Computerised Tomography

Mathematical Foundations of Signal Processing

Dr. Matthieu Simeoni

November 23, 2020

EPFL

Table of contents

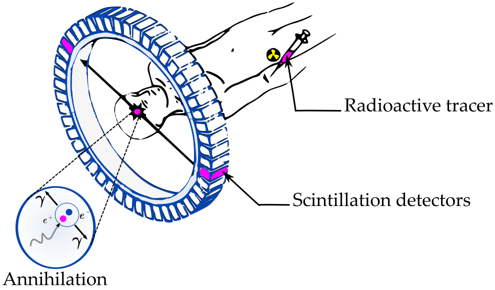
Background Concepts

Positron Emission Tomography (PET)

Definition: (Positron Emission Tomography)

Positron Emission Tomography (PET) is a medical diagnostic technique that enables a physician to study blood flow in and metabolic activity of an organ in a visual way.

- A biochemical metabolite labeled with a positron emitting radioactive material is introduced into the organ.
- The biochemical (typically sugar for the brain) concentrates in regions of high metabolic activity.
- Positron emissions occur randomly are counted by a PET scanner (ring of scintillation detectors).

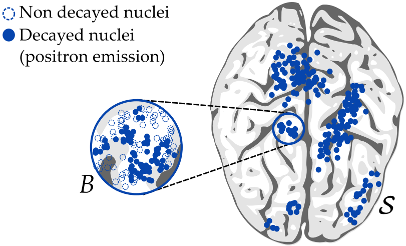


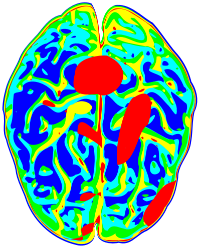
Positron Emissions as Poisson Process

- Positron emissions occur randomly in $\mathscr{S}$ according to a Poisson process.
- Rate of occurrence is characterised by an intensity function $\lambda : \mathscr{S} \to \mathbb{R}$.
- For a given region $B \subset \mathscr{S}$ of the organ, the expected number of positron emissions $N(B)$ is given by

$$\mathbb{E}[N(B)] = \int_B \lambda(\boldsymbol{x})\, d\boldsymbol{x}.$$

- The intensity function is assumed to be proportional to the metabolic activity of interest.



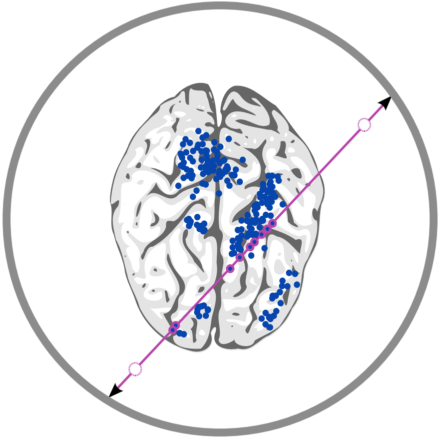


Indirectly Observed Poisson Process

- In practice, we cannot directly observe positron emissions.
- Instead, we observe gamma rays induced by annihilation with neighbouring electrons.
- Coincident gamma rays are recorded on a detector ring.
- We speak of an indirectly observed Poisson process. The number n_d of gamma rays recorded by each detector pair is Poisson distributed, with rate

$$n_d \sim \mathscr{P}(\check{\lambda}_d), \quad \check{\lambda}_d = \mathbb{E}[N(L_d)] = \int_{L_d} \lambda(\boldsymbol{x})\, d\boldsymbol{x}, \quad d = 1, \ldots, D,$$

  where D denotes the total number of detector pairs on the detector ring. $\check{\lambda}_d$ corresponds to the line integral of λ along the line L_d linking the d-th pair of detectors on the detector ring.

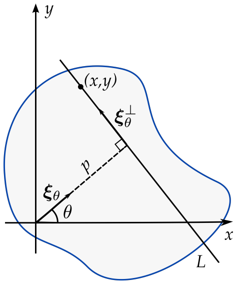


The Radon Transform

- Line integrals can be seen as samples of the so-called Radon transform.
- The chord linking two points on a ring can be parametrised as:

$$L = \{\boldsymbol{x} \in \mathbb{R}^2 : \langle \boldsymbol{x}, \boldsymbol{\xi}_\theta \rangle = p\},$$

  where $p \in \mathbb{R}$, $\theta \in [0, \pi)$ and $\boldsymbol{\xi}_\theta = [\cos(\theta), \sin(\theta)] \in \mathbb{S}^1$.
- The Radon transform maps a function λ onto its line integrals:



Definition (Radon Transform)

The Radon transform $\check{\lambda} : [0, \pi[\times \mathbb{R} \to \mathbb{R}$ of a function $\lambda : \mathbb{R}^2 \to \mathbb{R} \in \mathscr{L}^2(\mathbb{R}^2)$ is

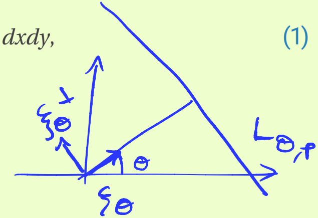
$$\check{\lambda}(\theta, p) = (\mathscr{R}\lambda)(\theta, p) := \int_{\mathbb{R}^2} \lambda(\boldsymbol{x}) \delta(p - \langle \boldsymbol{x}, \boldsymbol{\xi}_\theta \rangle) d\boldsymbol{x}.$$

Example: Radon Transform of a Gaussian

Example: Radon Transform of a Gaussian

Let $f(x,y) = e^{-\pi(x^2+y^2)}, \forall (x,y) \in \mathbb{R}^2$. Then,

$$\check{f}(\theta,p) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\pi(x^2+y^2)}\delta\left(p-\cos(\theta)x-\sin(\theta)y\right)dxdy, \tag{1}$$



We perform the following orthogonal transformation:

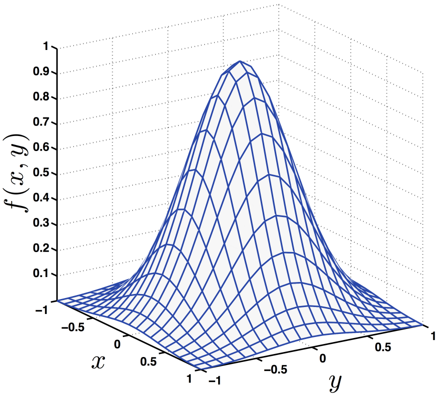
$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

