A new approach to interpret non-negative least squares (NNLS) T2relaxation results

Guilhem Pages, Amidou Traore, J.-M. Bonny

A new approach to interpret non-negative least squares (NNLS) T_2 relaxation results

Guilhem Pagès, Amidou Traoré, Jean-Marie Bonny

AgroResonance, UR 370 QuaPA–INRA, F-63122 Saint-Genès-Champanelle, France

The modelisation of an NMR signal decay in a sum of exponential terms is an ill-posed problem. Experimental data are not sufficient to find both relaxation times and amplitudes. Several solutions lead to minimal least-square distance between the model and the experimental data. To reduce the number of solutions, an efficient strategy consists in adding a constraint of positivity on all the parameters. Non-negative least-squares (NNLS) algorithm (1) is the most popular algorithm incorporating this constraint. The relaxation time values are *a priori* set in the decomposition basis (DB), the algorithm returning an unique solution of positive amplitudes.

To obtain a smooth amplitude distribution, a Tikhonov regularization is most often performed after the NNLS analysis. The choice of the regularization parameters is operator-dependent and is based on both prior-knowledge and T_2 distribution hypothesis (2).

Considering that only amplitude positivity is an indisputable *a priori*, we propose here to scrutinize in details the solutions provided by NNLS without further regularization. We show by simulations that interpreting NNLS results from the cumulative distribution function (cdf) leads to more robust analyses than an interpretation by probability density function (pdf) as usually done.

Probability *vs* Cumulative density functions

1. Simulation of T_2 signal decays for different models and addition of noise
2. Analysis of each of the 10,000 FIDs by NNLS without further processing

