

Getting the numbers right

How JASON helps you to quantify NMR measurements

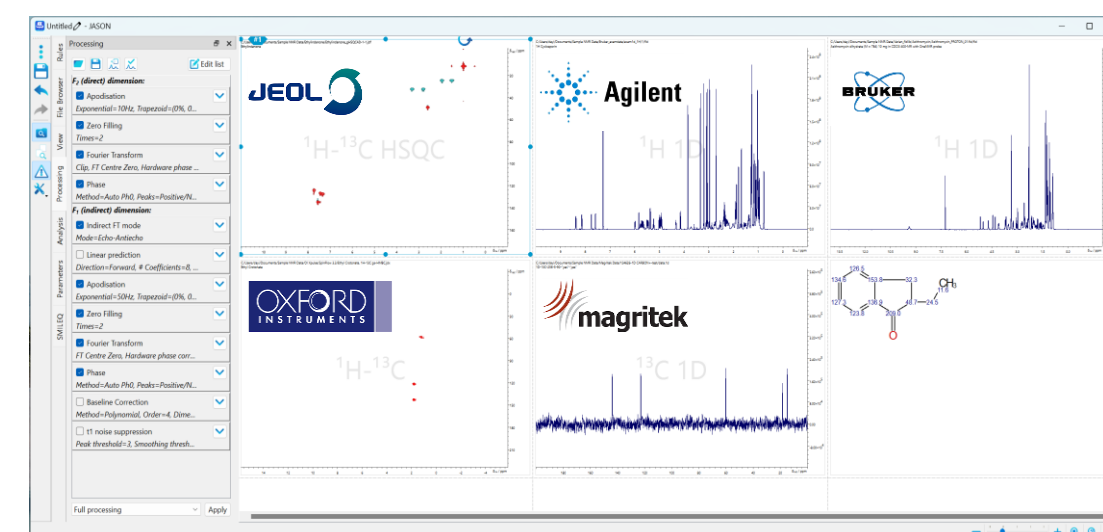
Vadim Zorin, Peter Kiraly

What is JASON



JASON is designed to streamline NMR data processing, analysis and reporting. It provides efficient tools for specialists and casual users alike, enabling them to navigate complex NMR data with ease. Additionally, JASON boasts features such as chemical drawing capabilities, accurate NMR chemical shift predictions, and seamless export options.

ChatGPT



JASON is vendor-agnostic



Why qNMR?

Pros:

- **Direct and Linear Response:** the area of an NMR signal is directly proportional to the concentration of the compound.
- **High Precision and Low Uncertainty:** When properly validated, qNMR offers high precision and accuracy.
- Unlike some other techniques, qNMR **does not require identical reference materials.**
- **Applicable to Various Nuclides.**
- **Structural Information Content:** NMR spectra provide rich qualitative information about the structure of compounds, making qNMR valuable for both quantitative and qualitative analysis.

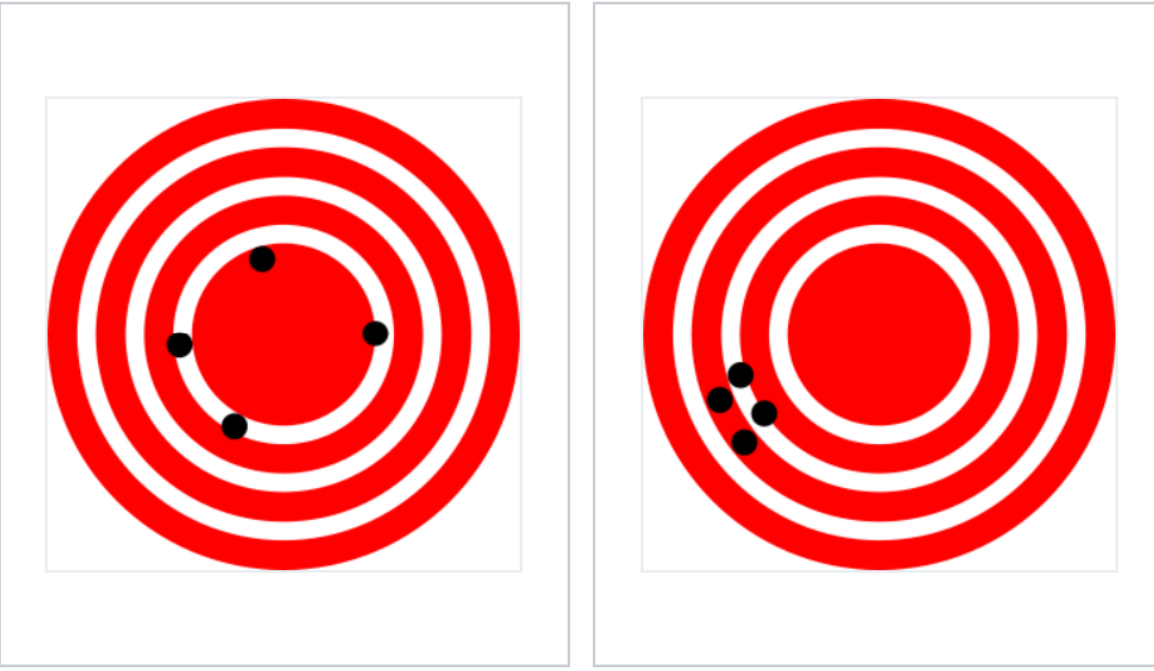
Cons:

- **Sensitivity and Detection Limit:** qNMR has low sensitivity compared to other analytical methods.
- **Overlapping Resonances.**
- **Ensuring consistency and accuracy across different laboratories can be challenging.**

Understanding of quantitation errors is essential!



Noise, errors, bias...



Low precision =
Stochastic error

Low trueness =
Bias


- Stochastic errors are easy to estimate, e.g. standard deviation
- Bias can be estimated ONLY if the true value is known

A simple math of propagation of errors:

If we measure a non-linear function $f(a, b)$ of two variable a and b :

$$\sigma_f^2 \approx \left[\left| \frac{\partial f}{\partial a} \right|^2 \sigma_a^2 + \left| \frac{\partial f}{\partial b} \right|^2 \sigma_b^2 \right] + \left[2 \frac{\partial f}{\partial a} \frac{\partial f}{\partial b} \sigma_{ab} \right]$$

where σ_f is the standard deviation of the function f ; σ_a and σ_b are the standard deviation of a and b ; $\sigma_{ab} = \sigma_a \sigma_b \rho_{ab}$ is the covariance between a and b .



- More complex manipulations/measurements we do to get the results – more error we introduce.
- But it's only strict if manipulations/measurements are totally independent.



What can we quantify in NMR?

Peak parameters (via peak modelling):

- Position
- Height
- Width
- Shape parameter(s)
- Area (could be derived from the parameters above)

Direct (sum) integrals

Concentration, purity, etc. (derived from the parameters above)

All these quantities have intrinsic measurement errors!



Peak detection and deconvolution in JASON

Two steps:

