Efficient Maximum Entropy Reconstruction of Nuclear Magnetic Resonance T1-T2 Spectra

Emilie Chouzenoux, Saïd Moussaoui, Jérôme Idier, *Member, IEEE,* and François Mariette

***Abstract*—This paper deals with the reconstruction of T1-T2 correlation spectra in Nuclear Magnetic Resonance relaxometry. The ill-posed character and the large size of this inverse problem are the main difficulties to tackle. While maximum entropy is retained as an adequate regularization approach, the choice of an efficient optimization algorithm remains a challenging task. Our proposal is to apply a truncated Newton algorithm with two original features. Firstly, a theoretically sound line search strategy suitable for the entropy function is applied to ensure the convergence of the algorithm. Secondly, an appropriate preconditioning structure based on a singular value decomposition of the forward model matrix is used to speed up the algorithm convergence. Furthermore, we exploit the specific structures of the observation model and the Hessian of the criterion to reduce the computation cost of the algorithm. The performances of the proposed strategy are illustrated by means of synthetic and real data processing.**

***Index Terms*—Nuclear magnetic resonance, T1-T2 spectrum, Laplace inversion, Maximum entropy, truncated Newton, line search, SVD preconditioning.**

I. Introduction

NUCLEAR magnetic resonance (NMR) relaxometry is a measurement technique used to analyze the properties of matter in order to determine its molecular structure and dynamics. After the immersion of the matter in a strong magnetic field, all the nuclear spins align to an equilibrium state along the field orientation. The application of a short magnetic pulse in resonance with the spin motion perturbates the spin orientation with a predefined angle Φ, called flip angle or pulse angle. The NMR experiment aims at analyzing the relaxation process which corresponds to the re-establishment of the spin into its equilibrium state.

This movement is decomposed into longitudinal and transverse dynamics, characterized by relaxation times T_1 and T_2 respectively. In practice, the longitudinal magnetization after τ_1 seconds of relaxation is measured by applying a 90° impulsion in the transverse plane. The transverse magnetization after $\tau_1 + \tau_2$ seconds of relaxation is obtained by a series of dephasing impulsions in the transverse plane ([1, Chap.4], [2, Chap.4], [3]). Classical NMR experiments are conducted to analyze the samples independently, either in terms of longitudinal or transverse relaxation, leading to one-dimensional (1D) distributions [4, 5]. On the contrary, joint measurements with respect to the two relaxation parameters allow to build two-dimensional (2D) T_1-T_2 spectra. Such spectra reveal couplings between T_1 and T_2 relaxations that are very useful for structure determination [6–8].

The physical model behind NMR relaxometry states that the measured NMR data $X(\tau_1, \tau_2)$ are related to the T_1-T_2 spectrum $S(T_1, T_2)$, according to a 2D Fredholm integral of the first kind

$$X(\tau_1, \tau_2) = \iint k_1(\tau_1, T_1) S(T_1, T_2) k_2(\tau_2, T_2) dT_1 dT_2 \quad (1)$$

where k_1 and k_2 are kernels modeling the longitudinal and transverse relaxations

$$\begin{array}{ll} k_1(\tau_1, T_1) & = 1 - \gamma e^{-\tau_1/T_1}, \\ k_2(\tau_2, T_2) & = e^{-\tau_2/T_2}, \end{array} \quad (2)$$

with $\gamma = 1 - \cos\Phi$. In practice, an uncertainty in this observation model can occur if the pulse angle Φ is not set exactly to its desired value.

The associated inverse problem involving the recovery of the continuous distribution $S(T_1, T_2)$ is known to be an ill-posed problem [9].

Experimental data are collected at $m_1 \times m_2$ discrete values in the τ_1-τ_2 domain. Thus, the data function $X(\tau_1, \tau_2)$ is replaced by a data matrix $\boldsymbol{X} \in \mathbb{R}^{m_1 \times m_2}$. Similarly, the kernels k_1 and k_2 are discretized as matrices $\boldsymbol{K}_1 \in \mathbb{R}^{m_1 \times N_1}$ and $\boldsymbol{K}_2 \in \mathbb{R}^{m_2 \times N_2}$. Equation (1) takes a discrete form $\boldsymbol{X} = \boldsymbol{K}_1 \boldsymbol{S} \boldsymbol{K}_2^{\mathrm{t}}$, where the spectrum $\boldsymbol{S}$ is a real-valued matrix of size $N_1 \times N_2$. In practice, measurements are modeled by

$$\boldsymbol{Y} = \boldsymbol{K}_1 \boldsymbol{S} \boldsymbol{K}_2^{\mathrm{t}} + \boldsymbol{E} \quad (3)$$

with $\boldsymbol{E}$ a noise term assumed white Gaussian. 2D NMR reconstruction amounts to estimating $\boldsymbol{S}$ given $\boldsymbol{Y}$ subject to $\boldsymbol{S} \succeq 0$ (in the sense $S_{ij} \geqslant 0 \ \forall i, j$). Attention must be paid to the size of the 2D NMR problem. Indeed, when converted to a standard one-dimensional representation, (3) reads

$$\boldsymbol{y} = \boldsymbol{K}\boldsymbol{s} + \boldsymbol{e} \quad (4)$$

with $\boldsymbol{y} = \mathrm{vect}\,[\boldsymbol{Y}]$, $\boldsymbol{s} = \mathrm{vect}\,[\boldsymbol{S}]$, $\boldsymbol{e} = \mathrm{vect}\,[\boldsymbol{E}]$, $\mathrm{vect}[\cdot]$ denoting a column vector obtained by stacking all the elements of a matrix in lexicographic order and

$$\boldsymbol{K} = \boldsymbol{K}_1 \otimes \boldsymbol{K}_2 \quad (5)$$

is the Kronecker product between matrices $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$. Matrix $\boldsymbol{K}$ is thus of size $m_1 m_2 \times N_1 N_2$. Typical values are $m_1 = 50$, $m_2 = 10^4$, $N_1 \times N_2 = 200 \times 200$, so $\boldsymbol{K}$ is a huge matrix whose

E. Chouzenoux, S. Moussaoui, and J. Idier are with IRCCyN (CNRS UMR 6597), Ecole Centrale Nantes, 44321 Nantes Cedex 03, France. (e-mail: emilie.chouzenoux@irccyn.ec-nantes.fr; said.moussaoui@irccyn.ec-nantes.fr; jerome.idier@irccyn.ec-nantes.fr)

F. Mariette is with Cemagref, UR TERE, F-35044 Rennes, France and Université européenne de Bretagne, France. (e-mail: francois.mariette@cemagref.fr)

explicit handling is almost impossible. It is one of the two main contributions of this paper to make use of the factored form (3) to solve this issue without any approximation.

Adopting the well-known least-squares approach would lead to define a spectrum estimate as the minimizer of

$$C(\boldsymbol{S}) = \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{K}_1\boldsymbol{S}\boldsymbol{K}_2^{\mathrm{t}}\|_F^2, \tag{6}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, under the positivity constraint $\boldsymbol{S} \succeq 0$. However, $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are rank-deficient and very badly conditioned matrices [10]. Therefore, such a solution is numerically unstable and regularized solutions must be sought instead. Given that the maximum entropy approach provides acknowledged methods for conventional (*i.e.*, one-dimensional) NMR [4, 11], this paper explores T_1-T_2 spectrum estimation based on maximum entropy regularization and proposes a specific descent algorithm. According to our experience, the barrier shape of the entropy function makes the minimization problem quite specific. In particular, general-purpose non-linear programming algorithms can be extremely inefficient in terms of convergence speed. More surprisingly, the more specific scheme adapted from [12] also turns out to be very slow to converge. This motivated us to devise an alternative optimization strategy that is provably convergent and shows a good trade-off between simplicity and efficiency. The proposed algorithm belongs to the truncated Newton algorithm family but possesses original features regarding the line search and the preconditioning strategy.

The rest of the paper is organized as follows: Section II gives an overview of different regularization strategies that can be applied to solve this problem. Section III proposes an efficient reconstruction method for maximum entropy regularization, based on a truncated Newton algorithm associated with an original line search strategy well suited to the form of the criterion. The computation cost of the algorithm is reduced by working directly with the factored form (6) to calculate quantities such as gradient and Hessian-vector products. In section IV, the efficiency of the proposed scheme is illustrated by means of synthetic and real data examples.

II. Problem statement and existing solutions

The mathematical methods developed to solve (1) can be classified into two groups. The first approach is to fit the decay curves with a minimal number of discrete exponentials terms. The parametric minimization is usually handled with the Levenberg-Marquardt algorithm [13]. In this paper, we rather focus on the second approach which analyzes the data in terms of a continuous distribution of relaxation components $S(T_1, T_2)$. This model gives rise to the linear equation (3). In this section, we give an overview of different inversion strategies for this problem.

A. Direct Resolution: TSVD and Tikhonov Methods

NMR reconstruction is a linear ill-posed problem. To tackle it, truncated singular value decomposition (TSVD) and Tikhonov penalization (TIK) are commonly used methods [9]. Each of them calls for its own regularization principle to compensate the ill-conditioned character of the observation matrix.

1) TSVD: The TSVD approach consists in replacing the inverse (or the generalized inverse) of $\boldsymbol{K}$ by a matrix of reduced rank, in order to avoid the amplification of noise due to the inversion of small nonzero singular values [14]. In practice, computing the TSVD requires the explicit decomposition of $\boldsymbol{K}$ in terms of singular elements, which can be numerically burdensome.

2) Tikhonov penalization: While TSVD tackles the ill-posed character by control of dimensionality, Tikhonov method follows a penalization approach by which a trade-off is sought between fidelity-to-data and regularity. It leads to the minimization of a mixed objective function

$$L(\boldsymbol{S}) = C(\boldsymbol{S}) + \lambda R(\boldsymbol{S}) \tag{7}$$

where the regularization parameter $\lambda > 0$ controls the respective weight of the two terms, C is a least-square term

$$C(\boldsymbol{S}) = \frac{1}{2}\left\|\boldsymbol{y} - \boldsymbol{K}\boldsymbol{s}\right\|^2 = \frac{1}{2}\left\|\boldsymbol{Y} - \boldsymbol{K}_1\boldsymbol{S}\boldsymbol{K}_2^{\mathrm{t}}\right\|_F^2 \tag{8}$$

and the additional term R is also a quadratic term. In the context of NMR reconstruction, the regularization functionnal R is usually chosen as the squared ℓ_2-norm of the spectrum ([5, 10, 15, 16])

$$R(\boldsymbol{S}) = \frac{1}{2}\left\|\boldsymbol{s}\right\|^2 = \frac{1}{2}\left\|\boldsymbol{S}\right\|_F^2. \tag{9}$$

Tikhonov solution is then obtained by solving the linear system $(\boldsymbol{K}^{\mathrm{t}}\boldsymbol{K} + \lambda\boldsymbol{I})\boldsymbol{s} = \boldsymbol{K}^{\mathrm{t}}\boldsymbol{y}$.

