Tomography reconstruction using CIL

Advanced Image Reconstruction Methods

12 September 2022

Evangelos Papoutsellis – STFC

Engineering and Physical Sciences Research Council

Science and Technology Facilities Council
Scientific Computing

MANCHESTER 1824
The University of Manchester

CoSeC
Computational Science Centre for Research Communities

CCPi
Tomographic Imaging

SyneRBI

evangelos.papoutsellis@stfc.ac.uk

Outline of talk

1. Introduction to CIL
2. Design of CIL
3. Optimisation Framework
4. Case Studies in CIL
5. Stochastic optimization in CIL

What is Core Imaging Library (CIL)?

- CIL is an open-source Python library for solving **Imaging Inverse Problems**
- Special emphasis on tomography applications with **challenging data sets**: low-count, non-standard geometries, incomplete, multi-channel
- **Highly flexible** and modular set of tools for different imaging problems
- **"Near-math"** specification and solution of **optimization problems**
- **Simple** to get started – **powerful** enough for large, real applications

- Funded by the **C**ollaborative **C**omputational **P**roject in Tomographic **I**maging **(CCPi)**
- Apache v2 license.
- Actively developed on GitHub: https://github.com/TomographicImaging/CIL

What is CCPi?

The Collaborative Computational Projects (CCPs) bring together leading UK expertise in key fields of computational research to tackle large-scale scientific software development, maintenance and distribution.

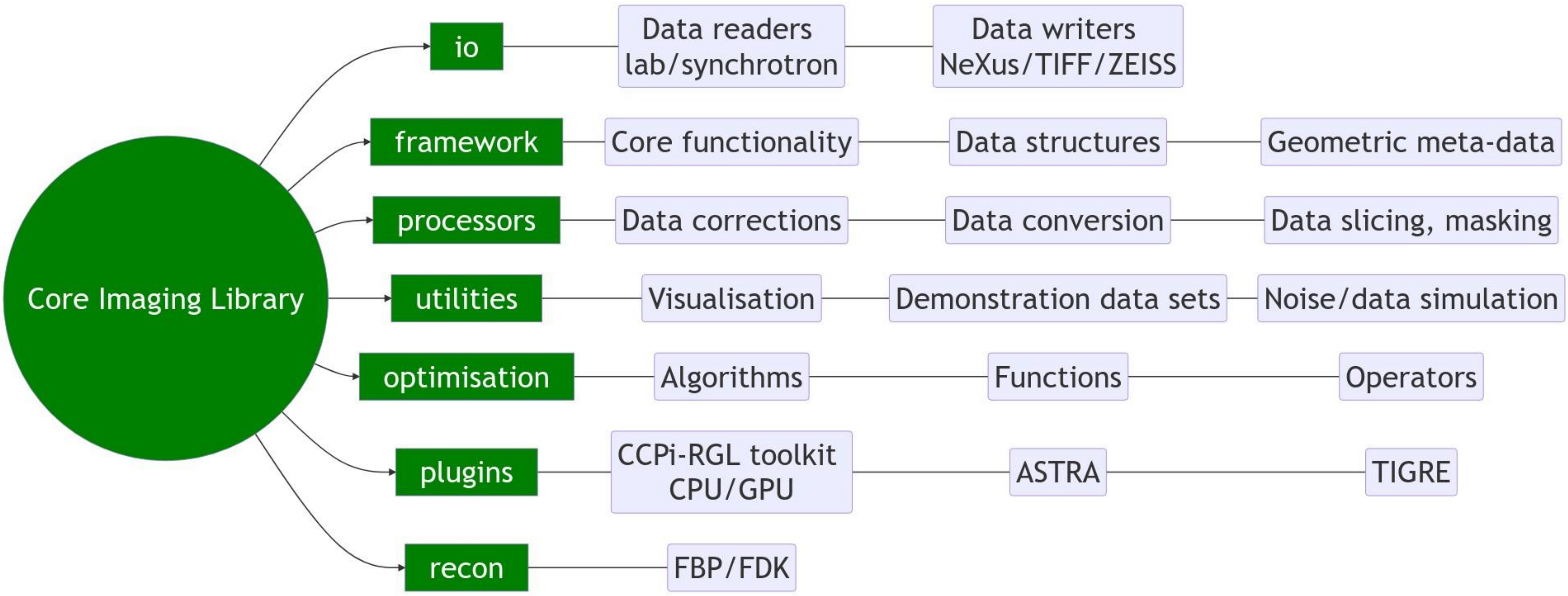
The Collaborative Computational Project in Tomographic Imaging (https://www.ccpi.ac.uk/) aims to provide the UK tomography community with a software toolbox of algorithms that increases the quality and level of information extracted by computed tomography.

- **PI:** Prof Philip Withers, Manchester
- **Phase III**: 2020 – 2025, new Co-Is:
  Martin Turner, Manchester
  Jakob Jørgensen, DTU
  Jay Warnett, Warwick
  Llion Evans, Swansea
- **Core Software Developers**:
  Edoardo Pasca, Gemma Fardell, Evangelos Papoutsellis and Laura Murgatroyd @ RAL

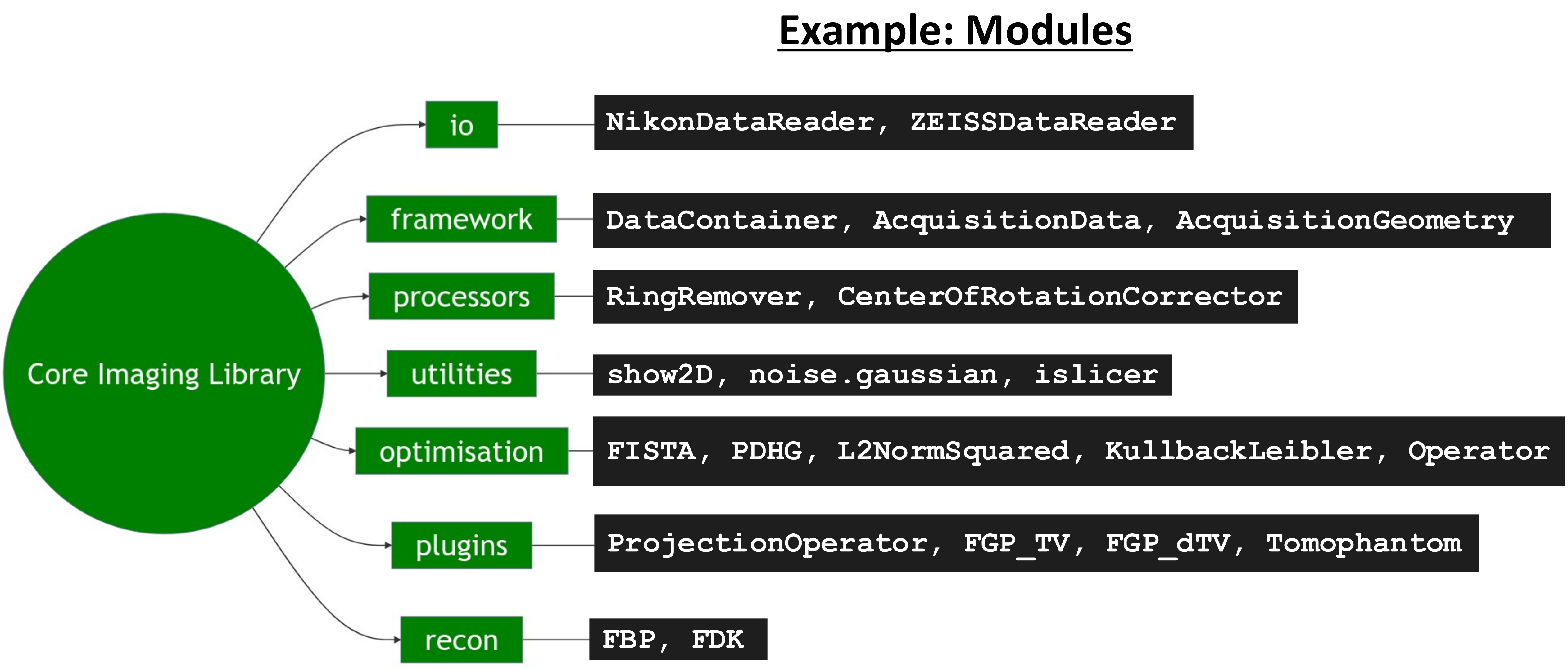
Who is CIL for?

- CT experimentalists
  - To quickly write CT processing pipelines
  - Optimised standard algorithms for large data
  - To utilise reconstruction algorithms for poor data quality or to handle novel imaging modalities
- Image processing specialists
  - to easily implement new reconstruction algorithms
  - assess them against existing ones.
- CIL:
  - is easy to use
  - Batch processing
  - Reproducibility