1. Hybrid peak detection:

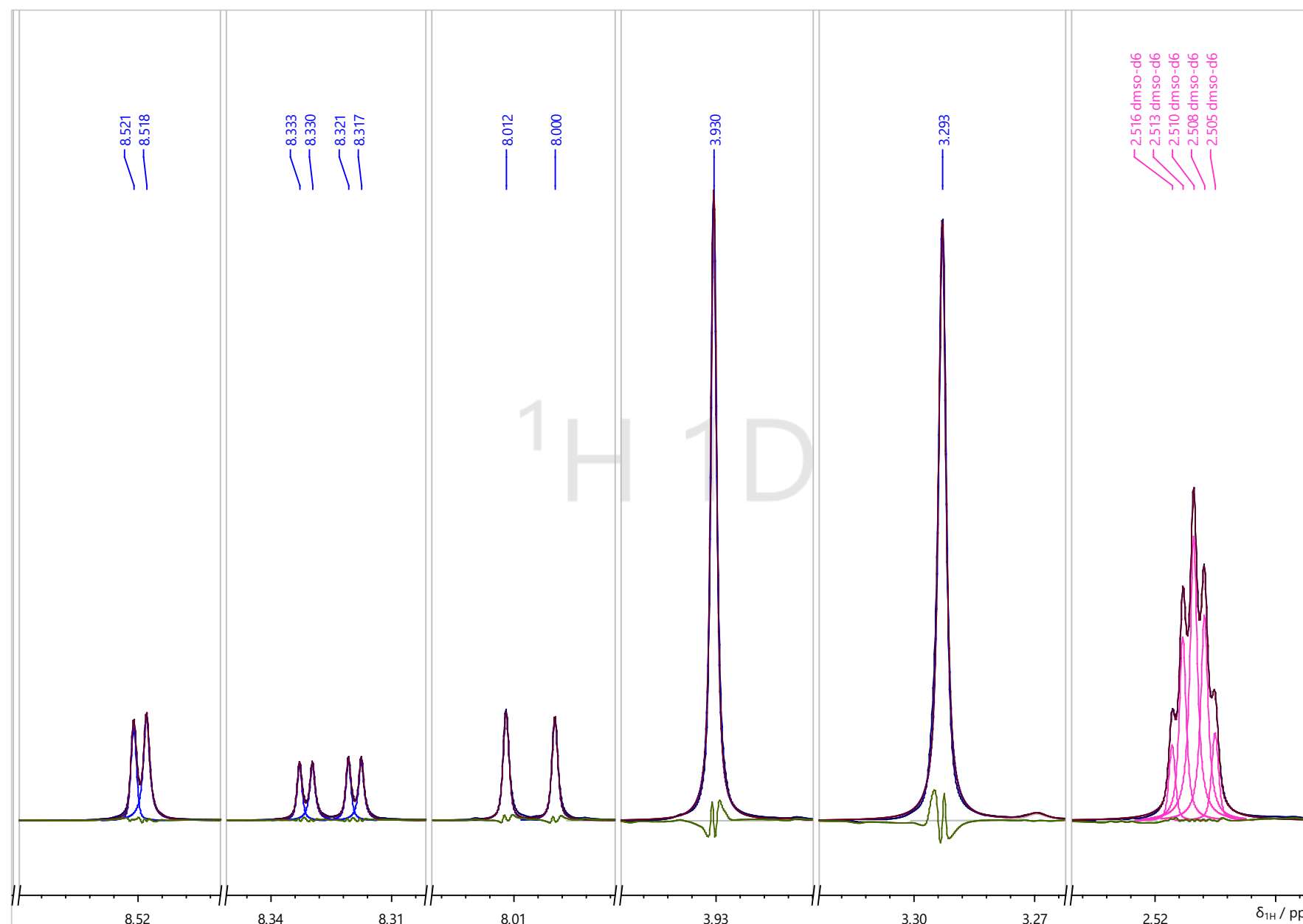
- Using derivative-based method and Savitzky–Golay filter. Good for resolving overlapping peaks
- Direct peak picking with multiple filter options. It ensures broad and weak peaks are detected

2. Peak model fitting:

- Non-linear fit for generalised Lorentzian or pseudo-Voigt lineshape models
- Robust to baseline distortions.
- A fast, peak-by-peak iterative method

3. Used in further analysis:

- Automatic solvent signals detection
- Automatic multiplet analysis
- Automatic qNMR analysis
- *etc.*



JASON detects, deconvolutes and visualises peaks, peak models, fitting sum and residuals



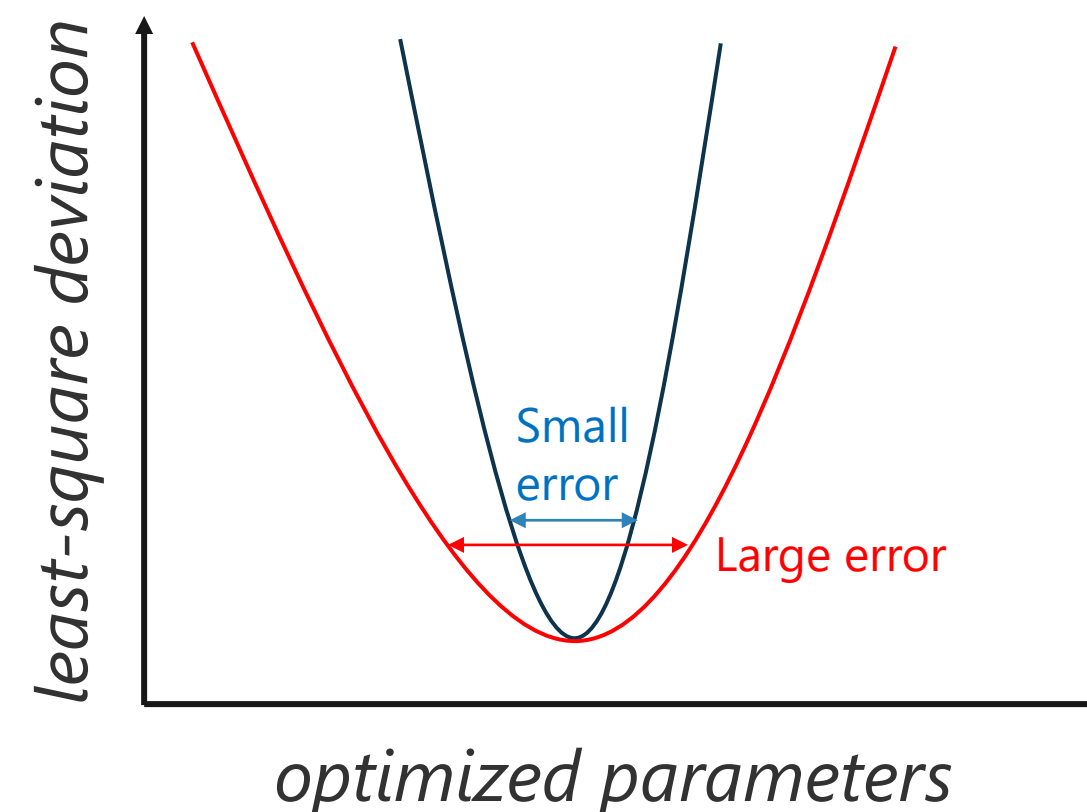
Errors of peak parameters

Peaks Table											
Pos (ppm)	σ -Pos (ppm)	Height	σ -Height	Width (Hz)	σ -Width (Hz)	Kurtosis	σ -Kurtosis	Area	σ -Area	Type	Label
8.521	4.34387e-06	4.72e+07	7.5543e+05	0.95	0.0091	1.69	0.0712	5.44e+07	1.4826e+06	compound	
8.518	2.78858e-05	5.34e+07	6.2209e+05	1.08	0.0394	0.63	0.0493	8.27e+07	3.4947e+06	compound	
8.333	4.73086e-05	3.01e+07	2.3301e+05	0.88	0.0450	0.48	0.0401	3.90e+07	2.0087e+06	compound	
8.330	7.02014e-06	2.84e+07	2.8691e+05	0.88	0.0130	0.92	0.0427	3.43e+07	7.7595e+05	compound	
8.321	9.56321e-06	3.13e+07	2.8039e+05	0.88	0.0162	0.76	0.0375	3.89e+07	9.4031e+05	compound	
8.317	9.02826e-06	3.10e+07	2.9742e+05	0.87	0.0157	0.80	0.0403	3.77e+07	9.2871e+05	compound	
8.012	4.12366e-06	5.69e+07	4.2747e+05	1.03	0.0081	1.13	0.0321	7.80e+07	1.1368e+06	compound	
8.000	3.55510e-06	5.21e+07	4.3531e+05	1.00	0.0072	1.29	0.0363	6.78e+07	1.0328e+06	compound	
3.931	3.81236e-06	3.29e+08	7.3576e+05	1.02	0.0059	0.69	0.0091	4.77e+08	3.3572e+06	compound	
3.293	2.26107e-05	3.00e+08	1.2045e+06	1.33	0.0250	0.54	0.0178	5.81e+08	1.1561e+07	compound	
2.516	8.95946e-06	3.91e+07	5.9816e+05	1.13	0.0177	1.15	0.0654	5.88e+07	1.7273e+06	NMR solvent	dms0-d6
2.513	1.91331e-05	8.88e+07	4.7254e+05	1.20	0.0487	0.36	0.0334	1.59e+08	5.9869e+06	NMR solvent	dms0-d6
2.510	1.52004e-06	1.45e+08	4.5669e+05	1.29	0.0040	0.33	0.0356	3.06e+08	1.1280e+06	NMR solvent	dms0-d6
2.508	4.84824e-06	9.75e+07	5.4215e+05	1.24	0.0141	0.04	0.0246	1.89e+08	1.9304e+06	NMR solvent	dms0-d6
2.505	8.44642e-06	3.86e+07	6.3139e+05	1.08	0.0169	1.20	0.0702	5.48e+07	1.6839e+06	NMR solvent	dms0-d6

Error estimates

JASON routinely reports error estimates for peak parameters using Peak Table as shown on a screenshot

- JASON estimates the errors for all peak parameters.
- The errors are generated when non-linear fit is converged. Calculation is based on Fisher Information Matrix.



A simplified representation of error estimates in non-linear fitting



Can we make peak deconvolution better?

- Is time-domain better than frequency domain deconvolution?
- Should we fit both, real and imaginary parts of the spectra?

