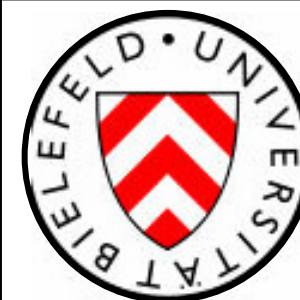


**The symbiosis between theory and experiment  
in gas phase structure determination:  
Methods for combination to combine  
the best from both worlds  
Norbert W. Mitzel**



UNIVERSITÄT  
BIELEFELD

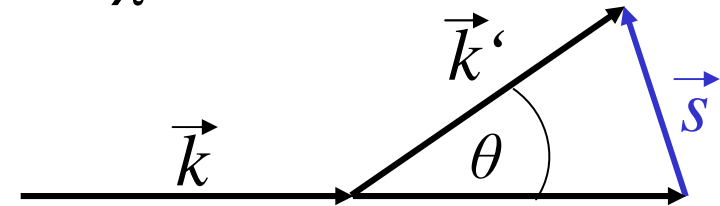
# Scattering theory

$$I = I_a + I_m + I_{\text{background}}$$

atomic molecular  
contribution

Change of momentum:

$$s = \frac{4\pi}{\lambda} \sin(\theta/2)$$



$$I_a = I_0 \frac{K^2}{R^2} \sum_{i=1}^N |f_i(s)|^2 \quad K = 8\pi^2 m e^2 / h^2$$

scattering amplitudes

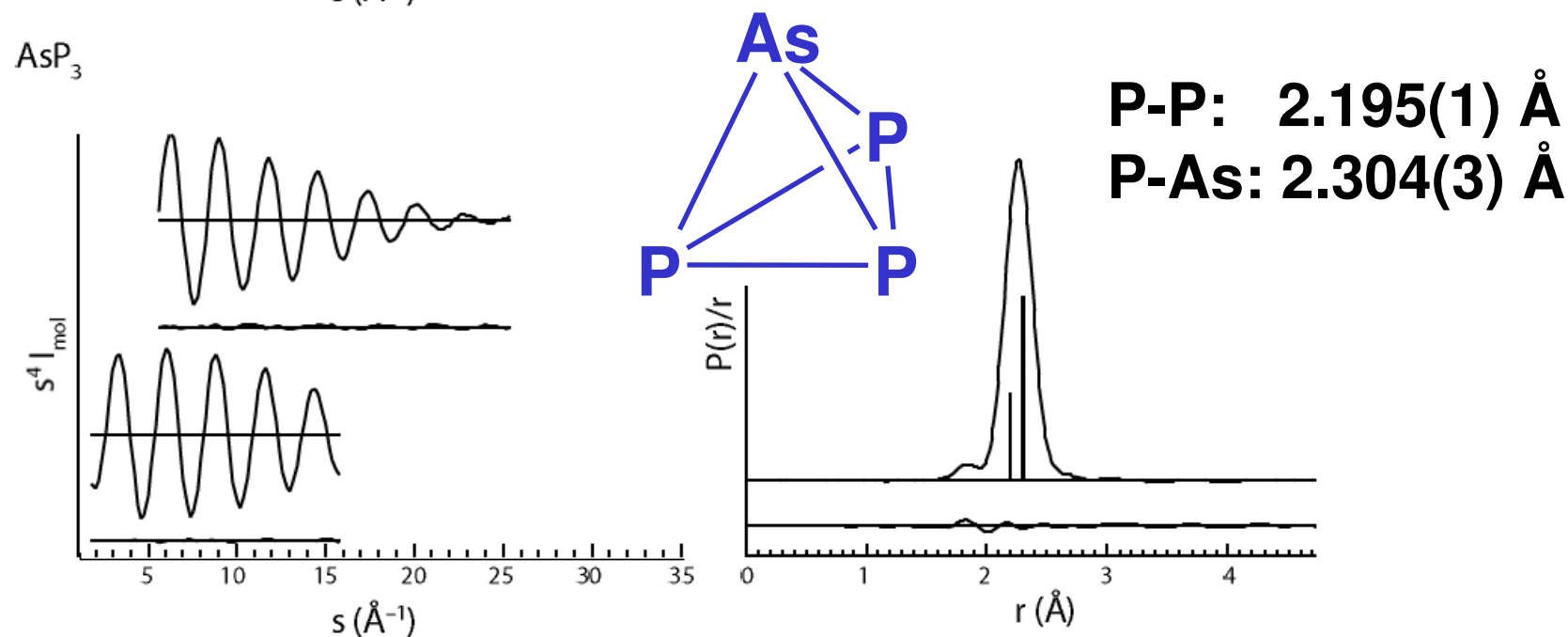
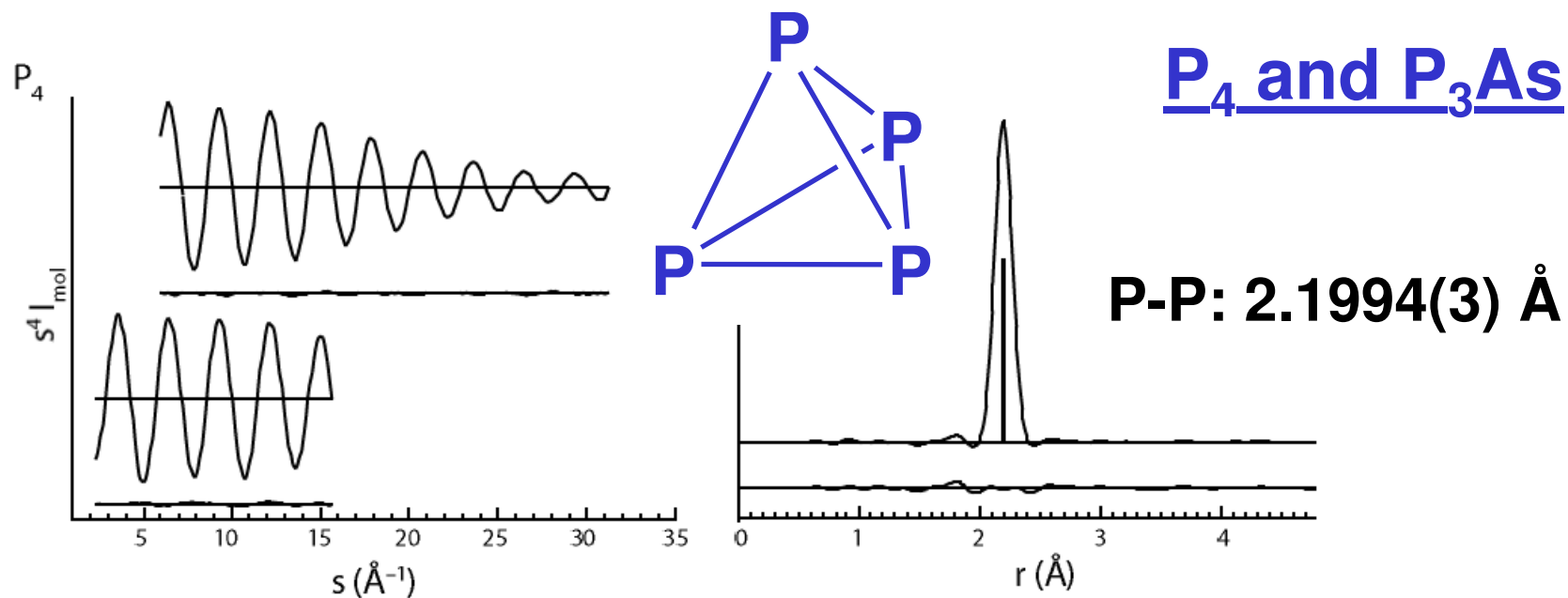
$$I_m = I_0 \frac{K^2}{R^2} \sum_{i=1}^N \sum_{\substack{k=1 \\ i \neq k}}^N |f_i(s)| |f_k(s)| \frac{\sin[s(\mathbf{r}_{ik} - \mathbf{k}_{ik} s^2)]}{s \mathbf{r}_{ik}} \cdot e^{-(1/2) \mathbf{l}_{ik} s^2} \cdot \cos[\eta_i(s) - \eta_k(s)]$$

