Volume 1 | Number 1 | Jan 2013 | Pages 1–100

Analyst

www.rsc.org/analyst

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Journal Name

ARTICLE TYPE

Monitoring polydispersity by NMR diffusometry with tailored norm regularisation and moving-frame processing

Mateusz Urbańczyk,[a,d] Diana Bernin,[b] Alan Czuroń,[c] and Krzysztof Kazimierczuk [*d]

Nuclear magnetic resonance (NMR) is currently one of the main analytical techniques applied to numerous branches of chemistry. Furthermore, NMR has been proven useful to follow in-situ reactions occurring on a time scale of hours and days. For complicated mixtures, NMR experiments providing diffusion coefficients are particularly advantageous. However, the inverse Laplace transform (ILT) used to extract the distribution of diffusion coefficients from an NMR signal is known to be unstable and vulnerable to noise. Numerous regularisation techniques have been proposed to circumvent this problem. In our recent study, we proposed a method based on sparsity-enforcing ℓ_1-norm minimisation. This approach, which is referred to as ITAMeD, has been successful but limited to samples with a 'discrete' distribution of diffusion coefficients. In this paper, we propose a generalisation of ITAMeD using a tailored ℓ_p-norm ($1 \leq p \leq 2$) to process in particular signals arising from 'polydisperse' samples. The performance of our method was tested on simulations and experimental datasets of polyethylene oxides with varying polydispersity index. Finally, we have applied our new method to monitor diffusion coefficient and polydispersity changes of heparin undergoing enzymatic degradation in real-time.

1 Introduction

Nuclear magnetic resonance spectroscopy (NMR) has found numerous applications in chemistry. In particular, following reactions in-situ has recently attracted much attention.[1] Changes of various spectral parameters, e.g. peak intensities, chemical shifts or diffusion coefficients, can be monitored. The latter ones, although less frequently used,[2,3] can be very valuable for studying degradation processes, when large molecules are fractionated into smaller fragments.

Commonly, the diffusion coefficient is estimated using the pulsed-field gradient (PGSE) technique, which is based on the signal attenuation during the time lapse between the encoding and decoding magnetic field gradient pulses. The attenuation correlates with the diffusion coefficient D of each compound in the sample as follows[4]

$$S(g) = S(0)e^{-Dg^2\gamma^2\delta^2\Delta'}, \qquad (1)$$

where $S(g)$ is the signal intensity for a given magnetic field gradient amplitude g, γ is the gyromagnetic ratio, δ is the duration of the magnetic field gradient pulse and Δ' is the effective diffusion time. For a continuous distribution of diffusion coefficients $A(D)$, one can modify Equation (1) as follows

$$\Psi = \frac{S(g)}{S(0)} = \int_{D_{min}}^{D_{max}} A(D)e^{-Dg^2\gamma^2\delta^2\Delta'} dD. \qquad (2)$$

Equation (2) describes the Laplace transform of the distribution of diffusion coefficients, $A(D)$, showing that the inverse Laplace transform (ILT) can be applied to obtain $A(D)$ from an experimental dataset Ψ. Unfortunately, this procedure is numerically unstable and highly prone to noise. Various methods have been proposed to circumvent this problem[5-14].

These methods can be divided into two groups based on the assumption regarding $A(D)$. The first group contains methods, which attempt to find only the diffusion coefficient of each compound and neglect the shape of $A(D)$. Applying a method of this group to polydisperse samples will bias the obtained diffusion coefficient as recently reported by Zhou et. al.[15]. However, they perform very well in case of monodisperse samples. Algorithms of

[a] Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw Żwirki i Wigury 101, 02-089 Warszawa, Poland
[b] Swedish NMR Centre, University of Gothenburg, BOX 465, 40530 Göteborg, Sweden
[c] Institute of Mathematics Polish Academy of Sciences ul. Śniadeckich 8, 00-656 Warszawa, Poland
[d] Centre of New Technologies University of Warsaw ul. Banacha 2C, 02-097 Warsaw, Poland, E-mail:k.kazimierczuk@cent.uw.edu.pl

the first group are e.g. Direct Exponential Curve Resolution (DE-CRA)[8], Speedy Component Resolution (SCORE)[9], Multivariate Curve Resolution MCR[16], Blind Source Separation[14] and mono-, and multi-exponential fitting[17].

The second group contains methods that can, to some extent, reconstruct the shape of $A(D)$. This can be utilitsed by fitting parameters of a strictly defined distribution (e.g. log-normal[18] and gamma distribution[19]), or by fitting a distribution of diffusion coefficients with additional constraints (Trust Region Algorithm for Inversion (TRAIn)[10], CONTIN[12], Maximum Entropy (MaxEnt)[11], and Iterative Thresholding Algorithm (ITAMeD)[13]).

