BAYESIAN DOSY:
A NEW APPROACH TO DIFFUSION DATA PROCESSING

MESTRELAB RESEARCH
NMR Solutions

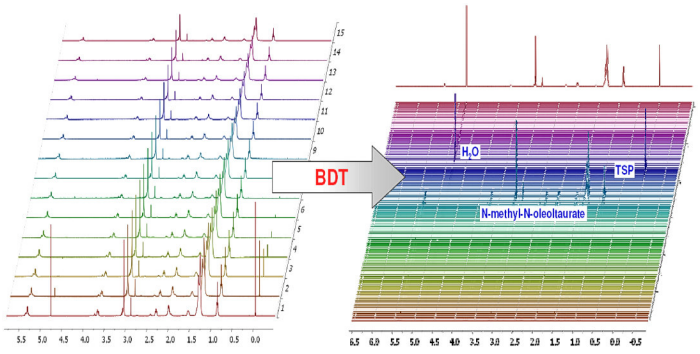
Carlos Cobas [1], **Stanislav Sýkora** [2]

[1] Mestrelab Research, Xosé Pasín 6, Santiago de Compostela, 15706 Spain; carlos@mestrec.com
[2] Extra Byte, Via Raffaello Sanzio 22/C, Castano Primo (Mi), Italy I-20022; sykora@ebyte.it

BDT: A novel Bayesian-based approach to the DOSY Transformation

NMR Diffusion Ordered Spectroscopy (**DOSY**) is one of the NMR techniques which present conceptual problems with the visual representation of the results in terms of the physical parameter associated with each spectral peak (diffusion coefficient). Current approaches are often little more than graphic sketches based on the evaluation of the decays at a selected set of spectral points (usually those corresponding to peak tops). Other, physically more acceptable approaches, such as the MaxEnt algorithm [1], unfortunately require very long processing times.

Here we present recent advances in the development of a new Bayesian Dosy Transform (BDT) algorithm for the processing of DOSY data which was initially presented at the ENC 2008 Conference. This algorithm is computationally **very efficient** and, at the same time, eminently **physically meaningful**. In particular, it treats all data points in the same way, and leads to **artifact-free**, **quantitative** results.



The Bayesian DOSY Transform (BDT) algorithm applied to an aqueous solution of *potassium N-methyl-N-oleoltaurate* with *TSP*.
The Varian FID file of this surfactant, measured at 23 ºC, has been obtained from the *VARIAN MR USER GROUP LIBRARY* (submitted by Brian Antalek to illustrate his DECRA algorithm).

The Bayesian approach in a nutshell

Let z be the gradient-dependent decay parameter associated with each DOSY spectrum and d the diffusion coefficient associated with a spectral line at some frequency f. Then the essence of a a Bayesian approach to the problem of transforming the experimental S[f,z] map to a probabilistic [f,d] map is embodied in the following question:

If I select an arbitrary point in the [f,d] map, can I assign to it an amplitude and a probability which give me an idea of how compatible (congruent) that particular point in the [f,d] map is with the experimental [f,z] map?

Having fixed f and d, one has only one parameter to fit, namely the decay-curve amplitude $a_k(d)$, k being the data index corresponding to the frequency f. This is a linear, one parameter fit which can be done explicitly. Assuming the decay curve

$$y_{k\alpha} = a_k(d)\exp(-dz_\alpha), \text{ where } \alpha \text{ is the index within the z-set,}$$

the value of a_k which gives the minumum of $\Sigma_\alpha (y_{k\alpha} - S_{k\alpha})^2$ is

$$a_k(d) = [\Sigma_\alpha S_{k\alpha}\exp(-dz_\alpha)] / [\Sigma_\alpha \exp(-2dz_\alpha)].$$