distance scattering point detector

atomic distances anharmonicity parameter

scattering amplitude phases

r.m.s.-vibrational amplitudes

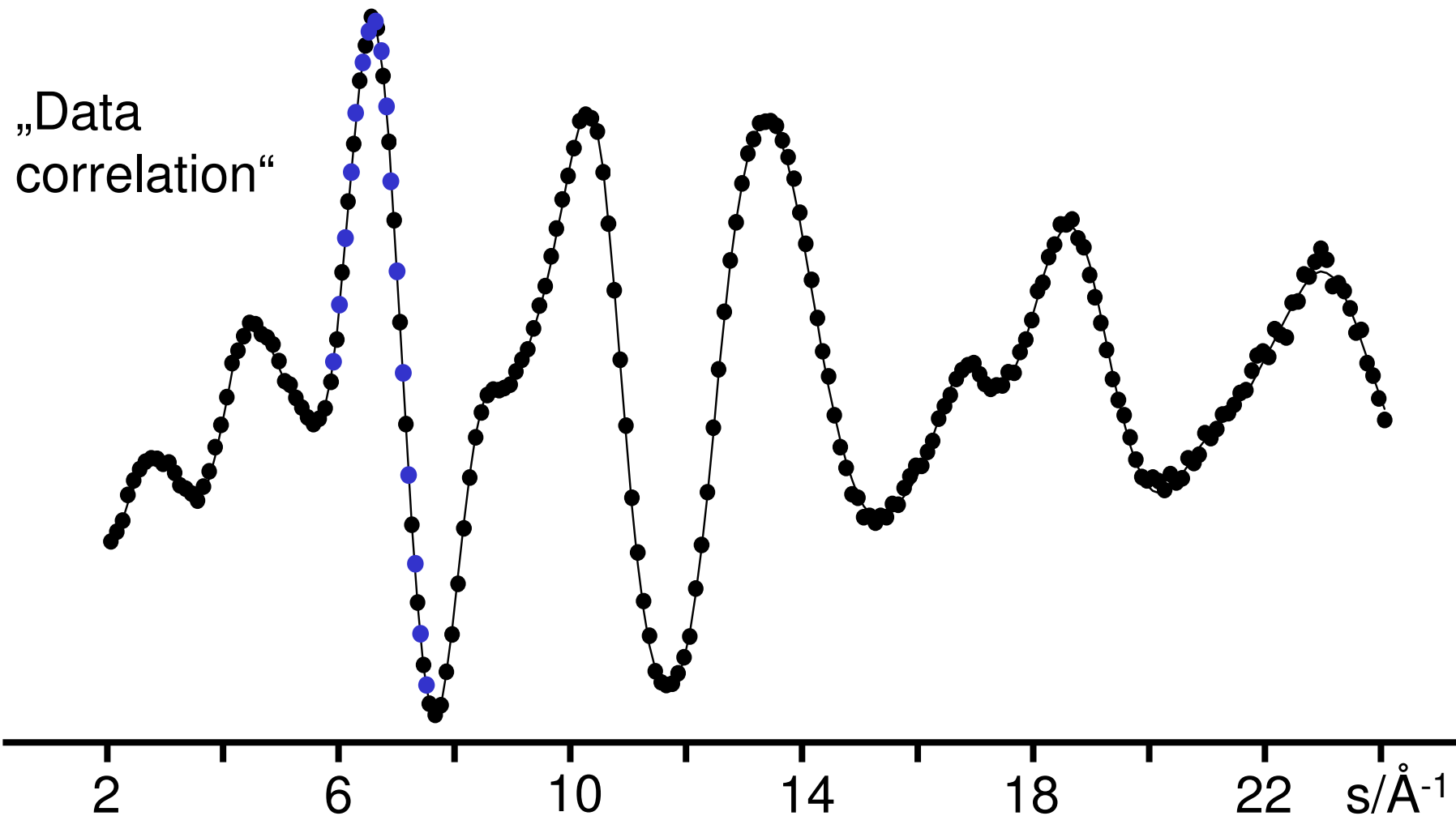


B. M. Cossairt, C. C. Cummins, A. R. Head, D. L. Lichtenberger, R. J. F. Berger, S. A. Hayes, N. W. Mitzel, G. Wu, *J. Am. Chem. Soc.* **2010**, 132, 8459.

## Structure description / Model / Degrees of freedom

A GED-Experiment gives about 150 to 300 data points

Are they really independent observations?



## Structure description / Model / Degrees of freedom

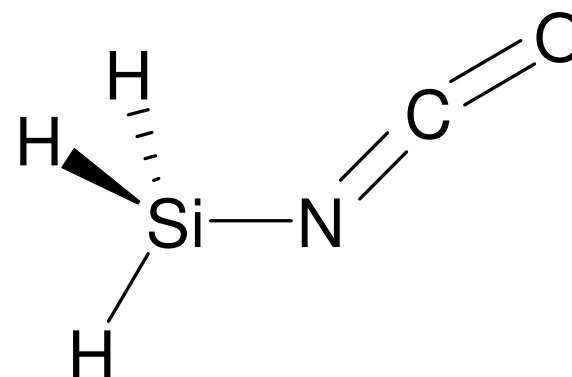
How many parameters are needed to describe a structure?

Degrees of freedom:  $3N - 6$

e.g.  $\text{H}_3\text{Si-N}=\text{C}=\text{O}$ ; 7 Atoms;  $3 \times 7 - 6 = 15$  Parameters

Z-Matrix:

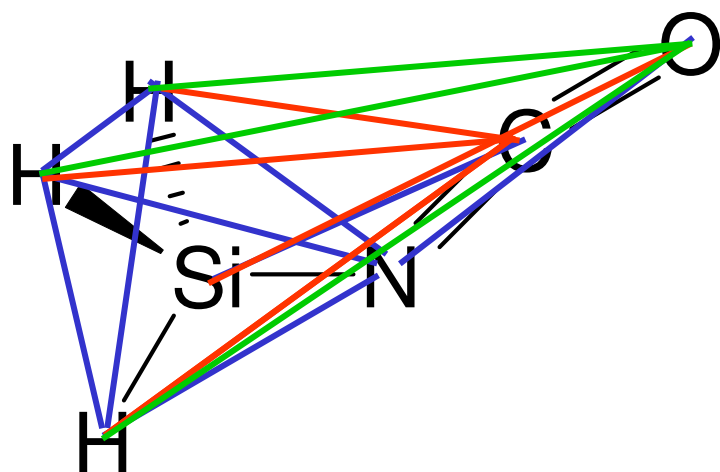
Distance	Angle	Torsion
$r(\text{Si-N})$		
$r(\text{N-C})$	$\angle(\text{Si-N-C})$	
$r(\text{C-O})$	$\angle(\text{N-C-O})$	$\tau(\text{Si-N-C-O})$
$r(\text{Si-H}^1)$	$\angle(\text{N-Si-H}^1)$	$\tau(\text{C-N-Si-H}^1)$
$r(\text{Si-H}^2)$	$\angle(\text{N-Si-H}^2)$	$\tau(\text{C-N-Si-H}^2)$
$r(\text{Si-H}^3)$	$\angle(\text{N-Si-H}^3)$	$\tau(\text{C-N-Si-H}^3)$



## But ... Information about vibrations missing

Again: 3 N - 6	Stretch	Bend	Torsion
	$r(\text{Si-N})$		
	$r(\text{N-C})$	$\angle(\text{Si-N-C})$	
	$r(\text{C-O})$	$\angle(\text{N-C-O})$	$\tau(\text{Si-N-C-O})$
	$r(\text{Si-H}^1)$	$\angle(\text{N-Si-H}^1)$	$\tau(\text{C-N-Si-H}^1)$
	$r(\text{Si-H}^2)$	$\angle(\text{N-Si-H}^2)$	$\tau(\text{C-N-Si-H}^2)$
	$r(\text{Si-H}^3)$	$\angle(\text{N-Si-H}^3)$	$\tau(\text{C-N-Si-H}^3)$

