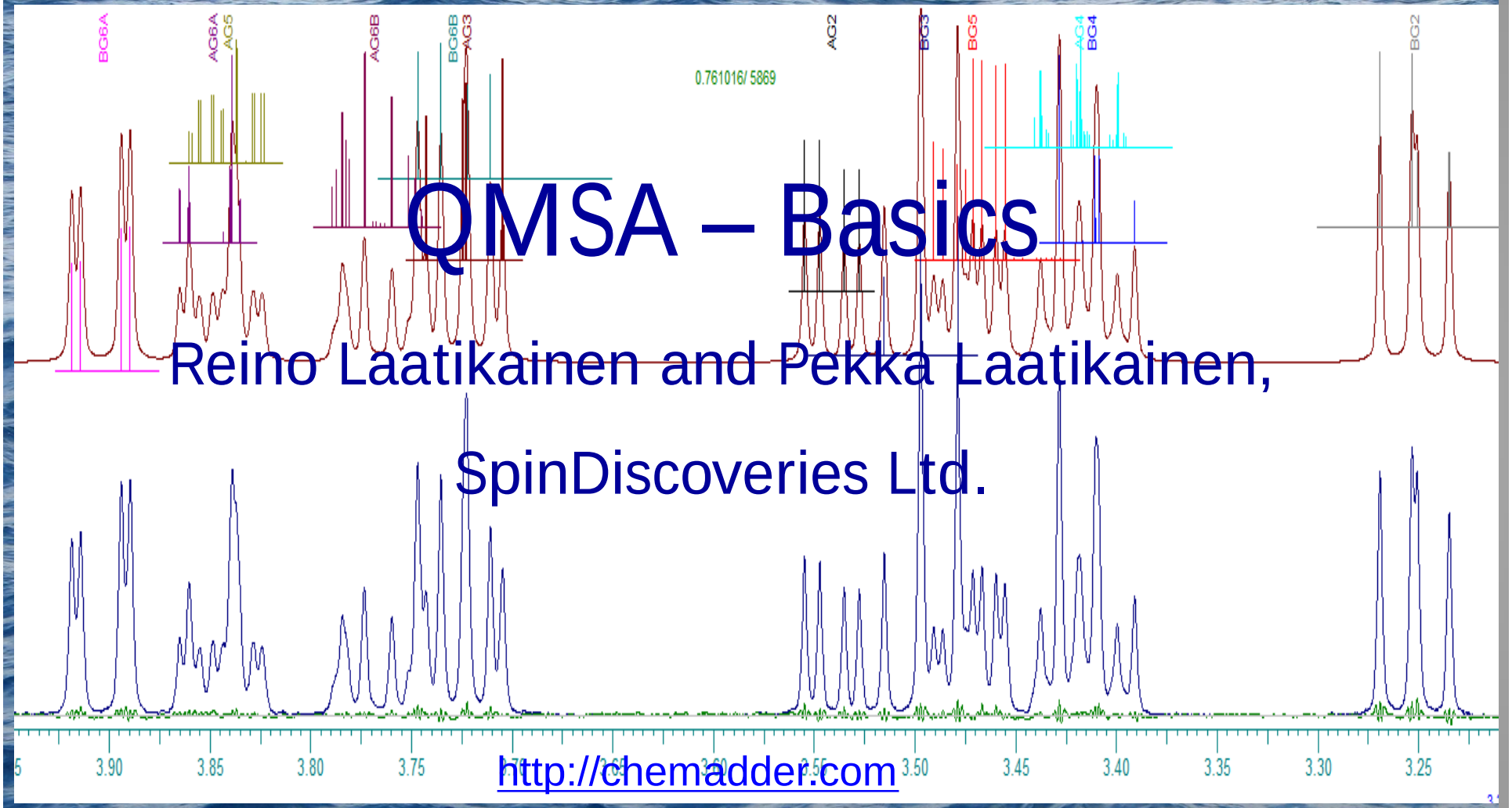


QMSA – Basics

Reino Laatikainen and Pekka Laatikainen,

SpinDiscoveries Ltd.

http://chemadder.com

Why Computational QMSA?

STRUCTURE ANALYSIS & VERIFICATION !

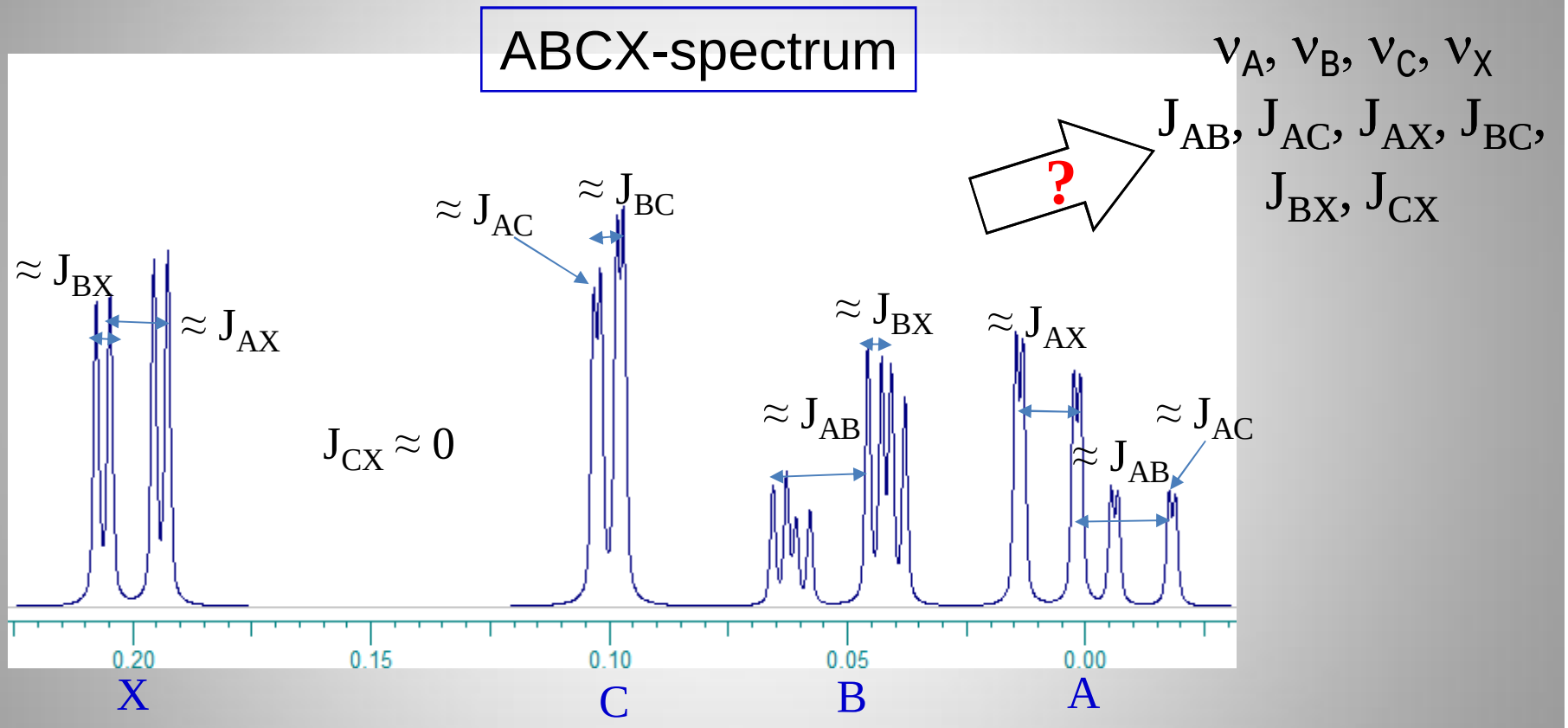
- Very accurate spectral parameters – diagnostic couplings
- To understand second-order spectra
- 3D and 4D structure

QUANTITATIVE ANALYSIS

- Adaptive Spectral Libraries; storage of spectra
- Profiling of mixtures
- Metabolomics
- Impurity analysis
- Physical chemistry applications: kinetics, thermodynamics, ...

