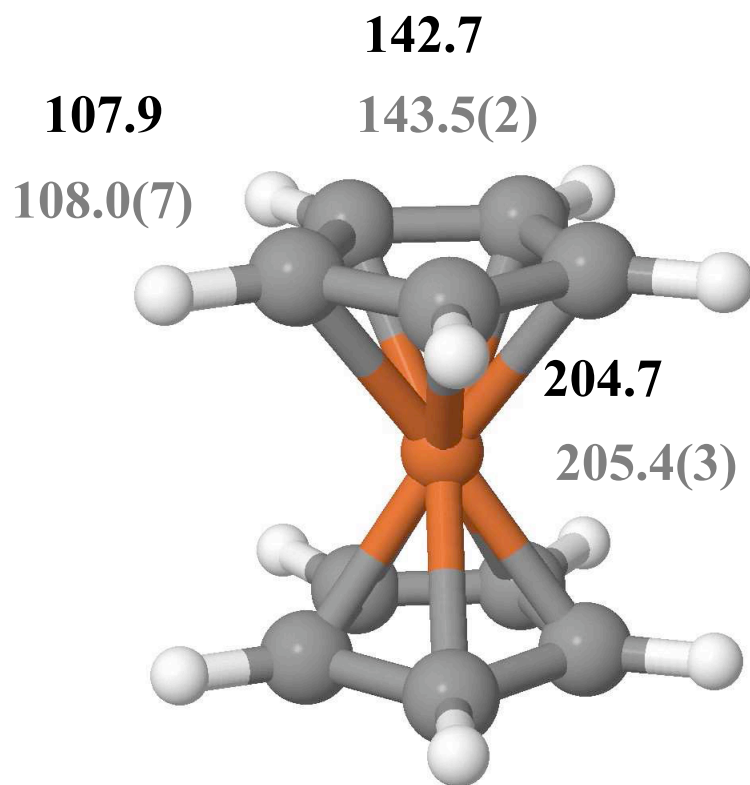


core correlation



core correlation essential for high accuracy
even more important for second-row elements

Equilibrium Structure of Ferrocene



CCSD(T)/cc-pwCVTZ

672 basis functions

96 correlated electrons

r(FeC5): 164.8 pm, \angle (C5H): 0.52°

r(FeC5): 166.1 pm, \angle (C5H): 3.7°

**gas-phase
electron diffraction**

Benchmark Study IV

12 closed-shell molecules

CH_2 , CH_4 , NH_3 , H_2O , HF , N_2 , $\text{C}\equiv\text{O}$, $\text{HC}\equiv\text{N}$,
 $\text{HN}\equiv\text{C}$, $\text{HC}\equiv\text{CH}$, $\text{O}=\text{C}=\text{O}$, F_2

wave function models

CCSD , CCSD(T) , CCSDT , CCSDTQ

basis sets

cc-p(C)VXZ , $X = \text{T, Q, 5, 6}$ + additivity assumptions

empirical equilibrium geometries

Beyond CCSD(T) Accuracy ...