The transformation is orthogonal, and thus $u^2+v^2 = x^2+y^2$. Moreover, in this new basis, the equation of the line $L = \{(x,y)\in\mathbb{R}^2 : \cos(\theta)x+\sin(\theta)y = p\}$ becomes simply: $L = \{(u,v)\in\mathbb{R}^2 : u = p\}$. Eq. (1) hence becomes:

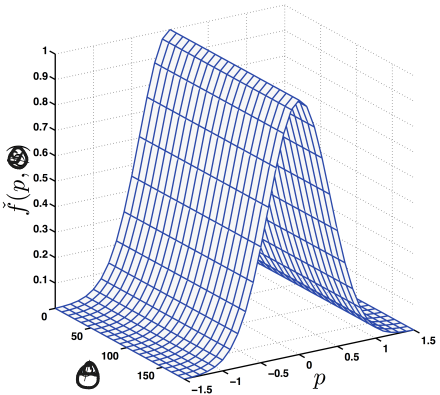
$$\check{f}(\theta,p) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\pi(u^2+v^2)}\delta(p-u)\,dudv = e^{-\pi p^2}\int_{-\infty}^{+\infty} e^{-\pi v^2}dv = e^{-\pi p^2},$$

where we have used the well known result $\int_{\infty}^{\infty} e^{-t^2}dt = \sqrt{\pi}$. Hence, we have: $\mathscr{R}\{e^{-\pi(x^2+y^2)}\} = e^{-\pi p^2}$.

Example: Radon Transform of a Gaussian



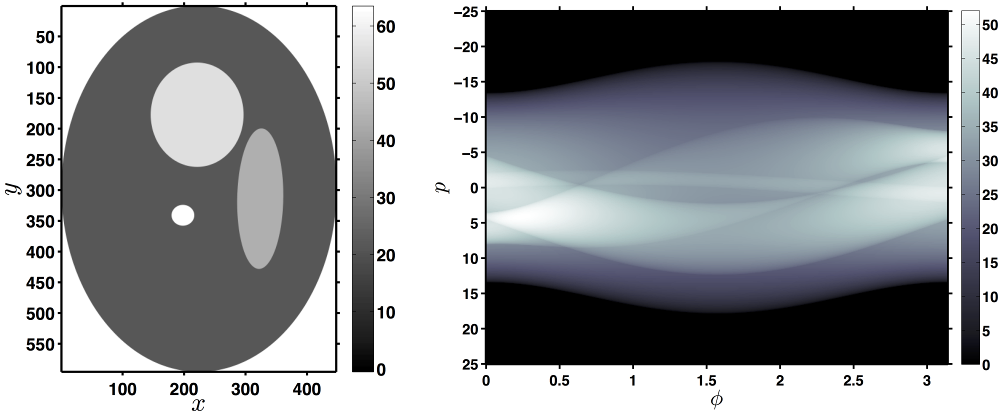
(a) The Gaussian distribution.



(b) Radon transform of the Gaussian distribution.

Figure: The Gaussian distribution and its Radon transform.

Example: Radon Transform of Ellispes [1, Example 2.2]



(a) Sum of ellipses' characteristic functions with heights given by the color of each ellipse.

(b) Radon transform.

Figure: The Radon transform of a sum of ellipses' characteristic functions.

Basic Properties of the Radon Transform

Proposition: (Properties of the Radon Transform)

- **Linearity:** Let f and g be two functions and $\alpha, \beta \in \mathbb{R}$. Then,

$$\mathscr{R}\{\alpha f + \beta g\} = \alpha \mathscr{R} f + \beta \mathscr{R} g.$$

- **Shifting Property:** Let $g(\boldsymbol{x}) = f(\boldsymbol{x} - \boldsymbol{a})$ for some $\boldsymbol{a} \in \mathbb{R}^2$. Then we have:

$$\mathscr{R} g(\theta, p) = \mathscr{R} f\left(\theta, p - \langle \boldsymbol{\xi}_\theta, \boldsymbol{a} \rangle\right), \qquad \forall (\theta, p) \in [0, \pi) \times \mathbb{R}.$$

- **Scaling Property:** Let $g(\boldsymbol{x}) = f(\alpha \boldsymbol{x})$ for some $\alpha \neq 0$. Then we have:

$$\mathscr{R} g(\theta, p) = \frac{1}{\alpha^2} \mathscr{R} f\left(\theta, \alpha p\right), \qquad \forall (\theta, p) \in [0, \pi) \times \mathbb{R}.$$

Proof:

- **Linearity:** $\mathscr{R}\{\alpha f + \beta g\} = \int \left(\alpha f(\mathbf{x}) + \beta g(\mathbf{x})\right) \delta(p - \boldsymbol{\xi} \cdot \mathbf{x})\, d\mathbf{x} = \alpha \check{f} + \beta \check{g}$.
- **Shifting Property:** $\mathscr{R}\{f(\mathbf{x} - \mathbf{a})\} = \int f(\mathbf{x} - \mathbf{a}) \delta(p - \boldsymbol{\xi} \cdot \mathbf{x})\, d\mathbf{x} = \int f(\mathbf{y}) \delta(p - \boldsymbol{\xi} \cdot \mathbf{a} - \boldsymbol{\xi} \cdot \mathbf{x})$.
- **Scaling Property:** $\mathscr{R}\{f(\alpha \mathbf{x})\} = \int f(\alpha \mathbf{x}) \delta(p - \boldsymbol{\xi} \cdot \mathbf{x})\, d\mathbf{x} = \frac{1}{\alpha} \int f(\mathbf{y}) \delta(p - \frac{1}{\alpha} \boldsymbol{\xi} \cdot \mathbf{y}) = \frac{1}{\alpha^2} \int f(\mathbf{y}) \delta(\alpha p - \boldsymbol{\xi} \cdot \mathbf{y})$.

Link with 2D Fourier Transform

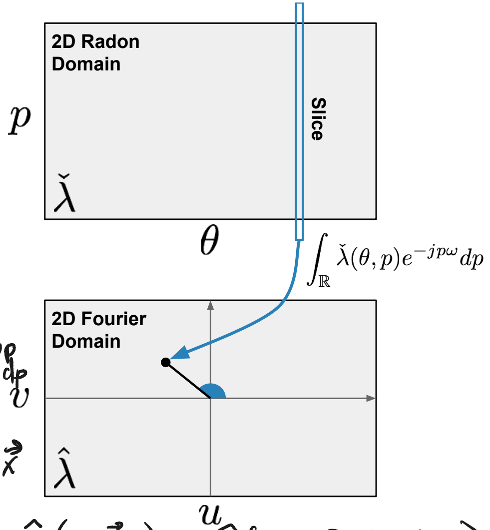
The Radon and Fourier transforms are linked by the projection-slice theorem:

Lemma: (Projection-Slice Theorem)

For any $\theta \in [0,\pi)$ we have:

$$\int_{\mathbb{R}} \mathscr{R}\lambda(\theta,p)e^{-j\omega p}dp = \hat{\lambda}(\omega\cos\theta, \omega\sin\theta),$$

where $\hat{\lambda}:\mathbb{R}^2 \to \mathbb{C}$ is the 2D Fourier transform of λ.