See for example, Laatikainen & al, in Encyclopedia of Magnetic Resonance in 2011, by John Wiley & Sons, Ltd.
DOI:10.1002/9780470034590.emrstm1226

NMR Spectral Analysis



Chemical shift = weight point of multiplet

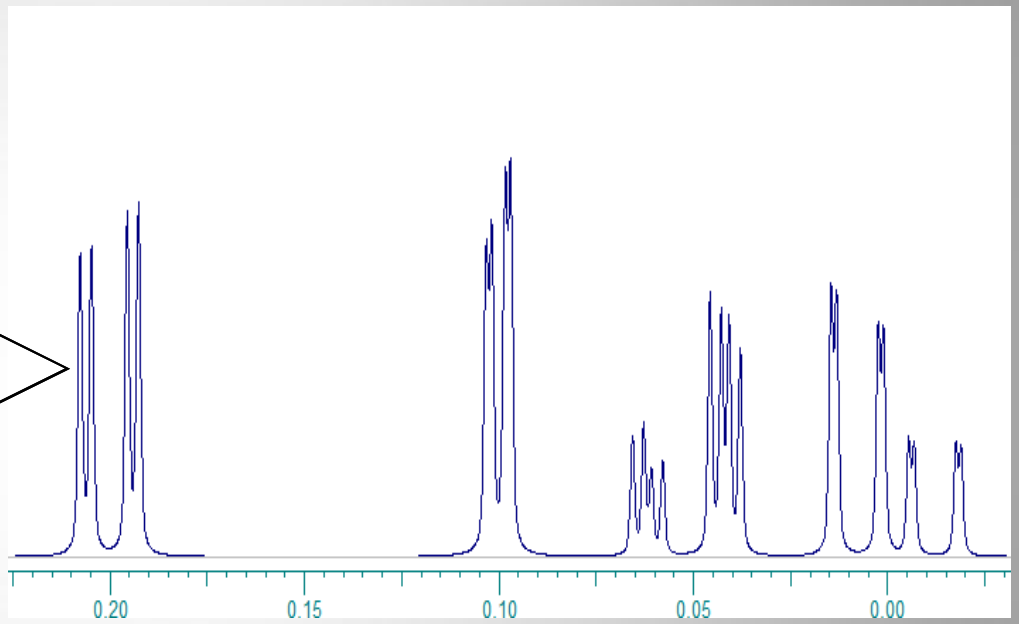
Coupling constant ≈ difference of two lines

The differences depend on field, couplings do not !

If chemical shifts, coupling constants & line-shape are known, spectrum can be simulated using QM (Quantum Mechanical) model !

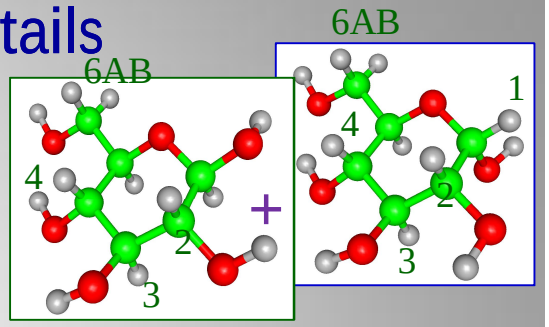
$\nu_A, \nu_B, \nu_C, \nu_X$
$J_{AB}, J_{AC}, J_{AX}, J_{BC},$
J_{BX}, J_{CX}

Simulation



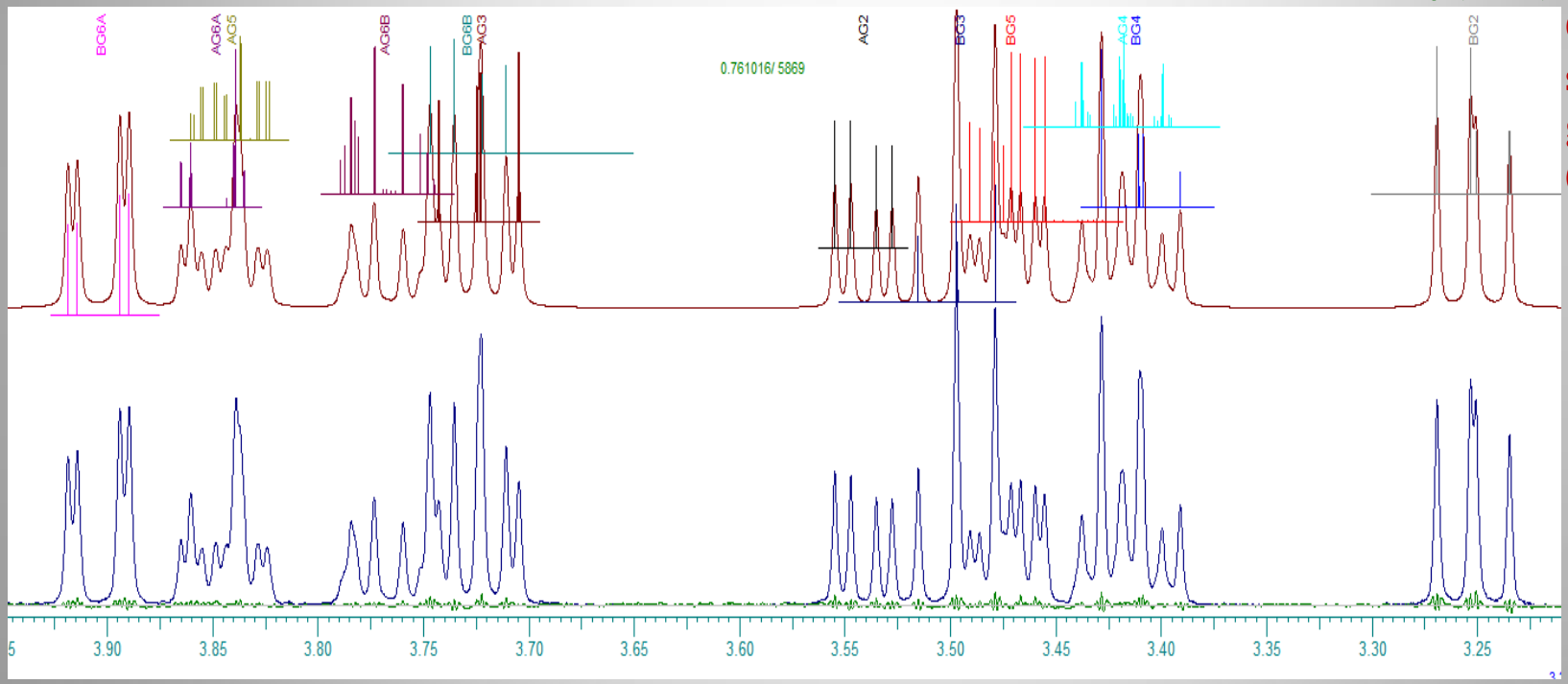
There is no explicit way to transform spectrum I(ν) to spectral parameters – but that can be done iteratively: guessing trial parameters, simulating spectrum with them and then trying to improve (iterate) the guess.

NMR spectra (relative signal positions and intensities) obey quantum mechanical rules – in very details



β(60%) α(40%)

Glucose (1-H proton signals excluded):



Calculated spectrum of glucoses (β+α)

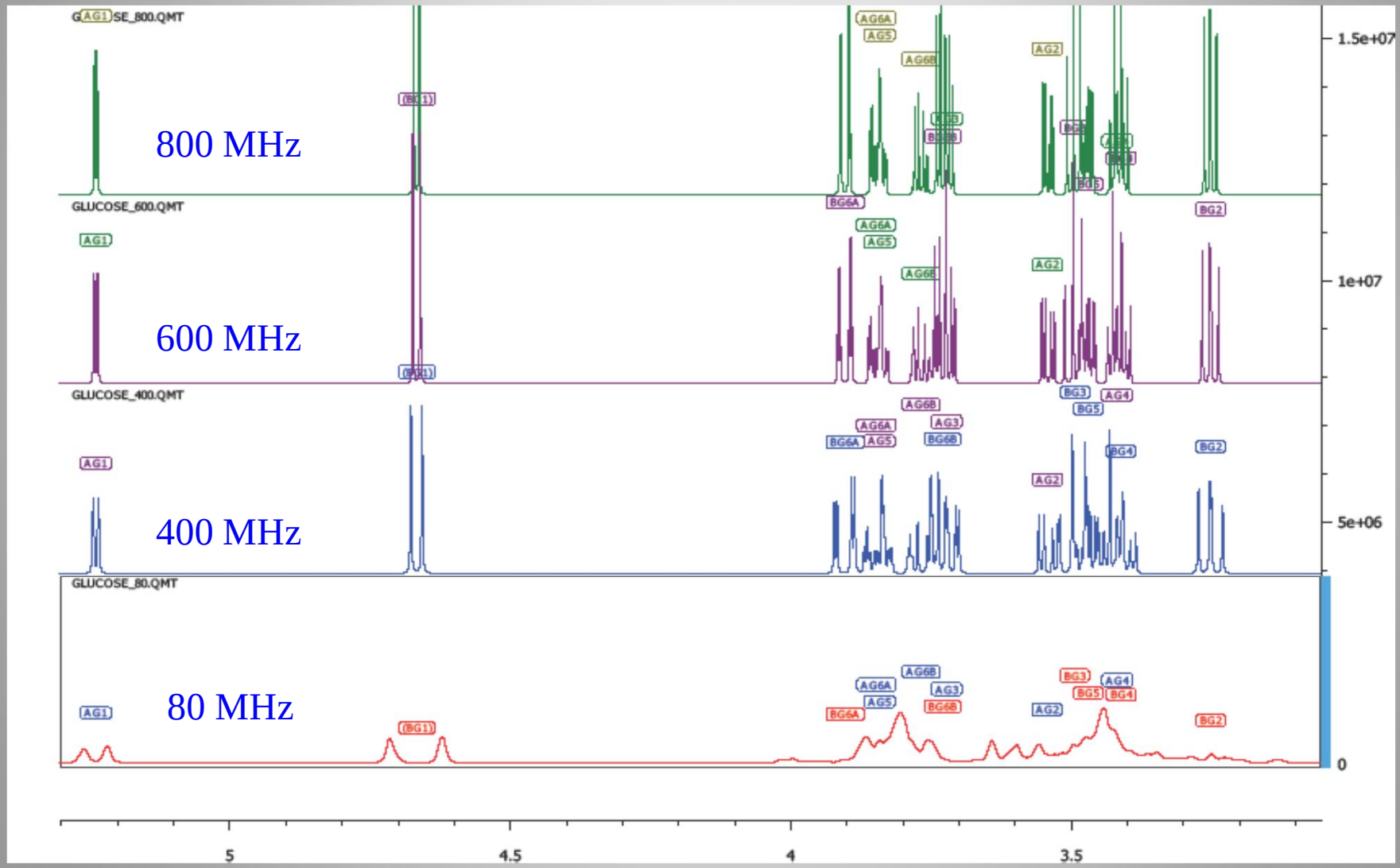
Observed spectrum & difference spectrum

768 theoretical transitions !

