Support Information for:
PALMA, an improved algorithm for the DOSY signal processing

Afef Cherni[a,b], Emilie Chouzenoux[b], Marc-André Delsuc[*,a]

Contents

a: Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), IN-SERM U596, CNRS UMR 7104 Université de Strasbourg, 67404 Illkirch-Graffenstaden, France.

b: Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE Paris, UPEM, Marne-la-Vallée, France.

* madelsuc@unistra.fr

1

1 Variational formulation

A well-known and very efficient strategy for solving ill-posed inverse problems is to adopt a penalization approach that provides an estimate $\hat{X} \in \mathbb{R}^N$ of the original signal $X \in R^N$, that is the solution of the constrained minimization problem:

$$\underset{X \in \mathbb{R}^N}{\text{minimize}} \ \Psi(X) \quad \text{subject to} \quad \|\boldsymbol{H}X - Y\| \leq \eta, \tag{1}$$

where Ψ is the regularization function and η is an estimate of the experimental noise. In this work, we choose to define Ψ for every $X \in \mathbb{R}^N$ and $\lambda \in [0,1]$, as:

$$\Psi(X) = \lambda \, \text{ent}(X, a) + (1 - \lambda)\ell_1(X), \tag{2}$$

where the Shannon entropy with a flat prior $a > 0$ and the ℓ_1 norm are defined, respectively, as:

$$(\forall X = (x_n)_{1 \leq n \leq N} \in \mathbb{R}^N) \quad \text{ent}(X, a) = \begin{cases} \sum_{n=1}^{N} \frac{x_n}{a} \log\left(\frac{x_n}{a}\right) & \text{if } x_n > 0 \\ 0 & \text{if } x_n = 0 \\ +\infty & \text{elsewhere} \end{cases}$$

and

$$(\forall X = (x_n)_{1 \leq n \leq N} \in \mathbb{R}^N) \quad \ell_1(X) = \sum_{n=1}^{N} |x_n|.$$

2 Proximity operator

To solve Problem (1), we propose to rely on a proximal optimization method, that makes use of the so-called proximity operator. Since Ψ is convex on $\mathbb{R}^N$, its proximity operator at a given point $X \in \mathbb{R}^N$ is defined as the unique minimizer of $\frac{1}{2}\|\cdot - X\|^2 + \Psi$. Since Ψ in (2) takes a separable form, i.e. it can be written as:

$$(\forall X = (x_n)_{1 \leq n \leq N} \in \mathbb{R}^N) \quad \Psi(X) = \sum_{n=1}^{N} \psi(x_n), \tag{3}$$

its proximity operator is given by:

$$\text{prox}_{\Psi}(x) = (p(x_n))_{1 \leq n \leq N}. \tag{4}$$

Hereabove, for every $n \in \{1, \ldots, N\}$, $p(x_n) \in \mathbb{R}$ is the unique minimizer of:

$$\phi : u \mapsto \frac{1}{2}(u - x_n)^2 + \psi(u). \tag{5}$$

Let us first consider the case when $\lambda \in]0, 1]$. Then,

$$\dot{\phi}(u) = 0 \Leftrightarrow u - x_n + \frac{\lambda}{a} \log(u) + \frac{\lambda}{a} - \frac{\lambda}{a} \log(a) + 1 - \lambda = 0$$

$$\Leftrightarrow \log(u) = \frac{a}{\lambda} \left(-u + x - \frac{\lambda}{a} + \frac{\lambda}{a} \log(a) - 1 + \lambda \right)$$

$$\Leftrightarrow u = \exp \left(-\frac{a}{\lambda} u + \frac{ax - a(1-\lambda)}{\lambda} + \log(a) - 1 \right).$$

Finally:

$$u = \frac{\lambda}{a} \mathcal{W} \left(\frac{a}{\lambda} \exp \left(\frac{ax - a(1-\lambda)}{\lambda} + \log(a) - 1 \right) \right). \tag{6}$$

In the above expression, $\mathcal{W}$ states for the Lambert function, also called Omega, defined as the inverse function of $f : x \to x \exp(x)$ for all $x \in \mathbb{C}$:

$$z = x \exp(x) \Leftrightarrow x = \mathcal{W}(z).$$

If $\lambda = 0$, we have:

$$x_n - u \in \partial |u| \Longleftrightarrow x_n - u \in \begin{cases} \operatorname{sign}(u) & \text{if} \quad u \neq 0 \\ [-1, 1] & \text{elsewhere} \end{cases}$$

$$\Longleftrightarrow u = \begin{cases} x_n - 1 & \text{if} \quad u > 1 \\ 0 & \text{if} \quad u \in [-1, 1] \\ x_n + 1 & \text{elsewhere.} \end{cases}$$

Finally, we can conclude that $(\forall X = (x_n)_{1 \leq n \leq N} \in R^N)$, $\operatorname{prox}_\Psi(X)$ is given by (4), where, for all $n \in \{1, \dots, N\}$,

$$p(x_n) = \begin{cases} \frac{\lambda}{a} \mathcal{W} \left(\frac{a}{\lambda} \exp \left(\frac{ax_n - a(1-\lambda)}{\lambda} + \log(a) - 1 \right) \right) & \text{if} \quad \lambda \in]0, 1] \\ \operatorname{sign}(x_n) \left(\max(|x_n| - 1, 0) \right) & \text{if} \quad \lambda = 0. \end{cases} \tag{7}$$

3   Asymptotic development

Let us emphasize that the Lambert function has many interesting properties. In particular:

$$\mathcal{W}(e^x) \xrightarrow[x \to +\infty]{} x \left(1 - \frac{\log(x)}{1+x} \right) \tag{8}$$

This asymptotic development is of main interest in our context. Indeed, when $\lambda \in]0, 1]$, for all $n \in \{1, \dots, N\}$,

$$p(x_n) = \frac{\lambda}{a} \mathcal{W}(\exp(c_n)), \tag{9}$$

with $c_n = \frac{ax_n - a(1-\lambda)}{\lambda} - 1 + 2\log(a) - \log(\lambda)$. When c_n is large (typically $c_n > 10^2$), standard numerical implementation of the Lambert function yields infinity as an output.