Unitary operations should not change information contents (lossless transforms). But data is different

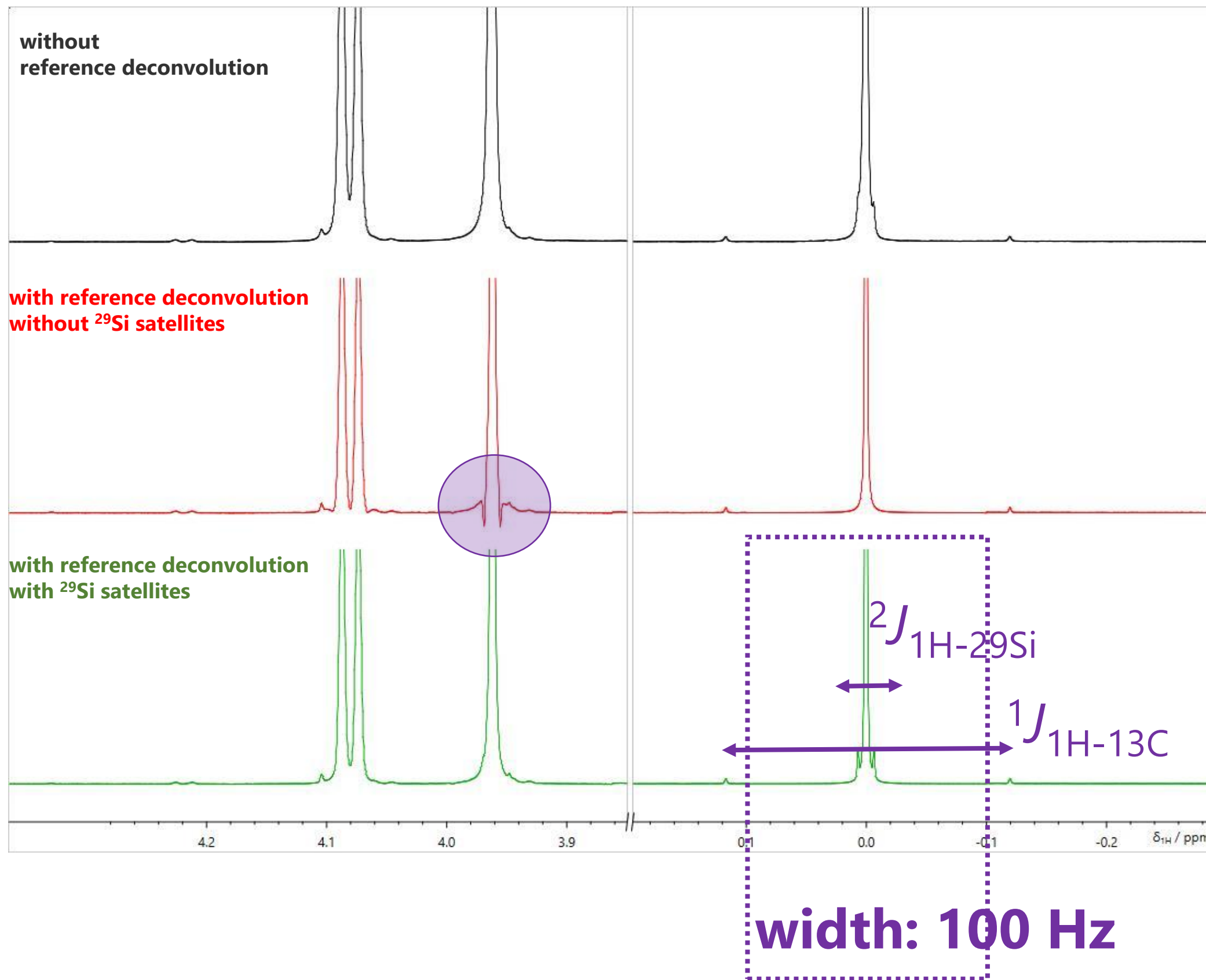
- Would resolution enhancement/apodization/etc. help?

Could be “lossy” transformations or, in opposite, adding information. Is the extra information trustable?

- Avoid peak deconvolution? For example: reference deconvolution?

A form of resolution enhancement. Adds information that all shapes in spectrum are identical

Reference deconvolution: satellites



Reference deconvolution

Reference

Position: 0.00000 ppm

Width: 100.00 Hz

Shift: 0.000 Hz

Lorentz width: 0.400 Hz

Gauss width: 0.600 Hz

Use satellites

Nuclides 1

Show #1 - +

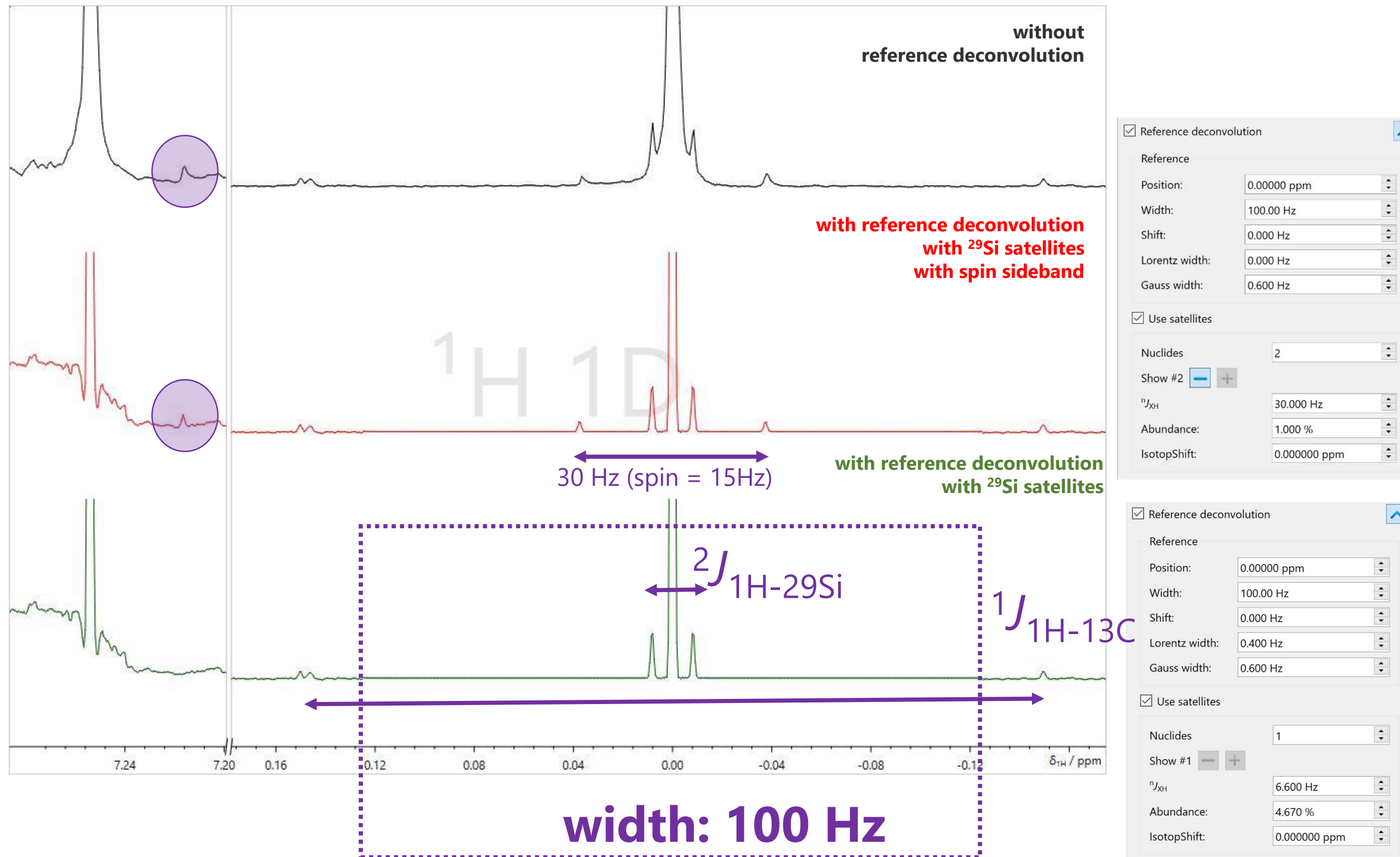
$^nJ_{\text{XH}}$ 6.600 Hz

Abundance: 4.670 %

IsotopShift: 0.000000 ppm

Satellites inside the reference region needs to be included in model, otherwise artefacts appear on all other peaks (see middle trace)

