Published in Image Processing On Line on 2023–03–21.
Submitted on 2022–06–30, accepted on 2022–10–13.
ISSN 2105–1232 © 2023 IPOL & the authors CC–BY–NC–SA
This article is available online with supplementary materials,
software, datasets and online demo at
https://doi.org/10.5201/ipol.2023.414

Electron Paramagnetic Resonance Image Reconstruction with Total Variation Regularization

Rémy Abergel[1], Mehdi Boussâa[1], Sylvain Durand[1], Yves-Michel Frapart[2]

[1]Université Paris Cité, CNRS, MAP5, Paris, France
[2]Université Paris Cité, CNRS, LCBPT, Paris, France
{Remy.Abergel, Mehdi.Boussaa, Sylvain.Durand, Yves.Frapart}@u-paris.fr

Communicated by Sung Ha Kang and Jérémy Anger *Demo edited by* Jérémy Anger

Abstract

This work focuses on the reconstruction of two and three dimensional images of the concentration of paramagnetic species from electron paramagnetic resonance (EPR) measurements. A direct operator, modeling how the measurements are related to the paramagnetic sample to be imaged, is derived in the continuous framework taking into account the physical phenomena at work during the acquisition process. Then, this direct operator is discretized to closely take into account the discrete nature of the measurements and provide an explicit link between them and the discrete image to be reconstructed. A variational inverse problem with total variation regularization is formulated and an efficient resolvant scheme is implemented. The setting of the reconstruction parameters is thoroughly studied and facilitated thanks to the introduction of appropriate normalization factors. Moreover, an a contrario algorithm is proposed to derive the optimal resolution at which the data should be acquired. Finally, an in-depth experimental study over real EPR datasets is done to illustrate the potential and limitations of the presented image reconstruction model.

Source Code

The reviewed source code and documentation for the algorithms presented in this article are available from the web page of this article[1]. Installation and usage instructions are included in the `README.org` file of the archive.

Supplementary Material

Some EPR datasets, to be used for further comparisons, are provided with the source code of the article and peer-reviewed. A Matlab graphical user interface (not reviewed and compatible with Matlab version 2019b or later versions) for EPR image reconstruction is also available for the user convenience.

Keywords: electron paramagnetic resonance imaging; total variation; variational models; inverse problems; Shannon sampling theory

---

[1] https://doi.org/10.5201/ipol.2023.414

Rémy Abergel, Mehdi Boussâa, Sylvain Durand, Yves-Michel Frapart, *Electron Paramagnetic Resonance Image Reconstruction with Total Variation Regularization*, Image Processing On Line, 13 (2023), pp. 90–139. https://doi.org/10.5201/ipol.2023.414

1 Introduction

Electron Paramagnetic Resonance is a powerful method for exploring paramagnetic species. It is based on the property of single electrons to absorb the energy of electromagnetic radiation when placed in a magnetic field. This process of absorption of energy is characteristic of the observed species. The acquired spectral signatures allow therefore to discriminate between different paramagnetic species. As in Nuclear Magnetic Resonance (NMR), the EPR phenomenon can be used for imaging purposes. By adding to the homogeneous magnetic field a field gradient whose orientation evolves in a plane (2D imaging) or in space (3D imaging), a sinogram is generated. Its formation can be mathematically modeled by the convolution between the Radon transform of the concentration of the present paramagnetic species and its spectral signature. The reconstruction of the concentration image from the sinogram is an ill-posed inverse problem which is the main focus of this article.

The standard method of reconstruction is made on two steps: a Wiener-like inverse filtering followed by a back-projection. This approach, that is taken by Xepr©[2] (a commercial software distributed by Bruker®), has however several drawbacks, including the well known "star effect" of the back-projection when there are too few projections. Moreover, the two-step processing leads to accumulating errors and involves the tuning of many parameters.

Several variational methods have been reported. As the sinogram is corrupted by an additive Gaussian white noise, they mostly consist in least squares with various regularizations such as entropy [18] or Tikhonov [40]. Other considered priors are built on the total variation (TV) as it allows to obtain images with sharp edges. In [19], the prior is the sum of the L^1-norm and the TV of the sought image. The minimization is done by a double-loop algorithm whose fixed point fails however to be the sought minimizer. This model is generalized in [2] where the gradient in TV is changed into a linear operator. The data fidelity is moreover expressed by a hard constraint. However, the numeral scheme is not given. In [35], a hard constraint is also used on the data fidelity, but for a noiseless model only. TV is smoothed as in [41] in order to perform a Steepest Descent.

Great improvements have been made in the minimization of functions involving TV thanks to the introduction of a fully proximal primal-dual algorithm by Chambolle and Pock [5]. It is applied to TV-reguralized least-squares in the context of EPR imaging in [10]. The inf-convolution of TV and the L^1-norm of weighted curvelet coefficients is also used as an alternative prior in order to obtain images with fine textures and sharp edges. To avoid a recursive use of the computationally expensive Radon transform and thus further speed up the algorithm, a Toeplitz hypermatrix structure is designed for the data fidelity as is done in [25] for Computed Tomography. Chambolle-Pock Algorithm is also used in [33, 34] where the data fidelity or the TV-prior is expressed under a hard constraint. This hard constraint on the data fidelity eases the choice of the parameter that controls the balance between data fidelity and regularization when the standard deviation of the noise is known. But it does not allow to make use of the Toeplitz structure alluded to above which is very useful to obtain a fast algorithm and thus efficiently process 3D images. This problem of 3D images is addressed in [24] where an algebraic method is developed to quickly reconstruct images from a large number of projections.

This article focuses on the method developed in [10]. In addition to making accessible the algorithm introduced therein, we propose several improvements. First, we give its 3D extension. As there is no curvelet transform for 3D images, only the TV-prior is considered. But an optional Huber-TV-prior is proposed to process textured images. Moreover, in order not to change the Toeplitz hypermatrix into a circulant Hypermatrix, the primal-dual algorithm of [5] is changed into its forward-backward version introduced in [7]. At last, an a contrario method is introduced to find the optimal step at which the sinogram must be sampled and thus speed up the acquisition process.

---

[2] Bruker, Xepr version 2.6b.36, 2009, https://www.bruker.com/de/products-and-solutions/mr/epr-instruments/epr-research-instruments/xepr-software.html.

2 Continuous Modeling of the EPR Acquisition System

2.1 Spectroscopy by EPR

The most widespread application of EPR is *spectroscopy*. The aim of EPR spectroscopy is to analyze the electronic structure of the paramagnetic chemical species contained in a sample through the microwave absorption occurring when the sample is exposed to a magnetic field. By applying a so-called *homogeneous* magnetic field with increasing intensity $t \mapsto B(t)$ in the cavity of the acquisition system, we excite the single electrons of the paramagnetic species present in the cavity. When the microwave energy applied to the sample equals the energy gap between the spin state energy levels of the single electrons, the latter change their levels of energies. This phenomenon is called *electron resonance* and comes with an absorption by the single electrons of the incident microwave energy. We denote by $t \mapsto \mathcal{E}_{\text{abs}}(B(t))$ the energy absorption profile, or also, by dropping the temporal variable,

$$B \mapsto \mathcal{E}_{\text{abs}}(B)\,.$$

The levels of energies of the single electrons, as well as the energy absorption profile $\mathcal{E}_{\text{abs}}$ involved by the electron resonance phenomenon, are intrinsically linked to the electronic organization of the paramagnetic species and can be used to characterize them. In practice we are able to measure the first derivative of the energy absorption profile $\mathcal{E}_{\text{abs}}$, this function $S : B \mapsto \mathcal{E}'_{\text{abs}}(B)$ is called a *spectrum*. Each paramagnetic component X has its own spectral response, denoted by h_X and called the *reference spectrum* of X. In presence of an amount Q_X of a paramagnetic component X in the cavity, the measured spectrum S is proportional to h_X, and we simply have

$$\forall B \in \mathbb{R}_+\,, \quad S(B) = Q_X \cdot h_X(B)\,. \tag{1}$$

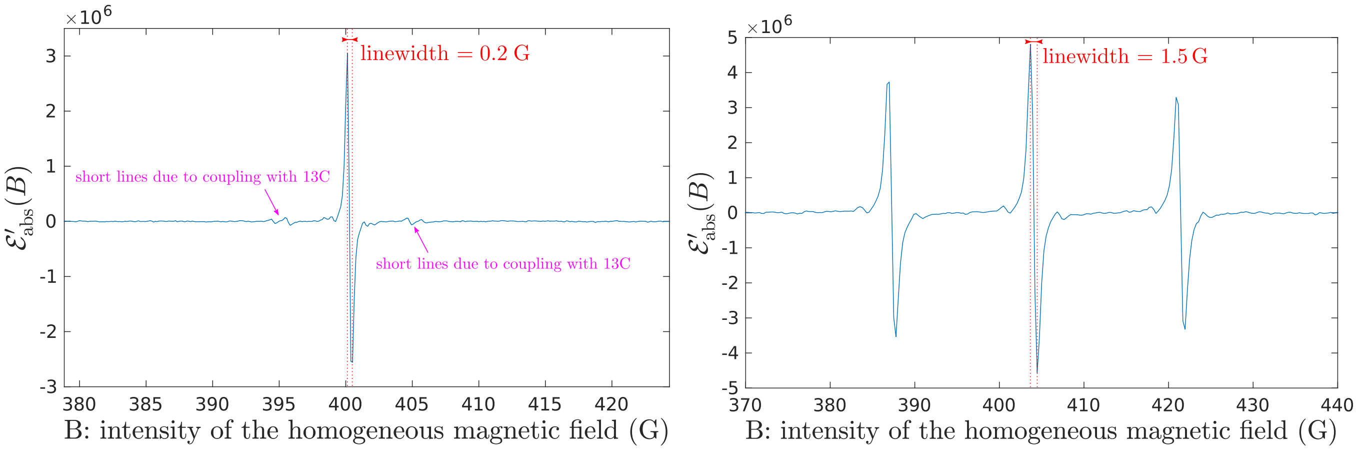
In Figure 1, we display the measured spectra of two paramagnetic species, the Tetrathiatriarylmethyl (TAM) and the 4-hydroxy-2,2,6,6-tetramethylpiperidin-1-oxyl (4OH-TEMPO). By visual inspection, and potentially with the help of EPR simulation algorithms (such as [38]), EPR spectroscopists are able to recognize a paramagnetic component, to estimate its amounts Q_X, and sometimes to measure some features related to the paramagnetic species interactions within the whole molecular environment (such as oxygen concentration, molecule rotation speed,...) [12, 13, 29, 39].

2.2 Spatialization of the EPR Measurements for Imaging Applications

The acquisition of EPR spectra like (1) makes the analysis of the paramagnetic species contained in a sample possible. However, such kind of signals do not provide any information about the spatial repartition of the different species within the sample. Indeed, at any given time t of the acquisition, the intensity of the homogeneous magnetic field $B(t)$ is the same within the whole cavity. Therefore, no spatial information is embedded within the measurement $t \mapsto S(B(t))$. Interestingly enough, the EPR measurements can be spatialized in a simple but yet ingenious way. Let us denote by $\Omega_c \subset \mathbb{R}^3$ the cavity domain, let $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi]$ denote a couple of polar and azimuthal angles in the spherical coordinate system, and let $e_{(\theta,\varphi)} = (\cos\theta \sin\varphi, \sin\theta \sin\varphi, \cos\varphi)^t$ (see Figure 2). Using magnetic coils, we can add to the homogeneous magnetic field a *field gradient* in the direction $e_{(\theta,\varphi)}$ which is an oriented magnetic field with linearly increasing intensity. Thanks to this field gradient, the intensity of the magnetic field in the cavity at the spatial location $x \in \Omega_c$ and time $t \geq 0$ is given by

$$B_{\mu,\theta,\varphi}(x, t) = B(t) + \langle \mu\, e_{(\theta,\varphi)}\,,\, x \rangle\,, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product and $\mu > 0$ represents the intensity of the field gradient. Therefore, in presence of a single paramagnetic specie X in the cavity, denoting by $U_X(x)$



(a) spectrum of a solution of TAM

(b) spectrum of solid sample soaked with 4OH-TEMPO

Figure 1: **EPR spectra of TAM and 4OH-TEMPO.** Left (a): measured EPR spectrum of a sample containing an aqueous solution of TAM. This paramagnetic component is dominated by carbon-12. Its spectrum is mainly made of one *line* (the central peak in (a)) with very short *linewidth* (0.2 G, as indicated in red in (a)). We can also resolve small lines due to the coupling with carbon-13 (see arrows in (a)), those lines have small amplitude due to the small natural abundance of 13C (around 1% of natural abundance). Right (b): measured EPR spectrum of a solid sample soaked with 4OH-TEMPO. This paramagnetic component is characteristic of the nitroxide class of radicals, its reference spectrum exhibits three main lines which are due to the coupling of the nitrogen-14 (14N) nuclear spin with the electronic spin. The three lines of 4OH-TEMPO molecules in fluid sample should have the same shape (same amplitude and linewidth). In practice, they exhibit important variation of shapes according to the rotation speed of the molecule within the sample. A fluid sample, allowing fast molecule rotation (with rotation period less than the nanosecond) leads to three identical lines, while slower rotation speed in solid samples (as here, with a rotation period close to the microsecond) leads to distorted lines with different shapes.

the quantity of X at any location $x \in \Omega_c$ and by integrating over the whole cavity domain, we obtain the measurement

$$\forall t \geq 0\,, \quad S_{\theta,\varphi}(B(t)) = \int_{\Omega_c} U_X(x)\, h_X(B(t) + \langle \mu\, e_{(\theta,\varphi)}\,,\, x\rangle)\, dx\,, \tag{3}$$

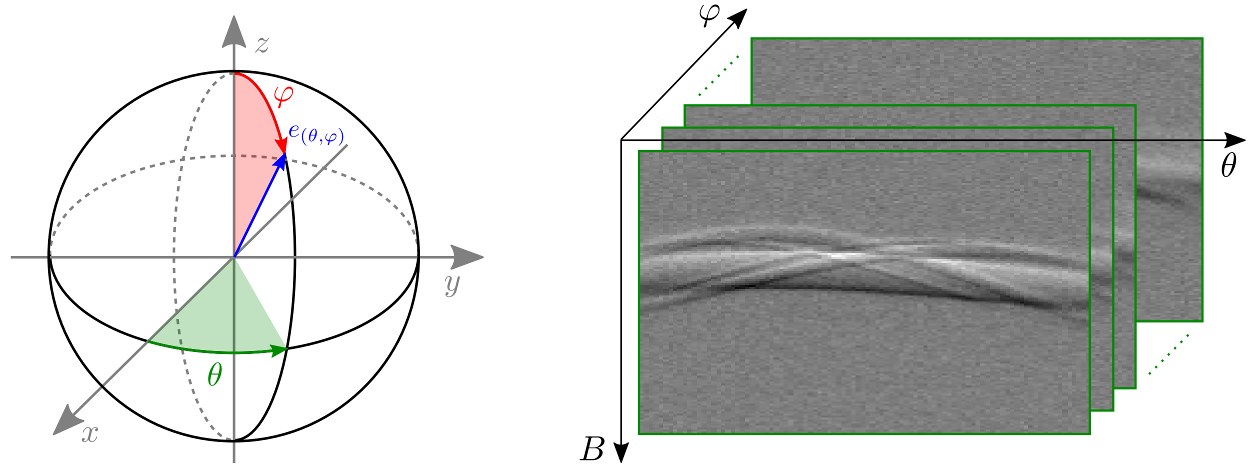
or, by dropping again the temporal variable,

$$\forall B \in \mathbb{R}_+\,, \quad S_{\theta,\varphi}(B) = \int_{\Omega_c} U_X(x)\, h_X(B + \langle \mu\, e_{(\theta,\varphi)}\,,\, x\rangle)\, dx\,. \tag{4}$$

The function $B \mapsto S_{\theta,\varphi}(B)$ is called the *projection in the direction* (θ, φ). We remark that whatever the value of (θ, φ), in absence of field gradient (i.e. when $\mu = 0$), the projection (4) boils down to (1) since we have $Q_X = \int_{\Omega_c} U_X(x)\, dx$. In this situation, the estimation of U_X is hopeless because the problem is too ill-posed. However, when $\mu \neq 0$, the estimation of U_X becomes possible provided that enough measurements $(S_{\theta,\varphi}(B))_{B,\theta,\varphi}$ are available [31].

By extending h_X (over $\mathbb{R} \setminus \mathbb{R}_+$) and U_X (over $\mathbb{R}^3 \setminus \Omega_c$) by zero and decomposing the space $\mathbb{R}^3$ as the sum of the monodimensional space $\mathrm{Span}(e_{(\theta,\varphi)})$ and its supplementary orthogonal hyperplane $\mathcal{H}_{\theta,\varphi} := \mathrm{Span}(e^{\perp}_{(\theta,\varphi)})$, we can write (4) as

$$\begin{aligned}
\forall B \in \mathbb{R}\,, \quad S_{\theta,\varphi}(B) &= \int_{\mathbb{R}} \int_{\mathcal{H}_{\theta,\varphi}} U_X(r\, e_{(\theta,\varphi)} + s)\, h_X(B + \mu r)\, ds\, dr \\
&= \int_{\mathbb{R}} \left(\int_{\mathcal{H}_{\theta,\varphi}} U_X(r\, e_{(\theta,\varphi)} + s)\, ds \right) h_X(B + \mu r)\, dr \\
&= \int_{\mathbb{R}} \mathcal{R}_{\theta,\varphi}(U_X)(r)\, h_X(B + \mu r)\, dr\,,
\end{aligned}$$



(a) polar and azimuthal angles θ and φ (b) example of EPR sinogram

Figure 2: **Spherical coordinate system and example of EPR sinogram.** We represent in (a) the polar and azimuthal angle (θ, φ) that we use to define the orientation of the field gradient (this orientation is represented by the blue unit vector $e_{(\theta,\varphi)}$). By construction in (2), at any acquisition time $t \geq 0$, the intensity of the magnetic field in the cavity is constant over any affine hyperplane orthogonal to $e_{(\theta,\varphi)}$ and its value evolves linearly along the $e_{(\theta,\varphi)}$ axis. We display in (b) an example of 3D sinogram acquired using our EPR imaging system.

where $\mathcal{R}_{\theta,\varphi}(U_X) = r \mapsto \int_{\mathcal{H}_{\theta,\varphi}} U_X(r\, e_{(\theta,\varphi)} + s)\, ds$ is the so-called *Radon transform* of the signal U_X in the direction (θ, φ). Then, setting $h_X^\mu(s) = h_X(-\mu s)$ for all $s \in \mathbb{R}$, we finally get

$$\forall B \in \mathbb{R}\,, \quad S_{\theta,\varphi}(B) = \left(h_X^\mu * \mathcal{R}_{\theta,\varphi}(U_X)\right)(-B/\mu)\ , \tag{5}$$

so the projection $S_{\theta,\varphi}$ is modeled as the convolution between the dilated reference spectrum h_X^μ and the Radon transform of U_X in the direction (θ, φ). The 3D signal $(B, \theta, \varphi) \mapsto S_{\theta,\varphi}(B)$ made of all the projections is called a *3D sinogram*. An example of 3D sinogram is displayed in Figure 2 (b).

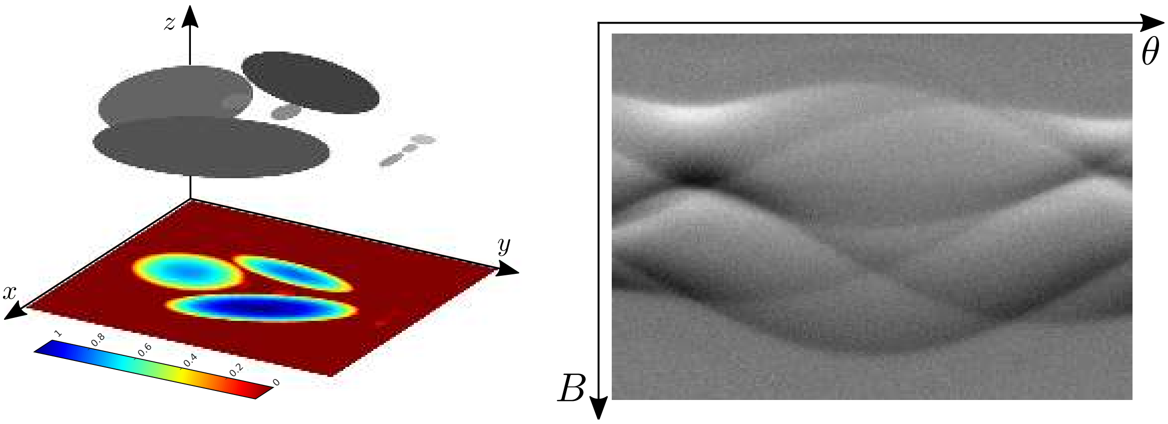
2.3 The 2D EPR Imaging Framework

The acquisition of a 3D sinogram like that displayed in Figure 2 (b) can be too long (several minutes per projection, i.e., per value of (θ, φ)) with respect to the targeted application. This is particularly the case for in-vivo imaging or for the study of species or chemical reactions with fast kinetic. In such situation, we may be interested in computing a 2D sinogram by keeping one angle (θ or φ) constant and letting the other vary. For sake of clarity, let us assume that we take $\varphi = \frac{\pi}{2}$ and keep θ variable. In this situation, the field gradient direction $e_{(\theta,\pi/2)} = (\cos\theta, \sin\theta, 0)^t$ lies in the XY-plane (see Figure 2). Therefore, the measured projections $S_\theta := B \mapsto S_{\theta,\pi/2}(B)$ satisfy, for all $B \in \mathbb{R}$ and all $\theta \in [0, \pi]$,

$$\begin{aligned} S_\theta(B) &= \int_{\mathbb{R}^3} U_X(x,y,z)\, h_X(B + \langle\, \mu\, e_{(\theta,\pi/2)}\,,\, (x,y,z)^t\,\rangle)\, dx\, dy\, dz \\ &= \int_{\mathbb{R}^2} V_X(x,y)\, h_X(B + \langle\, \mu\, e_\theta\,,\, (x,y)^t\,\rangle)\, dx\, dy, \end{aligned}$$

denoting by $e_\theta = (\cos\theta, \sin\theta)^t$ the direction of the field gradient in the XY-plane and

$$V_X(x,y) = \int_{\mathbb{R}} U_X(x,y,z)\, dz\,. \tag{6}$$



(a) from 3D sample to 2D imaging

(b) 2D sinogram $S_{\pi/2}$ (synthetic)

Figure 3: **From the 3D to the 2D setting.** We represent in gray colors in (a) the 3D repartition U_X of a single specie, the TAM, present in the cavity of the acquisition system at different concentrations (one concentration per ellipsoid). We also represent, in false colors, the 2D signal V_X obtained by summing U_X along the Z-axis (see (6)). In such situation, the acquisition of the 3D sinogram for a single azimuthal orientation of the field gradient, here $\varphi = \pi/2$, yields a 2D sinogram such as that displayed in (b). Each projection (or column) of this 2D sinogram is linked to the image V_X through (7) and 2D EPR imaging consists in reconstructing V_X from (b) by inversion. Notice that the image V_X does not correspond to any particular slice of U_X (due to the summation along the vertical axis). In this regard, 2D EPR imaging differs with computed tomography imaging systems that provide slice by slice image reconstruction capability.

Then, with the same methodology as in the 3D case, we show that

$$\forall B \in \mathbb{R}\,,\ \forall \theta \in [0, \pi]\,, \quad S_\theta(B) = \left(h_X^\mu * \mathcal{R}_\theta(V_X)\right)(-B/\mu)\,, \tag{7}$$

denoting by $\mathcal{R}_\theta(V_X) = r \mapsto \int_{\mathrm{Span}(e_\theta^\perp)} V_X(re_\theta + s)\, ds$ the Radon transform in the direction θ of the 2D signal V_X. Instead of estimating the 3D signal U_X, the goal of 2D EPR imaging is to estimate the 2D signal V_X from the 2D sinogram $(B, \theta) \mapsto S_\theta(B)$, as illustrated in Figure 3.

2.4 Filtering of the Projections by the Acquisition System

In practice, sinograms are acquired sequentially, projection by projection. During the acquisition of a projection, a low-pass filtering inherent to the EPR measurement system is done. This low-pass filtering operation can be modeled as a convolution with a low-pass filter which is characteristic of the acquisition system and the acquisition parameters (it may model for instance time integration operation). In the following, we will assume that this low pass filter is a perfect *band-cut filter* with bandwidth $\left[-\frac{\pi}{\nu}, \frac{\pi}{\nu}\right]$ (where $\nu > 0$). More precisely, denoting by $\mathcal{F}$ the Fourier transform in $L^1(\mathbb{R}^n)$, i.e.,

$$\forall f \in L^1(\mathbb{R}^n)\,,\ \forall \xi \in \mathbb{R}^n\,, \quad \mathcal{F}(f)(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-i\langle x, \xi \rangle}\, dx\,,$$

and $\mathbb{1}_X$ the indicator function of the set X (that takes the value 1 over X and zero outside), we assume from now that the low-pass filter is the function g_ν, denoting

$$\forall s > 0,\ \forall t \in \mathbb{R}\,, \quad g_s(t) = \mathcal{F}^{-1}\left(\mathbb{1}_{\left[-\frac{\pi}{s}, \frac{\pi}{s}\right]}\right)(t)\,. \tag{8}$$

Therefore, in the 2D setting, the acquired projection in a direction $\theta \in [0, \pi]$ is the signal $\widetilde{S}_\theta$ given by

$$\forall B \in \mathbb{R}\,, \quad \widetilde{S}_\theta(B) = (g_\nu * S_\theta)\,(B) = \Big(g_\delta * h_X^\mu * \mathcal{R}_\theta(V_X)\Big)(-B/\mu)\,, \tag{9}$$

denoting $\delta = \nu/\mu$ (as we shall see in the following this parameter δ plays an important role in practical experiments). Similarly, in the 3D setting, the acquired projection in a direction $(\theta,\varphi) \in [0,\pi]\times[0,2\pi]$ is the signal $\widetilde{S}_{\theta,\varphi}$ given by

$$\forall B \in \mathbb{R}\,, \quad \widetilde{S}_{\theta,\varphi}(B) = \left(g_\nu * S_{\theta,\varphi}\right)(B) = \Big(g_\delta * h^\mu_X * \mathcal{R}_{\theta,\varphi}(U_X)\Big)(-B/\mu)\,. \tag{10}$$

We can study further the effect of such filtering operation thanks to the *Central Slice Theorem* which provides, for instance in 2D (a similar formulation holds in the 3D setting), an interesting relationship between the Fourier transform of $\mathcal{R}_\theta(V_X)$ and that of V_X.

Theorem 1 (Central Slice Theorem, 2D setting)**.** *For any signal $V \in L^1(\mathbb{R}^2)$ and any $\theta \in \mathbb{R}$, we have*

$$\forall \xi \in \mathbb{R}\,, \quad \mathcal{F}(\mathcal{R}_\theta(V))(\xi) = \mathcal{F}(V)(\xi\cos\theta, \xi\sin\theta)\,. \tag{11}$$

Proof. See Appendix A. □

Corollary 1. *Denoting by $\mathcal{B}_{2\mathrm{D}}(0,\pi/\delta)$ the 2D centered ball with radius π/δ and $f_\delta = \mathcal{F}^{-1}\left(\mathbb{1}_{\mathcal{B}_{2\mathrm{D}}(0,\pi/\delta)}\right)$, for all $B \in \mathbb{R}$ and all $\theta \in [0,\pi]$, we have*

$$\widetilde{S}_\theta(B) = \left(\widetilde{h}^\mu_X * \mathcal{R}_\theta(\widetilde{V}_X)\right)(-B/\mu) \quad \textit{where} \quad \widetilde{h}^\mu_X = g_\delta * h^\mu_X \quad \textit{and} \quad \widetilde{V}_X = f_\delta * V_X\,. \tag{12}$$

Proof. Let $\theta \in [0,\pi]$ and $\xi \in \mathbb{R}$, we have $\mathcal{F}(g_\delta)(\xi) = \mathcal{F}(f_\delta)(\xi\cos\theta,\xi\sin\theta)$. Therefore, from the Central Slice Theorem, we get $\mathcal{F}(g_\delta * \mathcal{R}_\theta(V_X))(\xi) = \mathcal{F}(f_\delta * V_X)(\xi\cos\theta,\xi\sin\theta) = \mathcal{F}(\mathcal{R}_\theta(\widetilde{V}_X))(\xi)$, showing that $g_\delta * \mathcal{R}_\theta(V_X) = \mathcal{R}_\theta(\widetilde{V}_X)$. It follows that $\widetilde{h}^\mu_X * \mathcal{R}_\theta(\widetilde{V}_X) = g_\delta * h^\mu_X * g_\delta * \mathcal{R}_\theta(V_X)$. Then, using $g_\delta = g_\delta * g_\delta$ and the commutativity of the convolution product, we obtain the announced result. □

Interestingly enough, the proof of Corollary 1 points out that filtering by g_δ the Radon transform of the image V_X amounts to computing the Radon transform of the filtered image $\widetilde{V}_X$ which is *bandlimited*. As we shall see in the next section, this property provides a natural discretization scheme for the Radon transform operator.

3 Discretization of the Operators and the Measurements

During the acquisition of an EPR projection, the intensity B of the homogeneous magnetic field is discretized into a finite number of nodes. In most situations, those nodes are regularly sampled, as we shall describe now.

3.1 Measurement Sampling Grids

Let $N_B \in \mathbb{N}$ and let us denote by I_{N_B} and $\overline{I}_{N_B}$ the two sets of N_B consecutive integers defined by

$$I_{N_B} = \left[-\frac{N_B}{2}, \frac{N_B}{2}\right) \cap \mathbb{Z} \quad \text{and} \quad \overline{I}_{N_B} = \left(-\frac{N_B}{2}, \frac{N_B}{2}\right] \cap \mathbb{Z}\,.$$

In particular, when N_B is odd, say $N_B = 2n+1$, we have $I_{N_B} = \overline{I}_{N_B} = \{-n, -n+1, \dots, n-1, n\}$. We assume that the intensity of the homogeneous magnetic field B is regularly sampled over a set $B_{\text{CF}} + \left[-\frac{B_{\text{SW}}}{2}, \frac{B_{\text{SW}}}{2}\right]$ using

$$\forall m \in \overline{I}_{N_B}\,, \quad B_m = B_{\text{CF}} + m\,\delta_B \quad \text{where} \quad \delta_B = \frac{B_{\text{SW}}}{N_B}\,. \tag{13}$$

The quantities B_{CF} and B_{SW} involved above are usually called *center field* and *sweep width* (e.g. in Bruker instruments and XEPR software [17], see also [15]).

