Kernel-based Image Reconstruction from Radon data

Stefano De Marchi

Department of Mathematics
University of Padova - Italy

December 7th, 2016

\(^1\)Work done in several years with various collaborators
Main references


Part I

The problem and the first approach

Work with A. Iske, A. Sironi
Outline

1. Image Reconstruction from CT

2. Radon transform
   - Back-P and Filtered Back-P

3. Alg. Rec. Tech. (ART), Kernel approach
   - Numerical results
Description of CT
How does it work?

- Non-invasive medical procedure (X-ray equipment).
- X-ray beam is assumed to be:
  - monochromatic;
  - zero-wide;
  - not subject to diffraction or refraction.
- Produce cross-sectional images.
- **Transmission tomography** (emissive tomography, like PET and SPECT, are not considered here)
Description of CT
How does it work?

- $\ell(t, \theta) \rightarrow$ line along which the X-ray is moving;
- $(t, \theta) \in \mathbb{R} \times [0, \pi) \rightarrow$ polar coordinates of line-points;
- $f \rightarrow$ attenuation coefficient of the body;
- $I \rightarrow$ intensity of the X-ray.
X-rays

- Discovered by Wihelm Conrad Röntgen in 1895
- Wavelength in the range $[0.01, 10] \times 10^{-9} \text{ m}$
- Attenuation coefficient:
  \[
  A(x) \approx "\text{# photons absorbed/1 mm}" \\
  A : \Omega \rightarrow [0, \infty)
  \]

Figure: First X-ray image: Frau Röntgen left hand.
Computerized Tomography (CT)

A. Cormack and G. Hounsfield 1970

Reconstruction from X-ray images taken from 160 or more beams at each of 180 directions

Beer’s law (loss of intensity):

\[
\int_{x_0}^{x_1} A(x) \, dx = \ln \left( \frac{l_0}{l_1} \right)
\]

given by CT

Figure: First generation of CT scanner design.
# Outline

1. **Image Reconstruction from CT**

2. **Radon transform**
   - Back-P and Filtered Back-P

3. **Alg. Rec. Tech. (ART), Kernel approach**
   - Numerical results
A line \( l \) in the plane, perpendicular to the unit vector \( \mathbf{n} = (\cos \theta, \sin \theta) \) and passing through \( p = (t \cos \theta, t \sin \theta) = t \mathbf{n} \), can be characterized (by the polar coordinates \( t \in \mathbb{R}, \theta \in [0, \pi) \)), i.e. \( l = l_{t, \theta} \)

\[
l_{t, \theta} = \{ x := (t \cos \theta - s \sin \theta, t \sin \theta + s \cos \theta) = (x_1(s), x_2(s)) \mid s \in \mathbb{R} \} \]

**Figure** : A line in the plane.
The **Radon transform** of a given function $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ is defined for each pair of real number $(t, \theta)$, as line integral

$$Rf(t, \theta) = \int_{l_{t,\theta}} f(x) \, dx = \int_{\mathbb{R}} f(x_1(s), x_2(s)) \, ds$$

**Figure**: Left: image. Right: its Radon transform (**sinogram**)
A CT scan measures the X-ray projections through the object, producing a **sinogram**, which is effectively the **Radon transform** of the attenuation coefficient function $f$ in the $(t, \theta)$-plane.
Radon transform: another example

**Figure**: Shepp-Logan phantom.

**Figure**: Radon transform (sinogram).
First attempt to recover $f$ from $Rf$

The back projection of the function $h \equiv h(t, \theta)$ is the transform

$$Bh(x) = \frac{1}{\pi} \int_0^\pi h(x_1 \cos \theta + x_2 \sin \theta, \theta) \, d\theta$$

i.e. the average of $h$ over the angular variable $\theta$, where $t = x_1 \cos \theta + x_2 \sin \theta = x \cdot n$.

Figure: Back projection of the Radon transform.
Important theorems

Theorem (Central Slice Theorem)

For any suitable function $f$ defined on the plane and all real numbers $r, \theta$

$$F_2 f(r \cos \theta, r \sin \theta) = F(Rf)(r, \theta).$$

($F_2$ and $F$ are the 2-d and 1-d Fourier transforms, resp.).

Theorem (The Filtered Back-Projection Formula)

For a suitable function $f$ defined in the plane

$$f(x) = \frac{1}{2} B\{F^{-1}[|r|F (Rf)(r, \theta)]\}(x), \quad x \in \mathbb{R}^2.$$
Fundamental question

Fundamental question of image reconstruction.

Is it possible to reconstruct a function $f$ starting from its Radon transform $Rf$?

Answer (Radon 1917).