Reference deconvolution: spin sidebands and resolution boost



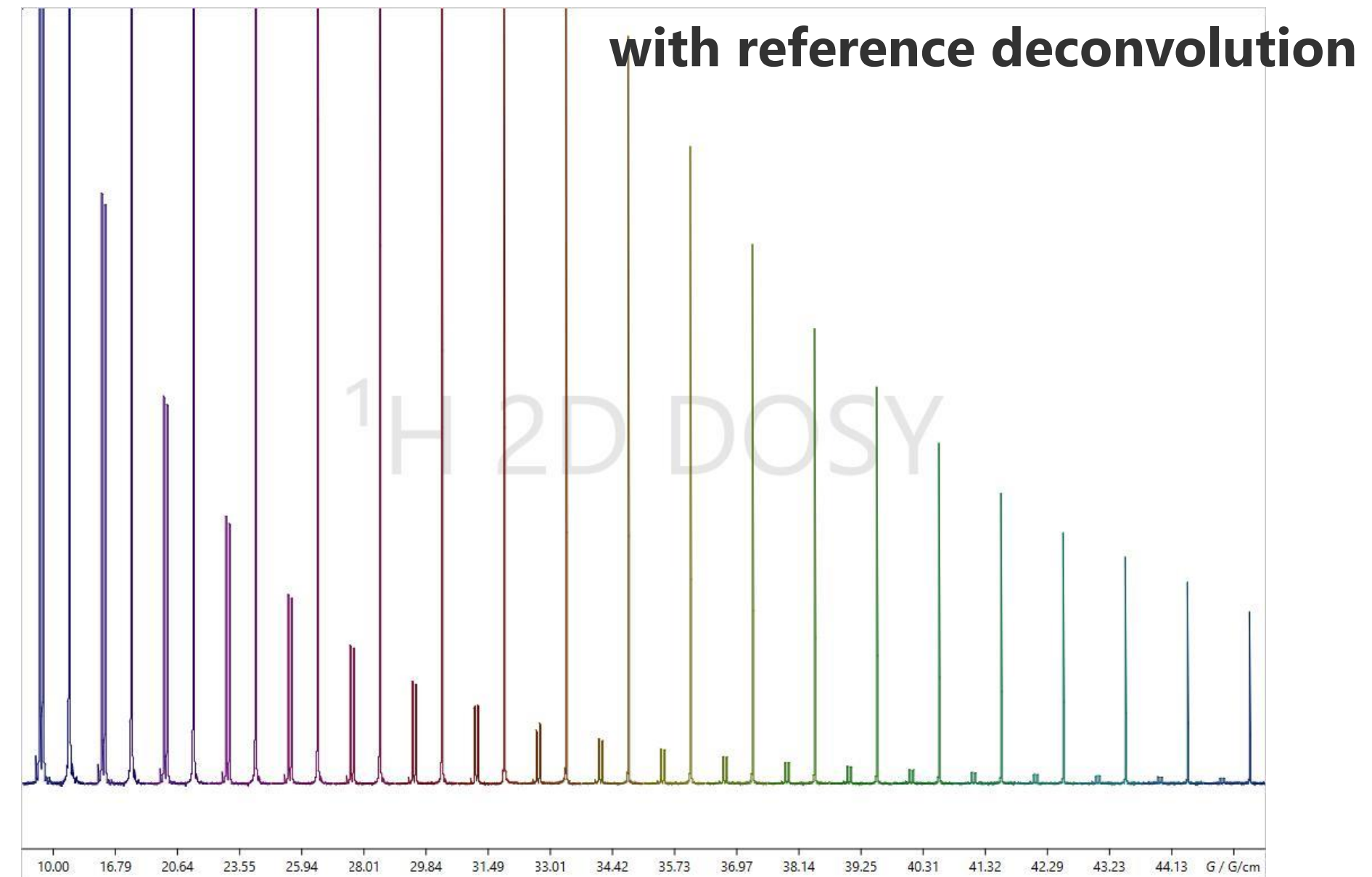
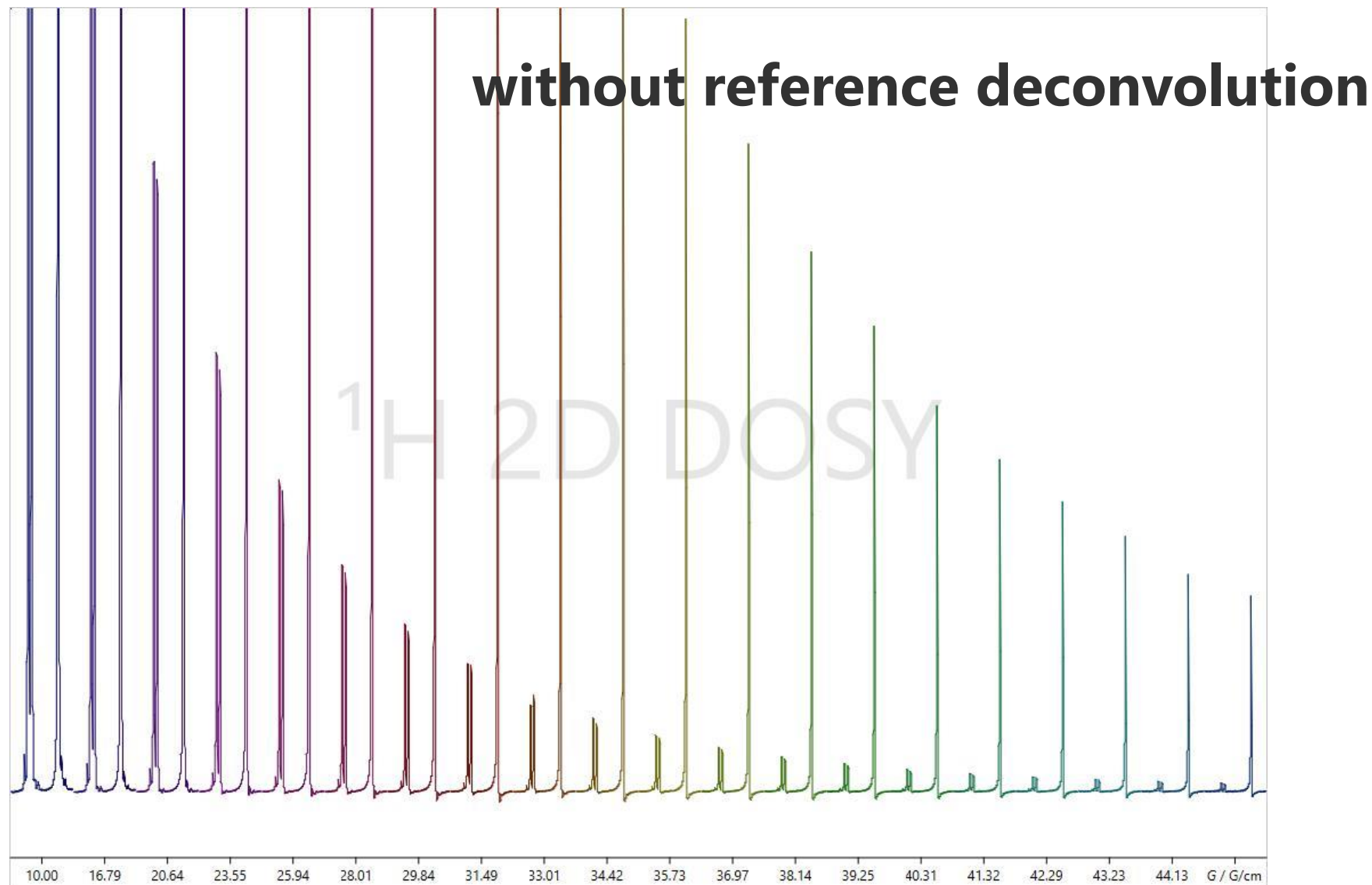
Included the spin sidebands in model, defined as a satellite doublet (see middle trace). This retains them in the corrected spectrum, see CHCl_3 solvent peak on the left

Conclusion: if any impurity peak or satellite appears in reference region, they need to be included in the reference model. Spin sidebands are the same in all signals so they can be removed, but not in solids spectra where their relative intensity is not the same at all peaks