We thus propose to use, for large c_n, the following approximation of the proximity operator, which is a consequence of the asymptotic development (8):

$$p(x_n) = \frac{\lambda}{a} \left(c_n - \frac{c_n}{1 + c_n} \log(c_n) \right).$$ (10)

Figure 8 illustrates relative error of the approximation, when $\lambda = a = 1$.
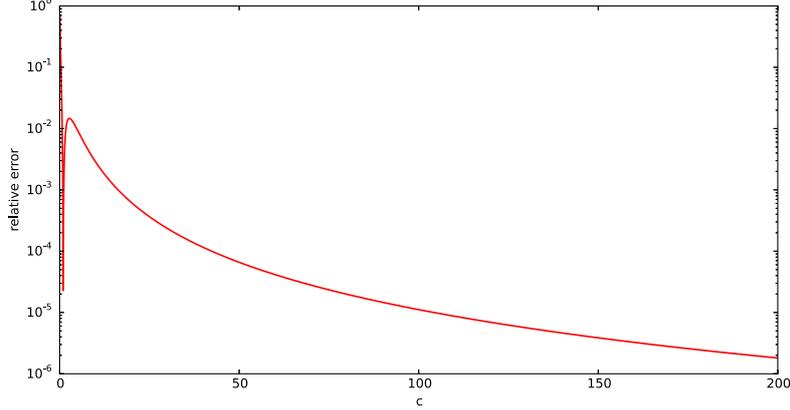

Figure S1: Relative error value between $\mathcal{W}(e^{c_n})$ and the approximation of $p(x_n)$ in 9. The relative error becomes lower than 10^{-3} for $c_n > 20$.

4 PALMA algorithm

Using the PPXA+ algorithm, we propose a new proximal algorithm to solve (1) that makes use of the proximal formula (7). The so-called PALMA algorithm, standing for "**P**roximal **A**lgorithm for **L**$_1$ combined with **MA**xent prior", is given below:

Initialization
$V^{(0,1)} \in \mathbb{R}^N, V^{(0,2)} \in \mathbb{R}^M$
$X^{(0)} = (\mathrm{I}_N + \boldsymbol{H}^\top \boldsymbol{H})^{-1}(V^{(0,1)} + \boldsymbol{H}^\top V^{(0,2)})$
$\boldsymbol{B} = (\mathrm{I}_N + \boldsymbol{H}^\top \boldsymbol{H})^{-1}$
$\gamma \in (0, 2),$
Minimization
For $k = 0, 1, \dots$
$\quad\begin{aligned} &Z^{(k,1)} = \mathrm{prox}_\Psi(V^{(k,1)}) \\ &Z^{(k,2)} = \mathrm{proj}_{\|\cdot - Y\| \le \eta}\left(V^{(k,2)}\right) \\ &U^{(k)} = \boldsymbol{B}(Z^{(k,1)} + \boldsymbol{H}^\top Z^{(k,2)}) \\ &X^{(k+1)} = X^{(k)} + \gamma(U^{(k)} - X^{(k)}) \\ &V^{(k+1,1)} = V^{(k,1)} + \gamma\left(2U^{(k)} - X^{(k)} - Z^{(k,1)}\right) \\ &V^{(k+1,2)} = V^{(k,2)} + \gamma\left(\boldsymbol{H}(2U^{(k)} - X^{(k)}) - Z^{(k,2)}\right) \end{aligned}$

4

where I_N is the identity matrix of $\mathbb{R}^N$ and the projection $\text{proj}_{\|\cdot - Y\| \leq \eta}$ is defined as follows:

$$\text{proj}_{\|\cdot - Y\| \leq \eta}(Z) = Z + (Z - Y)\min\left(\frac{\eta}{\|Z - Y\|}, 1\right) - Y \quad (\forall (Y, Z) \in (\mathbb{R}^N)^2).$$

5 Choice of processing parameters

As it is detailed in the previous theoretical section, the PALMA algorithm is controlled by several scaling parameters which have to be adapted to the current problem.

5.1 Choice of a

The entropy approach requires an expression of the prior knowledge on the system, which in the present case corresponds to *a priori* spectrum in the absence of experimental evidences. A general expression is thus a flat spectrum of intensity a, the value of which has to be adapted to the scaling of the experimental measurement. A natural approach consists in estimating the area under the signal expected signal: $\sum_n x_n$ and scaling the prior with this value. Due to the properties of the Laplace transform, we have chosen here $a = y_0 \approx \sum_n x_n$.