B. Iterative Minimization

Both TSVD and TIK solutions provide results of limited resolution. Moreover, they tend to exhibit oscillatory excursions, especially in the peripheral regions of the recovered peaks, which usually violate the positivity of the spectrum components [17]. Enforcing the positivity of the spectrum is obviously desirable from the viewpoint of physical interpretation, but it has also a favorable effect on the resolution of the estimated spectrum.

1) Tikhonov under positivity constraint (TIK$^+$): The positivity constraint $\boldsymbol{S} \succeq 0$ is naturally incorporated into Tikhonov approach by constraining the minimization of L to the positive orthant. However, there is no closed-form expression for the minimizer anymore, so the solution must be computed iteratively using a fixed-point algorithm.

Butler-Reeds-Dawson algorithm (BRD) [10] is a rather simple and efficient technique based on the resolution of the Karush-Kuhn-Tucker conditions [18]. Although commonly used in materials science, it is scarcely referenced in the quadratic programming literature. For the sake of clarification, Appendix A proposes a very simple interpretation of the BRD scheme as iteratively minimizing a dual function of the criterion in the sense of Legendre-Fenchel duality [19].

However, the BRD scheme requires the inversion of a system of size $m \times m$ at each iteration, where m is the number of measurements. In the case of 2D NMR problems, $m = m_1 m_2$, and usual values of m_1 and m_2 lead to a prohibitive computation cost. To solve this issue, a data compression step is proposed in [15], prior to the application of BRD. It relies on

strongly truncated singular value decompositions of $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$, $\boldsymbol{K}_i \approx \boldsymbol{U}_i\boldsymbol{\Sigma}_i\boldsymbol{V}_i^{\mathrm{t}}$, $i = 1, 2$, with $\tilde{m}_i = \mathrm{rank}(\boldsymbol{K}_i) \ll m_i$. The fidelity to data term is then approximated by

$$\tilde{C}(\boldsymbol{S}) = \frac{1}{2}\|\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{K}}_1\boldsymbol{S}\tilde{\boldsymbol{K}}_2^{\mathrm{t}}\|_F^2 \tag{10}$$

where $\tilde{\boldsymbol{K}}_1 = \boldsymbol{\Sigma}_1\boldsymbol{V}_1^{\mathrm{t}}$, $\tilde{\boldsymbol{K}}_2 = \boldsymbol{\Sigma}_2\boldsymbol{V}_2^{\mathrm{t}}$ and $\tilde{\boldsymbol{Y}} = \boldsymbol{U}_1^{\mathrm{t}}\boldsymbol{Y}\boldsymbol{U}_2$ are of size $\tilde{m}_1 \times N_1$, $\tilde{m}_2 \times N_2$ and $\tilde{m}_1 \times \tilde{m}_2$, respectively.

2) Maximum entropy: A different regularization approach will be considered here, based on Shannon entropy penalization $\phi(s) = -s\log s$. Maximum entropy (ME) [12, 20] is an acknowledged approach in the context of 1D NMR relaxometry [4, 11]. An interesting feature of entropy penalization is that it implicitly handles the positivity constraint since the norm of the gradient of the entropy term is unbounded at the boundary of the positive orthant. Thus, the minimizer of the resulting penalized least-square criterion cancels its gradient, and computing it is essentially similar to solving an unconstrained optimization problem.

Formally, the extension to the 2D case is easily obtained by minimization of

$$L(\boldsymbol{S}) = \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{K}_1\boldsymbol{S}\boldsymbol{K}_2^{\mathrm{t}}\|_F^2 + \lambda\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} S_{ij}\log S_{ij}. \tag{11}$$

However, the practical computation of the solution is clearly more difficult in the 2D case because the optimization problem is much larger-scale. The choice of a specific minimization scheme suited to maximum entropy 2D NMR reconstruction is a challenging task.

In the context of maximum entropy, [21] proposed the fixed-point multiplicative algebraic reconstruction technique (MART) that maximizes the entropy term subject to $\boldsymbol{K}\boldsymbol{s} = \boldsymbol{y}$. The simplicity of MART is attractive. However, as emphasized in [22], the presence of inherent noise in projection data makes this method less effective than an approach based on the minimization of the penalized criterion (11). In [12], an iterative minimization algorithm based on a quadratic approximation of the criterion over a low-dimension subspace is developped. However, according to [23, p. 1022], the convergence of this algorithm is not established. We have tested its behavior in the 2D NMR context. Our conclusions are that this algorithm does not ensure a monotonic decrease of the criterion, and that its convergence is very slow [24]. Finally, in a preliminary version of the present work, we have proposed to make use of a preconditioned nonlinear conjugate gradient algorithm [25]. Although the latter shows a good practical behavior, its theoretical convergence is not ensured, since the preconditioner is a variable matrix.

The goal of the next section is to derive an optimization algorithm that would benefit from stronger theoretical properties and sufficiently low computational cost to avoid any data compression step.

III. Proposed truncated Newton algorithm

A. Minimization Strategy

The truncated Newton (TN) algorithm [26, 27] is based on iteratively decreasing the objective function $L(\boldsymbol{s})$ by moving the current solution $\boldsymbol{s}_k$ along a descent direction $\boldsymbol{d}_k$

$$\boldsymbol{s}_{k+1} = \boldsymbol{s}_k + \alpha_k\boldsymbol{d}_k, \tag{12}$$

where $\alpha_k > 0$ is the stepsize and $\boldsymbol{d}_k$ is a search direction computed by solving approximately the Newton equations

$$\boldsymbol{H}_k\boldsymbol{d}_k = -\boldsymbol{g}_k \tag{13}$$

with $\boldsymbol{H}_k \triangleq \nabla^2 L(\boldsymbol{s}_k)$ and $\boldsymbol{g}_k \triangleq \nabla L(\boldsymbol{s}_k)$. The TN algorithm has been widely used in the context of interior point algorithms with logarithmic [28, 29] and entropic [22] barrier functions.

In practice, the TN method consists in alternating the construction of $\boldsymbol{d}_k$ and the computation of the stepsize α_k by a line search procedure. The direction $\boldsymbol{d}_k$ results from preconditioned conjugate gradient (PCG) iterations on (13) stopped before convergence. The stepsize α_k is obtained by iteratively minimizing the scalar function $\ell(\alpha) = L(\boldsymbol{s}_k + \alpha\boldsymbol{d}_k)$ until some convergence conditions are met [18, Chap.3]. Typically, the strong Wolfe conditions are considered

$$\ell(\alpha_k) \leqslant \ell(0) + c_1\alpha_k\dot{\ell}(0) \tag{14}$$

$$|\dot{\ell}(\alpha_k)| \leqslant c_2|\dot{\ell}(0)| \tag{15}$$

where $(c_1, c_2) \in (0, 1)$ are tuning parameters that do not depend on k. There exist several procedures to find an acceptable stepsize: exact minimization of $\ell(.)$, backtracking, approximation of $\ell(.)$ using cubic interpolations [18, 30] or quadratic majorizations [31, 32]. However, the entropic penalty term implies that the derivative of $\ell(\alpha)$ takes the value $-\infty$ as soon as any of the components of the vector $\boldsymbol{s}_k + \alpha\boldsymbol{d}_k$ vanishes, hence when α is equal to one of the two limit values

$$\alpha_- = \max_{i,\, d_{k,i}>0}\left(\frac{-s_i}{d_{k,i}}\right), \quad \alpha_+ = \min_{i,\, d_{k,i}<0}\left(\frac{-s_i}{d_{k,i}}\right). \tag{16}$$

The function ℓ is undefined outside (α_-, α_+), therefore, we must ensure that during the line search, the stepsize values remain in the interval (α_-, α_+). Moreover, because of the vertical asymptotes at α_- and α_+, standard methods using cubic interpolations or quadratic majorizations are not well suited. Our proposal is to adopt the specific majorization-based line search proposed in [33, 34] for barrier function optimization. Using an adequate form of majorization, we now derive an analytical stepsize formula preserving strong convergence properties.

B. Line Search Strategy

The minimization of $\ell(\cdot)$ using the Majorization-Minimization (MM) principle [35] is performed by successive minimizations of majorant functions for $\ell(.)$. Function $h(\alpha, \alpha')$ is said to be majorant for $\ell(\alpha)$ at α' if for all α,

$$\begin{cases} h(\alpha, \alpha') \geqslant \ell(\alpha) \\ h(\alpha', \alpha') = \ell(\alpha') \end{cases} \tag{17}$$

As illustrated in Fig.1, the initial minimization of $\ell(\alpha)$ is then replaced by a sequence of easier subproblems, corresponding to the MM update rule

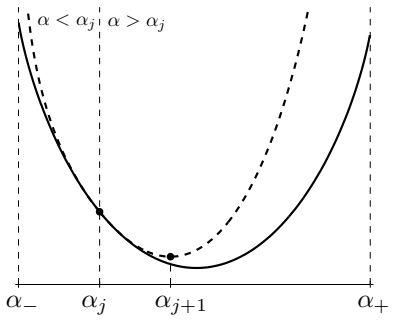
$$\begin{cases} \alpha_k^0 = 0, \\ \alpha_k^j = \arg\min_\alpha h^j(\alpha, \alpha_k^{j-1}), \quad j = 1, \ldots, J_k, \\ \alpha_k = \alpha_k^{J_k}. \end{cases} \tag{18}$$

Following [34], we propose a majorant function $h^j(.,\alpha_k^j)$ that incorporates barriers to account for the entropy term. It is piecewise defined under the following form (whenever unambiguous, the iteration index k will be dropped for the sake of simplicity)

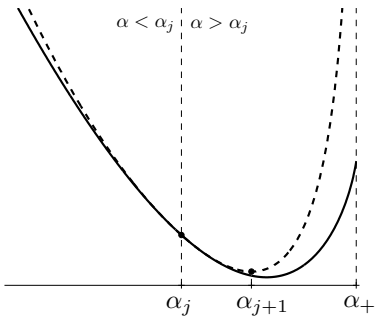
$$h^j(\alpha,\alpha^j)=\begin{cases} p_0^- + p_1^-\alpha + p_2^-\alpha^2 - p_3^-\log(\alpha-\alpha_-) \\ \qquad\qquad \text{for all } \alpha\in(\alpha_-;\alpha^j] \\ p_0^+ + p_1^+\alpha + p_2^+\alpha^2 - p_3^+\log(\alpha_+-\alpha) \\ \qquad\qquad \text{for all } \alpha\in[\alpha^j;\alpha_+) \end{cases} \tag{19}$$

The parameters $p_n^{\pm}$, $n=0,\ldots,3$ must be defined to ensure that $h^j(.,\alpha^j)$ is actually a majorant of $\ell(\cdot)$ at α^j (see Fig. 1(a) for an illustration). A direct application of [34, Prop. 2] allows to establish expressions for these parameters. The resulting form of $h^j(.,\alpha^j)$ is rather simple, though lengthy to express, so it is reported in Appendix B. According to [34, Lemma 2], it corresponds to a strictly convex, twice differentiable function in the set (α_-,α_+). Moreover, its unique minimizer takes an explicit form, the latter being also found in Appendix B.