Proof :

$$\int_{\mathbb{R}} \check{\lambda}(\theta,p)\, e^{-j\omega p}dp = \int_{\mathbb{R}}\left(\int_{\mathbb{R}^2} \lambda(\vec{x})\,\delta(p-\langle \vec{x},\vec{\xi}_\theta\rangle)\,d\vec{x}\right)e^{-j\omega p}dp$$

$$= \int_{\mathbb{R}^2} \lambda(\vec{x})\left(\int_{\mathbb{R}} e^{-j\omega p}\,\delta(p-\langle \vec{x},\vec{\xi}_\theta\rangle)\,dp\right)d\vec{x}$$

$$= \int_{\mathbb{R}^2} \lambda(\vec{x})\, e^{-j\omega\langle \vec{x},\vec{\xi}_\theta\rangle}\,d\vec{x} = \hat{\lambda}(\omega\,\vec{\xi}_\theta) = \hat{\lambda}(\omega\cos\theta, \omega\sin\theta)$$

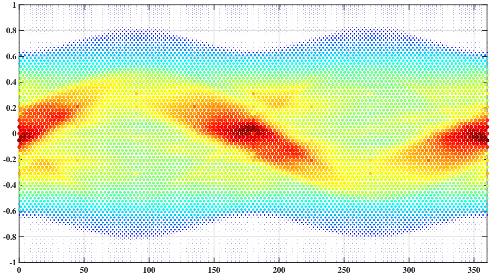
Back to PET

- A PET scanner samples the Radon transform of the metabolic activity.
- Samples are polluted by Poisson noise, resulting in the data:

$$n(\theta_d, p_d) \sim \mathscr{P}\left(\check{\lambda}(\theta_d, p_d)\right), \qquad d = 1, \ldots, D.$$

- It is customary to represent the data in the $(\theta, p) \in [0, \pi) \times \mathbb{R}$ plane, yielding a sinogram.

(a) Metabolic activity



(b) PET data as a Sinogram

Complications with Real-life Scanners

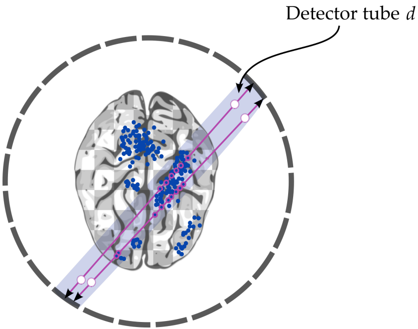
$$\chi_d(p) := \begin{cases} 1 & \text{if } |p_d - p| \leq \varepsilon_d \\ 0 & \text{otherwise} \end{cases}$$

- In practice, the detectors have a certain *width*. We have hence detector tubes instead of lines,

$$\begin{aligned}\mathcal{T}_d &= \left\{\boldsymbol{x} \in \mathbb{R}^2 : |\langle \boldsymbol{x}, \boldsymbol{\xi}_d \rangle - p_d| \leq \epsilon_d\right\}\\ &= \left\{\boldsymbol{x} \in \mathbb{R}^2 : \underline{\chi_d}\left(p_d - \langle \boldsymbol{x}, \boldsymbol{\xi}_{\theta_d} \rangle\right) = 1\right\}.\end{aligned}$$

- Actual data hence consist in tube integrals and not line integrals:

$$\check{\lambda}(\theta_d, p_d) = \mathbb{E}[N(\mathcal{T}_d)] = \int_{\mathbb{R}^2} \lambda(\boldsymbol{x}) \chi_d\left(p_d - \langle \boldsymbol{x}, \boldsymbol{\xi}_{\theta_d} \rangle\right) d\boldsymbol{x}.$$



- Formulating the data model in terms of the Radon transform is hence only valid in the limit for infinitely thin detectors.

Radon-based Tomographic Reconstruction

Inverting the Radon Transform

The Radon transform is invertible. Inversion formula in 2D is given by the filtered back-projection (FBP) formula [2]:

Theorem: (Filtered Back-Projection)

Let $\lambda : \mathbb{R}^2 \to \mathbb{R}$ be sufficiently smooth. Then we have

$$\lambda(\boldsymbol{x}) = \frac{1}{(2\pi)^2} \int_0^{\pi} \left[\check{\lambda}(\theta, \cdot) * h\right] (\langle x, \boldsymbol{\xi}_\theta \rangle) d\theta,$$

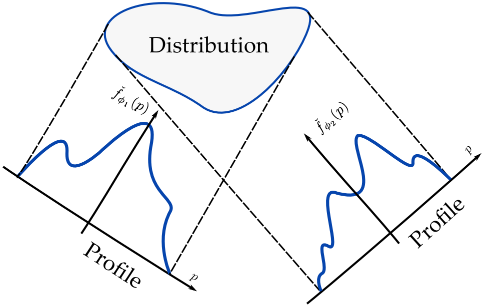
where $h : \mathbb{R} \to \mathbb{R}$ −called the Ramp filter− is defined in terms of its Fourier transform $\hat{h}(\omega) := |\omega|, \forall \omega \in \mathbb{R}$.

$$(f * h)(p) \xleftrightarrow[FT]{} |\omega| \, \hat{f}(\omega)$$

(multiplicat° theorem)

Inverting the Radon Transform

- Holographic result: Reconstruct 2D object from many 1D projections (profiles).
- The Ramp filter, with Fourier transform $\hat{h}(\omega) = |\omega|$, is a *roughening filter*, acting as a derivative.[a] It makes objects more singular, with sharper edges.
- The function λ must hence be sufficiently smooth for the inversion formula to be well-defined!



[a] Recall that in the Fourier domain differentiating accounts to multiplying by $j\omega$: $\mathscr{F}\{g\}(\omega) = j\omega\hat{g}(\omega)$.