How to get the best spectral parameters from the observed spectrum !??

Adaptive Spectrum: Field effects !

Glucose, simulated on the basis of the 500 MHz spectrum analysis (slice 5).



Iterative Quantum Mechanical NMR Spectral Analysis (QMSA)

NMR spectrum I(ν) is sum of *spectra* of chemical components (S) and *background (B)*:

$$I(\nu) = \Sigma x_n S_n(\nu) + B(\nu)$$

Each spectrum *S* is a function of spectral parameters $\underline{w}$ = chemical shifts, $\underline{J}$ = coupling constants and $\underline{\Delta}$ = *Line-widths*, $\underline{R}$= *Response factors and Line-Shape*:

$$S_n(\nu) = F_n(\nu, \underline{w}, \underline{J}, \underline{\Delta}, \underline{R}, \text{Line-shape})$$

Structure analysis: I(ν) => $\underline{w}$ & $\underline{J}$

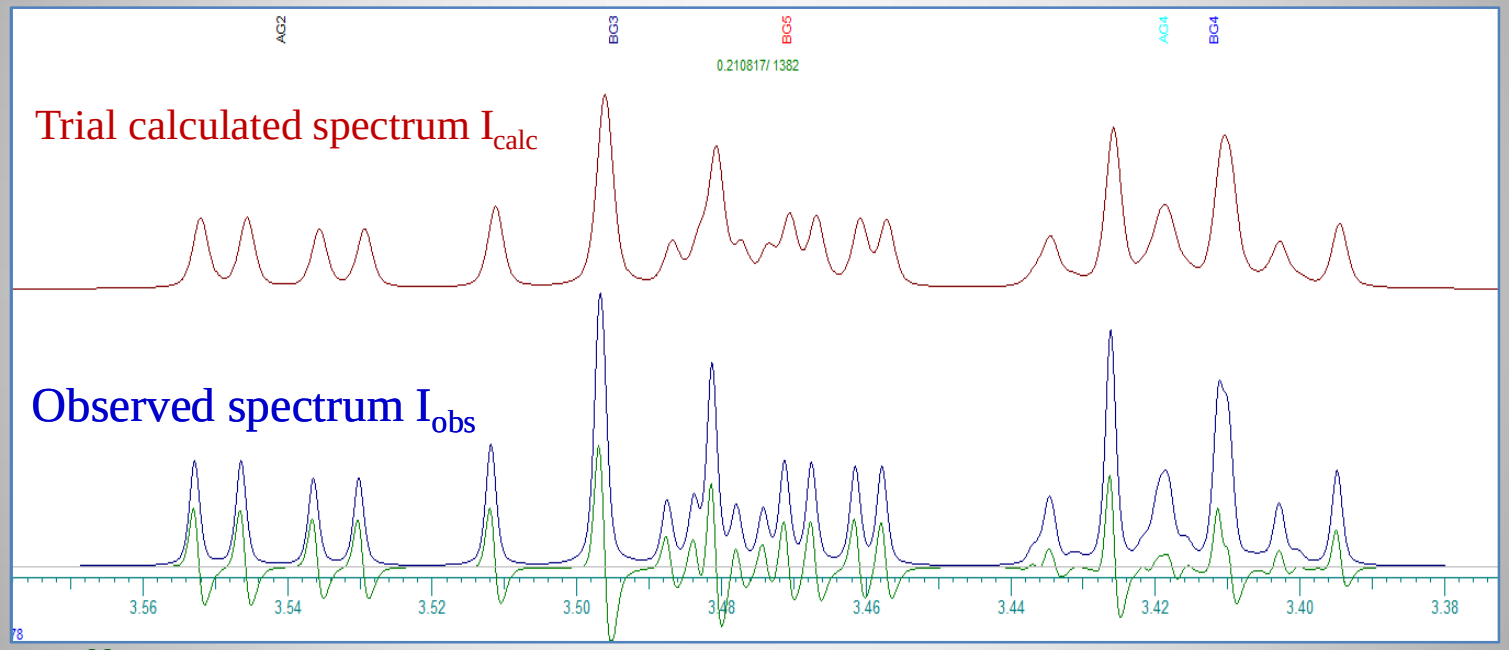
Quantitative NMR: I(ν) => x_n *(populations)*

Line-widths can be different for each spin-particle, in optimal case *response factors* are 1.0 but may vary significantly with solvent suppression, T2 edition, etc. The *line-shape* can be described by a combination of Lorenzian and Gaussian functions, and assumed to be the same for all lines.

A non-linear mathematical problem !!

The least-square iterative approach
Strongly non-linear problem: only iterative solution!

A piece of glucose 600 MHz spectrum:



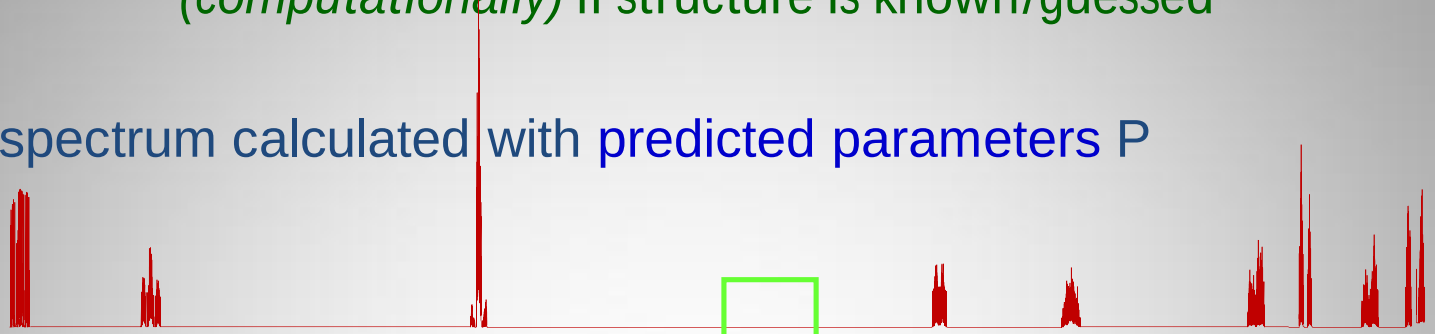
Difference spectrum $I_{obs} - I_{calc}$

Minimize Sum of Squares $SQ = \Sigma [I_{obs} - I_{calc}]^2$

Starting from a predicted spectrum

Spectra parameters can be estimated from the spectrum or *predicted (computationally)* if structure is known/guessed

Trial spectrum calculated with predicted parameters P



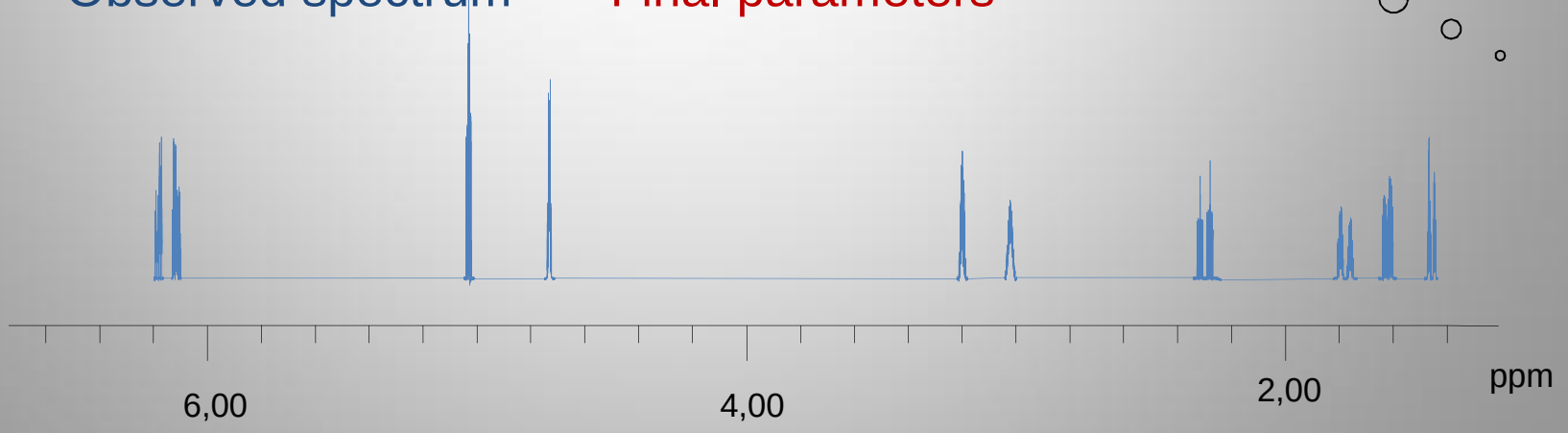
Iteration

Differentials $\partial(I_{obs}-I_{calc})/\partial P_i$ => corrections to P_i

Problem: the spectra do not overlap!?