Finally, (18) produces monotonically decreasing values $\{\ell(\alpha^j)\}$ and the series $\{\alpha^j\}$ converges to a stationary point of $\ell(\alpha)$ [36].



(a) Case α_- and α_+ finite



(b) Case $\alpha_-=-\infty$ and α_+ finite

Fig. 1. Schematic principle of the MM line search procedure. The tangent majorant function $h^j(\alpha,\alpha^j)$ (dashed line) for $\ell(\alpha)$ (solid line) at α_j is piecewise defined on the sets $(\alpha_-,\alpha_j]$ and $[\alpha_j,\alpha_+)$. The new iterate α_{j+1} is taken as the minimizer of $h_j(.,\alpha_j)$. Two cases are illustrated. The third and last case where α_- is finite and $\alpha_+=+\infty$ is the mirror image of case (b).

C. Convergence Result

Let us focus on the convergence of the truncated Newton algorithm when α_k is chosen according to the proposed MM strategy. A detailed analysis can be found in [34] in a more general framework. According to [34], the proposed line search procedure ensures that

$$\sum_k \frac{(\boldsymbol{g}_k^{\mathrm{t}}\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2}<\infty \tag{20}$$

and that the directions generated by the TN algorithm are *gradient related* in the sense of [37]. According to [38], inequality (20), known as *Zoutendijk condition*, is sufficient to prove the convergence of the algorithm in the sense $\lim_{k\to\infty}\|\boldsymbol{g}_k\|=0$. Finally, the objective function being strictly convex, the proposed algorithm converges to its unique minimizer.

D. Preconditioning

As emphasized in [39], the Hessian of the Shannon entropy regularization term is very ill-conditioned for points that are close to the boundary of the positive orthant since some of its eigenvalues tend to infinity. Furthermore, the exponential decays in kernels k_1 and k_2 imply that $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are also very ill-conditioned. Preconditioning is a well-known technique to obtain more clustered eigenvalues of the Hessian of the criterion and to accelerate the convergence of descent algorithms. The principle is to transform the space of original variables into a space in which the Hessian has more clustered eigenvalues by using a preconditioning matrix $\boldsymbol{P}_k$ that approximates the inverse $\boldsymbol{H}_k^{-1}$ of the Hessian. A good preconditioner achieves a trade-off between the approximation quality and the computation cost. General-purpose preconditioning strategies have been proposed in the literature including symmetric successive overrelaxation and incomplete LU or Cholesky factorizations ([40, Chap.10], [41]). In the context of ME optimization, [22] takes $\boldsymbol{P}_k$ as a diagonal matrix defined using the Hessian diagonal elements

$$\boldsymbol{P}_k=\left[\operatorname{diag}\left(\operatorname{diag}(\boldsymbol{K}^{\mathrm{t}}\boldsymbol{K})\right)+\lambda\operatorname{diag}(\boldsymbol{s}_k)^{-1}\right]^{-1} \tag{21}$$

We rather propose a more specific preconditioner. It is based on the fact that, as a consequence of (5), the singular value decomposition of $\boldsymbol{K}$ is given by $\boldsymbol{K}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{t}}$, with $\boldsymbol{U}=\boldsymbol{U}_1\otimes\boldsymbol{U}_2$, $\boldsymbol{V}=\boldsymbol{V}_1\otimes\boldsymbol{V}_2$, $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_1\otimes\boldsymbol{\Sigma}_2$, $\boldsymbol{U}_i\boldsymbol{\Sigma}_i\boldsymbol{V}_i^{\mathrm{t}}$ being the singular value decomposition of $\boldsymbol{K}_i$, $i=1,2$. Then, let us define

$$\boldsymbol{P}_k=\left[\tilde{\boldsymbol{V}}\tilde{\boldsymbol{\Sigma}}^2\tilde{\boldsymbol{V}}^{\mathrm{t}}+\lambda\operatorname{diag}(\boldsymbol{s}_k)^{-1}\right]^{-1} \tag{22}$$

where $\tilde{\boldsymbol{V}}$ and $\tilde{\boldsymbol{\Sigma}}$ correspond to truncated versions of $\boldsymbol{V}$ and $\boldsymbol{\Sigma}$. In the non-truncated case, $\tilde{\boldsymbol{V}}=\boldsymbol{V}$ and $\tilde{\boldsymbol{\Sigma}}=\boldsymbol{\Sigma}$, and $\boldsymbol{P}_k$ then is equal to the Hessian of $\boldsymbol{L}$ at $\boldsymbol{s}_k$. It remains to define the way we truncate the singular value decomposition of $\boldsymbol{K}$. Akin to [15, 42], we separately truncate the decompositions of $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ at ranks v_1, v_2 and we define $\tilde{\boldsymbol{V}}$ and $\tilde{\boldsymbol{\Sigma}}$ according to

$$\tilde{\boldsymbol{V}}=\tilde{\boldsymbol{V}}_1\otimes\tilde{\boldsymbol{V}}_2, \tag{23}$$

$$\tilde{\boldsymbol{\Sigma}}=\tilde{\boldsymbol{\Sigma}}_1\otimes\tilde{\boldsymbol{\Sigma}}_2. \tag{24}$$

Let us remark that the resulting approximation of $\boldsymbol{K}$ may slightly differ from the TSVD of $\boldsymbol{K}$. The reason is simple:

although $\tilde{\boldsymbol{\Sigma}}_1$ and $\tilde{\boldsymbol{\Sigma}}_2$ separately gather the largest singular values of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, $\tilde{\boldsymbol{\Sigma}}$ does not necessarily gather the largest singular values of $\boldsymbol{\Sigma}$. As a consequence, our approximation may be suboptimal compared to the TSVD, the latter being optimal in the least-square sense [43], but the fact that we maintain factored expressions for matrices $\tilde{\boldsymbol{V}}$ and $\tilde{\boldsymbol{\Sigma}}$ is essential in terms of computation cost.

E. Memory Storage and Computation Cost Reduction

The computation cost can be reduced by exploiting the factored form of the observation model. Three main operations are involved in the iterative optimization algorithm: the computation of the gradient vector $\boldsymbol{g}_k = \nabla L(\boldsymbol{s}_k)$, and the products of $\boldsymbol{P}_k$ and $\boldsymbol{H}_k$ with a vector. The three resulting quantities can be calculated using low cost operations, as described below.

1) Gradient: The gradient of the criterion can be computed without explicitly handling matrix $\boldsymbol{K}$, according to

$$\boldsymbol{g}_k = -\mathrm{vect}\left[\boldsymbol{K}_1^{\mathrm{t}}(\boldsymbol{Y} - \boldsymbol{K}_1\boldsymbol{S}_k\boldsymbol{K}_2^{\mathrm{t}})\boldsymbol{K}_2\right] + \lambda(1+\log \boldsymbol{s}_k). \quad (25)$$

2) Hessian: In the same manner, products between the Hessian matrix and any vector $\boldsymbol{w} = \mathrm{vect}\,[\boldsymbol{W}]$ can be computed as follows

$$\boldsymbol{H}_k\boldsymbol{w} = \mathrm{vect}\left[\boldsymbol{K}_1^{\mathrm{t}}\boldsymbol{K}_1\boldsymbol{W}\boldsymbol{K}_2^{\mathrm{t}}\boldsymbol{K}_2\right] + \lambda(\boldsymbol{w}./\boldsymbol{s}_k), \quad (26)$$

where $./$ denotes componentwise division.

3) Preconditioner: In order to compute products involving $\boldsymbol{P}_k$, the matrix inversion lemma is applied to (22). Thus,

$$\boldsymbol{P}_k = \boldsymbol{A}_k - \boldsymbol{A}_k\tilde{\boldsymbol{V}}(\tilde{\boldsymbol{\Sigma}}^{-2} + \tilde{\boldsymbol{V}}^{\mathrm{t}}\boldsymbol{A}_k\tilde{\boldsymbol{V}})^{-1}\tilde{\boldsymbol{V}}^{\mathrm{t}}\boldsymbol{A}_k, \quad (27)$$

with $\boldsymbol{A}_k = \lambda^{-1}\mathrm{diag}(\boldsymbol{s}_k)$. Moreover, the following factored expression can be deduced from (23) for the entries of matrix $\boldsymbol{M} = \tilde{\boldsymbol{V}}^{\mathrm{t}}\boldsymbol{A}_k\tilde{\boldsymbol{V}} \in \mathbb{R}^{v_1v_2\times v_1v_2}$

$$M_{ij} = \frac{1}{\lambda}\sum_{m=1}^{N_1}\sum_{n=1}^{N_2}(S_k)_{mn}\,(\tilde{V}_1)_{ma}\,(\tilde{V}_2)_{nb}\,(\tilde{V}_1)_{mc}\,(\tilde{V}_2)_{nd},$$

where (a,b) and (c,d) are row and column subscripts that correspond to the linear indexes i and j, respectively. Thus, the product $\boldsymbol{P}_k\boldsymbol{w}$ can be efficiently computed according to

$$\begin{aligned}\boldsymbol{P}_k\boldsymbol{w} &= \boldsymbol{b}_k - \boldsymbol{A}_k\tilde{\boldsymbol{V}}(\tilde{\boldsymbol{\Sigma}}^{-2} + \boldsymbol{M})^{-1}\tilde{\boldsymbol{V}}^{\mathrm{t}}\boldsymbol{b}_k,\\ &= \boldsymbol{b}_k - \boldsymbol{A}_k\mathrm{vect}\big[\tilde{\boldsymbol{V}}_1\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_2^{\mathrm{t}}\big] \quad (28)\end{aligned}$$

where $\boldsymbol{b}_k = \boldsymbol{A}_k\boldsymbol{w}$, $\boldsymbol{q}_k = (\tilde{\boldsymbol{\Sigma}}^{-2} + \boldsymbol{M})^{-1}\,\mathrm{vect}\big[\tilde{\boldsymbol{V}}_1^{\mathrm{t}}\boldsymbol{B}_k\tilde{\boldsymbol{V}}_2\big]$ and $\boldsymbol{Q}_k$, $\boldsymbol{B}_k$ denote the equivalent square matrix representations of $\boldsymbol{q}_k$ and $\boldsymbol{b}_k$ respectively.

F. Resulting algorithm

The resulting TN algorithm is given in Alg. 1. The algorithm convergence is checked using the following stopping rule ([18])

$$\|\boldsymbol{g}_k\|_\infty < \epsilon(1 + |L(\boldsymbol{s}_k)|), \quad (29)$$

and the PCG iterations in Alg. 2 are stopped when ([27])

$$\|\boldsymbol{g}_k + \boldsymbol{H}_k\boldsymbol{d}_k\| \leqslant \eta\|\nabla L(\boldsymbol{s}_k)\|. \quad (30)$$

Typical values of (ϵ,η) are $(10^{-8}, 10^{-4})$.