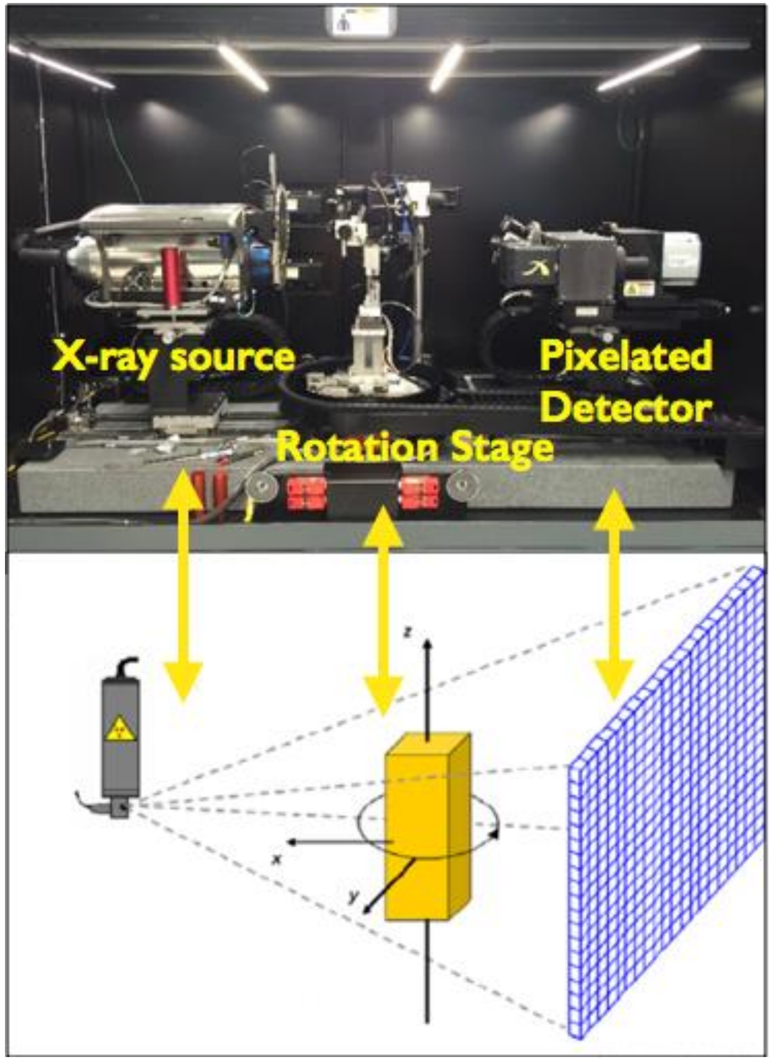
Design of CIL



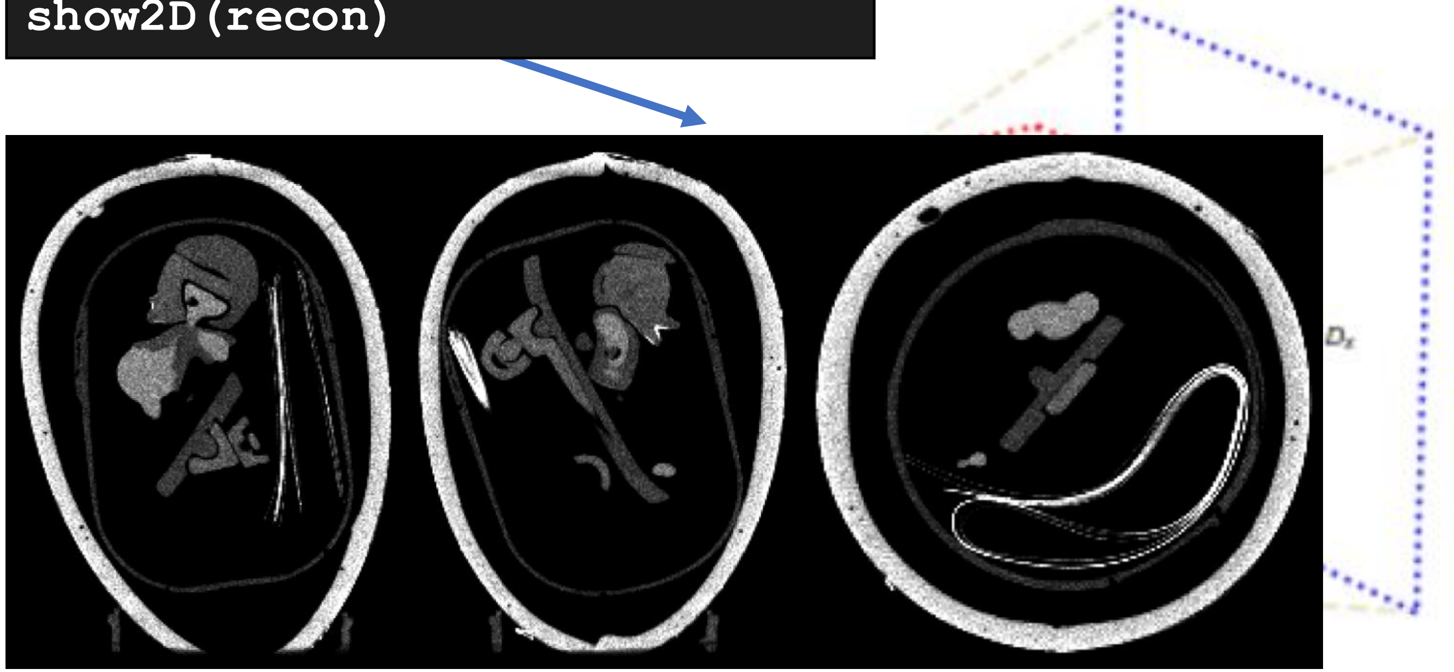
Design of CIL



CIL for X-Ray Computed Tomography (CT)




```
data = TXRMDataReader(filename).read()
data = TransmissionAbsorptionConverter()(data)
data = CentreOfRotatioCorrector.image_sharpness()(data)
recon = FBP(data).run()
show2D(recon)
```



Optimisation in CIL

$$\underbrace{\Phi(u)}_{\textbf{Objective}} = \underbrace{\mathcal{D}(\mathcal{A}u, d)}_{\textbf{Fidelity}} + \underbrace{\Psi(u)}_{\textbf{Regulariser}} \Rightarrow \min_{u} \Phi(u)$$

Optimisation in CIL

$$\Phi(u) = \mathcal{D}(\mathcal{A}u, d) + \Psi(u) \quad \Rightarrow \min_u \Phi(u)$$

name	description
BlockFunction	separable sum of multiple functions
ConstantFunction	function taking the constant value
OperatorCompositionFunction	compose function f and operator A: $f(Ax)$
IndicatorBox	indicator function for box (lower/upper) constraints
KullbackLeibler	Kullback–Leibler divergence data fidelity
L1Norm	L^1-norm: $\|x\|_1 = \sum_i \|x_i\|$
L2NormSquared	squared L^2-norm: $\|x\|_2^2 = \sum_i x_i^2$
LeastSquares	least-squares data fidelity: $\|Ax - b\|_2^2$
MixedL21Norm	mixed $L^{2,1}$-norm: $\|(U_1; U_2)\|_{2,1} = \|(U_1^2 + U_2^2)^{1/2}\|_1$
SmoothMixedL21Norm	smooth $L^{2,1}$-norm: $\|(U_1; U_2)\|_{2,1}^S = \|(U_1^2 + U_2^2 + \beta^2)^{1/2}\|_1$
WeightedL2NormSquared	weighted squared L^2-norm: $\|x\|_w^2 = \sum_i (w_i \cdot x_i^2)$

Optimisation in CIL

$$\Phi(u) = \mathcal{D}(\mathcal{A}u, d) + \Psi(u) \quad \Rightarrow \min_u \Phi(u)$$

name	description
BlockOperator	form block (array) operator from multiple operators
BlurringOperator	apply point spread function to blur an image
ChannelwiseOperator	apply the same operator to all channels
DiagonalOperator	form a diagonal operator from image/acquisition data
FiniteDifferenceOperator	apply finite differences in selected dimension
GradientOperator	apply finite difference to multiple/all dimensions
IdentityOperator	apply identity operator, i.e. return input
MaskOperator	from binary input, keep selected entries, mask out rest
SymmetrisedGradientOperator	apply symmetrized gradient, used in TGV
ZeroOperator	operator of all zeroes
ProjectionOperator	tomography forward/back-projection from ASTRA
ProjectionOperator	tomography forward/back-projection from TIGRE

Optimisation in CIL

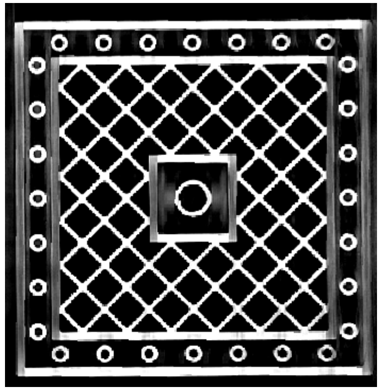
$$\Phi(u) = \mathcal{D}(\mathcal{A}u, d) + \Phi(u) \Rightarrow \min_{u} \Phi(u)$$

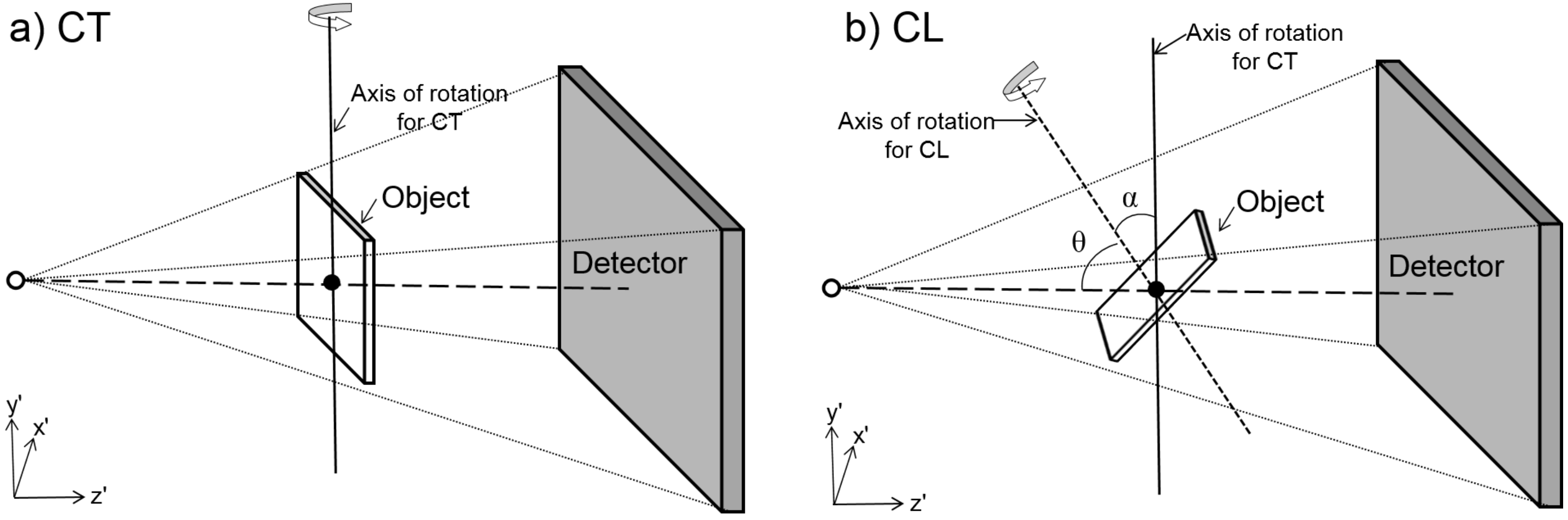
name	description	problem type solved
CGLS	conjugate gradient least squares	least squares
SIRT	simultaneous iterative reconstruction technique	weighted least squares
GD	gradient descent	smooth
FISTA	fast iterative shrinkage-thresholding algorithm	smooth + non-smooth
LADMM	linearized alternating direction method of multipliers	non-smooth
PDHG	primal dual hybrid gradient	non-smooth
SPDHG	stochastic primal dual hybrid gradient	non-smooth