→ Averaged NNLS behavior compared with the theoretical responses

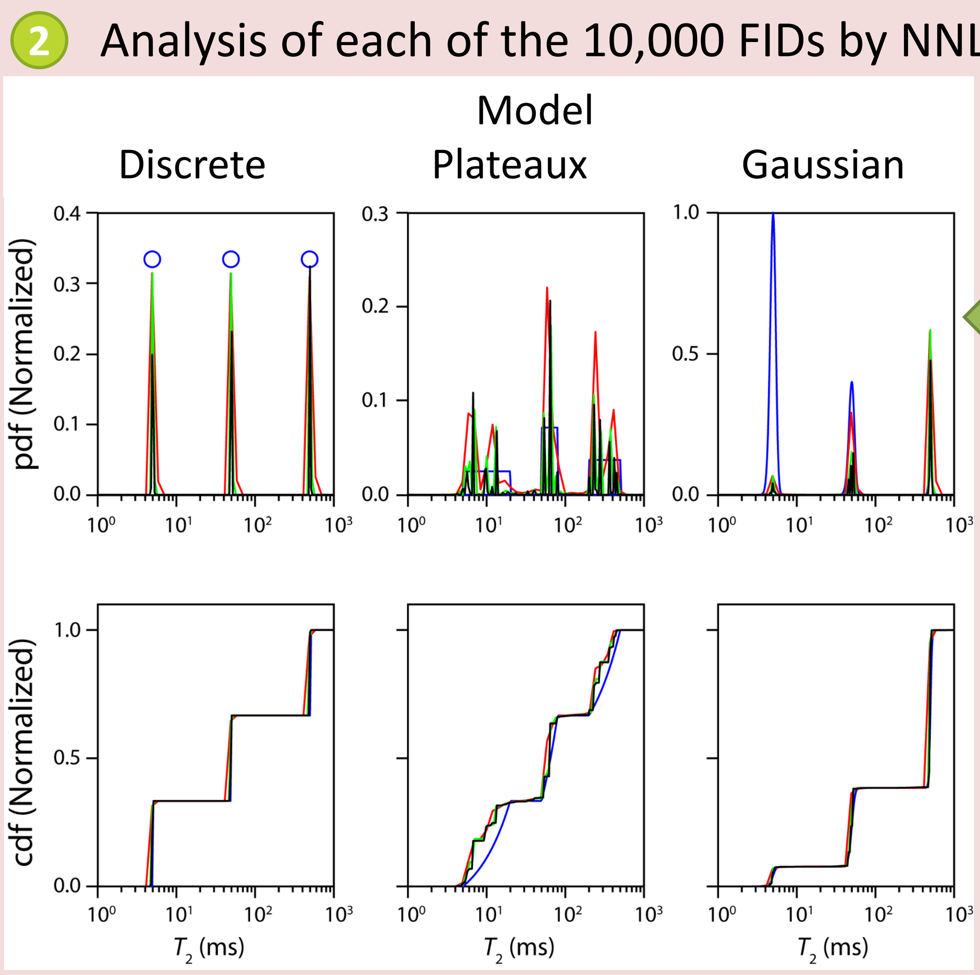
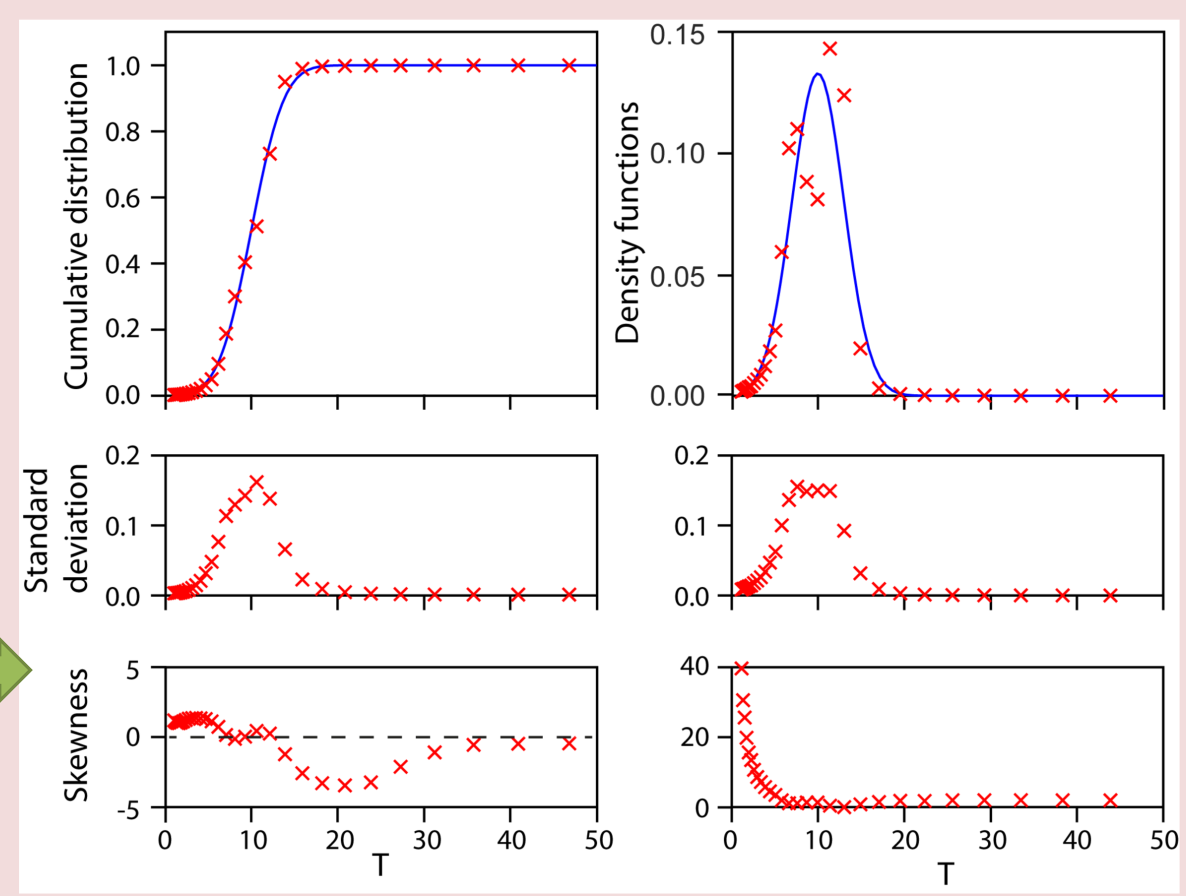


Figure 1: Averaged results for 10,000 simulated T_2 NMR signal decays (SNR 10^3) analysed with NNLS algorithm and a DB containing either 40 (blue) or 200 (red) T_2 values. Black lines represent the theoretical values. Representation as a pdf (top) and cdf (bottom).

Figure 2: Representation of the average values calculated from the 10,000 simulations for each of the 30 points of the DB as a cdf (top left) and pdf (top right). The middle plots represent the standard deviation for each point while the bottom ones the skewness. The blue lines represent the model.