**Yes, we can** if we know the value of the Radon transform for all $r, \theta$. 
Discrete problem

Ideal case
- $Rf(t, \theta)$ known for all $t, \theta$
- Infinite precision
- No noise

Real case
- $Rf(t, \theta)$ known only on a finite set $\{(t_j, \theta_k)\}_{j,k}$
- Finite precision
- Noise in the data
Fourier-based approach

- Sampling: \( Rf(t, \theta) \rightarrow R_Df(jd, k\pi/N) \)
- Discrete transform: e.g.

\[
B_Dh(x) = \frac{1}{N} \sum_{k=0}^{N-1} h(x \cos (k\pi/N) + y \sin (k\pi/N), k\pi/N)
\]

- Filtering (low-pass): \(|r| = F\phi(r), \text{ with } \phi \text{ band-limited function}\)
- Interpolation: \( \{f_k : k \in \mathbb{N}\} \rightarrow If(x), x \in \mathbb{R} \)
Discrete problem

- Back-Projection Formula

\[ f(x) = \frac{1}{2} B \{ F^{-1} [ |r| \cdot F(Rf(r, \theta)) ] \}(x) \]

- Filtering

\[ f(x) = \frac{1}{2} B \{ F^{-1} [ F(\phi(r)) \cdot F(Rf(r, \theta)) ] \}(x) = \]
\[ = \frac{1}{2} B \{ F^{-1} [ F(\phi * Rf(r, \theta)) ] \}(x) \]
\[ = \frac{1}{2} B [ \phi * Rf(r, \theta) ](x) \]

- Sampling and interpolation

\[ f(x_1^m, x_2^n) = \frac{1}{2} B_D l[ \phi * R_D f(r_j, \theta_k)](x_1^m, x_2^n) \]
Discrete problem: an example

**Figure**: Shepp-Logan phantom.

**Figure**: Reconstruction with linear interpolation and $180 \times 101 = 18180$ samples.
Outline

1. Image Reconstruction from CT
2. Radon transform
   - Back-P and Filtered Back-P
3. Alg. Rec. Tech. (ART), Kernel approach
   - Numerical results
Kernel-based Image Reconstruction

Differently from Fourier-based reconstruction, we fix a set $\mathcal{B} = \{ b_i \}_{j=1}^n$ of basis functions.

Example

A square image ($m = K^2$) can be expressed as

$$I(x) = \sum_{i=1}^{m} a_i b_i(x),$$

where

- $a_i$ is the color of the $i$-th pixel,
- $b_i$ the pixel basis, for $i = 1, \ldots, m$ given as

$$b_i(x) = \begin{cases} 
1 & \text{if } x \text{ lies in pixel } i \\
0 & \text{otherwise}
\end{cases}$$
Asking

\[ RL(t_j, \theta_j) = Rf(t_j, \theta_j), \quad j = 1, \ldots, n \]

we obtain the linear system

\[
\sum_{i=1}^{m} a_i Rb_i(t_j, \theta_j) = Rf(t_j, \theta_j), \quad j = 1, \ldots, n
\]

• Large but sparse linear system (usually rectangular)
• Solution by iterative methods (Kaczmarz, MLEM, OSEM, LSCG), or SIRT techniques (see also the Matlab package AIRtools by Hansen & Hansen 2012).
ART reconstruction: Example 1

Figure: Bull's eye phantom.

Figure: $64 \times 64 = 4096$ reconstructed image with 4050 samples by Kaczmarz.
ART reconstruction: Example 2

Figure: Shepp-Logan phantom.

Figure: The phantom reconstructed by MLEM in 50 iterations.
Hermite-Birkhoff (H-B) (generalized) interpolation

Let $\Lambda = \{\lambda_1, \ldots, \lambda_n\}$ be a set of linearly independent linear functionals and $f_\Lambda = (\lambda_1(f), \ldots, \lambda_n(f))^T \in \mathbb{R}^n$.

The solution of a general H-B reconstruction problem:

H-B reconstruction problem

find $g = \sum_{j=1}^n c_j g_j$ such that $g_\Lambda = f_\Lambda$, that is

$$\lambda_k(g) = \lambda_k(f), \quad k = 1, \ldots, n.$$  \(1\)

Being $\lambda_k := R_k f = Rf(t_k, \theta_k), \quad k = 1, \ldots, n$ the conditions (1)

$$\sum_{j=1}^n c_j \lambda_k(g_j) = \lambda_k(f), \quad k = 1, \ldots, n$$  \(2\)

that corresponds to the linear system $A c = b$ as before.
H-B interpolation: basis functions

- **Haar-Maierhuber-Curtis theorem (1904, 1956, 1959):** In the multivariate setting, the well-posedness of the interpolation problem of scattered data is guaranteed if we no longer fix in advance the set of basis functions.