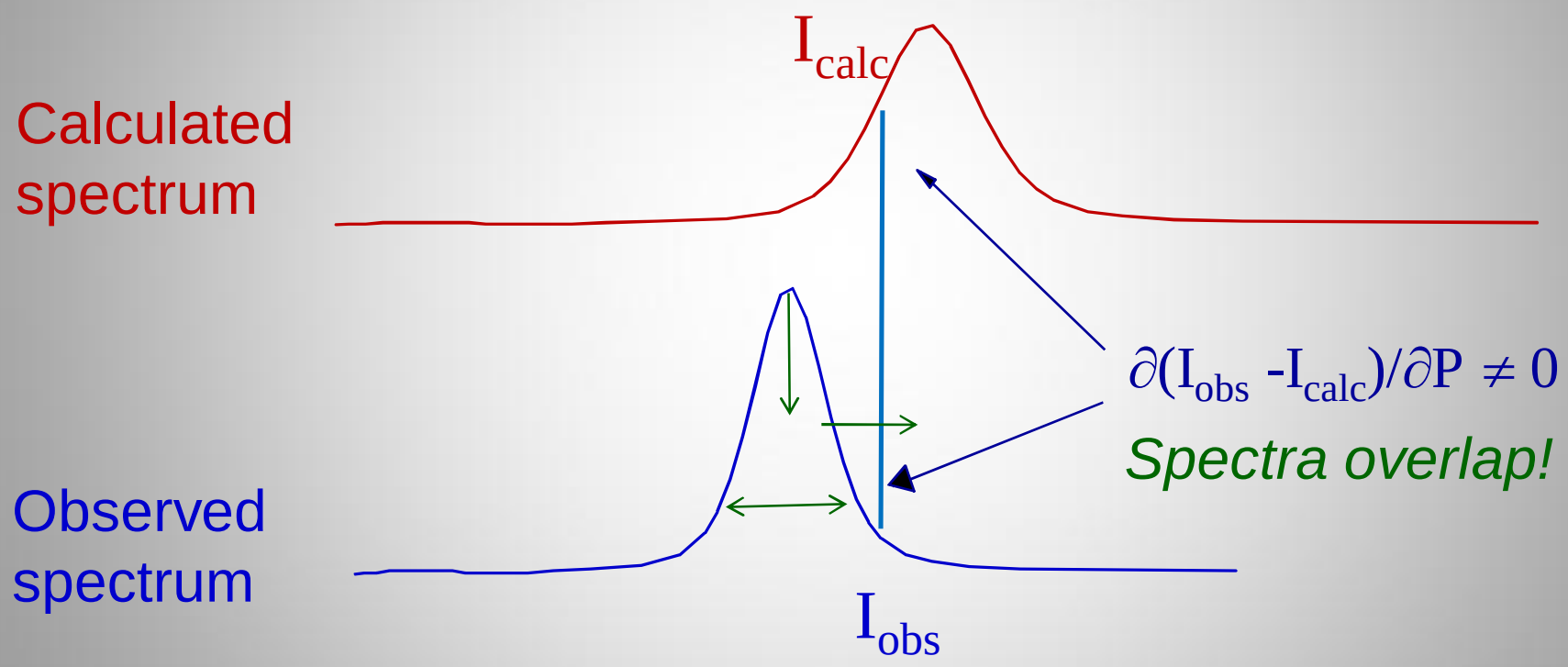
Observed spectrum

Final parameters



Total-Line-Shape Fitting

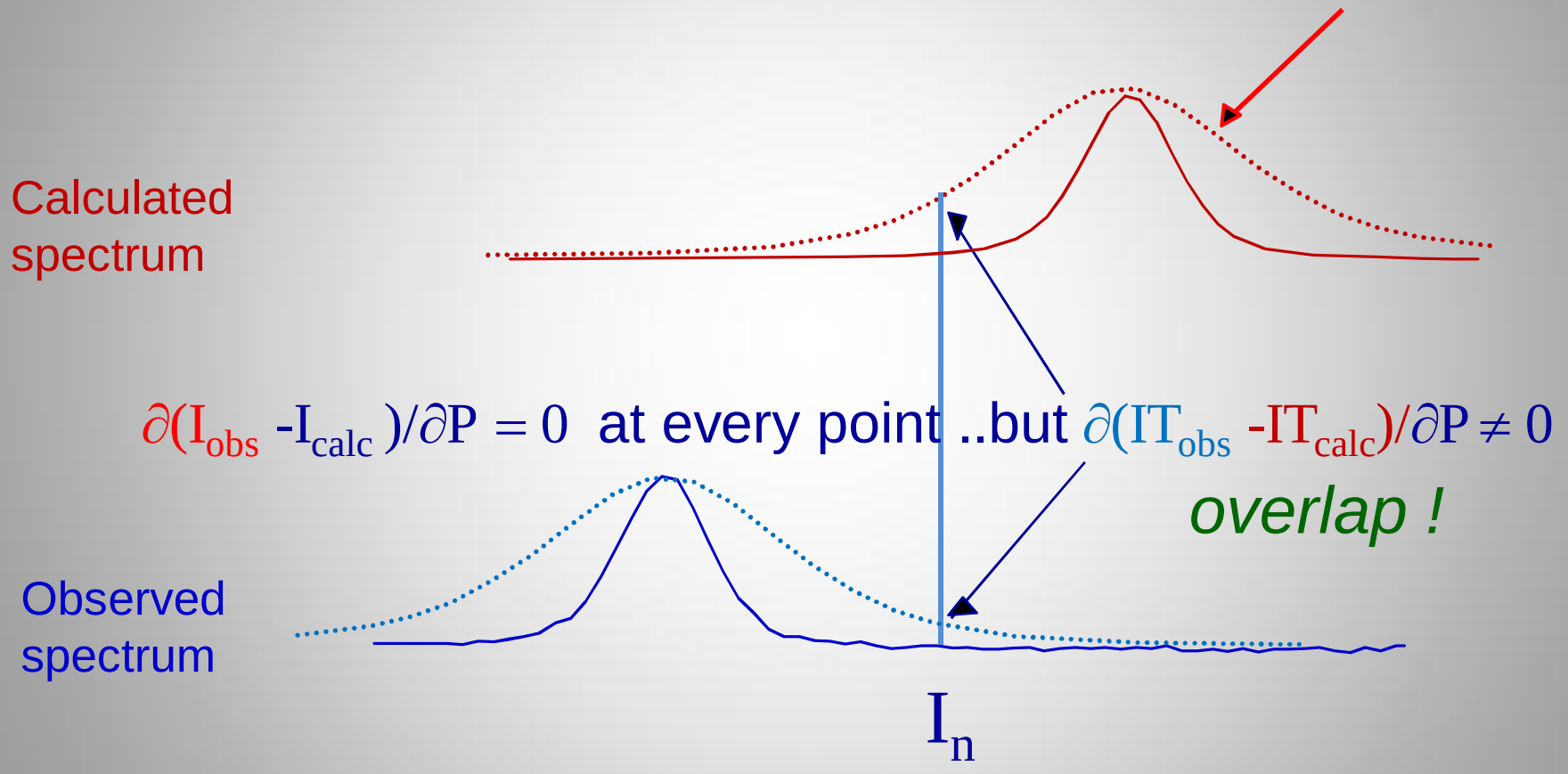
Minimize $SQ = \Sigma_n (I_{obs} - I_{calc})^2$ (n=spectral points)



Strongly non-linear problem: only iterative solution!