Interpretation of FBP

- The adjoint of the Radon transform $\mathscr{R}$ is given, for all $\mu : [0,\pi) \times \mathbb{R} \to \mathbb{R}$, by

$$\mathscr{R}^* \mu = \int_0^\pi \mu(\theta, \langle \boldsymbol{x}, \boldsymbol{\xi}_\theta \rangle)\, d\theta.$$

  Indeed,

$$\begin{aligned}\langle \mathscr{R}\lambda, \mu \rangle &= \int_0^\pi \int_{\mathbb{R}} \left[\int_{\mathbb{R}^2} \lambda(\boldsymbol{x}) \delta(p - \langle \boldsymbol{x}, \boldsymbol{\xi}_\theta \rangle)\, d\boldsymbol{x} \right] \mu(\theta, p)\, dp\, d\theta \\ &= \int_{\mathbb{R}^2} \lambda(\boldsymbol{x}) \left[\int_0^\pi \mu(\theta, \langle \boldsymbol{x}, \boldsymbol{\xi}_\theta \rangle)\, d\theta \right] d\boldsymbol{x} = \langle \lambda, \mathscr{R}^* \mu \rangle.\end{aligned}$$

- $\mathscr{R}^*$ performs a back-projection: all sinogram points to which $\boldsymbol{x}$ contributed to are summed together.
- The inverse Radon transform decomposes as:

$$\mathscr{R}^{-1} = \frac{1}{(2\pi)^2} \mathscr{R}^* \tilde{\lambda}, \quad \text{where} \quad \tilde{\lambda}(\theta, p) = \int_{\mathbb{R}} \check{\lambda}(\theta, p - t) h(t)\, dt = \frac{1}{2\pi} \int_{\mathbb{R}} |\omega| \hat{\lambda}(\omega \boldsymbol{\xi}_\theta) e^{j\omega p}\, d\omega,$$

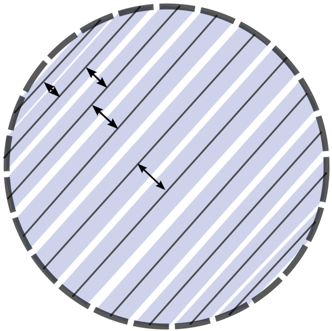
  where the last equality results from the projection-slice and convolution/multiplication theorems.
- In summary: first apply 1D filter h to slices $\check{\lambda}(\theta, \cdot)$ of the Radon transform, then back-project with $\mathscr{R}^*$. Hence the name: filtered back-projection.

Filtered Backprojection in Practice

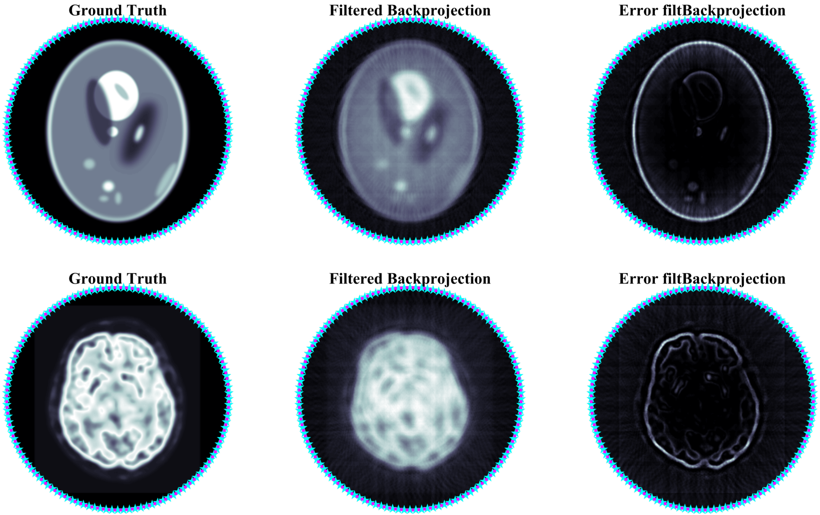
- Tubes are assumed *narrow* and approximated by lines with coordinates (θ_d, p_d).
- Samples are typically non-uniform and noisy.
- Discrete FBP for uniform samples in the p-direction is given by

$$\lambda_{FBP}(\boldsymbol{x}) = \frac{\Delta_\theta}{(2\pi)^2} \sum_{n=1}^{N_\theta} \left(\boldsymbol{\lambda}[\theta_n, \cdot] \circledast \boldsymbol{h}\right)\left[\lfloor \langle \boldsymbol{x}, \boldsymbol{\xi}_{\theta_n} \rangle / \Delta_p \rfloor\right], \quad \boldsymbol{x} \in \mathbb{R}^2.$$

  - Data must be gridded.
  - Convolution is approximated by a circular discrete convolution, efficiently implemented via FFT/iFFT.
  - The routine `skimage.transform.iradon` implements the discrete FBP in Python.



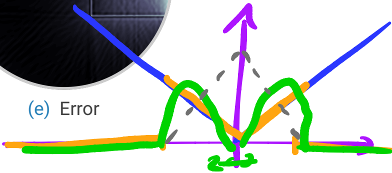
Example


Ground Truth
Filtered Backprojection
Error filtBackprojection
Ground Truth
Filtered Backprojection
Error filtBackprojection

Issues with FBP

- Gridding is expensive and ad-hoc.
- Discrete formula makes a lot of approximations (circular convolution, interpolation).
- Border effects due to circular convolution.
- Ramp filter boosts high frequencies, generally polluted by noise. FBP is unstable and must be regularised by truncating or more generally windowing the Ramp filter (optimal window?).

(c) Metabolic activity (d) FBP Estimate (e) Error



Sampling & Interpolation in PET

Sampling Operator

- Define $\mathscr{H} = \mathscr{L}^2(\mathbb{B}_1)$ with $\mathbb{B}_1$ the unit disk (the brain space).
- We can link the measurements $(n_1, \ldots, n_D)$ to the Poisson process of interest N through a sampling operator $\Phi^* : \mathscr{H} \to \mathbb{R}^D$:

$$\begin{bmatrix} n_1 \\ \vdots \\ n_D \end{bmatrix} = \Phi^* N = \begin{bmatrix} \langle N, \phi_1 \rangle_{\mathscr{H}} \\ \vdots \\ \langle N, \phi_D \rangle_{\mathscr{H}} \end{bmatrix},$$

where ϕ_d are the indicator functions of the detector tubes:

$$\phi_d(x) = \chi_d(x) = \begin{cases} 1 & \text{if } x \in \mathscr{T}_d, \\ 0 & \text{otherwise.} \end{cases}.$$

Sampling Operator

- Since Φ^* is linear, on expectation we have

$$\begin{bmatrix} \lambda_1^* \\ \vdots \\ \lambda_D^* \end{bmatrix} = \Phi^* \lambda = \begin{bmatrix} \langle \lambda, \phi_1 \rangle_{\mathcal{H}} \\ \vdots \\ \langle \lambda, \phi_D \rangle_{\mathcal{H}} \end{bmatrix}.$$

- Measurements give us *evidence about the components* of λ in $\mathcal{R}(\Phi) = \text{span}\{\phi_1, \ldots, \phi_D\}$.
- Goal: to solve this inverse problem and find an *estimate* λ.
- This is an ill-posed problem: any component in $\mathcal{N}(\Phi^*)$ is unaccessible to us

$$\Phi^* \lambda = \Phi^* \lambda_1 + \underbrace{\Phi^* \lambda_2}_{=0} = \Phi^* \lambda_1,$$

where $\lambda = \lambda_1 + \lambda_2 \in \mathcal{R}(\Phi) \oplus \mathcal{N}(\Phi^*) = \mathcal{H}$.