- Thus, the basis $g_j$ should depend on the data

$$g_j(x) = \lambda_j^y(K(x,y)) \ [= R^y[K(x,y)](t_k, \theta_k)] , \quad j = 1, \ldots, n$$

choosing the kernel $K$ such that the matrix

$$A = (\lambda_j^x[\lambda_k^y(K(x,y))]_{j,k}$$

is not singular $\forall (t_k, \theta_k)$. 
Positive definite radial kernels

We consider continuous kernels $K : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ s.t.

- Symmetric $K(x, y) = K(y, x)$
- Radial $K(x, y) = \Phi_\epsilon(\|x - y\|_2)$, $\epsilon > 0$
- Positive definite (PD)

$$
\sum_{k,j=1}^{n} c_j c_k \lambda_j^x \lambda_k^y K(x, y) \geq 0
$$

for all set of linear operators $\lambda_j$ and for all $c \in \mathbb{R}^n \setminus \{0\}$
Positive definite kernels: examples

- **Gaussian**
  \[
  \Phi_\epsilon(\|x\|) = e^{-\epsilon \|x\|^2}, \quad PD \ \forall \ x \in \mathbb{R}^2, \ \epsilon > 0
  \]

- **Inverse multiquadrics**
  \[
  \Phi_\epsilon(\|x\|) = \frac{1}{\sqrt{1 + (\epsilon \|x\|)^2}}, \quad PD \ \forall \ x \in \mathbb{R}^2, \ \epsilon > 0
  \]

- **Askey’s compactly supported (or radial characteristic function)**
  \[
  \Phi_\epsilon(\|x\|) = (1 - \epsilon \|x\|)_+^\beta = \begin{cases} 
  (1 - \epsilon \|x\|)^\beta & \|x\| < 1/\epsilon \\
  0 & \|x\| \geq 1/\epsilon 
  \end{cases}
  \]

  which are PD for any \( \beta > 3/2 \).
Lemma

Let $K(x, y) = \Phi(||x - y||_2)$ with $\Phi \in L^1(\mathbb{R})$. Then for any $x \in \mathbb{R}^2$ the Radon transform $R_y K(x, y)$ at $(t, \theta) \in \mathbb{R} \times [0, \pi)$ can be expressed

$$(R_y K(x, y))(t, \theta) = (R_y K(0, y))(t - x \cdot n, \theta).$$

$\rightarrow$ shift invariant property of the Radon transform.
Problem

- Inverse multiquadric kernel

\[ K(x, y) = \frac{1}{\sqrt{1 + \|x - y\|^2}}. \]

Applying the previous Lemma we have

\[ R_y[K(0, y)](t, \theta) = \int_{\mathbb{R}} \frac{1}{\sqrt{1 + t^2 + s^2}} \, ds = +\infty \]

Hence, the basis \( g_k \) and the matrix \( A \) are not well defined!
Multiplying the kernel $K$ by a **window function** $w$ such that

$$R[K(x, y)w](t, \theta) < \infty \ \forall \ (x, y), (t, \theta).$$

This corresponds to use the linear operator $R_w$ in place of $R$

$$R_w[f](t, \theta) = R[fw](t, \theta).$$

We consider $w$ radial, $w = w(\|\cdot\|_2)$
Example of window functions

- **Characteristic function**
  \[ w(x) = \chi_{[-L,L]}(\|x\|_2), \ L > 0 \]

- **Gaussian**
  \[ w(x) = e^{-\nu^2\|x\|_2^2}, \ \nu > 0 \]

- **Compactly supported (Askey’s family)**
  \[ w(x) = (1 - \nu^2\|x\|_2^2)_+, \ \nu > 0 \]
Example: Gaussian kernel

- Gaussian kernel, shape parameter $\varepsilon$
  
  $$K(x, y) = e^{-\varepsilon^2 \|x - y\|^2_2}, \quad \varepsilon > 0$$

- Basis function
  
  $$g_j(x) = R^y[K(x, y)](t_j, \theta_j) = \frac{\sqrt{\pi}}{\varepsilon} e^{-\varepsilon^2 (t_j - x \cdot v_j)^2}$$

  with $v_j = (\cos \theta_j, \sin \theta_j)$

- Matrix $A = (a_{k,j})$
  
  $$a_{k,j} = R[g_j](t_k, \theta_k) = +\infty, \quad \text{if } \theta_j = \theta_k$$
Example: Gaussian kernel (cont’)

- Considering the Gaussian window function

\[ w(x) = e^{-\nu^2\|x\|^2}, \quad \nu > 0 \]