starting point

CCSD(T)/cc-pV6Z

+

core correlation

CCSD(T)/cc-pCVQZ

+

full triples contribution

CCSDT/cc-pVTZ

+

full quadruples contribution

CCSDTQ/cc-pVDZ

relativistic and adiabatic effects for first-row elements negligible

Additivity Assumptions

energy and potential energy surface

$$E = E(\text{CCSD(T)/cc-pV6Z})$$

$$+ E(\text{CCSD(T)/cc-pCVQZ}) \text{ (all electron)}$$

core correlation

$$- E(\text{CCSD(T)/cc-pCVQZ}) \text{ (frozen core)}$$

$$+ E(\text{CCSDT/cc-pVTZ})$$

full T correction

$$- E(\text{CCSD(T)/cc-pVTZ})$$

$$+ E(\text{CCSDTQ/cc-pVDZ})$$

full Q correction

$$- E(\text{CCSDT/cc-pVDZ})$$

Additivity Assumptions

geometry optimization via following gradient

$$\mathbf{dE/dx} = \mathbf{dE(CCSD(T)/cc-pV6Z)/dx}$$

$$+ \mathbf{dE(CCSD(T)/cc-pCVQZ)/dx} \text{ (all electron)}$$

core correlation

$$- \mathbf{dE(CCSD(T)/cc-pCVQZ)/dx} \text{ (frozen core)}$$

$$+ \mathbf{dE(CCSDT/cc-pVTZ)/dx}$$

full T correction

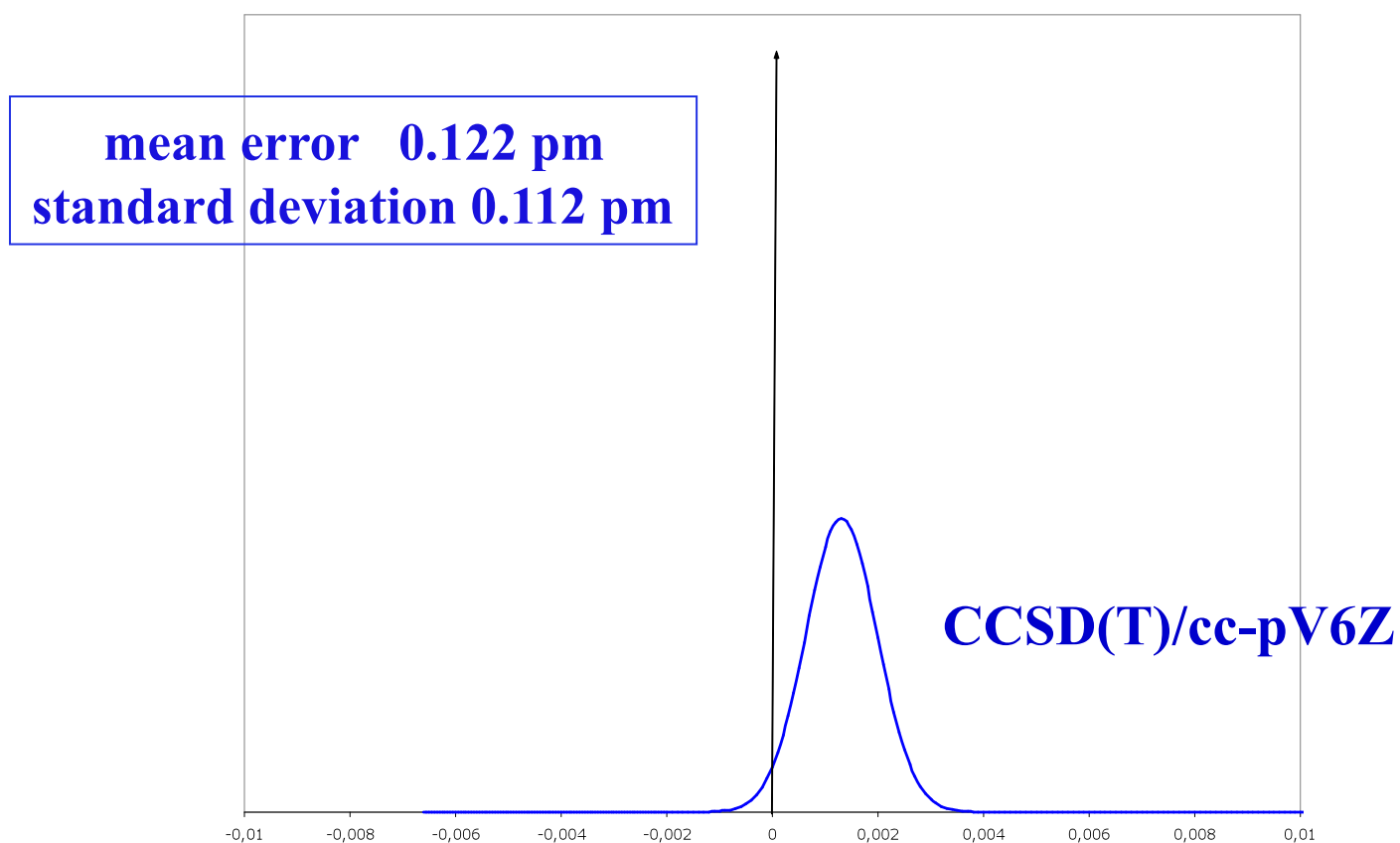
$$- \mathbf{dE(CCSD(T)/cc-pVTZ)/dx}$$

$$+ \mathbf{dE(CCSDTQ/cc-pVDZ)/dx}$$

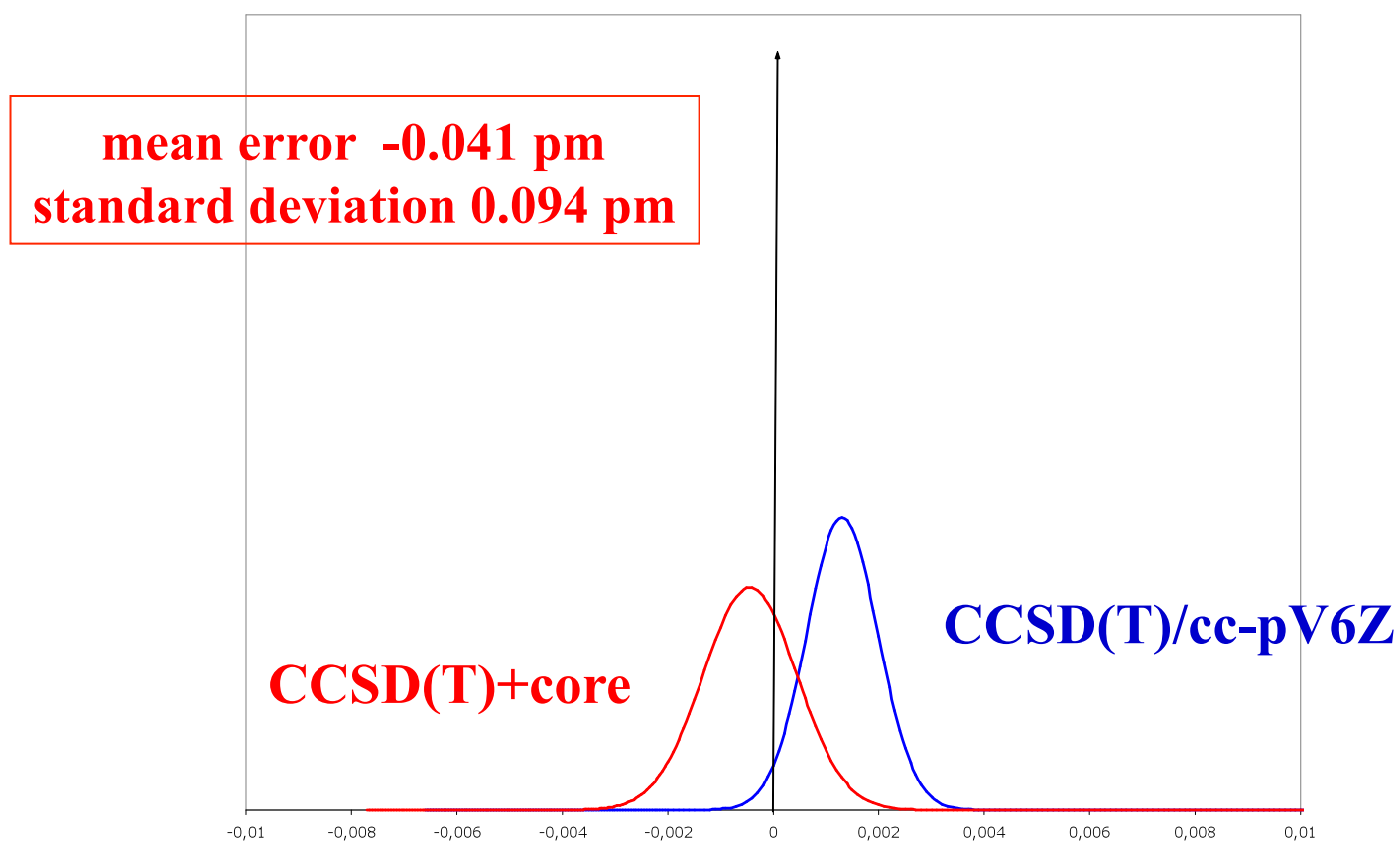
full Q correction

$$- \mathbf{dE(CCSDT/cc-pVDZ)/dx}$$

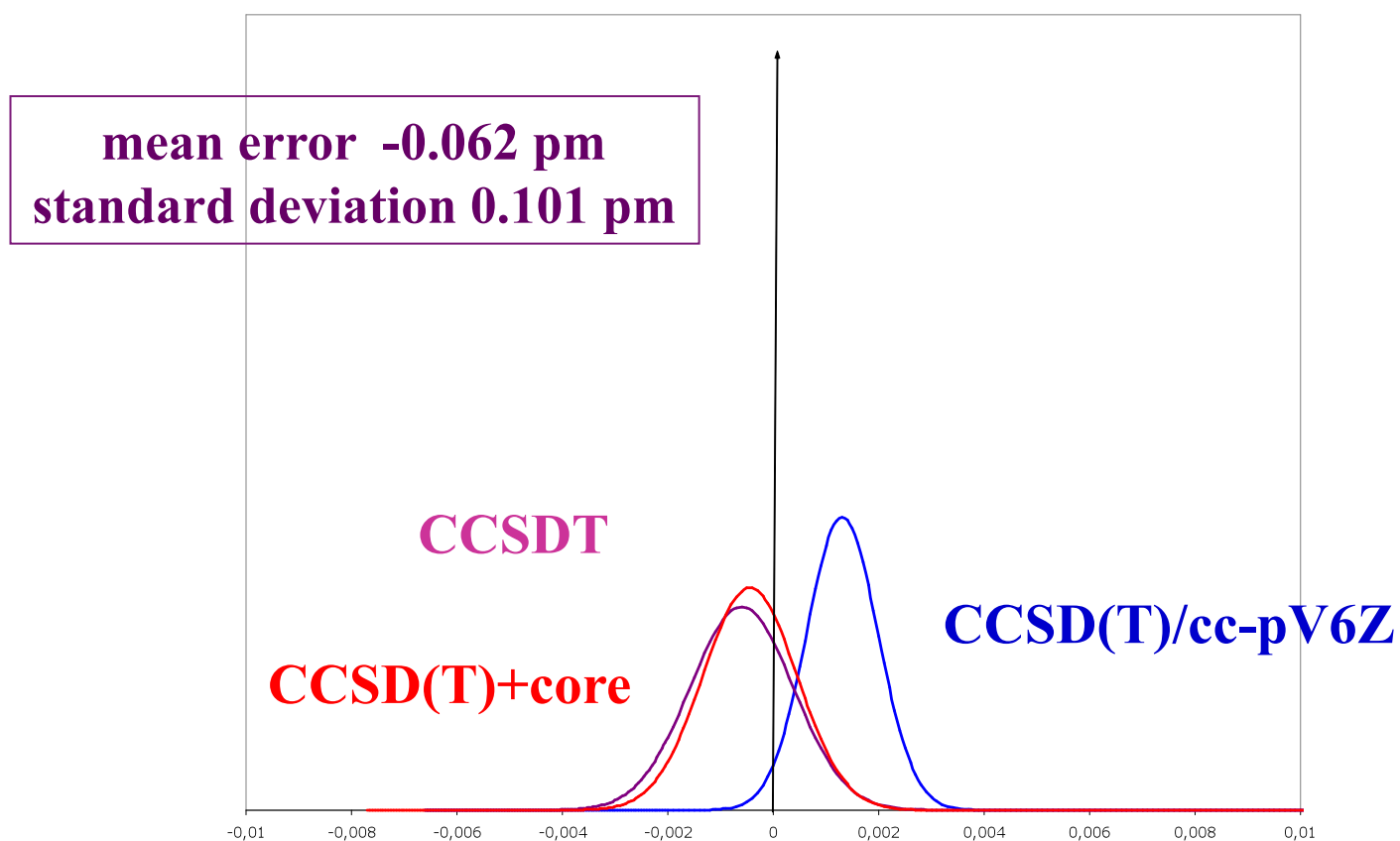
Importance of Higher Excitations



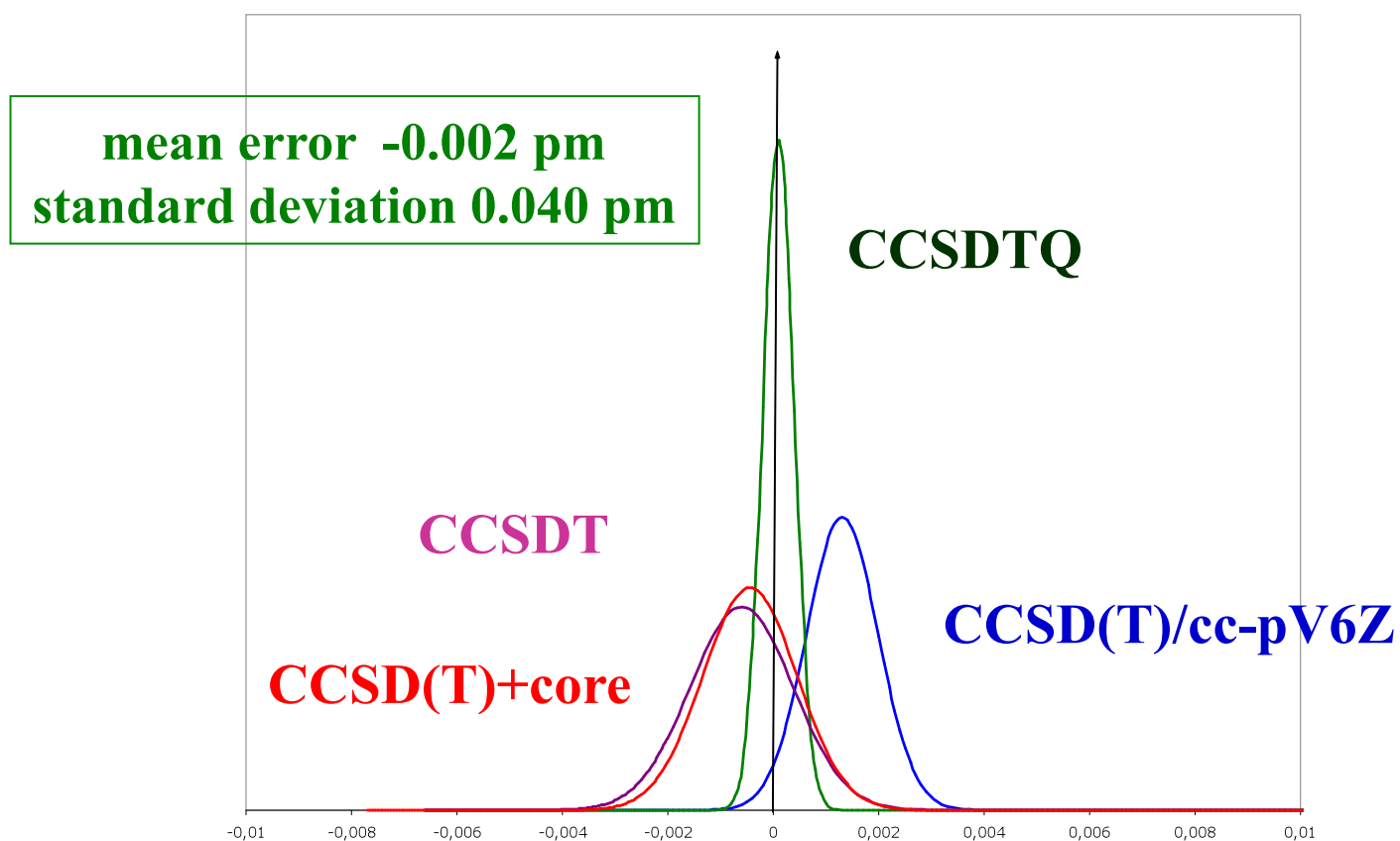