5.2 Choice of η

If we assume that the experimental values are tainted by an additive random noise: $y_m = \hat{y}_m + \varepsilon_m$, and if this noise is supposed to be centered and Gaussian i.i.d. with variance σ^2, then we can expect the residual of the fit to be $\eta \approx \sigma\sqrt{M}$. The value of σ can be measured here either from additional measures or from a simple polynomial fit of the y_m curve as it is done in the current implementation. The case of a correlated Gaussian noise ε, with covariance matrix Σ, is encompassed by our method. Indeed, the least squares constraint becomes:

$$(\forall X \in \mathbb{R}^N) \quad \left((\boldsymbol{H}X - Y)^\top \Sigma^{-1}(\boldsymbol{H}X - Y)\right)^{1/2} \leq \eta,$$

which is equivalent to $\|\boldsymbol{H}X - Y\|_2 \leq \eta$, up to any change of variable with the form $\boldsymbol{H} \longleftarrow \theta\Sigma^{-1/2}\boldsymbol{H}$ and $Y \longleftarrow \theta\Sigma^{-1/2}Y$, $\eta \longleftarrow \theta\eta$, for some $\theta > 0$. In order to facilitate the choice of η, we propose to take $\theta = \sigma_{\max}$, i.e. the maximal singular value of Σ, so that a suitable choice for η is $\sigma_{\max}\sqrt{M}$.

6 Choice of λ

In order to evaluate the influence of the λ factor on the balance between sparsity and the amount of information in the processed signal, we launch our PALMA algorithm for different signal of monodisperse and polysdisperse species (A, B, C1 and C2) and we trace on Figures S2 to S6, the recovered signal by varying λ from 0 to 1.

The associated reconstruction quality, measured in terms of signal to noise ratio (SNR) $= 10 \log 10 \left(\frac{\|X\|}{\|\hat{X} - X\|} \right)$ of each tested case, is reported in Table S1. **Larger value of quality of reconstruction correspondents to the best reconstruction.**

• **Signal A**

Signal A consists in three monodisperse components with diffusion coefficients $16 \, \mu m^2/s$, $63 \, \mu m^2/s$, and $230 \, \mu m^2/s$, with respective intensities $1.0, 0.33$ and 0.66. This data-set is equivalent to the simulation used in *Kazimierczuk et al.*

• **Signal B**

Signal B is a wide monodisperse distribution, simulated as a symmetric log-normal distribution centered at $35 \, \mu m^2/s$ and with variance 25.

• **Signals C1 & C2**

C1 and C2 signals are asymmetric distributions built from 15 log-normal components, ranging from 18 to $85 \, \mu m^2/s$, with intensities ranging from 0.1 to $10 \, s$. They have PDI estimated respectively to 1.79 and 1.32.

For C1, the intensities are chosen as $10.0/(1.4^i), i = 1, \ldots, 15$, and for C2, as $10.0/(1.4^{15-i}), i = 1, \ldots, 15$.

In all simulations, an additive zero-mean white Gaussian noise and standard deviation σ equals 0.1% of the initial point of $\boldsymbol{H}X$ was added to the observed data. The number of measurement is $M = 64$ and the original signal has the dimension $N = 256$.
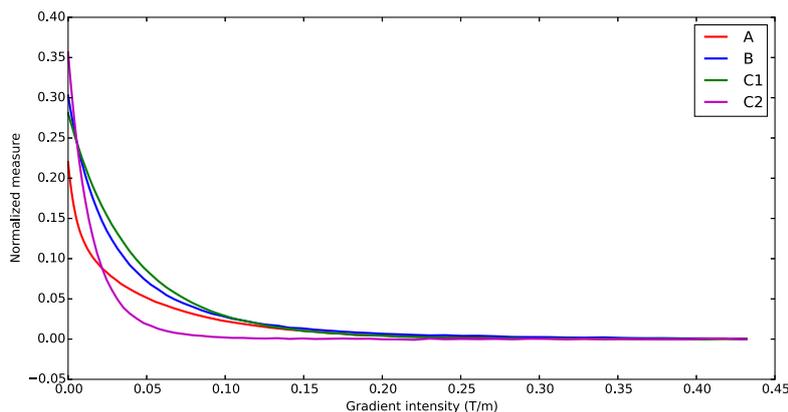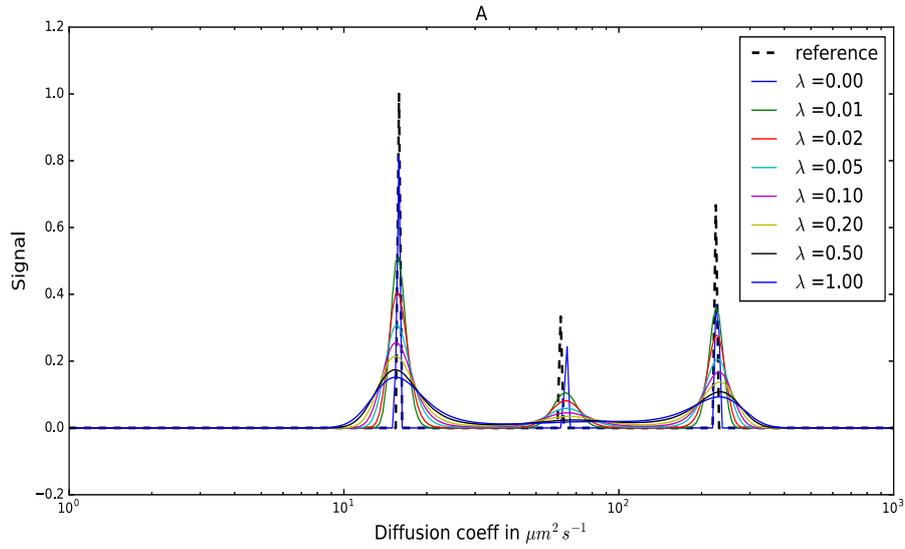