- Pdf analysis are highly subjected to bias, especially for the continuous distribution
- Cdfs give the true amplitudes when they reach the plateaus (i.e. pdf~0)
- Cdf amplitudes are almost independent of the DB used

Non-null amplitude on the first point: Edge artefact

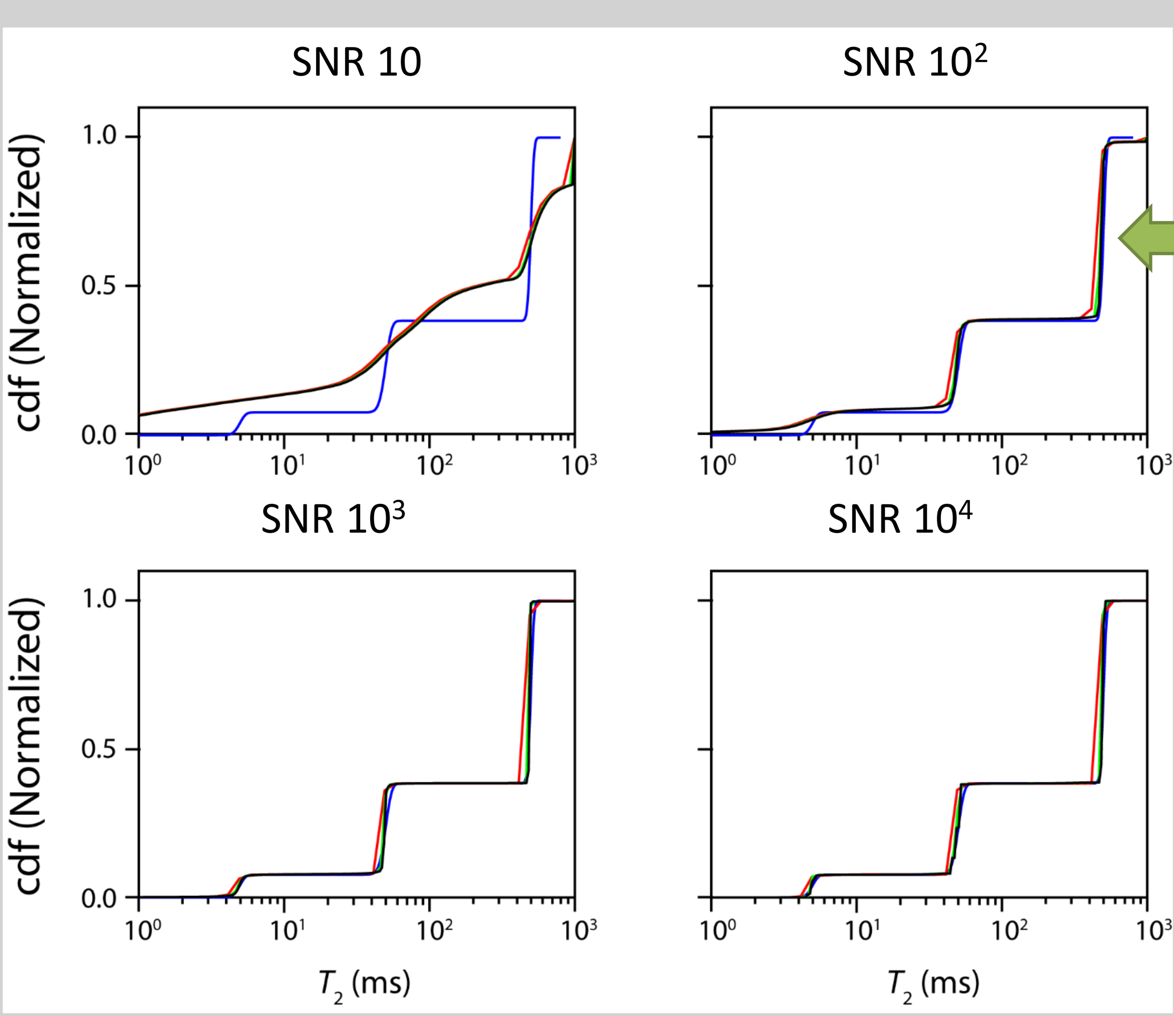
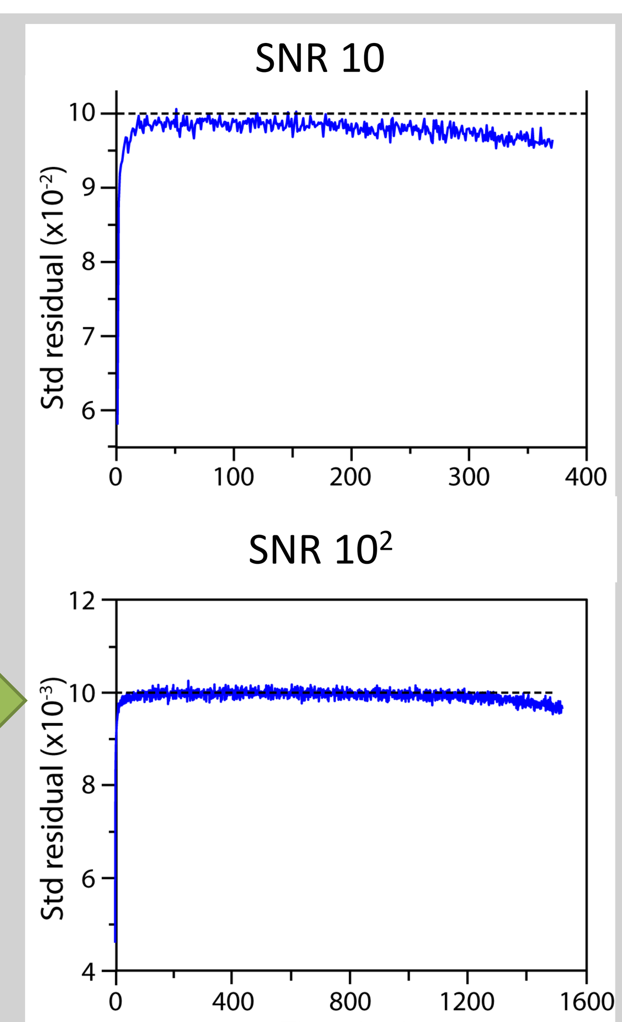