Require: Initial value $\boldsymbol{s}_0 \succeq 0$, parameters v_1, v_2, λ, J and accuracies ϵ, η.
Ensure: Resolution of (11)
  Compute the TSVD of $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ at ranks v_1, v_2.
  while (29) does not hold **do**
    Compute $\boldsymbol{g}_k, \boldsymbol{P}_k$ and $\boldsymbol{H}_k$ using (25), (26) and (27).
    Compute $\boldsymbol{d}_k$ using PCG algorithm (Table 2).
    Set α_k after J iterations of (18).
    Update $\boldsymbol{s}_k$ according to (12).
  end while

Algorithm 1: TN algorithm for ME optimization

Require: $\boldsymbol{g}_k, \boldsymbol{H}_k, \boldsymbol{P}_k, \eta$
Ensure: Approximate solution $\boldsymbol{d}_k$ of (13)
  $\boldsymbol{u}_0 \leftarrow \boldsymbol{0}$
  $\boldsymbol{r}_0 \leftarrow -\boldsymbol{g}_k - \boldsymbol{H}_k\boldsymbol{u}_0$
  $\boldsymbol{p}_0 \leftarrow \boldsymbol{P}_k\boldsymbol{r}_0$
  while (30) does not hold **do**
    $\theta_i \leftarrow (\boldsymbol{r}_i^{\mathrm{t}}\boldsymbol{P}_k\boldsymbol{r}_i)/(\boldsymbol{p}_i^{\mathrm{t}}\boldsymbol{H}_k\boldsymbol{p}_i)$
    $\boldsymbol{u}_{i+1} \leftarrow \boldsymbol{u}_i + \theta_i\boldsymbol{p}_i$
    $\boldsymbol{r}_{i+1} \leftarrow \boldsymbol{r}_i - \theta_i\boldsymbol{H}_k\boldsymbol{p}_i$
    $\beta_i \leftarrow (\boldsymbol{r}_{i+1}^{\mathrm{t}}\boldsymbol{P}_k\boldsymbol{r}_{i+1})/(\boldsymbol{r}_i^{\mathrm{t}}\boldsymbol{P}_k\boldsymbol{r}_i)$
    $\boldsymbol{p}_{i+1} \leftarrow \boldsymbol{P}_k\boldsymbol{r}_{i+1} + \beta_i\boldsymbol{p}_i$
    $\boldsymbol{d}_k \leftarrow \boldsymbol{u}_{i+1}$
  end while

Algorithm 2: PCG algorithm

IV. Experimental results

This section discusses the performances of the proposed method and illustrates its applicability. First, we consider synthetic data in order to discuss the influence of the tuning parameters on the algorithm behaviour. Then, the proposed method applicability is illustrated through the processing of real NMR data.

In NMR experiments, the pulse angle Φ may not be set exactly to its desired value. Therefore, we analyze the effect of a potential error in the value of γ in the observation model and propose an original strategy allowing to estimated this parameter.

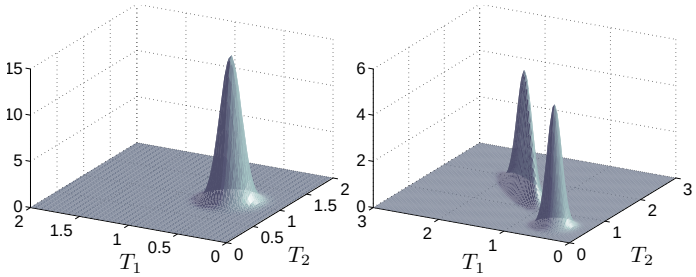
The different results are obtained with Matlab 7.5 running on an Intel Pentium 4 3.2 GHz, 3 GB RAM.

A. Synthetic Data

We consider two spectra A and B (Fig. 2) and the corresponding decays (Fig. 3) according to the observation model (4) with a signal to noise ratio (SNR) of 10 dB, $m_1 = 100$, $m_2 = 1000$ and $\gamma = 1$ (i.e, $\Phi = 90°$). The synthetic spectrum A has a symmetric Gaussian shape located at $[T_1, T_2] = [0.5\text{ s}, 1\text{ s}]$ while spectrum B is the sum of two Gaussian patterns. The first one is symmetric and located at $[T_1, T_2] = [0.5\text{ s}, 0.5\text{ s}]$. The second pattern is located at $[T_1, T_2] = [1.5\text{ s}, 1.5\text{ s}]$ and simulates a positive T_1-T_2 correlation. The reconstruction is performed for $N_1 = N_2 = 100$ and the algorithm is initialized with a uniform positive 2D spectrum. The regularization parameter λ is set to minimize the normalized quadratic error

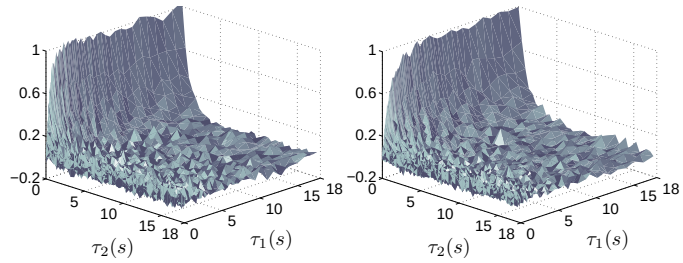
$$Q = 100\,\|\boldsymbol{s}(\lambda) - \boldsymbol{s}^o\|_2^2/\|\boldsymbol{s}^o\|_2^2, \quad (31)$$

and the preconditioner truncation parameters v_1, v_2 are set to the same value v.



(a) Dataset A (b) Dataset B

Fig. 2. Simulated 2D spectra



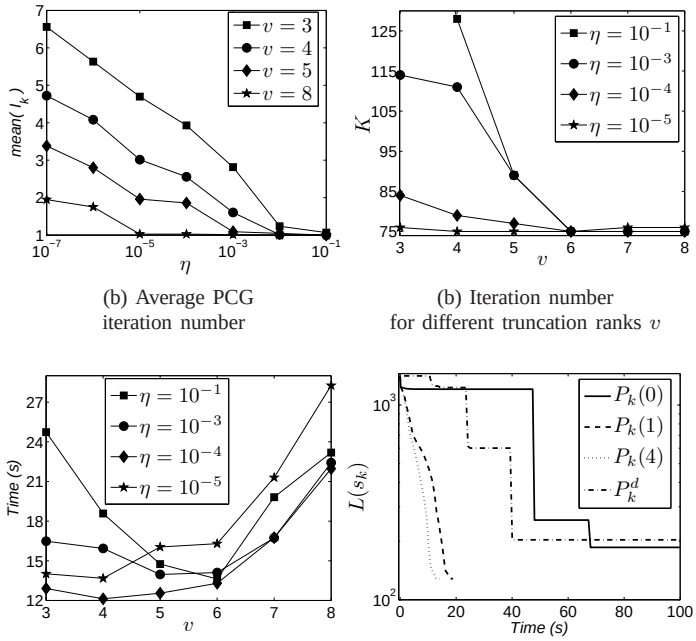
(a) Dataset A (b) Dataset B

Fig. 3. NMR decays

1) PCG subiterations: The parameter η controls the accuracy of the PCG minimization. The smaller it is, the more accurate the solving of (13). Here, several values are tested within the range $[10^{-7}, 10^{-1}]$. Let I_k denotes the number of PCG subiterations (inner loop) at iteration k. As expected, the average value of I_k generally increases with η (Fig. 4(a)) while the number of TN iterations K (outer loop) decreases (Fig. 4(b)). The number of PCG subiterations depends also on the truncation rank v of the SVD preconditioner, it can be noted that I_k decreases as this rank increases, corresponding to a more accurate approximation of the inverse Hessian matrix. The smallest overall minimization time is achieved when a tradeoff is reached between the number of outer iterations and the number of inner iterations (Fig. 4(c)). In this example, the best compromise is $(v, \eta) = (4, 10^{-4})$. This setting will be retained in the sequel.

2) Preconditioning: Fig. 4(d) illustrates the criterion evolution for different preconditioners: the proposed approximation $\boldsymbol{P}_k(v)$ given by (22) with $v_1 = v_2 = v = 0, 1, 4$ and the diagonal preconditioner $\boldsymbol{P}_k^d$ resulting from (21). The stopping criterion (29) is fulfilled after 93 and 80 iterations for $\boldsymbol{P}_k(1)$ and $\boldsymbol{P}_k(4)$ whereas it is not fulfilled after 1000 iterations neither for $\boldsymbol{P}_k(0)$ nor for $\boldsymbol{P}_k^d$. Moreover, according to Fig. 4(a), the TN iteration number decreases as the SVD truncation rank v increases. However, the choice of v involves a compromise between an acceleration of the algorithm and an increase of the computational cost (Fig. 4(b-c)).

3) Line search: Let us compare the performances of the algorithm when the stepsize is obtained either by the proposed MM line search or by Moré and Thuente's cubic interpolation



(b) Average PCG iteration number

(b) Iteration number for different truncation ranks v

(c) Computation time for different truncation ranks v

(d) Comparison between four preconditioning strategies

Fig. 4. Dataset A: Analysis of the TN algorithm performances for different PCG strategies. The SVD preconditioner with truncature parameter v was used for (a)-(c) while the truncature parameter η is set to 10^{-4} for (d). Moreover, in all cases, the stepsize results from $J = 1$ subiteration of MM line search.

procedure (MT) [30]. The latter performs an iterative minimization of $\ell(.)$ based on cubic interpolation until identifying α_k that fulfills the strong Wolfe conditions (14) and (15).

	c_1	c_2	K	T (s)
MT	10^{-1}	0.5	93	19.84
	10^{-1}	0.9	90	15.64
	10^{-1}	0.99	170	25.72
	10^{-3}	0.5	93	16.98
	10^{-3}	0.9	90	**15.36**
	10^{-3}	0.99	170	25.14
	J		K	T (s)
MM	1		79	**13.56**
	2		85	15.09
	3		84	15.06
	4		84	15.11
	5		85	15.31

TABLE I
DATASET A: COMPARISON BETWEEN MM AND MT LINE SEARCH STRATEGIES IN TERMS OF ITERATION NUMBER AND TIME BEFORE CONVERGENCE FOR THE TN ALGORITHM.

According to Table I, the TN algorithm with the MM line search performs better than with the MT line search with the best settings for c_1 and c_2. Concerning the choice of the sub-iteration number, it appears that $J = 1$ leads to the best results in terms of computation time which shows that an exact minimization of the scalar function $\ell(\alpha)$ during line search is not necessary.