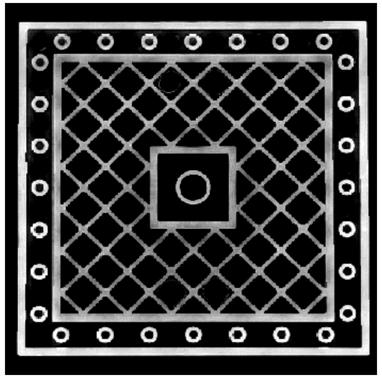
Case Studies in CIL : Laminography

Tomographic imaging with Laminography

- Planar samples, like composite panels and printed circuit boards are difficult to scan due to different exposure along views.
- Conventional CT scan gives limited-angle artifacts, missing edges. Laminography allows uniform exposure.
- **Dataset** : Planar LEGO-brick







Jørgensen. et al. 2021

Case Studies in CIL : Laminography

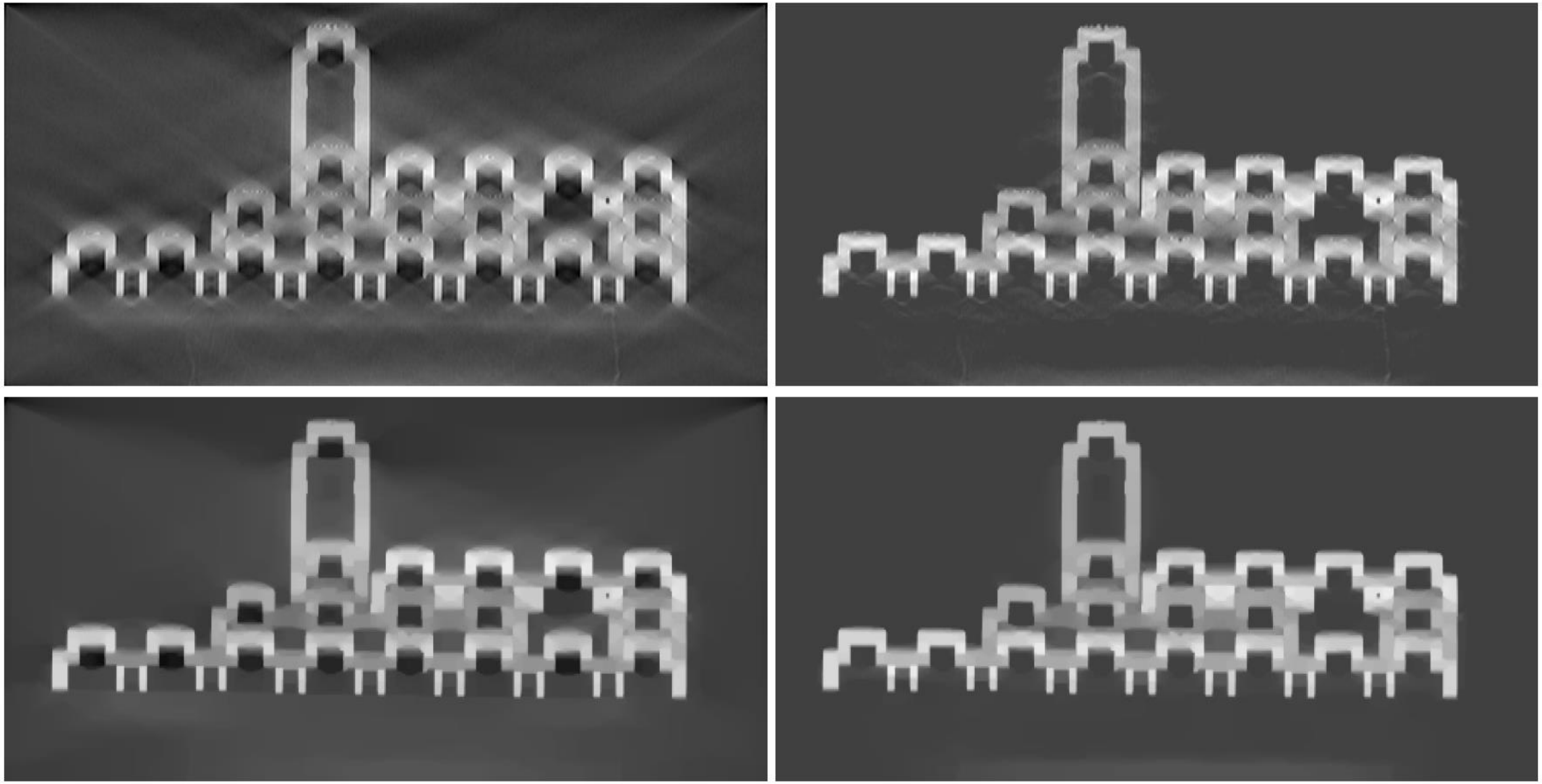
	Least Squares, unconstrained	Least squares, nonnegativity
No regu.	$\arg\min_u \frac{1}{2}\|Au - d\|^2$	$\arg\min_u \frac{1}{2}\|Au - d\|^2 + \mathbb{I}_{\{u>0\}}(u)$
TV regu.	$\arg\min_u \frac{1}{2}\|Au - d\|^2 + \alpha\,\mathrm{TV}(u)$	$\arg\min_u \frac{1}{2}\|Au - d\|^2 + \alpha\,\mathrm{TV}(u) + \mathbb{I}_{\{u>0\}}(u)$

Case Studies in CIL : Laminography

Least Squares, unconstrained | Least squares, nonnegativity

No regu.

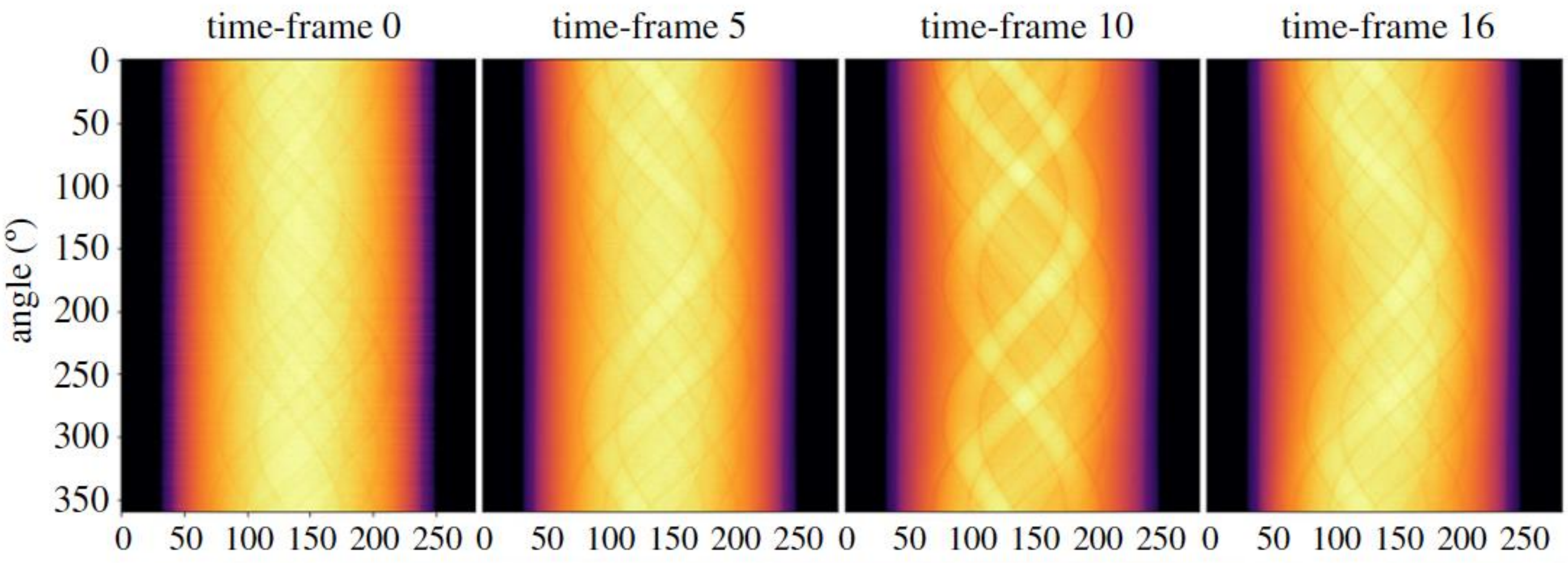
TV regu.