Or: amplitude of vibration for each atom-pair



— one-bond amplitudes	6
— two-bond amplitudes	8
— three-bond amplitudes	4
— four-bond amplitudes	3
	<hr/>
	21

# Least Squares

- linearisation of the functional relationship between observations  $y_i$  and parameters  $x_i$
- assuming a set of initial values  $x_i^{(0)}$
- set of equations of the form (truncated Taylor series)

$$r_i = y_i - \left[ f_i(x_1^{(0)}, \dots, x_m^{(0)}) + \sum_j \frac{\delta f_i}{\delta x_j} \bigg|_{x^{(0)}} (x_j - x_j^{(0)}) \right]$$

residual                      parameters  
   observation

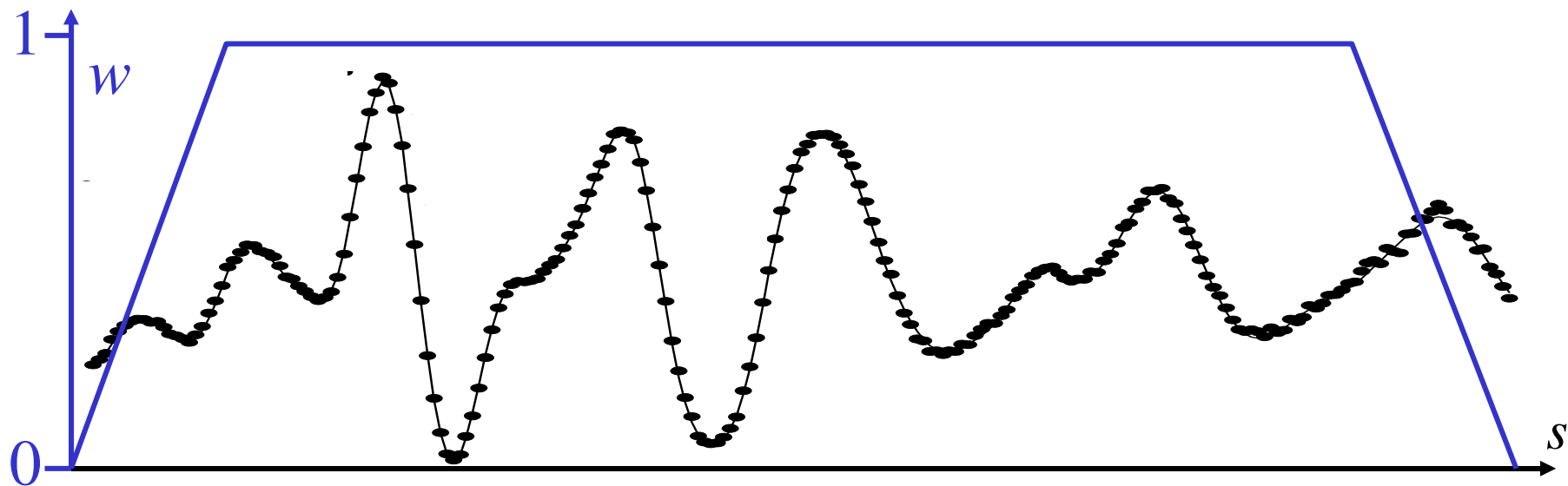
# Least Squares

Quantity to be minimised in GED:

$$\sum_{k=1}^n [s_k I_k^{(\text{exp})}(s_k) - s_k I_k^{(\text{model})}(s_k)]^2 w_k$$

Goodness of fit:

$$R = \sqrt{\frac{\sum_{k=1}^n [s_k I_k^{(\text{exp})}(s_k) - s_k I_k^{(\text{model})}(s_k)]^2 w_k}{\sum_{k=1}^n [s_k I_k^{(\text{exp})}(s_k)]^2}}$$





# Weighted Least Squares

**Restraint:** Condition with an assigned weight

$$w = 1/\sigma^2 \quad \sigma = \text{uncertainty}$$

**Constraint:** Condition that has to be met exactly

$$w = \infty$$

Quantity to be minimised in combined analyses:

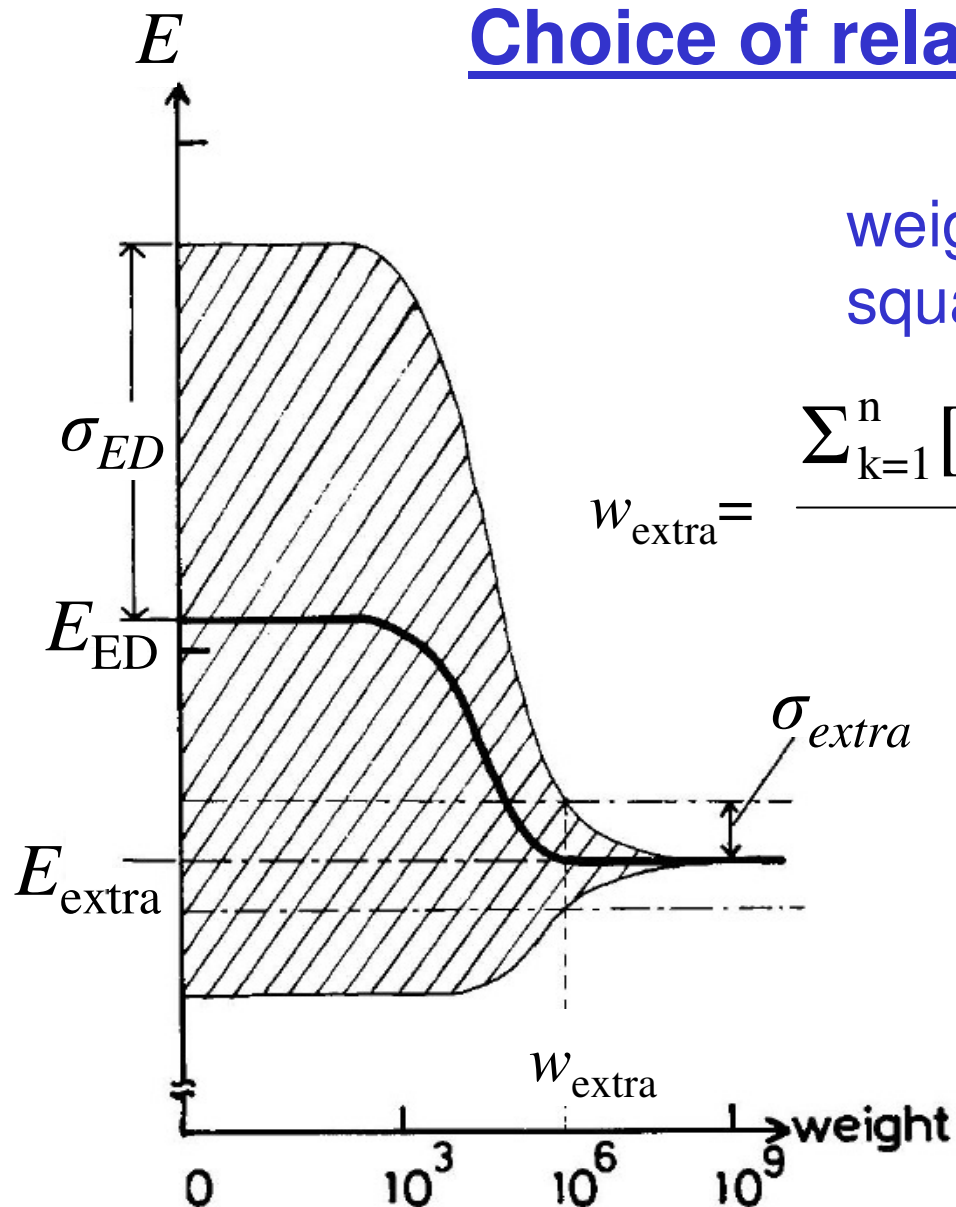
$$\sum_{k=1}^n [s_k I_k^{(\text{exp})}(s_k) - s_k I_k^{(\text{model})}(s_k)]^2 w_k + \sum_j [E_j^{(\text{exp})} - E_j^{(\text{model})}]^2 w_j$$

|  
extra information

## Choice of relative weights

weighted sum of squared residuals of n GED data

$$w_{\text{extra}} = \frac{\sum_{k=1}^n [s_k I_k^{(\text{exp})}(s_k) - s_k I_k^{(\text{model})}(s_k)]^2 w_k}{n \sigma_{\text{extra}}^2}$$



$n \sigma_{\text{extra}}^2$   
 ↑  
 σ of extra „observation“  
 (e.g. rotation constant,  
 calculated distance etc.)