4) Regularization term: As explained in the introduction, the application of BRD algorithm to 2D NMR reconstruction requires data compression. This preprocessing step calls for

the tuning of two additional parameters, $\tilde{m}_1$ and $\tilde{m}_2$. Table II illustrates the reconstruction quality and algorithmic properties of BRD method for different values of $\tilde{m}_i$. As expected, the computation cost decreases with $\tilde{m}_i$. However, according to Fig. 5, below a certain compression value $\tilde{m}_1^{\min}$, the reconstruction error fastly grows. We observe that $\tilde{m}_1^{\min} = 3$ for dataset A and $\tilde{m}_1^{\min} = 5$ for dataset B. The same behavior was observed when varying $\tilde{m}_2$. This shows that the compression tuning not only depends on spectral properties of matrices $\boldsymbol{K}_i$ [15], but also on the spectra shape. Therefore, the setting of these parameters may be problematic when processing real data.

	$\tilde{m}_1$	10	10	10	5	2	1
	$\tilde{m}_2$	100	50	10	5	5	1
A	Q	4.53	4.66	4.79	3.92	81.6	97.9
	K	31	30	20	22	13	12
	T (s)	43	29	4	2	< 1	< 1
B	Q	12.8	12.8	12.6	10.7	84.4	94.3
	K	19	18	18	20	11	2
	T (s)	50	27	3	< 1	< 1	< 1

TABLE II
RECONSTRUCTION QUALITY Q, ITERATION NUMBER K AND TIME BEFORE CONVERGENCE T FOR TIK$^+$-BRD RECONSTRUCTION WITH DIFFERENT LEVELS OF DATA COMPRESSION.

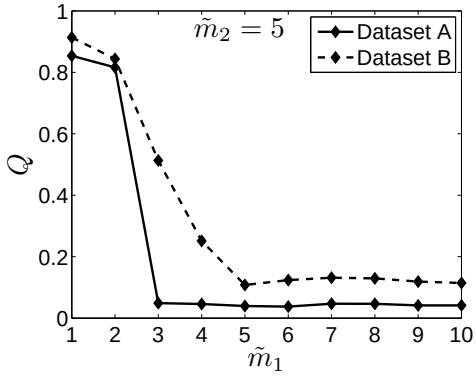


Fig. 5. TIK$^+$-BRD reconstruction quality of dataset A and B with different level of data compression. In both cases, the compression parameter $\tilde{m}_2$ is equal to 5 while $\tilde{m}_1$ is varying. (SNR = 10 dB)

In order to compare the ME and TIK$^+$ regularizations, we apply the same compression level $\tilde{m}_1 = \tilde{m}_2 = 5$. We have tested different noise realizations with SNR $= 5, 10$ and 15 dB. According to Fig. 6, the minimum value of $Q(\lambda)$ decreases with the noise level, for both ME and TIK$^+$ regularizations, as expected. Moreover, the two strategies lead to similar reconstruction minimum errors for the three noise levels. Furthermore, their sensitivity to λ is similar. However, as illustrated in Fig. 7 and Fig. 8, the entropy penalization leads to spectra whose shape is closer to the simulated one. More precisely, the ME spectra are smoother. This regularity is evaluated in Table III which compares the reconstructions in terms of the Euclidian norm of the first-order difference $\|\Delta \boldsymbol{s}\|$.

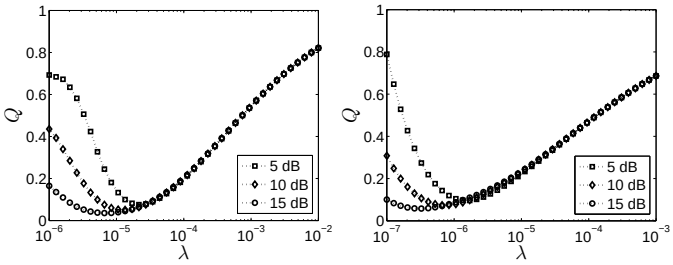


Fig. 6. Dataset A: Similarity error for ME (left) and TIK$^+$ (right) reconstructions. Average of Monte Carlo simulations with 100 random realizations for SNR = $5, 10$ and 15 dB ($\tilde{m}_i = 5$)

	Dataset A		Dataset B	
	ME	TIK$^+$	ME	TIK$^+$
$\|\Delta \boldsymbol{s}\|$	51.9	57.8	23.6	26.1
$\|\Delta \boldsymbol{s}\|/\|\boldsymbol{s}\|$	0.5484	0.5891	0.5256	0.5324

TABLE III
REGULARITY OF RECONSTRUCTED SPECTRA FOR ME AND TIK$^+$ RECONSTRUCTIONS (SNR = 10 DB AND $\tilde{m}_i = 5$). Δ IS THE FIRST ORDER DIFFERENCE MATRIX.

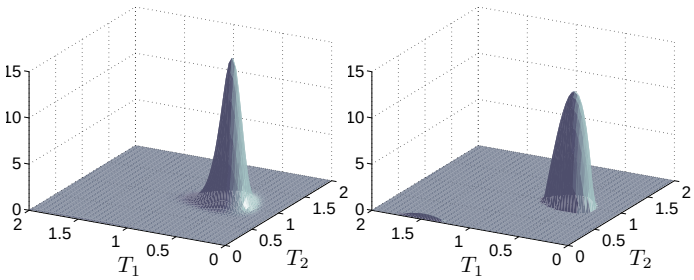


Fig. 7. Dataset A: Reconstructed spectra with optimal setting of λ for ME (left) and TIK$^+$ (right) regularization (SNR = 10 dB and $\tilde{m}_i = 5$).

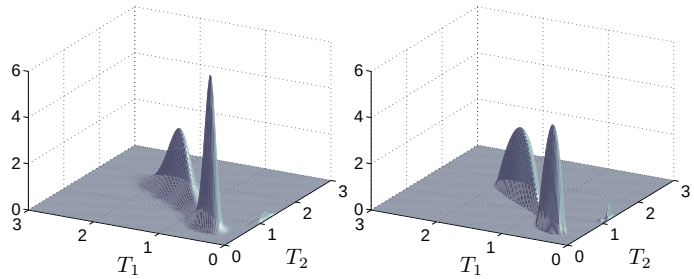


Fig. 8. Dataset B: Reconstructed spectra with optimal setting of λ for ME (left) and TIK (right) regularization (SNR = 10 dB and $\tilde{m}_i = 5$).

5) Hyperparameter estimation: In the previous experiments, the regularization parameter λ is tuned by minimizing a quadratic error whose evaluation requires the knowledge of the reference spectrum. This strategy is impractical in an experimental context but it can be replaced by different procedures proposed in the literature. In NMR reconstruction [4, 10, 15] and ME optimization [12, 44], a frequently used strategy is the Chi-square approach.

Given measurements $\boldsymbol{Y}$ and an estimate of the noise standard deviation $\hat{\sigma}$, statistical considerations state that the error

$$\chi^2(\boldsymbol{S}) = \|\boldsymbol{K}_1 \boldsymbol{S} \boldsymbol{K}_2^{\mathrm{t}} - \boldsymbol{Y}\|_F^2 / \hat{\sigma}^2 \tag{32}$$

follows a Chi-square distribution [45, 46]. In the limit of a

large number of independent measurements $m_1 m_2$, the latter tends to a standard normal distribution with expected value $m_1 m_2$ and variance $2m_1 m_2$.

Thus, a classical method for setting the regularization parameter and avoiding over-smoothed reconstructions ([44, 46]) is to find the value of λ allowing to reach

$$\chi^2_{\text{aim}} = m_1 m_2 - \sqrt{2m_1 m_2} \tag{33}$$

However, when the noise level is high or when the estimation of σ is too rough, one can have $\chi^2(\lambda) > \chi^2_{\text{aim}}$ for all values of the regularization parameter so that the Chi-square test cannot be achieved.

An alternative approach, based on the S-curve [47], consists in choosing λ such that its reduction does not lead to a significant decrease in $\chi^2(\lambda)$

$$\frac{\partial \log_{10} \chi^2(\lambda)}{\partial \log_{10} \lambda} \ll 1. \tag{34}$$

Here, we suggest to combine the two latter strategies for the determination of λ, as detailed in Alg. 3 and Fig. 9. We emphasize that the minimizations (35) can be performed at very low cost by initializing the TN algorithm of Alg. 1 with the solution at previous λ. Table IV illustrates the efficiency of the proposed scheme for finding λ.

Require: Initial values $\boldsymbol{s}_0 \succeq 0$, λ_0, parameter $\theta \in (0, 1)$ and accuracy η
Ensure: ME resolution with Chi-square tuned λ
  while (33) **and** (34) do not hold **do**
    Using Table 1, compute

$$\hat{\boldsymbol{S}} = \arg\min L(\boldsymbol{S}) + \lambda_n R(\boldsymbol{S}). \tag{35}$$

    Compute $\chi^2(\hat{\boldsymbol{S}})$ using (32).
    $\lambda_{n+1} \leftarrow \theta \lambda_n$
  end while

Algorithm 3: Chi-square method for regularization parameter estimation

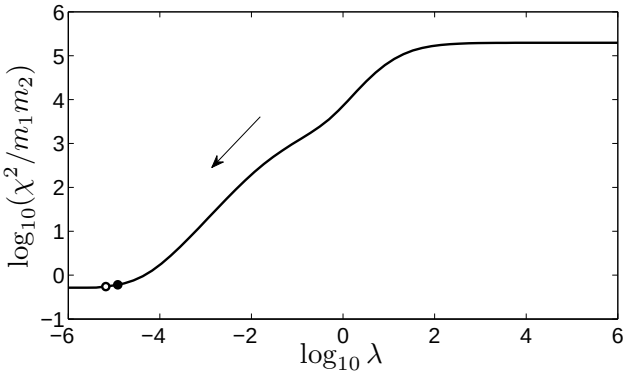


Fig. 9. Dataset A (SNR = 10 dB and $\tilde{m}_i = 5$): Estimation of the regularization parameter for ME reconstruction. The fulfillment of the Chi-square test (33) and the S-curve test (34) are illustrated by black and white dots respectively. According to Alg. 3, the result of the Chi-square test is retained.

	$-\log_{10} \lambda_Q$	$-\log_{10} \lambda_S$	$Q(\lambda_Q)$	$Q(\lambda_S)$
ME	4.92	5.05	2.05	2.43
TIK$^+$	6.19	5.91	3.92	4.67

(a) Dataset A

	$-\log_{10} \lambda_Q$	$-\log_{10} \lambda_S$	$Q(\lambda_Q)$	$Q(\lambda_S)$
ME	5.32	5.59	13.8	22.9
TIK$^+$	5.92	5.92	10.7	10.7

(b) Dataset B

TABLE IV
REGULARIZATION PARAMETER ESTIMATES (λ_Q, λ_S) OBTAINED RESPECTIVELY BY MINIMIZING Q AND BY APPLYING THE ALGORITHM SUMMARIZED IN ALG. 3. (SNR = 10 DB AND $\tilde{m}_i = 5$)

6) Pulse angle effect: In NMR experiments, the pulse angle Φ may not be set exactly to its desired value. This uncertainty introduces a potential error in the value of γ in the observation model. Let us first discuss the effect of an inexact value of this parameter on the reconstruction results. Several reconstructions using an observation model with $\gamma \neq 1$ have been performed. Fig. 10 shows the optimal value of the regularization parameter λ and the reconstruction error Q for different values of γ, for ME and TIK$^+$ algorithms. As expected, an error on the value of γ leads to a larger reconstruction error. Moreover, a larger value of λ has to be chosen to compensate the increase of the modelization error. We can conclude that the pulse angle parameter has an influence on the reconstruction results whatever the employed inversion algorithm.