Figure S2: Measurement of A, B, C1 and C2 signals

6

Figure S3: Reconstruction of signal A using PALMA with different λ. The minimal error is obtained for $\lambda = 0$ (Spectra obtained with $\lambda \neq 0$ have their intensities multiplied by 3 for clarity).
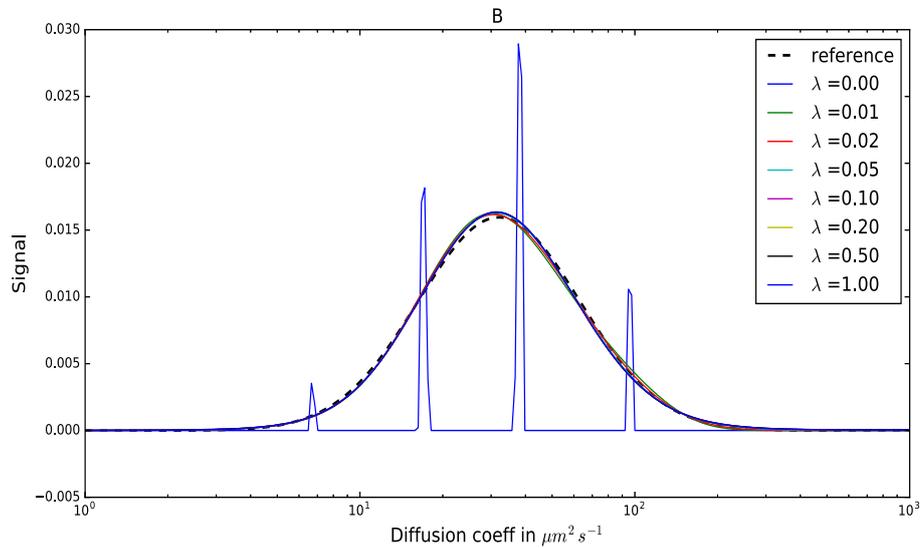

Figure S4: Reconstruction of signal B using PALMA with different λ. The minimal error is obtained for $\lambda = 0.05$ (Spectra obtained with $\lambda = 0$ have their intensities divided by 8 for clarity).
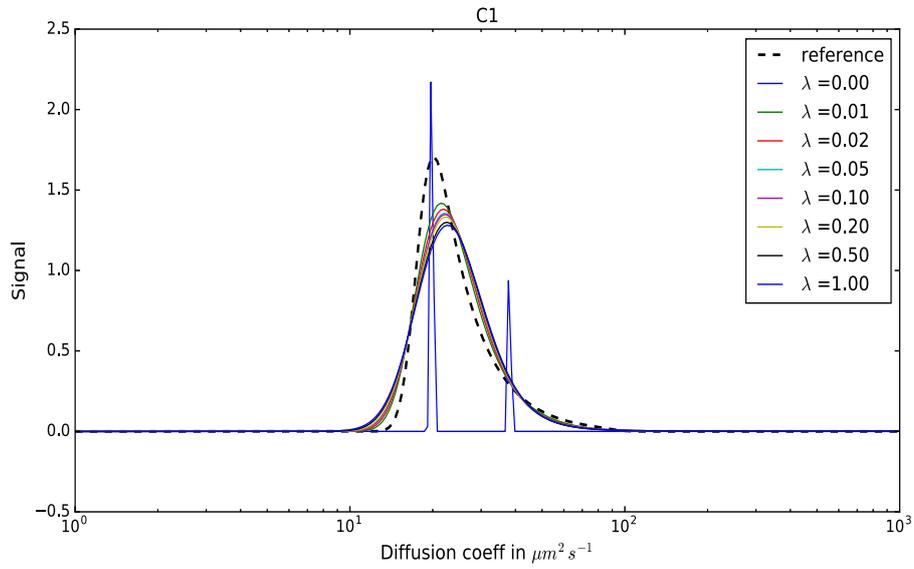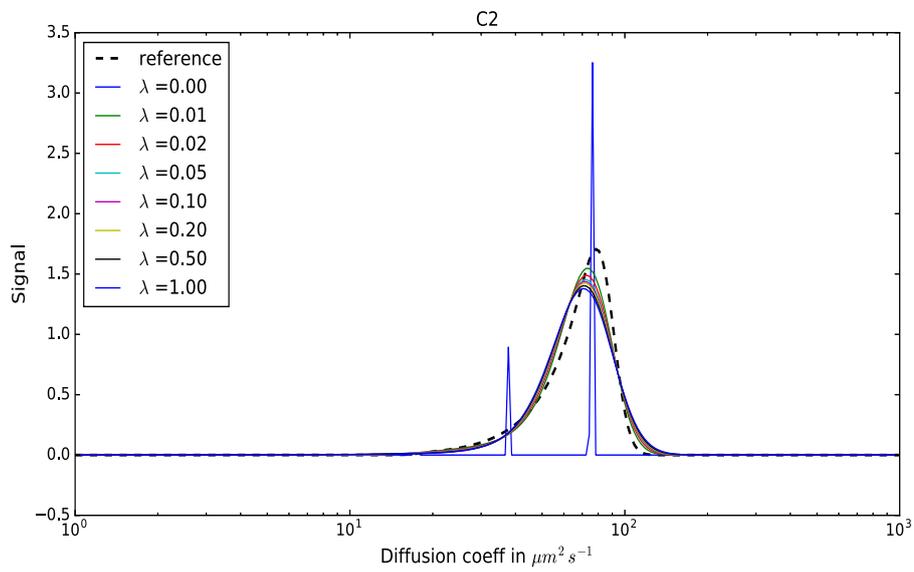
Figure S5: Reconstruction of signal C1 using PALMA with different λ. The minimal error is obtained for $\lambda = 0.01$ (Spectra obtained with $\lambda = 0$ have their intensities divided by 8 for clarity).