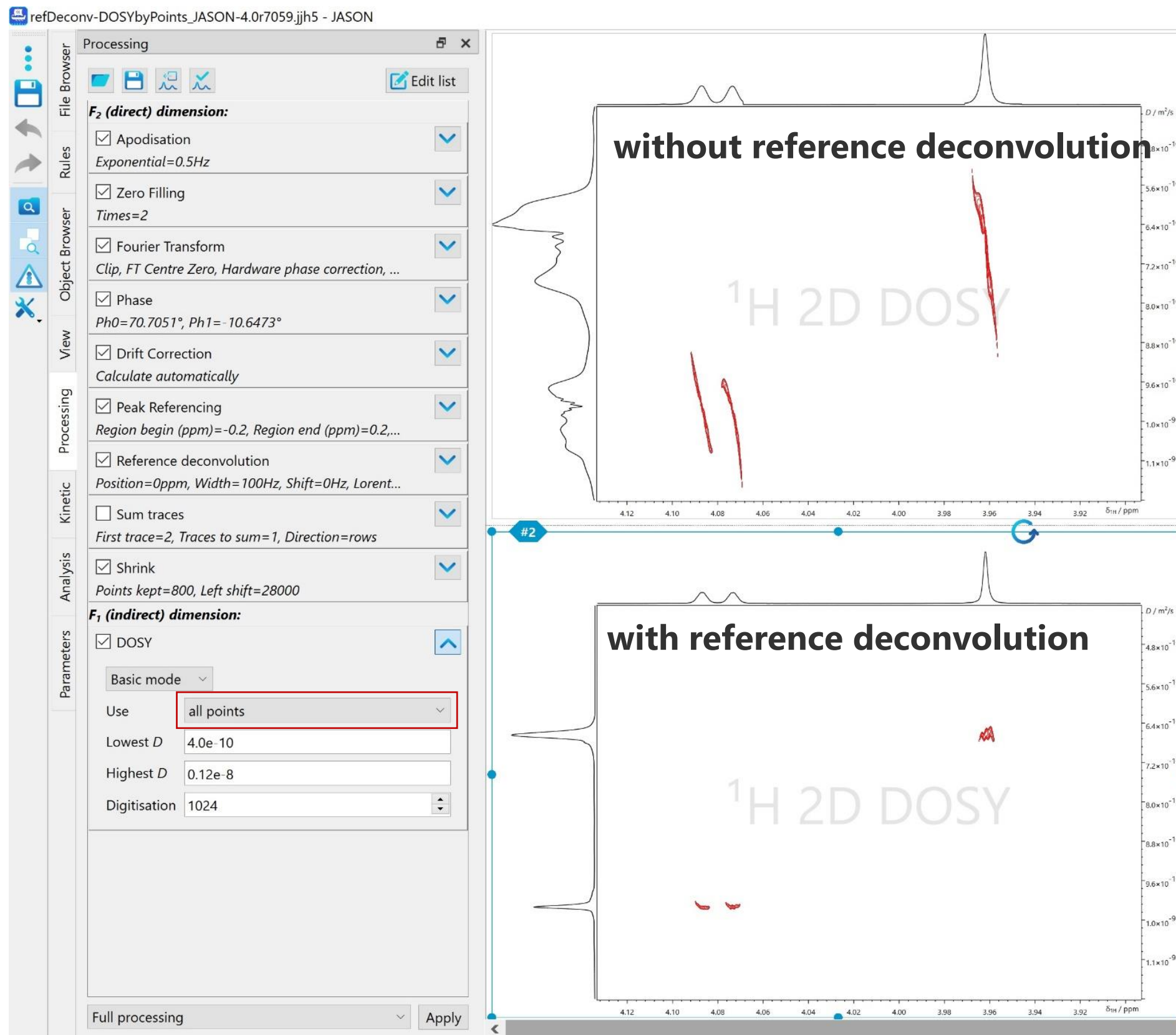
For the same reason, temperature variation induced distortions can't be corrected using reference deconvolution

Reference deconvolution: DOSY gradient array

Selected region shows two signals with experimental errors distorting the decay. Correctable by using a reference peak (TMS)
sample: quinine + camphor + geraniol mixture



Reference deconvolution: DOSY from all points

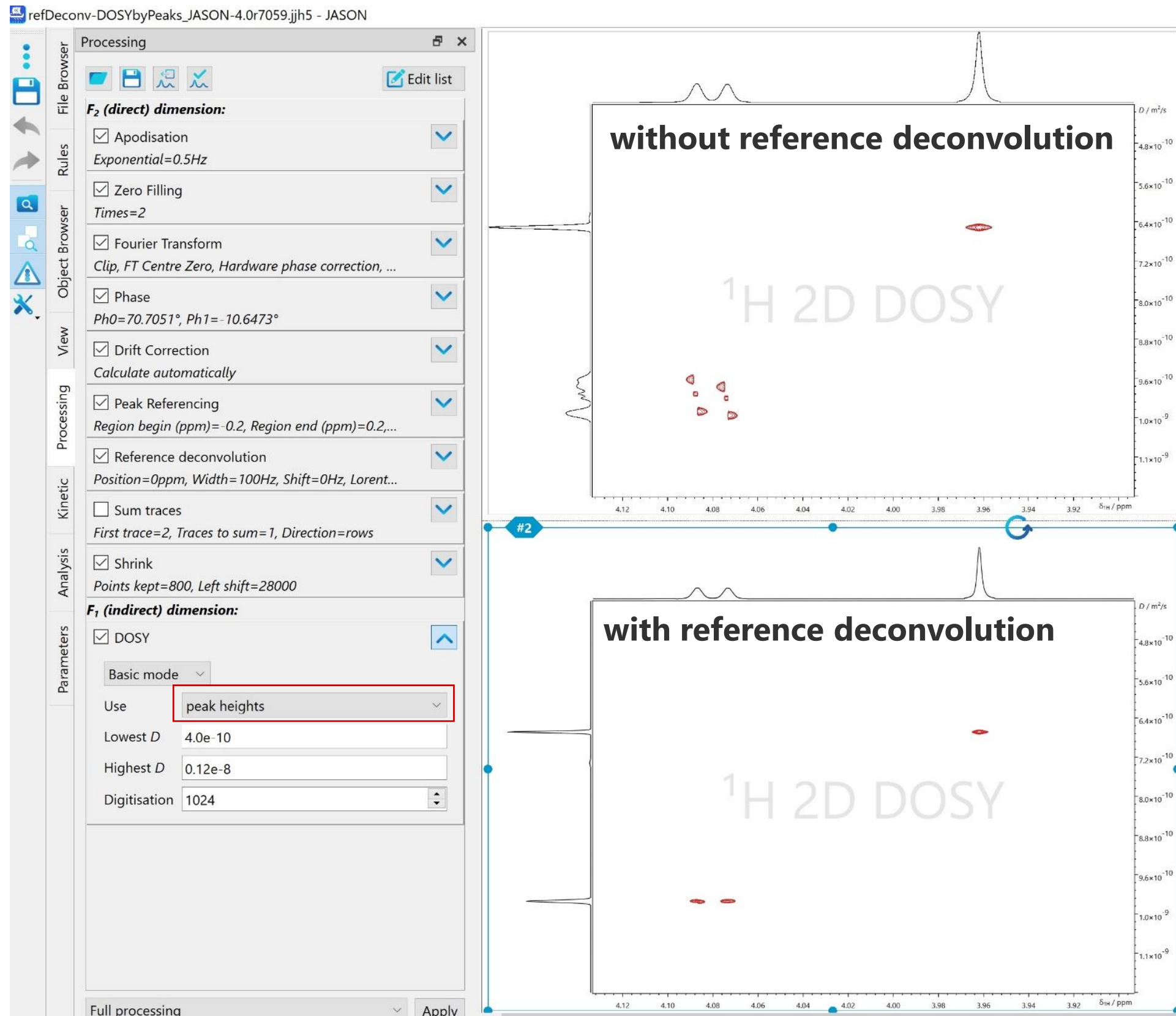


Using *all points* in DOSY highlights consequences of experimental imperfections

Spread of signals across the diffusion (vertical) domain is much less when reference deconvolution was utilized to improve data quality before fitting for DOSY

Note: high enough digitization is needed to visualize such DOSY plots without distortions in the horizontal projection. Just for demonstration, in practice we don't use this type of plots in real applications

Reference deconvolution: DOSY from peak heights



When using *peak heights* in DOSY

Individual lines of multiplets may not appear at the same diffusion coefficient. Their spread can be greater than what fitting error suggests for uncertainty

