

## Workshop Muon and Neutrino Tomography - Clermont Ferrand



### Inverse problems for reconstruction in tomography

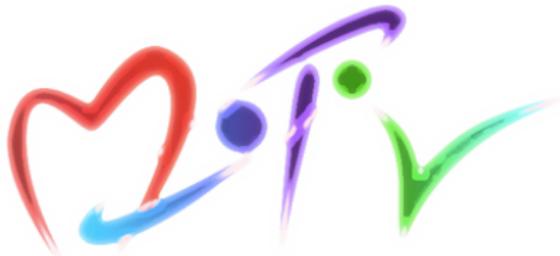
Fabien Momey<sup>1,2</sup> Jean-Marie Becker<sup>2</sup> Loïc Denis<sup>1,2</sup> Laurent Desbat<sup>3</sup>  
Catherine Mennessier<sup>2</sup> Eric Thiébaut<sup>1</sup>

<sup>1</sup>Centre de Recherche Astrophysique de Lyon - Observatoire de Lyon

<sup>2</sup>Laboratoire Hubert Curien - Université Jean Monnet Saint-Étienne

<sup>3</sup>TIMC-IMAG

- Fabien Momey, PhD student: **Inverse problems for reconstruction in dynamic tomography.**
- Makes part of the project **MiTIV**, supported by the french ANR: *Méthodes Inverse pour le Traitement en Imagerie du Vivant.*
  - Blind deconvolution for 3-D microscopy and cardiac imaging (coronarography).
  - Tomography.
  - Astronomy.



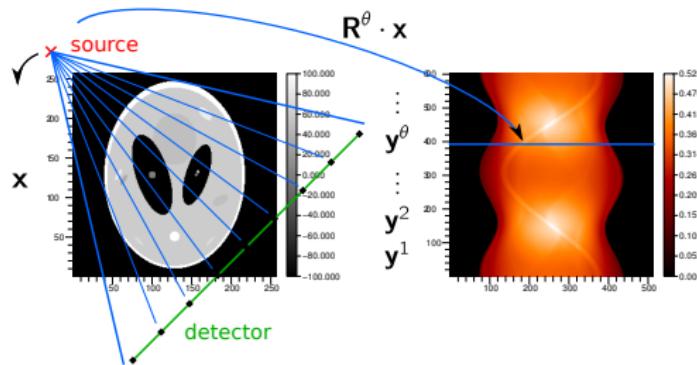
## Tomography : an inverse problem (1)

- **Tomography:** retrieve a 3D information  $x$  from a finite set of projections  $y^\theta$  (radiography principle), acquired on a detector, positioned around the object.

- **Model of the data:**

$$x \quad ? \quad \longleftrightarrow \quad \{y^1, y^2, \dots, y^\theta, \dots\}$$

$$y^\theta = R^\theta \cdot x + e$$



In X-ray tomography, the physical principle of data acquisition is based on the **Beer-Lambert law of absorption**  $\Rightarrow$  we reconstruct an "image" of the absorption coefficients of the object.

## Tomography : an inverse problem (2)

- We resolve the following **inverse problem**:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \underbrace{\sum_{\theta \in \Theta} \|\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x}\|_W^2}_{f_{\text{data}}(\mathbf{x})} + \mu \cdot f_{\text{prior}}(\mathbf{x})$$

This problem is → due to ↓	ill-posed	ill-conditioned
$\mathbf{y}^\theta$	<i>finite and often few number of projections</i>	-
$\mathbf{R}^\theta$	-	<i>approximated numerical model ⇒ an accurate model will be very important</i>

## Data criterion

- **A convex criterion** compares the true data  $\mathbf{y}^\theta$  to pseudo data  $\tilde{\mathbf{y}}^\theta \dots$

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \underbrace{\sum_{\theta \in \Theta} \|\mathbf{y}^\theta - \underbrace{\mathbf{R}^\theta \cdot \mathbf{x}}_{{f}_{\text{data}}(\mathbf{x})}\|_W^2 + \mu \cdot f_{\text{prior}}(\mathbf{x})}$$

Data residuals:  
Weighted least squares (linear and convex)

## Data modelization (1)

- . . . obtained from the reprojection of the current estimate  $\mathbf{x}$ , using a numerical model  $\mathbf{R}^\theta$ .