Case Studies in CIL : Dynamic CT

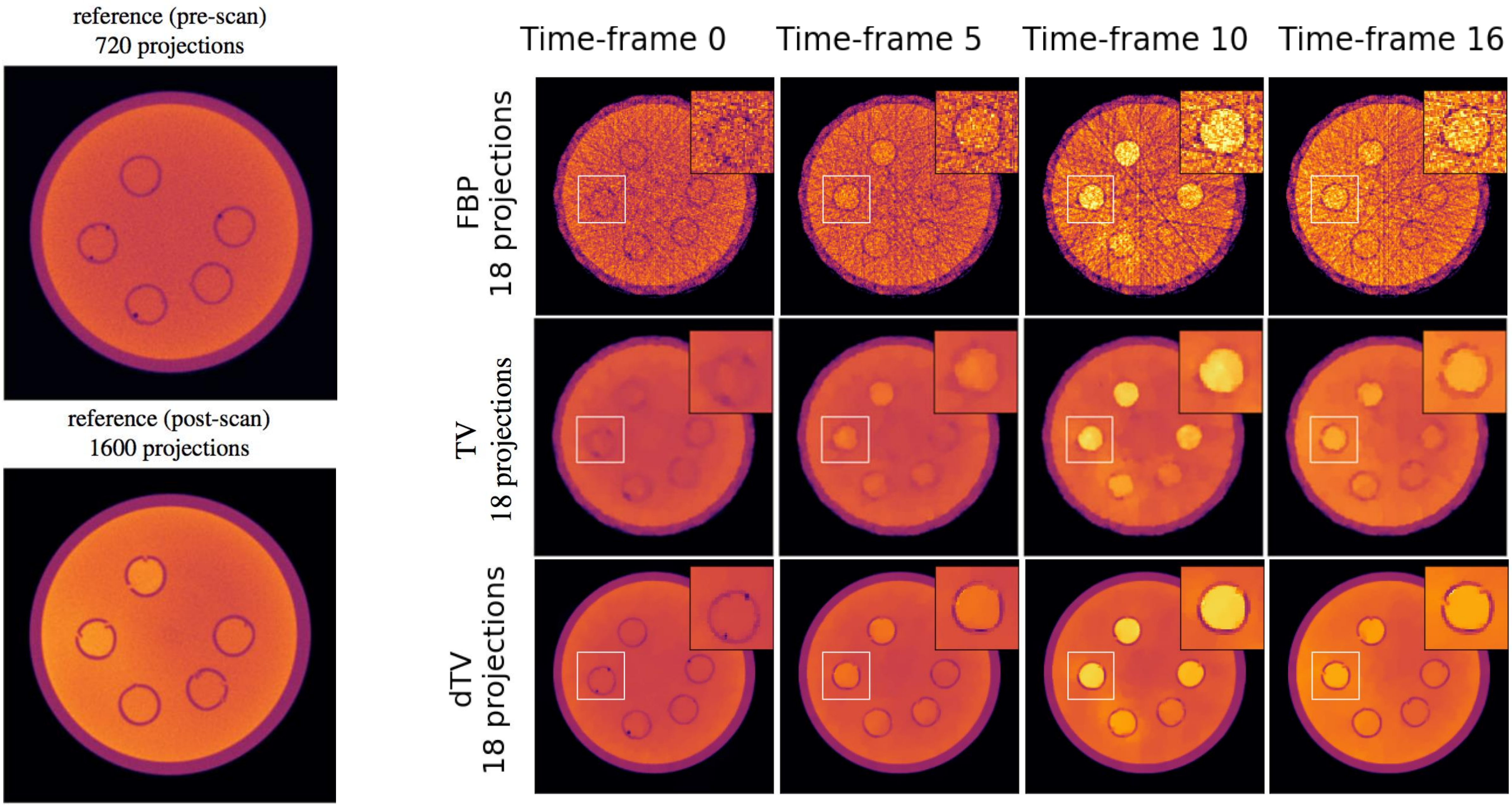
Dynamic X-Ray CT reconstruction

- Scan a sample in time that undergoes some change, internal/external.
- Use iterative reconstruction methods to increase temporal resolution and reduce radiation dose during a CT scan.
- **Dataset**: Simulate diffusion of liquids inside plant stems. Dynamic agarose-gel phantom perfused with a liquid contrast agent (iodine).



Papoutsellis et al. 2021

Case Studies in CIL : Dynamic CT



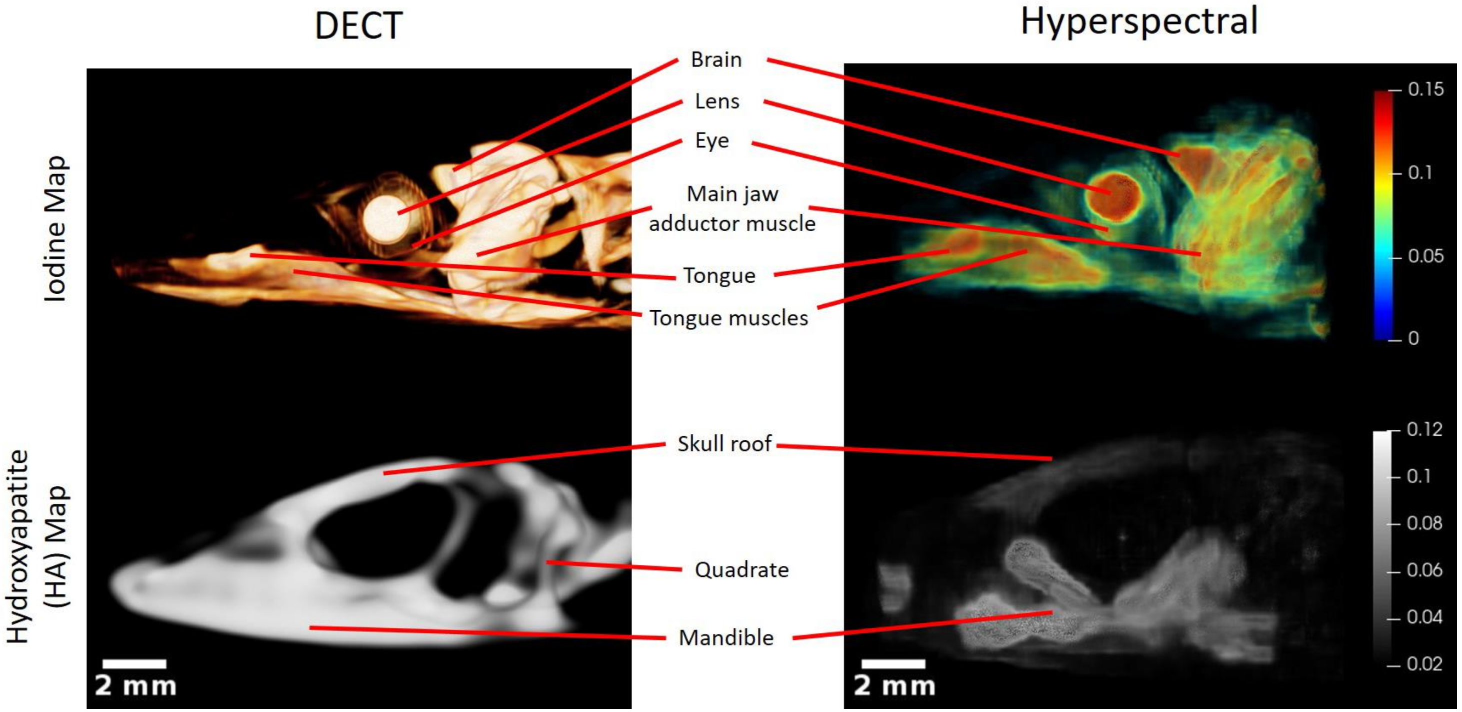
Case Studies in CIL : Hyperspectral CT

Hyperspectral Lab X-CT reconstruction Manchester Colour Bay

- In conventional X-ray CT, each detector pixel records the total number of detected photons, without energy information.
- Spectroscopic/Energy detectors can be used to extract additional, measuring both the **energy** and **position** of each incident photon.
- **Dataset**: Iodine-stained lizard head.

DECT

- **1051 projections**
- **15s exposure time**
- **4.4 hrs scan time**



Hyperspectral

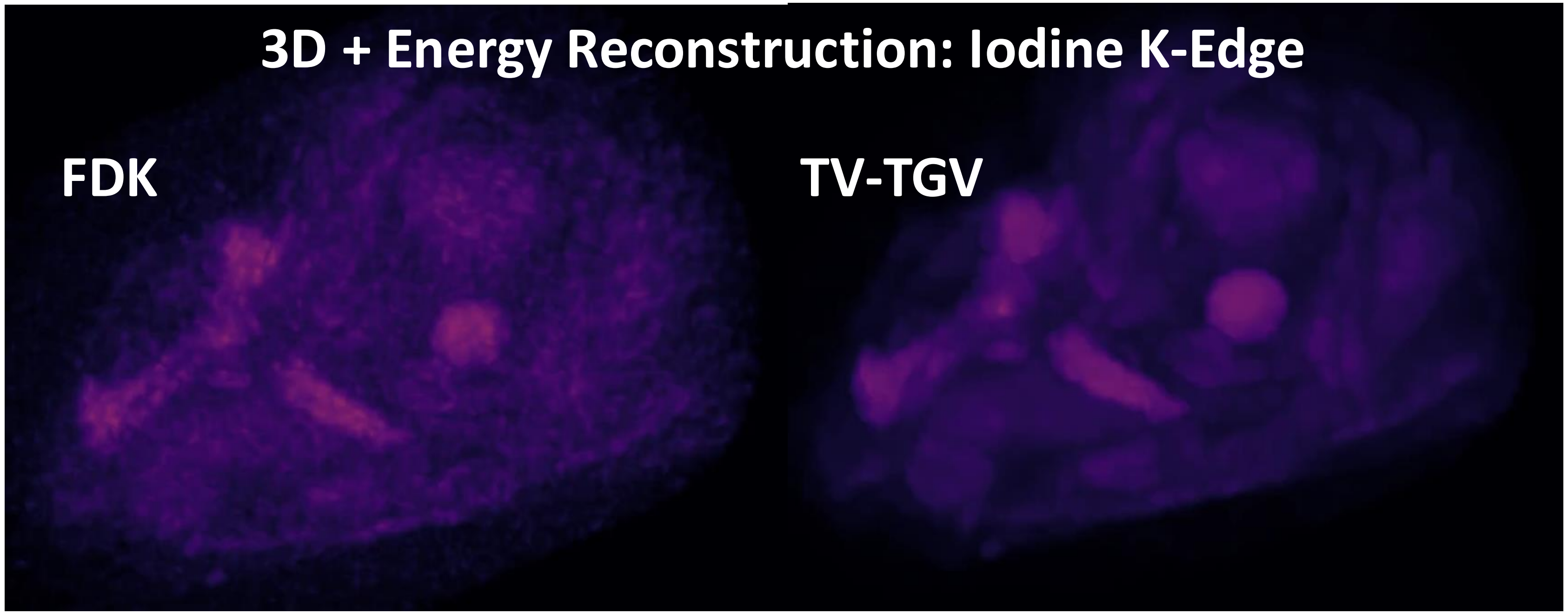
- **60 projections**
- **120s exposure time**
- **2.5 hrs scan time**

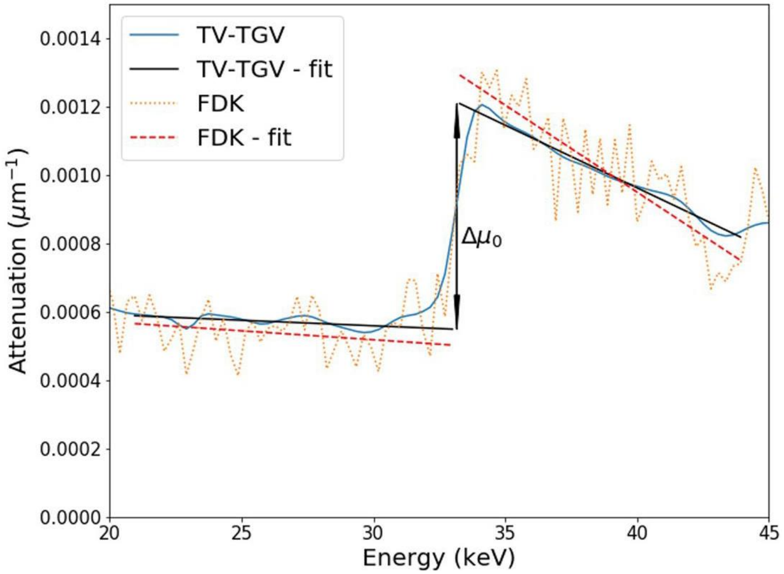
Warr et al. 2021

Case Studies in CIL : Hyperspectral CT

CIL Optimisation can build complex objectives through a mix & match setting of different existing or user-defined functions