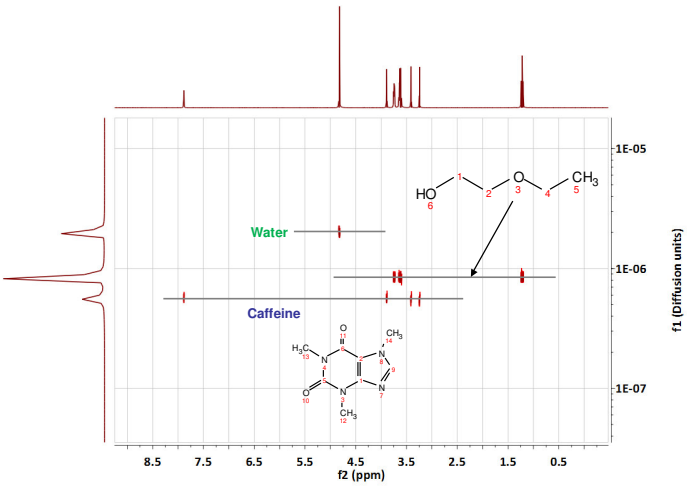
Even more important than the optimal amplitude is the Bayesian weight w_k associated with it. Denoting as σ the standard deviation of the experimental noise, the multiplicative contribution of the experimental data point $S_{k\alpha}$ to w_k equals $\exp(-(y_{k\alpha} - S_{k\alpha})^2/(2\sigma^2))$. Considering all z-points, we have

$$w_k = \Pi_\alpha \exp(-(y_{k\alpha} - S_{k\alpha})^2/\sigma^2) = \exp(-[\Sigma_\alpha (y_{k\alpha} - S_{k\alpha})^2]/\sigma^2),$$

Whose maximum $w_k(d)$ coincides with the $a_k(d)$ obtained by least-squares. Notice that, in principle, any d is legitimate, but the value of $w_k(d)$ is appreciable only when d is close to the 'correct' value. When the chosen d differs from the correct one, there are large systematic deviations (residues) and w(d) becomes very small. $w_k(d)$ is the proper 'vertical' value to be assigned to the (f,d) point in the DOSY 2D-plot, apart from the final normalization.

The final normalization, which is a characteristic step of any Bayesian evaluation, is very simple. Once the $y_k(d)$ values were calculated for all possible d-values, they are scaled so as to make their sum (vertical projection) equal to the intensity of the original spectrum at z = 0 (usually defined by S_{k0}).

An additional improvement



BayDOSY of a mixture of *Caffeine, 2-EthoxyEthanol,* and *Water*
The original set of FID's file has been acquired on a Bruker 400 MHz spectrometer by Andy Soper at University of Rhodes, South Africa

Bayesian approaches are particularly suitable for incorporating a-priori knowledge of all kinds. In the case of **DOSY** this knowledge consists in the fact that all signals arising from a particular molecule must have the same value of diffusion constant. Properly worked out (again in the Bayesian spirit), this leads to what we call the **LineSNAP** algorithm. When this repetitively applicable technique is combined with **BDT**, it gives rise to a new algorithm which we call **BayDOSY**.

Its effect is an improved resolution along the D-axis and a better alignment of the signals along horizontal lines – both aspects which greatly facilitate in chemist in distinguishing and interpreting the individual sample components.

Conclusions

Compared to other approaches, the **Bayesian method** has been incorporated into the Mnova software package under the name **BayDOSY** and turns out to be extremely satisfactory. It automatically avoids having exact, unnatural zeros anywhere in the resulting map since *every point of the 2D map has has a well defined value of statistical congruence with the data*. Moreover, the **BDT maps** show 'normal' line widths in the f-direction, correctly positioned and resolved peaks in the d-direction and **quantitatively correct horizontal and vertical projections** – a combination difficult to achieved by any other means. With the recent addition of the **LineSNAP** feature, it also places a desirable statistical premium on the alignment of spectral peaks along horizontal lines in the [f,d] plot and makes the NMR-DOSY method behave ever more as a true analytical separative method.

The approach will be now extended to non-exponential cases arising from **overlapping lines**.

UNIÓN EUROPEA
Fondo Europeo de Desenvolvemento Rexional

XUNTA DE GALICIA
CONSELLERÍA DE INNOVACIÓN, INDUSTRIA E COMERCIO
Dirección Xeral de Investigación e Desenvolvemento

DOI: 10.3247/sl2nmr08.010

Developed in collaboration with www.ebyte.it