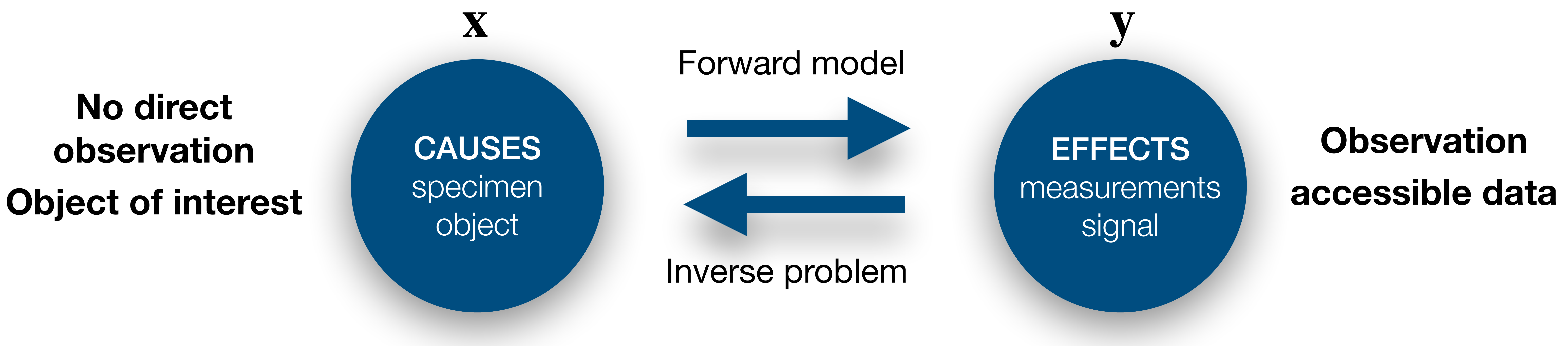
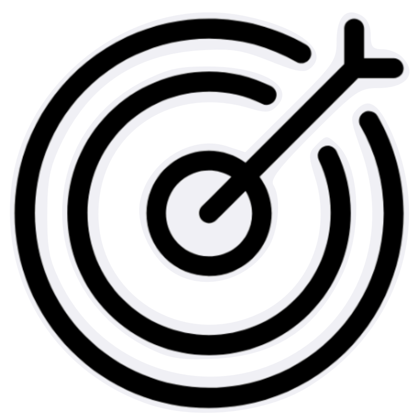
Daniel Sage & Arne Seitz

Course

Introduction to Inverse Problems

Inverse Problems



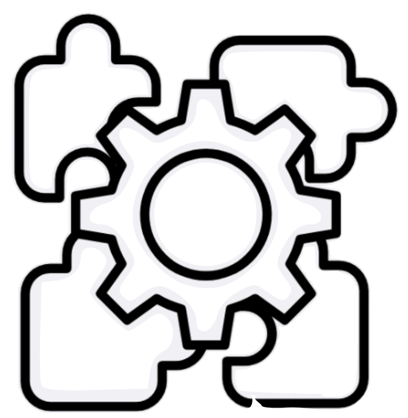


Objective

Find a cause from a consequence

Infer hidden quantities from indirect data

Translate measurements into physics

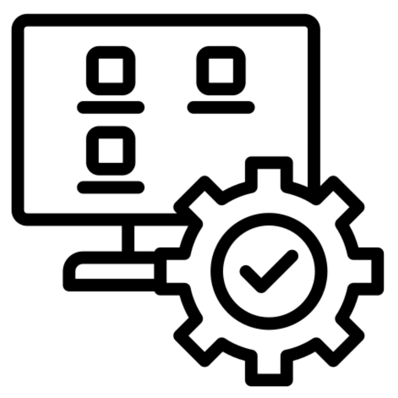


Reconstruction

Numerically recovering signal

Make use of physical model

Optimization for large data



Applications

Medical Imaging: CT, MRI, EEG

Seismology - Nondestructive testing

Microscopy for life science - Depth

Unknowns vs. Measurements

Sufficient measurements

$$\left.\begin{aligned} x_0 + x_1 &= 7 \\ x_0 - x_1 &= 1 \end{aligned}\right\} \Longrightarrow$$ [redacted]