## Bartell's predicate observations<sup>1</sup>

### **R. F. Curl: „diagnostic least-squares procedure“.<sup>2</sup>**

This used flexible restraints based on previous general knowledge, but was specifically applied to structures for which some rotation constants were available, but not enough to allow full structure determination.

### **L. S. Bartell: General application of restraints**

as auxiliary conditions in the optimisation procedure (least-squares procedure) for structural parameters

Restraints represent any knowledge about the problem  
e.g. the range of reasonable parameter values

1. L. S. Bartell, D. J. Romenensko, T. C. Wong, In: Molecular Structure by Diffraction Methods; Specialist Periodical Reports; The Chemical Society: **1975**; Vol. 3, p 72

2. R. F. Curl, *J. Comput. Phys.* **1970**, 6, 367

## MOCED

Molecular orbital constrained ED

Use of QM calculations  
as source of **constraints**

(tacitly assuming they are correct)

**GED:**

Weakness:

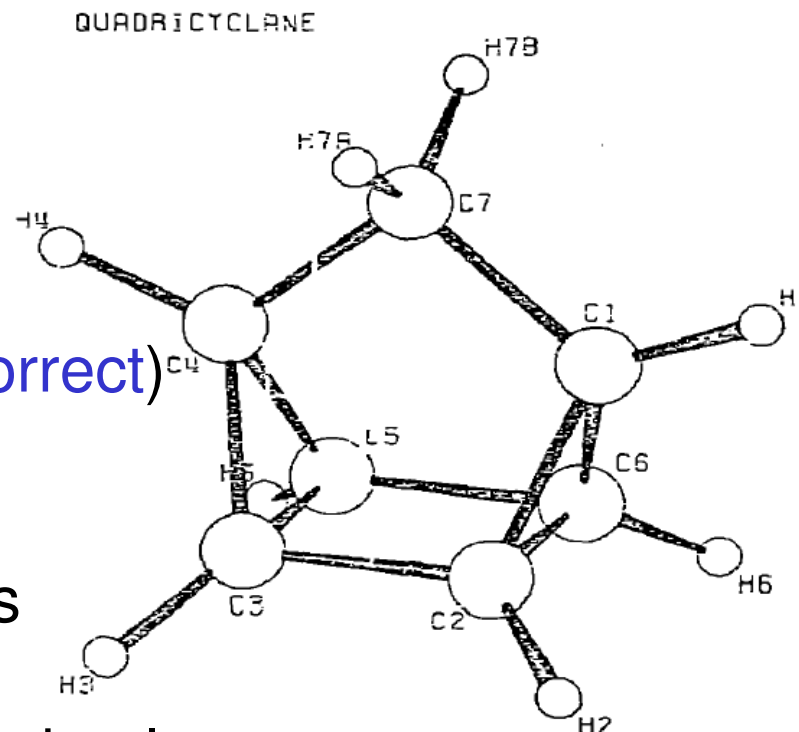
Resolving closely spaced distances

Strength:

Determination of mean values and absolute  
scaling

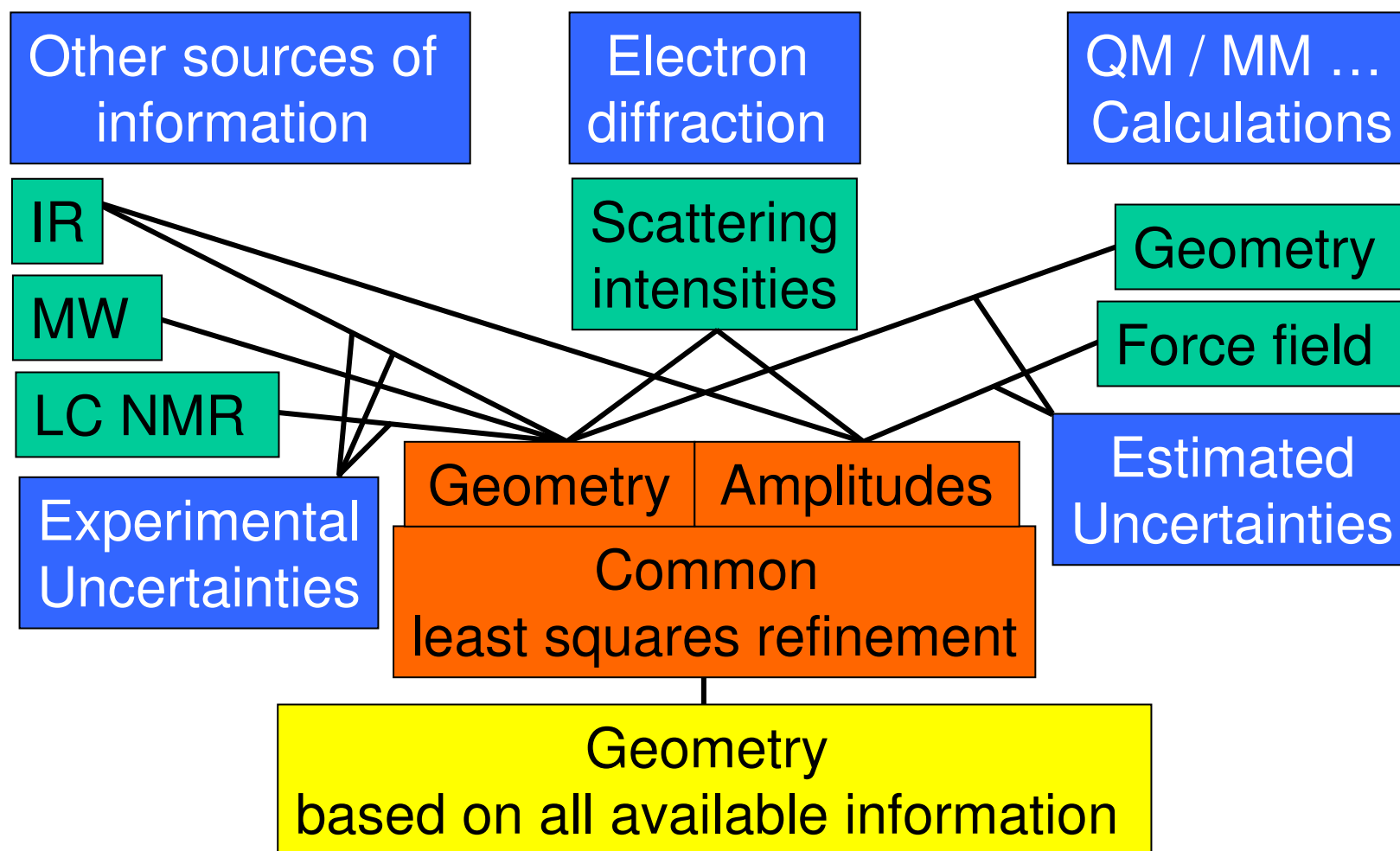
**QM calculations:**

Differences between comparable parameters  
resolved with particular accuracy



# SARACEN

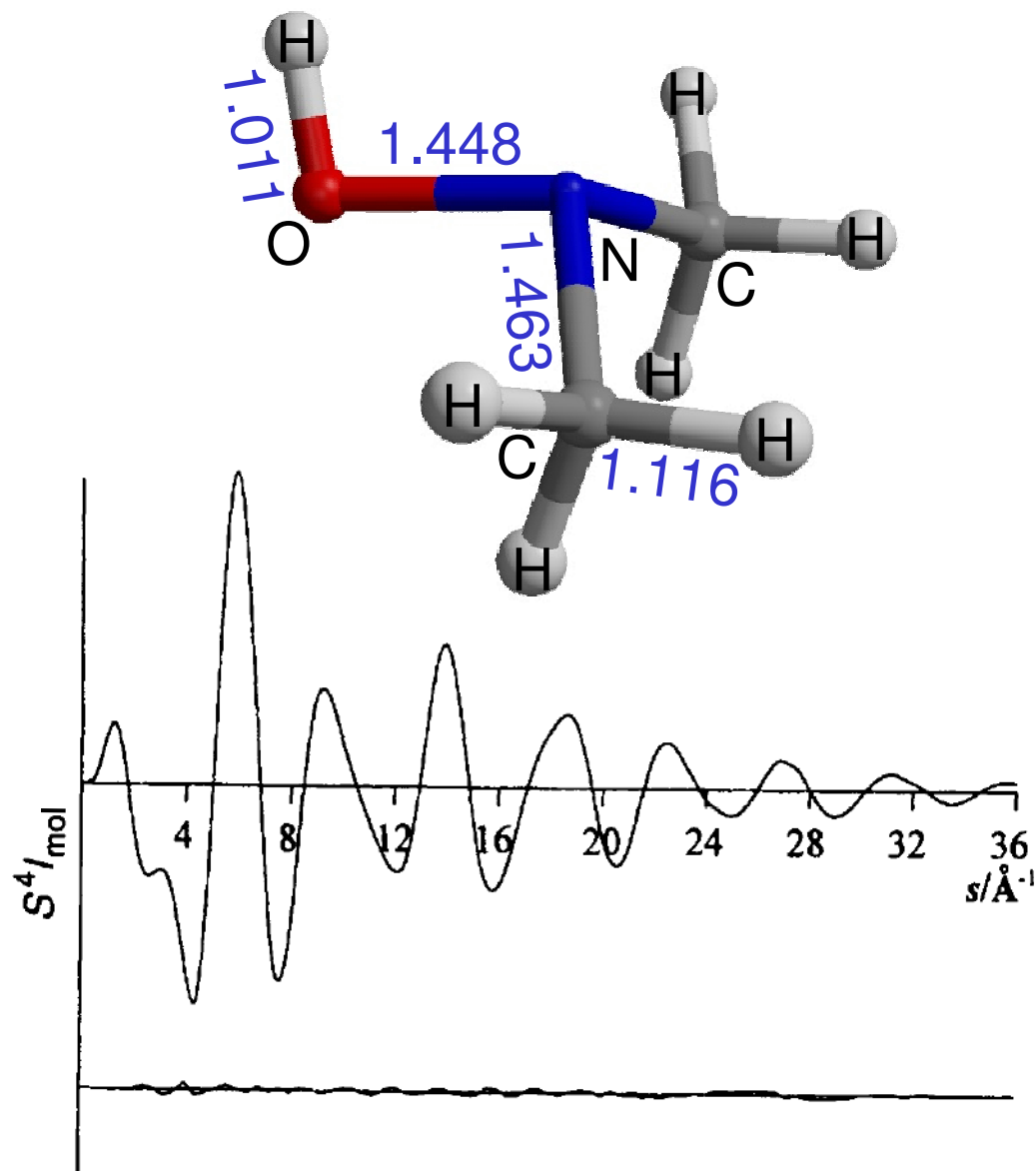
**Structural Analysis Restrained by Ab initio  
Calculations for Electron diffraction**



P. T. Brain, C. A. Morrison, D. W. H. Rankin, et al., *J. Phys. Chem.* **1996**, *100*, 12280

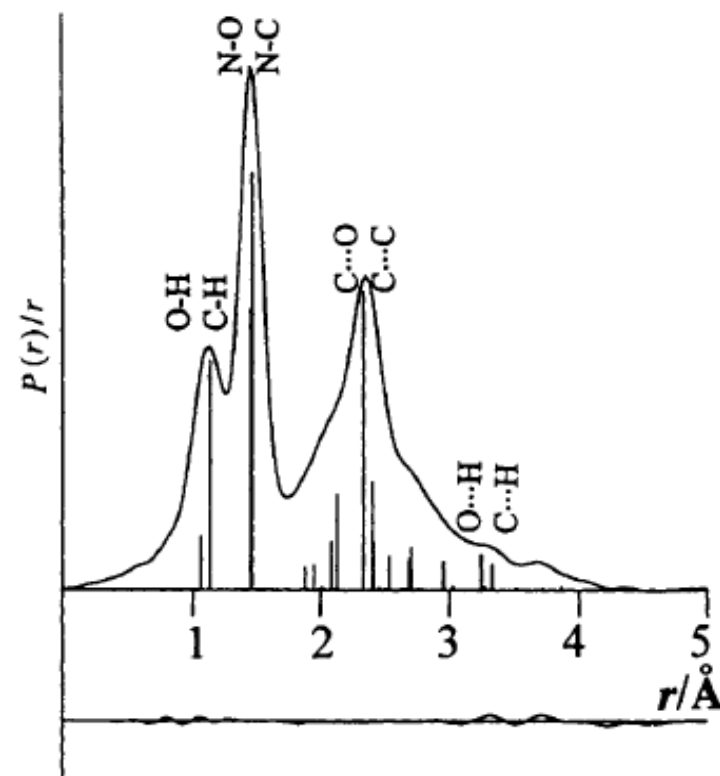
Review: N. W. Mitzel, D. W. H. Rankin, *Dalton Trans.* **2003**, 3650

## Example 1: Me<sub>2</sub>NOH



Model assumptions:

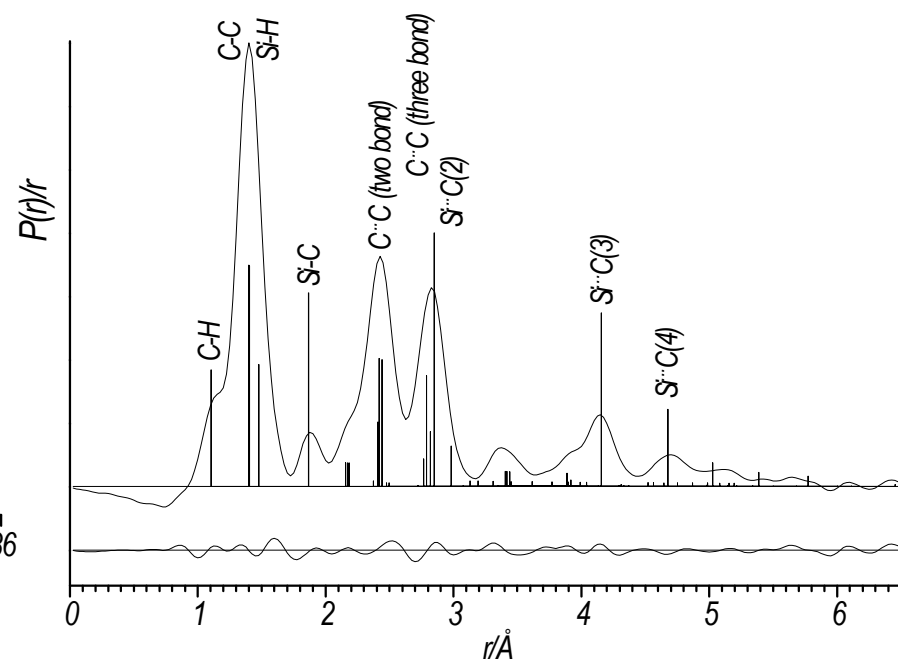
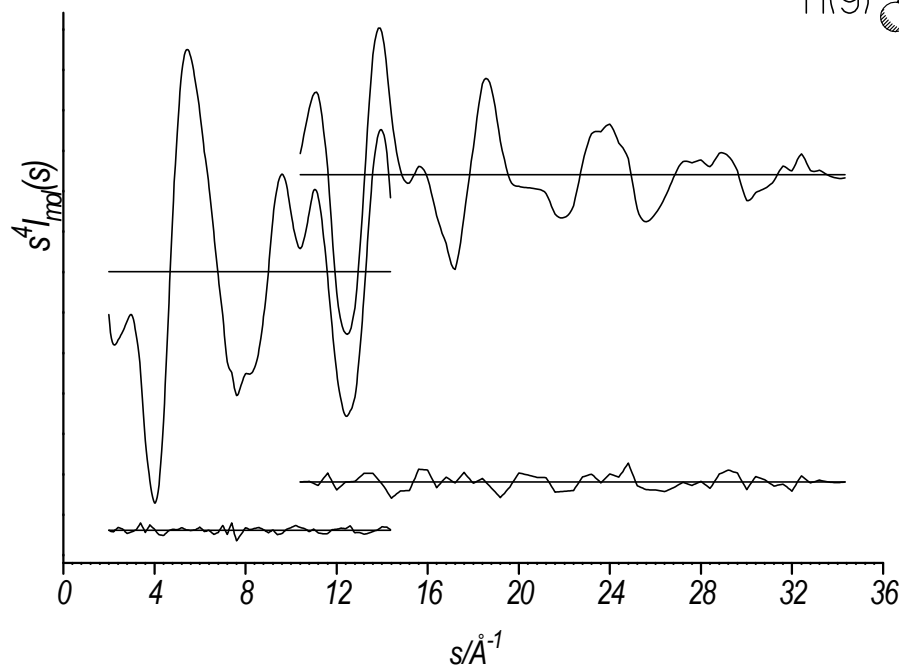
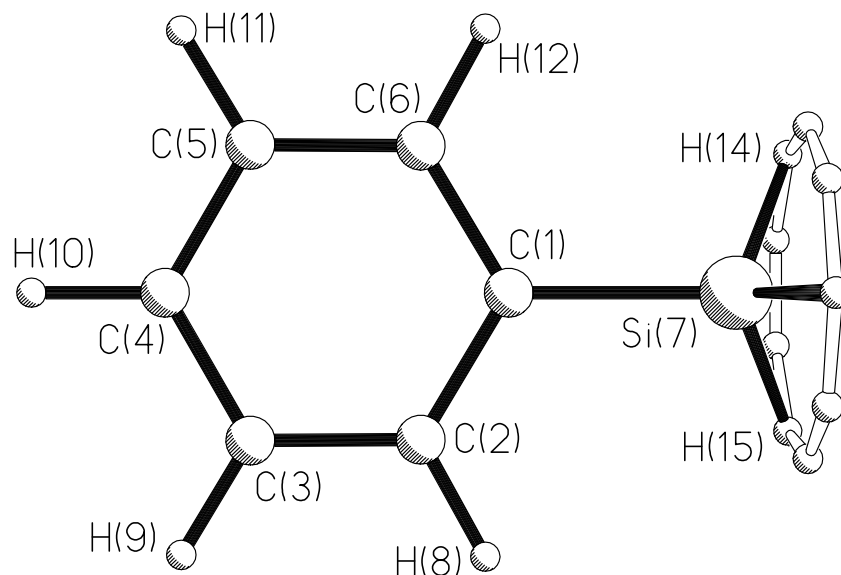
- $C_s$  symmetry
- all  $r(\text{CH})$  equal



parameter	without restraint $r_g$	SARACEN $r_g$	restraint value	MP2/ 6-311G(d,p) $r_e$
<i>independent:</i>				
$r_{\text{mean}}(\text{NC/NO})$	1.453(1)	1.458(1)		1.453
$\Delta r(\text{NC/NO})$	0.058(11)	-0.014(17)	-0.016(20)	-0.016
$r_{\text{av}}(\text{CH})$	1.122(2)	1.116(6)		1.091-1.101
$\Delta r(\text{CH/OH})$	-0.135 fix	-0.105(31)	-0.135(40)	-0.135
$\angle(\text{CNC})$	112.2(17)	110.4(15)		110.3
$\angle(\text{CNO})$	106.4(7)	106.7(3)		105.4
$\angle(\text{NOH})$	101.5 fix	101.4(15)	101.0(20)	101.0
$\angle_{\text{av}}(\text{NCH})$	110.0 fix	107.4(7)	110.0(15)	108.3-111.8
$\tau(\text{CH}_3)$	1.4(98)	9.3(13)		
$\delta(\text{CH}_3)$	-2.2 fix	-2.3(12)	-2.2(15)	
<i>dependent:</i>				
$r(\text{NO})$	1.492(8)	1.448(11)		1.442
$r(\text{NC})$	1.434(4)	1.463(7)		1.458
$r(\text{OH})$	0.987(2)	1.011(25)		0.960
$R$	5.62%	4.88%		

## Example 2: Phenylsilane

rotational barrier:  
74 J/mol



N.W. Mitzel, P.T. Brain, M. Hofmann,  
D.W.H. Rankin, R. Schröck, H. Schmidbaur, *Z. Naturforsch.* **2001**, 57b, 3650

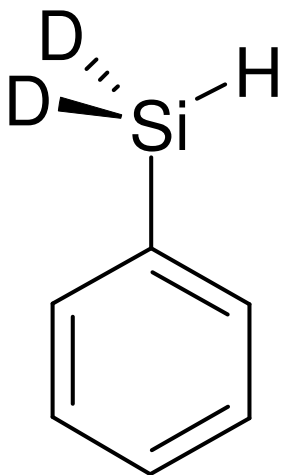


## Phenylsilane

Parameter	GED $r_{\alpha}^0$ (new)	MP2/ 6-311G(d,p) $r_e$	GED $r_a$ (old)*
C(1)-C(2)	1.403(3)	1.409	1.410(3)
C(2)-C(3)	1.396(2)	1.399	1.400 <sup>a</sup>
C(3)-C(4)	1.397(3)	1.399	1.400 <sup>a</sup>
Si(7)-C(1)	1.863(3)	1.876	1.870(4)
Si(7)-H	1.466(3)	1.478	1.497(7)
Si-C(1)-C(2)	120.9(2)	121.0	121.0
C(6)-C(1)-C(2)	118.2(2)	118.0	118.1(2)
C(1)-C(2)-C(3)	121.1(3)	121.1	121.1(1)
C(2)-C(3)-C(4)	119.7(4)	120.0	119.9(2)
C(3)-C(4)-C(5)	120.0(5)	119.8	120.0(3)