The last three methods can be discussed on the basis of regularisation

$$\min_{A \geq 0} \|\Phi A - \Psi\|_{\ell_2}^2 + \tau \Theta(A), \tag{3}$$

where Φ is the Laplace transform matrix, A is the vector of the distribution of diffusion coefficients, Θ is the regularisation term, and τ controls the ratio between first and second term. For MaxEnt, Θ is defined as

$$\Theta(A) = -\sum_i \frac{A_i}{\sum_j A_j} \log \frac{A_i}{\sum_j A_j}, \tag{4}$$

while CONTIN uses

$$\Theta(A) = \|\mathbf{L}A\|_{\ell_2}, \tag{5}$$

where $\mathbf{L}$ is a matrix that contains prior assumptions about the data[20]. ITAMeD utilises the following regularisation term

$$\Theta(A) = \|A\|_{\ell_1}. \tag{6}$$

All aforementioned regularisations $\Theta(A)$ are equivalents of certain assumptions about the shape of $A(D)$. For example, ITAMeD assumes that the resulting vector is sparse, CONTIN makes the assumption that the distribution is smooth, while MaxEnt prefers $A(D)$ with the highest entropy.

In fact, none of the regularisation terms is generally valid. ITAMeD will not give the correct reconstruction for polydisperse samples with broad diffusion coefficient distributions, while CONTIN and MaxEnt may provide an over-smoothed result for samples with very sparse $A(D)$[13].

To circumvent these limitations we propose a new method using a tailored regularisation term, which is automatically tuned. This regularisation term exploits the ℓ_p-norm with $1 \leq p \leq 2$, that allows balancing between sparsity and smoothness of the resulting distribution. The order p norm is defined as follows

$$\|\mathbf{A}\|_{\ell_p} = (|A|_1^p + |A|_2^p + \ldots + |A|_n^p)^{\frac{1}{p}}. \tag{7}$$

The proposed concept was tested on polymer samples with varying polydispersity and finally applied to monitor the degradation of heparin using PGSE NMR.

In contrary to previous NMR diffusometry studies of this reaction[2], we monitored not only the change of diffusion coefficient but also polydispersity.

Additionally, we introduced the "moving-frame" processing, which is known in the field of non-uniform sampling[21]. The method is implemented by performing a series of PGSE experiments with randomly permuted gradient values (known as p-DOSY[3]) and combining them into one large dataset. The dataset is then divided into overlapping subsets processed separately with regularised ILT (see Figure 1). The method provides detailed information, as it allows to obtain in principle a continuous time-profile of the process[22].

2 Methods

Fig. 1 The idea of the "moving-frame" processing is applied to time-resolved PGSE data. FID signals with different amplitudes of the diffusion-encoding gradient are acquired, while certain processes are occurring in the sample that change the distribution of diffusion coefficients. Overlapping data subsets are then processed using inverse Laplace transform with tailored ℓ_p-norm regularisation.
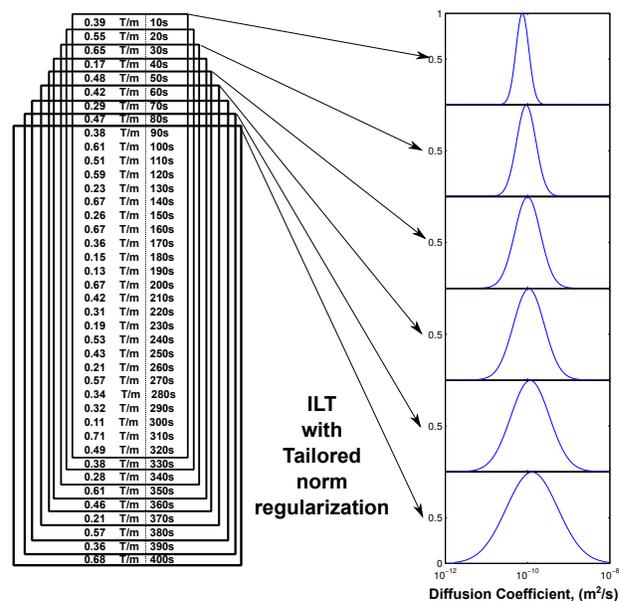
2.1 ℓ_p-norm regularization

The proposed method is using the regularisation term defined as

$$\Theta(A) = \|A\|_{\ell_p}, \tag{8}$$

where $1 \leq p \leq 2$ and thus the algorithm seeks for the following minimum

$$\min_{A \geq 0} \|\Phi A - \Psi\|_{\ell_2}^2 + \tau \|A\|_{\ell_p}. \tag{9}$$

For $p > 1$ the solution is "smoothened", i.e. the differences between the values of elements of a solution vector A are suppressed. In other words, the distributions of diffusion coefficients without significant "jumps" are preferred. The greater p is, the more pronounced is the smoothing effect. The minimized function presented in Equation (9) is composed of two terms: $\|\Phi A - \Psi\|_{\ell_2}^2$ and $\tau \|A\|_{\ell_p}$. The former term is ℓ_p-norm independent, while the latter changes with p. In case of $p > 1$ for any two vectors with the same mean value the smoother one (having smaller deviation from the mean value) has a smaller $\|A\|_{\ell_p}$. This general feature can be shown on the example of a two-element vector

Fig. 2 Comparison between the results of the optimal norm (dashed yellow line), ℓ_1-norm (dashed green line) and ℓ_2-norm (dashed red line) used in the processing of the simulated data set of various $A(D)$. The widths of the diffusion profiles and the corresponding optimal norms (σ/p) are: A - 0.5/1.6, B - 0.4/1.45, C - 0.3/1.3, D - 0.2/1.15, E - 0.1/1.1, F - 0.05/1.