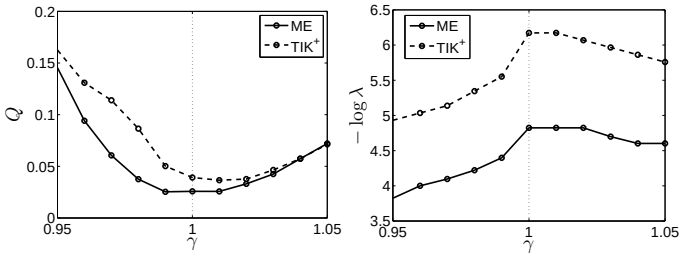
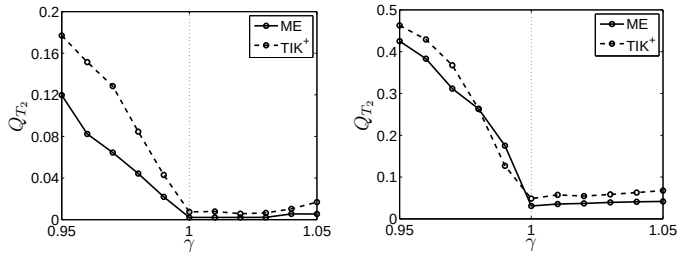


Fig. 10. Dataset A (SNR = 10 dB, $\gamma = 1$, $\tilde{m}_i = 5$): Sensitivity to a wrong estimation of γ in terms of reconstruction error Q (left) and optimal regularization parameter λ (right).



(a) Dataset A (b) Dataset B

Fig. 11. Sensitivity to a wrong estimation of γ in terms of error Q_{T_2} between the T_2 marginalized spectra and the reference T_2 spectrum.

7) Pulse angle estimation: In [47], some data preprocessing strategies are proposed to handle systematic errors, including pulse angle inaccuracy, in NMR experiments. An alternative strategy allowing to assess the pulse angle value is proposed

here. The basic idea is to use the reconstructed T_2 spectrum, obtained from T_2 relaxation data, as a reference spectrum. Since these data are obtained for high values of τ_1, the underlying spectrum is not affected by the value of γ. After performing several 2D reconstructions with different values of γ, we retain the pulse angle value maximizing the similarity between the marginalized T_2 spectrum and the reference T_2 spectrum.

Fig. 11 illustrates the relative euclidian distance Q_{T_2} between the 1D recovered T_2 spectrum and the marginalized T_2 spectra for several values of γ. The best matching is reached when γ equals its actual value, i.e. $\gamma = 1$.

B. Application to Experimental Data

Measurements have been performed on a plant matter sample (apple) to test the applicability of the proposed algorithm on experimental data. In the experiment, $m_1 = 50$ values of τ_1, non-uniformly spaced between $30\,\mathrm{ms}$ and $12\,\mathrm{s}$ were retained. In all cases, $m_2 = 10000$ echoes with a uniform time spacing of $800\,\mu\mathrm{s}$ between $600\,\mu\mathrm{s}$ and 8 s were acquired.

The proposed algorithm was applied to reconstruct a spectrum with $N_1 = N_2 = 200$ values of T_1 and T_2 relaxation times, equally spaced between 25 ms and 3 s.

1) Reconstruction algorithm tuning: The lowest computation time was reached when using only one sub-iteration of MM line search and computing the preconditioner with TSVDs at rank $v = 7$. The proposed strategy in Alg. 3 was used to set the regularization parameter.

2) Pulse angle parameter setting: Fig. 12 summarizes the reconstruction results for different values of γ between 0.9 and 1. It can be noted that the positions and the amplitudes of some peaks are highly affected by the pulse angle value. Therefore, the reconstruction of a reliable spectrum requires the use of an accurate value of this parameter. The same strategy as that proposed in subsection IV-A7 is used to set the pulse angle value. According to Fig. 13, the retained value corresponds to $\Phi \approx 85°$ (i.e., $\gamma = 0.92$).

3) Comparison of algorithms: Fig. 14 shows the reconstructed 2D spectrum for $\gamma = 0.92$. It can be noted that this spectrum allows to analyze the correlation between T_1 and T_2 relaxation times. This correlation appears, for example, in the peak located around $[T_1 = 1.4\,\mathrm{s}, T_2 = 0.9\,\mathrm{s}]$. Such information is very useful to obtain the T_1/T_2 ratio which gives insights related to the molecular structure of the analyzed sample [7]. Concerning the reconstruction algorithm performances, the computation time was $59\,\mathrm{s}$ for 67 iterations and the final value of λ was $1.3 \cdot 10^{-4}$.

Since there is no ground truth regarding the T_1-T_2 correlation spectrum of the apple, we compare the 1D distributions (T_1 and T_2) obtained by 1D inversion with the 1D distributions deduced by marginalization of the reconstructed 2D distribution. It can be noted from Fig. 15 the similarity between the 1D spectra which shows the relevance of the 2D spectrum.

We also compare these results with the ones obtained by the TIK$^+$ algorithm of [47]. This algorithm was tuned with a compression rank $\tilde{m}_i = 10$ and the same strategy as in [47] was used to determine the regularization parameter. The

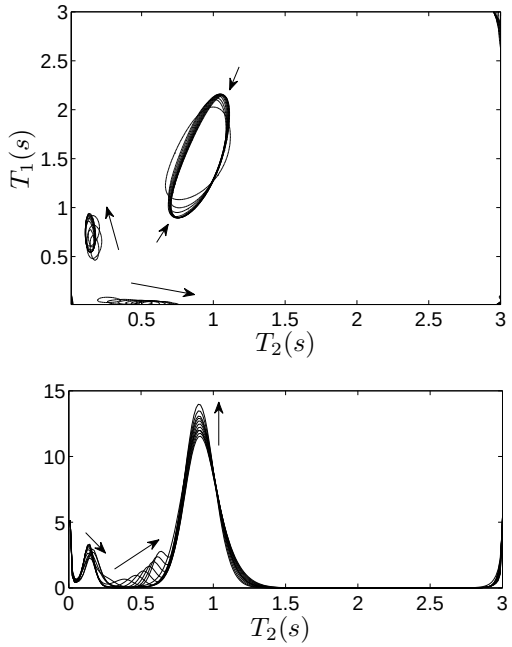


Fig. 12. 2D ME spectra (top) from experimental data and 1D distributions resulting from T_2 marginalization (bottom), for different values of the pulse angle parameter in the interval $[0.9, 1]$. The effect of increasing γ onto peak positions and amplitudes is indicated by arrows.

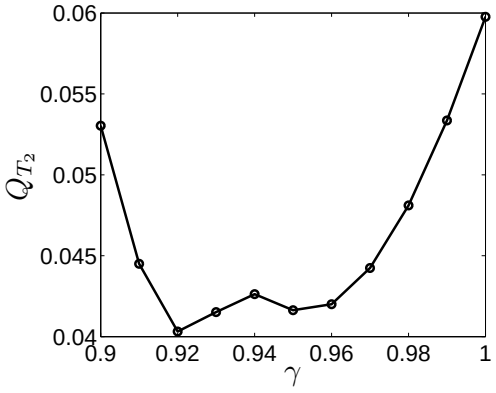


Fig. 13. Error between 1D ME T_2 reconstruction and 2D ME marginalized spectrum. The minimum is reached for $\gamma = 0.92$ which corresponds to $\Phi = 85°24'$.

algorithm requires a computation time of $11\,\mathrm{s}$ for 14 iterations and the final value of $\lambda = 2 \cdot 10^{-5}$. The reconstructed 2D spectrum and the corresponding 1D distributions are shown in Figs. 16 and 17. Even if the two reconstruction methods led to similar measurement data fit (98 %), a visual comparison reveals significant differences between the two spectra shapes in terms of regularity and amplitude.

V. Conclusion

The reconstruction of a T_1-T_2 spectrum in NMR requires a numerical inversion of a 2D Laplace transform. This is known to be an ill posed inverse problem. In this paper, we presented an efficient inversion method based on maximum entropy regularization and truncated Newton optimization. A second difficulty is related to the large scale of the 2D model. To

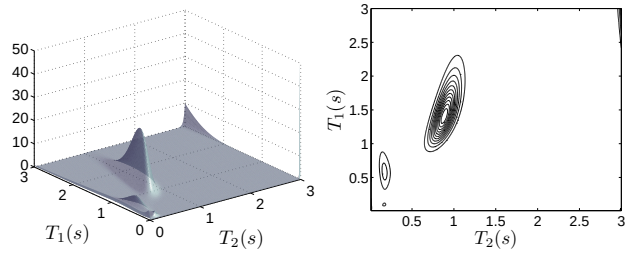
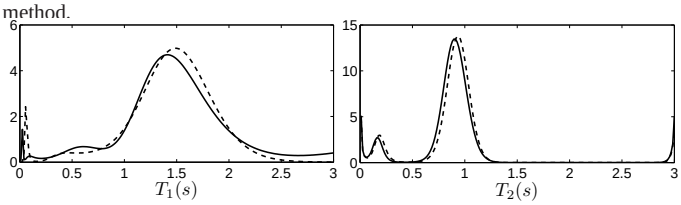


Fig. 14. Reconstructed spectrum from 2D NMR experimental data with ME method.



(a) T_1 spectrum (b) T_2 spectrum

Fig. 15. 1D distributions resulting from marginalization of the 2D ME spectrum (solid line) or 1D ME reconstruction (dashed line).

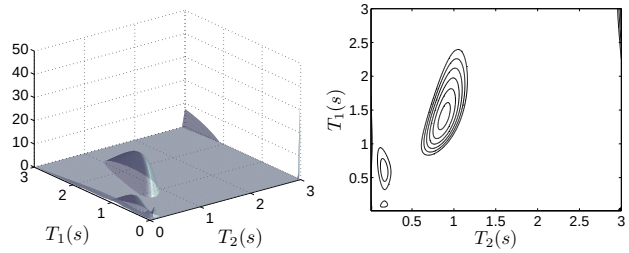
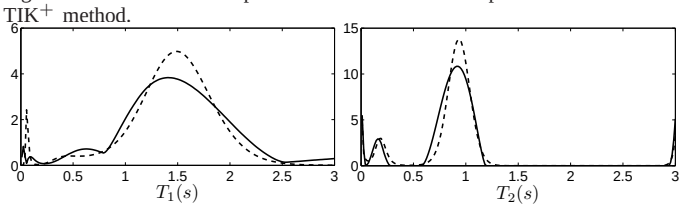


Fig. 16. Reconstructed spectrum from 2D NMR experimental data with TIK$^+$ method.