- The matrix \( A \) becomes

\[ a_{k,j} = R[g_j w](t_k, \theta_k) = \frac{\pi \exp \left[ -\nu^2 \left( t_k^2 + \frac{\varepsilon^2 \beta^2}{\varepsilon^2 \alpha^2 + \nu^2} \right) \right]}{\varepsilon \sqrt{\varepsilon^2 \alpha^2 + \nu^2}} \]

where \( \alpha = \sin (\theta_k - \theta_j) \) and \( \beta = t_j - t_k \cos (\theta_k - \theta_j) \)
Example: Gaussian kernel (cont’)

**Figure:** Crescent-shaped phantom.

**Figure:** $256 \times 256 = 65536$ reconstructed image with $n = 4050$ samples.
A numerical experiment

- Gaussian kernel $\Phi_\epsilon$ and gaussian weight $w_\nu$
- Comparison with the Fourier-based reconstruction (relying on the FBP)
- Reconstructions from scattered Radon data and noisy Radon data
- Root Mean Square Error

$$RMSE = \frac{1}{J} \sqrt{\sum_{i=1}^{J} (f_i - g_i)^2}$$

$J$ is the dimension of the image, $\{f_i\}, \{g_i\}$ the greyscale values at the pixels of the original and the reconstructed image.
Kernel-based vs Fourier based: I

Test phantoms

Figure: crescent shape
Figure: bull’s eye
Figure: Shepp-Logan
Figure: Left: parallel beam geometry, 170 lines (10 angles and 17 Radon lines per angle). Right: scattered Radon lines, 170 lines.
Kernel-based vs Fourier based: II

- Using parallel beam geometry, i.e.
  \[ \theta_k = k\pi/N, \quad k = 0, \ldots, N-1 \] and
  \[ t_j = jd, \quad j = -M, \ldots, M, \]
  with sampling spacing \( d > 0 \rightarrow (2M + 1) \times N \) regular grid of Radon lines. No noise on the data.

- With \( N = 45, \quad M = 40, \quad \epsilon = 60 \) we got

<table>
<thead>
<tr>
<th>Phantom</th>
<th>optimal ( \nu )</th>
<th>kernel-based</th>
<th>Fourier-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>crescent</td>
<td>0.5</td>
<td>0.102</td>
<td>0.120</td>
</tr>
<tr>
<td>bull’s eye</td>
<td>0.4</td>
<td>0.142</td>
<td>0.134</td>
</tr>
<tr>
<td>Shepp-Logan</td>
<td>1.1</td>
<td>0.159</td>
<td>0.177</td>
</tr>
</tbody>
</table>

Table: RMSE of kernel-based vs Fourier-based method
Kernel-based vs Fourier based: III

- Using **scattered Radon data**, with increasing randomly chosen Radon lines $n = 2000, 5000, 10000, 20000$. **No noise on the data.**
- With $\epsilon = 50$ and $\nu = 0.7$ we got

<table>
<thead>
<tr>
<th>Phantom</th>
<th>2000</th>
<th>5000</th>
<th>10000</th>
<th>20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>crescent</td>
<td>0.1516</td>
<td>0.1405</td>
<td>0.1431</td>
<td>0.1174</td>
</tr>
<tr>
<td>bull’s eye</td>
<td>0.1876</td>
<td>0.1721</td>
<td>0.2102</td>
<td>0.1893</td>
</tr>
</tbody>
</table>

**Table:** RMSE of kernel-based vs different number $n$ of Radon lines
Kernel-based vs Fourier based: IV

- These experiments are with noisy Radon data, i.e. we add a gaussian noise of zero mean and variance $\sigma = 1.e^{-3}$ to each of the three phantoms.

- With same parallel beam geometry and same $\epsilon$ and $\nu$

<table>
<thead>
<tr>
<th>Phantom</th>
<th>kernel-based</th>
<th>Fourier-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>crescent</td>
<td>0.1502</td>
<td>0.1933</td>
</tr>
<tr>
<td>bull’s eye</td>
<td>0.1796</td>
<td>0.2322</td>
</tr>
<tr>
<td>Shepp-Logan</td>
<td>0.1716</td>
<td>0.2041</td>
</tr>
</tbody>
</table>

Table: RMSE of kernel-based vs Fourier-based with noisy data

- With scattered Radon data and same $\epsilon$ and $\nu$

<table>
<thead>
<tr>
<th>Phantom</th>
<th>noisy</th>
<th>noisy-free</th>
</tr>
</thead>
<tbody>
<tr>
<td>crescent</td>
<td>0.2876</td>
<td>0.1820</td>
</tr>
<tr>
<td>bull’s eye</td>
<td>0.3140</td>
<td>0.2453</td>
</tr>
</tbody>
</table>

Table: RMSE with noisy and noisy-free data
Window function parameter

- Gaussian kernel; Gaussian window function

\[ K(x, y) = e^{-\varepsilon^2 \|x-y\|^2_2} \]
\[ w(x) = e^{-\nu^2 \|x\|^2_2} \]

Figure: Bull's eye phantom, \( \varepsilon = 30 \).