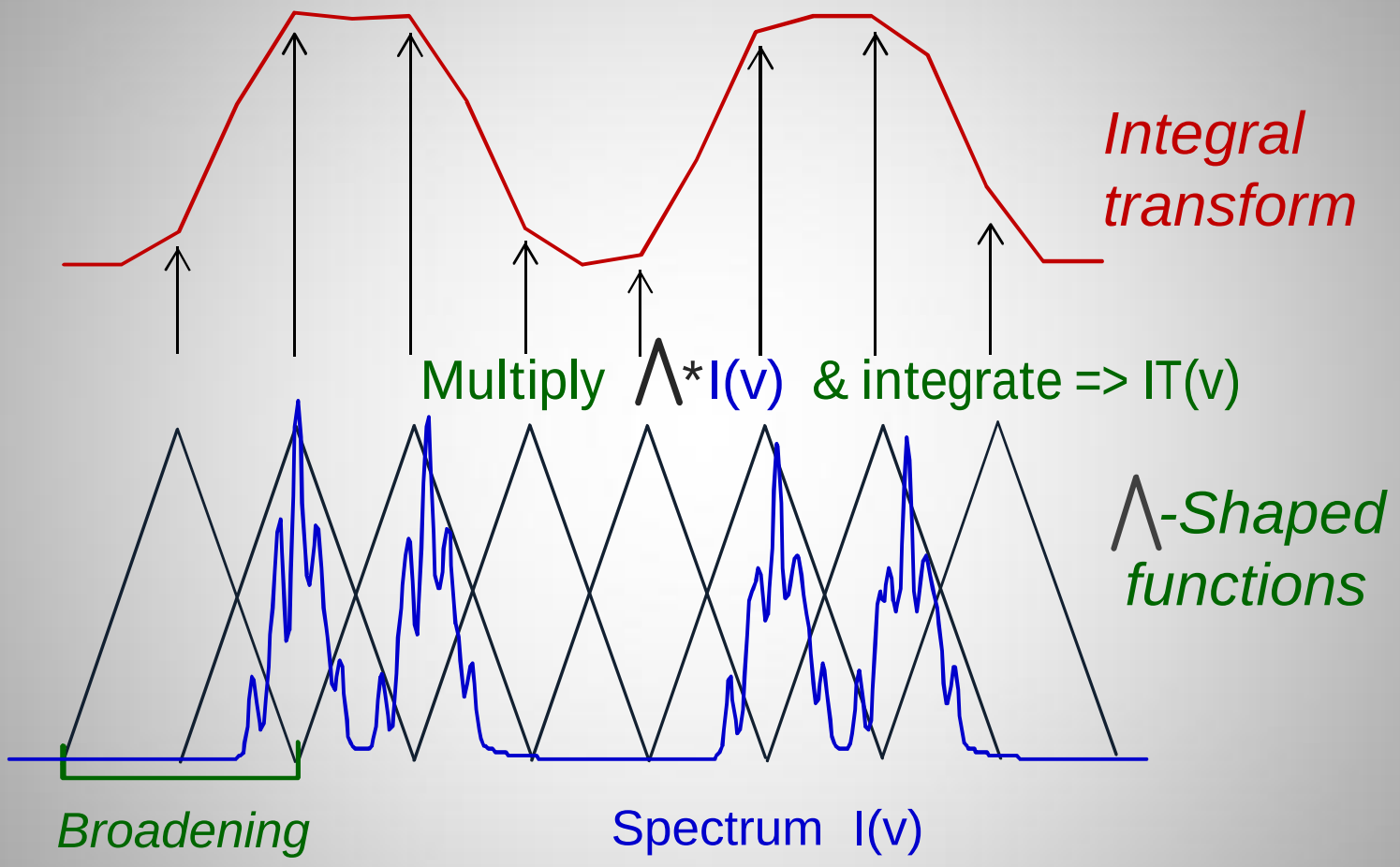
Problem: calculated and observed signals do not overlap, and derivative $\partial(I_{obs} - I_{calc})/\partial P$ is zero !

Solution: broaden lines doing *Integral Transform* (IT)



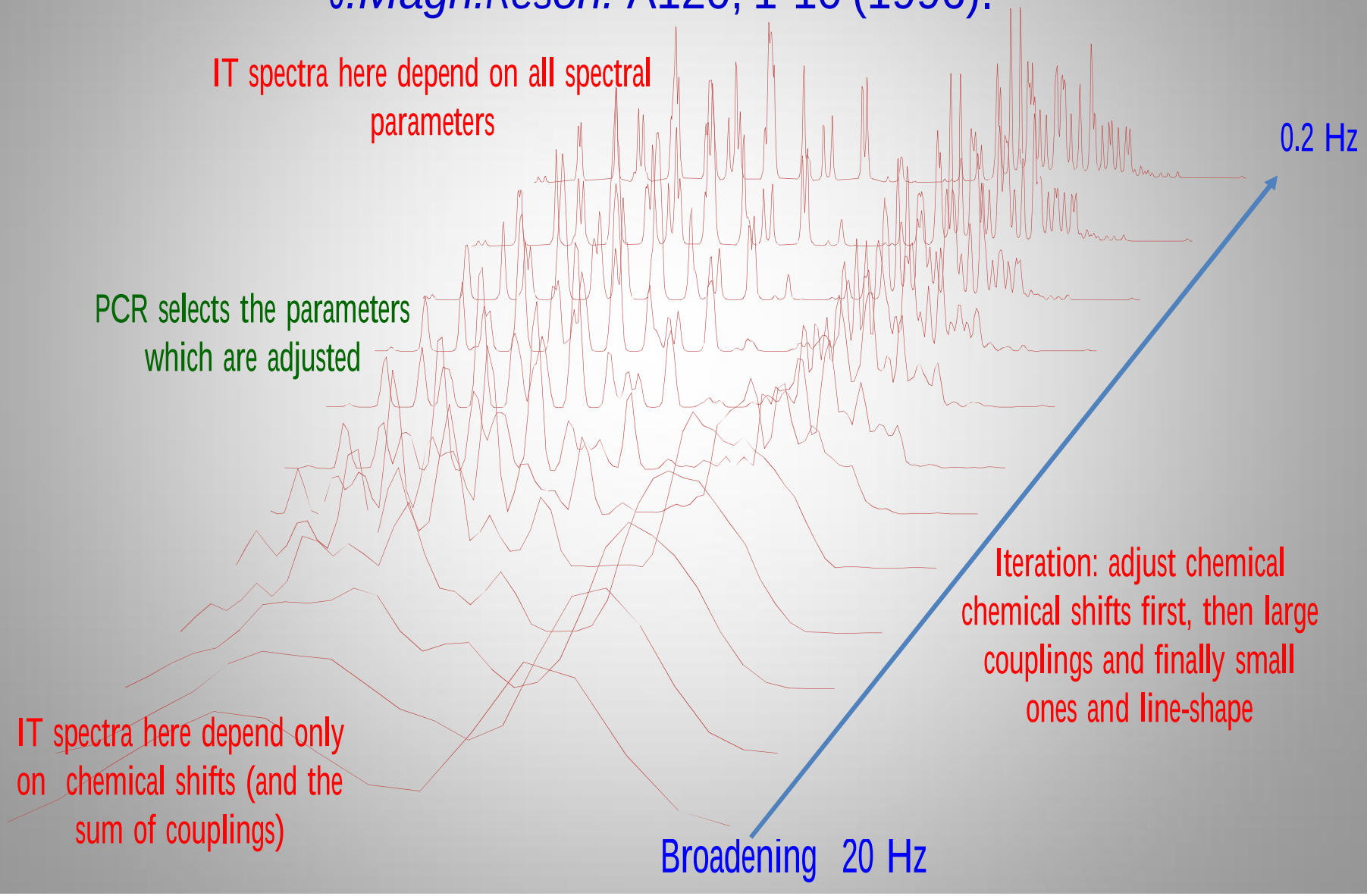
If IT's overlap, the iteration converges!

Broadening with Bartlett Integrals



Details smaller than *Broadening* are hidden

Iterative process based on *Integral Transform* (IT) spectra and *PCR* (= Principal Component Regression), see Laatikainen & al. *J.Magn.Reson.* A120, 1-10 (1996).



The RESPONSE problem

Experimental conditions (pulse sequence, solvent suppression, etc.) may lead significant deviations from theoretical intensity ratios of proton signals !!

Response factor R = Observed intensity/ Expected intensity (Reference)

R=1.0 means that the responses (signal area) of protons are equal, R= 0.80 means that the signal of a proton is only 80% of that of reference proton (the proton having the strongest signal).