(a) T_1 spectrum (b) T_2 spectrum

Fig. 17. 1D distributions resulting from marginalization of the 2D TIK$^+$ spectrum (solid line) or 1D ME reconstruction (dashed line).

handle this problem, rather than compressing the data matrix, we rely on an exact data model thanks to an iterative algorithm exploiting the separability of the convolution kernel. All required quantities such as gradient, Hessian-vector product are computed with reduced memory storage and computation time. Moreover, since the entropy criterion introduces a barrier in the criterion to minimize, an appropriate line search strategy is used. This procedure is fast and ensures the theoretical convergence of the truncated Newton algorithm. Finally, the convergence speed of the algorithm is increased by applying an adequate preconditioner using TSVDs of the convolution kernels. The applicability of the proposed method has been demonstrated through the processing of simulated and real data and a comparison with the constrained Tikhonov approach of [15]. Our conclusion is that the two methods produce reconstructions of similar quality. The constrained Tikhonov approach is noticeably faster, at the price of resorting to a data compression step that needs the tuning of two parameters. In contrast, our approach remains fast without data compression.

The processing of real data measurements allowed us to point out the difficulty of setting the pulse angle parameter appearing in the observation model. We have shown that an inaccurate value of this parameter tends to produce a significant error in peak positions and amplitudes. Up to our knowledge, this point is only partially addressed in NMR litterature where data preprocessing strategies are suggested. Therefore, we proposed an original strategy allowing to estimate this parameter. Although this strategy seems to give satisfying results in our tests, further investigations and experiments would be needed to validate this approach. Another perspective would be to build a criterion allowing to reduce the number of peaks in the reconstructed spectrum or to propose a strategy based on a parametric 2D reconstruction where the number of peaks will be imposed.

From the methodological point of view, we restricted our analysis to the case of separable convolution kernels. However, in some NMR measurement models [7], the separability is no longer valid. It would be interesting to generalize our approach by considering the case where the observation model can be expressed as a linear superposition of several separable kernels.

APPENDIX

A. Interpretation of BRD algorithm using Legendre-Fenchel duality

Let us consider the constrained minimization problem

$$\min_{\boldsymbol{s}\geqslant 0}\left\{L(\boldsymbol{s})=\frac{1}{2}\|\boldsymbol{K}\boldsymbol{s}-\boldsymbol{y}\|^2+\frac{\lambda}{2}\|\boldsymbol{s}\|^2\right\}. \tag{36}$$

The BRD algorithm [10] is based on the equivalence between the KKT conditions of problem (36) and the following unconstrained problem

$$\min_{\boldsymbol{c}\in\mathbb{R}^m}\left\{\chi(\boldsymbol{c})=\frac{1}{2}\boldsymbol{c}^{\mathrm{t}}\left(G(\boldsymbol{c})+\lambda\boldsymbol{I}\right)\boldsymbol{c}-\boldsymbol{c}^{\mathrm{t}}\boldsymbol{y}\right\} \tag{37}$$

with the reparametrization $\boldsymbol{s}=\max(\boldsymbol{0},\boldsymbol{K}^{\mathrm{t}}\boldsymbol{c})$ and

$$G(\boldsymbol{c})=\boldsymbol{K}^{\mathrm{t}}\mathrm{Diag}(\mathbb{H}(\boldsymbol{K}^{\mathrm{t}}\boldsymbol{c}))\boldsymbol{K}, \tag{38}$$

where $\mathbb{H}$ denotes a component-wise unit step function that takes the value zero for negative or zero arguments and one for positive arguments. Let us show that this equivalence can also be obtained from the Legendre-Fenchel conjugacy theory (see [19] for a reminder on Legendre-Fenchel theory).

First, let us introduce the Legendre-Fenchel conjugate f^* of the quadratic $f(\boldsymbol{u})=\frac{1}{2}\|\boldsymbol{u}-\boldsymbol{y}\|^2$, *i.e.*,

$$f^*(\boldsymbol{u})=\sup_{\boldsymbol{v}}\left(\boldsymbol{v}^{\mathrm{t}}\boldsymbol{u}-\frac{1}{2}\|\boldsymbol{v}-\boldsymbol{y}\|^2\right)=\frac{1}{2}\|\boldsymbol{u}\|^2+\boldsymbol{y}^{\mathrm{t}}\boldsymbol{u}. \tag{39}$$

According to the conjugacy theorem [19, Prop. 7.1.1],

$$L(\boldsymbol{s})=\sup_{\boldsymbol{u}\in\mathbb{R}^m}\left(\boldsymbol{s}^{\mathrm{t}}\boldsymbol{K}^{\mathrm{t}}\boldsymbol{u}-f^*(\boldsymbol{u})\right)+\frac{\lambda}{2}\|\boldsymbol{s}\|^2. \tag{40}$$

Moreover, according to the minimax theorem [19, Prop. 2.6.2], (40) implies

$$\begin{aligned}\min_{\boldsymbol{s}\geqslant 0} L(\boldsymbol{s}) &= \max_{\boldsymbol{u}\in\mathbb{R}^m} \min_{\boldsymbol{s}\geqslant 0} \left(\boldsymbol{s}^{\mathrm{t}}\boldsymbol{K}^{\mathrm{t}}\boldsymbol{u} - f^*(\boldsymbol{u}) + \frac{\lambda}{2}\|\boldsymbol{s}\|^2 \right), \\ &= \max_{\boldsymbol{u}\in\mathbb{R}^m} \left(\varphi(\boldsymbol{u}) - f^*(\boldsymbol{u})\right) \end{aligned} \tag{41}$$

where

$$\varphi(\boldsymbol{u}) = \min_{\boldsymbol{s}\geqslant 0} \left(\boldsymbol{s}^{\mathrm{t}}\boldsymbol{K}^{\mathrm{t}}\boldsymbol{u} + \frac{\lambda}{2}\|\boldsymbol{s}\|^2 \right). \tag{42}$$

The minimization problem (42) is convex, separable and the following expression of the minimizer is easy to derive

$$\boldsymbol{s}^*(\boldsymbol{u}) = \frac{1}{\lambda}\max\left(\boldsymbol{0}, -\boldsymbol{K}^{\mathrm{t}}\boldsymbol{u}\right) \tag{43}$$

where $\max$ is to be considered component-wise. Moreover, we have

$$\varphi(\boldsymbol{u}) = \left(\boldsymbol{s}^*(\boldsymbol{u})\right)^{\mathrm{t}} \boldsymbol{K}^{\mathrm{t}}\boldsymbol{u} + \frac{\lambda}{2}\left\|\boldsymbol{s}^*(\boldsymbol{u})\right\|^2 = \frac{1}{2}\left(\boldsymbol{s}^*(\boldsymbol{u})\right)^{\mathrm{t}} \boldsymbol{K}^{\mathrm{t}}\boldsymbol{u}, \tag{44}$$

the latter expression being a consequence of $(\max(0,x))^2 = x\max(0,x)$ for all $x\in\mathbb{R}$. Finally, given (39), (43) and (44), (41) also reads

$$\begin{aligned}\min_{\boldsymbol{s}\geqslant 0} L(\boldsymbol{s}) = \max_{\boldsymbol{u}\in\mathbb{R}^m} \Big(&-\frac{1}{2\lambda}\left(\max\left(\boldsymbol{0}, -\boldsymbol{K}^{\mathrm{t}}\boldsymbol{u}\right)\right)^{\mathrm{t}} \boldsymbol{K}^{\mathrm{t}}\boldsymbol{u} \\ &+ \frac{1}{2}\|\boldsymbol{u}\|^2 + \boldsymbol{y}^{\mathrm{t}}\boldsymbol{u}\Big) \\ = -\lambda \min_{\boldsymbol{c}\in\mathbb{R}^m} &\chi(\boldsymbol{c})\end{aligned}$$

where the last identity is obtained using the change of variable $\boldsymbol{c} = -\boldsymbol{u}/\lambda$. Thus, (36) and (37) are equivalent through Legendre-Fenchel duality, and $\boldsymbol{c}^*$ minimizes $\chi(\boldsymbol{c})$ in $\mathbb{R}^m$ if and only if $\boldsymbol{s}^* = \max(\boldsymbol{0}, \boldsymbol{K}^{\mathrm{t}}\boldsymbol{c}^*)$ minimizes $\boldsymbol{L}(\boldsymbol{s})$ in $\mathbb{R}^m_+$.

B. Expression of the majorant function $h^j(\cdot,\alpha^j)$ and of its minimizer

The majorant function $h^j(\cdot,\alpha^j)$ is piecewise defined, whether $\alpha\in(\alpha_-;\alpha^j]$ or $\alpha\in[\alpha^j;\alpha_+)$. In both cases, it takes the following form

$$\begin{aligned}h^j(\alpha,\alpha^j) = &\ell(\alpha^j) + (\alpha-\alpha^j)\dot{\ell}(\alpha^j) + \frac{1}{2}m^j(\alpha-\alpha^j)^2 \\ &+ \gamma^j\left[(\bar{\alpha}^j-\alpha^j)\log\frac{\bar{\alpha}^j-\alpha^j}{\bar{\alpha}^j-\alpha} - \alpha + \alpha^j\right]\end{aligned} \tag{45}$$

while the expressions of parameters $\bar{\alpha}^j$, m^j, and γ^j are specific to each case. The notation $\dot{\ell}$ refers to the derivative of ℓ, also defined as $\dot{\ell}(\alpha) = \boldsymbol{d}_k^T\nabla L(\boldsymbol{s}_k+\alpha\boldsymbol{d}_k)$.

1) Case $\alpha\in(\alpha_-;\alpha^j]$:

$$\begin{cases}\bar{\alpha}^j = \alpha_- \\ m^j = \boldsymbol{d}_k^{\mathrm{t}}\boldsymbol{K}^{\mathrm{t}}\boldsymbol{K}\boldsymbol{d}_k + \lambda\sum_{i|d_{k,i}<0}\phi_i(\alpha^j) \\ \gamma^j = \lambda(\alpha_- - \alpha^j)\sum_{i|d_{k,i}>0}\phi_i(\alpha^j)\end{cases} \tag{46}$$

2) Case $\alpha\in[\alpha^j;\alpha_+)$:

$$\begin{cases}\bar{\alpha}^j = \alpha_+ \\ m^j = \boldsymbol{d}_k^{\mathrm{t}}\boldsymbol{K}^{\mathrm{t}}\boldsymbol{K}\boldsymbol{d}_k + \lambda\sum_{i|d_{k,i}>0}\phi_i(\alpha^j) \\ \gamma^j = \lambda(\alpha_+ - \alpha^j)\sum_{i|d_{k,i}<0}\phi_i(\alpha^j)\end{cases} \tag{47}$$

where $\phi_i(\alpha) = d_{k,i}^2/(s_i+\alpha d_{k,i})$ in both cases.