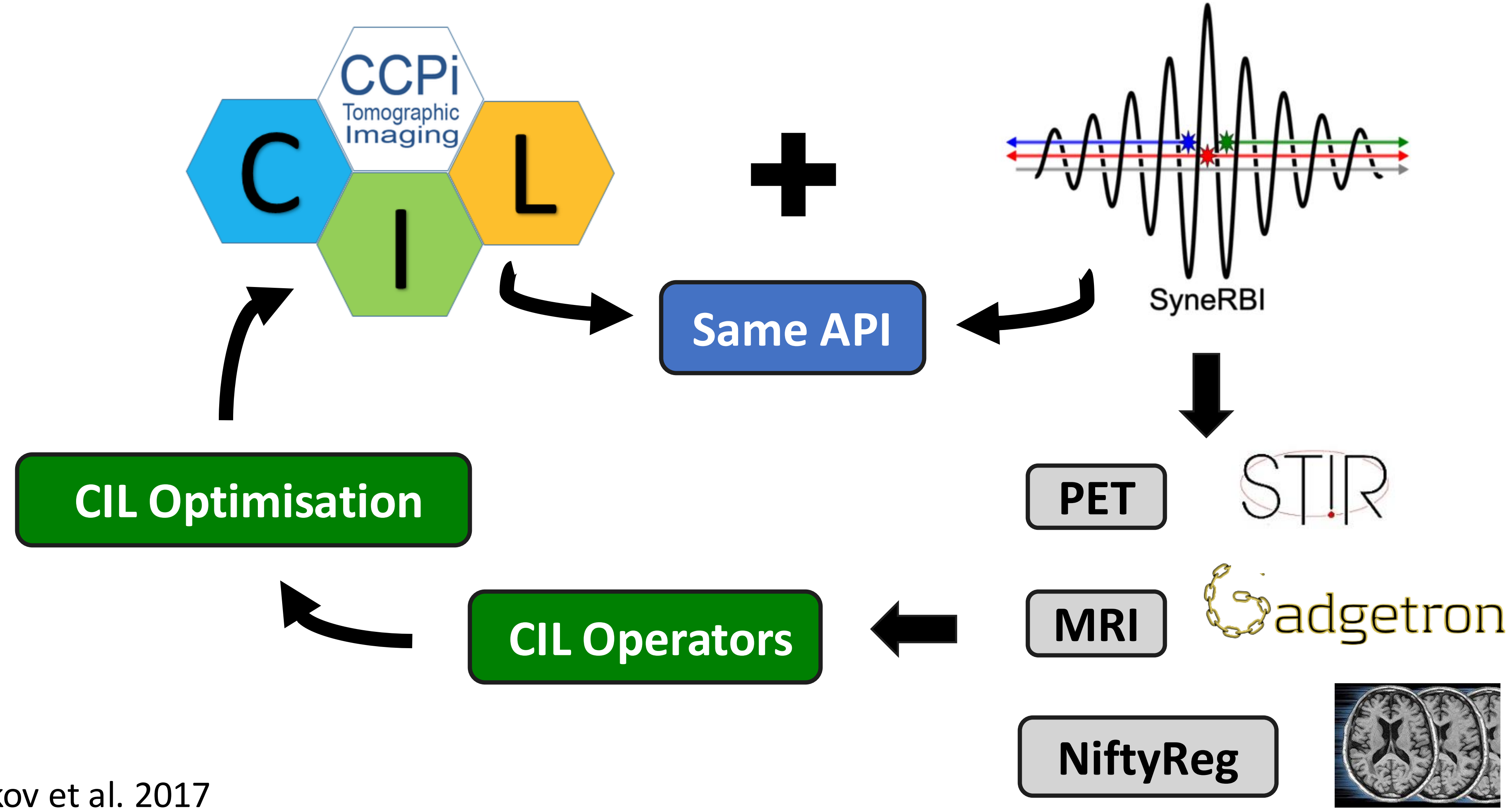
$$\arg\min_{u} \frac{1}{2}\|Au - b\|^2 + \alpha\,\mathrm{TV}(u_{space}) + \beta\,\mathrm{TGV}(u_{spectral})$$





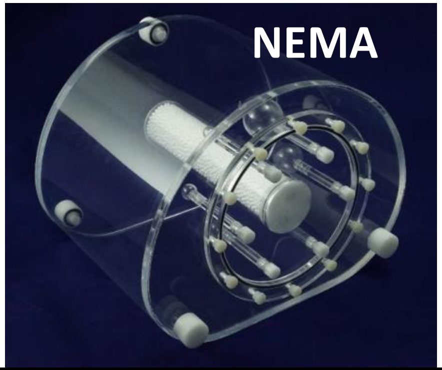
Case Studies in CIL : PET/MR Reconstruction

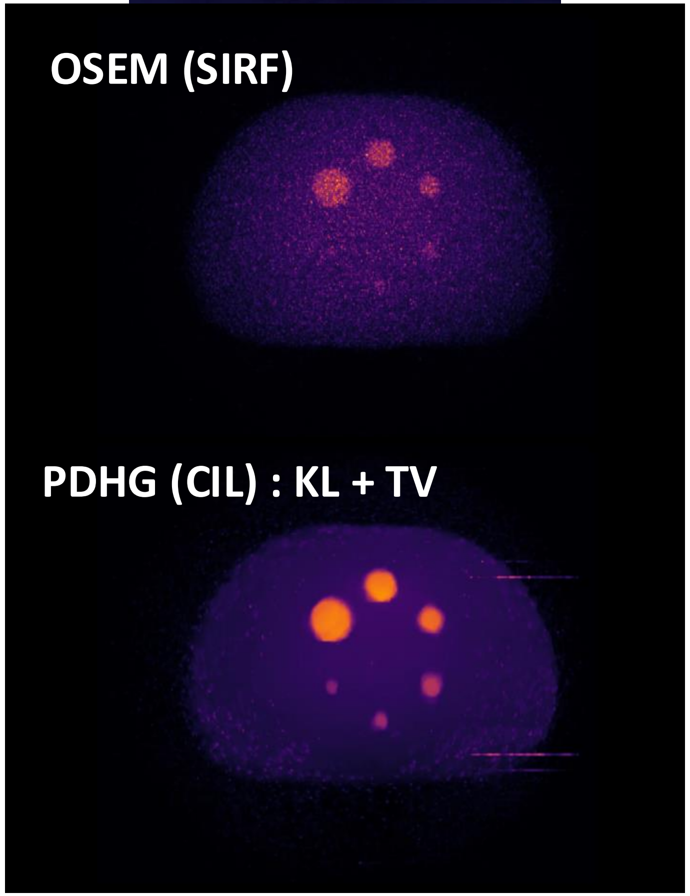
Synergy between two libraries: ***Core Imaging Library (CIL)*** *and* ***Synergestic Image Reconstruction Framework (SIRF)*** *for Medical Imaging Applications*



Ovtchinnikov et al. 2017

Case Studies in CIL : PET/MR Reconstruction

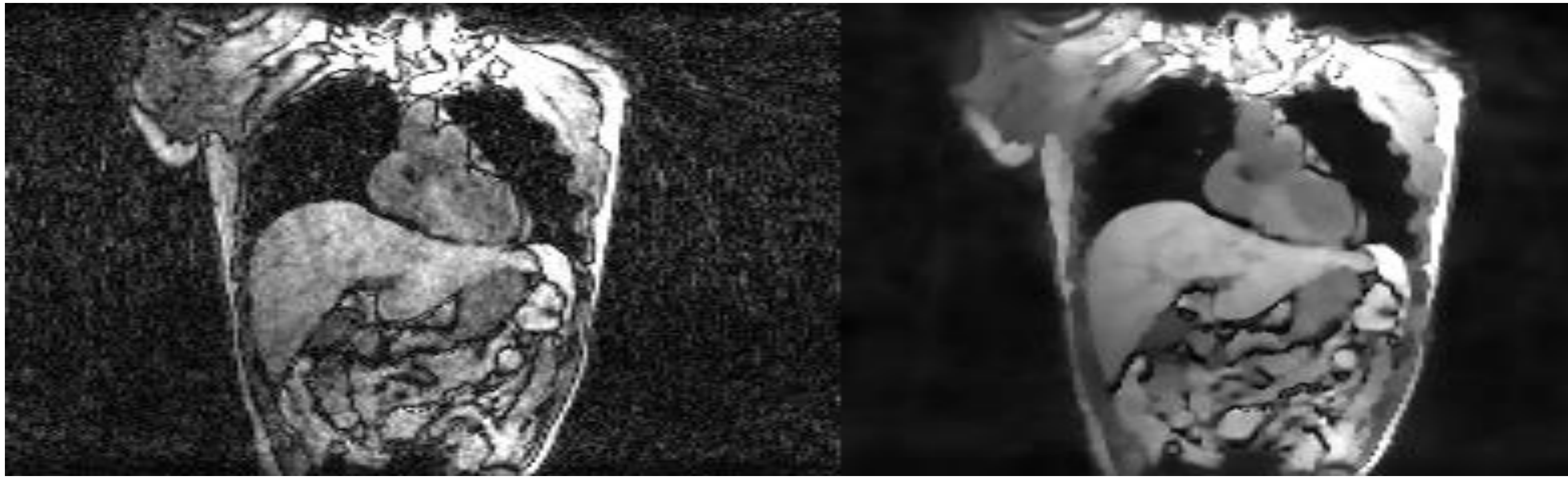




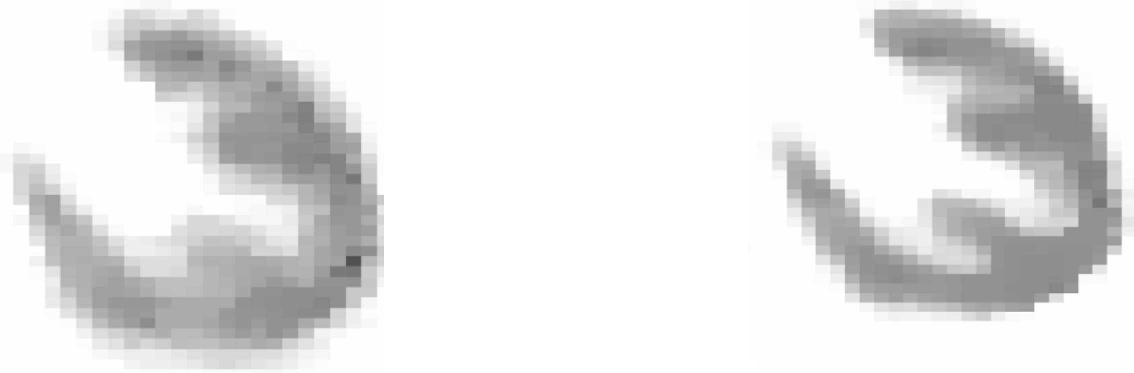
No motion correction & regularization

Motion correction with TV

MRI reconstruction



PET reconstructure Incorporate information from MRI



Brown et al. 2021

Stochastic Optimisation in CIL

Extend CIL Optimisation (Deterministic) framework to Stochastic Optimisation

Joint work : Kris Thielemans, Gillman Ashley, Tang Junqi, Zeljko Kereta, Imraj Singh, Gemma Fardell, Evgueni Ovtchinnikov, Matthias Ehrhardt, Laura Murgatroyd, Robert Twyman, Edoardo Pasca, Claire Delplancke, EP, Georg Schramm, Jakob Jørgensen, Sam Porter