R's can be optimized in SpinAdder, usually forcing them toward 1.0 (or to any default value) to avoid problems arising from overlap of signals

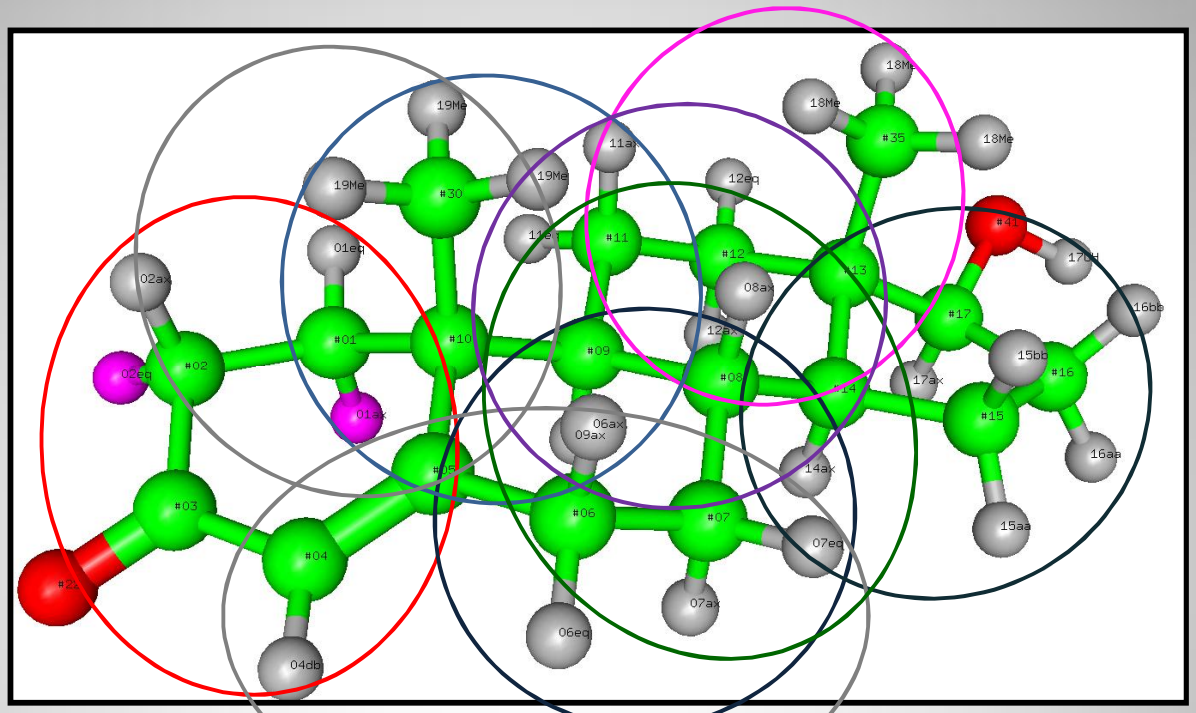
Response factors of a- and b-glucoses determined measured with different settings.

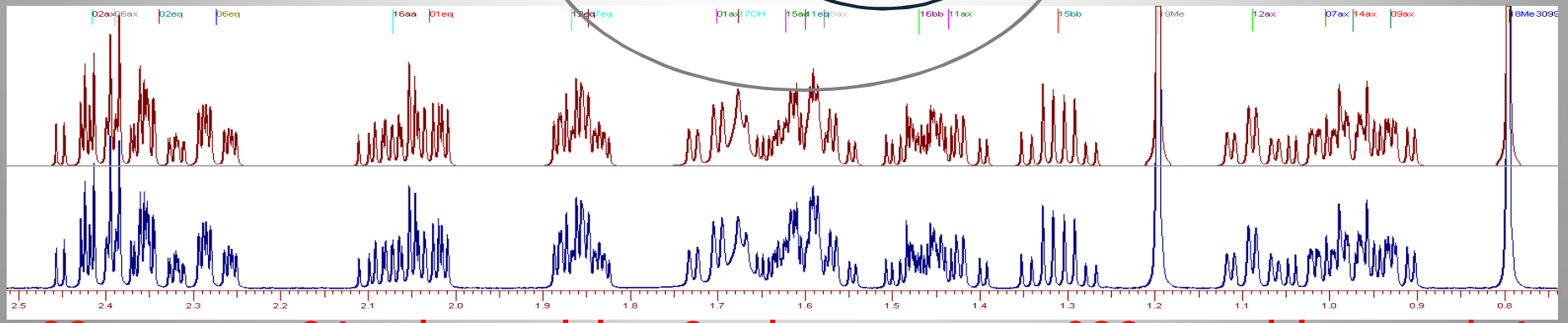
Proton	qH[a]	H[b]	qpresat[c]	presat[d]	qpresat[c]	presat[d]
	D_2O	D_2O	D_2O	D_2O	H_2O+D_2O	H_2O+D_2O
a-H1	0.962	0.875	0.960	0.880	0.950	0.924
a-H2	0.974	0.993	0.965	0.993	0.904	0.909
a-H3	1.000	0.910	1.000	0.920	0.969	1.000
a-H4	0.978	0.953	0.990	0.990	1.000	0.978
a-H5	0.965	0.997	0.975	1.000	0.850	0.885
a-H6A	0.977	0.997	0.953	0.994	0.884	0.868
a-H6B	0.975	1.000	0.955	0.981	**0.811**	0.840
b-H1	-	-	-	-	-	-
b-H2	0.988	0.869	0.949	0.840	1.000	0.993
b-H3	0.996	0.955	0.978	0.945	0.986	1.000
b-H4	0.986	0.959	0.951	0.926	0.952	0.954
b-H5	0.989	0.993	1.000	1.000	0.974	0.989
b-H6A	1.000	1.000	0.913	0.914	0.870	0.881
b-H6B	0.982	0.987	0.904	0.908	**0.845**	0.863

[a] Basic 1H spectrum (zg): ds=4, ns= 8, aq=7.7s, rd=52s and 90 pulse. [b] Basic 1H spectrum (zg): ds=4, ns=32, aq = 7.7s, d1 = 2.3s and 90 pulse. [c] Noesypresat (noesygppr1d): mt=10ms, ds=4, ns=8, aq=7.7s, d1=3s, d2=49s and 90 pulse. [d] As in c, but d2=0.

Glucose RESPONSE factors depend on experiment (JMR, 2014), the worst values are obtained with water suppression in 90% H_2O, in D_2O with sufficient relaxation delay the values are > 0.96 !

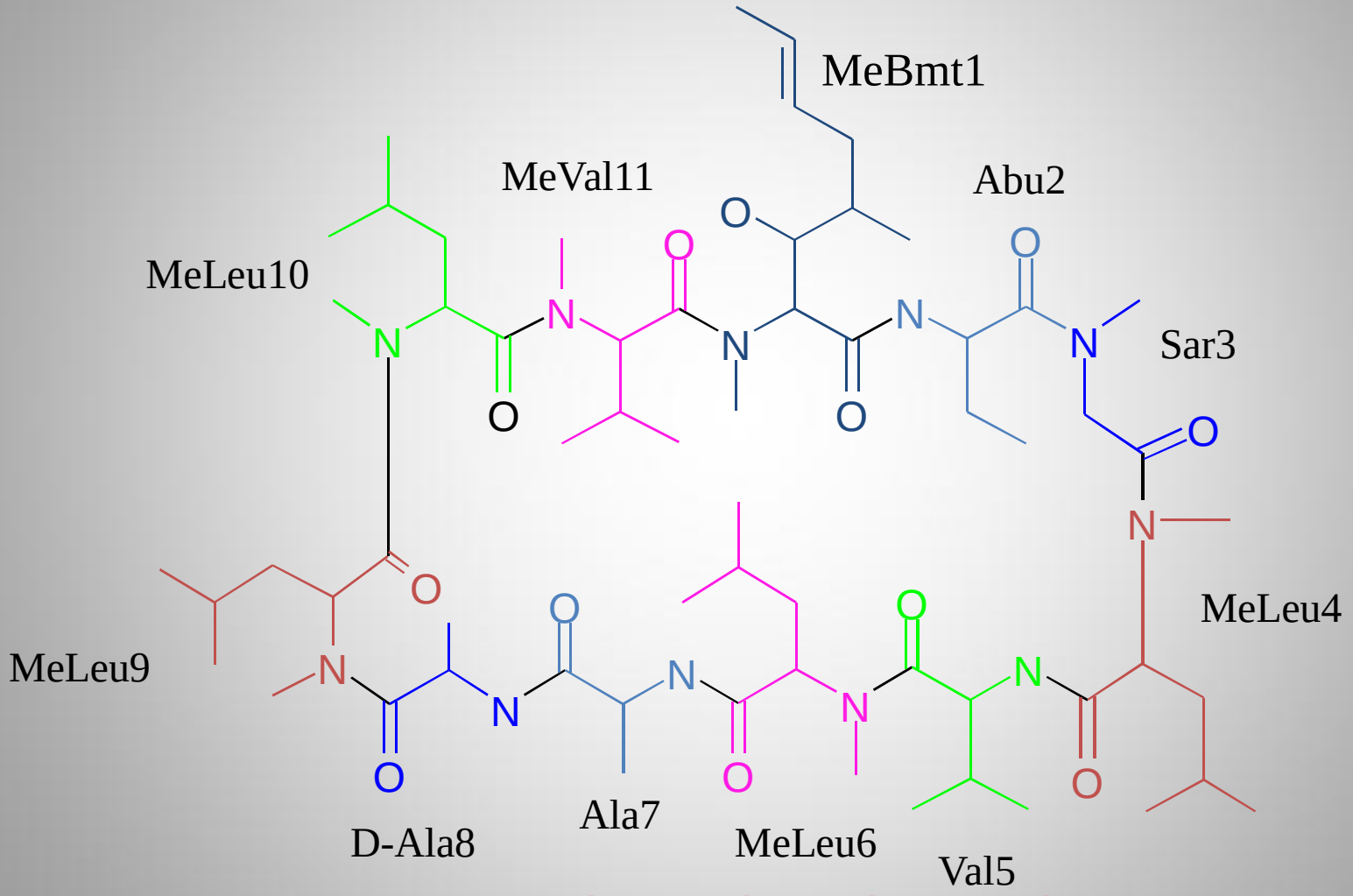
Large Spin-network: Testosterone





28 protons, 24-spin particles, 8 sub-systems => 688 transitions, only !
Simulation time ca. 2 s.