The minimizer of $h^j(\cdot,\alpha^j)$ can be expressed as follows

$$\alpha^j + \operatorname{sign}(-\dot{\ell}(\alpha^j))\frac{2\left|A_3\right|}{\left|A_2\right| + \sqrt{A_2^2 - 4A_1A_3}}, \tag{48}$$

with

$$\begin{cases}A_1 = -m^j \\ A_2 = \gamma^j - \dot{\ell}(\alpha^j) + m^j(\bar{\alpha}^j - \alpha^j) \\ A_3 = (\bar{\alpha}^j - \alpha^j)\dot{\ell}(\alpha^j)\end{cases}. \tag{49}$$

REFERENCES

[1] D. Canet, J.-C. Boubel, and E. Canet-Soulas, *La RMN. Concepts, méthodes et applications*, 2nd ed. Paris, France: Dunod, 2002.

[2] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, 2nd ed., ser. International Series of Monographs on Chemistry. Oxford, NY: Oxford University Press, 1997.

[3] F. J. M. Van de Ven, *Multidimensional NMR in Liquids. Basic Principles and Experimental Methods*. New York, NY: Wiley-VCH, 1995.

[4] F. Mariette, J. P. Guillement, C. Tellier, and P. Marchal, "Continuous relaxation time distribution decomposition by MEM," *Signal Treat. and Signal Anal. in NMR*, pp. 218–234, 1996.

[5] R. Lamanna, "On the inversion of multicomponent NMR relaxation and diffusion decays in heterogeneous systems," *Concepts Magn. Reson. Part A*, vol. 26A, no. 2, pp. 78–90, 2005.

[6] A. E. English, K. P. Whittall, M. L. G. Joy, and R. M. Henkelman, "Quantitative two-dimensional time correlation relaxometry," *Magn. Reson. Med.*, vol. 22, pp. 425–434, 1991.

[7] M. D. Hürlimann and L. Venkataramanan, "Quantitative measurement of two-dimensional distribution functions of diffusion and relaxation in grossly inhomogeneous fields," *J. Magn. Reson.* , vol. 157, pp. 31–42, 2002.

[8] M. Fleury and J. Soualem, "Quantitative analysis of diffusional pore coupling from T_2-store-T_2 NMR experiments," *J. Colloid Interf. Sci.*, vol. 336, pp. 250–259, 2009.

[9] E. Sternin, "Use of inverse theory algorithms in the analysis of biomembrane NMR data," in *Methods in Membrane Lipids*, H. Press, Ed., 2008, vol. 400, pp. 103–125.

[10] J. P. Butler, J. A. Reeds, and S. V. Dawson, "Estimating solutions of first kind integral equations with nonnegative constraints and optimal smoothing," *SIAM J. Num. Anal.*, vol. 18, no. 3, pp. 381–397, June 1981.

[11] E. D. Laue, J. Skilling, J. Staunton, S. Sibisi, and R. G. Brereton, "Maximum entropy method in nuclear magnetic resonance spectroscopy," *J. Magn. Reson.* , vol. 62, no. 3, pp. 437–452, 1985.

[12] J. Skilling and R. K. Bryan, "Maximum entropy image reconstruction: General algorithm," *Month. Not. Roy. Astr. Soc.*, vol. 211, pp. 111–124, 1984.

[13] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM J. Appl. Mathematics*, vol. 11, pp. 431–441, 1963.

[14] G. H. Golub and C. F. Van Loan, *Matrix computations*, 3rd ed. Baltimore: John Hopkins University Press, 1996.

[15] L. Venkataramanan, Y. Q. Song, and M. D. Hürlimann, "Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1017–1026, 2002.

[16] G. Bruckner and J. Cheng, "Tikhonov regularization for an integral equation of the first kind with logarithmic kernel," Weierstrass Institute for Applied Analysis and Stochastics, Tech. Rep., 1998. [Online]. Available: http://www.wias-berlin.de/publications/preprints/523/wias_preprints_523.pdf

[17] Y.-W. Chiang, P. P. Borbat, and J. H. Freed, "Maximum entropy: A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR," *J. Magn. Reson.* , vol. 177, no. 2, pp. 184–196, December 2005.

[18] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY: Springer-Verlag, 1999.
[19] D. P. Bertsekas, *Convex analysis and optimization*, 1st ed. Belmont, MA: Athena Scientific, 2003.
[20] P. P. B. Eggermont, "Maximum entropy regularization for Fredholm integral equations of the first kind," *SIAM J. Math. Anal.*, vol. 24, no. 6, pp. 1557–1576, 1993.
[21] R. Gordon, R. Bender, and G. T. Herman, "Algebric reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *J. Theor. Biol.*, vol. 29, pp. 471–481, 1970.
[22] C. A. Johnson and D. McGarry, "Maximum entropy reconstruction methods in electron paramagnetic resonance imaging," *Ann. Oper. Res.*, vol. 119, pp. 101–118, 2003.
[23] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York: Cambridge Univ. Press, 1992.
[24] E. Chouzenoux, S. Moussaoui, J. Idier, and F. Mariette, "Reconstruction d'un spectre RMN 2D par maximum d'entropie," in *GRETSI*, Dijon, , France, September 2009.
[25] ——, "Optimization of a maximum entropy criterion for 2D nuclear magnetic resonance reconstruction," in *ICASSP*, Dallas, Texas, March 2010.
[26] R. Dembo, S. C. Eisenstat, and S. Steihaug, "Inexact Newton methods," *SIAM J. Num. Anal.*, vol. 19, no. 2, pp. 400–408, April 1982.
[27] S. G. Nash, "A survey of truncated-Newton methods," *J. Comput. Appl. Math.*, vol. 124, pp. 45–59, 2000.
[28] S. G. Nash and A. Sofer, "On the complexity of a practical interior-point method," *SIAM J. Optimization*, vol. 8, no. 3, pp. 833–849, 1998.
[29] S. Bellavia, "Inexact interior-point method," *J. Optim. Theory Appl.*, vol. 96, pp. 109–121, 1998.
[30] J. J. Moré and D. J. Thuente, "Line search algorithms with guaranteed sufficient decrease," *ACM Trans. on Math. Soft.*, vol. 20, no. 3, pp. 286–307, 1994.
[31] J. Sun and J. Zhang, "Global convergence of conjugate gradient methods without line search," *Ann. Oper. Res.*, vol. 103, pp. 161–173, March 2001.
[32] C. Labat and J. Idier, "Convergence of conjugate gradient methods with a closed-form stepsize formula," *J. Optim. Theory Appl.*, vol. 136, no. 1, pp. 43–60, January 2008.
[33] E. Chouzenoux, S. Moussaoui, and J. Idier, "A majorize-minimize line search algorithm for barrier function optimization," in *EUSIPCO*, Glasgow, UK, August 2009.
[34] ——, "A new line search method for barrier functions with strong convergence properties," IRCCyN, Tech. Rep., 2009, http://hal.archives-ouvertes.fr/IRCCYN-ADTSI.
[35] D. R. Hunter and L. K., "A tutorial on MM algorithms," *Amer. Statist.*, vol. 58, no. 1, pp. 30–37, February 2004.
[36] M. Jacobson and J. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Trans. Image Processing*, vol. 16, no. 10, pp. 2411–2422, October 2007.
[37] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.
[38] Z.-J. Shi, "Convergence of line search methods for unconstrained optimization," *Appl. Math. and Comp.*, vol. 157, pp. 393–405, 2004.
[39] P. C. Hansen, *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Philadelphia, PA, USA: SIAM, 1998.
[40] Y. Saad, *Iterative methods for sparse linear systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2003.
[41] K. Chen, *Matrix preconditioning techniques and applications*, 1st ed. Cambridge, UK: Cambridge University Press, 2005.
[42] R. Piché, "Regularization operators for multidimensional inverse problems with kronecker product structure," in *ECCOMAS*, Jyväskylä, Finland, July 2004.
[43] C. Eckart and G. Young, "Multiplicative iterative algorithms for convex programming," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.
[44] C. Pichon and E. Thiébaut, "Non-parametric recontruction of distribution functions from observed galactic discs," *Month. Not. Roy. Astr. Soc.*, vol. 301, no. 2, pp. 419–434, 1998.
[45] H. Trussell, "Convergence criteria for iterative restoration methods," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 31, no. 1, pp. 129–136, 1983.
[46] N. Galatsanos and A. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Image Processing*, vol. 1, no. 3, pp. 322–336, July 1992.
[47] Y. Q. Song, L. Venkataramanan, M. D. Hürlimann, M. Flaum, P. Frulla, and C. Straley, "T1-T2 correlation spectra obtained using a fast two-dimensional Laplace inversion," *J. Magn. Reson.* , vol. 154, pp. 261–268, 2002.

Émilie Chouzenoux received the M. Sc. degree in signal processing and the engineering degree from École Centrale, Nantes, France in 2007. She is currently working towards the Ph. D. degree in signal processing at the Institut de Recherche en Communications et Cybernétique in Nantes (IRCCYN, UMR CNRS 6597). Her research interests are in optimization algorithms for large scale problems of image and signal reconstruction.

Saïd Moussaoui received the State engineering degree from Ecole Nationale Polytechnique, Algiers, Algeria, in 2001, and, in 2005, the Ph.D. degree in Automatic Control and Signal Processing from Université Henri Poincaré, Nancy, France. He is currently Assistant Professor at Ecole Centrale de Nantes. Since september 2006, he is with the Institut de Recherche en Communications et Cybernétique in Nantes (IRCCYN, UMR CNRS 6597). His research interests are in statistical signal and image processing including source separation, Bayesian estimation and their applications.

Jérôme Idier was born in France in 1966. He received the diploma degree in electrical engineering from École Supérieure d'Électricité, Gif-sur-Yvette, France, in 1988 and the Ph.D. degree in physics from University of Paris-Sud, Orsay, France, in 1991. Since 1991, he joined the Centre National de la Recherche Scientifique. He is currently a Senior Researcher at the Institut de Recherche en Communications et Cybernétique in Nantes. His major scientific interests are in probabilistic approaches to inverse problems for signal and image processing. He is serving as an Associate Editor for the IEEE Transactions on Signal Processing and for the Journal of Electronic Imaging, copublished by SPIE and IS&T.

François Mariette received the Ph.D. degree in Physical Chemistry from the University of Nantes, France in 1992. From 1993 to 1999 he was researcher in the Food Process Engineering research unit at the Cemagref research institute. In 2001 he was visitor scientist in the Physical Chemistry 1 Department at the Lund University, Sweden. Since 2004 he is Research Director and Leader of the NMR/MRI research team in the Food Process Engineering research unit at Cemagref. Since 2006, he is the coordinator of the national research platform PRISM (Structural and metabolic imaging and spectroscopy research platform). PRISM is a research platform located in Rennes under the responsibility of the University of Rennes 1, INRA and Cemagref. F. Mariette is permanent member of the scientific committee of a French association of dairy companies. He is author of 83 papers in scientific journals and chapters in textbook, and of 66 presentations to international conference in the field of NMR and MRI applied to food science and food processing.