- Trade-off principle (Schaback 1995)
Kernel shape parameter

- Multiquadric kernel, Gaussian window

\[ K(x, y) = \sqrt{1 + \rho^2 \|x - y\|^2} e^{-\varepsilon^2 \|x - y\|^2} \]

(a) Crescent-shaped phantom        (b) Shepp-Logan phantom

Figure: Optimal values depend on the data.
Comparison with FBP Formula

**Figure**: FBP and Gaussian kernel reconstruction (with optimal parameters $\varepsilon^*, \nu^*$).

**Figure**: Crescent-shaped: (a) FBP; (b) Gaussian kernel.
Comparison with FBP Formula

* RMSE of the same order (ok!)
* More computational time and memory usage (not so well!)

Figure: Computational time.
Part II

Fast implementation of ART

Work with F. Filbir, J. Frikel and M. Narduzzo
Inverse problem of CT
ART-solution

**Algebraic Reconstruction Technique (ART): generality**

- ART determines a solution in the **recovery subspace**

  \[
  S_\gamma := \text{span} \{ \Phi_\epsilon (\cdot - y_j) : 1 \leq j \leq J \} \subseteq N_\Phi(\Omega),
  \]

  where \( \gamma := \{y_1, \ldots, y_J\} \subseteq \Omega \) arbitrary, but **fixed**, set of **reconstruction points** and \( \{ \Phi_\epsilon (\cdot - y_j) \}_{j=1}^J \) translates of the basis function \( \Phi \) (RBFs).

- Search solution for \( R f = \bar{g} \) of the form

  \[
  \tilde{f}_\gamma = \sum_{j=1}^J \alpha_j \Phi_\epsilon (\cdot - y_j) \in S_\gamma,
  \]

  with \( \alpha = (\alpha_1, \ldots, \alpha_J)^T \in \mathbb{R}^J \) to be determined.
Inverse problem of CT
ART-solution

Thanks to the linearity of the Radon transform ...

**ART-problem**

Search a solution $\tilde{\alpha} \in \mathbb{R}^J$ for the linear system

$$A\alpha = \bar{g},$$

where $A \in \mathbb{R}^{KL \times J}$ is the collocation matrix defined as

$$A_{i(k,l),j} := R(\Phi(\cdot - y_j))(t_k, \theta_l)$$

for $1 \leq i(k, l) \leq KL$ and $1 \leq j \leq J$. 
Inverse problem of CT
ART-solution

**Theorem**

Let $A \in \mathbb{R}^{KL \times J}$ and $\bar{g} \in \mathbb{R}^{KL}$. Then,

1. There is at least one solution of the minimization problem

   $\min_{\alpha \in \mathbb{R}^{J}} |A\alpha - \bar{g}|$.

   There exists exactly one solution with minimal Euclidean norm (Moore-Penrose solution $\alpha^+$).

2. The solution comes from the system of normal equations

   $A^T A\alpha = A^T \bar{g}$. 
Inverse problem of CT

ART-solution

\[ M := A^T A \text{ symmetric and positive definite} \]

\[ M \text{ non-singular} \]

\[ \exists! \text{ solution for } M\alpha = \tilde{g} \]

with \( \tilde{g} := A^T \bar{g} \)

\[ M\alpha = \tilde{g} \text{ solve with e.g. CGM} \]
Inverse problem of CT
Conjugate Gradient Method

Theorem

Let $M \in \mathbb{R}^{J \times J}$ be a positive definite matrix. For any $\alpha^{(0)} \in \mathbb{R}^J$, the sequence $\{\alpha^{(k)}\}_{k \in \mathbb{N}}$, generated by the CGM, converges to the minimal norm solution $\alpha^+$ in at most $J$ steps.

The matrix-vector product performed in $O(J \cdot KL)$ floating-points operations.
Fast implementation of ART
Efficient choice of reconstruction points

Fast implementation

Choose a polar reconstruction grid

\[ \gamma_{D_N,\Theta} := \left\{ y_{j(n,\tilde{l})} = r_n(\cos(\theta_{\tilde{l}}), \sin(\theta_{\tilde{l}}))^T : 1 \leq j(n,\tilde{l}) \leq NL \right\}, \]

with

\[ D_N := \left\{ r_n \in [-r, r] : \Delta r > 0 \text{ for } 1 \leq n \leq N \right\} \]

and

\[ \Theta := \left\{ \theta_{\tilde{l}} \in [0, \pi) : \Delta \theta > 0 \text{ for } 1 \leq \tilde{l} \leq L \right\} \]

s.t. \( \Theta \) is the set of angular coordinates of the line-points.
Fast implementation of ART
Efficient choice of reconstruction points