From (2), for all orientation (θ,φ) of the field gradient, the intensity of the magnetic field at a given time t of any point x located in the intersection of the cavity with the affine hyperplane $r\, e_{(\theta,\varphi)} + \mathcal{H}_{\theta,\varphi}$ (with radial distance $r = \langle x, e_{(\theta,\varphi)}\rangle$ from the cavity center) is

$$B_{\mu,\theta,\varphi}(x,t) = B(t) + \mu\, r\,.$$

Therefore, if electron resonance occurs in the cavity, by increasing by step δ_B the intensity of the homogeneous magnetic field, we decrease by step $\delta_r = \delta_B/\mu$ the radial positions of the affine hyperplanes where this resonance phenomenon occurs. Thus, sampling with step δ_B the homogeneous magnetic field intensity boils down to spatially sampling the cavity along the direction $e_{(\theta,\varphi)}$ with radial step δ_B/μ, leading to the radial sampling nodes

$$\forall m \in I_{N_B}\,, \quad r_m = m\,\delta_r \quad \text{where} \quad \delta_r = \frac{\delta_B}{\mu} = \frac{B_{\text{SW}}}{\mu\, N_B}\,. \tag{14}$$

In the 2D setting, using (13) and (14) into (12), yields the the sampled measurements

$$\forall m \in I_{N_B}\,, \quad \widetilde{S}_\theta(B_{-m}) = \left(\widetilde{h}_X^\mu * \mathcal{R}_\theta(\widetilde{V}_X)\right)(-B_{\text{CF}}/\mu + m\,\delta_r)\,. \tag{15}$$

We call *field of view* (FOV) the ball delimiting the area covered by the radial measurements. This ball is centered on the cavity center and its diameter D_{FOV} is given by

$$D_{\text{FOV}} = N_B\,\delta_r = \frac{B_{\text{SW}}}{\mu}\,.$$

In practice the FOV may be smaller than the cavity itself but the setting of the sweep width B_{BW} is done to ensure that D_{FOV} is large enough so that the totality of the paramagnetic sample present in the cavity is contained in the FOV.

Once B_{SW} is set, the setting of N_B (and thus, that of the sampling steps δ_B and δ_r) is guided by Shannon Sampling Theorem that we recall below in dimension $n \geq 1$.

Theorem 2 (Shannon Sampling Theorem [37]). *Consider a positive real number $\Delta > 0$ and an absolutely integrable function $f : \mathbb{R}^n \to \mathbb{R}$ whose Fourier transform is supported in $\left[-\frac{\pi}{\Delta}, \frac{\pi}{\Delta}\right]^n$, i.e.,*

$$\forall \xi \notin \left[-\frac{\pi}{\Delta}, \frac{\pi}{\Delta}\right]^n\,, \quad \mathcal{F}(f)(\xi) = 0\,.$$

Then, f is continuous and uniquely determined by its values on $\Delta\,\mathbb{Z}^n$, as

$$\forall x \in \mathbb{R}^n\,, \quad f(x) = \sum_{k\in\mathbb{Z}^n} f(k\Delta)\,\mathrm{sinc}\left(\frac{x}{\Delta} - k\right)\,,$$

where the cardinal sine function is defined in $\mathbb{R}^n$ by

$$\mathrm{sinc}(x) = \prod_{i=1}^{n} \frac{\sin(\pi x_i)}{\pi x_i}\,,$$

with the continuity preserving convention $\frac{\sin(0)}{0} = 1$.

From (9), we see that the signal $r \mapsto (\widetilde{h}_X^\mu * \mathcal{R}_\theta(\widetilde{V}_X))(-B_{\text{CF}}/\mu + r)$ involved in (15) has its Fourier transform supported in the set $[-\frac{\pi}{\delta}, \frac{\pi}{\delta}]$. Therefore, the Shannon Sampling Theorem ensures that no loss of information occurs when sampling this signal on $\delta_r\,\mathbb{Z}$ using a radial sampling step $\delta_r \leq \delta$. From (14), this condition amounts to set $N_B \geq M := B_{\text{SW}}/(\mu\,\delta)$ and we assume that this condition is fulfilled in the following. For commodity, we will also assume that M is an integer (i.e., δ/δ_r is the rational number N_B/M). To simplify the study, the reader may consider $\delta_r = \delta$ (and $M = N_B$) but this is not a requirement that we impose in our model and algorithms (the reason for this choice will be discussed in Section 6).

3.2 Discretization of the Two-dimensional Operators

Let $\theta \in [0,\pi]$. From Theorem 1, we have

$$\forall \xi \in \mathbb{R}\,, \quad \mathcal{F}\left(\mathcal{R}_\theta(\widetilde{V}_X)\right)(\xi) = \mathcal{F}(\widetilde{V}_X)(\xi\cos\theta, \xi\sin\theta)\,. \tag{16}$$

Besides, from $\widetilde{V}_X = f_\delta * V_X$ with $f_\delta = \mathcal{F}^{-1}\left(\mathbb{1}_{\mathcal{B}_{2\mathrm{D}}(0,\pi/\delta)}\right)$, we see that $\mathcal{F}(\widetilde{V}_X)$ is supported in $\mathcal{B}_{2\mathrm{D}}(0,\pi/\delta)$ which is a subset of $\left[-\frac{\pi}{\delta},\frac{\pi}{\delta}\right]^2$. Thus, according to Theorem 2, we have

$$\forall (x,y) \in \mathbb{R}^2\,, \quad \widetilde{V}_X(x,y) = \sum_{(k,\ell)\in\mathbb{Z}^2} \widetilde{V}_X(k\delta,\ell\delta)\,\mathrm{sinc}\left(\frac{x}{\delta}-k\right)\mathrm{sinc}\left(\frac{y}{\delta}-\ell\right). \tag{17}$$

Then, using $\mathcal{F}\left((x,y)\mapsto \mathrm{sinc}((\frac{x}{\delta},\frac{y}{\delta})-(k,\ell))\right)(\xi_1,\xi_2) = \delta^2\, e^{-i(k\delta\xi_1+\ell\delta\xi_2)}\cdot \mathbb{1}_{\left[-\frac{\pi}{\delta},\frac{\pi}{\delta}\right]^2}(\xi_1,\xi_2)$, we get

$$\forall \xi \in \mathbb{R}\,, \quad \mathcal{F}(\widetilde{V}_X)(\xi\cos\theta,\xi\sin\theta) = \begin{cases} \delta^2 \displaystyle\sum_{(k,\ell)\in\mathbb{Z}^2} \widetilde{V}_X(k\delta,\ell\delta)\, e^{-i\xi(k\delta\cos\theta+\ell\delta\sin\theta)} & \text{if } |\xi| \le \frac{\pi}{\delta} \\ 0 & \text{otherwise}\,. \end{cases} \tag{18}$$

Equations (17) and (18) involve an infinite number of samples $\widetilde{V}_X(k\delta,\ell\delta)$ that we cannot handle in practical implementations. Let $\Omega_{2\mathrm{D}} = I_M \times I_M$ and $\widetilde{v}_X : \Omega_{2\mathrm{D}} \to \mathbb{R}$ denote the discrete image obtained by sampling $\widetilde{V}_X$ with step δ in the FOV domain, i.e.,

$$\forall (k,\ell) \in \Omega_{2\mathrm{D}}\,, \quad \widetilde{v}_X(k,\ell) = \widetilde{V}_X(k\delta,\ell\delta)\,.$$

By taking $\xi \in 2\pi\mathbb{Z}/(M\delta)$ and neglecting the terms $\widetilde{V}_X(k\delta,\ell\delta)$ outside the FOV in (18), we obtain

$$\forall \alpha \in \mathbb{Z}\,, \quad \mathcal{F}(\widetilde{V}_X)\left(\tfrac{2\pi\alpha\cos\theta}{M\delta}, \tfrac{2\pi\alpha\sin\theta}{M\delta}\right) \approx \begin{cases} \delta^2 \displaystyle\sum_{(k,\ell)\in\Omega_{2\mathrm{D}}} \widetilde{v}_X(k,\ell)\, e^{-2i\pi\frac{\alpha}{M}(k\cos\theta+\ell\sin\theta)} & \text{if } |\alpha| \le \frac{M}{2} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where the sum in the right-hand side corresponds to the Nonequispaced Discrete Fourier Transform (NDFT) of the discrete image $\widetilde{v}_X$ at the frequency $\left(\frac{\alpha}{M}\cos\theta, \frac{\alpha}{M}\sin\theta\right)$ and can be efficiently evaluated using Nonequispaced Fast Fourier Transform algorithms (NFFT) such as for instance that proposed in [21]. The Fourier transform of $\mathcal{R}_\theta(\widetilde{V}_X)$ being supported in $[-\frac{\pi}{\delta_r},\frac{\pi}{\delta_r}]$ (since $\delta_r \le \delta$), the same methodology yields the approximation

$$\forall \alpha \in I_{N_B}\,, \quad \mathcal{F}(\mathcal{R}_\theta(\widetilde{V}_X))\left(\tfrac{2\pi\alpha}{N_B\delta_r}\right) \approx \delta_r \sum_{m\in I_{N_B}} \mathcal{R}_\theta(\widetilde{V}_X)(m\delta_r)\, e^{-2i\pi\frac{\alpha m}{N_B}}\,, \tag{20}$$

where the right-hand side sum corresponds, up to the multiplicative factor δ_r, to the Discrete Fourier Transform (DFT) of the discrete signal $m \in I_{N_B} \mapsto \mathcal{R}_\theta(\widetilde{V}_X)(m\delta_r)$ at the (discrete) frequency α. Since $M\,\delta = N_B\,\delta_r$, we have $\frac{2\pi\alpha}{M\delta} = \frac{2\pi\alpha}{N_B\delta_r}$ and thanks to Theorem 1, the left-hand side terms of (19) and (20) are equal for any $\alpha \in I_{N_B}$. Then, using $\delta/\delta_r = N_B/M$, we get for all $\alpha \in I_{N_B}$,

$$\mathrm{DFT}\left(m \mapsto \mathcal{R}_\theta(\widetilde{V}_X)(m\,\delta_r)\right)(\alpha) \approx \begin{cases} \delta\cdot\dfrac{N_B}{M}\cdot \displaystyle\sum_{(k,\ell)\in\Omega_{2\mathrm{D}}} \widetilde{v}_X(k,\ell)\, e^{-2i\pi\frac{\alpha}{M}(k\cos\theta+\ell\sin\theta)} & \text{if } |\alpha| \le \frac{M}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

One can see that (21) provides an explicit relation between the discrete image $\widetilde{v}_X$ and the (continuous) Radon transform coefficients $\mathcal{R}_\theta(\widetilde{V}_X)(m\delta_r)$. In the particular case $\delta_r = \delta$ (and $N_B = M$), this yields the following discretization scheme for the Radon transform operator.

Definition 1 (2D discrete Radon transform)**.** *Let $v : I_M \times I_M \to \mathbb{R}$ be a discrete image with size $M \times M$, and let $\theta \in [0, \pi]$. We call the discrete Radon transform of v in the direction θ the discrete signal $\mathcal{R}^{\mathrm{d}}_\theta v : I_M \to \mathbb{R}$ whose DFT, denoted by $\widehat{\mathcal{R}^{\mathrm{d}}_\theta v}$, satisfies*

$$\forall \alpha \in I_M\,, \quad \widehat{\mathcal{R}^{\mathrm{d}}_\theta v}(\alpha) = \varepsilon(\alpha/M) \cdot \sum_{(k,\ell) \in I_M \times I_M} v(k,\ell)\, e^{-2i\pi\left(k\frac{\alpha}{M}\cos\theta + \ell\frac{\alpha}{M}\sin\theta\right)}\,, \tag{22}$$

denoting $\varepsilon = \mathbb{1}_{\left(-\frac{1}{2},\frac{1}{2}\right)}$ (i.e., $\varepsilon(x) = 1$ if $|x| < 1/2$ and $\varepsilon(x) = 0$ otherwise).

Remark. *For $|\alpha| < M/2$, we have $\varepsilon(\alpha/M) = 1$ and we can see that $\widehat{\mathcal{R}^{\mathrm{d}}_\theta v}(\alpha)$ and $\widehat{\mathcal{R}^{\mathrm{d}}_\theta v}(-\alpha)$ are conjugate to each other. When $\alpha = -M/2$ (which can only happen when M is even), we have $\varepsilon(\alpha/M) = 0$ and thus $\widehat{\mathcal{R}^{\mathrm{d}}_\theta v}(-M/2) = 0 = \widehat{\mathcal{R}^{\mathrm{d}}_\theta v}(M/2)$ by M-periodicity of $\widehat{\mathcal{R}^{\mathrm{d}}_\theta v}$. The role of the term $\varepsilon(\alpha/M)$ in (22) is to ensures Hermitian symmetry for $\widehat{\mathcal{R}^{\mathrm{d}}_\theta v}$ whatever the parity of M, so that the signal $\mathcal{R}^{\mathrm{d}}_\theta v$ is real valued.*

When $\delta_r = \delta$ (and $N_B = M$), taking the inverse discrete Fourier transform (IDFT) of (21) yields an explicit relation between the discrete and continuous Radon coefficients,

$$\forall m \in I_M\,, \quad \mathcal{R}_\theta(\widetilde{V}_X)(m\delta) \approx \delta\, \mathcal{R}^{\mathrm{d}}_\theta \widetilde{v}_X(m)\,. \tag{23}$$

When the radial sampling step δ_r is strictly less than δ, taking the IDFT of (21) yields

$$\forall m \in I_{N_B}\,, \quad \mathcal{R}_\theta(\widetilde{V}_X)(m\delta_r) \approx \delta\, Z_r \mathcal{R}^{\mathrm{d}}_\theta \widetilde{v}_X(m)\,. \tag{24}$$

denoting by $Z_r \mathcal{R}^{\mathrm{d}}_\theta \widetilde{v}_X : I_{N_B} \to \mathbb{R}$ the signal whose DFT coefficients $\widehat{Z_r \mathcal{R}^{\mathrm{d}}_\theta \widetilde{v}_X}(\alpha)$ are given by

$$\forall \alpha \in I_{N_B}\,, \quad \widehat{Z_r \mathcal{R}^{\mathrm{d}}_\theta \widetilde{v}_X}(\alpha) = \begin{cases} \frac{N_B}{M}\, \widehat{\mathcal{R}^{\mathrm{d}}_\theta \widetilde{v}_X}(\alpha) & \text{if } \alpha \in I_M \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

Now, using (15) and remarking that the Fourier transforms of the signals $r \mapsto \widetilde{h}^\mu_X(-B_{\text{CF}}/\mu + r)$ and $r \mapsto (\widetilde{h}^\mu_X * \mathcal{R}_\theta(\widetilde{V}_X))(-B_{\text{CF}}/\mu + r)$ are both supported in the set $\left[-\frac{\pi}{\delta_r}, \frac{\pi}{\delta_r}\right]$, we can derive with the same methodology an explicit relation between the projection samples $\widetilde{S}_\theta(B_{-m})$ and the discrete image $\widetilde{v}_X$, we end up with

$$\forall m \in I_{N_B}\,, \quad \widetilde{S}_\theta(B_{-m}) \approx \delta^2 \frac{M}{N_B} \cdot \left(h^{\mathrm{d}}_X \circledast Z_r \mathcal{R}^{\mathrm{d}}_\theta \widetilde{v}_X\right)(m)\,, \tag{26}$$

where $h^{\mathrm{d}}_X : I_{N_B} \to \mathbb{R}$ denotes the discrete signal defined by

$$\forall m \in I_{N_B}\,, \quad h^{\mathrm{d}}_X(m) := (g_\nu * h_X)(B_{-m}) \underset{(\text{when } \mu \neq 0)}{=} \widetilde{h}^\mu_X(-B_{\text{CF}}/\mu + m\,\delta_r)\,, \tag{27}$$

and $u \circledast v := \mathrm{IDFT}\left(\mathrm{DFT}(u) \cdot \mathrm{DFT}(v)\right)$ denotes the circular convolution between the two discrete signals $u : I_{N_B} \to \mathbb{R}$ and $v : I_{N_B} \to \mathbb{R}$. Details about the approximation (26) are given in Appendix B.

Given a sequence $(\theta_p)_{1 \le p \le N_\theta}$ made of N_θ polar angles in $[0, \pi]$ and denoting $\omega_{\text{2D}} = I_{N_B} \times \{1, 2, \ldots, N_\theta\}$, let us consider the 2D discrete sinogram $s_{\text{2D}} : (m, p) \in \omega_{\text{2D}} \mapsto S_{\theta_p}(B_{-m})$. Thanks to (26) we have

$$\forall (m,p) \in \omega_{\text{2D}}\,, \quad s_{\text{2D}}(m,p) \approx \mathcal{A}\widetilde{v}_X(m,p) := \delta^2 \frac{M}{N_B} \cdot \left(h^{\mathrm{d}}_X \circledast Z_r \mathcal{R}^{\mathrm{d}}_{\theta_p} \widetilde{v}_X\right)(m)\,, \tag{28}$$

where the linear operator $\mathcal{A} : \mathbb{R}^{\Omega_{\rm 2D}} \mapsto \mathbb{R}^{\omega_{\rm 2D}}$ models, up to the approximation errors we made in this section, the link between the discrete image $\widetilde{v}_X$ and the measured 2D discrete sinogram $s_{\rm 2D}$. For sake of completeness, we summarize in Algorithm 1 how the 2D discrete sinogram $\mathcal{A}\widetilde{v}_X$ can be simulated from the discrete image $\widetilde{v}_X$, although we are more interested in the reciprocal operation, i.e., reconstructing the image $\widetilde{v}_X$ from the 2D discrete sinogram $s_{\rm 2D}$. The inversion of $\mathcal{A}$ with a regularized least-squares strategy will be addressed in the next section. For that purpose, we will need to evaluate $\mathcal{A}^*$, the adjoint of $\mathcal{A}$. As the for direct operator $\mathcal{A}$, the computation of the adjoint operator $\mathcal{A}^*$ can be done efficiently using the NFFT library [21], as we describe in Algorithm 2.

Algorithm 1: compute the 2D discrete sinogram $\mathcal{A}v$ from a 2D discrete image v.

```
Inputs: a 2D discrete image $v : \Omega_{\rm 2D} \to \mathbb{R}$ with discrete domain $\Omega_{\rm 2D} = I_M \times I_M$, the discrete
    spectrum $h_X^{\rm d} : I_{N_B} \to \mathbb{R}$ with size $N_B \geq M$, a sequence containing $N_\theta$ polar angles
    $(\theta_p)_{1 \leq p \leq N_\theta}$.
Optional input: the radial sampling step $\delta_r = \frac{\delta_B}{\mu}$ (if not provided, set $\delta_r = 1$).
Output: the 2D discrete sinogram $s = \mathcal{A}v : I_{N_B} \times \{1, 2, \dots, N_\theta\} \to \mathbb{R}$ whose p-th column,
    denoted by $s_p$, corresponds to the discrete projection of $v$ in the direction $\theta_p$.
Core of the module:
$\delta \leftarrow \delta_r \cdot \frac{N_B}{M}$   // image sampling step (pixel size), same unit as $\delta_r$
forall $p \in \{1, 2, \dots, N_\theta\}$ do
    set $\widehat{r}_p : I_M \to \mathbb{C}$ the discrete signal[1] defined by
        $\forall \alpha \in I_M\,, \quad \widehat{r}_p(\alpha) = \sum_{(k,\ell) \in \Omega_{\rm 2D}} v(k,\ell)\, e^{-2i\pi\left(k \frac{\alpha}{M} \cos\theta_p + \ell \frac{\alpha}{M} \sin\theta_p\right)}\,;$
    set $\widehat{w}_p : I_{N_B} \to \mathbb{C}$ the discrete signal defined by
        $\forall \alpha \in I_{N_B}\,, \quad \widehat{w}_p(\alpha) = \begin{cases} \mathrm{DFT}(h_X^{\rm d})(\alpha) \cdot \widehat{r}_p(\alpha) & \text{if } |\alpha| < M/2 \\ 0 & \text{otherwise}\,; \end{cases}$
    for all $m \in I_{N_B}$ set $s_p(m) = \delta^2 \cdot \mathrm{IDFT}(\widehat{w}_p)(m)$
return the 2D discrete sinogram $s = \mathcal{A}v$.
```

---

[1] All coefficients $\widehat{r}_p(\alpha)$ can be computed efficiently at once by evaluating the NDFT of the discrete image v at the irregularly spaced frequency nodes $\left\{\left(\frac{\alpha}{M}\cos\theta_p, \frac{\alpha}{M}\sin\theta_p\right),\ \alpha \in I_M\,,\ 1 \leq p \leq N_\theta\right\}$ using a NFFT algorithm such as that proposed in [21]. Thank to this remark, we can efficiently compute all coefficients $\widehat{w}_p(\alpha)$ and afterward the sinogram samples $\{s_p(m),\ m \in I_{N_B}\,,\ 1 \leq p \leq N_\theta\}$ using the (monodimensional) FFT Algorithm [14].

3.3 Discretization of the Three-dimensional Operators

In the 3D setting, one can show that the measured projections $\widetilde{S}_{\theta,\varphi}$ given by (10) can be linked to the filtered signal $\widetilde{U}_X := \mathcal{F}^{-1}(\mathbb{1}_{\mathcal{B}_{\rm 3D}(0,\pi/\delta)}) * U_X$ (denoting by $\mathcal{B}_{\rm 3D}(0,\pi/\delta)$ the centered 3D ball with radius π/δ) through the relation

$$\forall B \in \mathbb{R}\,, \quad \widetilde{S}_{\theta,\varphi}(B) = \left(\widetilde{h}_X^{\mu} * \mathcal{R}_{\theta,\varphi}(\widetilde{U}_X)\right)(-B/\mu)\,. \tag{30}$$

With the same methodology as in the 2D setting, one can derive the following discretization scheme for the 3D Radon transform operator.

Algorithm 2: efficient evaluation of $\mathcal{A}^* s$

```
Inputs: a 2D discrete sinogram $s : I_{N_B} \times \{1, 2, \dots, N_\theta\} \to \mathbb{R}$ with size $N_B \times N_\theta$, the discrete spectrum $h_X^{\mathrm{d}} : I_{N_B} \to \mathbb{R}$ with size $N_B$, the sequence containing the $N_\theta$ polar angles $(\theta_p)_{1 \leq p \leq N_\theta}$, and a positive integer $M \leq N_B$ corresponding to the width (also equal to the height) of the output 2D discrete signal (with domain $\Omega_{\mathrm{2D}} = I_M \times I_M$).
Optional input: the radial sampling step $\delta_r = \frac{\delta_B}{\mu}$ (if not provided, set $\delta_r = 1$).
Output: the 2D discrete signal $\mathcal{A}^* s : \Omega_{\mathrm{2D}} \to \mathbb{R}$.
Core of the module:
$\delta \leftarrow \delta_r \cdot \frac{N_B}{M}$   // image sampling step (pixel size), same unit as $\delta_r$
For all $p \in \{1, 2, \dots, N_\theta\}$, denote by $s_p$ the $p$-th column of $s$, i.e., the discrete signal defined by
    $$\forall m \in I_{N_B}\,, \quad s_p(m) = s(m, p)\,,$$
  and set $\widehat{w}_p : I_{N_B} \to \mathbb{C}$ the signal defined by
    $$\forall \alpha \in I_{N_B}\,, \quad \widehat{w}_p(\alpha) = \overline{\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)} \cdot \mathrm{DFT}(s_p)(\alpha)\,,$$
  denoting by $\overline{z}$ the complex conjugate of $z \in \mathbb{C}$.
For all $(k, \ell) \in \Omega_{\mathrm{2D}}$, set[2]
    $$\mathcal{A}^* s(k, \ell) = \frac{\delta^2}{N_B} \cdot \sum_{\substack{\alpha \in \mathbb{Z},\ |\alpha| < M/2 \\ p \in \{1, 2, \dots, N_\theta\}}} \widehat{w}_p(\alpha)\, e^{2i\pi\left(k \frac{\alpha}{M} \cos\theta_p + \ell \frac{\alpha}{M} \sin\theta_p\right)}\,. \tag{29}$$
  return the 2D discrete signal $\mathcal{A}^* s$.
```

Definition 2 (3D discrete Radon transform)**.** *Let $u : \Omega_{\mathrm{3D}} \to \mathbb{R}$ be a discrete image with discrete domain $\Omega_{\mathrm{3D}} = I_M \times I_M \times I_M$, and let $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi]$. We call the discrete Radon transform of u in the direction (θ, φ) the discrete signal $\mathcal{R}_{\theta,\varphi}^{\mathrm{d}} u : I_M \to \mathbb{R}$ whose DFT, denoted by $\widehat{\mathcal{R}_{\theta,\varphi}^{\mathrm{d}} u}$, satisfies*

$$\forall \alpha \in I_M\,, \quad \widehat{\mathcal{R}_{\theta,\varphi}^{\mathrm{d}} u}(\alpha) = \varepsilon(\alpha/M) \cdot \sum_{(k,\ell,m) \in \Omega_{\mathrm{3D}}} u(k, \ell, m)\, e^{-2i\pi \frac{\alpha}{M}(k \cos\theta \sin\varphi + \ell \sin\theta \sin\varphi + m \cos\varphi)}\,.$$

Let $\widetilde{u}_X : \Omega_{\mathrm{3D}} \to \mathbb{R}$ be the discrete 3D image obtained by sampling $\widetilde{U}_X$ regularly with step δ along each dimension, more precisely, the discrete signal defined by

$$\forall (k, \ell, m) \in \Omega_{\mathrm{3D}}\,, \quad \widetilde{u}_X(k, \ell, m) = \widetilde{U}_X(k\delta, \ell\delta, m\delta)\,.$$

Then, one can link the discrete Radon transform of $\widetilde{u}_X$ to the continuous Radon transform of $\widetilde{U}_X$ through the approximation

$$\forall m \in I_{N_B}\,, \quad \mathcal{R}_{\theta,\varphi}(\widetilde{U}_X)(m\delta_r) \approx \delta^2 \cdot Z_r \mathcal{R}_{\theta,\varphi}^{\mathrm{d}} \widetilde{u}_X(m)\,,$$

denoting by $Z_r \mathcal{R}_{\theta,\varphi}^{\mathrm{d}} \widetilde{u}_X : I_{N_B} \to \mathbb{R}$ the discrete signal whose DFT coefficients $\widehat{Z_r \mathcal{R}_{\theta,\varphi}^{\mathrm{d}}} \widetilde{u}_X$ are given by

$$\forall \alpha \in I_{N_B}\,, \quad \widehat{Z_r \mathcal{R}_{\theta,\varphi}^{\mathrm{d}}} \widetilde{u}_X(\alpha) = \varepsilon(\alpha/M) \cdot \frac{N_B}{M} \cdot \sum_{(k,\ell,m) \in \Omega_{\mathrm{3D}}} \widetilde{u}_X(k, \ell, m)\, e^{-2i\pi \frac{\alpha}{M}(k \cos\theta \sin\varphi + \ell \sin\theta \sin\varphi + m \cos\varphi)}\,.$$

---

[2] All coefficients $\mathcal{A}^* s(k, \ell)$ can be computed at once by evaluating the adjoint NDFT of the discrete signal $(\alpha, p) \mapsto \widehat{w}_p(\alpha)$ for $\{\alpha \in \mathbb{Z}\,,\ |\alpha| < M/2\,,\ 1 \leq p \leq N_\theta\}$, associating each sample $\widehat{w}_p(\alpha)$ to the frequency node $\left(\frac{\alpha}{M} \cos\theta_p, \frac{\alpha}{M} \sin\theta_p\right)$. Such computation can be done efficiently using an adjoint NFFT algorithm such as that proposed in [21].

Eventually, given a sequence of polar and azimuthal angles (θ_p, φ_q) (for $1 \leq p \leq N_\theta$, and $1 \leq q \leq N_\varphi$), we can link the measured samples to the discrete image $\widetilde{u}_X$ through the relation

$$\begin{aligned}&\forall m \in I_{N_B}\,,\ \forall p \in \{1, 2, \dots, N_\theta\}\,,\ \forall q \in \{1, 2, \dots, N_\varphi\}\\ &\qquad s_{3\mathrm{D}}(m, p, q) := \widetilde{S}_{\theta_p, \varphi_q}(B_{-m}) \approx \mathcal{B}\,\widetilde{u}_X(m, p, q) := \delta^3 \frac{M}{N_B} \cdot \left(h_X^{\mathrm{d}} \circledast Z_r \mathcal{R}^{\mathrm{d}}_{\theta_p, \varphi_q} \widetilde{u}_X\right)(m)\,. \end{aligned} \tag{31}$$

The simulation of the 3D discrete sinogram $\mathcal{B}\,\widetilde{u}_X$ from the 3D discrete image $\widetilde{u}_X$ can be done using Algorithm 3, while the adjoint operation can be computed efficiently using Algorithm 4.