Figure S6: Reconstruction of signal C2 using PALMA with different λ. The minimal error is obtained for $\lambda = 0.01$ (Spectra obtained with $\lambda = 0$ have their intensities divided by 8 for clarity).

8

λ	Qlty reconstruction of Signal			
	A	B	C1	C2
0.0	5.65	-11.04	-9.87	-12.58
0.01	1.05	28.57	15.21	16.15
0.02	0.80	31.87	13.76	14.89
0.05	0.58	32.51	12.92	14.06
0.1	0.47	31.11	12.71	13.66
0.2	0.39	31.17	12.37	13.35
0.5	0.30	31.44	11.65	12.91
1	0.26	31.55	11.21	12.44

Table S1: Quality of reconstruction of A, B, C1 and C2 signals for various λ values.

7   Comparison with state-of-the-art algorithms

Several algorithms have been developed to solve the ill-posed problem in DOSY experience. In Table S2, we present the comparison results between our approach and several recent algorithms, namely ITAMeD, ITAMeD with ℓ_p, and TRAIN.

As an illustration, we present in Figures S7 to S10, the reconstruction of B ans C2 signals with different algorithms for 4 different noise levels.

The following settings have been used, which lead to the best performance in terms of both reconstruction quality and computational cost:

PALMA: $\lambda = 0.01$.

ITAMED : Regularization parameter $= 10^{-6}$.

ITAMeD with ℓ_p: Smoothing parameter $\tau = 10^{-7}$, ration between 1st and 2nd term $\epsilon = 10$.

TRAIn: $\tau = 1.02$

Signal	Noise level / Algorithm	Qlty reconstruction in dB			
		1%	0.1%	0.01%	0.001%
B	ITAMeD	3.37	18.65	29.04	29.40
	ITAMeD with ℓ_p	6.06	25.26	36.69	37.08
	TRAIn	24.75	28.63	26.53	19.47
	PALMA with $\lambda = 0.01$	20.54	28.57	41.69	53.25
	PALMA with $\lambda = 0.05$	24.01	32.51	48.28	51.37
C2	ITAMeD	1.57	15.13	14.36	14.06
	ITAMeD with ℓ_p	3.37	6.30	6.29	6.29
	TRAIn	8.62	15.39	23.36	25.5
	PALMA with $\lambda = 0.01$	10.6	12.72	17.72	23.24
	PALMA with $\lambda = 0.05$	7.62	10.97	16.59	20.75

Table S2: Quality of reconstruction of signals B and C2 with different algorithms for various noise levels. Here, $M = 64$

M	Noise level / Algorithm	Qlty reconstruction in dB			
		1%	0.1%	0.01%	0.001%
64	ITAMeD	1.57	15.13	14.36	14.06
	ITAMeD with ℓ_p	3.37	6.30	6.29	6.29
	TRAIn	8.62	15.39	23.36	25.5
	PALMA with $\lambda = 0.01$	10.6	12.72	17.72	23.24
	PALMA with $\lambda = 0.05$	7.62	10.97	16.59	20.75
32	ITAMeD	-3.72	8.89	12.89	13.91
	ITAMeD with ℓ_p	0.57	6.18	6.27	6.27
	TRAIn	3.68	6.60	9.396	18.04
	PALMA with $\lambda = 0.01$	11.09	14.60	19.48	23.04
	PALMA with $\lambda = 0.05$	8.69	13.09	18.92	20.56