- **1st Hackathon:** November 23-26, 2021
- **2nd Hackathon**: April 4-7, 2022
- **Organised**: CCP SyneRBI , CCPi , PET++

- **Implement selected randomized algorithms in CIL, e.g., SGD, SAG, SAGA, SVRG and more**
- **Implement subset data structures in STIR/SIRF**
- **Establish benchmarking framework for CT and PET applications**

Stochastic Optimisation in CIL

$$\min_{x} f(x) + g(x), \quad f\ : L\text{-smooth},\ g\ : \text{convex}$$

- In CIL, we have the following (Proximal) Gradient algorithms

GD	ISTA	FISTA
$x_{k+1} = x_k - \gamma_k \nabla f(x_k)$	$x_{k+1} = \operatorname{prox}_{\gamma_k g}(x_k - \gamma_k \nabla f(x_k))$	$x_{k+1} = \operatorname{prox}_{\gamma_k g}(y_k - \gamma_k \nabla f(y_k))$ $\alpha_{k+1} = \dfrac{1 + \sqrt{1 + 4\alpha_k^2}}{2}$ $y_k = x_k + \dfrac{\alpha_k - 1}{a_{k+1}}(x_k - x_{k-1})$

```
GD(initial, objective_function, step_size)
```

```
ISTA(initial, f, g, step_size)
```

```
FISTA(initial, f, g, step_size)
```

Stochastic Optimisation in CIL

$$\min_x \frac{1}{n}\sum_{i=1}^{n} f_i(x) + g(x), \qquad f_i\ : L_i\text{-smooth},\ g\ : \text{convex}$$

GD	ISTA	FISTA
$x_{k+1} = x_k - \gamma_k \nabla f_{i_k}(x_k)$	$x_{k+1} = \text{prox}_{\gamma_k g}(x_k - \gamma_k \nabla f_{i_k}(x_k))$	$x_{k+1} = \text{prox}_{\gamma_k g}(y_k - \gamma_k \nabla f_{i_k}(y_k))$ $\alpha_{k+1} = \frac{1+\sqrt{1+4\alpha_k^2}}{2}$ $y_k = x_k + \frac{\alpha_k - 1}{a_{k+1}}(x_k - x_{k-1})$

`SubsetSumFunction`

- CIL Class Function
- Selects (randomly) $i_k \in \{1, \ldots, n\}$
- Computes ∇f_{i_k}

Stochastic Optimisation in CIL

$$\min_x \frac{1}{n}\sum_{i=1}^{n} f_i(x) + g(x), \qquad f_i : L_i\text{-smooth},\ g : \text{convex}$$

GD	ISTA	FISTA
$x_{k+1} = x_k - \gamma_k \nabla f_{i_k}(x_k)$	$x_{k+1} = \mathrm{prox}_{\gamma_k g}(x_k - \gamma_k \nabla f_{i_k}(x_k))$	$x_{k+1} = \mathrm{prox}_{\gamma_k g}(y_k - \gamma_k \nabla f_{i_k}(y_k))$ $\alpha_{k+1} = \frac{1+\sqrt{1+4\alpha_k^2}}{2}$ $y_k = x_k + \frac{\alpha_k - 1}{a_{k+1}}(x_k - x_{k-1})$

$$\frac{1}{n}\sum_{i=1}^{n} f_i(x)$$ `SubsetSumFunction` $\Bigg\{$ `SGDFunction`

Stochastic Optimisation in CIL

$$\min_x \frac{1}{n}\sum_{i=1}^{n} f_i(x) + g(x), \qquad f_i\ : L_i\text{-smooth},\ g\ : \text{convex}$$

GD	ISTA	FISTA
$x_{k+1} = x_k - \gamma_k \tilde{\nabla} f_{i_k}(x_k)$	$x_{k+1} = \text{prox}_{\gamma_k g}(x_k - \gamma_k \tilde{\nabla} f_{i_k}(x_k))$	$x_{k+1} = \text{prox}_{\gamma_k g}(y_k - \gamma_k \tilde{\nabla} f_{i_k}(y_k))$ $\alpha_{k+1} = \frac{1+\sqrt{1+\alpha_k^2}}{2}$ $y_{k+1} = x_k + \frac{\alpha_k - 1}{\alpha_{k+1}}(x_k - x_{k-1})$

$$\frac{1}{n}\sum_{i=1}^{n} f_i(x)$$

`SubsetSumFunction`

- `SGDFunction`
- `SAGFunction`
- `SAGAFunction`
- `SVRGFunction`

Stochastic Optimisation in CIL

Plug and Play Framework - Different Stochastic Gradient Estimators

SubsetSumFunction \ Algorithms	GD	ISTA	FISTA
SGDFunction	SGD	Prox-SGD	Acc-Prox-SGD
SAGFunction	SAG	Prox-SAG	Acc-Prox-SAG
SAGAFunction	SAGA	Prox-SAGA	Acc-Prox-SAGA
SVRGFunction	SVRG	Prox-SVRG	Acc-Prox-SVRG

Add Composite part

$$\min_{x} f(x) + g(x) + h(Ax), \ \text{(PD3O, Condat-Vũ)}$$

Stochastic Optimisation in CIL

Extend to Stochastic Primal-Dual algorithms (Sum of three functions)

$$\min_x f(x) + g(x) + h(Ax), \;\; (\text{PD3O, Condat-Vũ})$$

$$f : \text{L–smooth}, \quad g : \text{proximable}, \quad h : \text{composite}$$

$$\begin{aligned}
x_k &= \operatorname{prox}_{\sigma g}(p_k) \\
w_k &= 2x_k - p_k - \sigma \nabla f(x_k) \\
y_{k+1} &= \operatorname{prox}_{\tau h^*}(y_k - \tau A(w_k - \sigma A^* y_k)) \\
p_{k+1} &= x_k - \sigma \nabla f(x_k) - \sigma A^* y_{k+1}
\end{aligned}$$

Stochastic Optimisation in CIL

Extend to Stochastic Primal-Dual algorithms (Sum of three functions)

$$\min_x f(x) + g(x) + h(Ax), \ \ (\text{PD3O, Condat-Vũ})$$

$$f: \text{L–smooth}, \quad g: \text{proximable}, \quad h: \text{composite}$$

- Yurtserver et al. 2016: Gradient Estimators, $A = \mathrm{I}$
- Zhao et al. 2018 : Gradient Estimators, $A \neq \mathrm{I}$
- Salim et al. 2022 : Variance-Reduced Estimators, $A \neq \mathrm{I}$

$$\begin{aligned} x_k &= \mathrm{prox}_{\sigma g}(p_k) \\ w_k &= 2x_k - p_k - \sigma \tilde{\nabla} f_{i_k}(x_k) \\ y_{k+1} &= \mathrm{prox}_{\tau h^*}(y_k - \tau A(w_k - \sigma A^* y_k)) \\ p_{k+1} &= x_k - \sigma \tilde{\nabla} f_{i_k}(x_k) - \sigma A^* y_{k+1} \end{aligned}$$

Stochastic Optimisation in CIL

Extend to Stochastic Primal-Dual algorithms (Sum of three functions)

$$\min_{x} f(x) + g(x) + h(Ax), \;\; (\text{PD3O, Condat-Vũ})$$

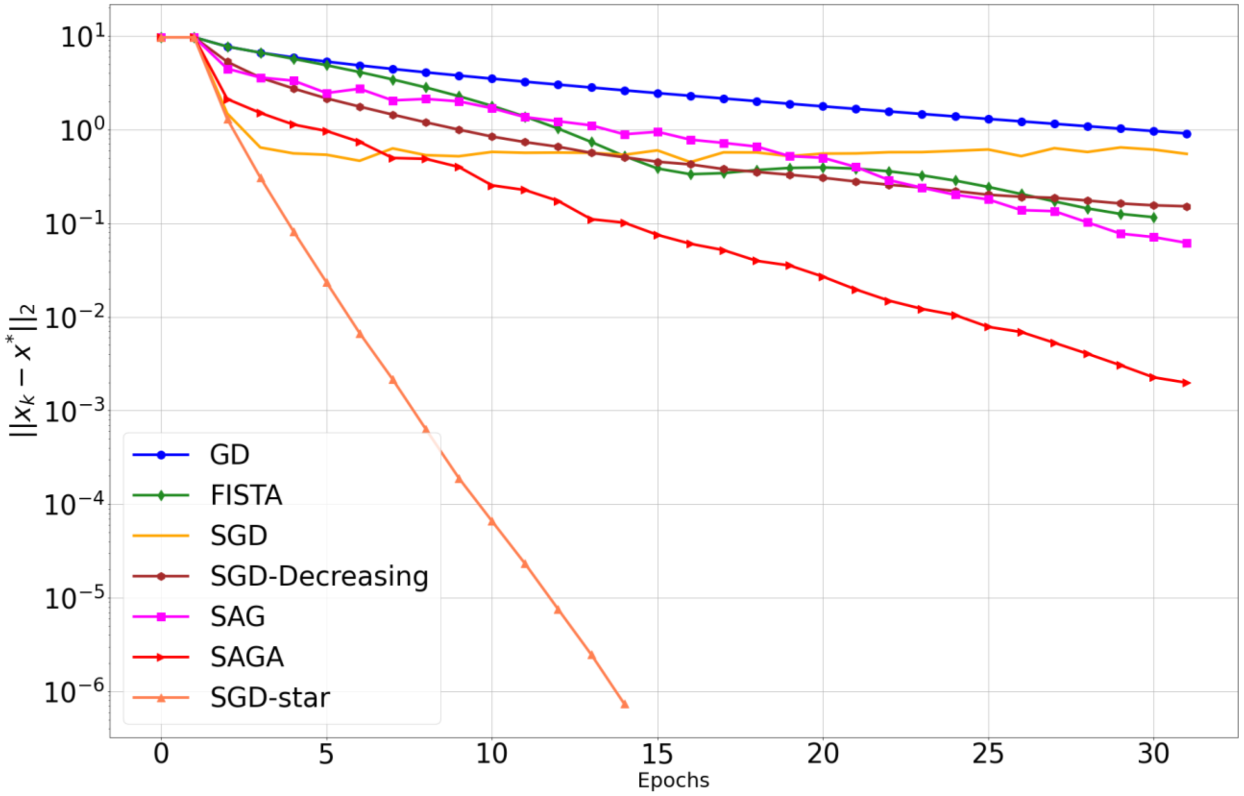
$$f \,:\, \text{L–smooth}, \quad g \,:\, \text{proximable}, \quad h \,:\, \text{composite}$$