Algorithm 3: compute the 3D discrete sinogram $\mathcal{B}u$ from the 3D discrete image u.

```
Inputs: a 3D discrete image $u : \Omega_{3\mathrm{D}} \to \mathbb{R}$ with discrete domain $\Omega_{3\mathrm{D}} = I_M \times I_M \times I_M$, the
  discrete spectrum $h_X^{\mathrm{d}} : I_{N_B} \to \mathbb{R}$ with size $N_B \geq M$, a sequence containing $N_\theta$ polar angles
  $(\theta_p)_{1 \leq p \leq N_\theta}$ and a sequence containing $N_\varphi$ azimuthal angles $(\varphi_q)_{1 \leq q \leq N_\varphi}$.
Optional input: the radial sampling step $\delta_r = \frac{\delta_B}{\mu}$ (if not provided, set $\delta_r = 1$).
Output: the 3D discrete sinogram $s = \mathcal{B}u : I_{N_B} \times \{1, 2, \dots, N_\theta\} \times \{1, 2, \dots, N_\varphi\} \to \mathbb{R}$.
Core of the module:
$\delta \leftarrow \delta_r \cdot \frac{N_B}{M}$   // image sampling step (pixel size), same unit as $\delta_r$
forall $(p, q) \in \{1, 2, \dots, N_\theta\} \times \{1, 2, \dots, N_\varphi\}$ do
    set $\widehat{r}_{p,q} : I_M \to \mathbb{C}$ the discrete signal defined by
        $\forall \alpha \in I_M\,,\quad \widehat{r}_{p,q}(\alpha) = \sum_{(k,\ell,m) \in \Omega_{3\mathrm{D}}} u(k, \ell, m)\, e^{-2i\pi\left(k\frac{\alpha}{M}\cos\theta_p \sin\varphi_q + \ell\frac{\alpha}{M}\sin\theta_p \sin\varphi_q + m\frac{\alpha}{M}\cos\varphi_q\right)}\,;$
    set $\widehat{w}_{p,q} : I_{N_B} \to \mathbb{C}$ the discrete signal defined by
        $\forall \alpha \in I_{N_B}\,,\quad \widehat{w}_{p,q}(\alpha) = \begin{cases} \mathrm{DFT}(h_X^{\mathrm{d}})(\alpha) \cdot \widehat{r}_{p,q}(\alpha) & \text{if } |\alpha| < M/2 \\ 0 & \text{otherwise}\,; \end{cases}$
    for all $m \in I_{N_B}$ set $s(m, p, q) = \delta^3 \cdot \mathrm{IDFT}(\widehat{w}_{p,q})(m)$
return the 3D discrete sinogram $s = \mathcal{B}u$.
```

4 A Total Variation Based Image Reconstruction Model

We derived in Section 3 some relations between the measurements ($(\theta, m) \mapsto \widetilde{S}_\theta(B_{-m})$ in the 2D setting, or $(\theta, \varphi, m) \mapsto \widetilde{S}_{\theta,\varphi}(B_{-m})$ in the 3D setting) and a discrete image ($\widetilde{v}_X$ or $\widetilde{u}_X$). We can address now the problem of the reconstruction of the discrete image from those measurements. We focus first on the 2D setting.

4.1 Two-dimensional Image Reconstruction

As done in [10], we address the reconstruction of $\widetilde{v}_X$ from $s_{2\mathrm{D}}$ using a total variation (TV) regularized least-squares strategy. For $\lambda > 0$, we aim to compute an image

$$v_\lambda \in \operatorname*{argmin}_{v \in \mathbb{R}^{\Omega_{2\mathrm{D}}}} J_\lambda^{2\mathrm{D}}(v) := \frac{1}{2}\|\mathcal{A}v - s_{2\mathrm{D}}\|_2^2 + \lambda\,\mathrm{TV}(v)\,, \tag{32}$$

Algorithm 4: efficient evaluation of $\mathcal{B}^* s$

Inputs: *a 3D discrete sinogram* $s : I_{N_B} \times \{1, 2, \ldots, N_\theta\} \times \times \{1, 2, \ldots, N_\phi\} \to \mathbb{R}$ *with size* $N_B \times N_\theta \times N_\varphi$*, the discrete spectrum* $h_X^{\mathrm{d}} : I_{N_B} \to \mathbb{R}$ *with size* $N_B \geq M$*, the sequences containing the* N_θ *polar angles* $(\theta_p)_{1 \leq p \leq N_\theta}$ *and the* N_φ *azimuthal angles* $(\varphi_q)_{1 \leq q \leq N_\varphi}$*, a positive integer* $M \leq N_B$ *corresponding to the width (also equal to the height and depth) of the output 3D discrete signal (with domain* $\Omega_{3\mathrm{D}} = I_M \times I_M \times I_M$*).*

Optional input: *the radial sampling step* $\delta_r = \frac{\delta_B}{\mu}$ *(if not provided, set* $\delta_r = 1$*).*

Output: *the 3D discrete signal* $\mathcal{B}^* s : \Omega_{3\mathrm{D}} \to \mathbb{R}$ *.*

Core of the module:

$\delta \leftarrow \delta_r \cdot \frac{N_B}{M}$ *// image sampling step (pixel size), same unit as* δ_r

For all $(p, q) \in \{1, 2, \ldots, N_\theta\} \times \{1, 2, \ldots, N_\varphi\}$, denote by $s_{p,q} : I_{N_B} \to \mathbb{R}$ the discrete signal defined by

$$\forall m \in I_{N_B}\,, \quad s_{p,q}(m) = s(m, p, q)\,,$$

and set $\widehat{w}_{p,q} : I_{N_B} \to \mathbb{C}$ the signal defined by

$$\forall \alpha \in I_{N_B}\,, \quad \widehat{w}_{p,q}(\alpha) = \overline{\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)} \cdot \mathrm{DFT}(s_{p,q})(\alpha)\,.$$

For all $(k, \ell, m) \in \Omega_{3\mathrm{D}}$, set

$$\mathcal{B}^* s(k, \ell, m) = \frac{\delta^3}{N_B} \cdot \sum_{\substack{\alpha \in \mathbb{Z},\ |\alpha| < M/2 \\ p \in \{1,2,\ldots,N_\theta\} \\ q \in \{1,2,\ldots,N_\varphi\}}} \widehat{w}_{p,q}(\alpha)\, e^{2i\pi\left(k\frac{\alpha}{M}\cos\theta_p \sin\varphi_q + \ell\frac{\alpha}{M}\sin\theta_p \sin\varphi_q + m\frac{\alpha}{M}\cos\varphi_q\right)}\,.$$

return the 3D discrete signal $\mathcal{B}^* s$.

---

where the TV of the discrete image $v \in \mathbb{R}^{\Omega_{2\mathrm{D}}}$ is defined as

$$\mathrm{TV}(v) = \sum_{(k,\ell) \in \Omega_{2\mathrm{D}}} \|\nabla_{2\mathrm{D}} v(k, \ell)\|_2\,, \tag{33}$$

denoting by $\nabla_{2\mathrm{D}} : \mathbb{R}^{\Omega_{2\mathrm{D}}} \to \mathbb{R}^{\Omega_{2\mathrm{D}}} \times \mathbb{R}^{\Omega_{2\mathrm{D}}}$ the classical *forward 2D discrete finite differences scheme* defined by $\nabla_{2\mathrm{D}} v = (\nabla_x^{2\mathrm{D}} v, \nabla_y^{2\mathrm{D}} v)$ with, for all $(k, \ell) \in \Omega_{2\mathrm{D}}$,

$$\nabla_x^{2\mathrm{D}} v(k, \ell) = \begin{cases} v(k+1, \ell) - v(k, \ell) & \text{if } (k+1, \ell) \in \Omega_{2\mathrm{D}} \\ 0 & \text{otherwise,} \end{cases}$$

$$\nabla_y^{2\mathrm{D}} v(k, \ell) = \begin{cases} v(k, \ell+1) - v(k, \ell) & \text{if } (k, \ell+1) \in \Omega_{2\mathrm{D}} \\ 0 & \text{otherwise.} \end{cases}$$

The parameter λ in (32) can be used to control the relative weight of the penalty term $\mathrm{HTV}_\alpha(v)$ with respect to the data-fidelity term $\frac{1}{2}\|\mathcal{A}v - s_{2\mathrm{D}}\|_2^2$ in the energy $J_\lambda^{2\mathrm{D}}(v)$. The practical tuning of λ is closely related to the EPR acquisition parameters and will be discussed in Section 6.4.

Modern convex analysis tools based on Legendre-Fenchel duality [11, 36] allowed the development of many iterative algorithms able to efficiently handle the minimization of convex but non-differentiable energies like $J_\lambda^{2\mathrm{D}}$ in the two last decades [5, 7, 9]. Recently, a class of algorithms particularly well suited to (32) was proposed in [7] (see also [42]) and was further studied and generalized in [6]. Those algorithms are able to efficiently take advantage of the presence of a *Lipschitz* *differentiable* (i.e. differentiable with Lipschitz-continuous gradient) term in the energy to minimize. Indeed, the data-fidelity term $f : v \mapsto \frac{1}{2}\|\mathcal{A}v - s_{\text{2D}}\|_2^2$ involved in (32) is differentiable over $\mathbb{R}^{\Omega_{\text{2D}}}$ with gradient ∇f given by

$$\forall v \in \mathbb{R}^{\Omega_{\text{2D}}}, \quad \nabla f(v) = \mathcal{A}^*\mathcal{A}v - \mathcal{A}^* s_{\text{2D}} \,. \tag{34}$$

Since ∇f is an affine operator, it is also Lipschitz-continuous with Lipschitz constant $L_f \leq |||\mathcal{A}^*\mathcal{A}|||$ (denoting by $|||\mathcal{A}^*\mathcal{A}|||$ the ℓ^2-induced norm of the operator $\mathcal{A}^*\mathcal{A}$). Then, applying to (32) the *Euclidean non-ergodic non-linear primal-dual algorithm* proposed in [6] boils down to the following numerical scheme. Given $v^0 \in \mathbb{R}^{\Omega_{\text{2D}}}$, $p^0 \in \mathbb{R}^{\Omega_{\text{2D}}} \times \mathbb{R}^{\Omega_{\text{2D}}}$ and two positive parameters (τ, σ), set $\overline{v}^0 = v^0$ and iterate for $n \geq 0$

$$\left\{ \begin{aligned} p^{n+1} &= \Pi_{\mathcal{B}_{\text{2D}}}\left(p^n + \sigma\lambda\nabla_{\text{2D}}\overline{v}^n\right), && \text{(35a)} \\ v^{n+1} &= v^n - \tau\left(\nabla f(v^n) - \lambda\, \mathrm{div}_{\text{2D}}\left(p^{n+1}\right)\right), && \text{(35b)} \\ \overline{v}^{n+1} &= 2\, v^{n+1} - v^n, && \text{(35c)} \end{aligned} \right.$$

where $\mathrm{div}_{\text{2D}} = -\nabla_{\text{2D}}^*$ denotes the opposite adjoint of the 2D forward finite differences operator ∇_{2D} (see Appendix C) and $\Pi_{\mathcal{B}_{\text{2D}}}$ denotes the orthogonal projection over the closed and convex set $\mathcal{B}_{\text{2D}} := \{p \in \mathbb{R}^{\Omega_{\text{2D}}} \times \mathbb{R}^{\Omega_{\text{2D}}}\,,\ \forall (k,\ell) \in \Omega_{\text{2D}}\,,\ \|p(k,\ell)\|_2 \leq 1\}$ which is more explicitly given by

$$\forall p_0 \in \mathbb{R}^{\Omega_{\text{2D}}} \times \mathbb{R}^{\Omega_{\text{2D}}}\,,\ \forall (k,\ell) \in \Omega_{\text{2D}}\,, \quad \Pi_{\mathcal{B}_{\text{2D}}}(p_0)(k,\ell) = \frac{p_0(k,\ell)}{\max(1, \|p_0(k,\ell)\|_2)}\,. \tag{36}$$

Besides, since we have the upper bound $|||\nabla_{\text{2D}}||| \leq \sqrt{8}$, Theorem 1 established in [6] ensures the convergence of the sequence $(v^n)_{n\in\mathbb{N}}$ toward a minimizer of J_λ^{2D} when the time-step parameters τ and σ satisfy

$$\left(\frac{1}{\tau} - |||\mathcal{A}^*\mathcal{A}|||\right)\frac{1}{\sigma} \geq 8\lambda^2\,. \tag{37}$$

Remark (adding constraints in (32))**.** *One can restrict the minimization of* J_λ^{2D} *to any convex set* $\mathcal{C} \subset \mathbb{R}^{\Omega_{\text{2D}}}$ *simply by replacing* (35b) *by*

$$v^{n+1} = \Pi_{\mathcal{C}}\left(v^n - \tau\left(\nabla f(v^n) - \lambda\, \mathrm{div}_{\text{2D}}(p^{n+1})\right)\right)\,,$$

denoting by $\Pi_{\mathcal{C}} : \mathbb{R}^{\Omega_{\text{2D}}} \to \mathcal{C}$ *the orthogonal projection over* $\mathcal{C}$*. In particular, we remarked that using* $\mathcal{C} = \mathbb{R}_+^{\Omega_{\text{2D}}}$ *in order to restrict the minimization of* J_λ^{2D} *to the set of nonnegatively-valued images may increase the convergence rate of the scheme* (35)*. Such nonnegativity constraint makes sense because, if* $\widetilde{v}_X$ *may assume some negative values, the latter are due to the oscillations involved by the filtering of* V_X *(which is nonnegatively valued) into* $\widetilde{V}_X = V_X * g_\delta$ *(which may assume negative values), and we usually want to avoid such oscillatory patterns in the reconstruction. However, in our experiments over real EPR data, we remarked that not imposing the positivity constraint and taking the positive part of the reconstructed image at the end of the process yielded visually more satisfactory results.*

We can see that the computation of $\nabla f(v^n) = \mathcal{A}^*\mathcal{A}v^n - \mathcal{A}^* s_{\text{2D}}$ is needed at each iteration $n \geq 0$ of the scheme in (35b). The term $\mathcal{A}^* s_{\text{2D}}$ never changes and can be computed once and for all using Algorithm 2. The term $\mathcal{A}^*\mathcal{A}v^n$ could be computed by using successively Algorithm 1 and Algorithm 2, however, this computation is rather costly since NFFTs are involved in both algorithms. In fact, the terms $\mathcal{A}^*\mathcal{A}v^n$ can be evaluated in a more efficient way thanks to Proposition 1.

Proposition 1. *Let* $v \in \mathbb{R}^{\Omega_{\text{2D}}}$ *and denote by* Zv *and* φ *the two-dimensional discrete signals defined over the augmented domain* $\Upsilon_{\text{2D}} = I_{2M} \times I_{2M}$ *by,*

$$\forall (k,\ell) \in \Upsilon_{\text{2D}}\,, \quad Zv(k,\ell) = \left\{ \begin{array}{cl} v(k,\ell) & \textit{if } (k,\ell) \in \Omega_{\text{2D}} \\ 0 & \textit{otherwise,} \end{array} \right. \tag{38}$$

$$\text{and} \quad \varphi(k,\ell) = \frac{\delta^4}{N_B} \cdot \sum_{\substack{\alpha\in\mathbb{Z},\ |\alpha|<M/2 \\ p\in\{1,2,\dots,N_\theta\}}} \overline{\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)} \cdot \mathrm{DFT}(h_X^{\mathrm{d}})(\alpha) \cdot e^{2i\pi\left(k\frac{\alpha}{M}\cos\theta_p + \ell\frac{\alpha}{M}\sin\theta_p\right)} . \tag{39}$$

Then, $\mathcal{A}^*\mathcal{A}v$ corresponds to the restriction of $\varphi \circledast Zv$ to $\Omega_{2\mathrm{D}}$, i.e., we have

$$\forall (k,\ell) \in \Omega_{2\mathrm{D}}\,, \quad \mathcal{A}^*\mathcal{A}v(k,\ell) = \left(\varphi \circledast Zv\right)(k,\ell)\,. \tag{40}$$

Proof. See Appendix D. ☐

In other words, the hypermatrix of $\mathcal{A}^*\mathcal{A}$ has a Toeplitz structure and $\mathcal{A}^*\mathcal{A}v_n$ can be computed using two sequential FFTs. This property was first used for Computer Tomography in [25] and for EPR imaging in [10]. It therefore allows to quickly compute the gradient of the data fidelity $u \mapsto \frac{1}{2}\|\mathcal{A}u - s_{2\mathrm{D}}\|_2^2$. Notice that the proximity operator of the latter can also be computed using Proposition 1 provided we can quickly invert $\mathcal{A}^*\mathcal{A} + I$ (where I denotes the identity operator). For this purpose, $\mathcal{A}^*\mathcal{A}$ is approximated by a circulant hypermatrix in [10]. In this article, we have instead chosen to carry out a forward gradient descent on the primal variable or, in other words, change the algorithm of [5] into that of [7, 42, 6].

Corollary 2. *We have the upper bound $|||\mathcal{A}^*\mathcal{A}||| \leq L'_f := \|\mathrm{DFT}(\varphi)\|_\infty$. Then, setting $\tau = 1/(2L'_f)$ and $\sigma = L'_f/(8\lambda^2)$ ensures that (37) is fulfilled.*

Proof. Thanks to Proposition 1, for any $v \in \mathbb{R}^{\Omega_{2\mathrm{D}}}$, we have $\|\mathcal{A}^*\mathcal{A}v\|_2^2 \leq \|\varphi \circledast Zv\|_2^2$. Then, Parseval's identity yields $\|\varphi \circledast Zv\|_2^2 = \frac{1}{4M^2}\sum_{(\alpha,\beta)\in\Upsilon_{2\mathrm{D}}} |\mathrm{DFT}(\varphi)(\alpha,\beta)\cdot\mathrm{DFT}(Zv)(\alpha,\beta)|^2 \leq {L'_f}^2\|Zv\|_2^2 = {L'_f}^2\|v\|_2^2$, from which the announced upper-bound $|||\mathcal{A}^*\mathcal{A}||| \leq L'_f$ follows. ☐

From the kernel φ and using Corollary 2, we are able to compute an explicit bound for $|||\mathcal{A}^*\mathcal{A}|||$, which yields an explicit setting of the time-step parameters (τ, σ) that ensures the convergence of the scheme toward a minimizer of $J_\lambda^{2\mathrm{D}}$. The computation of the kernel φ can be done using NFFT and since this kernel does not change during the scheme iterations, it can be computed once and systematically used to evaluate $\mathcal{A}^*\mathcal{A}v^n$ at each iteration of the scheme using (40). Finally, the whole procedure for the reconstruction of an image v_λ by minimization of $J_\lambda^{2\mathrm{D}}$ over the constraint set $\mathcal{C} = \mathbb{R}_+^{\Omega_{2\mathrm{D}}}$ or $\mathcal{C} = \mathbb{R}^{\Omega_{2\mathrm{D}}}$ is summarized in Algorithm 5.

4.2 Three-dimensional Image Reconstruction

In the 3D setting, we adopt the same methodology. Given the observed 3D sinogram $s_{3\mathrm{D}} \approx \mathcal{B}\widetilde{u}_X$ and a regularity parameter $\lambda > 0$, we address the TV-regularized (constrained of unconstrained) least-squares problem

$$u_\lambda \in \operatorname*{argmin}_{u\in\mathcal{C}'} \; J_\lambda^{3\mathrm{D}}(u) := \frac{1}{2}\|\mathcal{B}u - s_{3\mathrm{D}}\|_2^2 + \lambda\,\mathrm{TV}(u)\,, \tag{41}$$

where $\mathcal{C}' = \mathbb{R}^{\Omega_{3\mathrm{D}}}$ (unconstrained case) or $\mathcal{C}' = \mathbb{R}_+^{\Omega_{3\mathrm{D}}}$ (nonnegatively constrained case), and where

$$\forall u \in \mathbb{R}^{\Omega_{3\mathrm{D}}}\,, \quad \mathrm{TV}(u) = \sum_{(k,\ell,m)\in\Omega_{3\mathrm{D}}} \|\nabla_{3\mathrm{D}}u(k,\ell,m)\|_2\,,$$

denoting by $\nabla_{3\mathrm{D}}u = (\nabla_x^{3\mathrm{D}}u, \nabla_y^{3\mathrm{D}}u, \nabla_z^{3\mathrm{D}}u)$ the 3D forward differences scheme defined by, for all $(k,\ell,m) \in \Omega_{3\mathrm{D}}$,

$$\nabla_x^{3\mathrm{D}}u(k,\ell,m) = \begin{cases} u(k+1,\ell,m) - u(k,\ell,m) & \text{if } (k+1,\ell,m) \in \Omega_{3\mathrm{D}} \\ 0 & \text{otherwise;} \end{cases}$$

```
Algorithm 5: 2D discrete image reconstruction from a 2D discrete sinogram.
  Inputs: a 2D discrete sinogram $s_{2D} : I_{N_B} \times \{1, 2, \dots, N_\theta\} \to \mathbb{R}$ with size $N_B \times N_\theta$, the
    associated discrete spectrum $h_X^{\mathrm{d}} : I_{N_B} \to \mathbb{R}$ with size $N_B$, the associated sequence containing
    the $N_\theta$ polar angles $(\theta_p)_{1 \le p \le N_\theta}$, a positive integer $M \le N_B$ corresponding to the desired
    width (also equal to the height) for the output image (with domain $\Omega_{2D} = I_M \times I_M$), a
    regularity parameter $\lambda > 0$, and a number of iterations $N \ge 0$ for the numerical
    scheme (35).
  Optional inputs: a flag specifying whether we aim to minimize $J_\lambda^{2D}$ over the set $\mathcal{C} = \mathbb{R}^{\Omega_{2D}}$
    (unconstrained minimization) or over the set $\mathcal{C} = \mathbb{R}_+^{\Omega_{2D}}$ (nonnegativity constraint), an initial
    guess $v^0 \in \mathbb{R}^{\Omega_{2D}}$ for the looked for image (if not provided, use $v^0 = \mathbf{0}_{\Omega_{2D}}$ = the zero-valued
    image in $\mathbb{R}^{\Omega_{2D}}$), the radial sampling step $\delta_r = \frac{\delta_B}{\mu}$ (if not provided, set $\delta_r = 1$).
  Outputs: a numerical estimate of a minimizer $v_\lambda$ of $J_\lambda^{2D}$ over $\mathcal{C}$.
  Core of the module:
  // precompute $\mathcal{A}^* s_{2D}$ and the kernel $\varphi$, initialize other variables
  Astar_s ← $\mathcal{A}^* s_{2D}$  // using Algorithm 2 with $\delta_r$ as optional input
  phi ← $\varphi$  // using (39) with $\delta = \delta_r \cdot \frac{N_B}{M}$ and a NFFT algorithm like [21]
  dft_phi ← DFT(phi)  // using the FFT Algoritm [14]
  $L'_f$ ← $\max_{(\alpha,\beta) \in \Upsilon_{2D}}$ |dft_phi$(\alpha, \beta)$|
  $\tau \leftarrow 1/(2L'_f)$
  $\sigma \leftarrow L'_f/(8\lambda^2)$
  $(v, \overline{v}, v\_prev) \leftarrow (v^0, v^0, v^0)$
  $p \leftarrow (\mathbf{0}_{\Omega_{2D}}, \mathbf{0}_{\Omega_{2D}})$
  // main loop (iterations of the scheme (35))
  for $1 \le n \le N$ do
      // compute $\nabla f(v) = \mathcal{A}^* \mathcal{A} v - \mathcal{A}^* s_{2D}$
      dft_Zv ← DFT($Zv$)  // where $Zv$ is obtained from $v$ using (38)
      dft_w ← dft_phi · dft_Zv  // we have dft_w = DFT($\varphi \circledast Zv$)
      AstarA_v ← the restriction of IDFT(dft_w) to $\Omega_{2D}$
      gradf_v ← AstarA_v − Astar_s
      // perform one scheme iteration
      $p \leftarrow \Pi_{\mathcal{B}_{2D}}(p + \sigma \lambda \nabla_{2D} \overline{v})$  // using (36)
      $v \leftarrow v - \tau \cdot$ (gradf_v $- \lambda\, \mathrm{div}_{2D}(p)$)  // using (68)
      if $\mathcal{C} = \mathbb{R}_+^{\Omega_{2D}}$ then $v \leftarrow \max(\mathbf{0}_{\Omega_{2D}}, v)$
        // projection into $\mathcal{C}$ by taking the positive part
      $\overline{v} \leftarrow 2v - v\_prev$
      $v\_prev \leftarrow v$
  return the 2D discrete image v  // sampling step (or pixel size) is $\delta = \delta_r \cdot \frac{N_B}{M}$ (same unit as $\delta_r$)
```

$$\nabla_y^{3D} u(k, \ell, m) \quad = \begin{cases} u(k, \ell+1, m) - u(k, \ell, m) & \text{if } (k, \ell+1, m) \in \Omega_{3D} \\ 0 & \text{otherwise;} \end{cases}$$

$$\nabla_z^{3D} u(k, \ell, m) \quad = \begin{cases} u(k, \ell, m+1) - u(k, \ell, m) & \text{if } (k, \ell, m+1) \in \Omega_{3D} \\ 0 & \text{otherwise.} \end{cases}$$

We denote by $\operatorname{div}_{3\mathrm{D}} = -\nabla_{3\mathrm{D}}^*$ the opposite adjoint of $\nabla_{3\mathrm{D}}$ (see Appendix C) and $f' : u \mapsto \frac{1}{2}\|\mathcal{B}u - s_{3\mathrm{D}}\|_2^2$ the data fidelity term involved in (41). The *Euclidean non-ergodic non-linear primal-dual algorithm* proposed in [6] yields the following resolvant scheme for (41). Given $u^0 \in \mathbb{R}^{\Omega_{3\mathrm{D}}}$, $p^0 \in \mathbb{R}^{\Omega_{3\mathrm{D}}} \times \mathbb{R}^{\Omega_{3\mathrm{D}}} \times \mathbb{R}^{\Omega_{3\mathrm{D}}}$ and two positive parameters (τ, σ), set $\overline{u}^0 = u^0$ and iterate for $n \geq 0$

$$\left\{\begin{aligned} p^{n+1} &= \Pi_{\mathcal{B}_{3\mathrm{D}}}\left(p^n + \sigma\lambda\nabla_{3\mathrm{D}}\overline{u}^n\right), && \text{(42a)}\\ u^{n+1} &= \Pi_{\mathcal{C}'}\left(u^n - \tau\left(\nabla f'(u^n) - \lambda \operatorname{div}_{3\mathrm{D}}(p^{n+1})\right)\right), && \text{(42b)}\\ \overline{u}^{n+1} &= 2\,u^{n+1} - u^n, && \text{(42c)}\end{aligned}\right.$$

denoting by $\Pi_{\mathcal{C}'}$ the orthogonal projection into $\mathcal{C}'$ and $\Pi_{\mathcal{B}'}$ the orthogonal projection into the dual ball $\mathcal{B}_{3\mathrm{D}} := \{p \in \mathbb{R}^{\Omega_{3\mathrm{D}}} \times \mathbb{R}^{\Omega_{3\mathrm{D}}} \times \mathbb{R}^{\Omega_{3\mathrm{D}}},\ \forall (k,\ell,m) \in \Omega_{3\mathrm{D}},\ \|p(k,\ell,m)\|_2 \leq 1\}$, i.e.,

$$\forall p_0 \in \mathbb{R}^{\Omega_{3\mathrm{D}}} \times \mathbb{R}^{\Omega_{3\mathrm{D}}} \times \mathbb{R}^{\Omega_{3\mathrm{D}}},\ \forall (k,\ell,m) \in \Omega_{3\mathrm{D}},\quad \Pi_{\mathcal{B}_{3\mathrm{D}}}(p_0)(k,\ell,m) = \frac{p_0(k,\ell,m)}{\max\left(1, \|p_0(k,\ell,m)\|_2\right)}. \tag{43}$$

The gradient $\nabla f'$ of the data-fidelity term f' involved in (42b) satisfies

$$\forall u \in \mathbb{R}^{\Omega_{3\mathrm{D}}},\quad \nabla f'(u) = \mathcal{B}^*\mathcal{B}u - \mathcal{B}^* s_{3\mathrm{D}}. \tag{44}$$

The constant term $\mathcal{B}^* s_{3\mathrm{D}}$ can be computed once using Algorithm 4 and used along all the iterations of the scheme (42) as it keeps unchanged. The terms $\mathcal{B}^*\mathcal{B}u^n$ involved in (42b) through the term $\nabla f'(u^n)$ can be efficiently computed using a circular convolution as we can show that

$$\forall u \in \mathbb{R}^{\Omega_{3\mathrm{D}}},\quad \forall (k,\ell,m) \in \Omega_{3\mathrm{D}},\quad \mathcal{B}^*\mathcal{B}u(k,\ell,m) = \left(\psi \circledast Zu\right)(k,\ell,m), \tag{45}$$

where ψ and Zu are the discrete signals defined over the augmented domain $\Upsilon_{3\mathrm{D}} := I_{2M} \times I_{2M} \times I_{2M}$ by

$$\forall (k,\ell,m) \in \Upsilon_{3\mathrm{D}},\quad Zu(k,\ell,m) = \left\{\begin{array}{cl} u(k,\ell,m) & \text{if } (k,\ell,m) \in \Omega_{3\mathrm{D}}\\ 0 & \text{otherwise}\end{array}\right. \tag{46}$$

$$\text{and}\quad \psi(k,\ell,m) = \frac{\delta^6}{N_B} \cdot \sum_{\substack{\alpha\in\mathbb{Z}\ \ |\alpha|<M/2\\ p\in\{1,2,\dots,N_\theta\}\\ q\in\{1,2,\dots,N_\varphi\}}} \widehat{w}(\alpha) \cdot e^{2i\pi\left(k\frac{\alpha}{m}\cos\theta_p\sin\varphi_q + \ell\frac{\alpha}{M}\sin\theta_p\sin\varphi_q + m\frac{\alpha}{M}\cos\varphi_q\right)}. \tag{47}$$

denoting $\widehat{w}(\alpha) = \overline{\mathrm{DFT}(h_X^d)(\alpha)} \cdot \mathrm{DFT}(h_X^d)(\alpha)$. Last, using the upper-bounds $|||\nabla_{3\mathrm{D}}||| \leq \sqrt{12}$ and $|||\mathcal{B}^*\mathcal{B}||| \leq L'_{f'} := \|\mathrm{DFT}(\psi)\|_\infty$, the convergence of the scheme (42) is ensured for $\tau = 1/(2L'_{f'})$ and $\sigma = L'_{f'}/(12\lambda^2)$. Finally, the reconstruction of a 3D image u_λ from the 3D sinogram $s_{3\mathrm{D}}$ can be done using Algorithm 6.