Importance of Higher Excitations



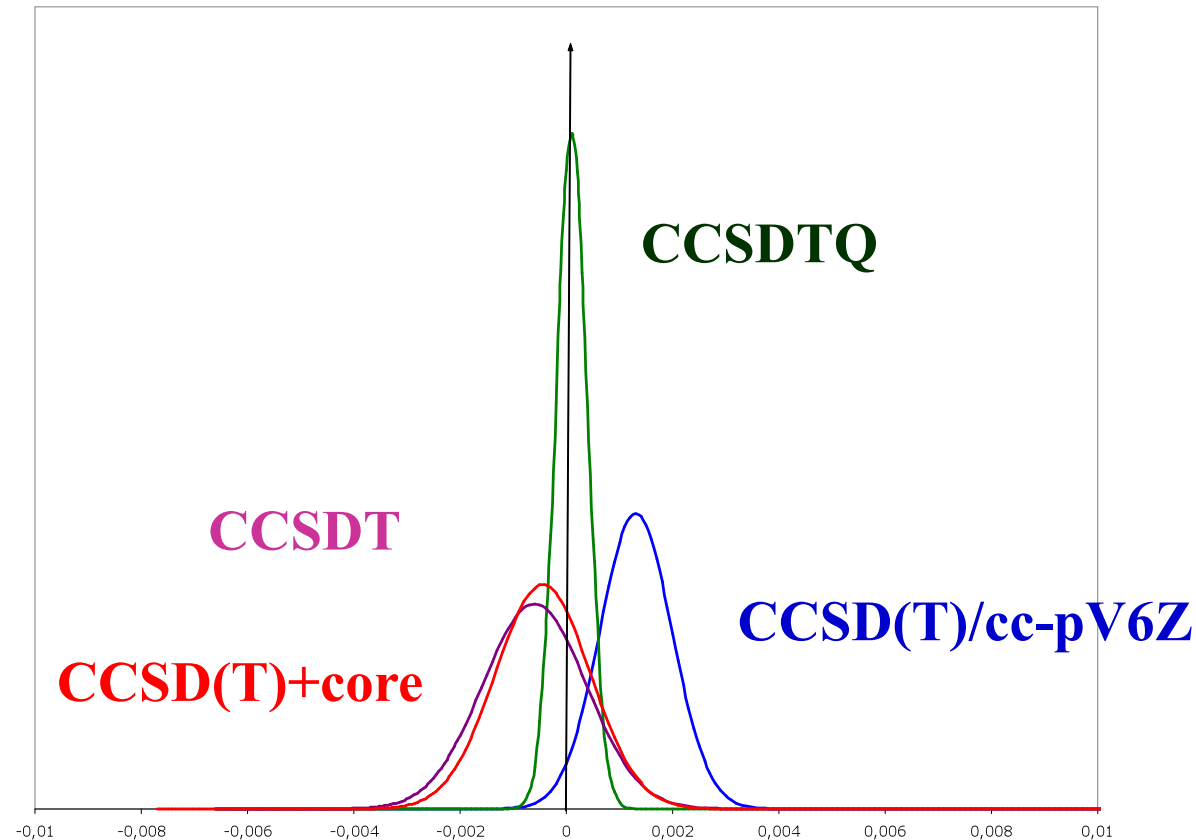
Importance of Higher Excitations



Importance of Higher Excitations



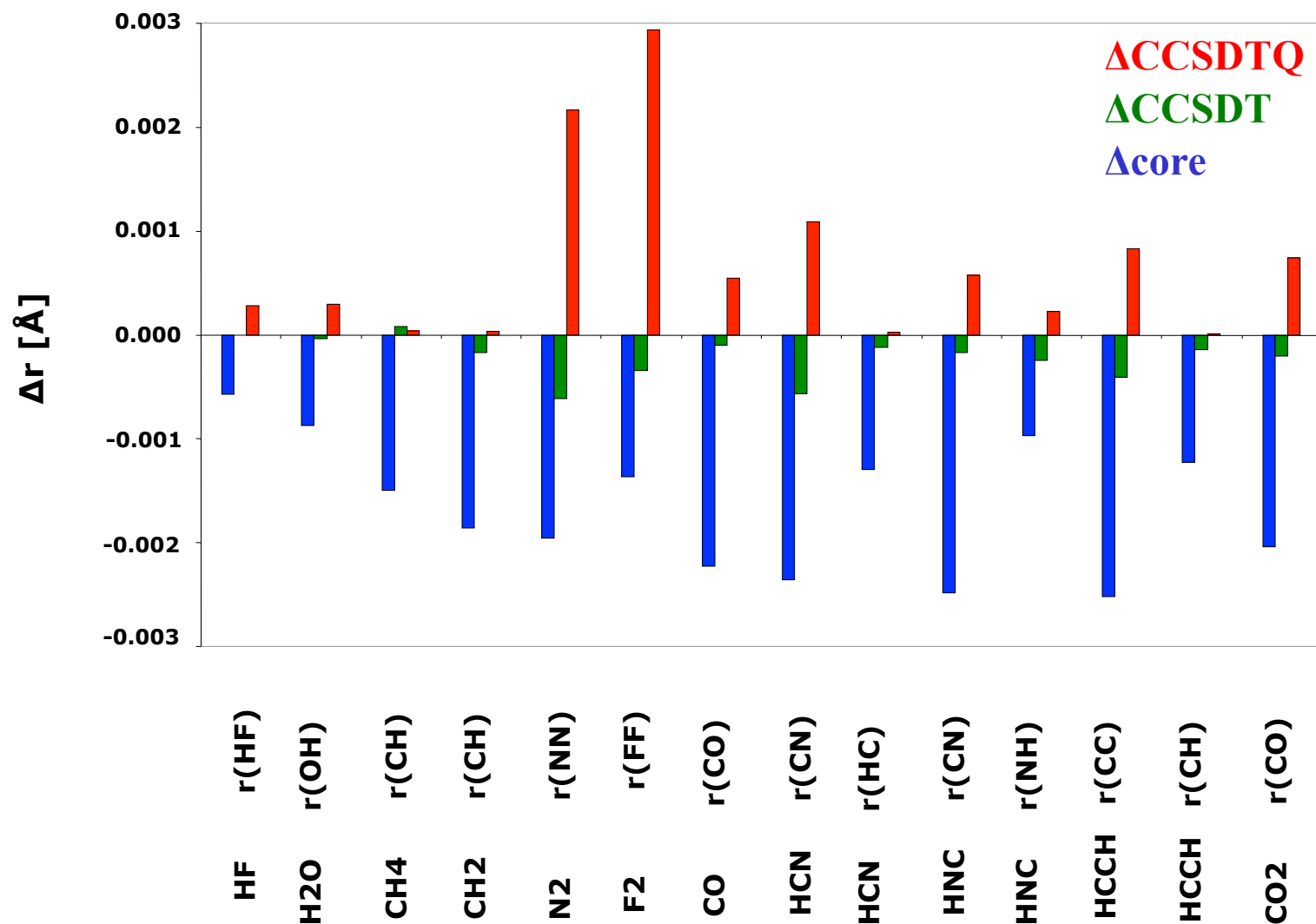
Importance of Higher Excitations



full CCSDT no improvement over CCSD(T)

CCSDTQ significantly more accurate than CCSD(T)

Analysis of Individual Contributions



Contribution of Pentuple Excitations

		Δr [pm]
HF	r(FH)	0.001
H₂O	r(OH)	0.002
CH₂	r(CH)	0.003
N₂	r(NN)	0.022
HCN	r(CN)	0.021
NO	r(NO)	0.021

**single
bonds**

**multiple
bonds**

CCSDTQP/cc-pVDZ contribution

Benchmark Study V

17 closed-shell and open-shell molecules