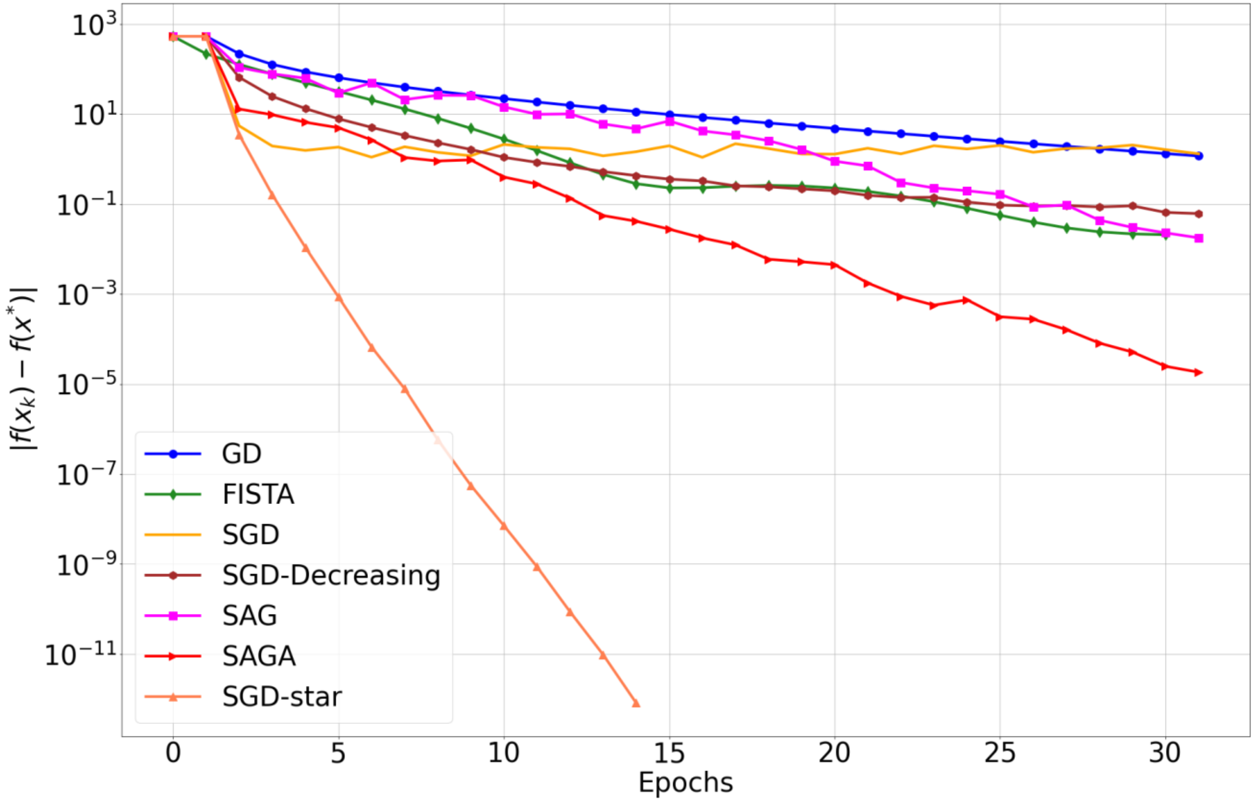
Algorithms / SubsetSumFunction	PD3O
SGDFunction	PD3O-SGD
SAGFunction	PD3O-SAG
SAGAFunction	PD3O-SAGA
SVRGFunction	PD3O-SVRG

Stochastic Optimisation in CIL

$$\min_{x\in\mathbb{R}^m}\left\{\frac{1}{2}\|Ax-d\|^2=\frac{1}{2n}\sum_{i=1}^{n}\|A_i x-d_i\|^2\right\},\ m=1000,\ n=100$$

x^* : CVXpy solution





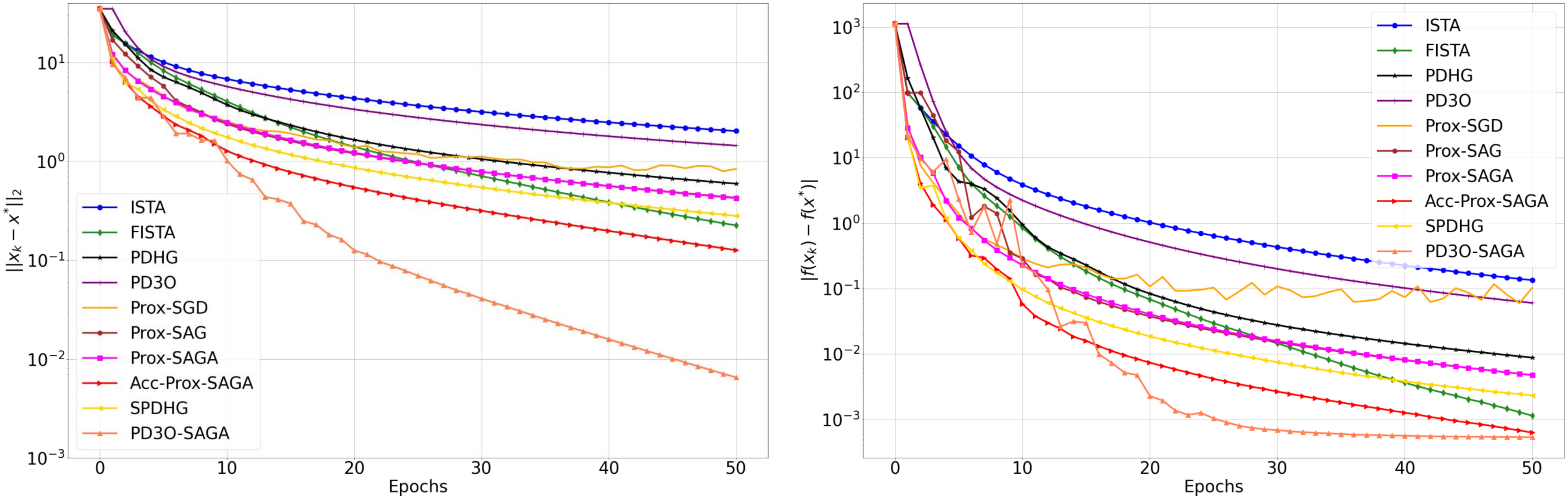
(*) SGD – star (Gorbunov et al. 2020)

$$v_k = \nabla f_{i_k}(x_k) - \nabla f_{i_k}(x^*) + \nabla f(x^*)$$

$$x_{k+1} = \mathrm{prox}_{\gamma_k g}(x_k - \gamma_k v_k)$$

Stochastic Optimisation in CIL

$$\min_{x} \left\{ \frac{1}{2}\|Ax - d\|^2 + \alpha\|\nabla x\|_{2,1} + \mathbb{I}_{\{x>0\}}(x) \right\}$$



Dataset : 3D Lizard Dataset, n = 30, initial = FBP reconstruction, x^* : FISTA 1000 epochs

Stochastic Optimisation in CIL

$$\min_{x} \left\{ \sum Ax - d\log(Ax+\eta) + \alpha\|\nabla x\|_{2,1} + \mathbb{I}_{\{x>0\}}(x) \right\}$$

Extend SPDHG with three operators splitting [WIP]

SPDHG: randomness in dual update h_i^* (Chambolle et al. 2018)

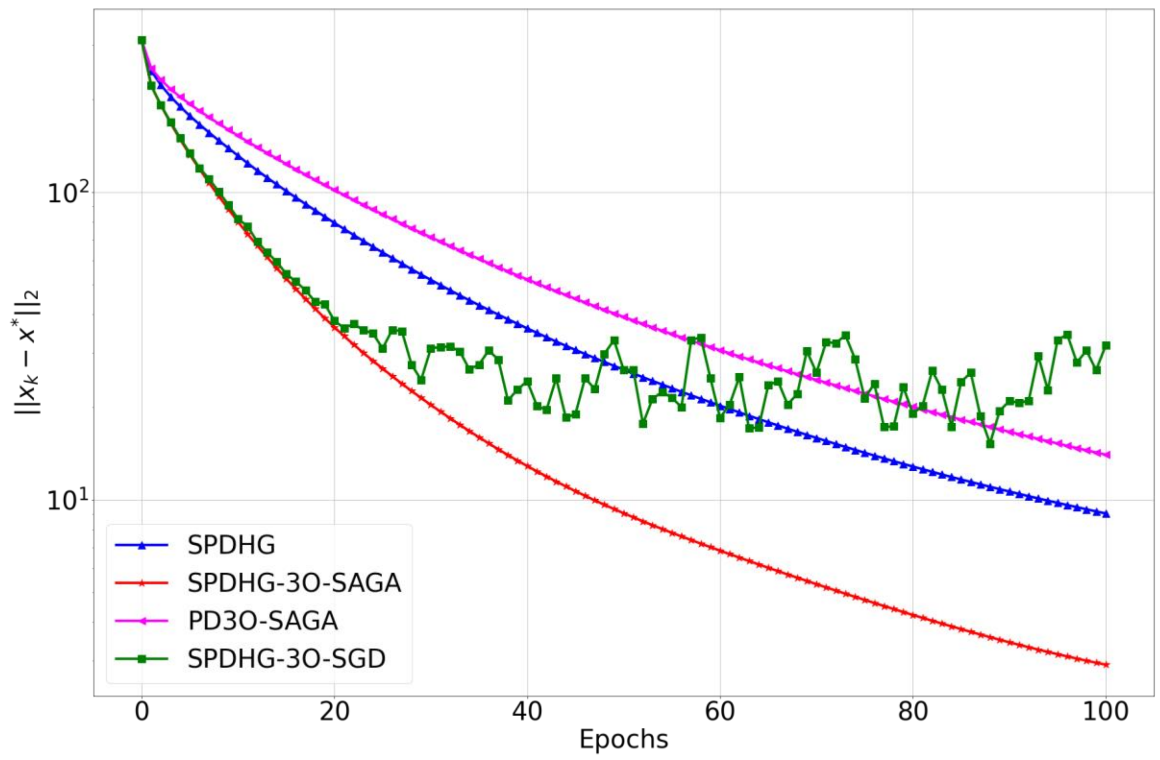
SPDHG – 3O: Split data fitting term in composite (proximal) and non-composite (gradient) terms (Combettes et al. 2019)