Least-Squares Estimate

- An estimate of λ can be obtained by solving

$$\lambda^\star \in \underset{\lambda \in \mathcal{H}}{\arg\min} \|\boldsymbol{n} - \Phi^* \lambda\|_2^2$$

  - No unique solution!
  - Impose minimal $\mathscr{L}_2$ norm and $\lambda \in \mathscr{R}(\Phi)$ for uniqueness.
  - Leads to the generalised Moore-Penrose pseudo-inverse solution:

$$\lambda^\star = (\Phi^*)^\dagger \boldsymbol{n}.$$

Ideally-Matched Interpolation

- Generalised pseudoinverse is given by

$$(\Phi^*)^\dagger = \Phi(\Phi^*\Phi)^{-1},$$

and the least-squares estimate is hence given by

$$\lambda^\star(x) = (\Phi(\Phi^*\Phi)^{-1}\boldsymbol{n})(x) = \sum_{d=1}^{D} \tilde{\boldsymbol{n}}\phi_d(x),$$

where $\tilde{\boldsymbol{n}} = (\Phi^*\Phi)^{-1}\boldsymbol{n} \in \mathbb{R}^D$.

- Recovery in two steps: apply *Gram correction* to the data, and interpolate using the synthesis operator Φ, adjoint of Φ^*

$$\Phi : \begin{cases} \mathbb{R}^D \to \mathscr{R}(\Phi), \\ \boldsymbol{y} \mapsto (\Phi\boldsymbol{y})(x) = \sum_{d=1}^D y_d\phi_d(\boldsymbol{x}). \end{cases}$$

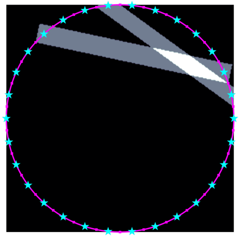
- We have consistency $\Phi^*(\Phi^*)^\dagger = \boldsymbol{I}_D$ and hence $(\Phi^*)^\dagger\Phi^*$ is an orthogonal projection.

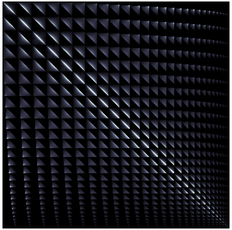
About the Gram Matrix

- The quantity $\Phi^*\Phi \in \mathbb{R}^{D\times D}$ is the Gram matrix.
- An element of the Gram matrix is given by

$$(\Phi^*\Phi)_{ij} = \langle \phi_i, \phi_j \rangle.$$

- Need to compute *areas of parallelograms*! Can be efficiently computed analytically.
- Dense matrix! Basis elements are not localised...
- Dense = ill-conditioned (often)





Regularisation

Eigenfunctions of the Integral Operator

- We have
$$\Phi(\Phi^*\Phi)^{-1}\Phi^*\Phi\boldsymbol{\alpha} = \Phi\boldsymbol{\alpha}.$$
- Hence any element of $\mathscr{R}(\Phi)$ is an eigenfunction with eigenvalue 1. To get orthogonal eigenspaces, we need to find $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_D\} \subset \mathbb{R}^D$ s.t.
$$\langle \Phi\boldsymbol{\alpha}_i, \Phi\boldsymbol{\alpha}_j \rangle = \boldsymbol{\alpha}_j^T(\Phi^*\Phi)\boldsymbol{\alpha}_i = 0.$$
- Choosing $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_D\}$ eigenvectors of $\Phi^*\Phi$ yields the spectral decomposition:
$$\Phi(\Phi^*\Phi)^{-1}\Phi^* = \sum_{d=1}^{D} \frac{1}{\eta_d}(\Phi\boldsymbol{\alpha}_d)(\Phi\boldsymbol{\alpha}_d)^*,$$

  where $\eta_d = \|\Phi\boldsymbol{\alpha}_d\|_2^2 = \boldsymbol{\alpha}_d^T(\Phi^*\Phi)\boldsymbol{\alpha}_d$.

Noisy Measurements

- We can hence re-write the pseudo-inverse estimate as

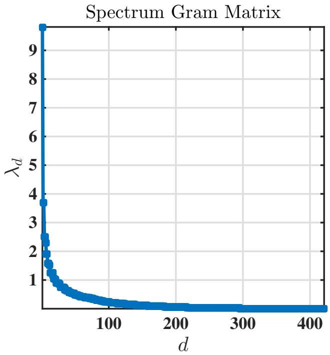
$$\lambda^\star = \sum_{d=1}^{D} \frac{\langle \lambda, \Phi \boldsymbol{\alpha}_d \rangle}{\eta_d} (\Phi \boldsymbol{\alpha}_d) = \sum_{d=1}^{D} \frac{\boldsymbol{\alpha}_d^T \boldsymbol{n}}{\eta_d} (\Phi \boldsymbol{\alpha}_d).$$

- In practice, measurements are noisy (Poisson noise). For high rates we can approximate

$$\boldsymbol{n} = \Phi^* \lambda + \epsilon,$$

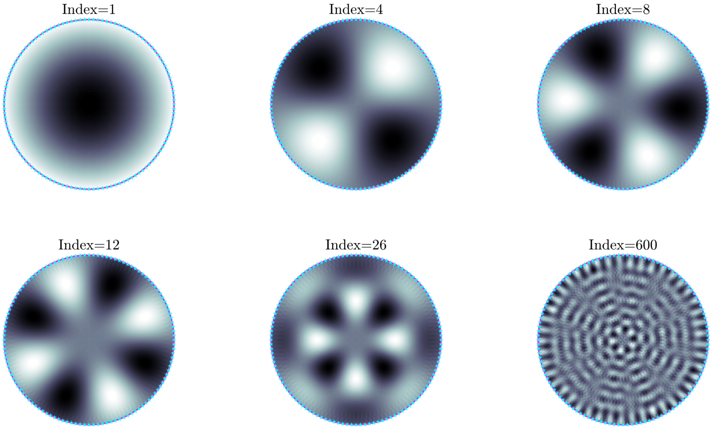
with $\epsilon \sim \mathcal{N}(0, \Sigma)$.