Diffusion Coefficient, (m^2/s)
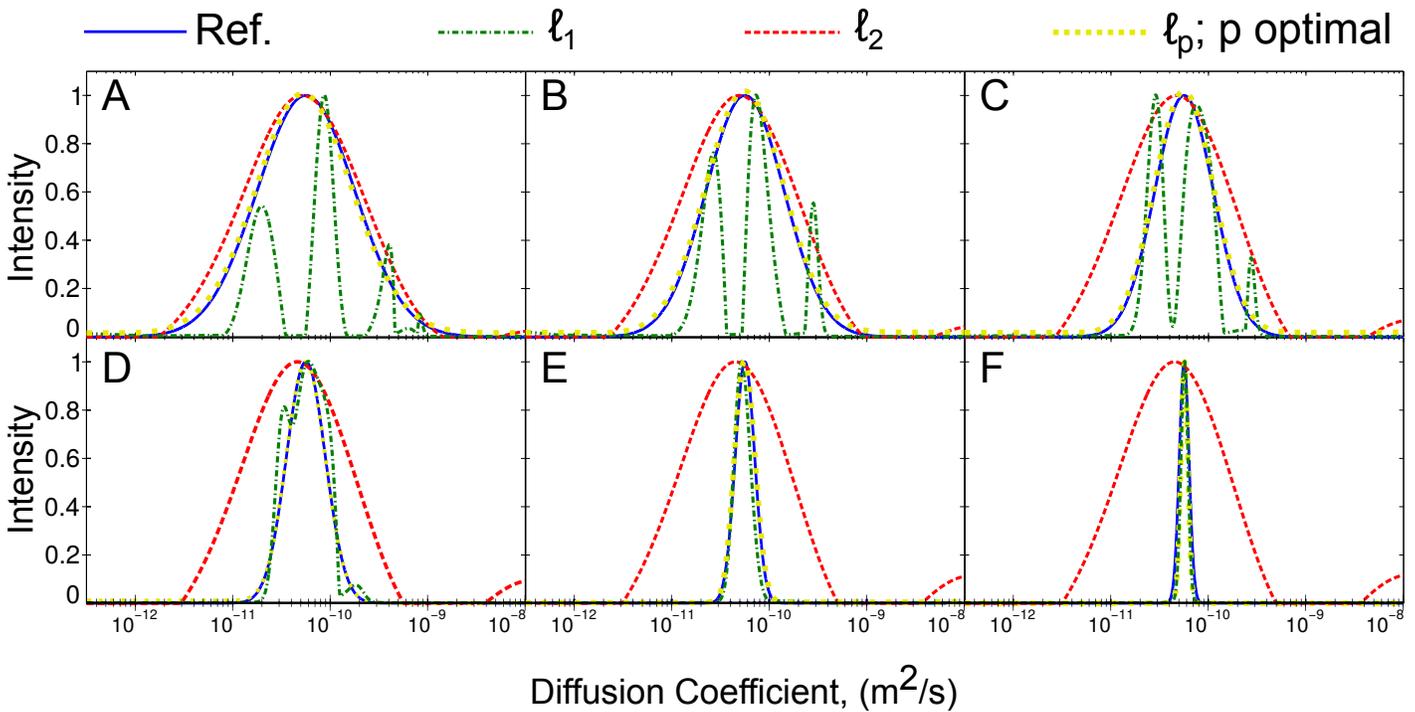
and ℓ_2-norm. Let $v_1 = [M, M]$ and $v_2 = [M - X, M + X]$, where M is the mean value of those vectors and X is a non-zero real number. Then it can be shown:

$$||v_2||_{\ell_2} = \sqrt{(M-X)^2 + (M+X)^2} = \tag{10}$$

$$= \sqrt{M^2 + X^2} > \sqrt{M^2} = ||v_1||_{\ell_2}.$$

Therefore $p = 2$ (and actually every $p > 1$) promotes the smoothened result. It can be shown in a similar way that for $p = 1$ the smoothing effect is not observed.

A more detailed explanation can be found in the Supplementary Information.

We have chosen the iteratively re-weighted least squares (IRLS) algorithm [23–25] as a method to implement the ILT regularised by ℓ_p-norm with arbitrary p. The description of the algorithm together with the Matlab code can be found in Supplementary Information.

2.2 Simulations

The method was tested on six different Gaussian distributions of diffusion coefficients (see Figure 2). Each distribution was generated using a logarithmically sampled diffusion coefficient grid with the centre of the peak at $\log_{10}(D) = -10.25$ for various widths σ. The distributions were converted to an exponential decay, composed of 64 logarithmically sampled points and 0.1% white noise (as in the following references [10,11,13]). Then, each simulated decay was processed using the IRLS method (1024 points, total computational time $\sim$12 s) and the ℓ_p-norm set to

Noise level (%)	σ	$\log_{10}(D)$ ($\frac{m^2}{s}$)
Ref	0.2	-10.25
0.001%	0.1999±0.03%	-10.250±0.0006%
0.005%	0.1998±0.16%	-10.249±0.0014%
0.01%	0.1998±0.45%	-10.2498± 0.0049%
0.05%	0.2020±3.54%	-10.250± 0.0300 %
0.1%	0.2072±8.60%	-10.249±0.0884%
0.2%	0.2128±11.46%	-10.248± 0.1009%
0.3%	0.2124±10.28%	-10.247±0.1463%
0.5%	0.2122±18.86%	-10.2466±0.1834%
1%	0.2117±22.72%	-10.2363±0.3309%

Table 1 Results of the reconstruction of the diffusion coefficient distribution from Figure 2.D for varying levels of random white noise. The mean and standard deviation of the results from 100 reconstructions for each noise level are shown.