\*F. A. Keidel, S. H. Bauer, *J. Chem. Phys.* **1956**, 25, 1218

## Excuse: Isolated SiH stretching frequencies



$$r_e(\text{SiH}) = 1.9089 - 0.0001998 \nu(\text{SiH})$$

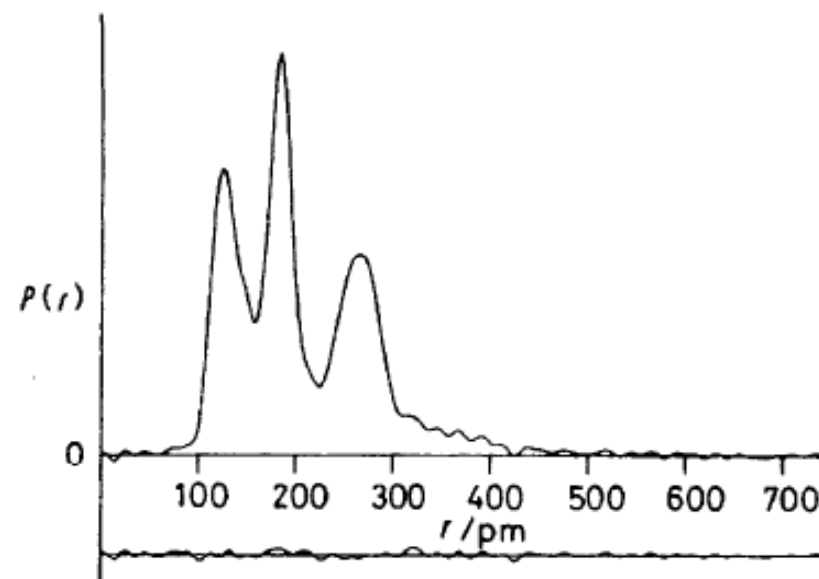
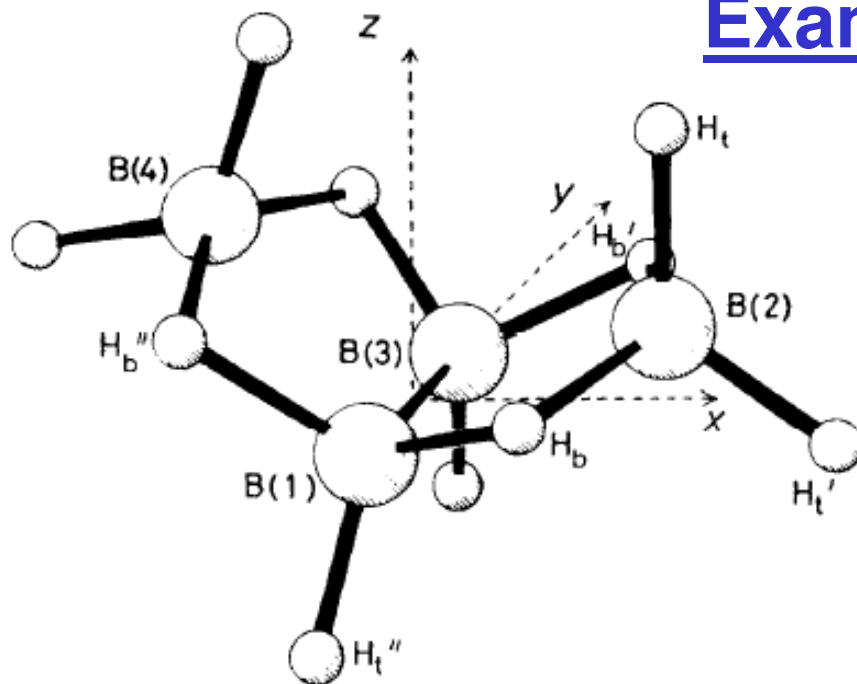
Originally developed for CH:  
„The solitary CH bond stretching frequency is effectively decoupled from all other motions in the molecule and is unaffected by the customary Fermi resonances“

## Phenylsilane - Restraints

No.	description	value, $r_\alpha^0$ Å, °	restraint	source
1	$rC(1)-C(2)-rC(2)-C(3)$	0.007(5)	0.010(5)	MP2
2	$rC(3)-C(4)-rC(1)-C(2)$	0.001(5)	0.0(5)	MP2
3	$r(\text{Si-H})$	1.466(3)	1.466(3)	IR
4	$\angle[C(1)-C(2)-H(8)]$	120.2(13)	119.9(15)	MP2
5	$\angle[C(2)-C(3)-H(9)]$	121.4(12)	119.9(15)	MP2
6	$\angle[C(6)-C(1)-C(2)]$	118.2(2)	$p_7-p_{11}=-3.1(5)$	MP2
$B_z$	(MHz)			
	$B$	1499.234	1499.234(34)	MW
	$C$	1187.058	1187.058(7)	MW

$B_z$  calcd with old GED structure:  
B: 1515 and C: 1198 MHz

## Example 3: Tetraborane(10)

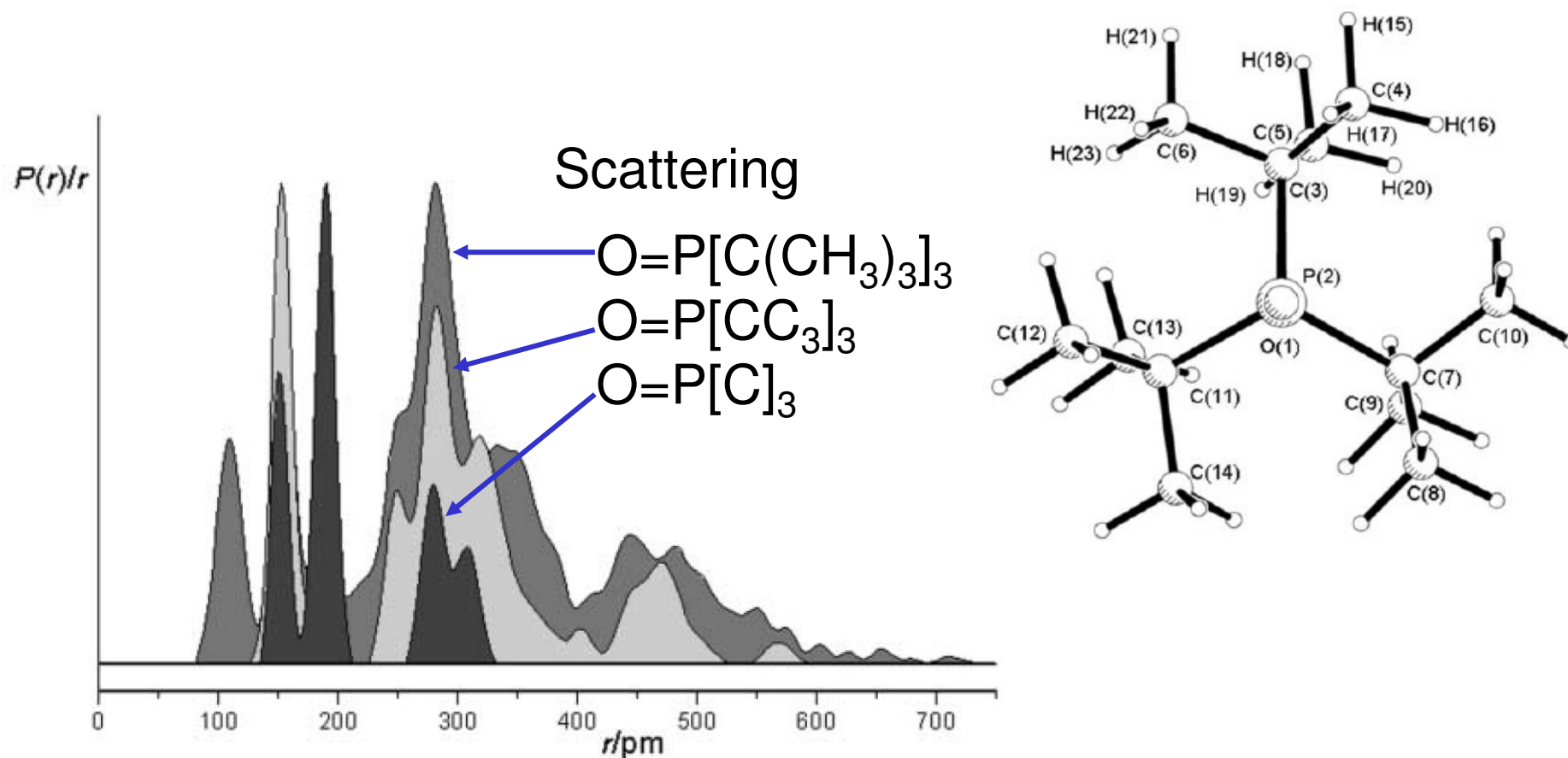


Parameter	GED $r_a$	Rot.- const	measured	calculate GED/MW
B(1)-B(3)	1.705(12)			
B(1)-B(2)	1.856(4)	$B_z$	5592.817(21)	5594.171
B(1)-B(2)-B(3)	54.7(6)	$B_y$	6198.643(23)	6200.627
B(2)-B(1)-B(3)	62.7(6)	$B_x$	11013.388(19)	11015.213
inter-planar angle	170.3(16)			