- Small η_d may lead to numerical instability! Need regularisation...



Back to the Eigenfunctions

- Small η_d correspond to *high-frequency eigenfunctions*! Without regularisation the estimate will be dominated by those eigenfunctions and hence very wiggly...



Regularisation

- Two avenues: Spectral truncation vs. Tikhonov regularization.
  - Spectral truncation:

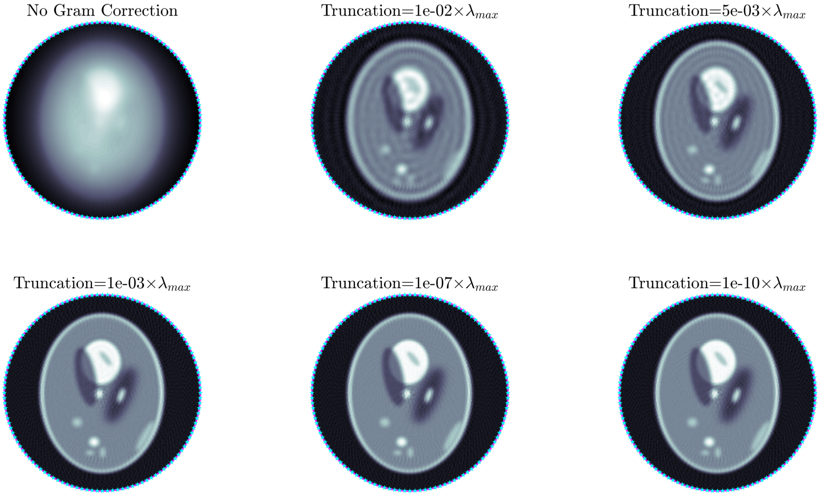
$$\lambda^\star_{ST} = \sum_{d=1}^{\tau} \frac{\boldsymbol{\alpha}_d^T \boldsymbol{n}}{\eta_d} (\Phi \boldsymbol{\alpha}_d),$$

    with $\tau \leq D$ some integer, truncation parameter.
  - Tikhonov regularization:

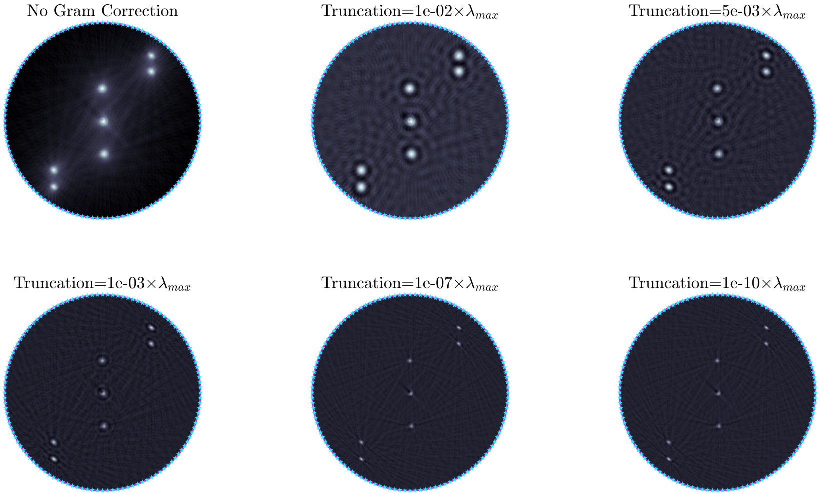
$$\lambda^\star_{\rho} = \sum_{d=1}^{D} \frac{\boldsymbol{\alpha}_d^T \boldsymbol{n}}{\eta_d + \rho} (\Phi \boldsymbol{\alpha}_d),$$

    with $\rho > 0$, called the ridge parameter.
- Act as smoothers. Tikhonov performs better (in terms of consistency), but spectral truncation is more economic computationally (less terms in the summation).

Spectral Truncation

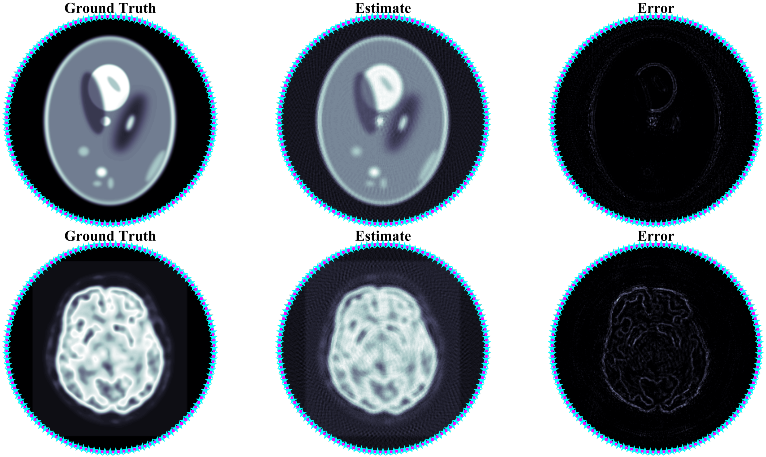


Spectral Truncation (Point Spread Function)