Line-points with radius 1 (asterisks). Polar reconstruction grid in $B_1(0) \subseteq \mathbb{R}^2$ (small circles).
Fast implementation of ART
Efficient storage of the matrix

**BLOCK CIRCULANT STRUCTURE for A**

\[
A = \begin{bmatrix}
A_{11} & \cdots & A_{1N} \\
\vdots & \ddots & \vdots \\
A_{K1} & \cdots & A_{KN}
\end{bmatrix} \in \mathbb{R}^{KL \times NL}
\]

where every block-matrix \( A_{kn} \in \mathbb{R}^{L \times L} \) takes the form

\[
A_{kn} = \begin{bmatrix}
a_1 & a_2 & \cdots & \cdots & a_{L-1} & a_L \\
a_L & a_1 & a_2 & \cdots & \cdots & a_{L-1} \\
\vdots & a_L & a_1 & \cdots & \cdots & \vdots \\
a_3 & \cdots & \cdots & \cdots & \cdots & a_2 \\
a_2 & a_3 & \cdots & \cdots & a_L & a_1
\end{bmatrix} \in \mathbb{R}^{L \times L}
\]
Fast implementation of ART
Fast matrix-vector product using a circulant matrix

**Theorem (Main theorem, Narduzzo master’s thesis)**

Let $C \in \mathbb{R}^{L \times L}$ be a circulant matrix with first column $c \in \mathbb{R}^L$. Further, let $F_L \in \mathbb{C}^{L \times L}$ be the rescaled Fourier matrix $\left(F_{i,j} = \frac{1}{\sqrt{L}} \mu_L^{(i-1)(j-1)} \right)$ with $\mu_L := e^{-\frac{2\pi i}{L}}$. Let $F_L^*$ be its conjugate transpose. Then, it holds

$$C v = F_L^* [(F_L c) \odot (F_L v)] \quad \forall v \in \mathbb{R}^L. \quad (3)$$

Here $\odot$ is the component-wise multiplication operator: $x \odot y = (x_1 y_1, \ldots, x_L y_L)$, $x, y \in \mathbb{R}^L$.

Matrix-vector product now performed in

$$O(NKL \log(L))$$ floating-points operations through the use of **FFT** and **IFFT**.
Choice of the RBF

**RBF:** Gaussian with shape parameter $\epsilon > 0$

$$\Phi_\epsilon(x - y) := e^{-\epsilon^2 \|x-y\|^2} \quad \forall x, y \in \mathbb{R}^2.$$ 

Entries of the collocation matrix:

$$A_{i(k,l), j(n,\tilde{l})} = R\Phi_\epsilon(\cdot - y_{j(n,\tilde{l})})(t_k, \theta_l)$$

$$= \frac{\sqrt{\pi}}{\epsilon} e^{-\epsilon^2( t_k - r_n \cos(\theta_l - \theta_{\tilde{l}}))^2}$$

for $1 \leq i(k,l) \leq KL$ and $1 \leq j(n,\tilde{l}) \leq NL$. 
Fast implementation of ART
Experiment 1: Computational efficiency and accuracy of FCGM

Experiment 1

Shepp-Logan phantom.

$L \times K = 360 \times 569$ (angular $\times$ radial)
$N = 150$ ($\Delta r$ cost);
$\epsilon = 150$;
$iter = 30$;
$R = 400$ (resolution).
Numerical results
Experiment 1: Computational efficiency and accuracy of FCGM

Computational efficiency...

CPU time plot (in sec) as a function of the available Radon data, \( \tilde{R} \)
Numerical results
Experiment 1: Computational efficiency and accuracy of FCGM

\[ R = 400; \ L \times \ K = 360 \times 569. \]
Numerical results
Experiment 1-Equispaced vs fast Leja radii

\[
RMSE = 0.076764 \\
\text{max-err} = 1.045491 \\
CPU \text{ time} = 274.5 \text{ sec}
\]

Error plot (equispaced radii)

\[
RMSE = 0.080900 \\
\text{max-err} = 1.086950 \\
CPU \text{ time} = 284.9 \text{ sec}
\]

Error plot (fast Leja radii)
Numerical results
Experiment 2: Trade-off numerical stability vs accuracy

**Experiment 2**

\[ \kappa(B) := \ell^2 \text{-condition number}; \]
\[ q_\gamma := \text{separation distance} \] between reconstruction points;
\[ h_{\mathcal{X}, \Delta} := \text{fill-distance line-points } \mathcal{X}; \]
\[ q_{\mathcal{X}, \gamma} := \text{separation distance} \] between line-points and reconstruction points.