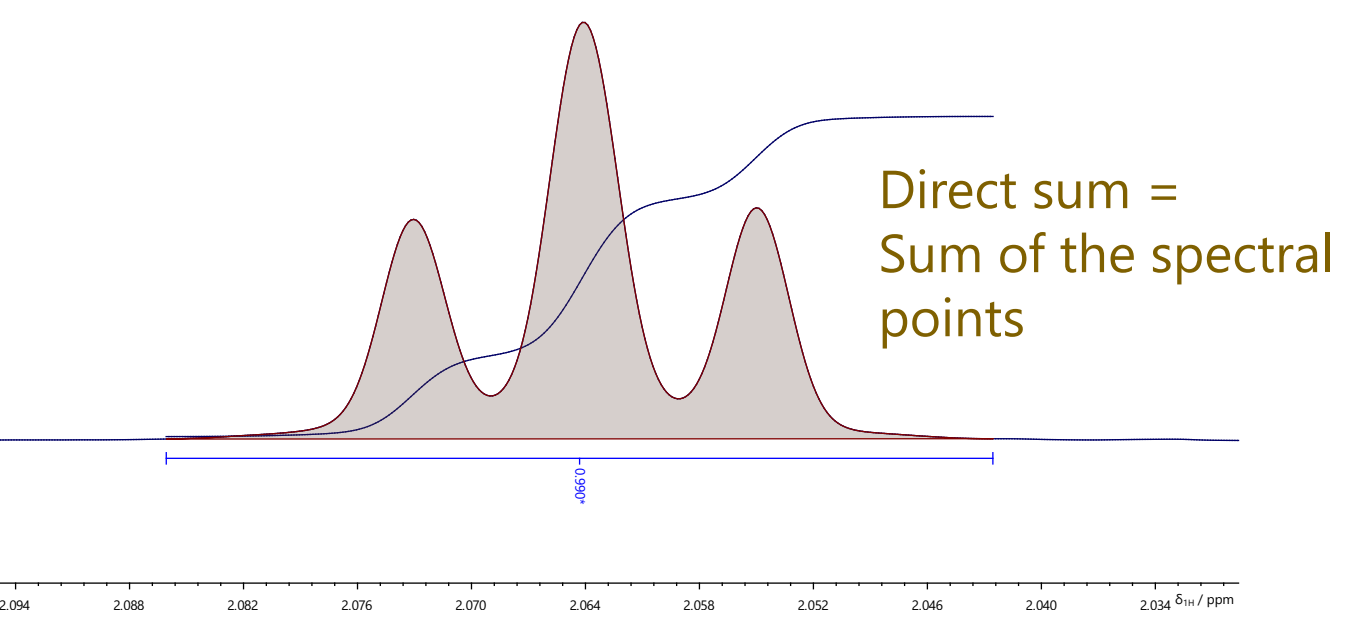
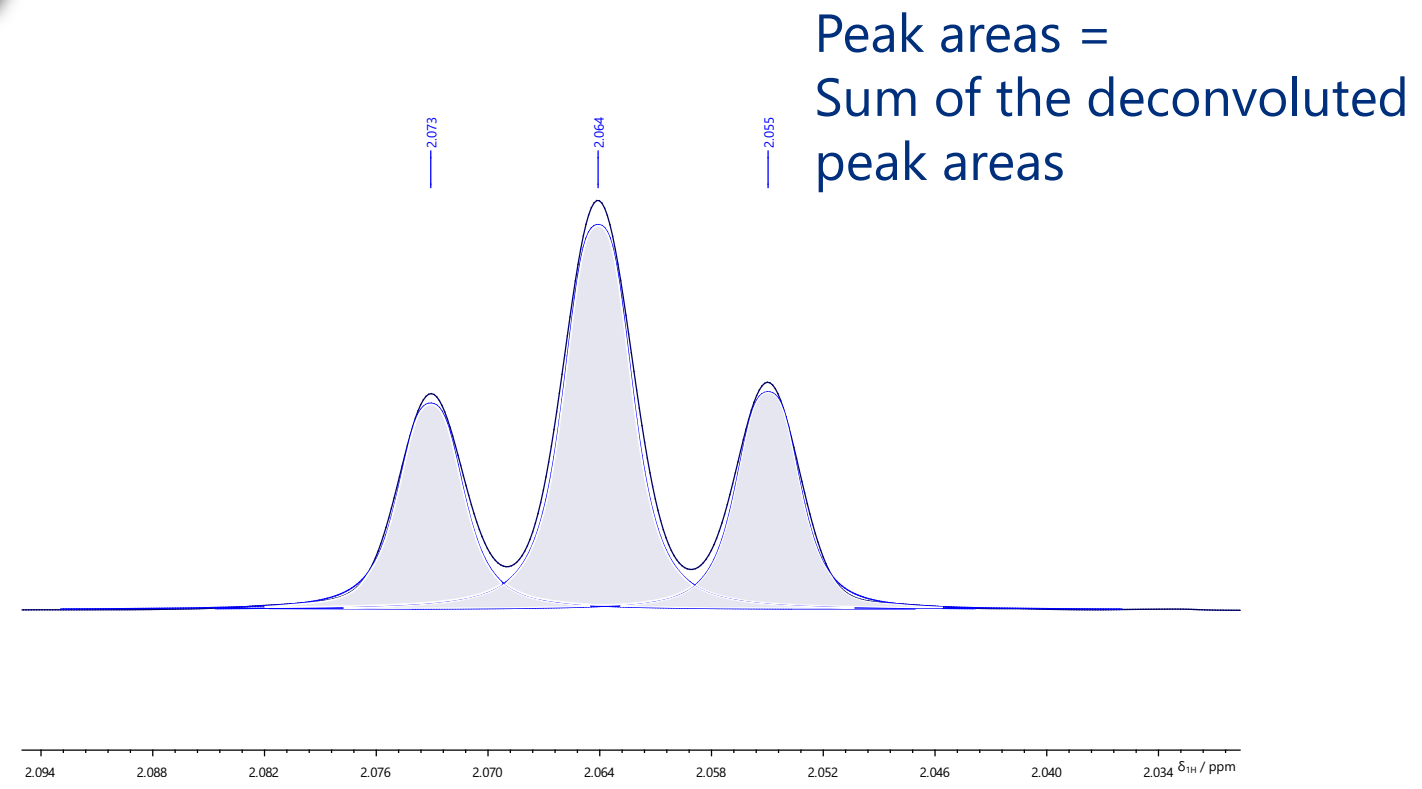
Reference deconvolution improves significantly.

Align of peaks is much better, and also for the singlet the sharper appearance of the signal along the diffusion (vertical) axis is due to the improved fitting.

Linewidth along the NMR domain (horizontal) reflects the use of apodization



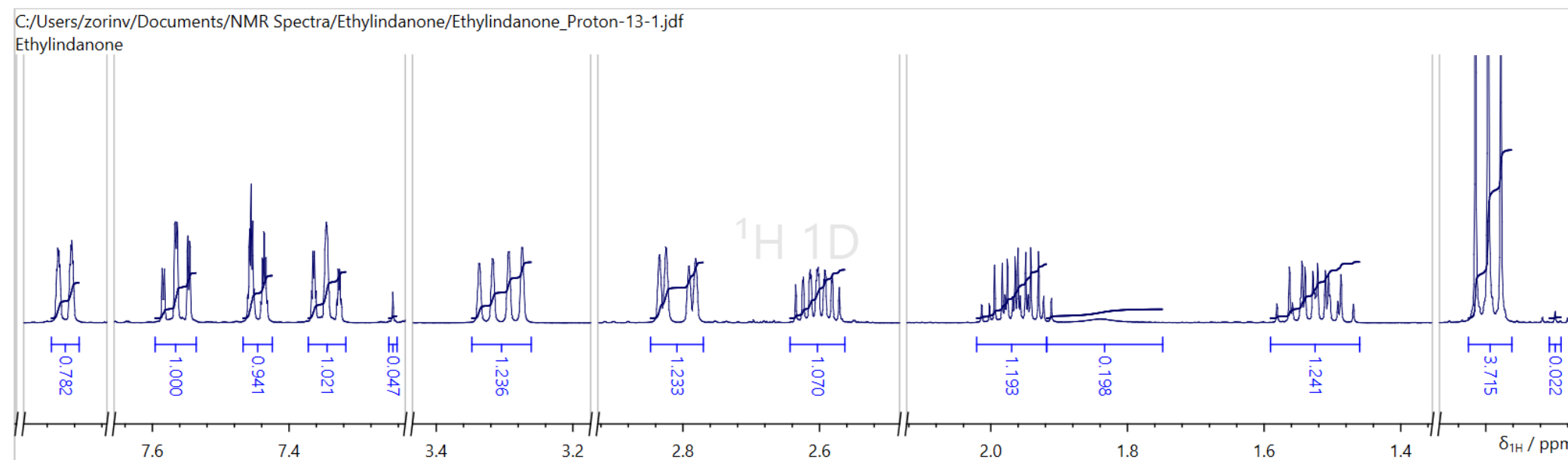
Direct (sum) integration vs peak areas



Direct sum	Peak Areas
Straightforward and simple	Require a well-performed peak deconvolution analysis
Depends on the region. Changing region changes the value – the larger region increases the integrals as the “wings” contribution is accounted for	Analytical equations for peak areas are used, so no regions required – only the peaks should be selected. Peak areas could be a bit larger than the sum integrals
Affected by the baseline distortions	As peak models are compensated for the baseline offset, their areas can be estimated without the baseline effect
Affected by peak overlaps	Peaks in overlaps are separated and their area estimated individually
If properly performed, can have high accuracy (of ca. 0.5%) for the values estimates	Accuracy is affected by the quality of the model (how good are the residuals). Usually worse than sum integrals – around 1-2%