$p = 1$, $p = 2$ and $p = 1, 1.05, 1.1 \ldots 2$ (20 values). The optimum p value was the one that provides A with minimal residuum of the fit

$$||\Phi A - \Psi||_{\ell_2}. \tag{11}$$

Thus, the method does not require any prior knowledge about the width of the distribution of diffusion coefficients. The residuum values for various p and σ values are shown in Figure 3. For $p = 1$, the approach is equivalent to the ITAMeD method, while $p = 2$ corresponds to the CONTIN method with **L** set to the identity matrix.

The robustness to noise was tested by repeating the simulation from Figure 3D 100 times for each of the nine white noise levels varying from 0.001% to 1% of the first data point. For each

simulation, the optimal p was estimated and the Gaussian curve was fitted to calculate the corresponding diffusion coefficient D and σ. Interestingly, variable p values compensate for wrong

Fig. 3 The residuum of the fit $\|\Phi A - \Psi\|_{\ell_2}$ as a function of p settings for the simulated data set shown in the Figure 2. The yellow dots correspond to the optimal p values.
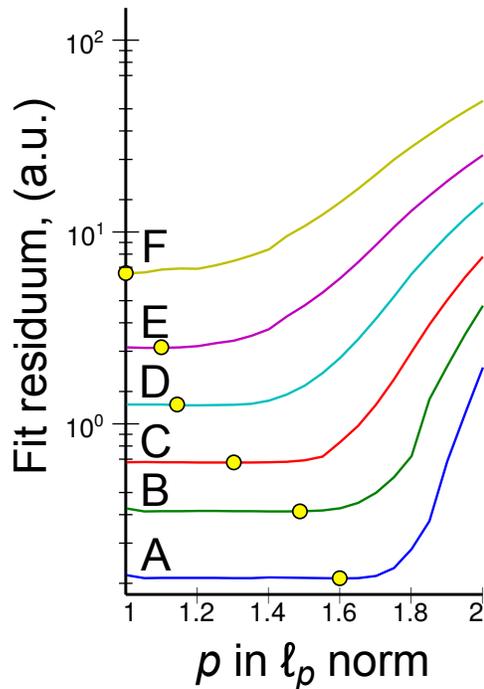
guesses of the manually entered parameter τ.

Furthermore, as shown in the Supplementary Information (Figure SI.1), the value of τ can be changed by several orders of magnitude without a significant difference in the result of reconstruction.

2.3 Test experiments on PEO polymers

Twelve polyethylene oxide polymers (PEOX21K, PEOX600K, PEOX900K, PEOX30K, PEOX150K, PEOX250K, PEOX50K, PEOX85K, PEOX90K, PEOX500K, PEOX120K and PEOX200K) with various polydispersity indexes (PDI) were supplied by American Polymer Standards Corporation. An appropriate amount of the polymer was dissolved in D_2O achieving a concentration of 0.1% w/w. The polymer solutions were run at 298 K on Bruker 600 MHz (Bruker, Germany) spectrometer equipped with a Diff30 diffusion probe and GREAT40 gradient amplifiers. The signal attenuation of the PEO peak at 3.6 ppm was obtained using a stimulated echo pulse sequence and 32 linearly spaced gradient amplitudes g, $\Delta = 100$ms, and $\delta = 2$ms. Each dataset was processed using (i) the IRLS method (1024 points of diffusion coefficient grid, $\tau = 10^{-6}$, total computational time $\sim$12 s) and the ℓ_p-norm set to $p = 1$, $p = 2$ and $p = 1, 1.05, 1.1 \ldots 2$ (20 values) and (ii) the log-normal fitting described in [19].

2.4 Monitoring the fractioning of heparin

12.3 mg of heparin sodium salt from porcine intestinal mucosa (Sigma-Aldrich) was dissolved in 1 ml of 90% D_2O, 7.0 pH phosphate buffer. Next, 0.3 mg of heparinase I and III Blend from Flavobacterium heparinum (Sigma Aldrich) was added. The reaction mixture was then transfered into an NMR tube, which was put into a 700 MHz Agilent spectrometer equipped with HCN probe temperature controlled at 25°C .

Each spectrum was obtained using a Bipolar Pulse Pair STE with Watergate and the signal was accumulated 64 times, g, $\Delta = 200$ms, and $\delta = 3$ms. The experiment was acquired using randomly shuffled sampling of the diffusion decay as in the p-DOSY experiment.[3] We used the list of 800 gradient values, which was constructed using 25 repetitions of a permuted 32 gradient array (The sampling schedule can be generated using the online interface at: http://itamed.spektrino.com). The experiments lasted in total for 59 hours.