Table S3: Quality of reconstruction of signal C2 for various noise levels and problem sizes.
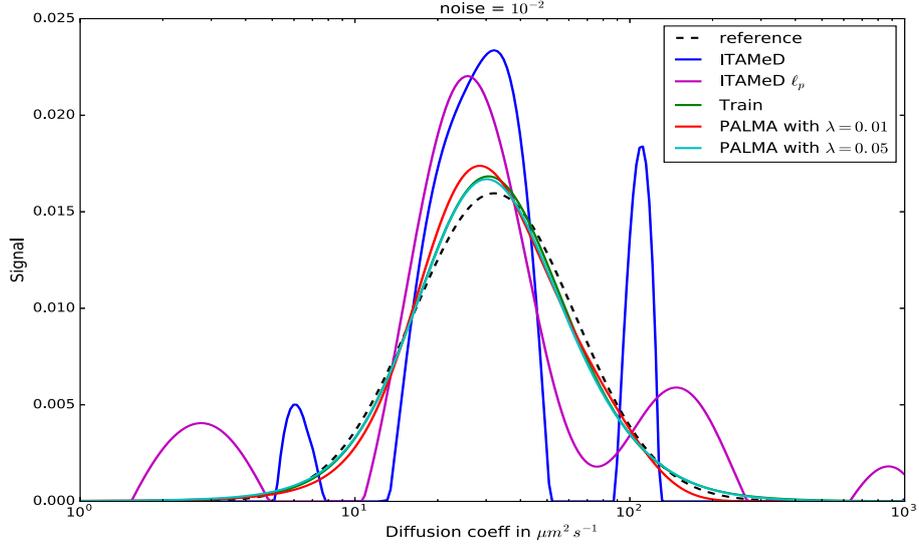
Figure S7: Reconstruction of signal B, using different algorithms. Here, $M = 64$ and noise level = 0.001%.
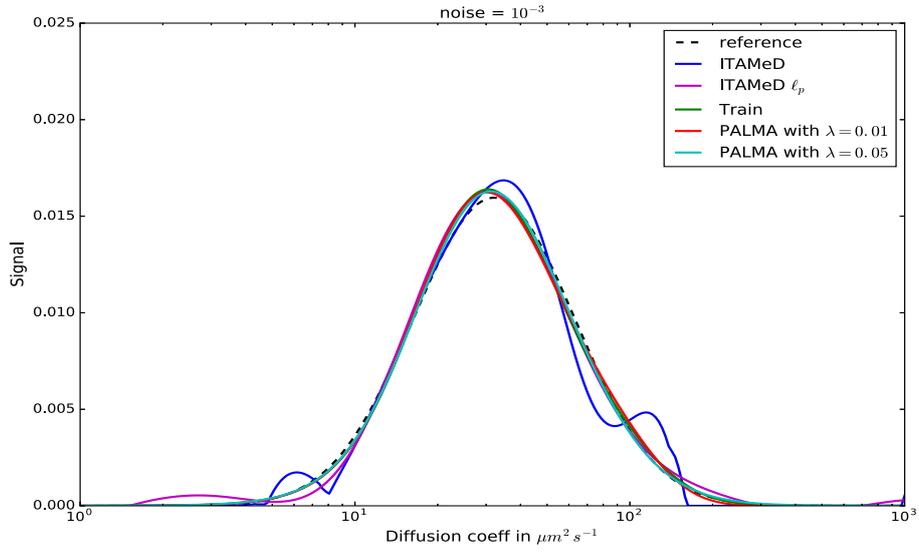
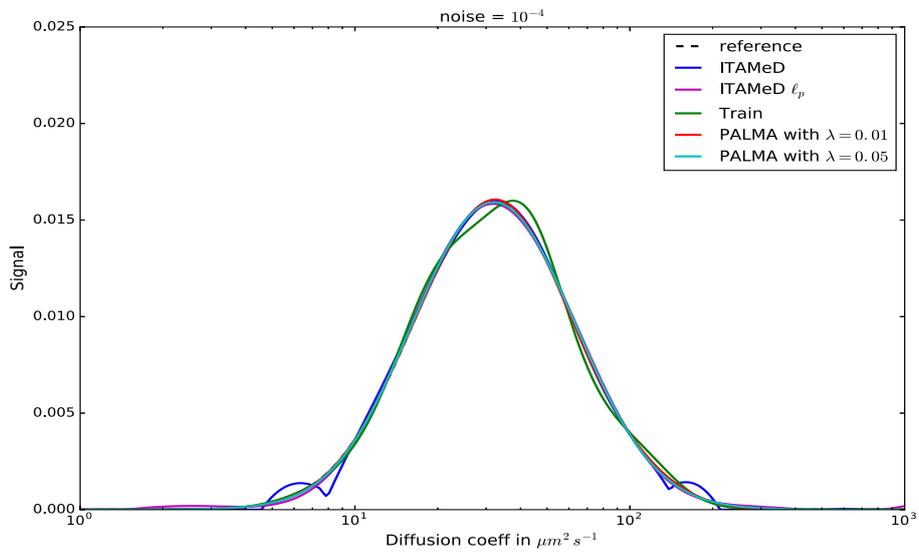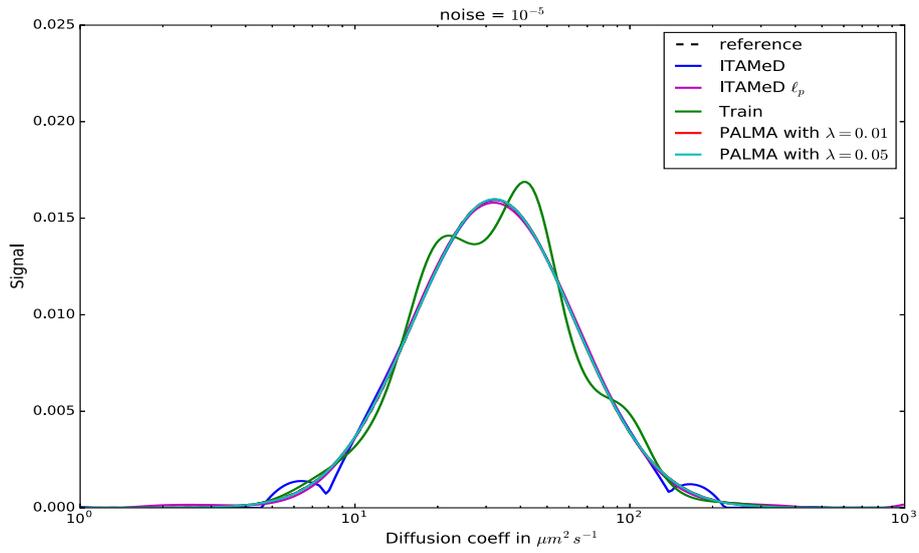Figure S8: Reconstruction of signal B, using different algorithms. Here, $M = 64$ and noise level $= 0.001\%$.
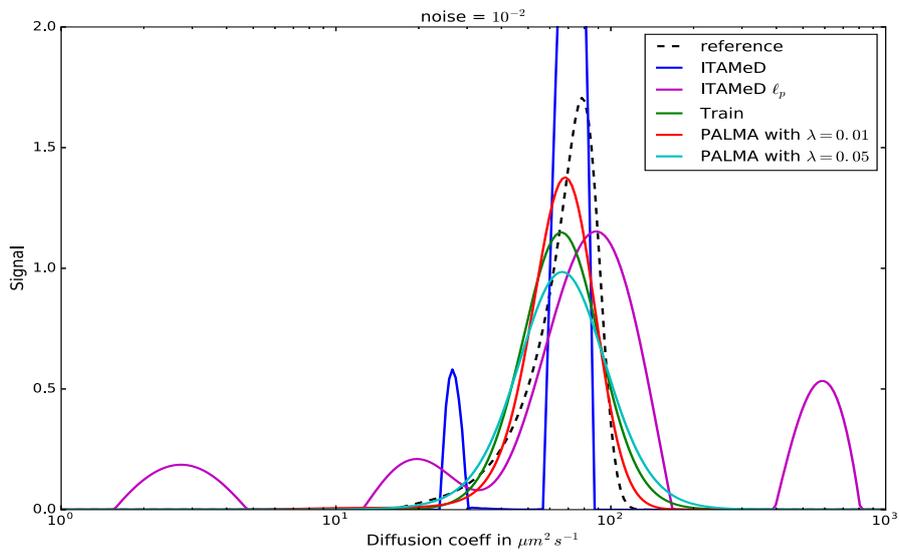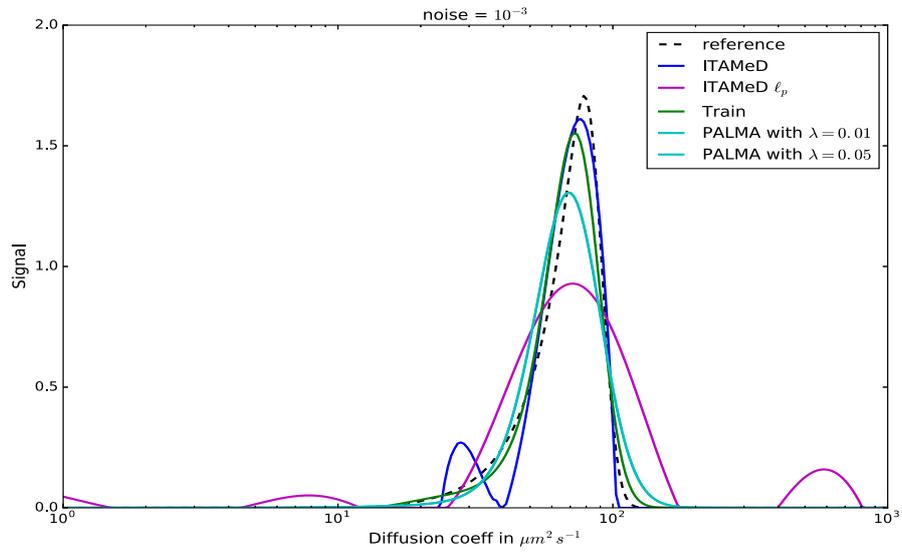