5 Extension to Huber Total Variation Regularization

A classical issue encountered with the minimization of TV regularized energies like $J_\lambda^{2\mathrm{D}}$ (see (32)) or $J_\lambda^{3\mathrm{D}}$ (see (41)) is the so called *staircasing artifact* which corresponds to the creation of piecewise constant regions with artificial boundaries in the reconstructed image. The latter is due to the excessive promotion of piecewise constant images by the TV regularity term in the minimization process. The staircasing issue has been reported and addressed in many different ways in the literature. A popular strategy consists in introducing in the energy to minimize some higher order terms [4, 3], or texture promoting terms based on wavelets or curvelets [30, 10]. A quite elegant (but computationally expensive) strategy to avoid staircasing was proposed in [26, 27] and is based on the computation

Algorithm 6: 3D discrete image reconstruction from a 3D discrete sinogram.

```
Inputs: a 3D discrete sinogram $s_{3D} : I_{N_B} \times \{1, 2, \dots, N_\theta\} \times \{1, 2, \dots, N_\varphi\} \to \mathbb{R}$ with size
  $N_B \times N_\theta \times N_\varphi$, the associated discrete spectrum $h_X^{d} : I_{N_B} \to \mathbb{R}$ with size $N_B$, the associated
  sequences containing the $N_\theta$ polar angles $(\theta_p)_{1 \le p \le N_\theta}$ and the $N_\varphi$ azimuthal angles
  $(\varphi_q)_{1 \le q \le N_\varphi}$, a positive integer $M \le N_B$ corresponding to the desired width (also equal to the
  height and depth) for the output image (with domain $\Omega_{3D} = I_M \times I_M \times I_M$), a regularity
  parameter $\lambda > 0$, and a number of iterations $N \ge 0$ for the numerical scheme.
Optional inputs: a flag specifying whether we aim to minimize $J_\lambda^{3D}$ over the set $\mathcal{C}' = \mathbb{R}^{\Omega_{3D}}$
  (unconstrained minimization) or over the set $\mathcal{C}' = \mathbb{R}_+^{\Omega_{3D}}$ (nonnegativity constraint), an
  initial guess $u^0 \in \mathbb{R}^{\Omega_{3D}}$ for the looked for image (if not provided, use $u^0 = \mathbf{0}_{\Omega_{3D}}$ = the
  zero-valued image in $\mathbb{R}^{\Omega_{3D}}$), the radial sampling step $\delta_r = \frac{\delta_B}{\mu}$ (if not provided, use $\delta_r = 1$).
Outputs: a numerical estimate of a minimizer $u_\lambda$ of $J_\lambda^{3D}$ over $\mathcal{C}'$.
Core of the module:
  // precompute $\mathcal{B}^* s_{3D}$ and the kernel $\psi$, initialize other variables
  Bstar_s ← $\mathcal{B}^* s_{3D}$   // using Algorithm 4 with $\delta_r$ as optional input
  psi ← $\psi$   // using (47) with $\delta = \delta_r \cdot \frac{N_B}{M}$ and a NFFT algorithm like [21]
  dft_psi ← DFT(psi)   // using the 3D FFT Algoritm [14]
  $L'_{fprime}$ ← $\max_{(\alpha,\beta,\gamma) \in \Upsilon_{3D}} |$dft_psi$(\alpha, \beta, \gamma)|$
  $\tau$ ← $1/(2L'_{fprime})$
  $\sigma$ ← $L'_{fprime}/(12\lambda^2)$
  (u, $\overline{u}$, u_prev) ← $(u^0, u^0, u^0)$
  p ← $(\mathbf{0}_{\Omega_{3D}}, \mathbf{0}_{\Omega_{3D}}, \mathbf{0}_{\Omega_{3D}})$
  // main loop (iterations of the scheme (42))
  for $1 \le n \le N$ do
      // compute $\nabla f'(u) = \mathcal{B}^*\mathcal{B}v - \mathcal{B}^* s_{3D}$
      dft_Zu ← DFT($Zu$)   // where $Zv$ is obtained from $v$ using (46)
      dft_w ← dft_psi · dft_Zu   // we have dft_w = DFT($\psi \circledast Zu$)
      BstarB_u ← the restriction of IDFT(dft_w) to $\Omega_{3D}$
      gradfprime_u ← BstarB_u − Bstar_s
      // perform one scheme iteration
      p ← $\Pi_{\mathcal{B}_{3D}}$(p + $\sigma\lambda\nabla_{3D}\overline{u}$)   // using (43)
      u ← u − $\tau$ · (gradfprime_u − $\lambda\,\mathrm{div}_{3D}$(p))   // using (69)
      if $\mathcal{C}' = \mathbb{R}_+^{\Omega_{3D}}$ then u ← max($\mathbf{0}_{\Omega_{3D}}$, u)
      // projection into $\mathcal{C}'$ by taking the positive part
      $\overline{u}$ ← 2u − u_prev
      u_prev ← u
  return the 3D discrete image u   // sampling step (or pixel size) is $\delta = \delta_r \cdot \frac{N_B}{M}$ (same unit as $\delta_r$)
```

of the posterior mean of a TV-based distribution. A very interesting and computationally efficient variant of this latter approach was also proposed in [28] but remains, to this date, restricted to image denoising. In this section, we would like to briefly focus on a very simple variant of TV, called the Huber TV, which can be used to avoid staircasing and generate more natural images than those obtained with the classical TV [5]. In the 2D framework, the Huber TV of a discrete image $v \in \mathbb{R}^{\Omega_{2D}}$

is defined as

$$\mathrm{HTV}_\alpha(v) = \sum_{(k,\ell)\in\Omega_{\text{2D}}} \mathcal{H}_\alpha^{\text{2D}}(\nabla_{\text{2D}} v(k,\ell))\,, \tag{48}$$

where $\alpha \geq 0$ is a smoothing parameter and $\mathcal{H}_\alpha^{\text{2D}} : \mathbb{R}^2 \to \mathbb{R}$ is the two-dimensional *Huber function* defined by

$$\forall z \in \mathbb{R}^2\,, \quad \mathcal{H}_\alpha^{\text{2D}}(z) = \begin{cases} \frac{\|z\|_2^2}{2\alpha} & \text{if } \|z\|_2 < \alpha \\ \|z\|_2 - \frac{\alpha}{2} & \text{otherwise.} \end{cases} \tag{49}$$

Function $\mathcal{H}_\alpha^{\text{2D}}$ corresponds to a smooth approximation of the ℓ^2 norm in $\mathbb{R}^2$ (more precisely, to the Moreau-Yoshida envelope with parameter α of the ℓ^2 norm). Notice that for $\alpha = 0$, (49) yields $\mathcal{H}_0 = \|\cdot\|_2$, and thus, $\mathrm{HTV}_0 = \mathrm{TV}$. When $\alpha > 0$, we can simply handle the minimization of the Huber-TV regularized energy

$$J_{\lambda,\alpha}^{\text{2D}} : v \mapsto \frac{1}{2}\|\mathcal{A}(v) - s_{\text{2D}}\|_2^2 + \lambda\,\mathrm{HTV}_\alpha(v) \tag{50}$$

using the numerical scheme (35) provided that we replace the dual update (35a) by

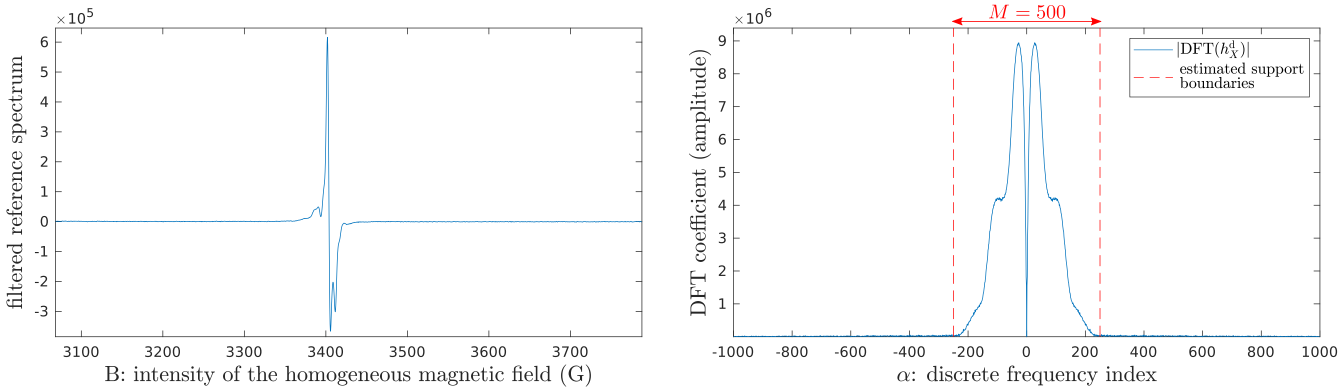
$$p^{n+1} = \Pi_{\mathcal{B}_{\text{2D}}}\left(\frac{p^n + \sigma\lambda\nabla_{\text{2D}}\overline{v}^n}{1 + \lambda\alpha\sigma}\right). \tag{51}$$

Consequently, the minimization of $J_{\lambda,\alpha}^{\text{2D}}$ can be done using Algorithm 5 provided that we replace the update $\mathtt{p} \leftarrow \Pi_{\mathcal{B}_{\text{2D}}}(\mathtt{p} + \sigma\lambda\nabla_{\text{2D}}\overline{\mathtt{v}})$ by $\mathtt{p} \leftarrow \Pi_{\mathcal{B}_{\text{2D}}}\left(\frac{\mathtt{p}+\sigma\lambda\nabla_{\text{2D}}\overline{\mathtt{v}}}{1+\lambda\alpha\sigma}\right)$. The same extension can be done in the 3D framework by replacing the update $\mathtt{p} \leftarrow \Pi_{\mathcal{B}_{\text{3D}}}(\mathtt{p} + \sigma\lambda\nabla_{\text{3D}}\overline{\mathtt{u}})$ by $\mathtt{p} \leftarrow \Pi_{\mathcal{B}_{\text{3D}}}\left(\frac{\mathtt{p}+\sigma\lambda\nabla_{\text{3D}}\overline{\mathtt{u}}}{1+\lambda\alpha\sigma}\right)$ in Algorithm 6. TV and Huber-TV image reconstruction models will be compared in Section 7.2.2.

6 Acquisition Process and Parameters Setting

6.1 Acquisition of the Reference Spectrum

An EPR experiment usually starts with the acquisition of the signal h_X^{d} defined in (27). This signal corresponds to the filtered and sampled reference spectrum of the specie X. It can be obtained (up to a multiplicative factor corresponding to the total amount Q_X of substance X present in the cavity) by running an EPR acquisition in presence of the homogeneous magnetic field and without any field gradient (i.e., using $\mu = 0$, as in the spectroscopy framework). The acquired signal can be renormalized afterward to get rid of the multiplicative factor Q_X, but we will ignore this step as we do not focus on quantitative image reconstruction (e.g. concentration estimation) and it has no impact on the final image quality. The acquisition of the signal h_X^{d} requires from the operator to adjust the center field B_{CF}, the sweep width B_{SW} and the number of sampling points N_B. For that purpose, the operator tunes the acquisition range $B_{\text{CF}} + \left[-\frac{B_{\text{SW}}}{2}, \frac{B_{\text{SW}}}{2}\right]$ so that all the signal of interest (i.e., all the rays of the reference spectrum) is captured, as displayed in Figure 4 (a). During the first acquisition attempts, the number of sampling points N_B must be set large enough to capture the signal oscillations. At this step of the acquisition process, we recommend to the operator to set a voluntary *too large* value for N_B (i.e., to perform an oversampled acquisition) as we will explain how this parameter can be tuned from h_X^{d} afterward. As mentioned earlier, during this adjustment step, care should be taken to ensure that once the field gradient is applied, the totality of the sample remains included in the FOV (recall that the FOV diameter is $D_{\text{FOV}} = B_{\text{SW}}/\mu$). Also, it must be noted that the acquired signal is corrupted by noise. In practice, several other parameters that we did not explicitly take into account in our model (such as, for instance, time integration parameters,



(a) filtered and sampled reference spectrum h_X^{d} (b) DFT amplitude of h_X^{d} and support estimation

Figure 4: **Acquisition of the filtered reference spectrum and frequency support estimation.** We display in (a) the spectrum acquisition h_X^{d} of an irradiated distal phalanx that will be presented and studied in more details in Section 7.2. The spectrum h_X^{d} was sampled using $N_B = 2000$ points and a sweep width B_{sw} equal to 719.44 G, leading to a sampling step $\delta_B = \frac{B_{\mathrm{sw}}}{N_B} \approx 0.36$ G. We display in (b) the DFT amplitude of h_X^{d}. We can see that the amplitude of the DFT coefficients of h_X^{d} rapidly drop below the noise level so that we can consider that they vanish outside from the *estimated DFT support* delimited by the red dashed lines. The size of the estimated DFT support is $M = 500$ which represents a proportion of $\frac{M}{N_B} = \frac{1}{4}$ of the total number of samples. This means that the signal h_X^{d} was oversampled with a factor four and that another acquisition of this signal with four times less samples (i.e., using $\delta_B = \frac{B_{\mathrm{sw}}}{M} \approx 1.44$ G) should theoretically not cause any loss of relevant information compared to the oversampled acquisition (a).

amplitude and frequency modulation parameters, etc.) may strongly influence the signal to noise ratio (SNR) of the acquired signal h_X^{d}.

6.2 Optimal Number of Samples and Sinogram Acquisition

At this time of the acquisition process, we performed the acquisition of h_X^{d} with a voluntary *too large* number of samples N_B. Although, in general, the acquisition of the spectrum h_X^{d} can be done in a reasonably short time despite the too large setting of N_B, this is usually not the case for the upcoming sinogram acquisition. Indeed, with N_B (and all other acquisition parameters) kept unchanged, the sinogram acquisition time roughly corresponds to the time necessary to acquire h_X^{d} multiplied by the number of projections (typically several tens to several hundred). As the acquisition time is often a limiting factor for EPR imaging applications, the setting of N_B must be optimized before starting the sinogram acquisition. On the one hand, the number of sampling points N_B must be set large enough to ensure that the sampling step $\delta_B = B_{\mathrm{sw}}/N_B$ is small enough to perform a correct sampling of the signal (more precisely δ_B must be below the critical sampling step ν imposed by the acquisition system cut-band filtering g_ν described in Section 2.4). On the other hand, N_B must also be kept small enough to avoid too long acquisition time for the sinogram. Besides, too large values for N_B yield a subcritical sampling step $\delta_B < \nu$ and correspond to oversampling situations where the acquired signal h_X^{d} does not contain more information than the signal acquired using $\delta_B = \nu$ (see Theorem (2)). Therefore, a somehow optimal setting for N_B corresponds to the situation where $\delta_B = \nu$, i.e.,

$$\delta_B = \frac{B_{\mathrm{sw}}}{N_B} = \nu \quad \Leftrightarrow \quad N_B = \frac{B_{\mathrm{sw}}}{\nu} \underset{\text{for any } \mu \neq 0}{=} \frac{B_{\mathrm{sw}}}{\mu\,\delta} = M \tag{52}$$

since $\delta = \nu/\mu$ (see Section 2.4). Finally we see that the optimal value for N_B is simply equal to M (which is not surprising since $M = N_B$ corresponds to the critical sampling situation). However, with our EPR acquisition system, the actual value of ν used during the acquisition is unknown (we believe this is also the case for other EPR setups), so that we cannot directly compute the actual value of M. Thankfully, it can be estimated from the signal h_X^{d} (acquired with $N_B > M$). Indeed, thanks to (27), in the Fourier domain, we have

$$\forall \alpha \in I_{N_B}\,, \quad \frac{1}{\delta_B} \cdot \mathrm{DFT}(h_X^{\mathrm{d}})(\alpha) \approx \mathcal{F}(g_\nu * h_X)\left(\frac{2\pi\alpha}{\delta_B N_B}\right) = \begin{cases} \mathcal{F}(h_X)\left(\frac{2\pi\alpha}{\delta_B N_B}\right) & \text{if } \left|\frac{2\pi\alpha}{\delta_B N_B}\right| \leq \frac{\pi}{\nu} \\ 0 & \text{otherwise.} \end{cases}$$

Since $\left|\frac{2\pi\alpha}{\delta_B N_B}\right| \leq \frac{\pi}{\nu}$ is equivalent to $|\alpha| \leq M/2$, we can estimate M (and thus ν) by finding the boundaries of the support of $\mathrm{DFT}(h_X^{\mathrm{d}})$. In practice, due to the presence of noise corrupting the signal h_X^{d}, its DFT coefficients do not vanish. Therefore, we suggest to consider that the support of $\mathrm{DFT}(h_X^{\mathrm{d}})$ corresponds to the positions $\alpha \in I_{N_B}$ such as $\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)$ is significantly above the noise level, as illustrated in Figure 4 (b). This can be done by visual inspection of the signal $|\mathrm{DFT}(h_X^{\mathrm{d}})|$, but we also provide an automatic frequency support detection algorithm for that purpose. The latter is based on the *a contrario* methodology [8] and is described in Appendix E. Once M is known, we can set $N_B = M$ and run the sinogram acquisition. Notice that since the estimated value of M is dependent from the noise level, it is important to keep the same noise level as for h_X^{d} during the sinogram acquisition (i.e., except for the number of samples, all the acquisition parameters must be kept unchanged). Regarding the filtered reference spectrum h_X^{d}, it can be acquired again with $N_B = M$, or computed from the oversampled acquisition by means of a frequency cutoff.

6.3 Optimal Resolution for the Image Reconstruction

The value of M corresponds to the number of pixels along each direction of the discrete images $\widetilde{v}_X$ (in the 2D setting) and $\widetilde{u}_X$ (in the 3D setting) obtained by sampling the bandlimited $\widetilde{V}_X$ and $\widetilde{U}_X$ over the FOV domain using the critical sampling step δ. The value of the critical sampling step δ can be easily derived from M using (52), or equivalently, using

$$\delta = \delta_r \frac{N_B}{M} \quad \text{with} \quad \delta_r = \frac{B_{\mathrm{sw}}}{\mu N_B}\,, \tag{53}$$

as we already mentioned before. As we will check experimentally, there is no valuable reason for computing an image with spatial sampling step $\delta' < \delta$ since such image cannot contain more details than the critically sampled one. On the contrary, computing an image with spatial sampling step $\delta' > \delta$ yields in general a less detailed image than the critically sampled one. In practice, in Algorithm 5 and Algorithm 6, we let to the user the ability to select the image reconstruction size M' (and thus the image sampling step $\delta' = \delta_r \frac{N_B}{M'}$) for the following reasons:

(i) Most of our EPR acquisitions were done in oversampled situations (with oversampling factor up to 4, as in Figure 4), leaving door for the setting of a reconstruction size $M' > M$. We want to be able to deal with such oversampled acquisitions and check that the reconstruction of an image with sampling step $\delta' < \delta$ does not yield a better level of details compared to the image obtained using $M' = M$ (and sampling step $\delta' = \delta$).

(ii) On the contrary, setting an image reconstruction size $M' < M$ amounts to select a sub-optimal sampling step $\delta' > \delta$ which may lead to a degraded level of details for the computed image. However, such situation also yields a faster reconstruction, leaving door for efficient multiscale image reconstruction schemes. In practice, the fast reconstruction of a coarse scaled image is helpful to tune the regularity parameter λ before refining the processing at finer scales.

(iii) The actual optimal level of details that we can obtain in the reconstructed image actually depends on other acquisition parameters (such as the number of projections) that are not taken into account here. Thus, the reconstruction size M corresponds in practice to an upper bound of the actual optimal reconstruction size.

We believe that offering to the user an accurate estimate of the optimal reconstruction size M from an oversampled spectrum acquisition like the one displayed in Figure 4 (a) is an important step forward regarding the optimization of the acquisition and reconstruction process. From the optimal value of M, the value of the radial sampling step δ_r and the value of the field gradient intensity μ, the user is able to compute the value of the critical sampling step δ which can be roughly interpreted as the smallest detail size that we can hope to observe in the reconstructed image and which is often a crucial information for the EPR imagist. Besides, knowing the optimal value for M is also helpful to avoid spending unnecessary large computation time in the reconstruction of an image with too large dimensions.

6.4 Tuning of the Regularity Parameter

The reconstruction of a 2D image using Algorithm 5 (respectively of a 3D image using Algorithm 6) requires the setting of the regularity parameter λ involved in the energy J_λ^{2D} defined in (32) (respectively in the energy J_λ^{3D} defined in (41)). This parameter controls the relative importance of the least-squares data-fidelity term with respect to the TV regularity term in the energy minimization process. In general, the range of *appropriate* values for λ (that is, leading to satisfactory image reconstructions) strongly depends on the amounts of noise corrupting the data, but also on modeling errors, on the content of the (unknown) image to reconstruct, and also on more complex factors such as some user needs regarding the reconstruction. Therefore, deriving an automatic setting for this parameter is a complex task that we will not address here. However, we believe that the tuning of λ can be drastically eased provided that we apply an appropriate rescaling of the energy to minimize.

Let us focus first on the 2D setting. We assume that the polar angle grid $(\theta_p)_{1\leq p\leq N_\theta}$ associated to the 2D sinogram s_{2D} is regularly spaced with sampling step δ_θ. This assumption is not mandatory but simplifies the discussion and corresponds to the classical acquisition setting using Bruker® instruments. Then, instead of tuning λ, we suggest to set

$$\lambda = 10^5 \frac{\lambda'\delta}{\delta_B\delta_\theta}, \tag{54}$$

and let to the user the task of tuning λ'. The reason for this choice is that, minimizing J_λ^{2D} with λ satisfying (54) is equivalent to minimize the normalized energy $\widetilde{J}_{\lambda'}$ defined by

$$\forall v \in \Omega_{\mathrm{2D}}\,, \quad \widetilde{J}_{\lambda'}^{\mathrm{2D}}(v) = \frac{\delta_B\delta_\theta}{2\cdot 10^5}\left\|\mathcal{A}v - s_{\mathrm{2D}}\right\|_2^2 + \lambda'\delta\,\mathrm{TV}(v), \tag{55}$$

where the normalized least-squares term $\frac{\delta_B\delta_\theta}{2}\|\mathcal{A}v - s_{\mathrm{2D}}\|_2^2$ and the normalized regularity term $\delta\,\mathrm{TV}(v)$ are less dependent on the sampling steps δ, δ_B and δ_θ than the unnormalized least-squares term $\frac{1}{2}\|\mathcal{A}v - s_{\mathrm{2D}}\|_2^2$ and the unnormalized regularity term $\mathrm{TV}(v)$. Practically speaking, this means that an appropriate tuning of the normalized regularity parameter λ' can be used almost unchanged for reconstructions at different sampling steps δ from the same sinogram s_{2D}, or from a sinogram acquisition to another with different sampling steps δ_B and δ_θ. Notice that the multiplicative factor 10^5 in (54) has no physical meaning and was introduced to reduce the amplitude of the appropriate settings for λ' over our own EPR acquisitions.

In the 3D setting, we add the assumption that the azimuth angle grid $(\varphi_q)_{1\leq q\leq N_\varphi}$ is regularly spaced with sampling step δ_φ. Then, we set

$$\lambda = 10^8 \frac{\lambda'\delta^2}{\delta_B\,\delta_\theta\delta_\varphi}, \tag{56}$$

so that minimizing J_λ^{3D} with λ satisfying (56) is equivalent to minimizing the normalized energy $\widetilde{J}_{\lambda'}^{\text{3D}}$ defined by

$$\forall u \in \Omega_{\text{3D}}\,, \quad \widetilde{J}_{\lambda'}^{\text{3D}}(u) = \frac{\delta_B \delta_\theta \delta_\varphi}{2 \cdot 10^8} \|\mathcal{B}u - s_{\text{3D}}\|_2^2 + \lambda' \delta^2 \, \text{TV}(u)\,, \tag{57}$$

and provides the same practical advantages as those presented in the 2D setting in terms of invariance of the tuning of the normalized parameter λ' with respect to the sampling steps δ, δ_B, δ_θ and δ_φ.

6.5 Huber Smoothing Parameter

Using the Huber TV image reconstruction model described in Section 5 requires the setting of the Huber-smoothing parameter α. As the λ parameter, α is unnormalized in the sense that its influence on the visual aspect of the reconstructed image strongly depends on the value of the spatial sampling step δ used for the reconstruction. Such dependence complicates the parameter tuning process when reconstructions with different sampling steps must be done. To avoid this phenomenon, we introduce the rescaling

$$\alpha = \alpha' \delta \quad \text{with} \quad \alpha' \geq 0, \tag{58}$$

letting to the user the task of setting the value of the normalized parameter α'. Combining the rescaling operations (54) and (58) in (50) provides, for a given choice of (λ', α'), images with similar dynamic and comparable visual aspect for a large range of values for the spatial sampling δ.

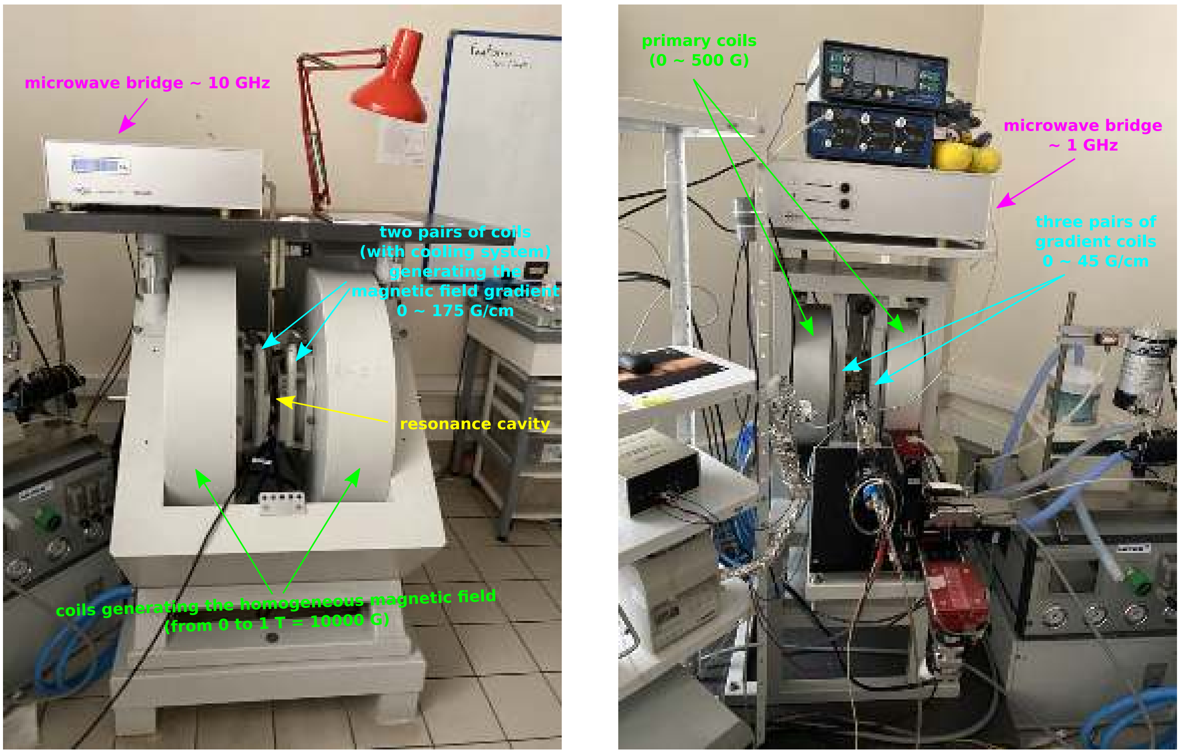
7 Experiments over Real EPR Data

7.1 EPR Acquisition Systems

In Figure 5, we display the pictures the two EPR spectrometers (and imaging systems) that we used to perform the acquisition of all the datasets studied in this section. In both instruments, the resonance cavity is surrounded by two large primary coils generating the homogeneous magnetic field inside the cavity. Secondary coils are placed in between the primary coils and are used to generate the magnetic field gradient needed for EPR imaging purpose. A microwave bridge is used to apply an electromagnetic wave to the sample and detect energy absorption caused by electron resonance. The spectrometer displayed in Figure 5 (a) is operated in X-Band, meaning that a microwave with frequency of roughly 10 GHz is applied to the sample, yielding a single electron resonance magnetic field characteristic value of around 3500 G. The EPR acquisition system displayed in Figure 5 (b) is operated in L-band, with a microwave around 1 GHz, leading to a characteristic resonance magnetic field of around 350 G. Due to dielectric losses, the penetration of the microwave in sample containing water is around 1 mm at X-Band and few cm at L-band. The L-band spectrometers are used for in vivo experiments and samples with large dimensions, while X-band ones are used for in vitro experiments in flat cells and capillary tubes.

7.2 2D Image Reconstruction of an Irradiated Phalanx

In this section, we focus on 2D image reconstruction of an irradiated distal phalanx (see Figure 6 (a)). Irradiations generate paramagnetic damages of the bone lattice microarchitecture that can be observed using EPR imaging [23, 10]. The spectrum displayed in Figure 4 (a) corresponds to a real X-band EPR acquisition obtained from this phalanx sample and a 2D sinogram acquisition is also displayed in Figure 6 (b). A non exhaustive list and values (or range of values) of the acquisition parameters that we used in our experiments is displayed in Table 1. The exact setting for the parameters μ, N_B and $\delta_r = \frac{B_{\text{sw}}}{\mu N_B}$ will be precised for each experiment. As regard the reconstruction



(a) X-band EPR 2D imaging system (b) L-band EPR 2D and 3D imaging system

Figure 5: **EPR spectrometers and imaging systems installed at Université Paris Cité.** We display some pictures of the X-band (a) and L-band (b) spectrometers hosted at LCBPT. In (a), the primary coils generating the homogeneous magnetic field are indicated with green arrows. Secondary coils (two pairs packaged together are indicated with blue arrows) can be used to apply a planar magnetic field gradient with intensity up to $175\,\mathrm{G/cm}$ in the cavity for EPR imaging purposes. As we shall see, such large field gradient intensity allows for high resolution image reconstruction. However, the space in between the primary coils is insufficient to stand the third pair of gradient coils needed to achieve non-planar 3D field gradient orientations. Therefore the X-band device (a) is restricted to 2D imaging. The L-band spectrometer (b) is equipped with three pairs of gradient coils and can be used for 3D EPR imaging acquisitions. However, the signal intensity in L-band spectrometer is a magnitude lower than the one in X-band system because the lower homogeneous magnetic field generated by its primary coils (up to only $0.05\,\mathrm{T} = 500\,\mathrm{G}$) leads to lower separation of the energy levels of the single electrons. Besides, the field gradient intensity cannot be larger than $45\,\mathrm{G/cm}$ on this L-band instrument. In general, EPR images reconstructed from L-band acquisition are less resolved than those obtained with X-band.