Figure 3: Averaged cdf results obtained from 10,000 NNLS analysis of noisy repetitions of the Gaussian model. The SNR was set to 10^1, 10^2, 10^3 and 10^4. For both lowest SNRs, non-null amplitudes can be seen for the first point, i.e. edge artefacts.

Figure 4: Standard deviation of the residuals of the 10,000 NNLS analysis (blue). While the standard deviation is at the noise level for most of the points, it is significantly lower for the first points. It demonstrates that NNLS uses the shortest T_2 values of the DB to adjust noise.



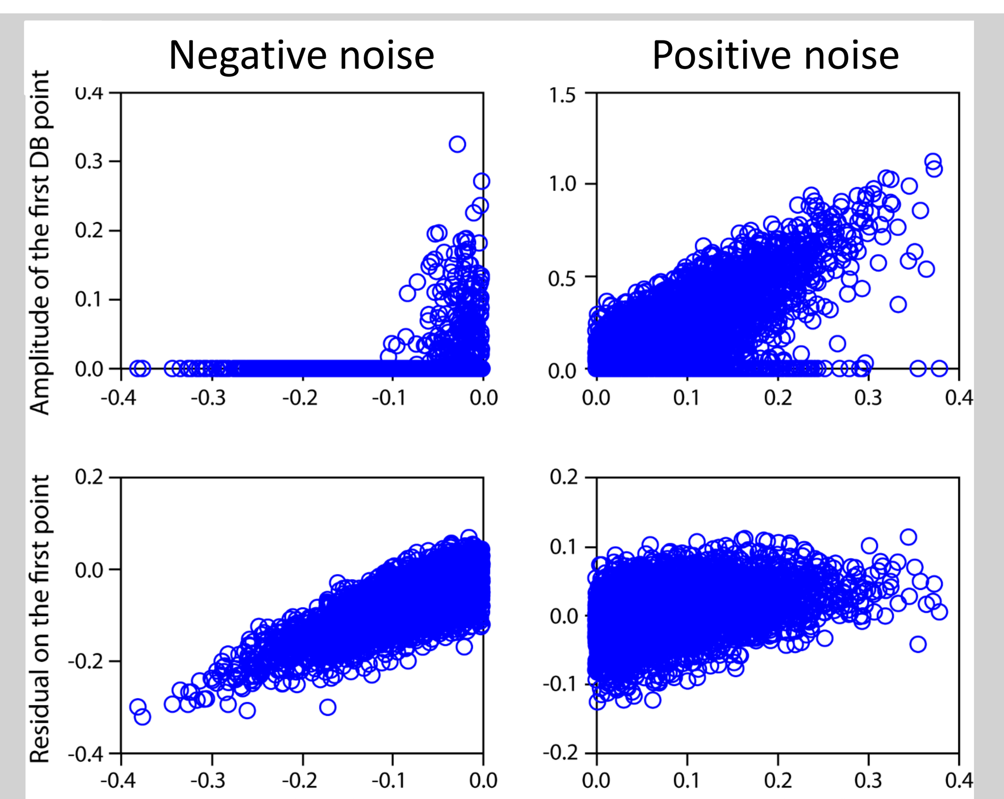


Figure 5: Analysis of the amplitude (top) and residual (bottom) on the first point of the DB (SNR 10) in function of the sign of the noise. On one hand, the correlation between the amplitude and the positive noise highlights the adjustment of the noise by the algorithm. On the other hand, negative noise is propagated into the residual signal due to the non-negativity constraint.

→ Edge-artefact is explained by the fitting of positive noise

Improving NNLS stability

How to set the first DB T_2 value (i.e. $T_{2,\min DB}$)

To be able to estimate a T_2 component, two following points of the signal must be at least higher than the signal (S) to noise (σ) ratio (k), i.e.:

$$S(2TE) > k\sigma$$

Considering A the amplitude of this $T_{2,\min DB}$ value, it leads to