### 1. Accurate and fast numerical projector

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \underbrace{\sum_{\theta \in \Theta} \|\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x}\|_W^2 + \mu \cdot f_{\text{prior}}(\mathbf{x})}_{f_{\text{data}}(\mathbf{x})}$$

Data residuals:  
Weighted least squares (linear and convex)

## Data modelization (2)

### Object representation

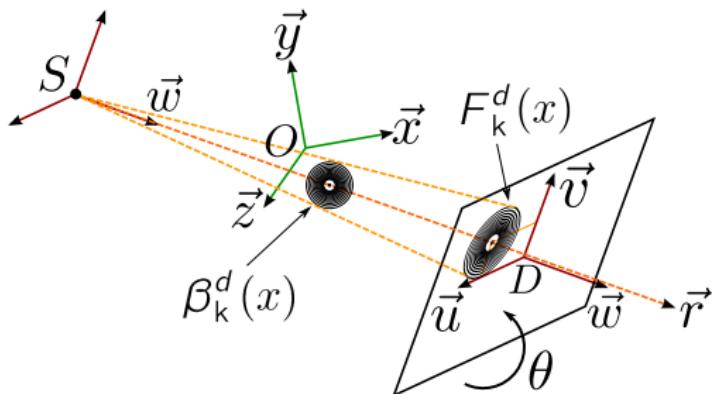
$$f(x) = \sum_{k \in \mathbb{Z}^n} x_k \beta_k^d(x) = \sum_{k \in \mathbb{Z}^n} x_k \beta^d(x - x_k)$$

$$\mathbf{x} = (x_k)^T, \text{ with } x_k = \sum_{k'} x_{k'} \beta_{k'}^d(x_k)$$

### Numerical projector

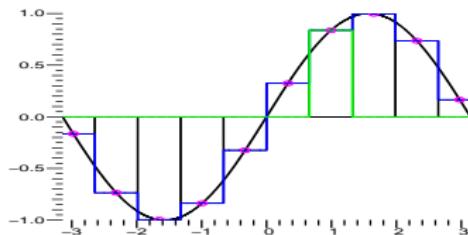
$$\mathbf{y}^\theta = \mathbf{R}^\theta \cdot \mathbf{x}, \text{ with } \mathbf{y}^\theta = (y_q^\theta)^T$$

$$y_q^\theta = \sum_{k \in \Omega_q^\theta} \mathbf{R}_{qk}^\theta \cdot x_k = \sum_{k \in \Omega_q^\theta} \mathbf{R}_q^\theta (\beta_k^d) \cdot x_k$$



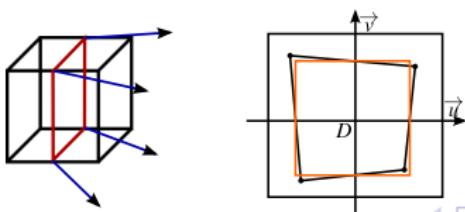
## Data modelization (3)

- Standard modelizations use **staircase voxels**  $\Rightarrow$  coarse representation of the continuity.



**Figure:** Representation (in blue) of the sinus function on  $[-\pi, \pi]$  from 10 regularly spaced staircase basis functions.

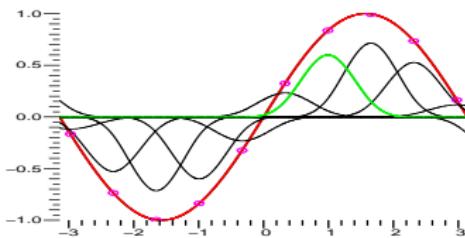
- Example of *Distance Driven* (DeMan & Basu, 2004)  $\Rightarrow$  projection of a cross section of a staircase voxel, approximated by a rectangular footprint on the detector plane.



## Data modelization (4)

- **A better way : modelizing with B-splines**

- Better *modelization of continuity*.
- *Compact support* keeps sparsity of the projection matrix  $\mathbf{R}^\theta \Rightarrow$  fast calculation.
- “*Quasi-sphericity*”  $\Rightarrow$  almost isotropic projection.