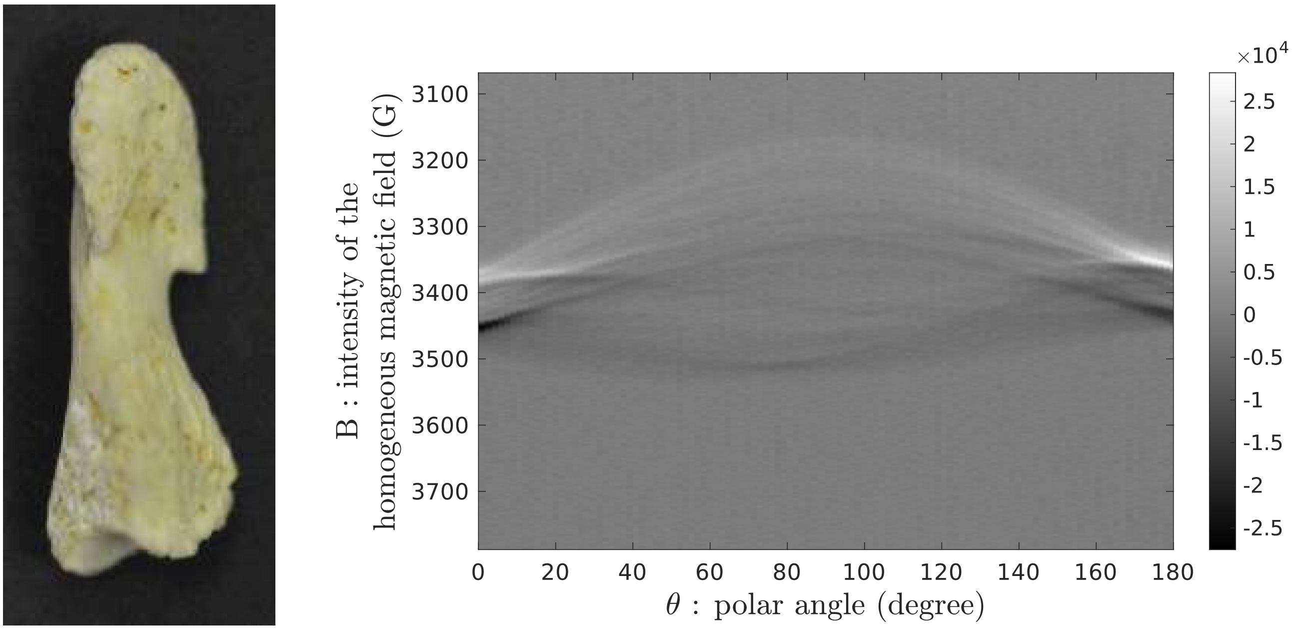
parameters, we will provide for each experiment the value of the spatial sampling step δ and the value of the normalized regularity parameter λ'. The effective value of the (unnormalized) parameter λ can be obtained using (54) (with δ_B expressed in Gauss unit, δ_θ in radian unit and δ in millimeter unit). Notice that all images displayed in this section correspond to the positive part of those returned by Algorithm 5.

7.2.1 Setting of the TV Regularity Parameter

As a first dataset, we consider the spectrum displayed in Figure 4 (a) and the sinogram displayed in Figure 6 (b). This dataset was acquired using $N_B = 2000$ samples per projection, a field gradient intensity $\mu = 168$ G/cm and a radial sampling step $\delta_r \approx 21.4$ µm. In Figure 7, we display the images reconstructed from this dataset for various values of normalized regularity parameter λ'. This experiments illustrates the benefits (and limits) of TV regularization in the image reconstruction process.

name	description	value (or range)	unit
μ	field gradient intensity	$42 \sim 168$	G/cm
N_B	number of samples per projection	$500 \sim 2000$	–
δ_r	radial sampling step	$21.4 \sim 85.6$	µm
B_{CF}	center field	3428.20	G
B_{SW}	sweep width	$\mu\, N_B\, \delta_r \cdot 10^{-4}$	G
δ_B	homogeneous magnetic field intensity sampling step	$\frac{B_{\text{SW}}}{N_B}$	G
N_θ	number of projections	113	–
δ_θ	angular sampling step	1.6	degree
τ_c	time constant	163	ms
–	conversion time	same as τ_c	ms
–	microwave frequency	9.56	GHz
–	microwave power	9.99	mW
–	power attenuation	33	dB
–	amplitude modulation	3	G
–	frequency modulation	100	kHz

Table 1: **Main acquisition parameters used with the phalanx sample (X-band).** We provide here the values (or range of values) of the acquisition parameters that we used to perform spectra and 2D sinogram acquisitions presented in this section. The first height parameters ($\mu, N_B, \delta_r, B_{\text{CF}}, B_{\text{SW}}, \delta_B, N_\theta, \delta_\theta$) were explicitly taken into account in the mathematical models described in this work. This is not the case for the remaining parameters although they are in practice important to the EPR spectroscopists. Their accurate tuning strongly impacts the measurements quality, and thus, indirectly, the overall image reconstruction.

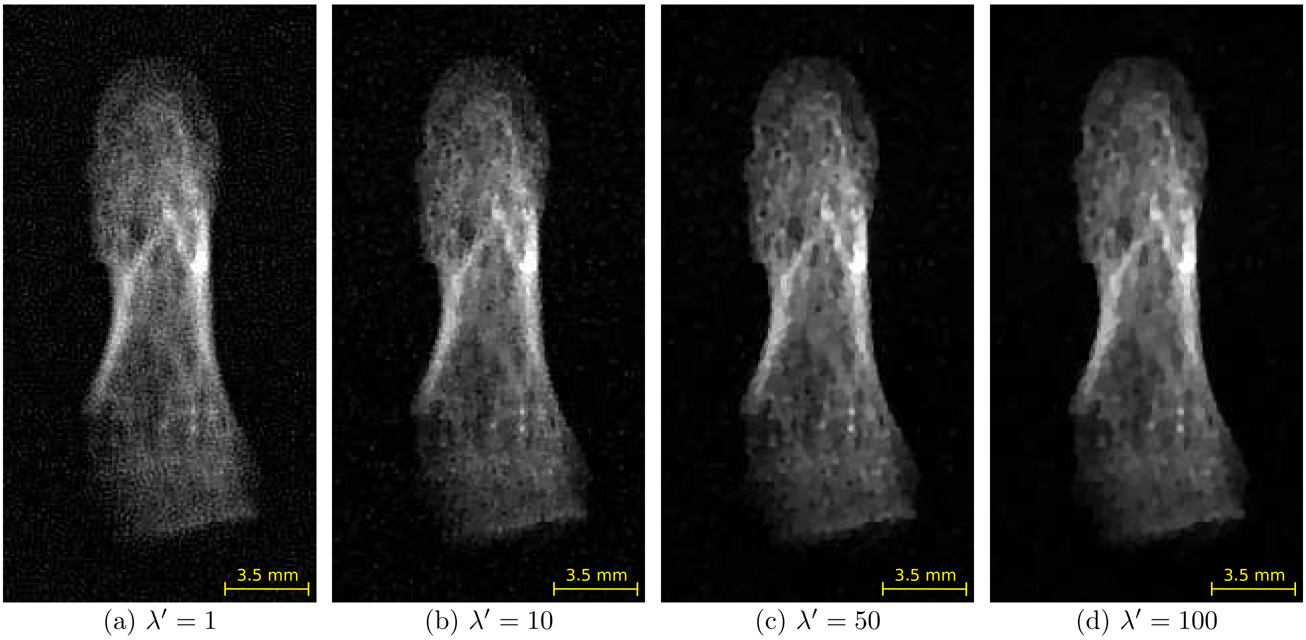


(a) irradiated distal phalanx (b) example of measured 2D sinogram

Figure 6: **Irradiated phalanx and a 2D sinogram acquisition.** We display (a) a picture of the distal phalanx and (b) a 2D sinogram acquisition obtained using $N_B = 2000$ samples per projections, a field gradient intensity $\mu = 168$ G/cm and a radial sampling step $\delta_r = \frac{B_{\text{SW}}}{\mu N_B} \approx 21.4$ µm. A filtered and sampled reference spectrum (linewidth=3.8 G) associated to this 2D sinogram acquisition is displayed in Figure 4 (a).

7.2.2 Huber-TV Based Image Reconstruction

We explained in Section 5 how Algorithm 5 could be extended to handle the minimization of the Huber-TV regularized energy (50). This model involves the tuning of another (normalized) parameter



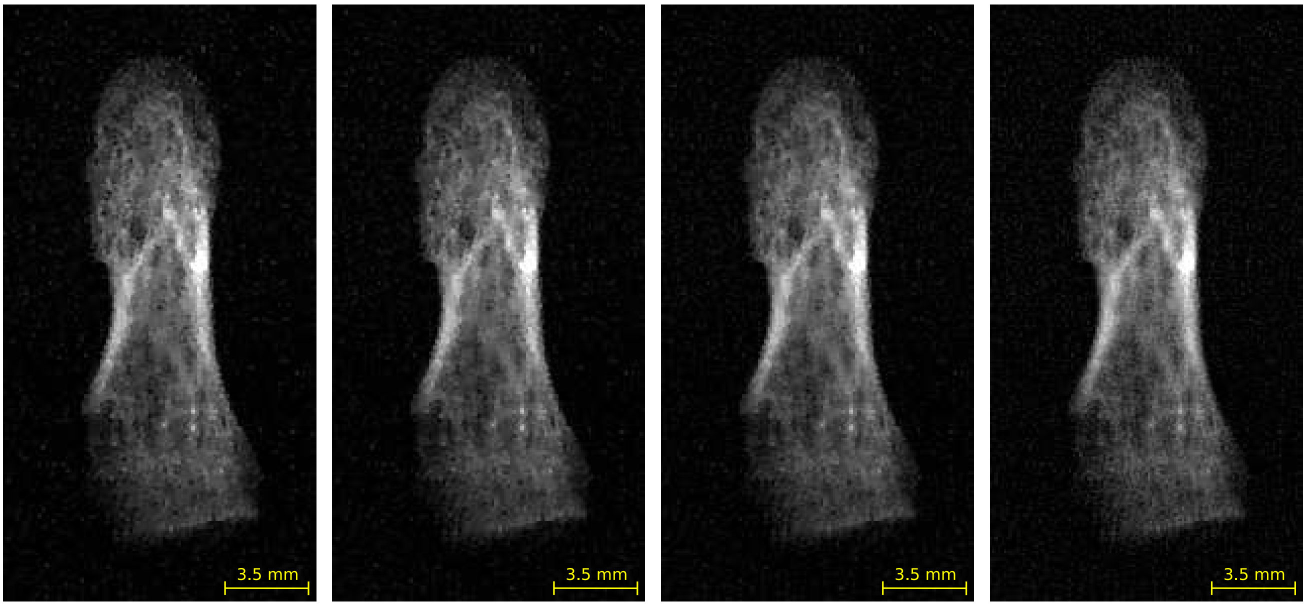
(a) $\lambda' = 1$ (b) $\lambda' = 10$ (c) $\lambda' = 50$ (d) $\lambda' = 100$

Figure 7: **Setting of the regularity parameter.** We used Algorithm 5 to process the spectrum and the sinogram displayed in Figure 4 (a) and Figure 6 (b) for various values of λ'. All reconstructions were made using $M = 500$ (which corresponds to a spatial sampling step $\delta = \delta_r \frac{N_B}{M} \approx 85.6$ µm and is the recommended setting for this dataset according to Section 6.3). One can see that the setting $\lambda' = 1$ (see (a)) yields an image with important residual noise due to a lack of regularization. Increasing λ' yields images with increasing regularity (see (b), (c), and (d)) but also involves a loss of details for too large values of λ'. In the present experiment, the setting $\lambda' = 10$ (see (b)) seems to achieve a correct tradeoff between image regularity and detail preservation.

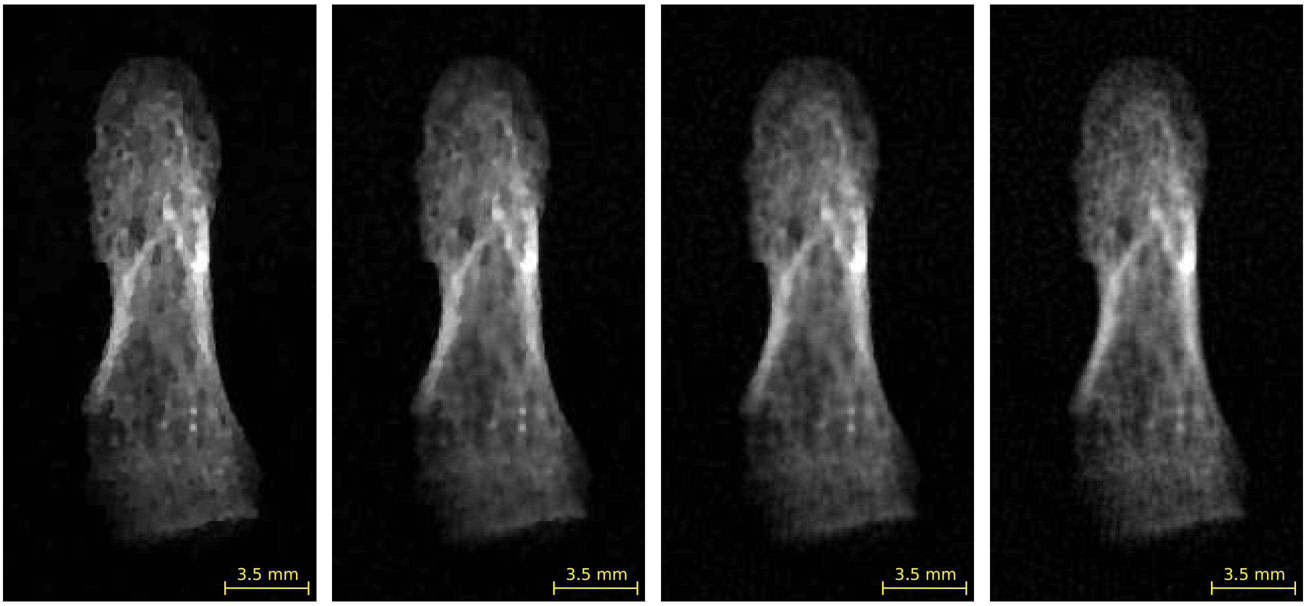
α' (see Section 6.5), and can be viewed as a generalization of (32) since both models are identical when $\alpha' = 0$. In Figure 8, we used the same dataset as in Figure 7 to illustrate how the Huber smoothing parameter may impact the visual aspect of the reconstruction. This experiment illustrates that, for an appropriate setting of (λ', α'), the Huber-TV generalization of the image reconstruction model can lead to images with more natural aspect. However the improvements in terms of texture rendering also come with two main drawbacks. First, the setting $\alpha' > 0$ yields more blurry images than the setting $\alpha' = 0$. Second, the tuning of α' is dependent from that of λ' (and reciprocally), so that, in practice, the joint tuning of (λ', α') is a more complicated task than the tuning of the single parameter λ' in the case $\alpha' = 0$. Besides, the value of the field gradient intensity ($\mu = 168$ G/cm) used for this dataset is close to the highest value allowed by the acquisition system. Most EPR acquisitions are done with smaller values of field gradient, leading to less detailed images where micro-texture rendering is not a relevant issue (this will be illustrated in Section 7.2.5). In practice, we believe that one would probably prefer to set $\alpha' = 0$ rather than jointly tuning (λ', α') for processing most EPR acquisitions.

7.2.3 Setting of the Reconstruction Resolution

As in the previous section, we consider the dataset made of the spectrum displayed in Figure 4 (a) and the sinogram displayed in Figure 6 (b), acquired using $N_B = 2000$ samples per projections, a field gradient intensity $\mu = 168$ G/cm and a radial sampling step $\delta_r \approx 21.4$ µm. In Figure 9, we display images reconstructed from this dataset using Algorithm 5 with various settings of the image reconstruction size M, corresponding to various settings for the spatial sampling step $\delta = \delta_r \frac{N_B}{M}$. This experiment illustrates that the optimal reconstruction size for this dataset is close to $M = 500$ (which corresponds to a spatial sampling step $\delta \approx 85.6$ µm), as predicted in Figure 4 (b) using the methodology presented in Section 6.3. In particular, it must be noted that image reconstruction using



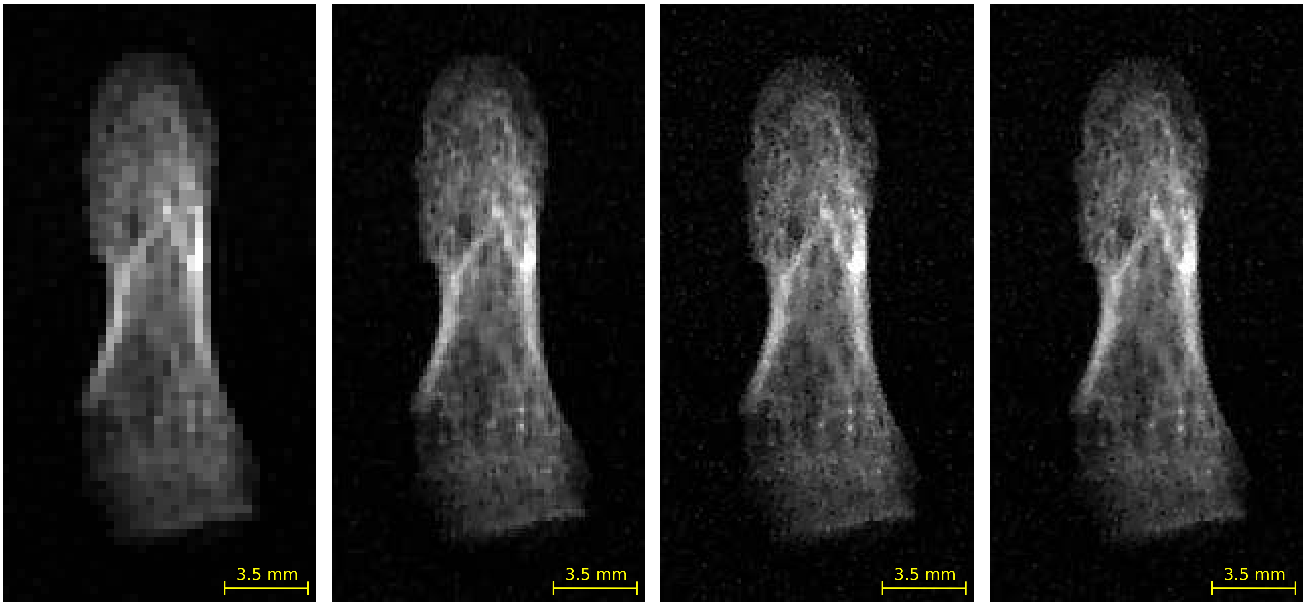
(a) $\lambda' = 10,\ \alpha' = 0$ (b) $\lambda' = 10,\ \alpha' = 0.01$ (c) $\lambda' = 10,\ \alpha' = 0.02$ (d) $\lambda' = 10,\ \alpha' = 0.05$



(e) $\lambda' = 100,\ \alpha' = 0$ (f) $\lambda' = 100,\ \alpha' = 0.01$ (g) $\lambda' = 100,\ \alpha' = 0.02$ (h) $\lambda' = 100,\ \alpha' = 0.05$

Figure 8: **Huber-TV based image reconstruction.** In this experiment, we considered the extension of Algorithm 5 presented in Section 5, which addresses image reconstruction using a Huber-TV regularized model. The latter relies on an additional parameter $\alpha = \alpha'\delta$ (see Section 6.5), with $\alpha' \geq 0$ that must be set by the user. The TV-regularized image reconstruction model corresponds to the setting $\alpha' = 0$ (since $\mathrm{HTV}_0 = \mathrm{TV}$), and leads to the images displayed in Figure 7. In the first row, we display the images obtained from the same dataset for $\lambda' = 10$ and several values of α'. The staircasing effect can be observed in (a), almost no difference can be observed in (b) while we can see in (c) that the setting $\alpha' = 0.02$ yields a slightly more textured image. However, for $\alpha' \geq 0.05$, the Huber TV regularizer is not able to efficiently deal with the noise (see (d)). Better handling of the noise remains possible by increasing the value of λ', as illustrated in the second row. This time, the setting $\alpha' \in \{0.01, 0.02\}$ allows to get rid of the strong staircasing effect observed in (e). For $\alpha' = 0.05$ (see (h)), the large regularization weight ($\lambda' = 100$) avoids the noise amplification observed in (d) and yields a somehow more natural image than (a) or (e), but also a more blurry reconstruction. Notice that this dataset was acquired using a strong field gradient ($\mu = 168$ G/cm) to allow for high resolution reconstruction of this highly textured phalanx sample. This makes the staircasing artifacts especially visible in the reconstruction.

$M > 500$ is computationally expensive and does not lead to a more detailed reconstruction at the end of the process. Note also that the reconstructions with different sizes and sampling steps displayed in Figure 9 were obtained using the common setting $\lambda' = 10$ for the normalized regularity parameter leading to images with very similar dynamics and SNR. In practice, and as announced in Section 6.4, this means that the user can tune the value of λ' at a coarse resolution scale (taking benefit from fast computations) before performing the reconstruction of an image with finer resolution keeping λ' unchanged.

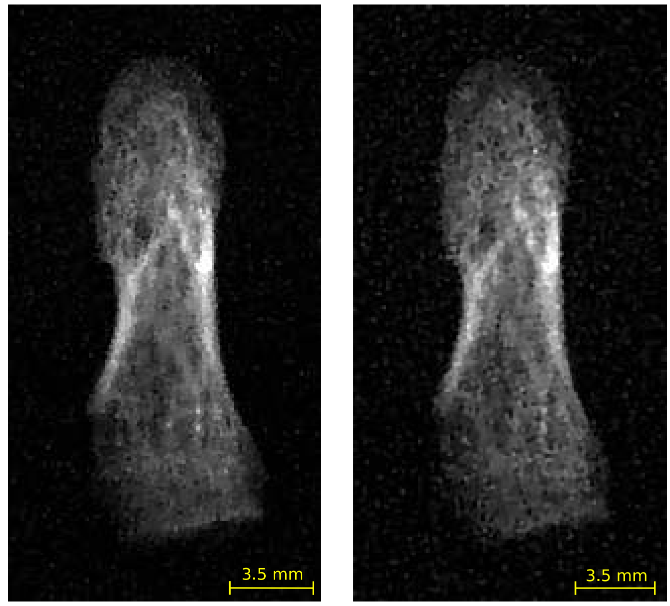


(a) $\delta = 340$ μm (M=126) (b) $\delta = 171$ μm (M=250) (c) $\delta = 85.6$ μm (M=500) (d) $\delta = 42.8$ μm (M=1000)

Figure 9: **Reconstructions at different resolutions from the same sinogram acquisition.** We used Algorithm 5 to process the discrete spectrum h_X^{d} displayed in Figure 4 (a) and the 2D sinogram $s_{2\mathrm{D}}$ displayed in Figure 6 (b). The reconstruction was done for various image reconstruction sizes M using the same setting $\lambda' = 10$ for the normalized regularity parameter and $\alpha' = 0$ for the normalized Huber smoothing parameter. The spatial sampling step δ associated to each reconstructed image is $\delta = \delta_r \frac{N_B}{M}$. According to Section 6.3 and Figure 4 (b), the recommended reconstruction size for this dataset is $M = 500$ (which corresponds to $\delta \approx 85.6$ µm). Comparing (a) to (b) and (c) we can see that the level of details in the reconstructed images increases as δ decreases. However, we observe no perceptual improvements in (d) compared to (c), as predicted in Figure 4 (b).

7.2.4 Tuning of the Number of Samples for the Sinogram Acquisition

In Section 6.2, we claimed that the value of M estimated from the oversampled spectrum acquisition also corresponds to the optimal value for N_B (the number of samples per projection). Typically, the acquisition of the sinogram with $N_B = 2000$ samples per projection displayed in Figure 6 (b) lasted in around 11 hours. Running the acquisition with $N_B = 500$ instead of $N_B = 2000$ and keeping all the other acquisition parameters unchanged yields a four time faster acquisition (≈ 2.75 hours). In Figure 10, we put this claim to the test and show that, although we did not expect a sinogram acquisition made with $N_B > 500$ samples per projection to contain more relevant information than a fast acquisition done using $N_B = M = 500$ samples per projection, in practice, a slight degradation is observed in the reconstruction obtained from the fast acquisition. Possible explanations for this phenomenon are proposed in Figure 11. Fortunately enough, the image degradation observed in the image reconstructed from the fast sinogram acquisition seems quite moderate and we believe that the fast acquisition remains competitive regarding the reduction of the acquisition time (by a factor 4) that it offers.



(a) $N_B = 2000$ samples/proj. acquisition time $\approx 11\,\mathrm{h}$ (b) $N_B = 500$ samples/proj. acquisition time $\approx 2.75\,\mathrm{h}$

Figure 10: **Optimization of the number of samples per projection.** We considered two datasets, both acquired using $\mu = 168$ G/cm and $B_{\text{sw}} = 719.44$ G. The first dataset was acquired using $N_B = 2000$ samples per projection (leading to $\delta_r = \frac{B_{\text{sw}}}{\mu N_B} \approx 21.4$ µm) while $N_B = 500$ samples per projection ($\delta_r \approx 85.6$ µm) were used for the second one. We used Algorithm 5 to reconstruct an image with size $M = 500$ (and spatial sampling step $\delta \approx 85.6$ µm) from each dataset using $\lambda' = 10$ and $\alpha' = 0$. Resulting images are displayed in (a) (first dataset) and (b) (second dataset). From Section 6.2, we expected two reconstructions with indiscernible quality. In practice, we observe that (b) seems slightly less detailed than (a) and exhibits a larger amounts of residual noise than (a). Possible explanations for this phenomenon are proposed in Figure 11. Nevertheless, the quality degradation observed in (b) compared to (a) remains moderate and the reduction (by a factor 4) of the acquisition time for the second dataset can be greatly appreciated in practice.

7.2.5 Influence of other Acquisition Parameters

In this section, we study how several acquisition parameters, such as the number of projections N_θ, the magnetic field gradient intensity μ or the integration time constant τ_c may impact the quality of the reconstruction.

In Figure 12, we compare reconstructions obtained from sinograms with various numbers of projections N_θ. As expected, the diminution of the number of projections deteriorates the quality of the reconstruction. Interestingly enough, we can see that the characteristic size of the thin structures becomes significantly larger than 85.6 µm as N_θ decreases. This illustrates that the estimate value $\delta = 85.6$ µm obtained from the reference spectrum (common to all reconstructions displayed in Figure 12) is a lower bound on the actual details size that we can hope to observe in the reconstructed image. This lower bound may not be attained, in particular, when N_θ is not large enough.

In Figure 13, we display reconstructions obtained from acquisitions with different integration time constant parameter τ_c. This parameter impacts the acquisition time as well as the level of noise in the acquired signal. Decreasing τ_c yields reduced acquisition time but also smaller SNR and, consequently, degraded image reconstructions (with degraded level of details). However, the τ_c parameter does not only affect the SNR of the sinogram acquisition but also that of the reference spectrum acquisition. Consequently, it impacts the value of δ computed from the reference spectrum using the methodology described in Section 6.3. We observe that δ increases as τ_c decreases and provides a realistic estimation of the smallest detail size that we can observe in the reconstructed image.