* From several experiments we obtained \((C > 0, \tau > \frac{d}{2})\)

\[
\left\| \tilde{f}_\gamma - f \right\|_2 \leq C (h_{\mathcal{X}, \Delta} \cdot q_{\mathcal{X}, \gamma})^\tau \left( 1 + \sqrt{\kappa(B)} \right) \| f \|_{N_{\Phi_\epsilon}(\Omega)},
\]

* From theoretical results on RBF we know \((C_d > 0)\) \([\text{cf. Fasshauer’s book}]\)

\[
\kappa(B) = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}},
\]

\[
\lambda_{\text{max}} \leq (NL) \cdot \Phi_\epsilon(0), \text{(Gerschgorin’s theorem)}
\]

\[
\lambda_{\text{min}} \geq C_d \left( \sqrt{2\epsilon} \right)^{-d} e^{-40.71d^2/(q_\gamma \epsilon)^2} q_\gamma^{-d}.
\]
Numerical results

Experiment 2: Trade-off numerical stability vs accuracy (non stationary case)

Smiley phantom;
$L \times K = 144 \times 171$. 

$iter = 30$;
$R = 100$.

$N$ variable ($\Delta r$ cost); $\bar{\epsilon} = 60$

<table>
<thead>
<tr>
<th>$N$ n.rec.circles</th>
<th>$L^2$ error</th>
<th>$\lambda_{\min}$ min.eigenvalue</th>
<th>$\kappa(B)$ cond.number</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>41.081787</td>
<td>1.081460 $\cdot 10^{-19}$</td>
<td>5.032868 $\cdot 10^{20}$</td>
</tr>
<tr>
<td>50</td>
<td>35.247683</td>
<td>3.592425 $\cdot 10^{-30}$</td>
<td>1.948570 $\cdot 10^{31}$</td>
</tr>
<tr>
<td>55</td>
<td>36.214269</td>
<td>1.998062 $\cdot 10^{-30}$</td>
<td>7.207981 $\cdot 10^{31}$</td>
</tr>
<tr>
<td>60</td>
<td>33.359540</td>
<td>1.292076 $\cdot 10^{-31}$</td>
<td>6.739194 $\cdot 10^{32}$</td>
</tr>
<tr>
<td>65</td>
<td>33.087770</td>
<td>4.161096 $\cdot 10^{-32}$</td>
<td>3.469419 $\cdot 10^{33}$</td>
</tr>
<tr>
<td>80</td>
<td>33.476398</td>
<td>9.725620 $\cdot 10^{-33}$</td>
<td>1.276718 $\cdot 10^{34}$</td>
</tr>
</tbody>
</table>
Numerical results

Experiment 2: Trade-off numerical stability vs accuracy (stationary case)

Smiley phantom;
$L \times K = 144 \times 171$.

$\text{iter} = 30$;
$R = 100$.

$\epsilon > 0$ variable; $\tilde{N} = 30$ ($\Delta r \text{ cost}$)

<table>
<thead>
<tr>
<th>$\epsilon$ (shape par.)</th>
<th>$L^2$ error</th>
<th>$\lambda_{min}$ (min. eigenvalue)</th>
<th>$\kappa(B)$ (cond. number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>43.746117</td>
<td>$2.792255 \cdot 10^{-31}$</td>
<td>$3.054914 \cdot 10^{32}$</td>
</tr>
<tr>
<td>35</td>
<td>44.580586</td>
<td>$4.402922 \cdot 10^{-31}$</td>
<td>$1.616392 \cdot 10^{32}$</td>
</tr>
<tr>
<td>40</td>
<td>45.708660</td>
<td>$9.237581 \cdot 10^{-28}$</td>
<td>$6.632798 \cdot 10^{28}$</td>
</tr>
<tr>
<td>45</td>
<td>47.903512</td>
<td>$2.764322 \cdot 10^{-25}$</td>
<td>$1.950951 \cdot 10^{26}$</td>
</tr>
<tr>
<td>50</td>
<td>52.583766</td>
<td>$2.059636 \cdot 10^{-24}$</td>
<td>$2.341632 \cdot 10^{25}$</td>
</tr>
</tbody>
</table>
Numerical results
Experiment 3: Choice of reconstruction parameters

Smiley phantom;
$L \times K = 360 \times 811$. 