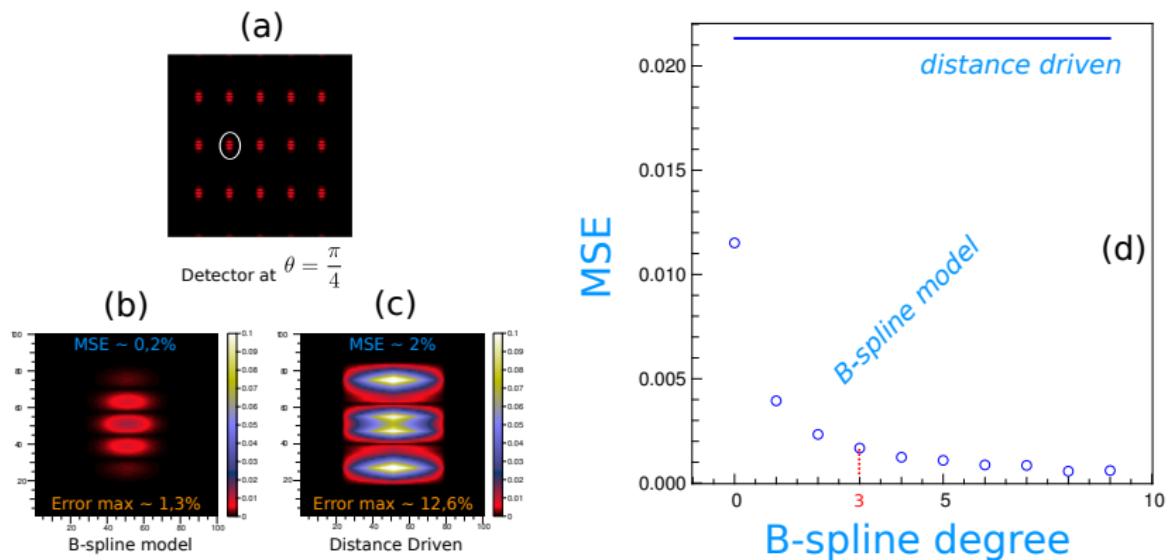


**Figure:** Representation (in red) of the sinus function on  $[-\pi, \pi]$  from 10 regularly spaced B-splines of degree 3.

- **Approximations** are made to calculate the projection of such functions on the detector.

## Data modelization (5)

- Approximation errors of the projection of a basis function  $\beta_k^d$ .



## Regularization (1)

- A regularization term, based on a prior knowledge of the object, compensates for the ill-posedness of the problem, the lack of data, and controls the noise amplification in the solution.

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \underbrace{\sum_{\theta \in \Theta} \|\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x}\|_W^2}_{f_{\text{data}}(\mathbf{x})} + \mu \cdot f_{\text{prior}}(\mathbf{x})$$

1. Accurate and fast numerical projector

2. Regularization term

Data residuals:  
Weighted least squares (linear and convex)

## Regularization (2)

- The regularization term may have different purposes:
  - Pure smoothness of the solution: e.g. a **quadratic penalty of the gradient**.

$$f_{\text{prior}}(\mathbf{x}) = \|\nabla \mathbf{x}\|_{L2}^2 = \sum_{\mathbf{k} \in \mathbb{Z}^n} \left( \left( \frac{\partial \mathbf{x}_k}{\partial x} \right)^2 + \left( \frac{\partial \mathbf{x}_k}{\partial y} \right)^2 \right)$$

- Relaxed smoothness to get a solution preserving high contrasts: e.g.  **$L2 - L1$  smoothness** penalizes quadratically small contrasts and linearly higher ones.

$$f_{\text{prior}}(\mathbf{x}) = \frac{|\nabla \mathbf{x}|}{\epsilon} - \log \left( 1 + \frac{|\nabla \mathbf{x}|}{\epsilon} \right)$$

where  $\epsilon$  is a chosen threshold.

- Get piecewise constant functions: e.g. minimizing the **total variation** ( **$L1$  norm of the gradient**) (Rudin et al, 1992) is very suitable (**but non linear**).

$$f_{\text{prior}}(\mathbf{x}) = \|\nabla \mathbf{x}\|_{L1} = |\nabla \mathbf{x}| = \sum_{\mathbf{k} \in \mathbb{Z}^n} \sqrt{\left( \frac{\partial \mathbf{x}_k}{\partial x} \right)^2 + \left( \frac{\partial \mathbf{x}_k}{\partial y} \right)^2}$$

⇒ this type of regularization suits very well to X-ray tomography.

## Regularization (3)

- The “dose” of regularization has to be **controlled**. A **compromise** has to be found with data accordance.