In Figure 14, we study the influence of the field gradient intensity μ on the reconstruction. This

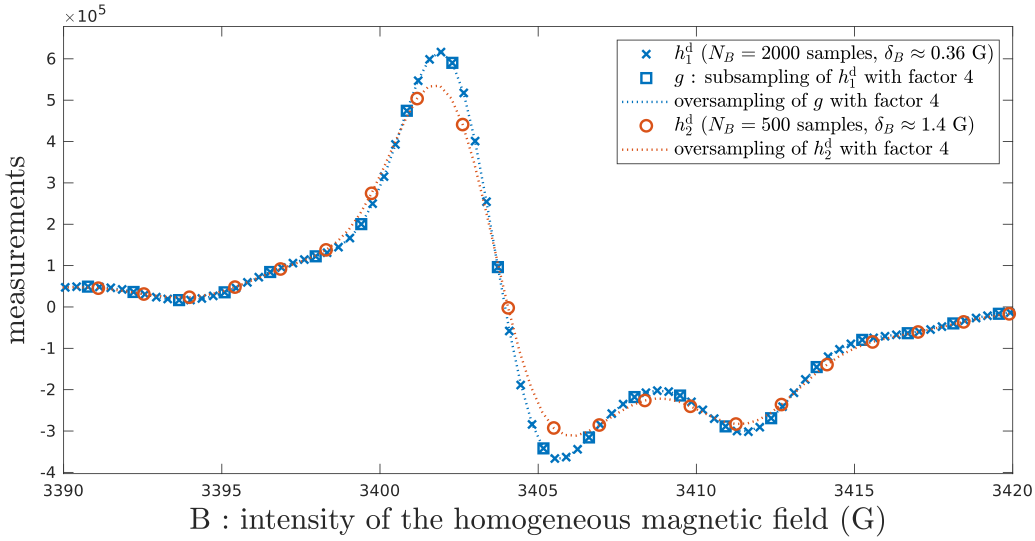
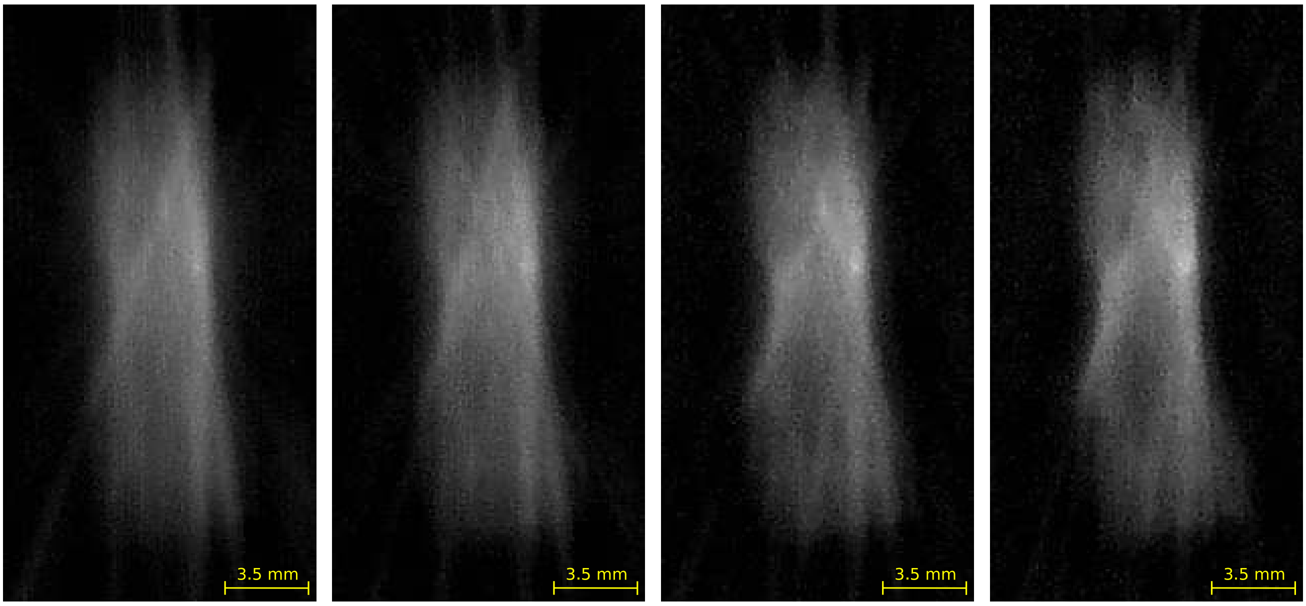


Figure 11: **Reference spectrum acquisitions with different sampling steps.** We display the reference spectra h_1^{d} and h_2^{d} associated to the two datasets considered in Figure 10. The spectrum h_1^{d} corresponds to the same acquisition as that displayed in Figure 4 (a) and was acquired using a sampling step $\delta_B \approx 0.36\,\mathrm{G}$ which is, according to Figure 4 (b), roughly four times below the estimated critical sampling step $\nu \approx 1.4\,\mathrm{G}$. Subsampling h_1^{d} with a factor four and resampling (using *zero-padding* Shannon interpolation) this subsampled signal yields the blue dotted line signal that almost perfectly fits with h_1^{d} samples. This confirms that an acquisition with $\delta_B = 1.4\,\mathrm{G}$ contains as much information as h_1^{d}. However, running an acquisition using $\delta_B = 1.4$ G (and keeping all other acquisition parameters unchanged) yielded the signal h_2^{d} which significantly misfits with h_1^{d}. The most likely explanation for this phenomenon is that a signal attenuation due to filtering occurred during the acquisition of h_2^{d} (yielding orange circle samples with incorrect positions along the Y-axis of the graph). Another explanation could be that the sampling rate is too fast for the instrument, yielding inaccurate sampling positions (i.e. orange circle samples with incorrect positions along the X-axis of the graph). As a matter of fact, the setting $N_B = 500$ yields an acquisition with reduced SNR compared to the setting $N_B = 2000$ which explains the image degradation observed in Figure 10.

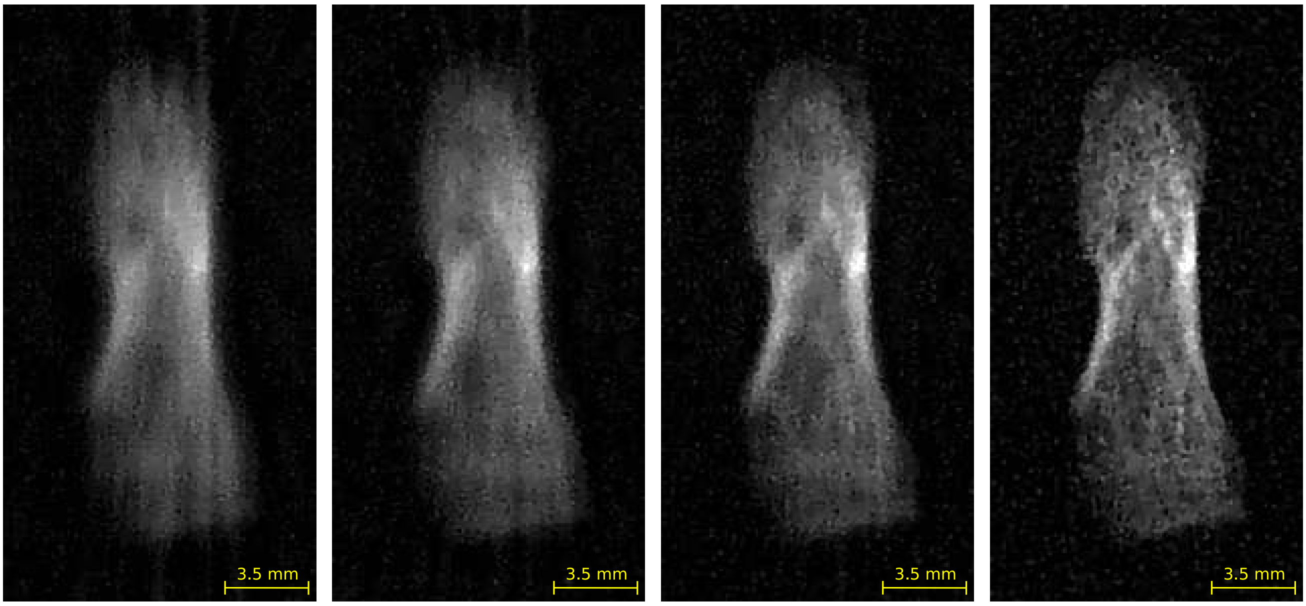
parameter strongly impacts the resolution since it is linked to the radial and spatial sampling steps through the relations $\delta_r = \frac{\delta_B}{\mu}$ and $\delta = \frac{\delta_B}{\mu} \cdot \frac{N_B}{M}$. The field gradient parameter μ is not involved in the acquisition of the reference spectrum (which is always done without field gradient, or equivalently, using $\mu = 0$). However, in Figure 14, by imposing the same FOV for all sinogram acquisitions, and since the FOV diameter satisfies $D_{\mathrm{FOV}} = \frac{B_{\mathrm{sw}}}{\mu}$, the sweep-width parameter B_{sw} and the homogeneous magnetic field intensity sampling step $\delta_B = \frac{B_{\mathrm{sw}}}{N_B}$ were set differently for each dataset acquisition, leading to different estimated values for δ. As expected, we see in Figure 14 that the image resolution increases with μ. However, we can also remark that increasing μ yields reconstructions with larger amount of residual noise (especially visible in the image background). Increasing τ_c jointly with μ could be a way to keep the noise level constant. However, setting arbitrary large field gradient intensity μ is not possible due to instrumental limitations (especially heating issues). Besides, for too large values of μ, the physical modeling described in Section 2 is not valid anymore. Therefore, increasing further the image resolution would involve important instrumental and modeling improvements.

7.2.6 Comparing EPR and CT-Scan Reconstructions

To conclude our experiments on the phalanx sample, we investigate in Figure 15 how far the structures that we reconstructed using Algorithm 5 are reliable. For that purpose, we performed a Computed Tomography Scan (CT-Scan) of the phalanx. A CT-Scan acquisition consists in rotating a X-rays beam around the sample and measuring, for different beam orientations, the X-rays attenuation. The latter is proportional to the density of matter from which the X-rays pass through, and can be used to reconstruct a 3D density image of the sample [20]. Since EPR is only sensitive to paramagnetic



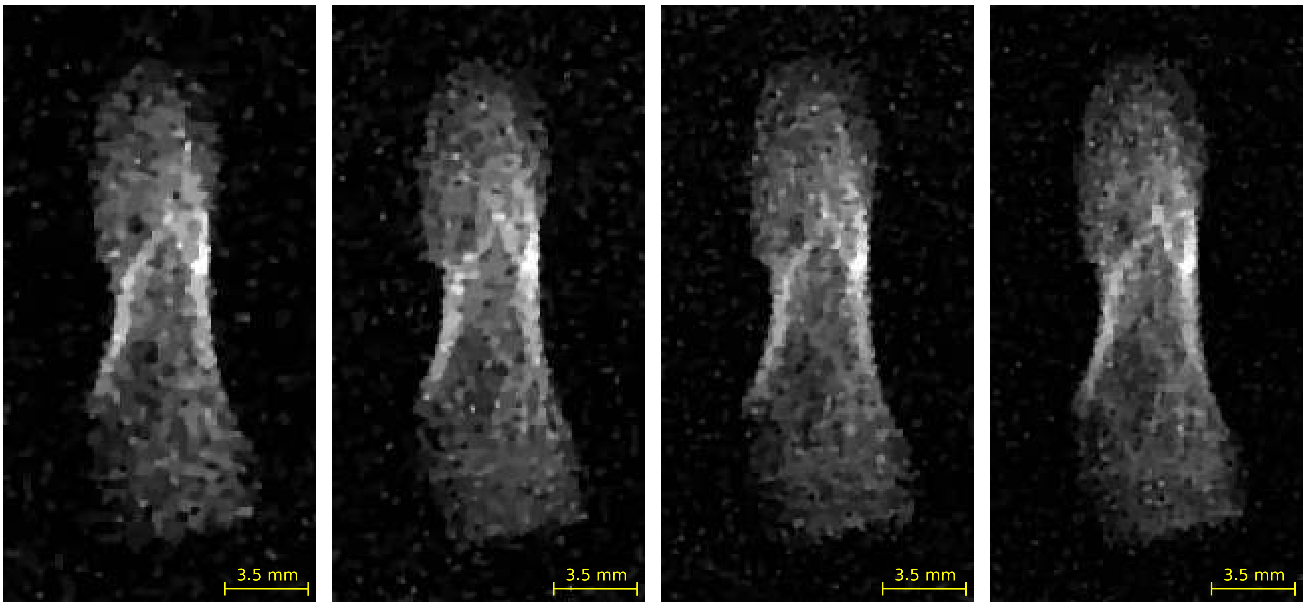
(a) 10 projections (b) 13 projections (c) 19 projections (d) 23 projections



(e) 29 projections (f) 38 projections (g) 57 projections (h) 113 projections

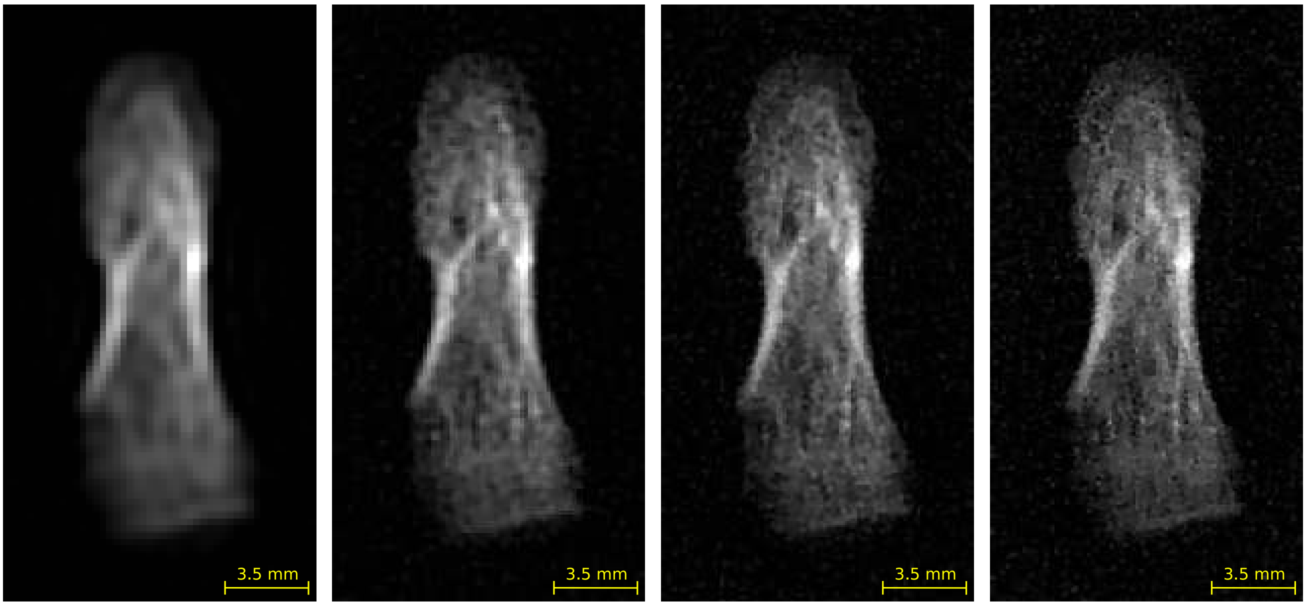
Figure 12: **Influence of the number of projections.** In this experiment, we used the sinogram $s_{2\mathrm{D}}$ made of $N_\theta = 113$ projections displayed in Figure 6 (b). By removing one out of k columns (for $k \in \{1,2,3,4,5,6,9,12\}$) from $s_{2\mathrm{D}}$, we generate sinogram acquisitions with lower number of projections (and associated angular sampling step $\delta_\theta \approx 1.6\,k$ degrees). We processed each 2D sinogram using Algorithm 5 taking as reference spectrum that displayed in Figure 4 (a), and setting $M = 500$ (i.e., $\delta \approx 85.6$ µm) and $\lambda' = 10$. We can see that the image quality is severely impacted by the diminution of the number of projections. In (a), the phalanx shape is barely reconstructed, then, the level of details progressively increases with N_θ. Since all datasets share the same reference spectrum from which we estimated $\delta \approx 85.6$ µm as optimal spatial sampling step, all reconstructions were done using this recommended setting for δ. In practice, we see that details with typical size at best equal to 0.2 mm can be recovered when $N_\theta \leq 38$ so that the estimated value 85.6 µm only corresponds to a lower bound on the actual minimal size of the details that we can hope to reconstruct. The actual optimal value for δ depends on many parameters (N_θ being obviously one of those) that are not involved in the spectrum acquisition alone.

species, images obtained from EPR and CT-Scan are usually different and not easily comparable. However, in the particular case of irradiated bones, paramagnetic damages caused by irradiations are more likely to occur in high density areas, making the images reconstructed with both EPR and CT-Scan modalities surprisingly similar.



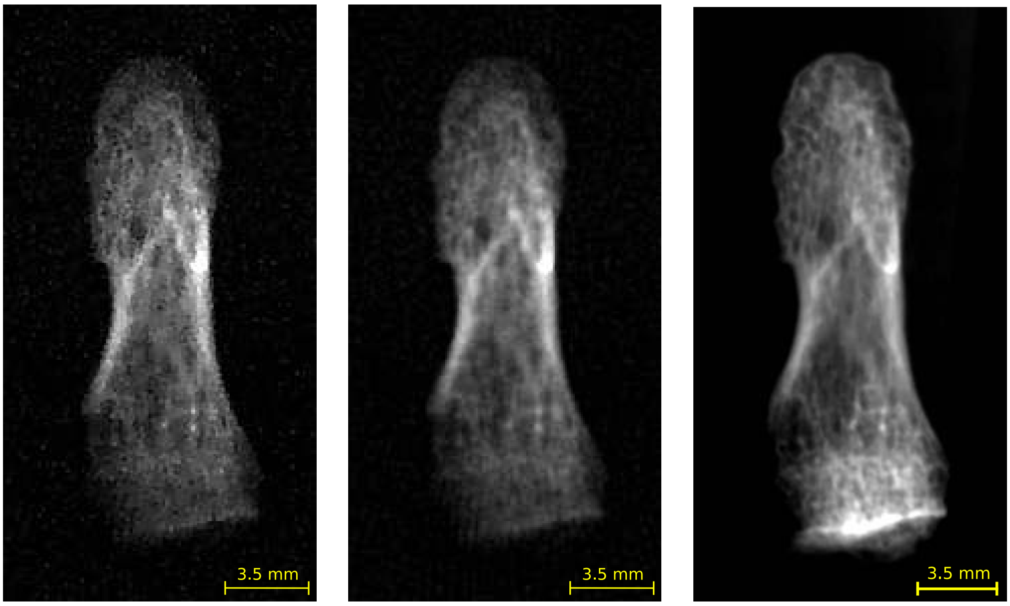
(a) $\tau_c = 10.24\,\text{ms}$ ($\lambda' = 500$) $M = 344$, $\delta \approx 125\,\mu\text{m}$
(b) $\tau_c = 20.48\,\text{ms}$ ($\lambda' = 250$) $M = 374$, $\delta \approx 115\,\mu\text{m}$
(c) $\tau_c = 40.96\,\text{ms}$ ($\lambda' = 100$) $M = 388$, $\delta \approx 110\,\mu\text{m}$
(d) $\tau_c = 81.92\,\text{ms}$ ($\lambda' = 50$) $M = 398$, $\delta \approx 108\,\mu\text{m}$

Figure 13: **Influence of the time constant parameter.** The time constant parameter τ_c of the Bruker spectrometer roughly corresponds to integration time between two sample acquisitions. Increasing τ_c improves the SNR of the acquired signal but also increases the acquisition time. We performed dataset (one reference spectrum and one sinogram) acquisitions using $N_B = 500$ samples per projection, $\mu = 168$ G/cm, $\delta_r \approx 85.6$ µm and different τ_c values. Then, we reconstructed an image for each dataset computing δ as detailed in Section 6.3 and tuning λ' by visual inspection. Unsurprisingly, the value of τ_c impacts the quality of the reconstruction as well as the thinness of the details that we can observe in the image.



(a) $\mu = 42$ G/cm $M = 152$, $\delta \approx 282\,\mu\text{m}$
(b) $\mu = 84$ G/cm $M = 230$, $\delta \approx 186\,\mu\text{m}$
(c) $\mu = 126$ G/cm $M = 326$, $\delta \approx 132\,\mu\text{m}$
(d) $\mu = 168$ G/cm $M = 410$, $\delta \approx 105\,\mu\text{m}$

Figure 14: **Reconstructions from acquisitions with different field gradient intensities.** We performed dataset acquisitions using different values of μ, $N_B = 500$ samples per projection and $\delta_r = \frac{B_{\text{sw}}}{\mu N_B} \approx 85.6$ µm (by changing B_{sw} according to the value of μ). Image reconstruction was done for each dataset computing δ as detailed in Section 6.3 and setting $\lambda' = 10$. We see that, as μ increases, the estimated value of δ decreases and the level of details increases in the reconstructed image. However, increasing μ (and keeping the integration time τ_c unchanged) also yields more noise during the acquisition.



(a) TV ($\lambda' = 10$) (b) Huber-TV ($\lambda' = 100, \alpha' = 0.05$) (c) CT-Scan

Figure 15: **EPR and CT-Scan image reconstructions of the phalanx.** We display in (a) and (b) two EPR reconstructions of the phalanx obtained using Algorithm 5. We also performed a micro-CT imaging of the phalanx sample using a Skyscan 1178 tomograph. The raw micro-CT images were processed using the Skyscan software suite (Nrecon, CTAn) to compute a 3D image with spatial sampling step $\delta_{\text{CT}} = 59\,\mu\text{m}$. Rotating appropriately this 3D image and summing along the third axis yields the 2D image (c) that can be compared to the 2D EPR images (see the link between 3D and 2D in Section 2.3) displayed in (a) and (b). We can see that most of the structures present in (a) and (b) are also visible in (c). The absence of staircasing in (b) is also highly appreciable due to the large amount of texture in the phalanx sample, making the aspect of image (b) surprisingly close to that of the micro-CT image (c).

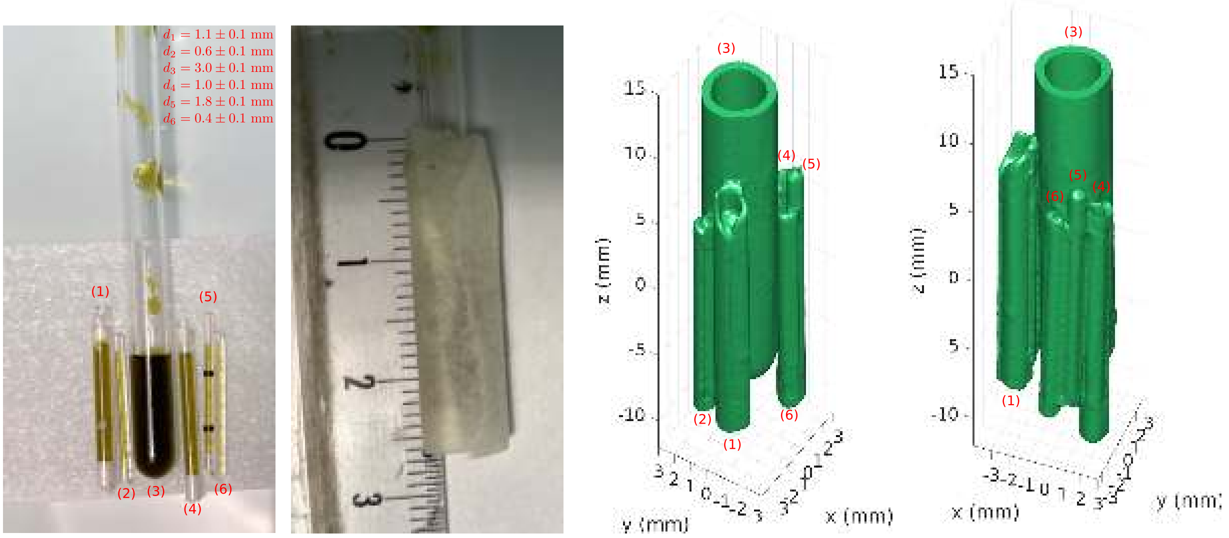
7.3 Reconstruction of 3D Images

In this section, we focus on the reconstruction of 3D EPR images using Algorithm 6. Notice that all the phalanx 2D datasets studied in Section 7.2 were acquired with the X-band 2D EPR acquisition system presented in Figure 5 (a). For 3D EPR imaging purpose, the L-band EPR imager presented in Figure 5 (b) can be used. As mentioned earlier, this L-band EPR imager is able to provide field gradient intensities up to 45 G/cm. With such setting, we cannot hope achieving as detailed reconstructions as those obtained at X-band over the phalanx sample, making the 3D study of the phalanx not so relevant anymore. In this section, we will illustrate the ability of Algorithm 6 to perform 3D EPR image reconstruction using two different samples, one made of capillary tubes filled with TAM and the second made of a fusillo pasta soaked with 4OH-TEMPO.

7.3.1 Capillary Tubes Filled with TAM

We filled some capillary tubes with an aqueous TAM solution with concentration 1 millimolaire ($1\,\text{mM} = 10^{-3}\,\text{mol.L}^{-1}$). The tubes were sealed using wax sealed plate (Fischer scientific) and attached together using masking tape. Some pictures and 3D visualizations (obtained from CT-Scan) of the sample are displayed in Figure 16 and Figure 17.

We performed the acquisition of a dataset made of one reference spectrum and a 3D sinogram. This dataset was acquired in L-band (with a microwave around 1.10 GHz) using $N_B = 360$ samples per projection, $B_{\text{SW}} = 45.6\,\text{G}$ and $B_{\text{CF}} = 401.65\,\text{G}$, leading to $\delta_B = \frac{B_{\text{SW}}}{N_B} \approx 0.127\,\text{G}$ for the sampling step of the homogeneous magnetic field. We used $\mu = 20\,\text{G/cm}$ as field gradient intensity during the sinogram acquisition, leading to $\delta_r = \frac{\delta_B}{\mu} \approx 63\,\mu\text{m}$ for the radial sampling step. The polar and



(a) capillaries filled with TAM (b) 3D visualization (from CT-Scan)

Figure 16: **Capillary tubes filled with TAM.** We filled six tubes with a solution of TAM with concentration 1 mM. The tubes, numbered from one to six, and their internal diameters $(d_i)_{1\leq i\leq 6}$, are visible in (a) (left). The tubes were wrapped together using masking tape, as displayed in (a) (right). We performed a CT-Scan acquisition of the sample, leading to the reconstruction of a 3D image of the wrapped tubes. An isosurface of the 3D image is displayed in (b) under two different orientations, showing the tubes organization inside the masking tape (the latter being transparent to X-rays). Some slices of the 3D image are also displayed in Figure 17, where we can visualize the TAM solution contained inside the tubes.

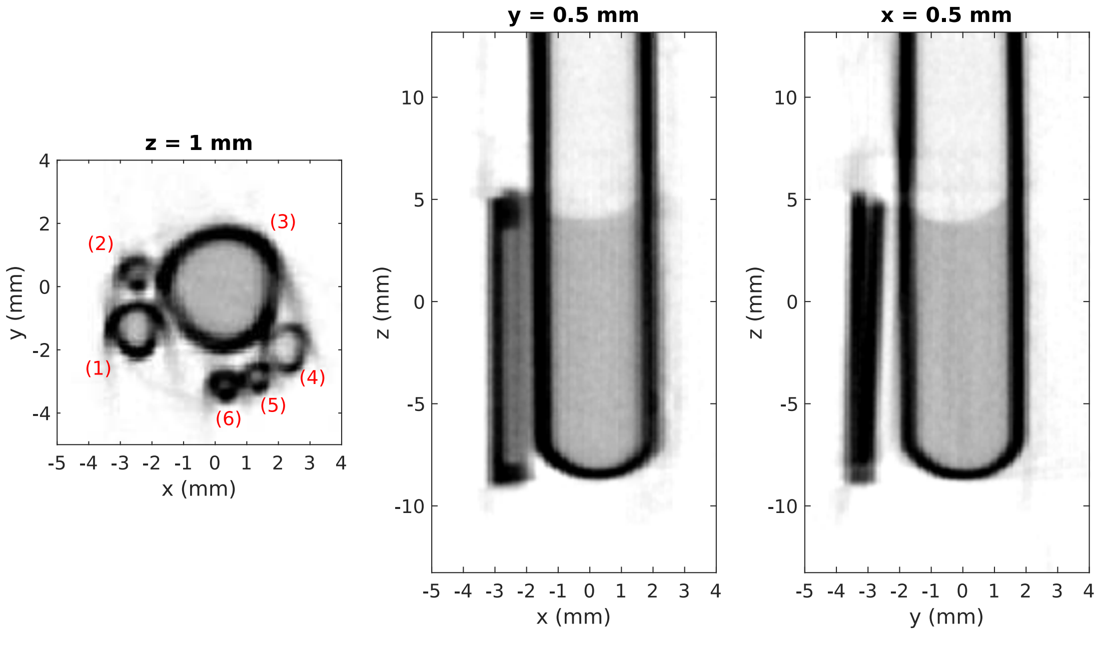
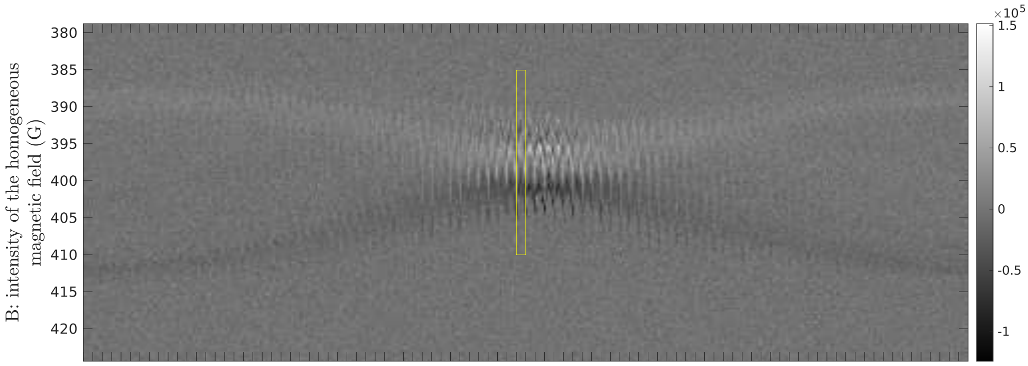
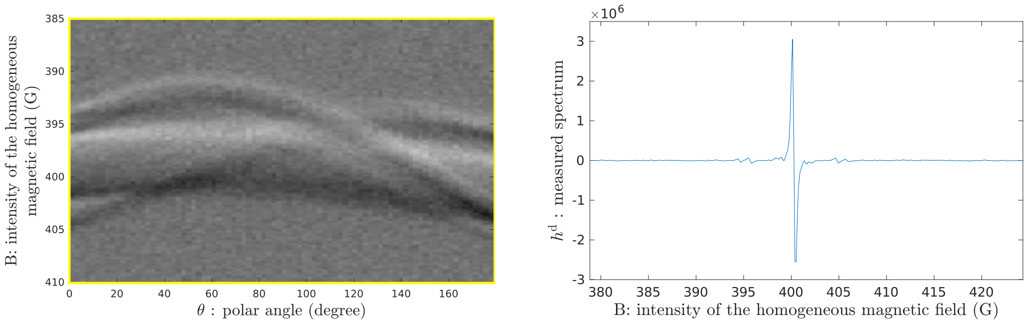


Figure 17: **Several slices extracted from the CT-Scan 3D image.** We display three slices of the 3D image reconstructed from the CT-Scan acquisition (see also Figure 16 (b)). In those slice images, large pixel values are represented in dark and low pixel values are represented in white. Darkest areas corresponds to the walls of the tubes (that are the areas with highest matter density in the sample). The gray area inside the tubes corresponds to the TAM solution. Note that we can see the meniscus of the TAM solution inside tube (3) in the XZ (middle) and YZ (right) slices.

azimuthal angular position were discretized into 94 samples regularly spaced in $[0°, 180°]$, leading to $N_\theta = N_\varphi = 94$ and $\delta_\theta = \delta_\varphi = 1.914°$, and a total number of $N_\theta \times N_\varphi = 8836$ projections. The time constant parameter τ_c was set to 20.48 ms and the same value as τ_c was used for the conversion time parameter. This lead to an acquisition time of $N_B\,\tau_c \approx 7$ seconds per projection and the acquisition of the whole 3D sinogram lasted in around $N_\theta N_\varphi N_B\,\tau_c \approx 18$ hours. The acquisition was done using an amplitude modulation of $0.5\,\mathrm{G}$, a frequency modulation of $100\,\mathrm{KHz}$, a microwave power of $1.005\,\mathrm{mW}$ with $26\,\mathrm{dB}$ of power attenuation. The reference spectrum and the 3D sinogram acquired as detailed above are displayed in Figure 18. At the end of the dataset acquisition, we noticed that tubes (1) and (6) where badly sealed. Tube (1) was partially empty (around one third of the solution remained inside the tube thanks to the presence of an air bubble visible in Figure 16 (a)), and tube (6) was totally empty. We used Algorithm 6 to reconstruct a 3D image from the dataset displayed in Figure 18. An isosurface of the 3D reconstruction is displayed in Figure 19, where we can see the content of the tubes, in particular, the meniscus of the TAM solution in tube (3), the remaining amount of solution in tube (1), while no paramagnetic signal is reconstructed in tube (6).