$$T_{2,\min DB} = \frac{2TE}{ln\left(\frac{A}{\sigma}\right) - ln(k)}$$

SNR	$T_{2,\min DB}$
10^1	2 TE
10^2	TE
10^3	½ TE
10^4	¼ TE

Parsimony of the DB

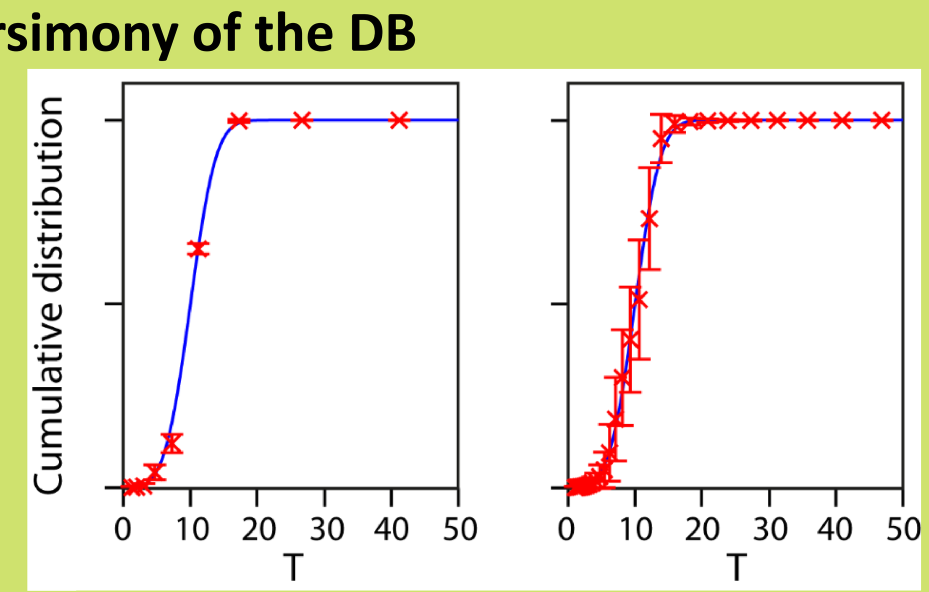


Figure 6: Average cdfs after simulation and analysis of 10,000 noisy FIDs (SNR 10^3). The DB contained 10 (left) and 30 (right) points. The standard errors on the cdf during a change in amplitude is significantly more important when the DB density increases.

→ Rational approach to set the user-dependent parameters

Conclusions

- To limit user-inputs into NNLS analysis, we push the idea that the cdf distributions are sufficient for obtaining useful information provided that the analysis focusses on the discernable plateaus
- Result interpretation based on cdf led to more robust results as shown by simulations
- The edge-artefact is explained by the fitting of positive noise at the beginning of the signal
- The first value of the DB should depend on both TE and SNR. We mathematically justified its value of $T_{2,\min DB}$
- The number of points in the DB should be limited. This is a change of paradigm in regard of the current practice (dense DB)

AgroResonance, UR370 QuaPA – INRA
F-63122 Saint-Genès-Champanelle, France
http://www6.inra.fr/agroresonance

(1) C.L. Lawson, R.J. Hanson, Solving least squares problems, Prentice-Hall, 1974, 23, 161.
(2) K.P. Whittall, A.L. MacKay, J. Magn. Reson., 1989, 84, 134.