The obtained heparin dataset was Fourier transformed and processed using nmrPipe[26] and imported into Matlab (MathWorks Inc.). The region of 3.20-3.66 ppm, which arises from protons of the sugar rings was integrated. Integration values were used as an input for the ILT with tailored norm regularisation. The data was processed using the "moving frame" method (see section 2.5). The size of the "frame" was set to 32, 64 and 128 points and diffusion coefficient grid to 256 points. Alse, we repeated the processing, using "standard" p-DOSY approach with series of 32-point subsets. Other processing parameters were the same as for the PEO samples. Each obtained diffusion profile was fitted to a Gaussian, whose centre corresponded to the mean diffusion coefficient and the width reflects the polydispersity. Both parameters were plotted as a function of time (Figure 6). Additionally, we analysed the peak intensity at $\sim$5.9 ppm, which corresponds to digested heparin fragments[2].

2.5 "Moving-frame" p-DOSY

Randomly shuffled sampling of the gradient domain, referred to as p-DOSY, has been recently reported as a good solution to study dynamically changing samples with PGSE NMR[3]. The use of p-DOSY allows to avoid bias in diffusion coefficient that could be caused by coherent variations of signal intensity due to reasons other than diffusion. We propose a slight modification of p-DOSY approach, conceptually similar to time-resolved non-uniform sampling, that has recently found numerous applications[21,22,27]. The series of experiments with differently permuted p-DOSY sampling schedules are performed and combined into one large dataset. Then, as shown in Figure 1, the dataset is divided into overlapping subsets, which are processed separately with ILT. The resulting stack of spectra forms temporal pseudo-dimension in whcih $A(D)$ changes. The size of the single subset is post-acquisition parameter, that has to compromise between signal-to-noise problems (small frames) and averaging of studied effects within the frame (large frames). If artifacts associated with particular schedule are observed, which can be the case of ILT or NUS reconstructions, then moving-frame processing has an advantage over standard "serial" p-DOSY. It provides time-profiles

237 with more points and thus imperfections can average out e.g. dur-
238 ing curve fitting. We exploit this feature in the analysis of kinetic parameters of heparin depolymerization.

Fig. 4 Optimal norm p as a function of the width of the diffusion coefficient distribution σ obtained for the simulated data set. The line fit with a coefficient of determination $R^2 = 0.982$ shows that the correlation is linear.
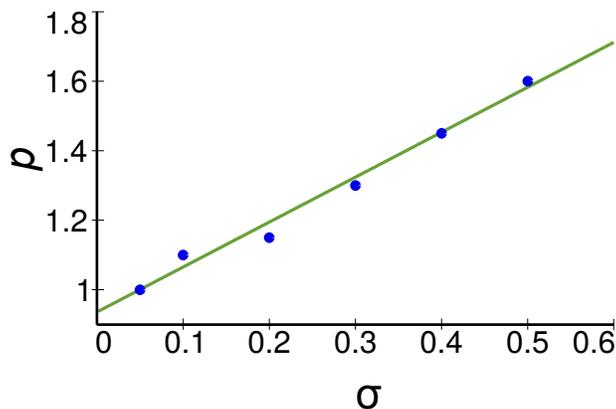
239
240 ## 3　Results and Discussion

241 The simulation (see Figure 2) showed that the optimal ℓ_p-norm
242 reconstructs the signal with a clearly better fidelity compared to
243 the over-sparsifing ℓ_1-norm or the over-smoothing ℓ_2-norm. The
244 over-sparsifing behaviour manifests itself in approximation of a
245 broad Gaussian peak made of a set of false narrow peaks. On the
246 other hand, over-smoothing leads to an artificial broadening of
247 the peak.

248 The optimal norm is found by repeating the simulation for dif-
249 ferent values of p and checking the residuals of the fit in the signal
250 domain (Ψ). Figure 3 shows that there is a wide range of similarly
251 behaving norms, especially for broad distributions of diffusion co-
252 efficients. In fact, the global minimum can be found (marked in
253 Figure 3), which provides the optimal p for all examples used in
254 this study.

255 It may happen, that for noisy signals the residual vs. p function
256 is not as smooth and there are many local minima. For this case,
257 one should choose p corresponding to right side of the "valley"
258 (Figure 3) as smoothed regularisations are less prone to instabili-
259 ties caused by noise.

260 Interestingly, the plots shown in Figure 3 and the way the op-
261 timal p is found resemble the L-curve approach used before for
262 the automatic setting of regularisation parameter.[28] Even very
263 recently, Scotti et. all.[29] demonstrated method for determining
264 polydispersity in light scattering based on CONTIN with this L-
265 curve criterion for finding τ. We have observed that both ℓ_p-norm
266 selection for constant τ and τ selection for constant ℓ_2-norm be-
267 have similarly for highly polydispersed samples. However, the
268 L-curve with constant ℓ_2-norm does not give a proper result for
269 samples with low polydispersity (See Supplementary Information
270 Fig.SI.2), in contrast to the tailored ℓ_p-norm. This is due to the
271 smoothing behaviour of ℓ_2-norm, which is explained in Section 4
272 of Supplementary Information. All the examples presented here
273 follow the model of an unimodal (although polydisperse) decay.