Figure S9: Reconstruction of signal B, using different algorithms. Here, $M = 64$ and noise level $= 0.001\%$.

Figure S10: Reconstruction of signal B, using different algorithms. Here, $M = 64$ and noise level $= 0.001\%$.
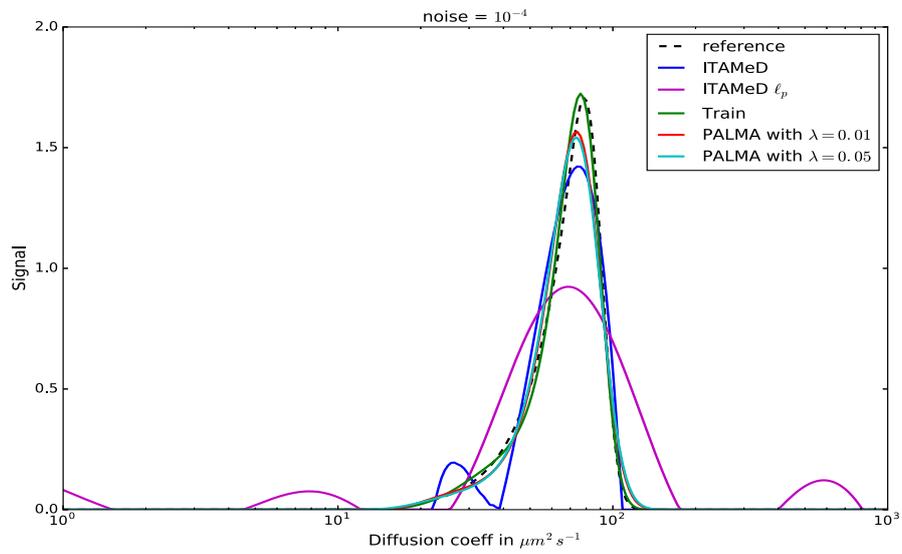

Figure S11: Reconstruction of signal C2, using different algorithms. Here, $M = 64$ and noise level $= 1\%$.

Figure S12: Reconstruction of signal C2, using different algorithms. Here, $M = 64$ and noise level $= 0.1\%$.