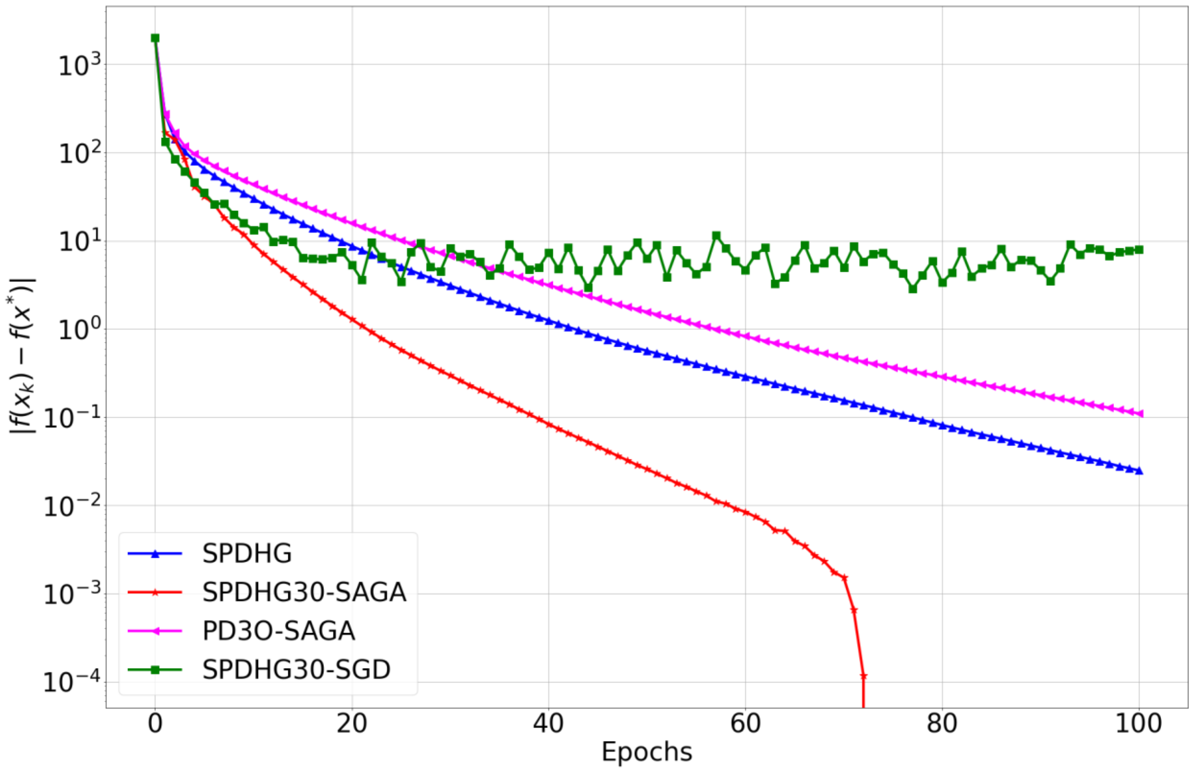
$$f(x) = \sum_{i=0}^{m-1} \mathrm{KL}(d_i, A_i x + \eta_i)\ ,\ h(Ax) = \sum_{i=m}^{n-1} \mathrm{KL}(d_i, A_i x + \eta_i)$$

1) **randomness in dual update**
2) **randomness in primal update (stochastic gradient/ variance reduced estimators)**

Stochastic Optimisation in CIL

$$\min_{x} \left\{ \sum Ax - d\log(Ax+\eta) + \alpha\|\nabla x\|_{2,1} + \mathbb{I}_{\{x>0\}}(x) \right\}$$





Dataset : Simulated 2D Thorax Slice, n = 8, initial = OSEM, x^* : FISTA 1000 epochs

Summary

- CIL is an open-source Python library for solving **Imaging Inverse Problems**
- Special emphasis on tomography applications with **challenging data sets**: low-count, non-standard geometries, incomplete, multi-channel.
- **Highly flexible** and modular for different imaging problems: CT, PET, MRI
- **Modular design for Plug and Play Stochastic optimisation**
- **Simple** to get started – **powerful** enough for large, real applications

• Website	https://www.ccpi.ac.uk/CIL
• Docs	https://tomographicimaging.github.io/CIL
• Discord	https://discord.gg/kmBcU2kebB
• Contact	epapoutsellis@gmail.com

Thank you! Questions?

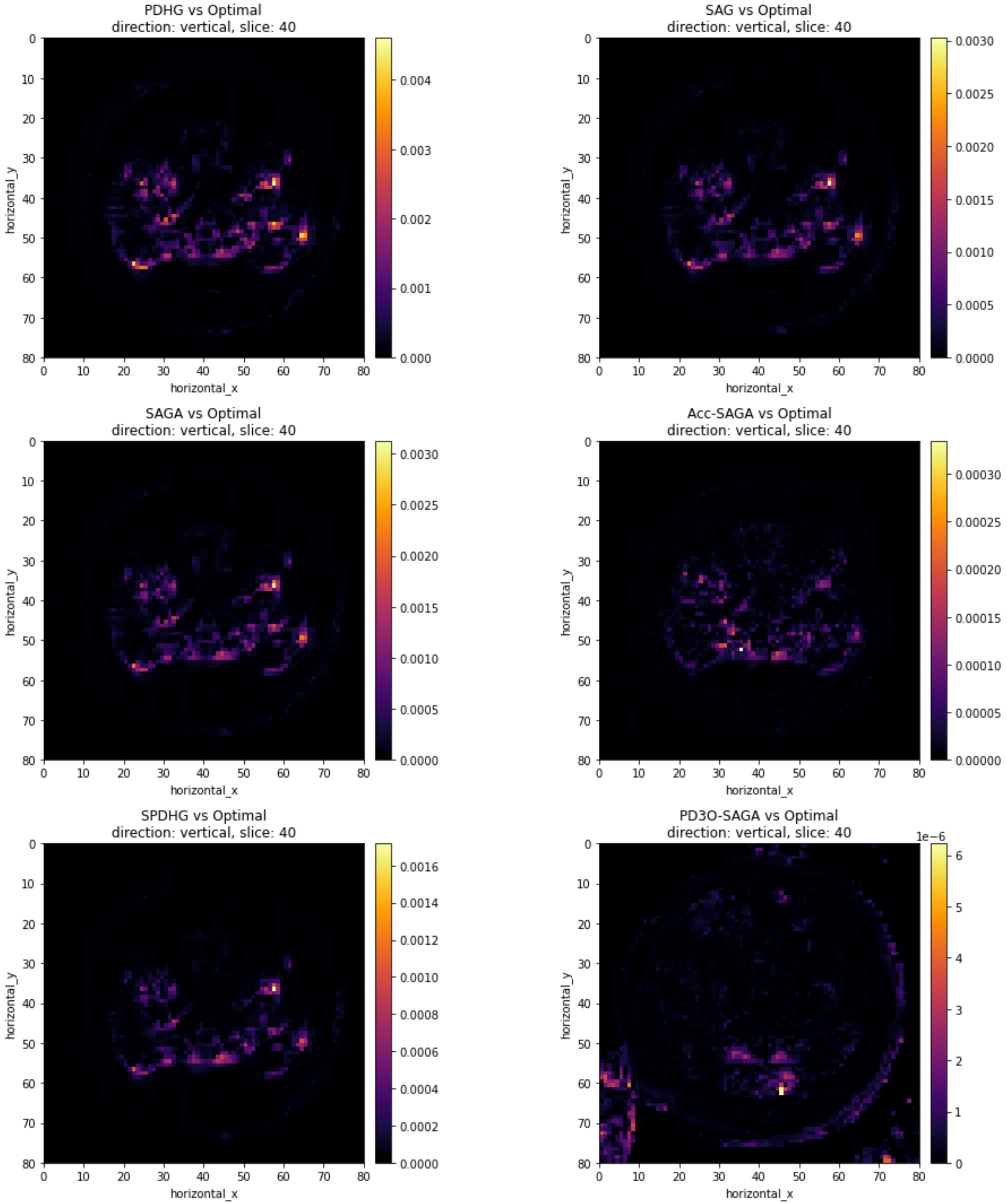
References

- Jørgensen. et al. 2021, *Core Imaging Library - Part I: a versatile Python framework for tomographic imaging*
- Papoutsellis et al. 2021, *Core Imaging Library - Part II: multichannel reconstruction for dynamic and spectral tomography*
- Ovtchinnikov et al. 2017, *SIRF: Synergistic Image Reconstruction Framework*
- Brown et al. 2021: *Motion estimation and correction for simultaneous PET/MR using SIRF and CIL*
- Defazio et al. 2014, *SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives*
- Schmidt et al. 2017, *Minimizing finite sums with the stochastic average gradient*
- Gorbunov et al. 2020, *A Unified Theory of SGD: Variance Reduction, Sampling, Quantization and Coordinate Descent*
- Salim et al. 2022, *Dualize, Split, Randomize: Toward Fast Nonsmooth Optimization Algorithms*

References

- Yan, 2018, *A new primal-dual algorithm for minimizing the sum of three functions with a linear operator*
- Salim, 2019, *On the performance of the stochastic FISTA*
- Driggs et al. 2020, *Accelerating variance-reduced stochastic gradient methods*
- Yurtserver et al. 2016, *Stochastic Three-Composite Convex Minimization*
- Zhao et al. 2018, *Stochastic Three-Composite Convex Minimization with a Linear Operator*
- Combettes et al. 2019, *Proximal Activation of Smooth Functions in Splitting Algorithms for Convex Image Recovery*
- Chambolle et al. 2018, *Stochastic Primal-Dual Hybrid Gradient Algorithm with Arbitrary Sampling and Imaging Applications*
- Warr et al. 2021, *Enhanced hyperspectral tomography for bioimaging by spatiospectral reconstruction*

Lizard: Absolute difference from optimal



Thorax: Absolute difference from optimal