Fig. 5 Correlation between σ and $\sqrt{\log(\text{PDI})}$ for the PEOs. A - the reconstructions with the optimal ℓ_p-norm B - with log-normal fitting. The small case letters correspond to different polymer samples described in Table 2. The line fit $\sigma = \alpha\sqrt{\log(PDI)}$ gave the slope $\alpha = 0.56$ (A) and $\alpha = 0.53$ (B) with coefficients of determination: $R^2 = 0.948$ (A), $R^2 = 0.886$ (B).
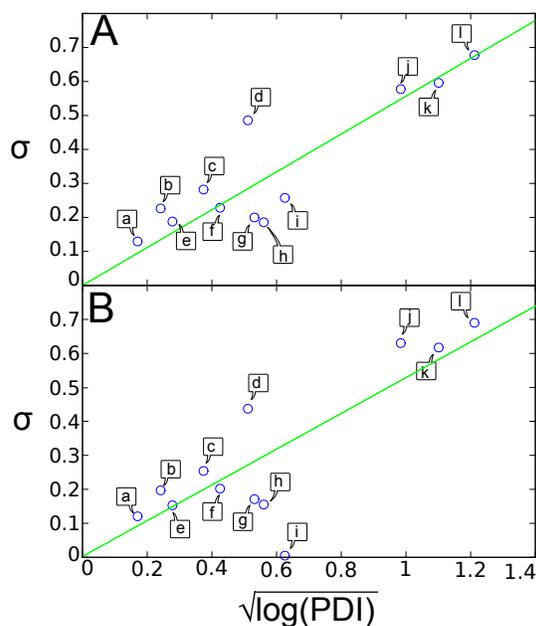
274 It is worth emphasising that the optimal norm can be found only
275 for signals where the diffusion peaks do not differ in polydisper-
276 sity. This is usually not the case for polymodal signals. However,
277 as shown in SI, the performance of the tailored norm is very simi-
278 lar to methods dedicated to deal with such problems, or performs
279 even better for noisy data.

280 The linear correlation between the optimal p and the actual
281 width of the Gaussian distribution of diffusion coefficients σ is
282 significant ($R^2 = 0.982$) and the dependence is shown in Figure 4.

283 Furthermore, Table 1 shows that the method is stable and not
284 very vulnerable to noise. In particular, the line widths are well
285 preserved even for quite high noise levels.

286 The results of extensive simulations allowed to test the effec-
287 tiveness of our method on a set of polydisperse samples. To ex-
288 perimentally verify the accuracy of the reconstruction with the
289 proposed tailored norm, we used polymer samples with varying
290 polydispersity. The polydispersity index (PDI) of the polymer can
291 be defined as $\text{PDI} = \frac{M_w}{M_n}$, where M_w is the weight-average, and M_n
292 is the number-average.[30] PDI values used for comparison were
293 taken from the certificate provided by the supplier of the poly-
294 mers and obtained using Gel Permeation Chromatography (GPC).
295 It can be compared with the σ value of the reconstructed distri-
296 bution. The value was obtained by fitting the log-normal curve to
297 the result of reconstruction

$$A(D, \mu, \sigma) = \frac{1}{D\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log(D) - \mu)^2}{2\sigma^2}\right). \quad (12)$$

298 As reported previously,[19] one correlates the PDI of a polymer with

Sample	Polymer	PDI	D from optimal p $\left(\frac{m^2}{s}\right)$	D log-normal $\left(\frac{m^2}{s}\right)$
a	PEOX21K	1.029	4.6E-11	4.57E-11
b	PEOX600K	1.060	5.53E-12	5.55E-12
c	PEOX900K	1.150	4.47E-12	4.51E-12
d	PEOX30K	1.299	5.26E-11	5.11E-11
e	PEOX150K	1.080	1.44E-11	1.43E-11
f	PEOX250K	1.198	1.03E-11	1.04E-11
g	PEOX50K	1.326	3.36E-11	3.37E-11
h	PEOX85K	1.369	2.40E-11	2.38E-11
i	PEOX90K	1.479	2.21E-11	2.24E-11
j	PEOX500K	2.632	1.15E-11	1.13E-11
k	PEOX120K	3.362	2.80E-11	2.79E-11
l	PEOX200K	4.340	2.21E-11	2.28E-11

Table 2 Description for Figure 5. D corresponds to the diffusion coefficient at the centre of the peak obtained from both methods.
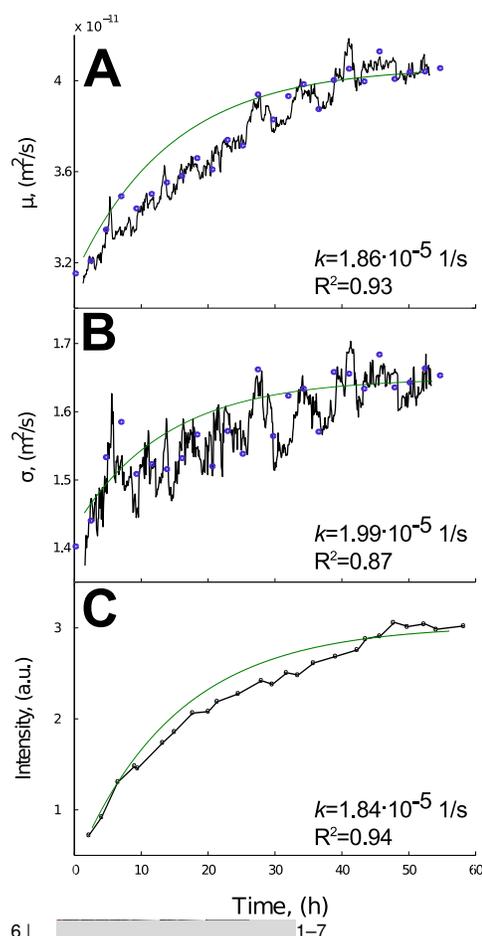
σ as follows