Figure S13: Reconstruction of signal C2, using different algorithms. Here, $M = 64$ and noise level $= 0.01\%$.
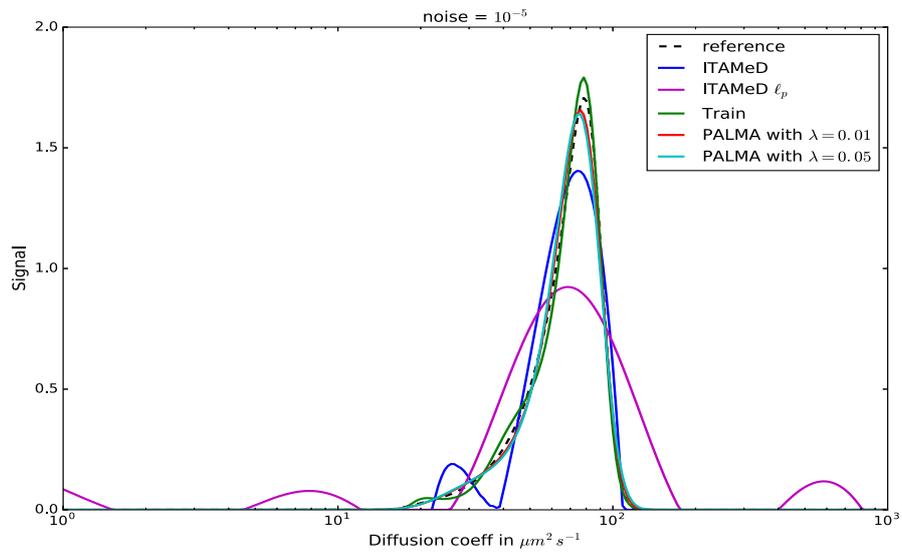
Figure S14: Reconstruction of signal C2, using different algorithms. Here, $M = 64$ and noise level = 0.001%.

8 Robustness against noise

We propose to evaluate the quality of reconstruction of signal C2 for different λ values and for 4 different noise levels. Figure S15 shows that the best quality reconstruction is obtained when $\lambda = 0.01$ whatever the different noise level. For this optimal λ PALMA algorithm ensures a great quality of reconstruction as it is illustrated in Figure S16.
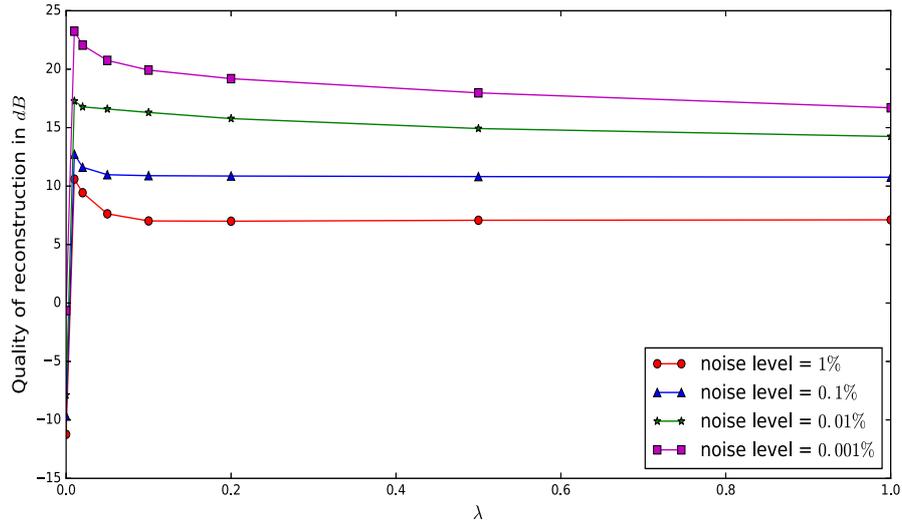

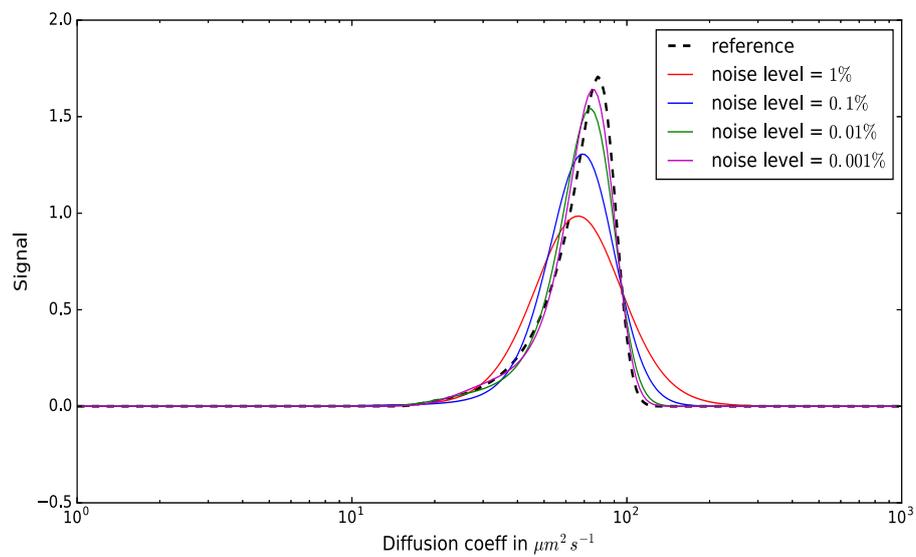Figure S15: Quality of reconstruction of signal C2 with different λ.

Figure S16: Reconstruction of signal C2 with $\lambda_{optimal}$ for different noise levels.