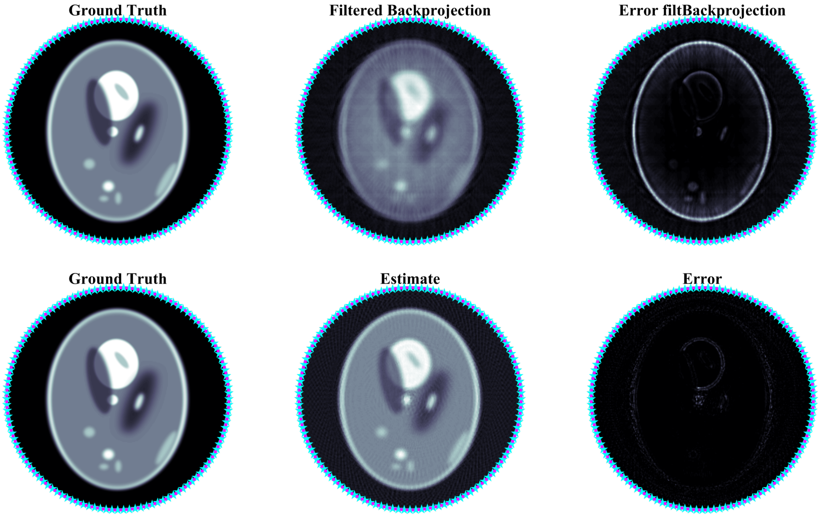


Final Estimate

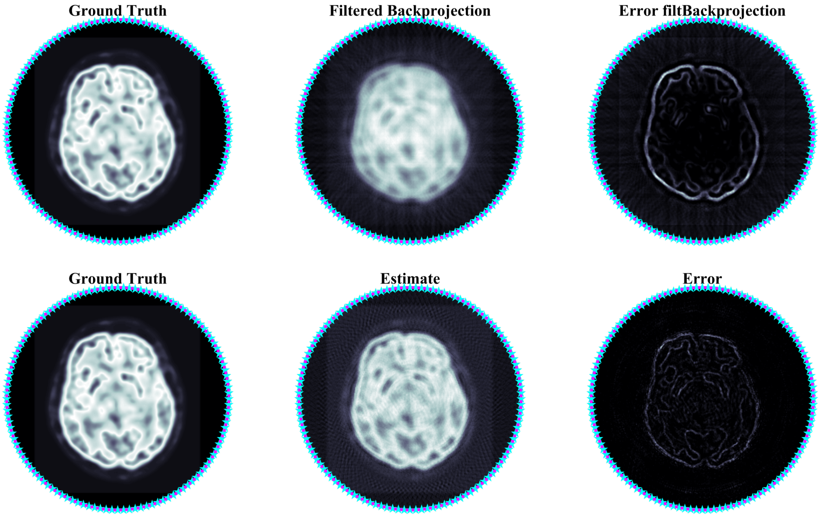
- Optimal truncation parameter chosen according to the width of point spread function main lobe.



Comparison: FBP vs. Interpolation



Comparison: FBP vs. Interpolation



Comparison: FBP vs. Interpolation

- Interpolation produces more accurate images than state of the art methods.
- Estimate can be sampled and displayed at any resolution (continuous estimate).
- Filtered backprojection scales however better with the number of detectors.
- Indeed the Gram matrix is expensive to compute and invert.
- Need to investigate dimensionality reduction via sketching (random projections):

$$\Psi^* = W^H \Phi^*,$$

with $W^H : \mathbb{R}^D \to \mathbb{R}^L$, $L \ll D$.

Gaussian Sketching

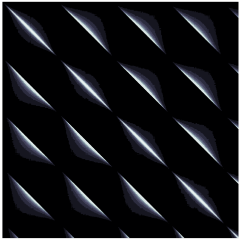
- Choose $\psi_m(x) = \sum_{d=1}^{D} W_{dm}\phi_d(x)$, with

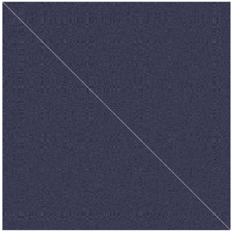
$$W_{dm} \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma).$$

- New Gram is $\boldsymbol{G}_\Psi = W^H \Phi^* \Phi W$.
- If $\sigma = 1/\text{trace}(\Phi^*\Phi)$ we can show that

$$\mathbb{E}\left[W^H \Phi^* \Phi W\right] = I_{L\times L}.$$

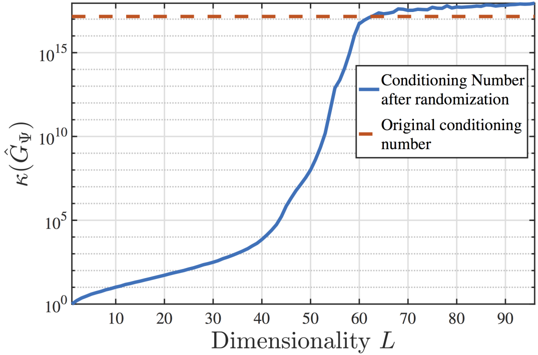
- Sketching acts as a preconditioner (improves conditioning). In expectation, the Gram is identity...
- Basis functions are less coherent.



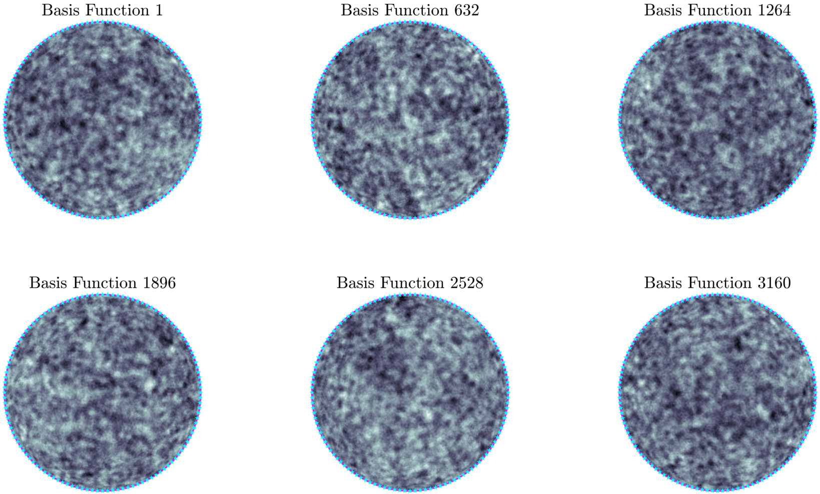


Compression Factor

- In practice we only have one random realization of $W^H \Phi^* \Phi W$.
- No guarantee it would fall near the mean!
- For small dimensions, this is more likely to be the case.
- Example: for $D = 7140$ and $L = 3160$, we go from $\kappa(G_\Phi) = 1.9458 \times 10^{21}$ to $\kappa(G_\Psi) = 7476$.
- Huge improvement (18 orders of magnitude)!