$R = 600$;
$\epsilon = ?$;
$N = ?$;
$\text{iter} = ?$. 
Numerical results
Experiment 3: Choice of best shape parameter

1) Best shape parameter... \( \text{iter} = 30 \)

\[ N = 75 \ (\Delta r \ 	ext{cost}) \]

\[ N = 150 \ (\Delta r \ 	ext{cost}) \]

RMSE as a function of \( \epsilon \) for equispaced reconstruction radii \( N = 75 \) (left) and \( N = 150 \) (right).
Numerical results
Experiment 3: Choice of best shape parameter

...Heuristic rule.. \((\Delta r := \frac{1}{N})\)

\[ \epsilon \approx -9828 \cdot \Delta r + 216 \]
2) **Best number of reconstruction circles...**

\[ \text{iter} = 20; \Delta r \text{ cost} \]

<table>
<thead>
<tr>
<th>(N) n. rec.circles</th>
<th>(N\cdot L) n. rec.points</th>
<th>(\epsilon) shape param.</th>
<th>RMSE</th>
<th>maximal error</th>
<th>CPU time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>18 000</td>
<td>19.440</td>
<td>0.408572</td>
<td>4.053001</td>
<td>101.7</td>
</tr>
<tr>
<td>75</td>
<td>27 000</td>
<td>84.960</td>
<td>0.317973</td>
<td>4.409125</td>
<td>134.0</td>
</tr>
<tr>
<td>100</td>
<td>36 000</td>
<td>117.720</td>
<td>0.300916</td>
<td>4.237258</td>
<td>181.8</td>
</tr>
<tr>
<td>125</td>
<td>45 000</td>
<td>137.376</td>
<td>0.295240</td>
<td>3.958031</td>
<td>237.7</td>
</tr>
<tr>
<td>150</td>
<td>54 000</td>
<td>150.480</td>
<td>0.297035</td>
<td>4.309017</td>
<td>309.1</td>
</tr>
<tr>
<td>175</td>
<td>63 000</td>
<td>159.840</td>
<td>0.296148</td>
<td>3.543423</td>
<td>384.0</td>
</tr>
<tr>
<td>200</td>
<td>72 000</td>
<td>166.860</td>
<td>0.295511</td>
<td>3.624649</td>
<td>486.0</td>
</tr>
<tr>
<td>225</td>
<td>81 000</td>
<td>172.320</td>
<td>0.295141</td>
<td>3.512590</td>
<td>594.3</td>
</tr>
<tr>
<td>250</td>
<td>90 000</td>
<td>176.688</td>
<td>0.296305</td>
<td>3.638046</td>
<td>668.8</td>
</tr>
</tbody>
</table>

RMSE and CPU time for increasing larger sets of reconstruction points. 

**For \(R = 600\), from \(N = 150\) to \(N = 200\)**
3) Different iteration steps...

\[ N = 200 (\Delta r \ cost); \quad \epsilon = 160 \]
Numerical results
Experiment 3: Choice of number of iterations

$$\text{RMSE} = 0.291745 \quad \text{max-err} = 3.614396$$
$$\text{CPU time} = 527.8 \text{ sec}$$

Error plot for $\text{iter} = 28$. 
...Stopping condition...

\[ N = 200 \ (\Delta r \ cost); \quad \epsilon = 160 \]

Convergence history of residual and RMSE respect to the number of iterations.

\[ tol = 10^{-3} \] for the residual decrease
Numerical results
Experiment 4: Noisy data

**Experiment 4**

10% white Gaussian noise

Shepp-Logan phantom;
$L \times K = 360 \times 711$.

$N = 150 \ (\Delta r \ cost);$  
$\epsilon = 150.48;$  
$iter = 25;$  
$R = 500.$
Numerical results
Experiment 4: Noisy data

\[ \text{RMSE} = 0.101538 \]
\[ \text{residual-decrease} = 10^{-2} \]
\[ \text{max-err} = 1.310516 \]
\[ \text{CPU time} = 298.3 \text{ sec} \]
Numerical results

Experiment 5: Real experimental data

human pelvis;
\( L \times K = 2304 \times 736 \).

\( N = 200 \ (\Delta r \ cost); \)
\( \epsilon = 166.86; \)
\( tol = 10^{-3}; \)
\( R = 520. \)

Data provided by Department of Diagnostic and Interventional Radiology at TUM
Numerical results
Experiment 5: Real experimental data

\[ NRMSE(A_{\alpha}^{(k)}, \bar{g}) = 0.034964 \]
\[ \text{iter} = 22 \]
\[ \text{residual-decrease} = 10^{-3} \]
\[ \text{CPU time} = 14027.2 \text{ sec} \]
Conclusions...

- BLOCK CIRCULANT STRUCTURE
- FAST MATRIX-VECTOR PRODUCT
- FAST CGM
- FAST ART

MORE COMPUTATIONAL EFFICIENCY!
SAME IMAGE RECONSTRUCTION CAPABILITY!
Thank you for your attention!