Direct integration errors



Integrals Table from #5

Pos (ppm)	Start (ppm)	End (ppm)	Normalised	Sum Integral	σ Sum Integral
7.734	7.713	7.754	0.782	2.68e+04	9.14
7.566	7.536	7.596	1.000	3.43e+04	11.09
7.446	7.424	7.467	0.941	3.23e+04	9.39
7.344	7.317	7.372	1.021	3.50e+04	10.60
7.248	7.242	7.254	0.047	1605.47	5.05
3.304	3.261	3.348	1.236	4.24e+04	13.32
2.809	2.770	2.847	1.233	4.23e+04	12.56
2.603	2.563	2.643	1.070	3.67e+04	12.79
1.970	1.918	2.021	1.193	4.09e+04	14.45
1.833	1.748	1.918	0.198	6795.93	18.59
1.525	1.460	1.591	1.241	4.25e+04	16.29
1.266	1.259	1.272	0.014	475.70	5.27
1.133	1.127	1.139	0.027	910.09	5.16
0.994	0.962	1.025	3.715	1.27e+05	11.39
0.898	0.890	0.907	0.022	758.62	5.99

Error estimates

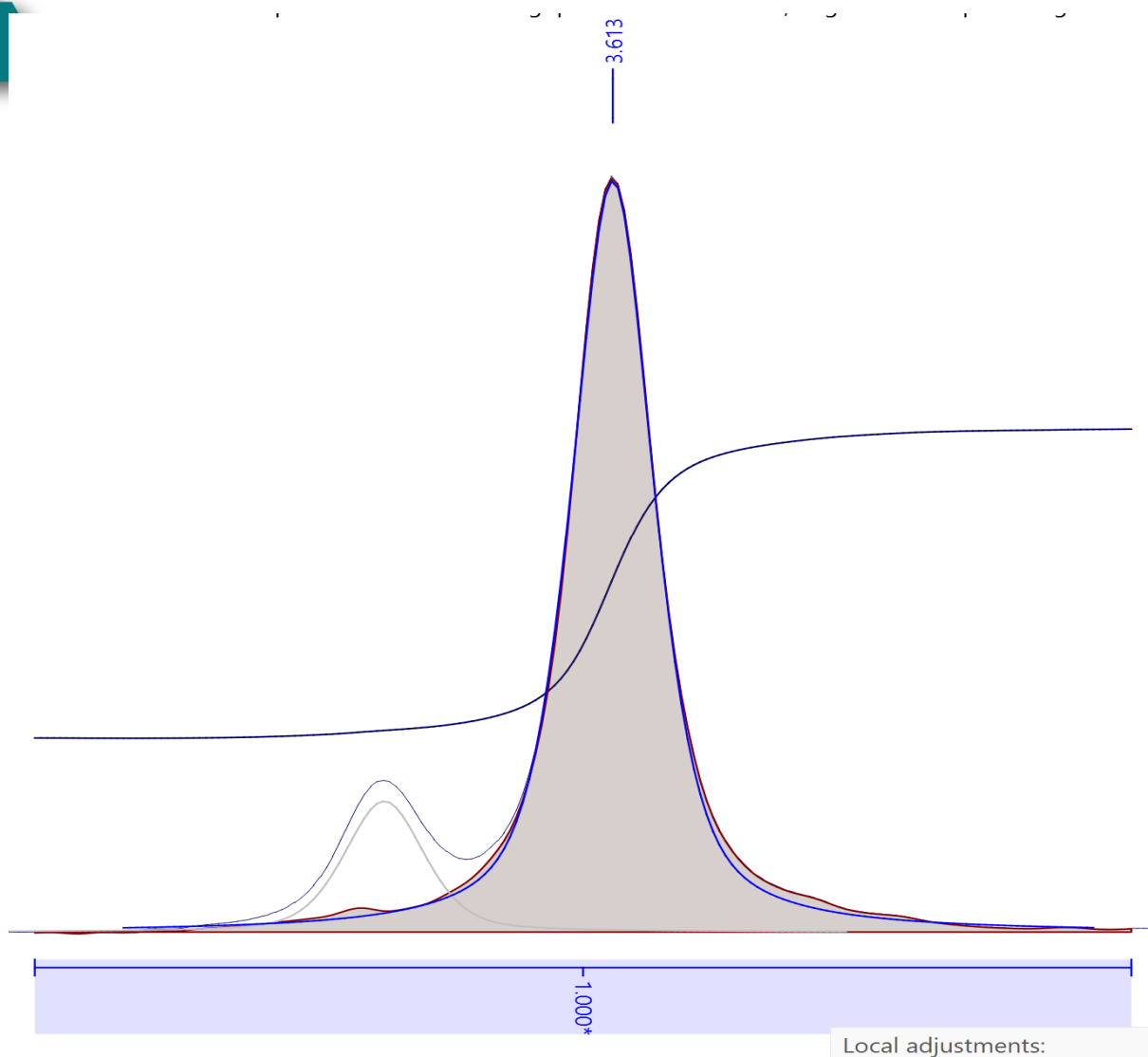
Direct integration error can be estimated by the error propagation rule from the error (noise) of each spectral point:

$$\sigma_{integral} = \sqrt{n \cdot \sigma^2} = \sigma\sqrt{n}$$

where σ – spectral noise, n – number of points in the integral

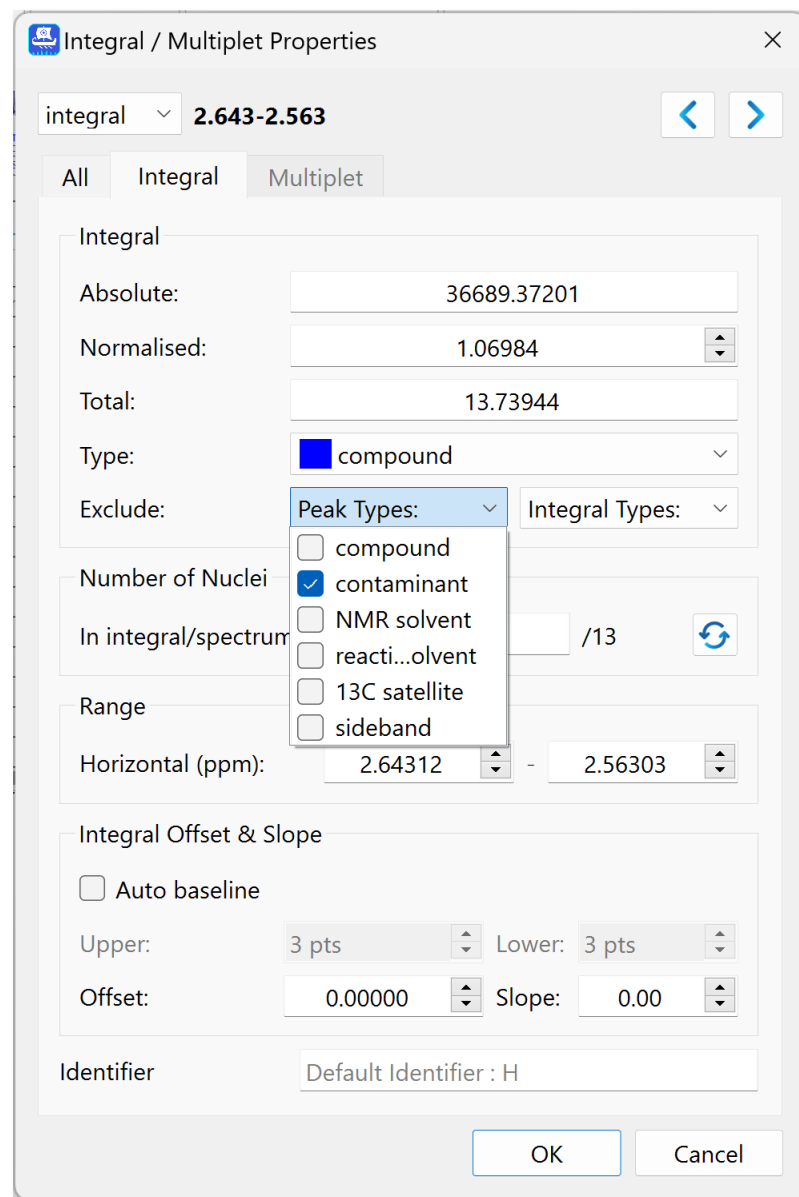
Estimates only stochastic error.
Biases: baseline distortion, overlapping impurities, phase error, insufficient region