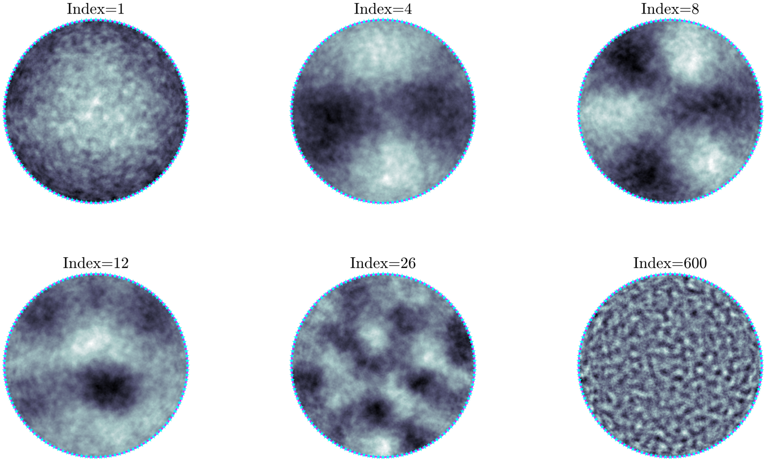


Results ($D = 7140$, $M = 3160$)

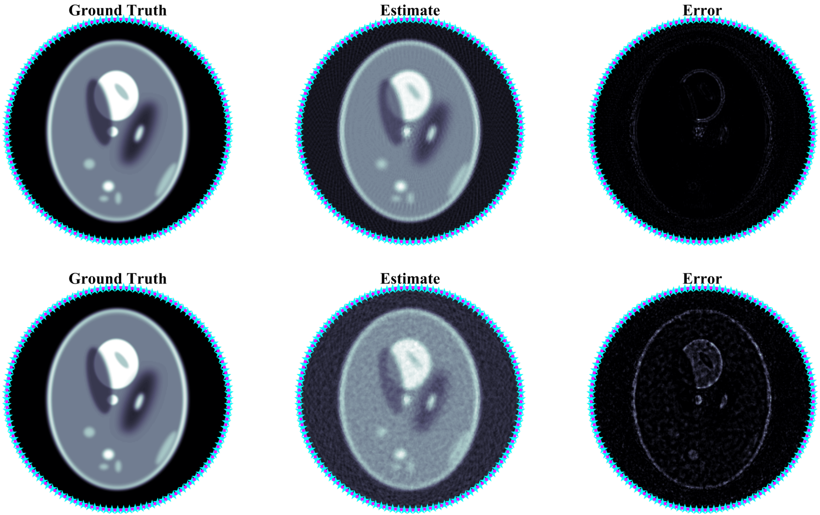


Results ($D = 7140$, $M = 3160$)

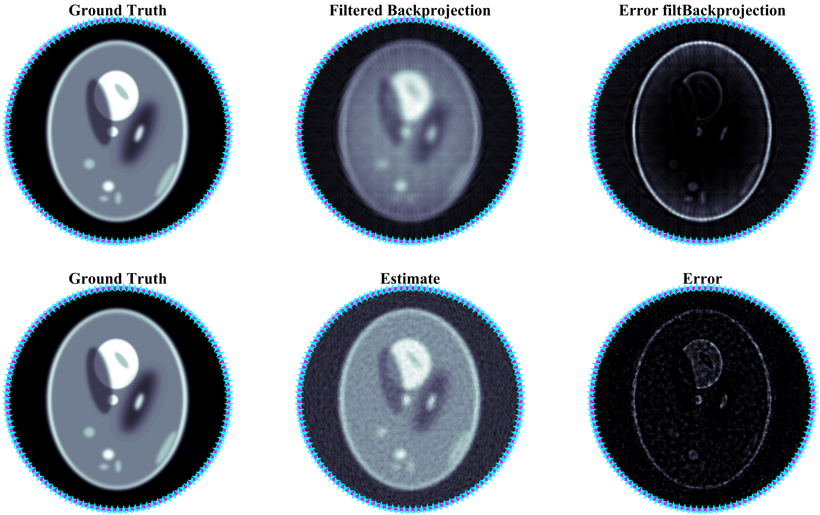
- Eigenfunctions of the sketched sampling operator Ψ^*:



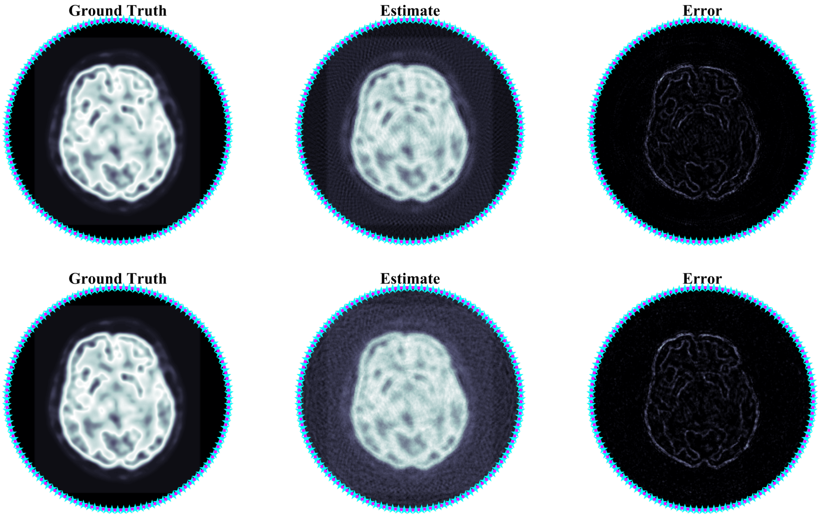
Comparison without/with Sketching



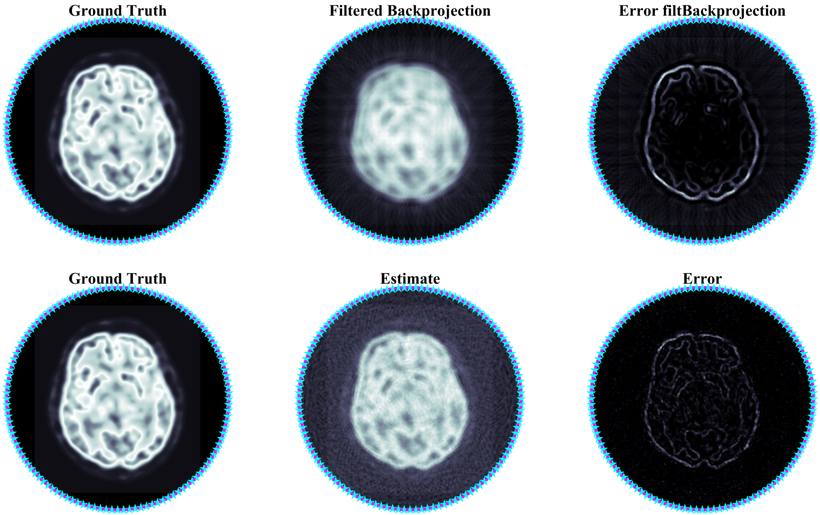
Comparison Sketching vs. FB



Comparison without/with Sketching



Comparison Sketching vs. FB



References I

[1] Matthieu Martin Jean-Andre Simeoni.
Statistical inference in positron emission tomography.
Technical report, 2014.

[2] Stanley R Deans.
***The Radon transform and some of its applications*.**
Courier Corporation, 2007.