(a) concatenation of 2D sinograms



(b) close-up view of (a)

(c) reference spectrum

Figure 18: **Reference spectrum and 3D sinogram of the capillary tubes filled with TAM.** The 3D sinogram acquisition consisted in the acquisition of $N_\varphi = 94$ 2D sinograms, each containing $N_\theta = 94$ projections. We display in (a) the concatenation of all $N_\theta N_\varphi = 8836$ projections, and a close-up view of a single 2D sinogram is displayed in (b). The reference spectrum acquired together with the 3D sinogram is displayed in (c). The whole acquisition process lasted in around 18 hours.

Our reconstruction was done using MATLAB® R2019b and a standard laptop operated under Ubuntu 19.10 and equipped with an Intel® Core™ i7-7920HQ CPU ($3.10\,\mathrm{GHz} \times 4$ cores) and $31.2\,\mathrm{GB}$ of RAM. The 3D image displayed in Figure 19 was reconstructed with a pixel sampling size $\delta = 100\,\mu\mathrm{m}$ ($M = 228$). Algorithm 6 was used to perform $N = 1000$ iterations of the scheme (42) and the latter was initialized using the zero-valued 3D image $0_{\Omega_{3\mathrm{D}}}$. The reconstruction at such resolution lasted in almost one hour and involved the allocation of up to $25\,\mathrm{GB}$ of RAM. Note that most of the RAM usage was due to the computation of $\Upsilon_{3\mathrm{D}}$ (which has size $2M \times 2M \times 2M$ and involves 3D NFFT computations). Once this kernel was computed, the amounts of RAM usage dropped to around $7\,\mathrm{GB}$). Notice that a reconstruction with coarser scale, $\delta = 285\,\mu\mathrm{m}$ ($M = 80$), yields a less detailed but still acceptable reconstruction (with well separated tube contents) and involves less computational resources (up to $4\,\mathrm{GB}$ of RAM for the computation of $\Upsilon_{3\mathrm{D}}$ and less than two minutes for a reconstruction using $N = 1000$ scheme iterations).

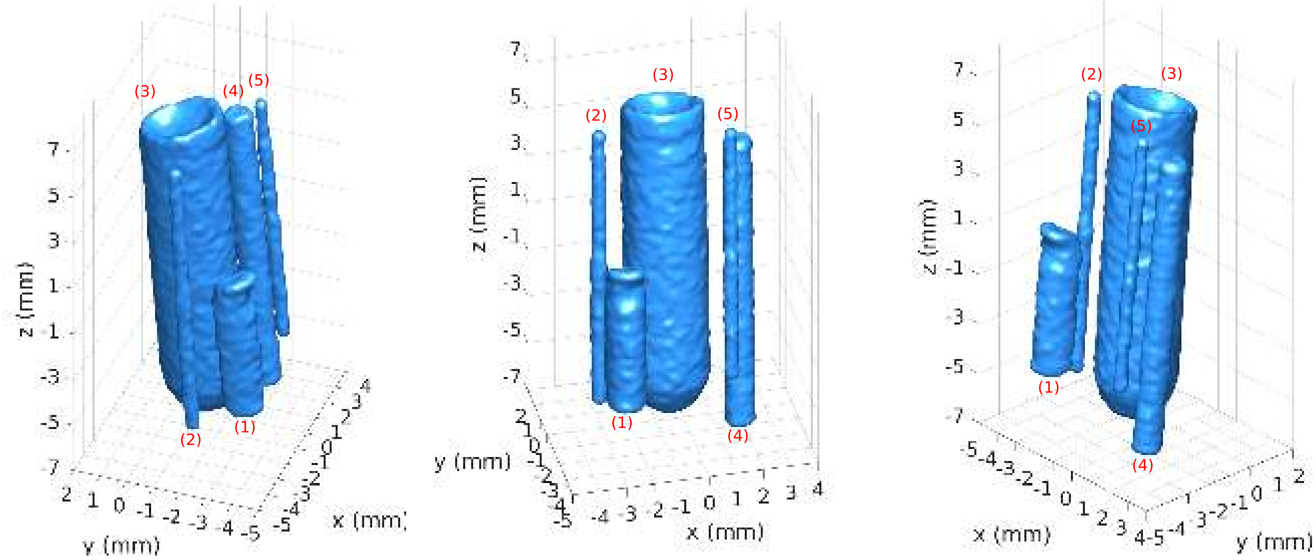
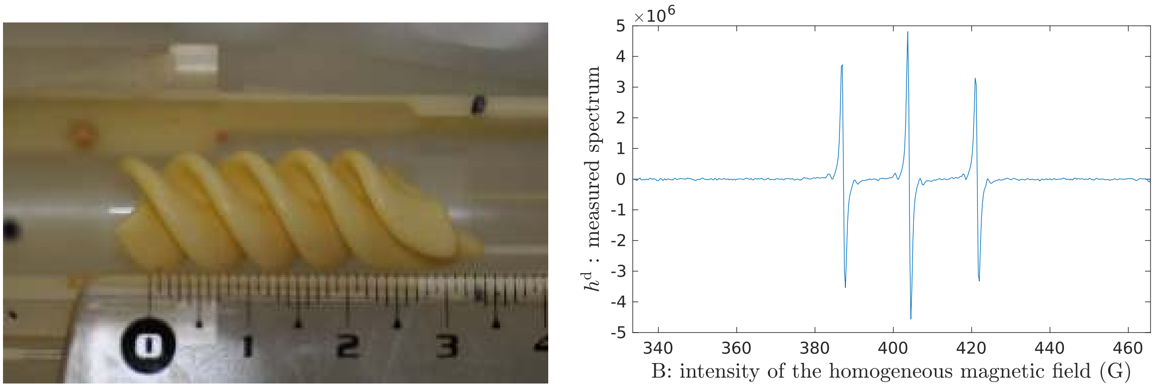


Figure 19: **Reconstruction of 3D capillary tubes filled with TAM.** Algorithm 6 was used with $\lambda' = 2$ (the value of the unnormalized parameter λ was computed using (18)) and $M = 228$ (leading to the spatial sampling step $\delta = 100\,\mu\mathrm{m}$) to reconstruct a 3D image. An isosurface is displayed here under three different orientations. As expected, only the TAM solution present inside the tubes is reconstructed. We can observe that the solution meniscus inside tube (3) is nicely reconstructed. We can also observe that tube (1) is partially empty and tube (6) is totally empty. Indeed liquid leaks occurred during manipulations prior to acquisition (see details above).

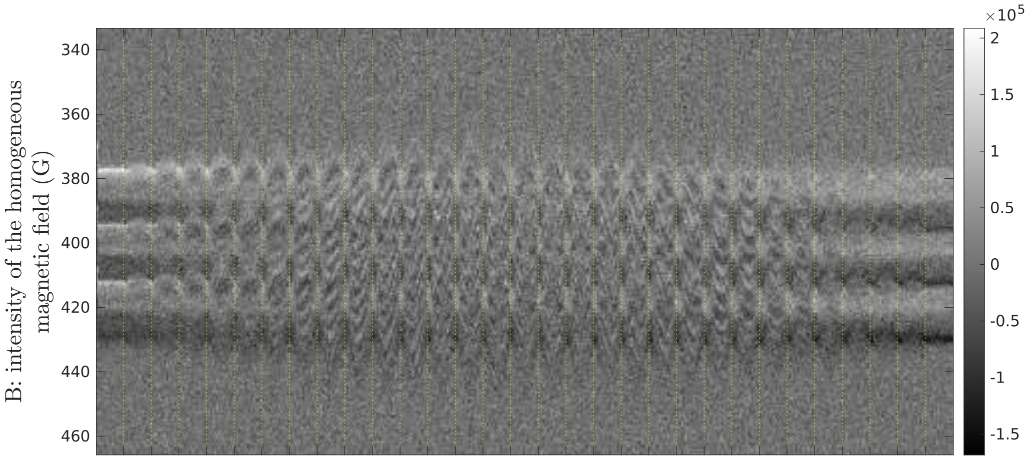
7.3.2 Fusillo Soaked with TEMPO

A fusillo (Barilla™) with helical shape was put into a 4 mM aqueous solution of 4OH-TEMPO for a night, then it was dried using paper and put inside the L-band EPR resonator. A 3D dataset was acquired, setting $B_{\mathrm{SW}} = 132.5\,\mathrm{G}$, $B_{\mathrm{CF}} = 399.7\,\mathrm{G}$ and $N_B = 500$, leading to $\delta_B = \frac{B_{\mathrm{SW}}}{N_B} \approx 0.265\,\mathrm{G}$. The sinogram acquisition was done using $\mu = 14\,\mathrm{G/cm}$, leading to the radial sampling step $\delta_r = \frac{\delta_B}{\mu} \approx 189\,\mu\mathrm{m}$. The polar and azimuthal angles were discretized into $N_\theta = N_\varphi = 31$ values regularly spaced in $[0°, 180°]$, leading to $\delta_\theta = \delta_\varphi = 5.806°$ and a total of $N_\theta \times N_\varphi = 961$ acquired projections. The integration time constant parameter τ_c was set to 2.56 ms, leading to an acquisition time of $N_\theta N_\varphi N_B \tau_c \approx 20$ minutes. The amplitude modulation was set to 0.6 G, the frequency modulation to 100 KHz, the microwave power to 20.05 mW with a power attenuation of 13 dB. A picture of the Fusillo sample as well as the reference spectrum and the 3D sinogram composing the EPR dataset are displayed in Figure 20.

In our numerical simulations, we observed that the initialization of the numerical scheme (42) by the zero-valued image $0_{\Omega_{3\mathrm{D}}}$ yielded very slow numerical convergence. This is probably due to the three-lines shape of the 4OH-TEMPO reference spectrum which makes the underlying deconvolution problem numerically difficult to address. In order to achieve a satisfactory 3D image reconstruction within a reasonable computational time, we adopted a multiscale reconstruction strategy to process this dataset. Algorithm 6 was run to achieve a very coarse scaled reconstruction, using $M = 12$ ($\delta \approx 7.9\,\mathrm{mm}$) and the zero-valued image as initial guess. The resulting image was upscaled (using cubic spline interpolation) to a finer scale ($M = 24$, $\delta \approx 3.9\,\mathrm{mm}$) and used as initializer to perform another reconstruction at this new scale. This process was repeated at scales $M \in \{12, 24, 50, 74, 100, 124, 150, 174, 200\}$, using 1000 to 5000 iterations depending on the scale, to achieve the final reconstruction displayed in Figure 21.



(a) picture of the *fusillo* (b) acquired reference spectrum



(c) concatenation of 2D sinograms

Figure 20: **Fusillo soaked with 4OH-TEMPO.** We display in (a) a picture of the fusillo sample. The latter was soaked with a solution of 4OH-TEMPO with concentration $4\,\mathrm{mM}$. The EPR dataset acquired over this sample is made of the reference spectrum displayed in (b) and the 3D sinogram displayed in (c) as a concatenation of 2D sinograms separated by the yellow dotted lines. The whole acquisition process lasted in around 20 minutes, which is rather fast for a 3D dataset. However, one can see that the acquired projections are corrupted with a large amount of noise.

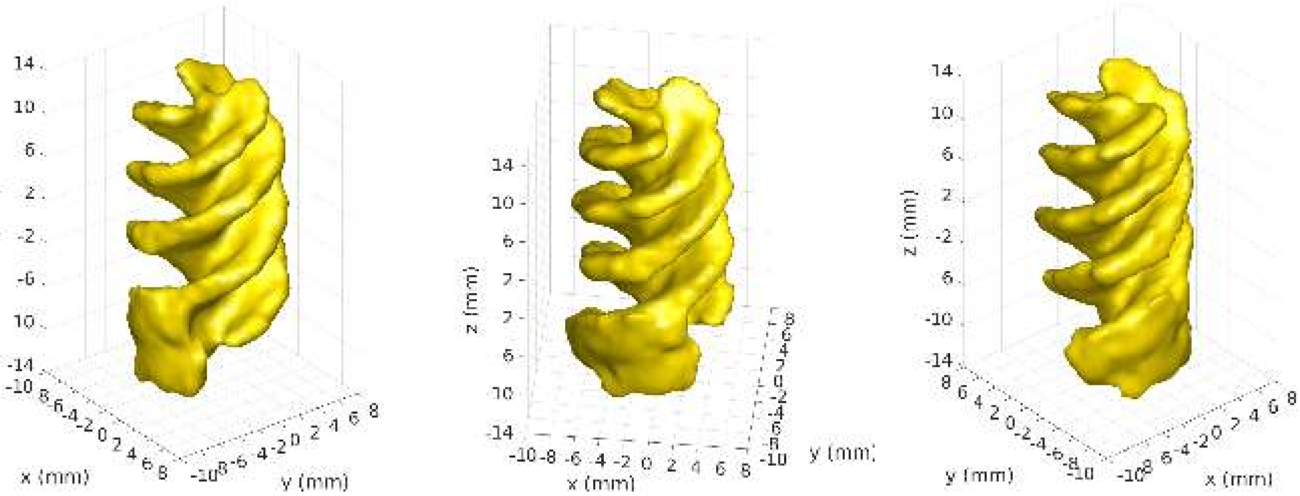


Figure 21: **Reconstruction of 3D fusillo soaked with 4OH-TEMPO.** We display an isosurface (from three different viewpoints) of the 3D image obtained using Algorihtm 6 over the fusillo dataset displayed in Figure 20. The reconstruction was done using $\lambda' = 250$ and $M = 200$ (leading to $\delta = 473.2\,\mu\mathrm{m}$). Despite the noise corrupting the sinogram, we achieve the reconstruction of regular 3D image, where we can recognize the helical shape of the fusillo.

8 Conclusion and Perspectives

In conclusion, the TV regularized least squares method presented in this article enables to properly reconstruct 2D and 3D EPR images. However, some avenues for improving its efficiency, as well as possible extensions, have not yet been exploited and will be the subjects of future works.

- First, the mathematical model can be refined. In particular, the filtering profile g_δ of the acquisition system could be calibrated in order to fine-tune the a contrario estimate of the optimal sampling step.
- The main drawback of the TV prior is the well-known staircase effect. Huber-TV prior allows to remove it, but at the expense of a blurrier image. More sophisticated priors could be considered. However, an efficient way to remove this artifact is to estimate the posterior expectation of the TV model instead of maximizing the posterior probability. When introduced in [26], this method was implemented in a computationally expensive algorithm, but later on, a very fast algorithm was proposed for image denoising only [28]. An extension of this algorithm to EPR imaging would be very useful.
- Another well-known drawback of TV-regularized least squares is a bias in grey level values of restored images. The latter becomes a concern when trying to accurately quantify the concentration of paramagnetic species. A series of methods have been proposed to correct this bias in line with [32] and could be applied to EPR imaging.
- At last, a slight change in the model allows to deal with cases where several paramagnetic species are present in the cavity. Recently conducted experiments have led to encouraging results and are the purpose of a forthcoming article. (See [22] for preliminary results.)

Image Credits

All images in this manuscript were produced by the authors (license CC-BY-SA).

A Proof of Theorem 1

Let $V \in L^1(\mathbb{R}^2)$, $\theta \in \mathbb{R}$ and $\xi \in \mathbb{R}$.

$$\forall r \in \mathbb{R}\,, \quad \mathcal{R}_\theta(V)(r) = \int_{\mathbb{R}} V(re_\theta + \rho e_\theta^\perp)\, d\rho = \int_{\mathbb{R}} V(r\cos\theta - \rho\sin\theta, r\sin\theta + \rho\cos\theta)\, d\rho\,.$$

Thus, taking the Fourier transform of $\mathcal{R}_\theta(V)$ at the point ξ yields

$$\mathcal{F}(\mathcal{R}_\theta(V))(\xi) = \int_{\mathbb{R}} \mathcal{R}_\theta(V)(r)\, e^{ir\xi}\, dr = \int_{\mathbb{R}}\int_{\mathbb{R}} V(r\cos\theta - \rho\sin\theta, r\sin\theta + \rho\cos\theta)\, e^{-ir\xi}\, d\rho\, dr\,. \tag{59}$$

Now, let us compute the Fourier transform of V at point $(\xi\cos\theta, \xi\sin\theta)$. We have

$$\mathcal{F}(V)(\xi\cos\theta, \xi\sin\theta) = \int_{\mathbb{R}}\int_{\mathbb{R}} V(x,y)\, e^{-i(x\cos\theta + y\sin\theta)\xi}\, dx\, dy\,. \tag{60}$$

Using the variable change

$$\left\{\begin{array}{ll} r & = x\cos\theta + y\sin\theta \\ \rho & = -x\sin\theta + y\cos\theta \end{array}\right. \quad \Leftrightarrow \quad \left\{\begin{array}{ll} x & = r\cos\theta - \rho\sin\theta \\ y & = r\sin\theta + \rho\cos\theta \end{array}\right.$$

with unitary Jacobian determinant, (60) becomes

$$\mathcal{F}(V)(\xi\cos\theta,\xi\sin\theta) = \int_{\mathbb{R}}\int_{\mathbb{R}} V(r\cos\theta - \rho\sin\theta, r\sin\theta + \rho\cos\theta)\, e^{-ir\xi}\, d\rho\, dr\,,$$

which is equal to (59).

B Details about Approximation (26)

Let $\theta \in [0,\pi]$, let $\widetilde{s}_\theta : r \mapsto \left(\widetilde{h}_X^\mu * \mathcal{R}_\theta(\widetilde{V}_X)\right)(-B_{\text{CF}}/\mu + r)$ and let $\widetilde{h} = r \mapsto \widetilde{h}_X^\mu(-B_{\text{CF}}/\mu + r)$. From (12), we have

$$\forall r \in \mathbb{R}\,, \quad \widetilde{s}_\theta(r) = \left(\widetilde{h} * \mathcal{R}_\theta(\widetilde{V}_X)\right)(r) \quad \text{and} \quad \forall \xi \in \mathbb{R}\,, \quad \mathcal{F}(\widetilde{s}_\theta)(\xi) = \mathcal{F}(\widetilde{h})(\xi)\cdot\mathcal{F}(\mathcal{R}_\theta(\widetilde{V}_X))(\xi)\,. \tag{61}$$

Since $\widetilde{h}_X^\mu = g_\delta * h_X^\mu$ and $\delta_r \le \delta$, the Fourier transform of $\widetilde{h}_X^\mu$ is supported in $[-\frac{\pi}{\delta_r}, \frac{\pi}{\delta_r}]$, and so as the Fourier transform of $\widetilde{h}$ (since for all $\xi \in \mathbb{R}$, we have $\mathcal{F}(\widetilde{h})(\xi) = \mathcal{F}(\widetilde{h}_X^\mu)(\xi)\cdot e^{-i\xi B_{\text{CF}}/\mu}$) and that of $\widetilde{s}_\theta$ (thanks to (61)). Therefore, Theorem 2 yields

$$\forall r \in \mathbb{R}\,, \quad \widetilde{s}_\theta(r) = \sum_{m\in\mathbb{Z}} \widetilde{s}_\theta(m\delta_r)\,\mathrm{sinc}\left(\frac{r}{\delta_r} - m\right) \quad \text{and} \quad \widetilde{h}(r) = \sum_{m\in\mathbb{Z}} \widetilde{h}(m\delta_r)\,\mathrm{sinc}\left(\frac{r}{\delta_r} - m\right). \tag{62}$$

Since $\mathcal{F}\left(r \mapsto \mathrm{sinc}(\frac{r}{\delta_r} - m)\right)(\xi) = \delta_r \cdot e^{-i\xi m\delta_r}\cdot \mathbb{1}_{[-\frac{\pi}{\delta_r},\frac{\pi}{\delta_r}]}(\xi)$, taking the Fourier transform of $\widetilde{s}_\theta$ in (62) yields

$$\forall \xi \in \mathbb{R}\,, \quad \mathcal{F}(\widetilde{s}_\theta)(\xi) = \begin{cases} \delta_r \sum_{m\in\mathbb{Z}} \widetilde{s}_\theta(m\delta_r)\, e^{-i\xi m\delta_r} & \text{if } |\xi| \le \pi/\delta_r \\ 0 & \text{otherwise.} \end{cases} \tag{63}$$

Denoting $\widetilde{s}_\theta^{\mathrm{d}} : m \in I_{N_B} \mapsto \widetilde{s}_\theta(m\delta_r)$, by taking $\xi \in 2\pi I_{N_B}/(N_B\delta_r)$ and neglecting the terms $\widetilde{s}_\theta(m\delta_r)$ for $m \notin I_{N_B}$ in (63), we obtain the approximation

$$\forall \alpha \in I_{N_B}\,, \quad \mathcal{F}(\widetilde{s}_\theta)\left(\frac{2\pi\alpha}{N_B\delta_r}\right) \approx \delta_r \cdot \mathrm{DFT}(\widetilde{s}_\theta^{\mathrm{d}})(\alpha)\,. \tag{64}$$

Proceeding similarly for $\widetilde{h}$, and remarking that (27) yields $\widetilde{h}(m\delta_r) = h_X^{\mathrm{d}}(m)$ for all $m \in I_{N_B}$, we get

$$\forall \alpha \in I_{N_B}\,, \quad \mathcal{F}(\widetilde{h})\left(\frac{2\pi\alpha}{N_B\delta_r}\right) \approx \delta_r \cdot \mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)\,. \tag{65}$$

Besides, combining (20) with (24) and using $\delta_r/\delta = M/N_B$ yields

$$\forall \alpha \in I_{N_B}\,, \quad \mathcal{F}(\mathcal{R}_\theta(\widetilde{V}_X))\left(\frac{2\pi\alpha}{N_B\delta_r}\right) \approx \delta^2\,\frac{M}{N_B}\cdot \mathrm{DFT}(Z_r\mathcal{R}_\theta^{\mathrm{d}}\widetilde{v}_X)(\alpha)\,. \tag{66}$$

Then, using the approximations (64), (65) and (66) in (61) yields

$$\forall \alpha \in I_{N_B}\,, \quad \mathrm{DFT}(\widetilde{s}_\theta^{\mathrm{d}})(\alpha) \approx \delta^2\,\frac{M}{N_B}\cdot \mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)\cdot \mathrm{DFT}(Z_r\mathcal{R}_\theta^{\mathrm{d}}\widetilde{v}_X)(\alpha)\,. \tag{67}$$

Eventually, taking the IDFT of (67) and remarking that $\widetilde{s}_\theta^{\mathrm{d}}(m) = \widetilde{S}_\theta(B_{-m})$ for all $m \in I_{N_B}$ leads exactly to (26).

C Discrete Divergence Operator

In the 2D setting, the discrete divergence operator $\mathrm{div}_{\mathrm{2D}} = -\nabla_{\mathrm{2D}}^*$ satisfies

$$\forall p = (p_x, p_y) \in \mathbb{R}^{\Omega_{\mathrm{2D}}} \times \mathbb{R}^{\Omega_{\mathrm{2D}}}\,, \quad \mathrm{div}_{\mathrm{2D}}(p) = \mathrm{div}_x^{\mathrm{2D}}(p_x) + \mathrm{div}_y^{\mathrm{2D}}(p_y)\,, \tag{68}$$

where, for all $(k,\ell) \in \Omega_{\mathrm{2D}}$,

$$\mathrm{div}_x^{\mathrm{2D}}(p_x)(k,\ell) = \begin{cases} p_x(k,\ell) & \text{if } k = -\lfloor M/2 \rfloor \\ p_x(k,\ell) - p_x(k-1,\ell) & \text{if } -\lfloor M/2 \rfloor < k < M - \lfloor M/2 \rfloor - 1 \\ -p_x(k-1,\ell) & \text{if } k = M - \lfloor M/2 \rfloor - 1 \end{cases}$$

$$\mathrm{div}_y^{\mathrm{2D}}(p_y)(k,\ell) = \begin{cases} p_y(k,\ell) & \text{if } \ell = -\lfloor M/2 \rfloor \\ p_y(k,\ell) - p_y(k,\ell-1) & \text{if } -\lfloor M/2 \rfloor < \ell < M - \lfloor M/2 \rfloor - 1 \\ -p_y(k,\ell-1) & \text{if } \ell = M - \lfloor M/2 \rfloor - 1\,, \end{cases}$$

denoting by $\lfloor x \rfloor$ the lower integer part of x. Similarly, in the 3D setting, the discrete divergence operator $\mathrm{div}_{\mathrm{3D}} = -\nabla_{\mathrm{3D}}^*$ satisfies

$$\forall p = (p_x, p_y, p_z) \in \mathbb{R}^{\Omega_{\mathrm{3D}}} \times \mathbb{R}^{\Omega_{\mathrm{3D}}} \times \mathbb{R}^{\Omega_{\mathrm{3D}}}\,, \quad \mathrm{div}_{\mathrm{3D}}(p) = \mathrm{div}_x^{\mathrm{3D}}(p_x) + \mathrm{div}_y^{\mathrm{3D}}(p_y) + \mathrm{div}_z^{\mathrm{3D}}(p_z)\,, \tag{69}$$

where, for all $(k,\ell,m) \in \Omega_{\mathrm{3D}}$,

$$\mathrm{div}_x^{\mathrm{3D}}(p_x)(k,\ell,m) = \begin{cases} p_x(k,\ell,m) & \text{if } k = -\lfloor M/2 \rfloor \\ p_x(k,\ell,m) - p_x(k-1,\ell,m) & \text{if } -\lfloor M/2 \rfloor < k < M - \lfloor M/2 \rfloor - 1 \\ -p_x(k-1,\ell,m) & \text{if } k = M - \lfloor M/2 \rfloor - 1 \end{cases}$$

$$\mathrm{div}_y^{\mathrm{3D}}(p_y)(k,\ell,m) = \begin{cases} p_y(k,\ell,m) & \text{if } \ell = -\lfloor M/2 \rfloor \\ p_y(k,\ell,m) - p_y(k,\ell-1,m) & \text{if } -\lfloor M/2 \rfloor < \ell < M - \lfloor M/2 \rfloor - 1 \\ -p_y(k,\ell-1,m) & \text{if } \ell = M - \lfloor M/2 \rfloor - 1 \end{cases}$$

$$\mathrm{div}_z^{\mathrm{3D}}(p_z)(k,\ell,m) = \begin{cases} p_z(k,\ell,m) & \text{if } m = -\lfloor M/2 \rfloor \\ p_z(k,\ell,m) - p_z(k,\ell,m-1) & \text{if } -\lfloor M/2 \rfloor < m < M - \lfloor M/2 \rfloor - 1 \\ -p_z(k,\ell,m-1) & \text{if } m = M - \lfloor M/2 \rfloor - 1\,. \end{cases}$$

D Proof of Proposition 1

Let $v \in \mathbb{R}^{\Omega_{\mathrm{2D}}}$. For all $p \in \{1,2,\dots,N_\theta\}$, let us denote by $\mathcal{A}_p v : I_{N_B} \to \mathbb{R}$ the projection in the direction θ_p, i.e.,

$$\forall m \in I_{N_B}\,, \quad \mathcal{A}_p v(m) := \mathcal{A}v(m,p) = \delta^2 \frac{M}{N_B} \cdot \left(h_X^{\mathrm{d}} \circledast Z_r \mathcal{R}_{\theta_p}^{\mathrm{d}} v \right)(m)\,.$$

Then, from (29), for all $(k,\ell) \in \Omega_{\mathrm{2D}}$, we have

$$\mathcal{A}^*\mathcal{A}v(k,\ell) = \frac{\delta^2}{N_B} \cdot \sum_{\substack{\alpha\in\mathbb{Z},\ |\alpha|\le M/2 \\ p\in\{1,2,\dots,N_\theta\}}} \overline{\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)} \cdot \mathrm{DFT}(\mathcal{A}_p v)(\alpha) \cdot e^{2i\pi\left(k\frac{\alpha}{M}\cos\theta_p + \ell\frac{\alpha}{M}\sin\theta_p\right)}\,. \tag{70}$$

Besides, for any $\alpha \in \mathbb{Z}$ such as $|\alpha| < M/2$, we have

$$\mathrm{DFT}(\mathcal{A}_p v)(\alpha) = \delta^2 \frac{M}{N_B}\,\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)\,\mathrm{DFT}(Z_r\mathcal{R}_{\theta_p}^{\mathrm{d}} v)(\alpha) = \delta^2\,\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)\,\mathrm{DFT}(\mathcal{R}_{\theta_p}^{\mathrm{d}} v)(\alpha)\,, \tag{71}$$

and from (22) we have

$$\mathrm{DFT}(\mathcal{R}^{\mathrm{d}}_{\theta_p} v)(\alpha) = \sum_{(k',\ell')\in\Omega_{2\mathrm{D}}} v(k',\ell')\, e^{-2i\pi\left(k'\frac{\alpha}{M}\cos\theta_p + \ell'\frac{\alpha}{M}\sin\theta_p\right)} . \tag{72}$$

Using (71) and (72) into (70), we obtain

$$\forall (k,\ell)\in\Omega_{2\mathrm{D}}\,, \quad \mathcal{A}^*\mathcal{A}v(k,\ell) = \sum_{(k',\ell')\in\Omega_{2\mathrm{D}}} v(k',\ell')\,\varphi(k-k',\ell-\ell')\,, \tag{73}$$

where φ is the kernel defined in (39) over the augmented domain $\Upsilon_{2\mathrm{D}} = I_{2M}\times I_{2M}$. Now, let us consider the circular convolution between φ and Zv, $\varphi\circledast Zv = \mathrm{IDFT}(\mathrm{DFT}(\varphi)\cdot\mathrm{DFT}(Zv))$. We have,

$$\forall (k,\ell)\in\Upsilon_{2\mathrm{D}}\,, \quad (\varphi\circledast Zv)\,(k,\ell) = \sum_{(k',\ell')\in\Upsilon_{2\mathrm{D}}} Zv(k',\ell')\,\varphi_{\mathrm{per}}(k-k',\ell-\ell')\,,$$

where we have denoted by φ_{per} the $\Upsilon_{2\mathrm{D}}$-periodical extension of φ to the domain $\mathbb{Z}^2$. Since Zv is zero outside from $\Omega_{2\mathrm{D}}$, we can restrict the latter sum to $(k',\ell')\in\Omega_{2\mathrm{D}}$, leading to

$$\forall (k,\ell)\in\Upsilon_{2\mathrm{D}}\,, \quad (\varphi\circledast Zv)\,(k,\ell) = \sum_{(k',\ell')\in\Omega_{2\mathrm{D}}} Zv(k',\ell')\,\varphi_{\mathrm{per}}(k-k',\ell-\ell')\,.$$

Besides, when (k,ℓ) lies in $\Omega_{2\mathrm{D}}$, then, for all $(k',\ell')\in\Omega_{2\mathrm{D}}$, we have $(k-k',\ell-\ell')\in\Upsilon_{2\mathrm{D}}$, and thus, $\varphi_{\mathrm{per}}(k-k',\ell-\ell') = \varphi(k-k',\ell-\ell')$. Therefore, we have

$$\forall (k,\ell)\in\Omega_{2\mathrm{D}}\,, \quad (\varphi\circledast Zv)\,(k,\ell) = \sum_{(k',\ell')\in\Omega_{2\mathrm{D}}} Zv(k',\ell')\,\varphi(k-k',\ell-\ell')\,,$$

which corresponds to (73), showing that the restriction of $\varphi\circledast Zv$ to $\Omega_{2\mathrm{D}}$ is equal to $\mathcal{A}^*\mathcal{A}v$, as announced.