CH_2 , CH_4 , NH_3 , H_2O , HF , N_2 , $\text{C}\equiv\text{O}$, $\text{HC}\equiv\text{N}$,

$\text{HN}\equiv\text{C}$, $\text{HC}\equiv\text{CH}$, $\text{O}=\text{C}=\text{O}$, F_2

OH , $\text{CH}_2(^3\text{B}_1)$, NH_2 , CN , NO

wave function models

CCSD(T) , CCSDT , CCSDTQ

basis sets

cc-p(C)VXZ , X=T,Q,5,6

basis-set extrapolation + additivity assumptions

Basis-Set Extrapolation Techniques

extrapolation of HF-SCF energy

$$E^{\infty}(\text{HF-SCF}) = E(\text{HF-SCF/cc-pVXZ}) + a \exp(-bX)$$

extrapolated value



a , b , and $E^{\infty}(\text{HF-SCF})$ from
3 HF-SCF energies

Halkier et al.,
Chem. Phys. Lett. 302, 437 (1999)

extrapolation of CC energy

$$\Delta E^{\infty}(\text{CC}) = \Delta E(\text{CC/cc-pVXZ}) - \frac{c}{X^3}$$

extrapolated value



a and $\Delta E^{\infty}(\text{CC})$ from
2 CC energies

Helgaker et al.,
J. Chem. Phys. 106, 9639 (1997)

Basis-Set Extrapolation Techniques

extrapolation of HF-SCF gradient

$$\frac{dE^\infty(\text{HF-SCF})}{dx} = \frac{dE(\text{HF-SCF/cc-pVXZ})}{dx} + \frac{da}{dx} \exp(-bX) - \frac{db}{dx} a X \exp(-bX)$$

extrapolated value

da/dx , db/dx , and $dE^\infty(\text{HF-SCF})/dx$
from 3 HF-SCF gradients

extrapolation of CC gradient

$$\frac{d\Delta E^\infty(\text{CC})}{dx} = \frac{d\Delta E(\text{CC/cc-pVXZ})}{dx} - \frac{dc}{dx} \frac{1}{X^3}$$