$$\text{PDI} = \exp\left(\frac{\sigma^2}{\alpha^2}\right), \qquad (13)$$

Fig. 6 Increase of the diffusion coefficient (A - μ) and the polydispersity (B- σ) following the enzymatic degradation reaction of bovine heparin by heparinase. Additionally, for comparison the intensity of the peak at 5.9 ppm is shown (C). Intensity was calculated as integration value of the peak at the lowest gradient strength (0.11 T/m.) The green lines are the fit of the first order reaction with reaction rate constant (k) and goodness of fit (R^2) written for each parameter. The line was fitted to "moving-frame" curve. Frame size of 64 points was used for A and B. Blue points show the result obtained from ILT processing based on p-DOSY processing.

where α is a scaling factor (typically between 0.5 and 0.6 for a good solvent).[31] Therefore, one would expect a linear correlation between σ and $\sqrt{\log(\text{PDI})}$

$$\sigma = \alpha\sqrt{\log(\text{PDI})} \qquad (14)$$

and the results, which are presented in Figure 5 confirm the linear correlation. The value obtained by a linear fit ($\alpha = 0.56$) is within the theoretical boundaries. The result of the well established method to evaluate the polydispersity of polymers, the log-normal distribution fitting[18,19] was compared with the tailored regularisation. As shown in Figure 5 the dependence of σ as a function of $\sqrt{\log \text{PDI}}$ deviates less for the optimal ℓ_p-norm compared to the log-normal fitting (R^2 = 0.948 for tailored vs. R^2 = 0.886 for log-normal).

Having established the robustness of the tailored norm regularisation for samples of different polydispersity, we were able to apply this technique to investigate the process of enzymatic degradation of heparin. Heparin, which is a long-chain polysaccharide with a inherently heterogeneous chain length, was fractionated with time by heparinase to form oligosacharides of different molecular weight i.e. chain length. As shown in Figure 6 both the diffusion coefficient and the width of the diffusion distribution are increasing during the enzymatic degradation as expected. Degradation products diffuse faster compared to heparin because of the lower molecular weight, which increases the observed mean diffusion coefficient. In addition, the average chain length of heparin is shortened and populations with different chain lengths are formed, which increases the polydispersity. The progress of the reaction can be independently monitored using the peak at 5.9 ppm corresponding to oligosaccharide fragments of digested heparin.

The heparin concentration is certainly below the K_D of heparinase, which is high in the absence of calcium.[32] Thus, we can assume a first order kinetic behaviour and regress the following equation $A(1 - exp(-k(t - t_0)))$ on the experimental data (where k is kinetic constant). The result is shown in Figure 6. A was assumed to be an average of the last 32 data points, while t_0 was back-extrapolated from the first 40 points of the curve. The obtained reaction rate constants are gathered in Table 3. It can be observed that all three time-profiles reveal similar kinetic

338 constants for a processing time frame size of 64 points ("moving 339 frame"), which gives also the largest R^2 values. For 32 points, the 340 decreased signal-to-noise ratio played role, while for 128 points 341 an averaging of the dynamics within the frame was observed. In 342 general, similar results were obtained but most important, the 343 frame size is a processing parameter and can be adjusted after 344 recording the experiment. It is noteworthy, that both goodness 345 of the fitting and obtained parameters for frame of 32 points, are 346 better than for standard ("serial") p-DOSY.

Table 3 Kinetic constants calculated for different frame sizes and spectral parameters.

Parameter	$k\left(\frac{1}{s}\right)$	R^2
Intensity of peak at 5.9 ppm	$1.84 \cdot 10^{-5}$	0.94
σ for frame size 32	$1.73 \cdot 10^{-5}$	0.87
σ for frame size 64	$1.99 \cdot 10^{-5}$	0.87
σ for frame size 128	$1.93 \cdot 10^{-5}$	0.80
σ from p-DOSY processing	$1.95 \cdot 10^{-5}$	0.7
μ for frame size 32	$2.03 \cdot 10^{-5}$	0.89
μ for frame size 64	$1.86 \cdot 10^{-5}$	0.93
μ for frame size 128	$2.21 \cdot 10^{-5}$	0.87
μ from p-DOSY processing	$2.06 \cdot 10^{-5}$	0.87

4   Conclusions

347
348 We have discussed the choice of ℓ_p-norm ($1 \leq p \leq 2$) as a regu- 349 larization for the inverse Laplace transform applied to diffusion 350 NMR spectroscopy. The iteratively re-weighted least squares al- 351 gorithm allowed us to implement ILT with an arbitrarily chosen 352 regularisation term. Both simulations and experiments showed 353 that proper reconstructions are obtained, when p is balanced be- 354 tween the sparsifying ($p = 1$) and smoothing ($p = 2$) variant. The 355 optimum can be found automatically by minimising the norm of 356 the residual with respect to p. The proposed method is tailored 357 to samples with various diffusion profiles and thus is an impor- 358 tant extension of previously introduced approach based on a plain 359 sparsity restraint. Additionally, we proved that ℓ_p-norm can be 360 used for monitoring reactions in situ where the change in poly- 361 dispersity plays a crucial role for a mechanistic explanation. The 362 moving-frame variant of p-DOSY method allowed to obtain time- 363 resolved data of high accuracy.