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \sum_{\theta \in \Theta} \underbrace{\|\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x}\|_W^2}_{f_{\text{data}}(\mathbf{x})} + \mu \cdot f_{\text{prior}}(\mathbf{x})$$

**1. Accurate and fast numerical projector**  
**2. Regularization term tuning parameter**

Data residuals:  
 Weighted least squares (linear and convex)

## Dealing with the noise (1)

- Another way to address the noise precisely is to **weight the data residuals** with its inverse covariance matrix  $\mathbf{W}$  (particularly if the noise is non-stationary): a **noise model** is needed.

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \sum_{\theta \in \Theta} \|\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x}\|_{\mathbf{W}}^2 + \mu \cdot f_{\text{prior}}(\mathbf{x})$$

**1. Accurate and fast numerical projector**

Data residuals:  
Weighted least squares (linear and convex)

**2. Regularization term tuning parameter**

- **3. The weighting matrix: inverse of the noise covariance**  
 (the noise model can depend function of the data itself)

$$\Rightarrow \sum_{\theta \in \Theta} (\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x})^T \cdot \mathbf{W} \cdot (\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x})$$

## Dealing with the noise (2)

- Tomography data = photon counting  $\Rightarrow$  Poisson statistics (**non uniform**).
- Data precorrections (e.g. diffusion) make measurement statistics not Poisson anymore (Fessler, 1994)  $\Rightarrow$  we consider a non stationary Gaussian noise.

$$\mathbf{W} = \begin{bmatrix} \frac{1}{\sigma_1^2} & & & \\ & \frac{1}{\sigma_2^2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_q^2} \\ & & & & \ddots \end{bmatrix}$$

## The optimization algorithm (1)

- The minimization of the global criterion is performed by a **convex optimization algorithm**.

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \sum_{\theta \in \Theta} \|\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x}\|_W^2 + \mu \cdot f_{\text{prior}}(\mathbf{x})$$

**1. Accurate and fast numerical projector**

**4. Optimizer**

Data residuals:  
Weighted least squares (linear and convex)

**2. Regularization term**  
**tuning parameter**

- **3. The weighting matrix: inverse of the noise covariance**  
 (the noise model can depend function of the data itself)

$$\Rightarrow \sum_{\theta \in \Theta} (\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x})^T \cdot \mathbf{W} \cdot (\mathbf{y}^\theta - \mathbf{R}^\theta \cdot \mathbf{x})$$

## The optimization algorithm (2)

- Optimisation based on the

Newton's method  $\Rightarrow$  minimization based on the second order Taylor expansion of the cost function  $f$ .

$$f(\mathbf{x} + \mathbf{p}) = f(\mathbf{x}) + \mathbf{p}^T \nabla f(\mathbf{x}) + \frac{1}{2} \mathbf{p}^T \cdot \nabla^2 f(\mathbf{x}) \cdot \mathbf{p} + o(\mathbf{p}^2)$$

$\mathbf{p}$  is the descent direction

$$\nabla f(\mathbf{x} + \mathbf{p}) = 0 \Leftrightarrow \mathbf{p} \approx -\frac{\nabla f(\mathbf{x})}{\nabla^2 f(\mathbf{x})} = -\mathbf{H}^{-1} \cdot \nabla f(\mathbf{x})$$

$\mathbf{H} = \nabla^2 f(\mathbf{x})$  is the Hessian of  $f$  in  $\mathbf{x}$ .

- Iterative reconstruction scheme:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{H}^{-1} \cdot \nabla f(\mathbf{x})$$

- Quasi-Newton algorithm: L-BFGS (Nocedal, 1980)  $\Rightarrow$  approximation  $\mathbf{B}_n$  of the Hessian  $\mathbf{H}$  (re-calculated at each iteration).