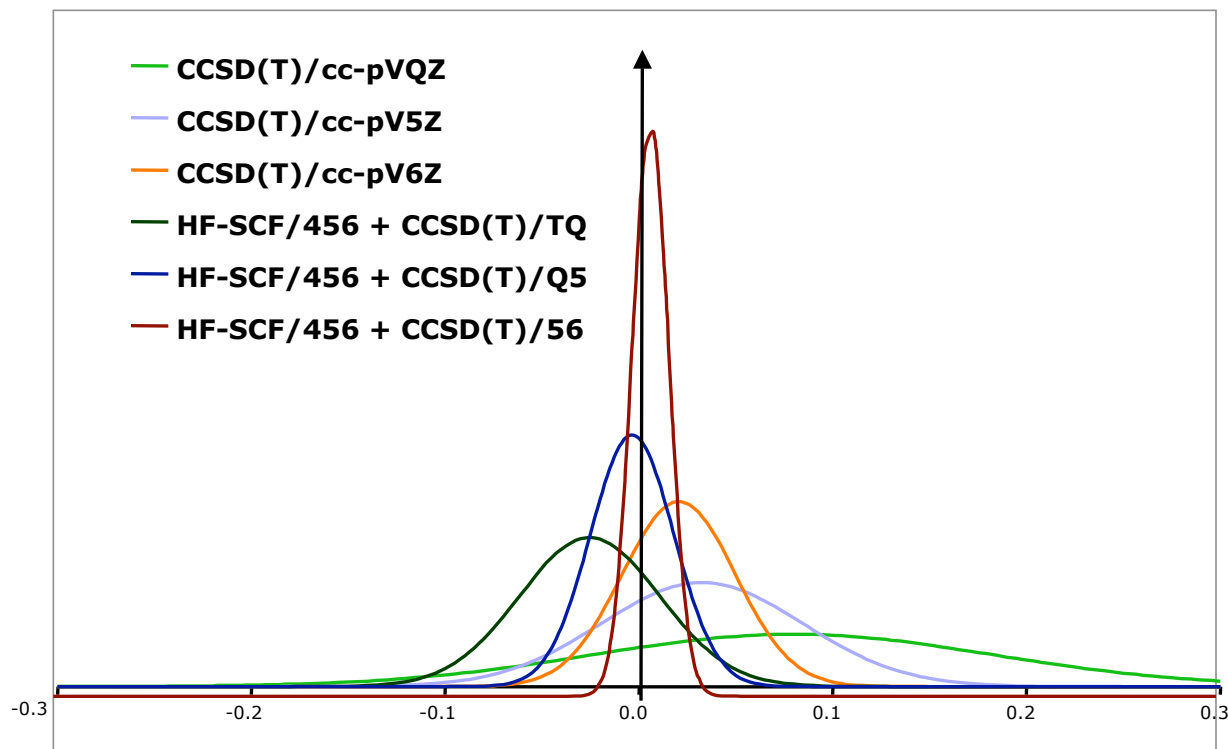
extrapolated value

da/dx and $\Delta dE^\infty(\text{CC})/dx$
from 2 CC gradients

Heckert et al., *J. Chem. Phys.* 125, 044108 (2006)