E A Contrario Detection of Frequency Support

Assuming an additive Gaussian noise model for the reference spectrum acquisition, we have

$$\forall k\in I_{N_B}\,, \quad h^{\mathrm{d}}_X(k) = \overline{h}^{\mathrm{d}}_X(k) + \varepsilon(k)\,, \tag{74}$$

where $\overline{h}^{\mathrm{d}}_X$ represents the noise-free reference spectrum and $(\varepsilon(k))_{k\in I_{N_B}}$ are independent and identically distributed realizations of a Gaussian random variable with zero mean and variance σ^2. When the value of σ is unknown, it can be empirically estimated from the values of h^{d}_X near the boundaries of I_{N_B}. In oversampling situations, i.e., when $N_B > M$, the DFT coefficients $\mathrm{DFT}(\overline{h}^{\mathrm{d}}_X)(\alpha)$ vanish for $M/2 < |\alpha| \leq N_B/2$. Due to the presence of the noise in (74), this is not the case for the DFT coefficients of the observed spectrum h^{d}_X. In order to estimate the value of M from h^{d}_X, we propose to use the a contrario methodology [8]. The latter is based on the *Helmoltz principle* for visual perception which states that perceived structures are those that are unlikely to happen by chance in a random configuration. The a contrario methodology can be viewed as a mathematical formulation of this principle and consists in the design of detectors based on a rejection principle of randomness. An a contrario detector relies on two main ingredients: (i) a random model $\mathcal{H}_0$ describing pure noise data in which no structure must be found, (ii) a measurement function that highlight structures that are unlikely to happen by chance in $\mathcal{H}_0$. In the present case, pure noise measurements correspond to signals h^{d}_X obtained when $\overline{h}^{\mathrm{d}}_X \equiv 0$ in (74). We define the $\mathcal{H}_0$ model exactly in this way.

Definition 3 ($\mathcal{H}_0$ model)**.** *We say that* $h_X^{\mathrm{d}} : I_{N_B} \to \mathbb{R}$ *follows* $\mathcal{H}_0$*, and we note* $h_X^{\mathrm{d}} \sim \mathcal{H}_0$*, if and only if* $\{h_X^{\mathrm{d}}(k)\}_{k\in I_{N_B}}$ *is a family of independent random variables following a Gaussian distribution with zero-mean and variance* σ^2*.*

Now, let us design a measurement function able to highlight signals h_X^{d} with frequency coefficients with significantly higher intensity than expected when $h_X^{\mathrm{d}} \sim \mathcal{H}_0$.

Proposition 2 (measurement function and its distribution in $\mathcal{H}_0$)**.** *Let* $h_X^{\mathrm{d}} : I_{N_B} \to \mathbb{R}$ *and let* $Z : [1, \frac{N_B}{2}] \cap \mathbb{Z} \to \mathbb{R}_+$ *the measurement function associated to* h_X^{d} *defined by*

$$\forall m \in \left[1, \frac{N_B}{2}\right] \cap \mathbb{Z}\,, \quad Z(m) = \sum_{\alpha=0}^{m-1} \frac{\left|\mathrm{DFT}(h_X^{\mathrm{d}}(\alpha))\right|^2}{\sigma^2 N_B}\,. \tag{75}$$

Then, when $h_X^{\mathrm{d}} \sim \mathcal{H}_0$*, we have*

$$\forall m \in \left[0, \frac{N_B}{2}\right] \cap \mathbb{Z}\,, \quad \forall z \geq 0\,, \quad \mathbb{P}_{\mathcal{H}_0}\left(Z(m) \geq z\right) = \frac{\Gamma(m,z)}{\Gamma(m)}\,, \tag{76}$$

where

$$\Gamma(m) = \int_0^{+\infty} s^{m-1}\, e^{-s}\, ds \quad \text{and} \quad \Gamma(m,z) = \int_z^{+\infty} s^{m-1}\, e^{-s}\, ds \tag{77}$$

denote respectively the complete and (upper) incomplete Gamma integrals.

Proof. Let $h_X^{\mathrm{d}} \sim \mathcal{H}_0$, and let $X_\alpha = \mathrm{Re}\left(\mathrm{DFT}(h_X^{\mathrm{d}}(\alpha))\right)$ and $Y_\alpha = \mathrm{Im}\left(\mathrm{DFT}(h_X^{\mathrm{d}}(\alpha))\right)$. Then, the set $\{X_\alpha, Y_\alpha\}_{0\leq\alpha\leq\frac{N_B}{2}}$ is made of independent Gaussian random variables with zero mean and satisfying $\mathbb{E}\left(X_\alpha^2\right) + \mathbb{E}\left(Y_\alpha^2\right) = \sigma^2 N_B$ (more precisely, we have $\mathbb{E}\left(X_\alpha^2\right) = \mathbb{E}\left(Y_\alpha^2\right) = \frac{\sigma^2 N_B}{2}$ when $0 < \alpha < \frac{N_B}{2}$, and we have $\mathbb{E}\left(X_\alpha\right) = \sigma^2 N_B$ and $\mathbb{E}\left(Y_\alpha^2\right) = 0$ when $\alpha \in \left\{0, \frac{N_B}{2}\right\}$). Therefore, the random variables $\frac{X_\alpha^2+Y_\alpha^2}{\sigma^2 N_B}$ are independent and follow an exponential distribution with unitary mean. Consequently, for all $m \in \left[1, \frac{N_B}{2}\right] \cap \mathbb{Z}$, the random variable $Z(m) = \sum_{\alpha=0}^{m-1} \frac{X_\alpha^2+Y_\alpha^2}{\sigma^2 N_B}$ follows a Gamma distribution with shape m and unit scale, with complementary cumulative distribution function given by (76). ☐

The measurement $Z(m)$ defined in (75) consists in grouping together the m first DFT coefficients of h_X^{d} and computing (up to the normalization factor $\sigma^2 N_B$) the sum of their intensities. The measurement function Z is suited to discriminate pure noise signals following $\mathcal{H}_0$ from a real-life reference spectrum like that displayed in Figure 4. Indeed, measurements $Z(m)$ made on signals h_X^{d} like in Figure 4 (a) are significantly larger than those made on signals following $\mathcal{H}_0$. The complementary cumulative distribution function (76) can be used to quantify the amounts of surprise associated to the measurement $Z(m) = z$. Multiplying this quantity by $N_B/2$ (which upper bounds the number of tested groupings), we obtain a Number of False Alarms (NFA),

$$\forall m \in \left[1, \frac{N_B}{2}\right] \cap \mathbb{Z}\,, \quad \mathrm{NFA}(m) = \frac{N_B}{2} \cdot \frac{\Gamma(m, Z(m))}{\Gamma(m)}\,, \tag{78}$$

and we say that the grouping made of the m first DFT coefficients of h_X^{d} is ε-meaningful when $\mathrm{NFA}(m) \leq \varepsilon$. Thanks to [16, Proposition 2], the NFA (78) satisfies the so-called *NFA-property*,

$$\forall \varepsilon > 0\,, \quad \mathbb{E}_{\mathcal{H}_0}\left[\#\left\{m \in \left[1, \frac{N_B}{2}\right] \cap \mathbb{Z}\,,\ \mathrm{NFA}(m) \leq \varepsilon\right\}\right] \leq \varepsilon\,, \tag{79}$$

denoting by # the cardinality of the set in (79). The NFA property ensures that, when $h_X^{\mathrm{d}} \sim \mathcal{H}_0$, in average, we find less than ε groupings that are ε-meaningful. This means that, detecting ε-meaningful groupings by thresholding the NFA with threshold ε yields (in average) less than ε false detections when h_X^{d} is a pure noise following $\mathcal{H}_0$. This gives a tangible sense to the detection threshold ε which corresponds to the maximal number of (false) detections allowed in pure noise data. In particular, a common setting for this threshold parameter is $\varepsilon = 1$ so that, in average, at most one false detection is allowed in $\mathcal{H}_0$. In practice, we are interested in the detection of the most meaningful grouping (i.e., of the grouping with smallest NFA),

$$\overline{m} = \operatorname*{argmin}_{1 \leq m \leq \frac{N_B}{2}} \; \mathrm{NFA}(m)\,. \tag{80}$$

Then, when $\mathrm{NFA}(\overline{m}) \leq \varepsilon$, extending the frequency grouping by Hermitian symmetry yields the detected frequency support $\alpha \in [-\overline{m}, \overline{m}] \cap I_{N_B}$, with size $M = \min(N_B, 2\,\overline{m} + 1)$. Notice that the NFFT implementation [21] is limited to signals with even dimensions, so that in practice, we set

$$M = 2 \cdot \min\left(\left\lfloor \frac{N_B}{2} \right\rfloor, \overline{m} + 1\right)\,, \tag{81}$$

denoting by $\lfloor \cdot \rfloor$ the lower integer part. Note also that some particular care must be taken to avoid underflow when evaluating $\mathrm{NFA}(m)$ using (78). Numerical underflow typically occurs when $Z(m)$ is large and leads to $\mathrm{NFA}(m) = 0$ in standard double precision. This may happen for many values of m and this is problematic to compute $\overline{m}$ using (80). Numerical underflow can be avoided using an appropriate normalization for the upper incomplete Gamma function [1] (see also the `scale` optional parameter of the `gammainc` function in Matlab language) allowing its evaluation in logarithmic scale. Finally, the estimation of M is summarized in Algorithm 7.

Algorithm 7: support size estimation from a discrete reference spectrum

```
Inputs: a discrete reference spectrum $h_X^{\mathrm{d}} : I_{N_B} \to \mathbb{R}$ with size $N_B$, the standard deviation $\sigma$
  of the noise corrupting $h_X^{\mathrm{d}}$, the NFA thresholding parameter $\varepsilon > 0$ (default setting $\varepsilon = 1$).
Output: an estimate of the size $M$ of the frequency support of $h_X^{\mathrm{d}}$.
Core of the module:
 // compute NFA in logarithmic scale
for $1 \leq m \leq \frac{N_B}{2}$ do
    $Z(m) \leftarrow \sum_{\alpha=1}^{m} \frac{|\mathrm{DFT}(h_X^{\mathrm{d}})(\alpha)|^2}{\sigma^2 N_B}$
    log_nfa$(m) \leftarrow \log\left(\frac{N_B}{2}\right) + \log\left(\Gamma(m, Z(m))\right) - \log\left(\Gamma(m)\right)$
 // find the most meaningful grouping
$\overline{m} \leftarrow \operatorname*{argmin}_{1 \leq m \leq \frac{N_B}{2}}$ log_nfa$(m)$  // take highest value of $m$ in case of multiple minimizers
 // compute the size of the frequency support
if log_nfa$(\overline{m}) \leq \log(\varepsilon)$  // meaningful frequency grouping detected
 then
    $M \leftarrow 2 \cdot \min\left(\left\lfloor \frac{N_B}{2} \right\rfloor, \overline{m} + 1\right)$ // ensure $M \leq N_B$ and even value for $M$
else  $M \leftarrow 0$  // no meaningful frequency grouping detected

return $M$
```

F Using the Companion Online Demo

An online demonstrator for 2D EPR image reconstruction is provided with this article. The latter can be used to apply Algorithm 5 to several example EPR datasets containing a reference spectrum and a 2D sinogram (see Figure 22), or to custom datasets provided by the user in Bruker BES3T format (through the *upload data* button visible in Figure 22). The setting of the reconstruction parameters through the demonstrator interface is detailed in Figure 23. More advanced operations (such as the optimal tuning of the resolution and the number of samples to acquire) as well as computationally intensive reconstruction tasks (such as 3D EPR image reconstruction using Algorithm 6) are not covered by this demonstrator but can be achieved using the (documented) source code published with this work.

Figure 22: **Input datasets.** Several example datasets are provided in the demonstrator. Those datasets are made of a reference spectrum h_X^{d} and a 2D sinogram $s_{2\mathrm{D}}$. Custom datasets in Bruker BES3T format can also be considered using the *Upload data* button.

(a) Reconstruction size (% of the number of samples per projection) 10 Max: 100 Min: 1 The size of the output image (M = width = height) will be computed as the specified percentage of the number of samples per projection.

(b) TV weight λ' 50 Max: 10000 Min: 1e-10 Normalized TV weight: this parameter controls the relative weight of the TV penalty with respect to the least-square data-fidelity term in the reconstruction process. The larger this parameter is, the smoother the reconstructed image will be.

(c) Huber parameter α' 0 Max: 1 Normalized Huber parameter: nonzero value for this parameter can be used to replace the standard (cartoon image promoting) TV penalty by its so-called Huber variant. Large values for this parameter would yield more textured but also blurrier images.

(d) Number of iterations 300 Max: 5000 Min: 1 Number of iterations for the minimization scheme

(e) Apply affine baseline correction to the input dataset (preprocessing) ☑ Perform a baseline correction of the dataset (useful for datasets with nonzero baseline due to acquisition issues); the specified percentage of points at both extremities of the reference spectrum will be used to estimate the affine correction.

(f) Baseline correction percentage 10 Max: 49 Percentage of points at both extremities of the reference spectrum to estimate the affine correction.

Figure 23: **Parameters setting.** The *Parameters* section of the demonstrator interface contains sliders that can be used to (directly or indirectly) tune the values of the parameters of Algorithm 5. The effect of each slider is detailed below.

(a) This slider can be used to set the reconstruction size M (in pixels) of the image to reconstruct as a percentage of the number of samples per projection (N_B) of the input sinogram $s_{2\mathrm{D}}$. Denoting by p the entered slider value, we set

$$M' = \left\lceil \frac{p}{100} N_B \right\rceil \quad \text{and} \quad M = \max\left(2, M' - (M' \bmod 2)\right)$$

where $\lceil x \rceil$ denotes the upper integer part of x. This ensures that M is a nonzero even integer (which is a requirement for the NFFT library [21] upon which our implementation relies on). We recall that the spatial sampling step (or pixel size) of the image to be computed is $\delta = \frac{\delta_B}{\mu} \cdot \frac{N_B}{M}$, where μ denotes the field gradient intensity and δ_B denotes the sampling step of the homogeneous magnetic field used to acquire the dataset. The value of the pixel size δ is systematically printed out (see the *Information* section of the demonstrator interface) after the reconstruction.

(b) This slider can be used to set the value of the normalized regularity parameter λ'. Then, (54) is used to compute the value of the λ input of Algorithm 5. The effect of this parameter on the reconstructed image is illustrated in Section 7.2.1.

(c) We explained in Section 5 how Algorithm 5 could be extended to handle the minimization of the Huber TV regularized energy (50). This slider can be used to set the value of the normalized Huber smoothing parameter α'. Then, (58) is used to compute the value of the Huber smoothing parameter α involved in (50). The effect of the Huber smoothing parameter on the reconstructed image is illustrated in Section 7.2.2. Notice that, when $\alpha' = 0$ (which is the default setting), the more classical TV regularized least-squares problem (32) is addressed rather than its Huber variant.

(d) This slider can be used to set the value of the number N of iterations to perform in Algorithm 5.

(e) In some situations, the acquired reference spectrum and the projections (in the acquired sinogram) can be corrupted by an affine offset leading to non-horizontal baseline. This affine baseline usually causes undesirable low frequency artifacts in the reconstruction. When this checkbox is enabled, affine regression is used to estimate the slope and the offset of the affine baseline from the reference spectrum (see Figure 24). Then, the estimated affine baseline is subtracted to h_X^{d} as well as to each projection of the sinogram $s_{2\mathrm{D}}$ before running the image reconstruction using Algorithm 5.

(f) This slider is visible only when the checkbox (e) is enabled. It can be used to set the percentage of samples of the reference spectrum to be used to estimate the slope and the offset of the affine pattern. Denoting by p the value of the slider, we set $n = \max(1, \lfloor p \cdot N_B \rceil)$ (denoting by $\lfloor x \rceil$ the rounded integer part of x) and select the n first samples and the n last samples observed in h_X^{d} to estimate the affine baseline.

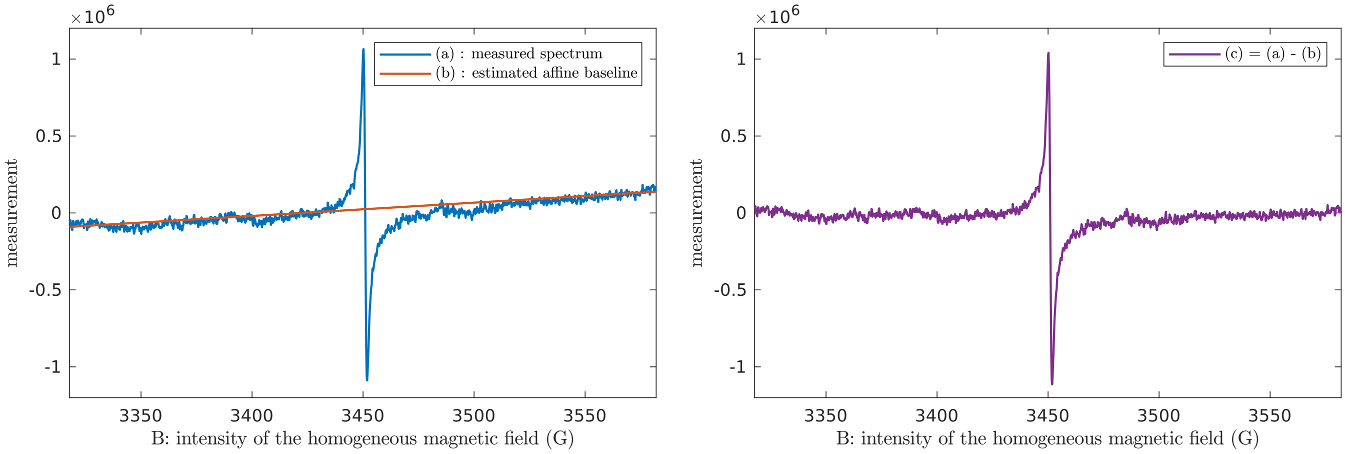


Figure 24: **Affine baseline correction.** The blue curve (a) represents a spectrum acquisition with shifted baseline. Affine least-squares regression was applied to the first and last 10% of the spectrum samples to estimate the baseline, leading to the orange curve (c). Subtracting the estimated baseline to the measured spectrum yields the corrected spectrum, displayed using a purple curve on the right-hand side of this figure.

References

[1] R. Abergel and L. Moisan, *Algorithm 1006: Fast and accurate evaluation of a generalized incomplete gamma function*, ACM Transactions on Mathematical Software, 46 (2020). https://doi.org/10.1145/3365983.

[2] R. Ahmad, A. Samouilov, and J. Zweier, *Accelerated dynamic EPR imaging using fast acquisition and compressive recovery*, Journal of Magnetic Resonance, 273 (2016), pp. 105–112. https://doi.org/10.1016/j.jmr.2016.10.001.

[3] K. Bredies, K. Kunisch, and T. Pock, *Total generalized variation*, SIAM Journal on Imaging Sciences, 3 (2010), pp. 492–526. https://doi.org/10.1137/090769521.

[4] A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numerische Mathematik, 76 (1997), pp. 167–188. https://doi.org/10.1007/s002110050258.

[5] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 120–145. https://doi.org/10.1007/s10851-010-0251-1.

[6] A. Chambolle and T. Pock, *On the ergodic convergence rates of a first-order primal–dual algorithm*, Mathematical Programming, 159 (2016), pp. 253–287. https://doi.org/10.1007/s10107-015-0957-3.

[7] L. Condat, *A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms*, Journal of Optimization Theory and Applications, 158 (2013), pp. 460–479. https://doi.org/10.1007/s10957-012-0245-9.

[8] A. Desolneux, L. Moisan, and J.-M. Morel, *From Gestalt Theory to Image Analysis. A Probabilistic Approach*, Springer-Verlag, collection “Interdisciplinary Applied Mathematics”, 2008. https://doi.org/10.1007/978-0-387-74378-3.

[9] Y. Drori, S. Sabach, and M. Teboulle, *A simple algorithm for a class of nonsmooth convex–concave saddle-point problems*, Operations Research Letters, 43 (2015), pp. 209–214. https://doi.org/10.1016/j.orl.2015.02.001.

[10] S. Durand, Y.-M. Frapart, and M. Kerebel, *Electron paramagnetic resonance image reconstruction with total variation and curvelets regularization*, Inverse Problems, 33 (2017), p. 114002. https://doi.org/10.1088/1361-6420/aa8412.

[11] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Society for Industrial and Applied Mathematics, 1999. https://epubs.siam.org/doi/abs/10.1137/1.9781611971088.

[12] B. Epel, G. Redler, and H. J. Halpern, *How in vivo EPR measures and images oxygen*, in Oxygen Transport to Tissue XXXVI, Harold M. Swartz, David K. Harrison, and Duane F. Bruley, eds., New York, NY, 2014, Springer New York, pp. 113–119. https://doi.org/10.1007/978-1-4939-0620-8_15.

[13] P. Fattibene and F. Callens, *EPR dosimetry with tooth enamel: A review*, Applied Radiation and Isotopes, 68 (2010), pp. 2033–2116. https://doi.org/10.1016/j.apradiso.2010.05.016.

[14] M. Frigo and S. G. Johnson, *The design and implementation of FFTW3*, Proceedings of the IEEE, 93 (2005), pp. 216–231. Special issue on “Program Generation, Optimization, and Platform Adaptation”. `https://doi.org/10.1109/JPROC.2004.840301`.

[15] N. Grinberg and S. Rodriguez, *Ewing’s analytical instrumentation handbook*, CRC Press, fourth ed., March 2019. ISBN 9781482218671.

[16] B. Grosjean and L. Moisan, *A-contrario detectability of spots in textured backgrounds*, Journal of Mathematical Imaging and Vision, 33 (2009), pp. 313–337. `https://doi.org/10.1007/s10851-008-0111-4`.

[17] J. Jiang and R. T. Weber, *ELEXSYS E 500 User’s Manual Basic Operations*, Bruker BioSpin Corporation, EPR Division, Billerica, MA USA, April 2001. Version 2.0, Part Number 8637060.

[18] C. Johnson, D. McGarry, J. Cook, N. Devashayam, J. Mitchell, S. Subramanian, and M. Krishna, *Maximum entropy reconstruction methods in electron paramagnetic resonance imaging*, Annals of Operations Research, 119 (2003), pp. 101–118. `https://doi.org/10.1023/A:1022978322046`.

[19] D. Johnson, R. Ahmad, G. He, A. Samouilov, and J. Zweier, *Compressed sensing of spatial electron paramagnetic resonance imaging*, Magnetic Resonance in Medecine, 72 (2014), pp. 893–901. `https://doi.org/10.1002/mrm.24966`.

[20] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, 2001. `https://doi.org/10.1137/1.9780898719277.fm`.

[21] J. Keiner, S. Kunis, and D. Potts, *Using NFFT 3—A Software Library for Various Nonequispaced Fast Fourier Transforms*, ACM Transactions on Mathematical Software, 36 (2009). `https://doi.org/10.1145/1555386.1555388`.

[22] M. Kerebel, *Méthodes variationnelles pour l’imagerie en résonance paramagnétique électronique*, theses, Université Sorbonne Paris Cité, 2017. `https://tel.archives-ouvertes.fr/tel-02122784`.

[23] N. Khan, B. B. Williams, and H. M. Swartz, *Clinical applications of in vivo EPR: rationale and initial results*, Applied magnetic Resonance, 30 (2006), pp. 185–199. `https://doi.org/10.1007/BF03166718`.

[24] D.A. Komarov, A. Samouilov, R. Ahmad, and J. Zweier, *Algebraic reconstruction of 3D spatial EPR images from high numbers of noisy projections: An improved image reconstruction technique for high resolution fast scan EPR imaging*, Journal of Magnetic Resonance, 319 (2020). `https://doi.org/10.1016/j.jmr.2020.106812`.

[25] C. Kostov, *Toeplitz structure in slant-stack inversion*, Society of Exploration Geophysicists, (1990). `https://doi.org/10.1190/1.1890075`.

[26] C. Louchet and L. Moisan, *Total variation denoising using posterior expectation*, in European Signal Processing Conference (EUSIPCO), 2008, pp. 1–5. `https://hal.archives-ouvertes.fr/hal-00258849`.

[27] C. Louchet and L. Moisan, *Posterior expectation of the total variation model: Properties and experiments*, SIAM Journal on Imaging Sciences, 6 (2013), pp. 2640–2684. `https://doi.org/10.1137/120902276`.

[28] C. Louchet and L. Moisan, *Total variation denoising using iterated conditional expectation*, in European Signal Processing Conference (EUSIPCO), 2014, pp. 1592–1596. https://hal.archives-ouvertes.fr/hal-01214735.

[29] H. S. Mchaourab, M. A. Lietzow, K. Hideg, and W. L. Hubbell, *Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics*, Biochemistry, 35 (1996), pp. 7692–7704. https://doi.org/10.1021/bi960482k.

[30] Y. Meyer, *Wavelets and Operators*, vol. 37, Cambridge university press, 1992. ISBN 9780521420006.

[31] K. Ohno, *Two-dimensional ESR imaging for paramagnetic species with anisotropic parameters*, Journal of Magnetic Resonance (1969), 64 (1985), pp. 109–114. https://doi.org/10.1016/0022-2364(85)90036-8.

[32] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, SIAM Journal on Multiscale Modeling and Simulation, 4 (2005). https://doi.org/10.1137/040605412.

[33] Z. Qiao, D. Liang, S. Tang, and H. Halpern, *Optimization-based image reconstruction from fast-scanned, noisy projections in EPR imaging*, IEEE Access, 7 (2019), pp. 19590–19601. https://doi.org/10.1109/ACCESS.2019.2897140.

[34] Z. Qiao, G. Redler, B. Epel, and H. Halpern, *A balanced total-variation-Chambolle-Pock algorithm for EPR imaging*, Journal of Magnetic Resonance, 328 (2021). https://doi.org/10.1016/j.jmr.2021.107009.

[35] Z. Qiao, G. Redler, B. Epel, Y. Qian, and H. Halpern, *3D pulse EPR imaging from sparse-view projections via constrained total variation minimization*, Journal of Magnetic Resonance, 258 (2015), pp. 49–57. https://doi.org/10.1016/j.jmr.2015.06.009.

[36] R. T. Rockafellar, *Convex analysis (Princeton mathematical series)*, Princeton University Press, 46 (1970), p. 49. https://doi.org/10.1515/9781400873173.

[37] C. E. Shannon, *Communication in the presence of noise*, Proceedings of the Institute of Radio Engineers, 37 (1949), pp. 10–21. https://doi.org/10.1109/JRPROC.1949.232969.

[38] S. Stoll and A. Schweiger, *EasySpin, a comprehensive software package for spectral simulation and analysis in EPR*, Journal of Magnetic Resonance, 178 (2006), pp. 42–55. https://doi.org/10.1016/j.jmr.2005.08.013.

[39] H. M. Swartz, H. Hou, N. Khan, L. A. Jarvis, E. Y. Chen, B. B. Williams, and P. Kuppusamy, *Advances in probes and methods for clinical EPR oximetry*, in Oxygen Transport to Tissue XXXVI, Harold M. Swartz, David K. Harrison, and Duane F. Bruley, eds., New York, NY, 2014, Springer New York, pp. 73–79. https://doi.org/10.1007/978-1-4939-0620-8_10.

[40] M. Tseitlin, T. Czechowski, S. S. Eaton, and G. R. Eaton, *Regularized optimization (RO) reconstruction for oximetric EPR imaging*, Journal of Magnetic Resonance, 194 (2008), pp. 212–221. https://doi.org/10.1016/j.jmr.2008.07.002.

[41] C.R. Vogel and M.E. Oman, *Iterative methods for total variation denoising*, SIAM Journal on Scientific Computing, 17 (1996), pp. 227–238. https://doi.org/10.1137/0917016.

[42] B. C. Vũ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Advances in Computational Mathematics, 38 (2013), pp. 667–681. `https://doi.org/10.1007/s10444-011-9254-8`.