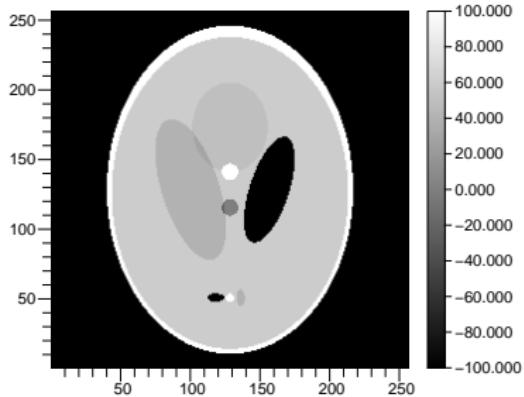
$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{B}_n^{-1} \cdot \nabla f(\mathbf{x})$$

- Convex linear optimizer  $\Rightarrow$  total variation regularization has to be **relaxed**:

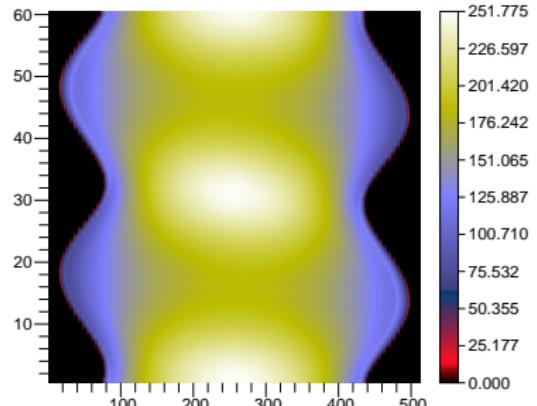
$$f_{\text{prior}}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^n} \sqrt{\left( \frac{\partial \mathbf{x}_k}{\partial x} \right)^2 + \left( \frac{\partial \mathbf{x}_k}{\partial y} \right)^2 + \epsilon^2}$$

## 2D static reconstructions (1)

- 60 simulated projections on a  $2\pi$  range.

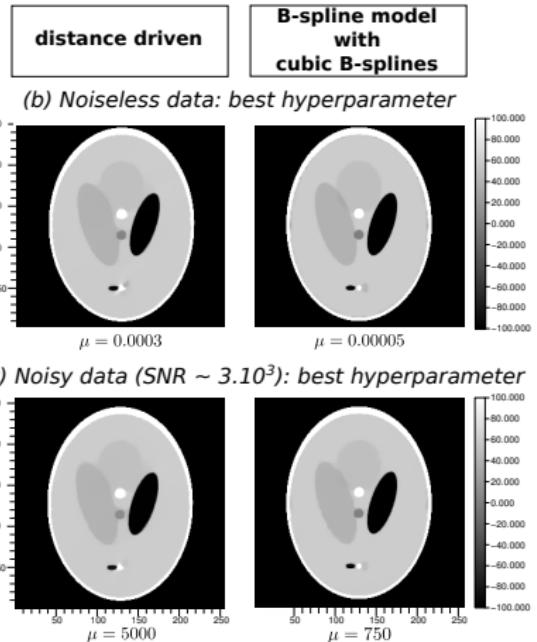
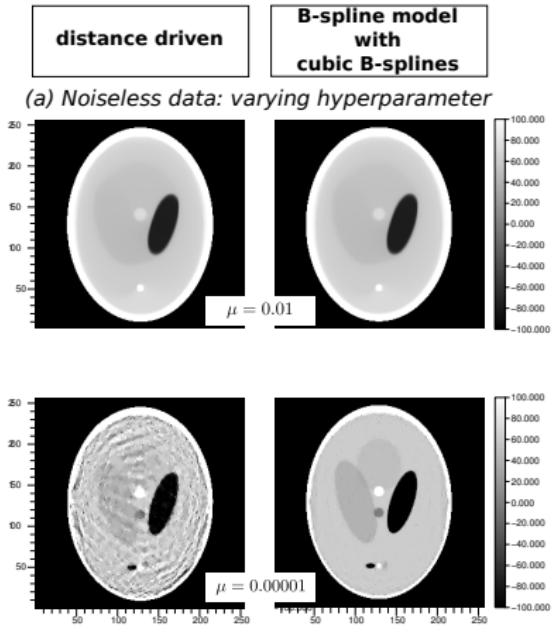