Acknowledgement

Supplementary Information

373 Algorithm description, Matlab Code of the IRLS algorithm , test 374 of the robustness to mis-setting of τ, behaviour of ℓ_2-norm regu- 375 larisation with different τ values,comparison with TRAIn method 376 for assymetric, bimodal distribution and heparin depolymeriza- 377 tion, detailed explanation of smoothing features of ℓ_p-norm and 378 ^{1}H spectra of heparin and PEO are shown.

References

1 F. Susanne, D. S. Smith and A. Codina, *Organic Process Research & Development*, 2012, **16**, 61–64.
2 J. F. K. Limtiaco, S. Beni, C. J. Jones, D. J. Langeslay and C. K. Larive, *Anal. Bioanal. Chem.*, 2011, **399**, 593–603.
3 M. Oikonomou, J. Asencio-Hernández, A. H. Velders and M.-A. Delsuc, *J. Magn. Reson.*, 2015, **258**, 12–16.
4 E. O. Stejskal and J. E. Tanner, *J. Chem. Phys.*, 1965, **42**, 288.
5 W. S. Price, *NMR Studies of Translational Motion: Principles and Applications*, Cambridge University Press, Leiden, 2009.
6 P. T. Callaghan, *Translational dynamics and magnetic resonance: principles of pulsed gradient spin echo NMR*, Oxford University Press, 2011.
7 C. Johnson Jr, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1999, **34**, 203–256.
8 Van Gorkom LCM and T. Hancewicz, *J. Magn. Reson.*, 1998, **130**, 125–30.
9 Nilsson, Mathias and Morris, Gareth A., *Anal. Chem.*, 2008, **80**, 3777–3782.
10 K. Xu and S. Zhang, *Anal. Chem.*, 2014, **86**, 592–599.
11 M. a. Delsuc and T. E. Malliavin, *Anal. Chem.*, 1998, **70**, 2146–2148.
12 S. Provencher, *Comput. Phys. Commun.*, 1982, **27**, 229–242.
13 M. Urbańczyk, D. Bernin, W. Koźmiński and K. Kazimierczuk, *Anal. Chem.*, 2013, **85**, 1828–1833.
14 I. Toumi, B. Torrésani and S. Caldarelli, *Anal. Chem.*, 2013, **85**, 11344–11351.
15 X. Zhou, K. Xu and S. Zhang, *J. Magn. Reson.*, 2015, **252**, 114–119.
16 W. Windig, B. Antalek, L. Sorriero, S. Bijlsma, D. Louwerse and A. Smilde, *Journal of Chemometrics*, 1999, **13**, 95–110.
17 Nilsson, Mathias and Connell, Mark A. and Davis, Adrian L. and Morris, Gareth A., *Anal. Chem.*, 2006, **78**, 3040–3045.
18 P. T. Callaghan and D. N. Pinder, *Macromolecules*, 1985, **18**, 373–379.
19 M. Röding, D. Bernin, J. Jonasson, A. Särkkä, D. Topgaard, M. Rudemo and M. Nydén, *J. Magn. Reson.*, 2012, **222**, 105–11.
20 I. J. Day, *J. Magn. Reson.*, 2011, **211**, 178–85.
21 M. Mayzel, J. Rosenlöw, L. Isaksson and V. Y. Orekhov, *J. Biomol. NMR*, 2014, **58**, 129–139.
22 R. Dass, W. Koźmiński and K. Kazimierczuk, *Anal. Chem.*, 2015, **87**, 1337–43.
23 E. J. Candes and M. B. Wakin, *IEEE Signal Process. Mag.*, 2008, **25**, 21–30.
24 I. Daubechies, R. Devore, M. Fornasier and S. Güntürk, *Commun. Pure Appl. Math.*, 2010, **LXIII**, 1–38.
25 K. Kazimierczuk and V. Y. Orekhov, *Angew Chem Int Ed Engl*, 2011, **50**, 5556–5559.
26 F. Delaglio, S. Grzesiek, G. Vuister, G. Zhu, J. Pfeifer and A. Bax, *Journal of Biomolecular NMR*, 1995, **6**, 277–293.
27 W. Bermel, R. Dass, K.-P. Neidig and K. Kazimierczuk, *ChemPhysChem*, 2014, **15**, 2217–2220.
28 P. C. Hansen, *SIAM Review*, 1992, **34**, 561–580.
29 A. Scotti, W. Liu, J. S. Hyatt, E. S. Herman, H. S. Choi, J. W. Kim, L. A. Lyon, U. Gasser and A. Fernandez-Nieves, *J. Chem. Phys.*, 2015, **142**, 234905.
30 D. F. Evans and H. Wennerström, *The colloidal domain: where physics, chemistry, biology, and technology meet*, Wiley-Vch New York, 1999.
31 H. Walderhaug, O. Söderman and D. Topgaard, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **56**, 406–25.
32 R. Sasisekharan, G. Venkataraman, R. Godavarti, S. Ernst, C. L. Cooney and R. Langer, *Journal of Biological Chemistry*, 1996, **271**, 3124–3131.