Excluding impurities



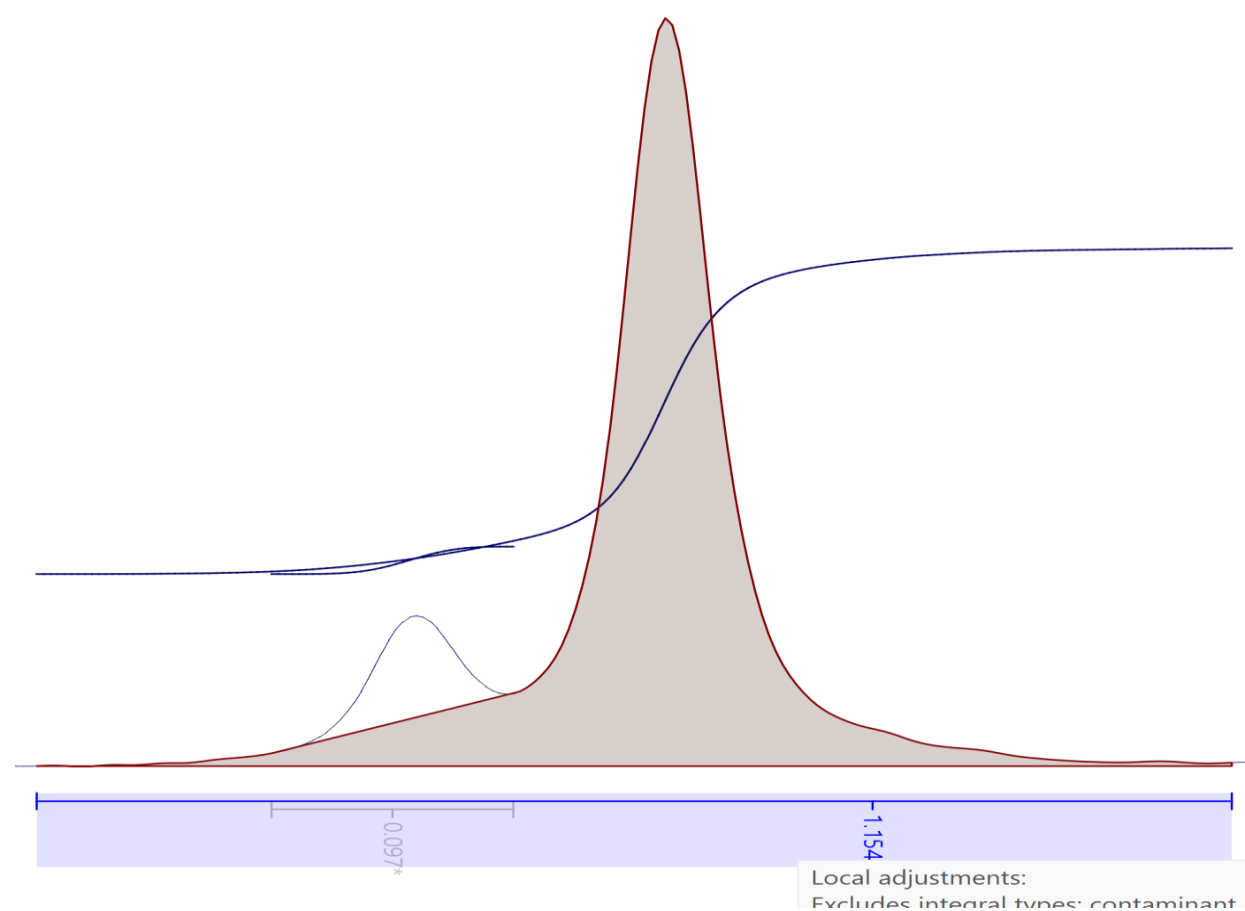
For each integral a peak areas of the specified type (e.g. contaminants) can be excluded:

- Requires good peak deconvolution

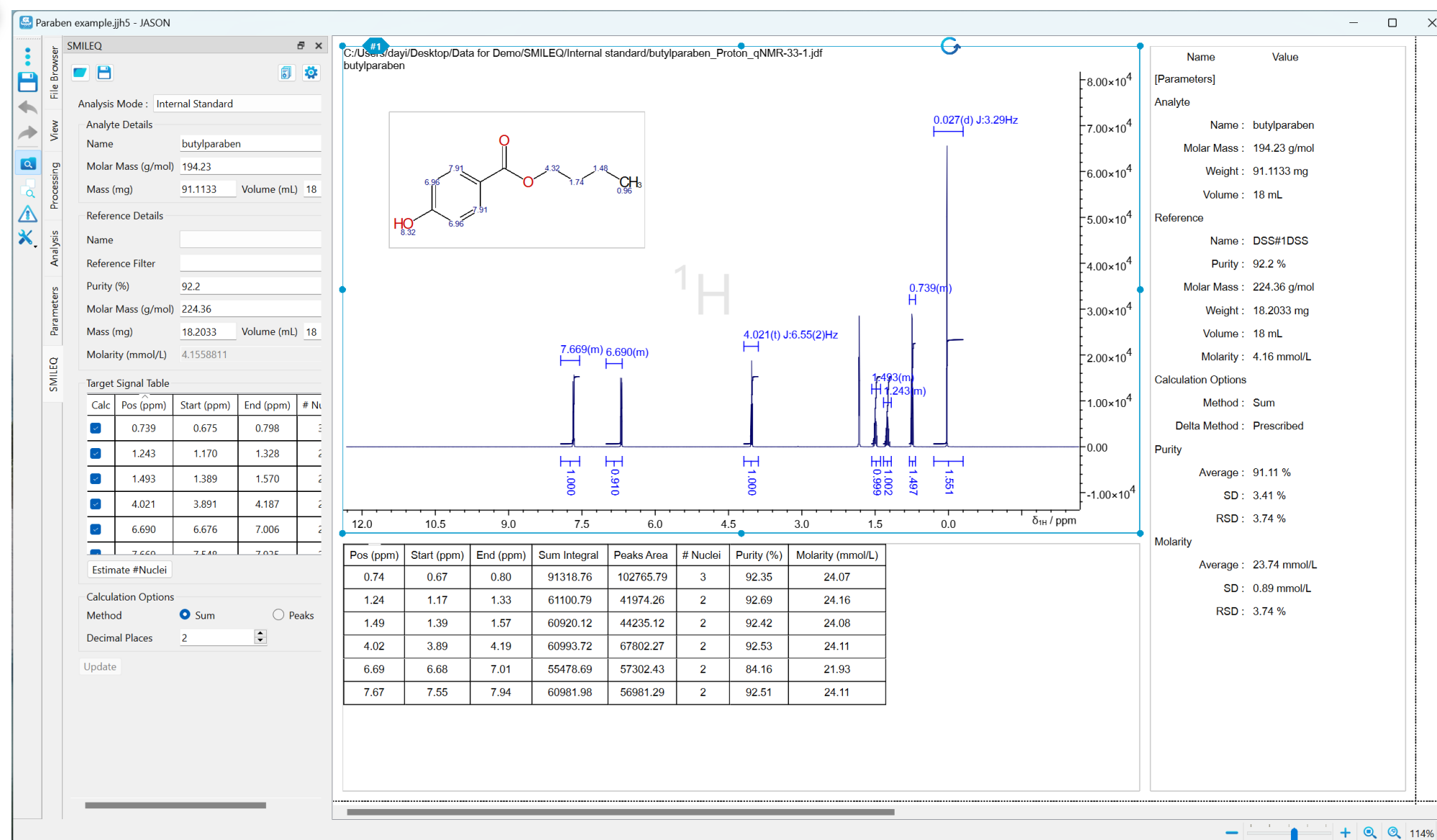


For each integral an incorporated integral of the specified type (e.g. contaminants) can be excluded (chromatographic integration):

- Could be less precise and depends on baseline setting



qNMR with SMILEQ



- Internal or External standard methods
 - PULCON and SOLCOR corrections
- Report generation
- Batch mode for multiple repetitions

- ISO 24583 compliant uncertainty reporting
- Seamless integration with Delta



Out of the scope: what we didn't discuss

JASON has a lot to offer!

- Power NMR processing engine
- Automated analysis
- Easy reporting
- Processing, analysis and layout "Rules" to get NMR analysis and report done as soon as you open a spectrum
- Chemical shift prediction, manual and automatic assignment
- Charts and fitting (relaxation, DOSY, reaction monitoring and beyond)
- DOSY and ROSY plots and analysis
- Open architecture for easy integration with external scientific tools
- BeautifulJASON: a powerful automation in Python
- Extendable plugin architecture
- MAGRES: a plugin for NMR crystallography
- **Affordable!**

Try it yourself:

<https://jeoljason.com>