Object

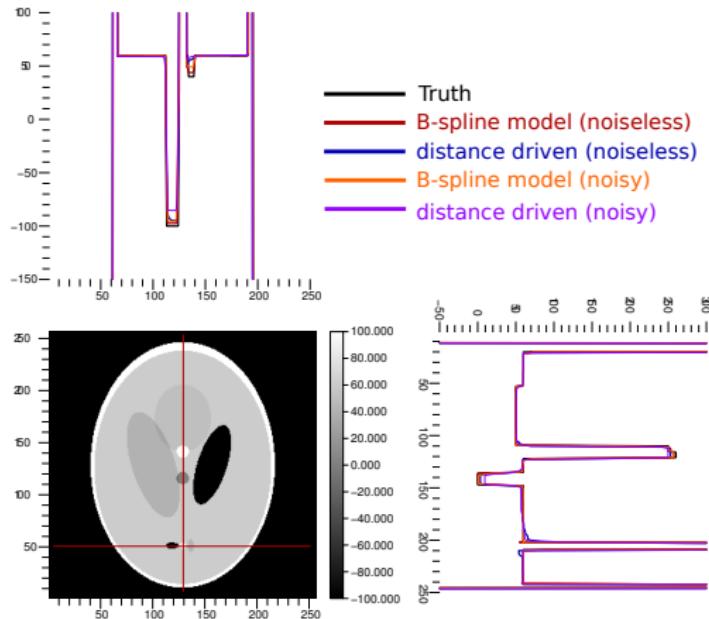


Data

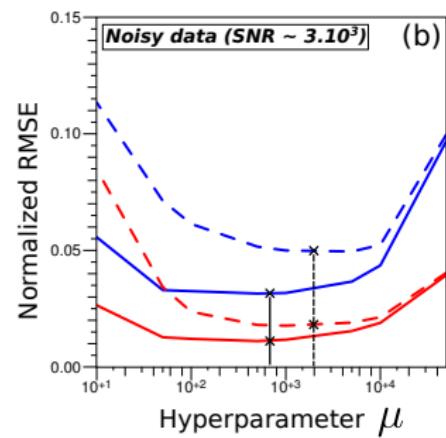
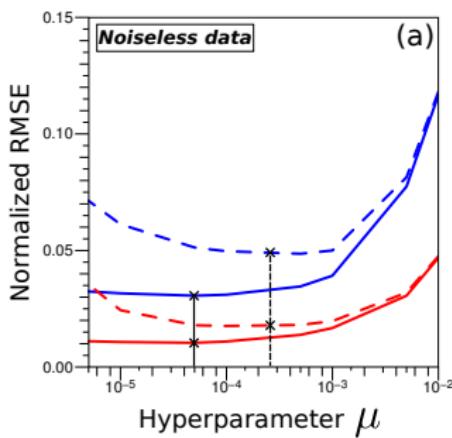
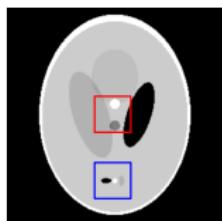
## 2D static reconstructions (2)