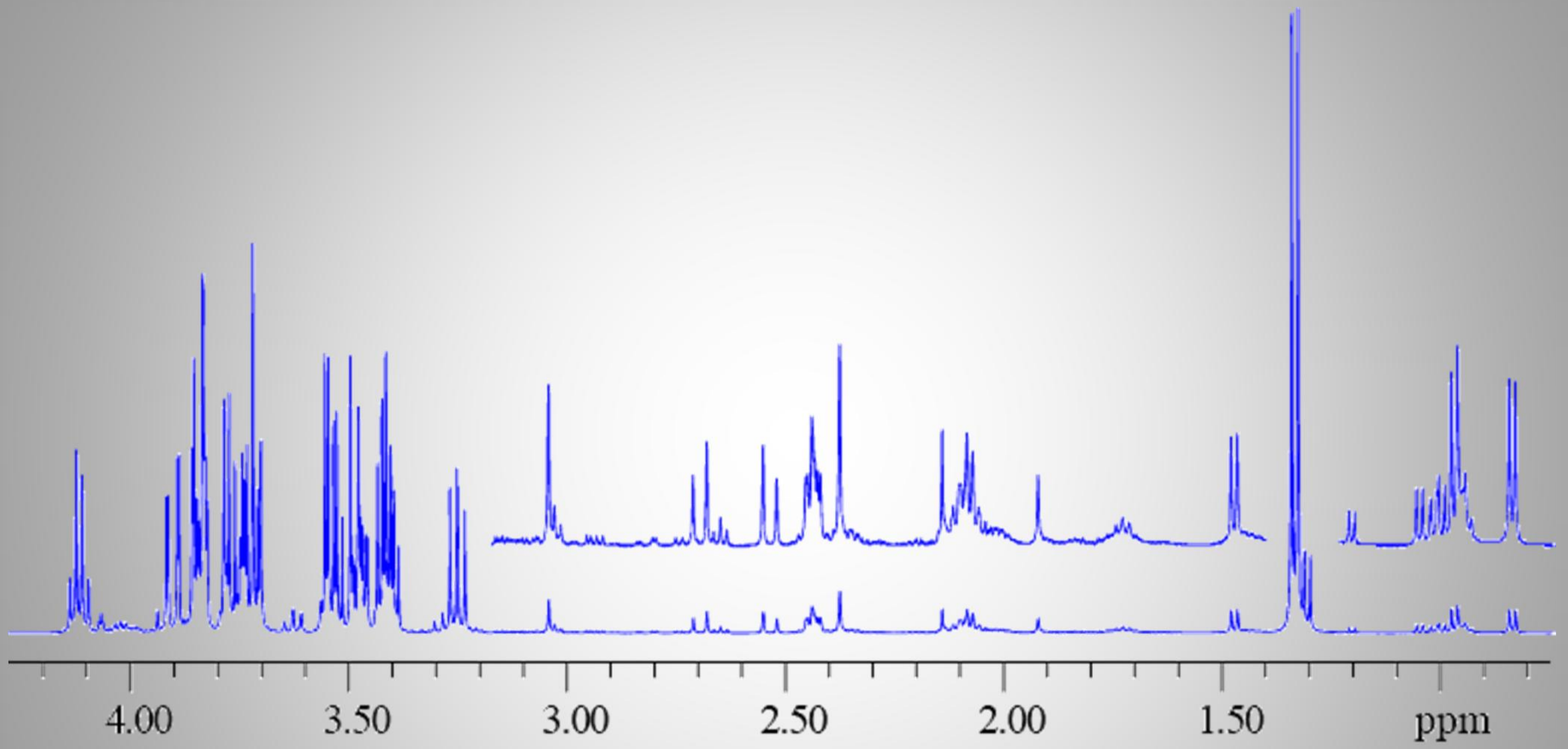
Sub-systems: Cyclosporine A



93 protons, 57 spin-particles, in 11 sub-systems.

Simulation time < 2 s.

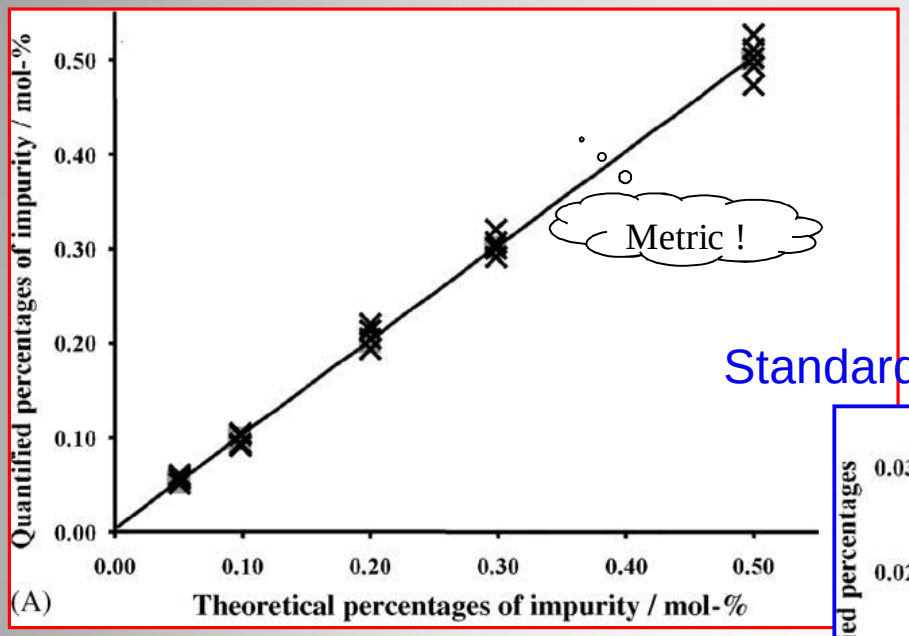
qNMR: the problem



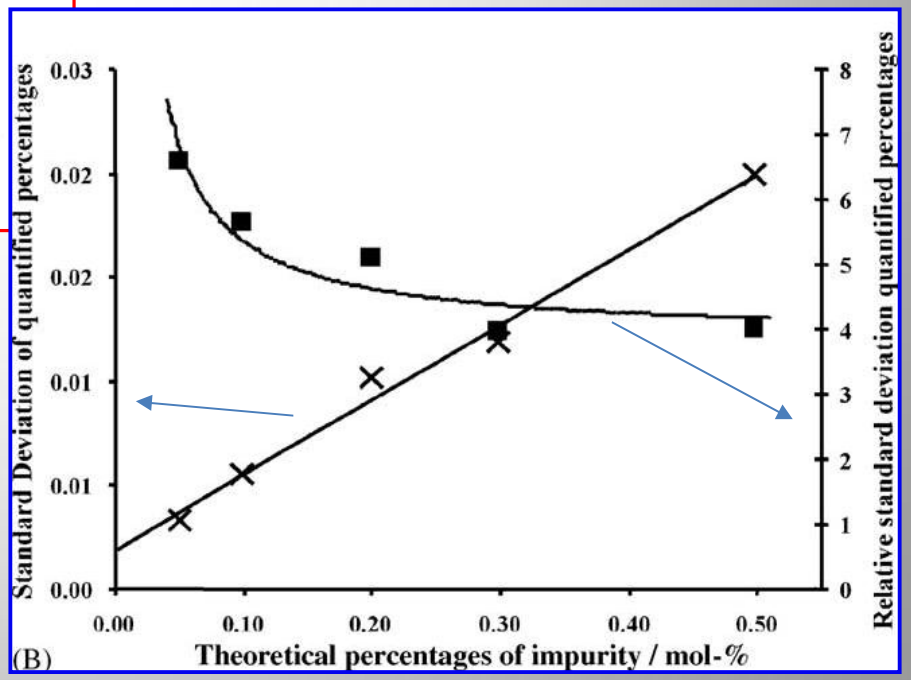
=> CONCENTRATIONS ?

Linearity & confidence limits

No calibration necessary, suits for impurity analysis, with << 0.1 mol% impurities !



Standard deviation vs. mol% %



Calculated vs. real impurity concentrations (in mol%)
$R^2 = 0.995$

See *Anal.Chim.Acta* (2005) 542, 178-185.