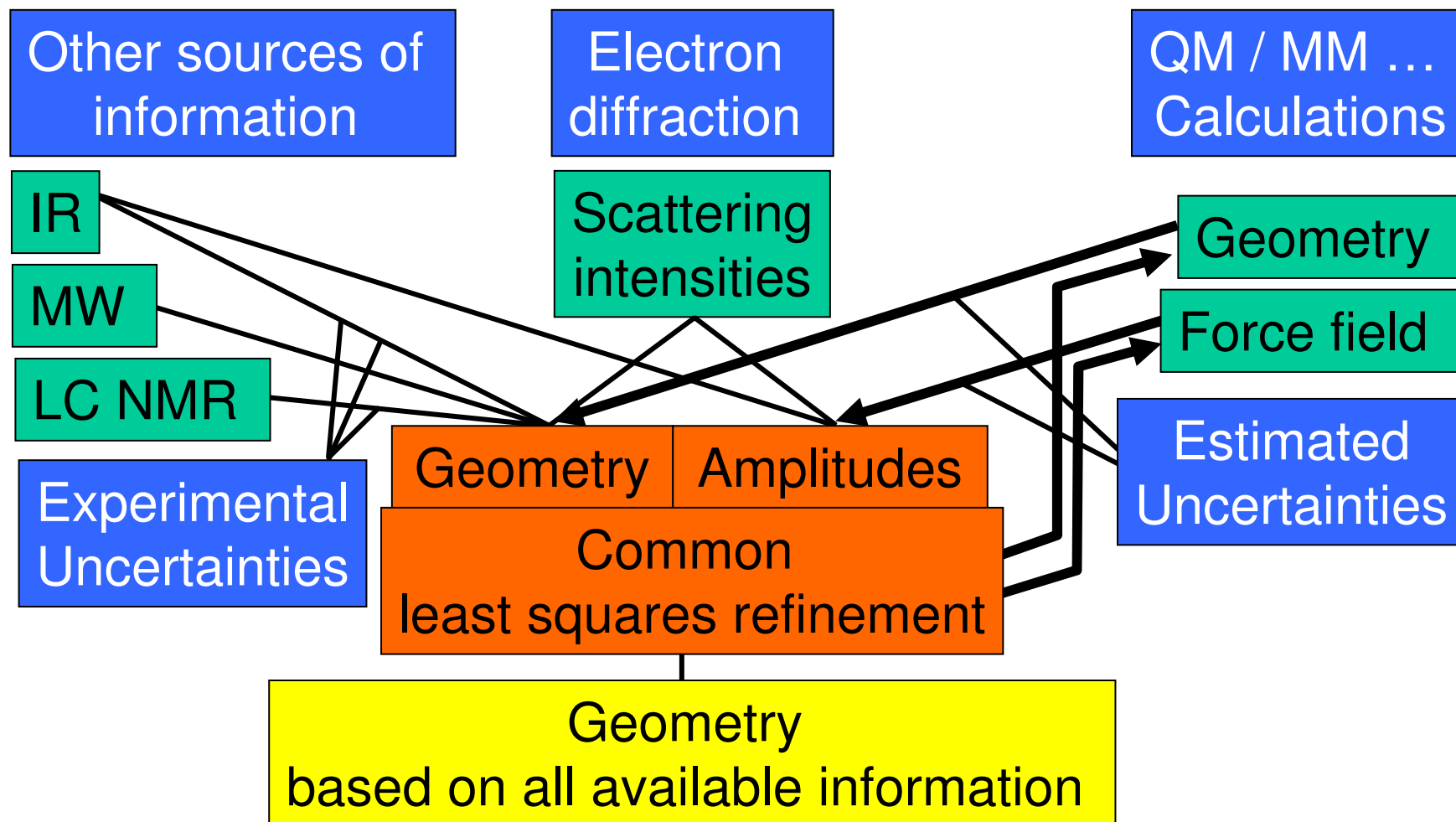
# DYNAMIC Interaction of Theory and Experiment

## DYNAMITE

Example:  $\text{O}=\text{P}[\text{C}(\text{CH}_3)_3]_3$



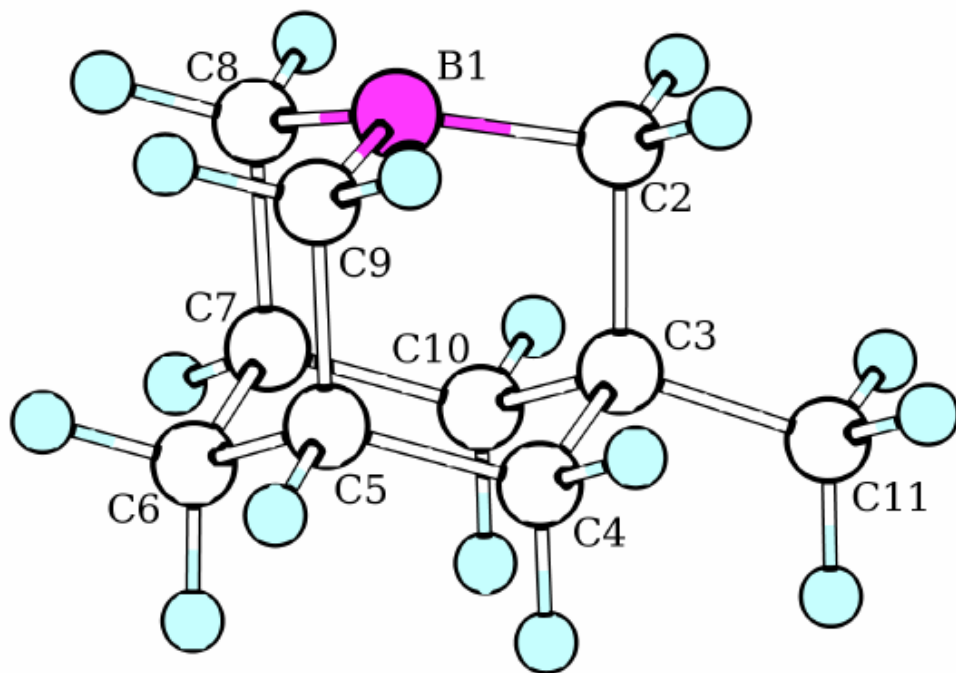
# DYNAMITE



## Model independence: The regularisation approach

- refinement of Cartesian coordinate model
- all available geometrical QM information used
- optimisation of weight  $\alpha$  for minimum QM contribution and yet full stable refinement

$$Q = \sum [sM(s)^{exp} - sM(s)^{model}]^2 + \alpha \sum w_i (p_i^0 - p_i)^2 \rightarrow \min$$



Details:  
see  
Yuri Vishnevskiy's  
poster



## Conclusion

- ▶ The combined analysis of various sources of experimental information leads to a much more accurate structural information than information based solely on one method
- ▶ Avoiding constraints gives more realistic description of errors
- ▶ The inclusion of calculated values helps to resolve problems in cases where insufficient experimental information is available
- ▶ In this case the structures represent the sum of the presently available knowledge, but are no longer purely of experimental nature
- ▶ Care is suggested in interpretation of such data, when theory and „experiment“ are to be compared



## Further reading:

The **SEMTEX** approach:

Like DYNAMITE, but using a feedback circle with  
ab initio calculations instead of molecular mechanics

See: G. R. Kafka, S. L. Masters, D. W. H. Rankin,  
*J. Phys. Chem. A* **2007**, *111*, 5913