## 2D static reconstructions (3)

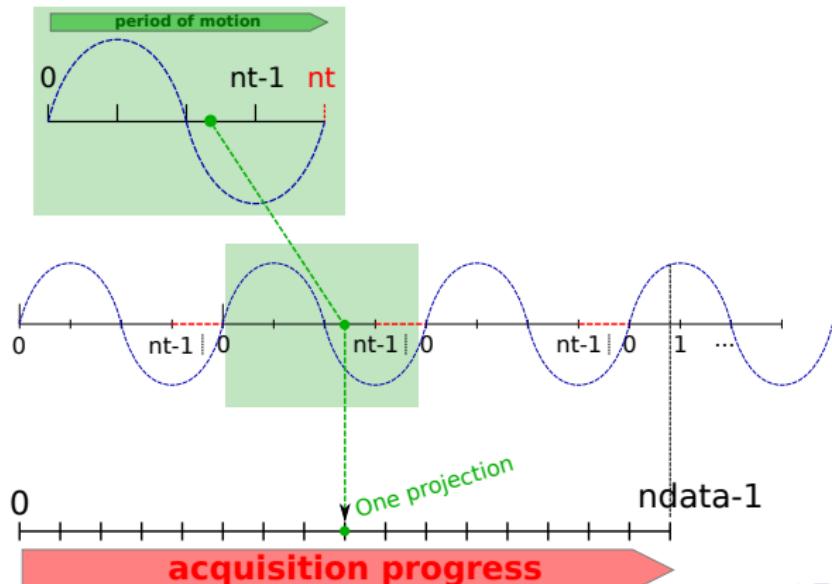


## 2D static reconstructions (4)



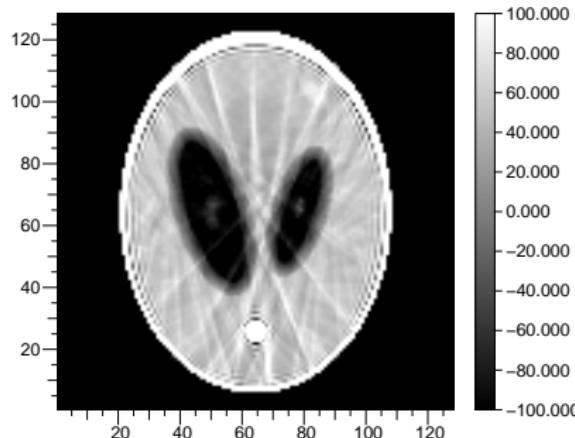
## Application to dynamic X-ray tomography (1)

- The acquired object is not static over time anymore, but animated by a periodic motion.



## Application to dynamic X-ray tomography (3)

- 601 simulated projections on a  $2\pi$  range, during 120 seconds.
- “Static” reconstruction using the whole projections.



## Application to dynamic X-ray tomography (3)

- 601 simulated projections on a  $2\pi$  range, during 120 seconds.
- Reconstruction of 25 temporal frames  $\Rightarrow$  24 “static” projections per frame.
- Spatio-temporal regularization used to take the temporal continuity into account.

using Distance Driven model	using B-spline model

## Bibliography (1)

- B. DeMan and S. Basu. *Distance driven projection and backprojection in three dimensions*. Physics in Medicine and Biology, 2004.
- J.A. Fessler. *Penalized weighted least-squares image reconstruction for positron emission tomography*. IEEE Transactions on Medical Imaging, 1994.
- J.A. Fessler. *Iterative methods for image reconstruction*. IEEE International Symposium on Biomedical Imaging Tutorial, Arlington Virginia, 2006.
- Y. Long, J.A. Fessler and J.M. Balter. *3D Forward and Back-Projection for X-Ray CT Using Separable Footprints*. 10th Intl. Mtg. on Fully 3D Im. Rec. in Rad. and Nuc. Med., 2009.
- S. Horbelt, M. Liebling and M. Unser. *Discretization of the Radon transform and of its inverse by spline convolutions*. IEEE Transactions on Medical Imaging, 2002.
- P.M. Joseph. *An improved algorithm for reprojecting rays through pixel images*. IEEE Transactions on Medical Imaging, 1982.
- M. Slaney and A. Kak. *Principles of computerized tomographic imaging*. SIAM, Philadelphia, 1988.
- R.M. Lewitt. *Multidimensional digital image representations using generalized Kaiser-Bessel window functions*. JOSA A, 1990.
- R.M. Lewitt. *Alternatives to voxels for image representation in iterative reconstruction algorithms*. Physics in Medicine and Biology, 1992.

## Bibliography (2)

- S. Matej and R.M. Lewitt. *Image representation and tomographic reconstruction using spherically symmetric volume elements.* IEEE Transactions on Medical Imaging, 1993.
- S. Matej and R.M. Lewitt. *Efficient 3D grids for image reconstruction using spherically-symmetric volume elements.* IEEE Transactions on Medical Imaging, 1995.
- S. Matej and R.M. Lewitt. *Practical considerations for 3D image reconstruction using spherically symmetric volume elements.* IEEE Transactions on Medical Imaging, 1996.
- J. Nocedal. *Updating quasi-Newton matrices with limited storage.* Mathematics of computation, 1980.
- L.I. Rudin, S. Osher and E. Fatemi. *Nonlinear total variation based noise removal algorithms.* Physica D: Nonlinear Phenomena, 1992.
- P. Thévenaz, T. Blu and M. Unser. *Interpolation Revisited.* IEEE Transactions on Medical Imaging, 2000.
- M. Unser., A. Aldroubi and M. Eden. *B-spline signal processing: Part I-Theory.* IEEE Transactions on Signal Processing, 1993.
- M. Unser., A. Aldroubi and M. Eden. *B-spline signal processing: Part II-Efficient design and applications.* IEEE Transactions on Signal Processing, 1993.

## Bibliographie (3)

- M. Unser. *Splines: A perfect fit for signal and image processing*. IEEE Signal Processing Mag., 1999.
- M. Unser. *Sampling-50 years after Shannon*. Proceedings of the IEEE, 2000.