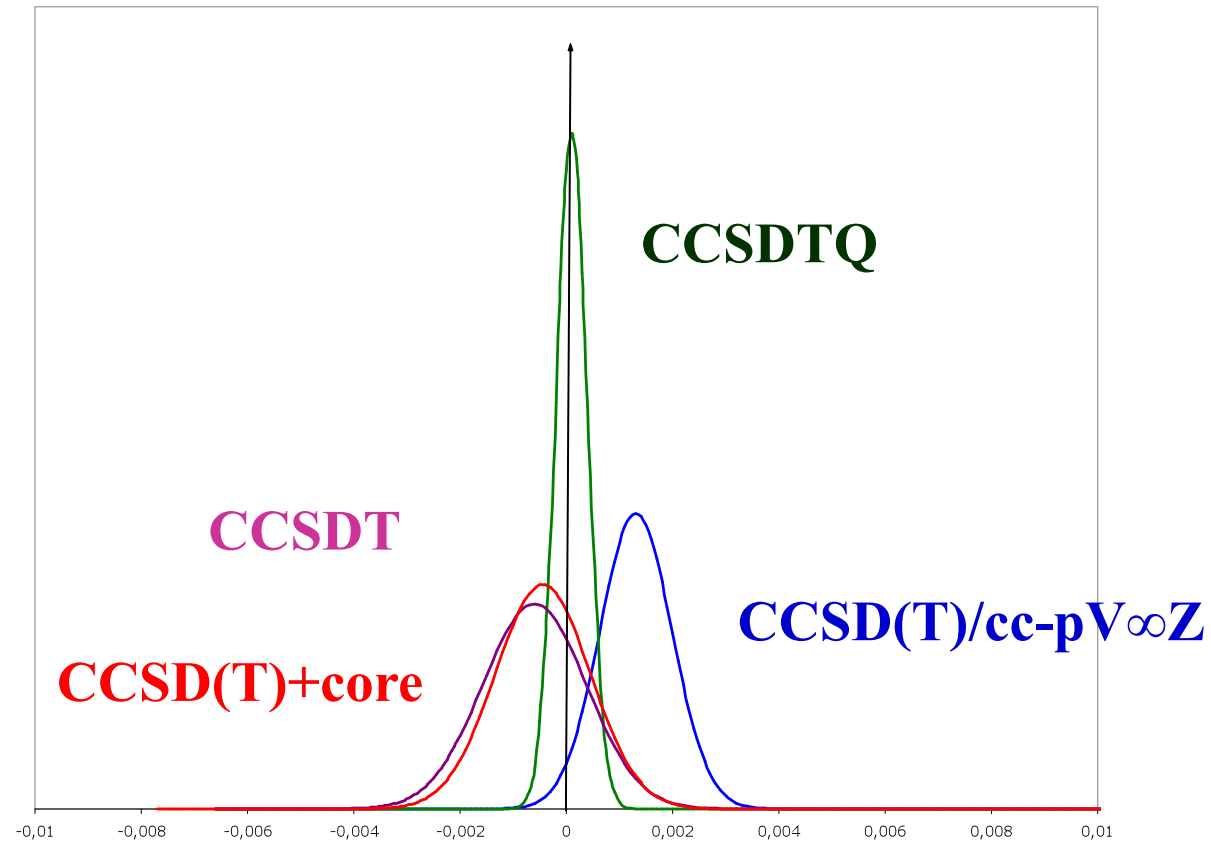
BSE in CCSD(T) Geometry Optimizations

comparison with CCSD(T)-R12



CCSD(T)/56: mean error 0.008 pm, stand. dev 0.009 pm

Core Correlation and Higher Excitations



mean error -0.006 pm, stand. dev 0.033 pm

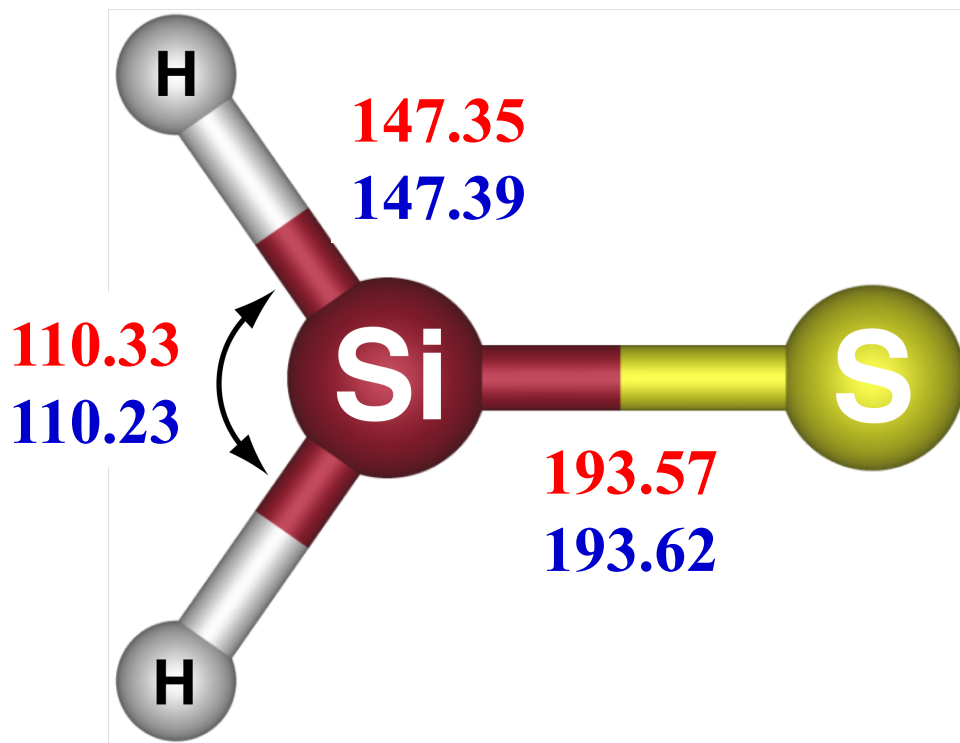
Comparison of Theory and Experiment

hydrogen cyanide (HCN)

	$\Delta r(\text{CN})$	$r(\text{CN})$
HF/∞		112.32
+ $\Delta\text{CCSD(T)}/\infty$	3.18	115.50
+ core correlation	-0.24	115.26
+ full triples	-0.06	115.20
+ quadruples	0.11	115.31
+ pentuples	0.02	115.33
Experiment		115.34

all values in pm

Example: Equilibrium Structure of $\text{H}_2\text{Si}=\text{S}$



experimental structure using CCSD(T)/cc-pV(Q+d)Z vibrational corrections

fc-CCSD(T)/cc-pV ∞ Z + Δ T/cc-pVTZ + Δ Q/cc-pVDZ + Δ core/cc-pCV5Z

Summary

HF/cc-pVDZ	~ 2 pm
MP2/cc-p(C)VTZ	~ 0.6 pm
CCSD(T)/cc-p(C)VQZ	~ 0.2 pm
CCSDTQ/cc-p(C)V∞Z	< 0.05 pm

make your choice for your calculations!