Quantitative Quantum Mechanical Spectral Analysis (qQMSA)

NMR spectrum I(ν) is sum of *spectra (S)* of the chemical components and *background (B)*:

$$I(\nu) = \Sigma x_n S_n(\nu) + B(\nu)$$

where each spectrum is a function of spectral parameters

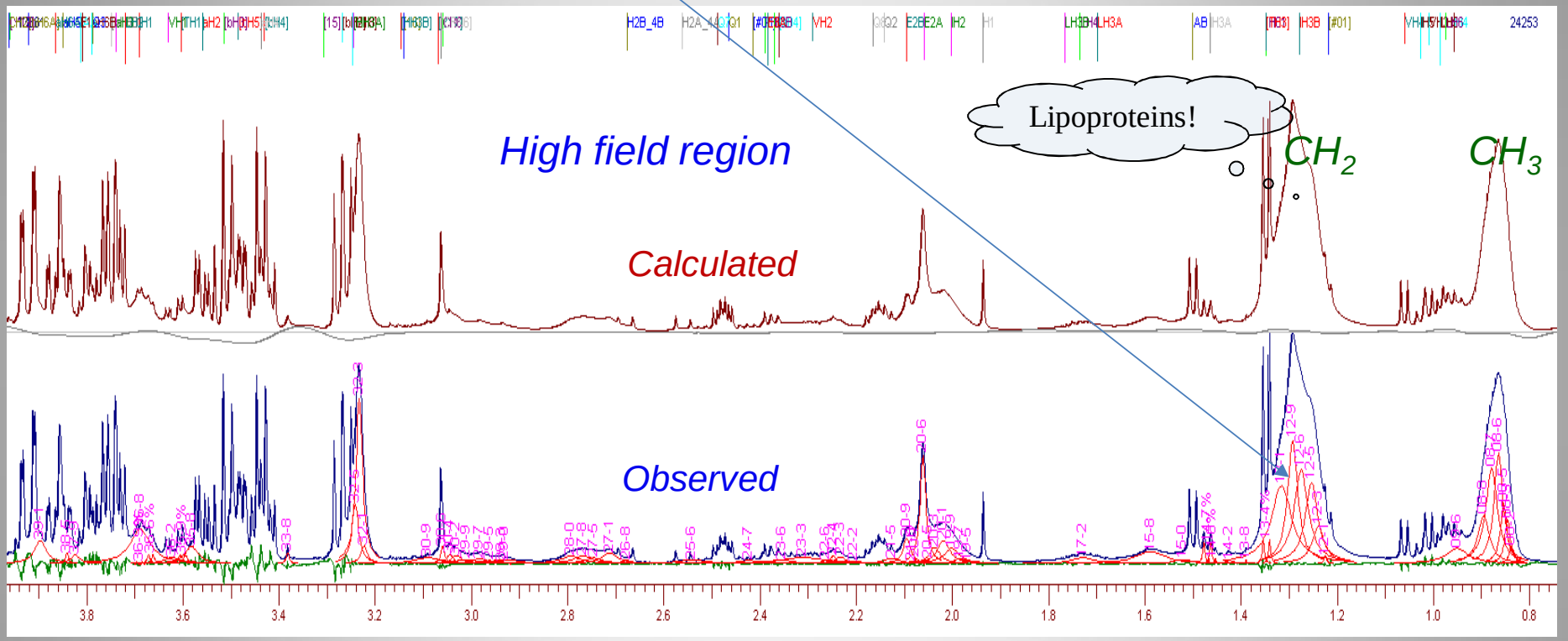
$$S_n(\nu) = F_n(\nu, \underline{w}, \underline{J}, \underline{\Delta}, \underline{R}, \textit{Line-shape})$$

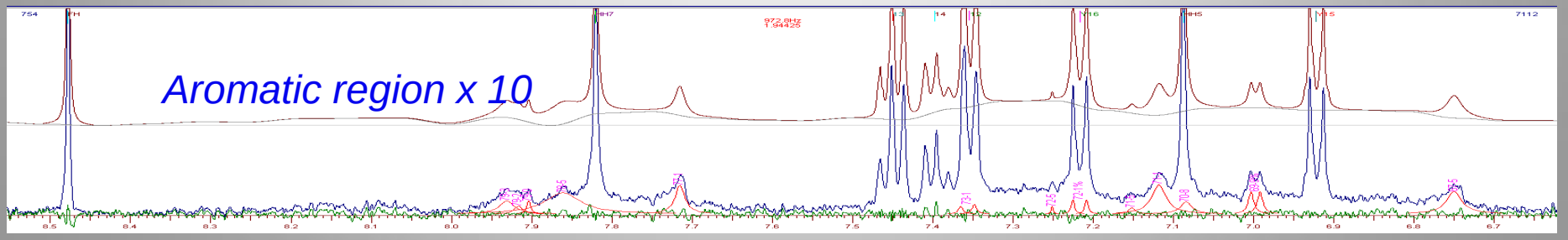
qQMSA: $I(\nu) \Rightarrow x_n$

Strict quantum mechanical rules between the positions and intensities of the observed lines: <u>maximum *prior knowledge* !</u>

qQMSA of T2 edited spectrum of serum

Some signals are described by 'EXTRA' lines: even the smallest details have an interpretation





qQMSA - Limitations

- > 100 metabolites?
- Dynamic range of 0.1-100 %
- Applications:
  - Any mixtures and impurity analysis
  - Plasma, CSF, lipid extracts of serum, urine, ...
  - Bioextracts, juices, ...

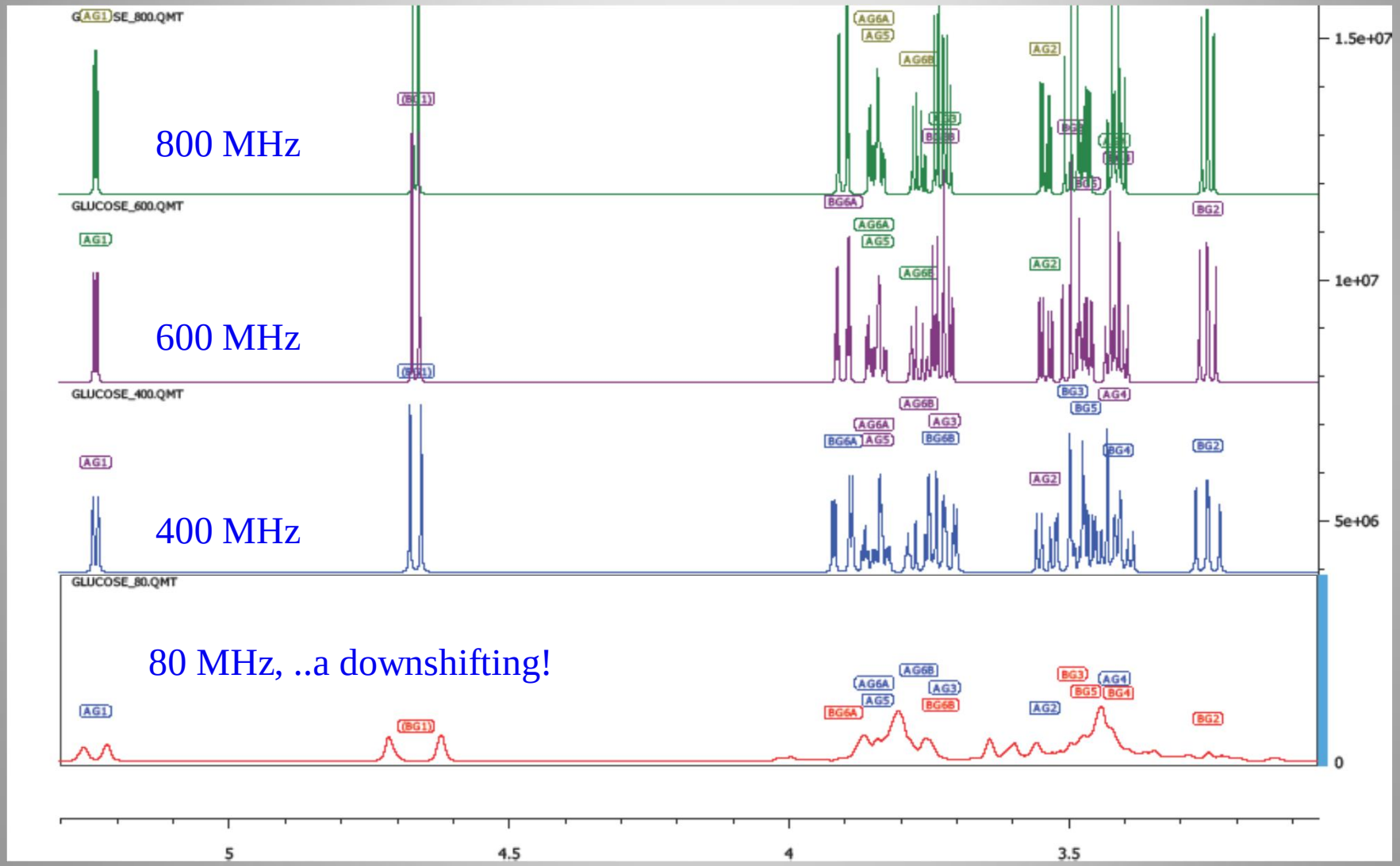
Adaptive Spectral Libraries (ASL)

- Spectral parameters obtained from QMSA
- Minimum storage space & fast
- Spectra can be simulated
  - at any field
  - with any line-width and line-shape
  - at any parameterized condition
- Free of artefacts, impurities etc.
- qQMSA with IT's
- Targeted ASL's

Perfect spectra from poor data !!

Adaptive Spectrum:

Glucose, simulated on the basis of the 500 MHz spectrum analysis (slice 5).
Line-width was 1.0 Hz in all the spectra.



ChemAdder, jess!

http://chemadder.com