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}$$

$$\mathbf{H}\,\mathbf{x} = \mathbf{y}$$

Well posed problem

Existence

Uniqueness

Continuous dependency on data

What if we had less?

$$x_0 + x_1 = 7$$

$x_0 - x_1 =$ **unknown**

$$(-1,8)\ldots(3,4),(4,3)\ldots(10,17)\ldots$$

➡ **Add prior knowledge**
e.g difference of A and B is small

$\Longrightarrow$ [redacted]

➡ **Add hard constraint**
e.g. A is larger than B

$\Longrightarrow$ [redacted]

➡ **Learn from previous experiences**

noisy me? model mismatch?

$$x_0 + (1 \pm \epsilon)x_1 = 7 \pm \alpha$$

$$x_0 - (1 \pm \epsilon)x_1 = 1 \pm \alpha$$

Real life: physical world

Ill posed problem

measures << # unknowns

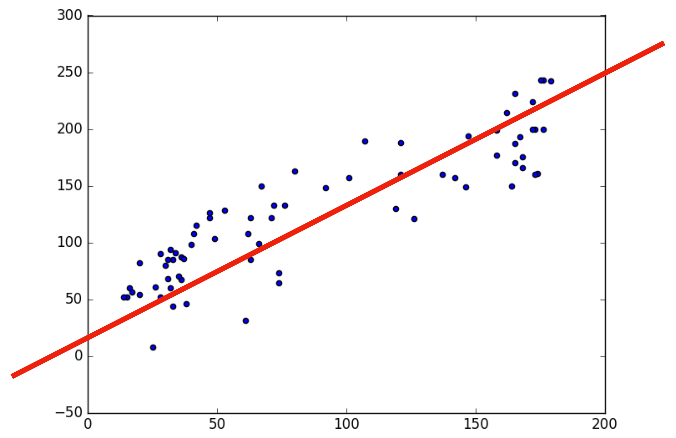
Degraded measures, noise

$\mathbf{H}$ approximation of a physical device

High-Dimensional Inverse Problem

Overdetermined Problem

Close-form solution



Observation (y_1, x_1) (y_2, x_2) ... (y_N, x_N)

Forward Model $\mathbf{y} = a\,\mathbf{x} + b$

Least-square solver

$$\xi = \sum (y_i - a x_i - b)^2$$

Underdetermined Problem

ξ **to minimize** $\quad \tilde{x} = \text{argmin}(\xi)$

Objective: Energy, criteria, loss, error, cost

Variational Optimizer

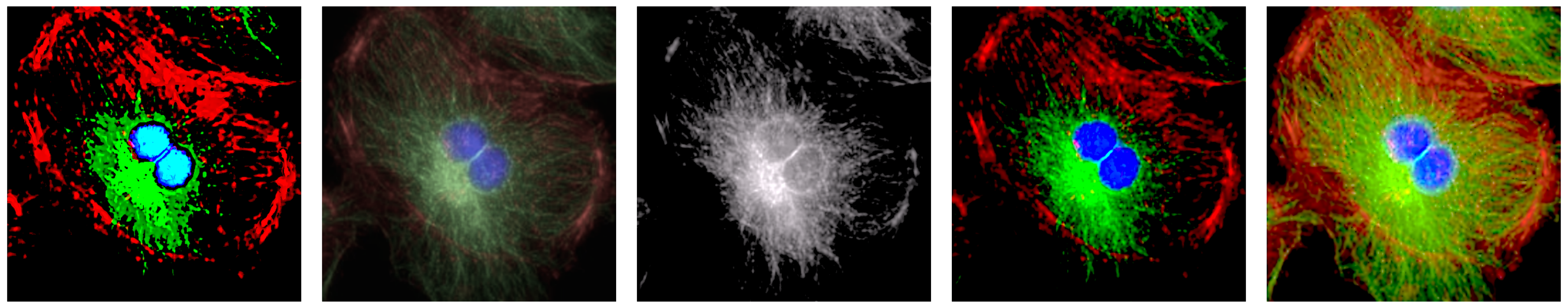
Gradient descent optimization

Variational solvers may converge to local minima

$\mathbf{x}$

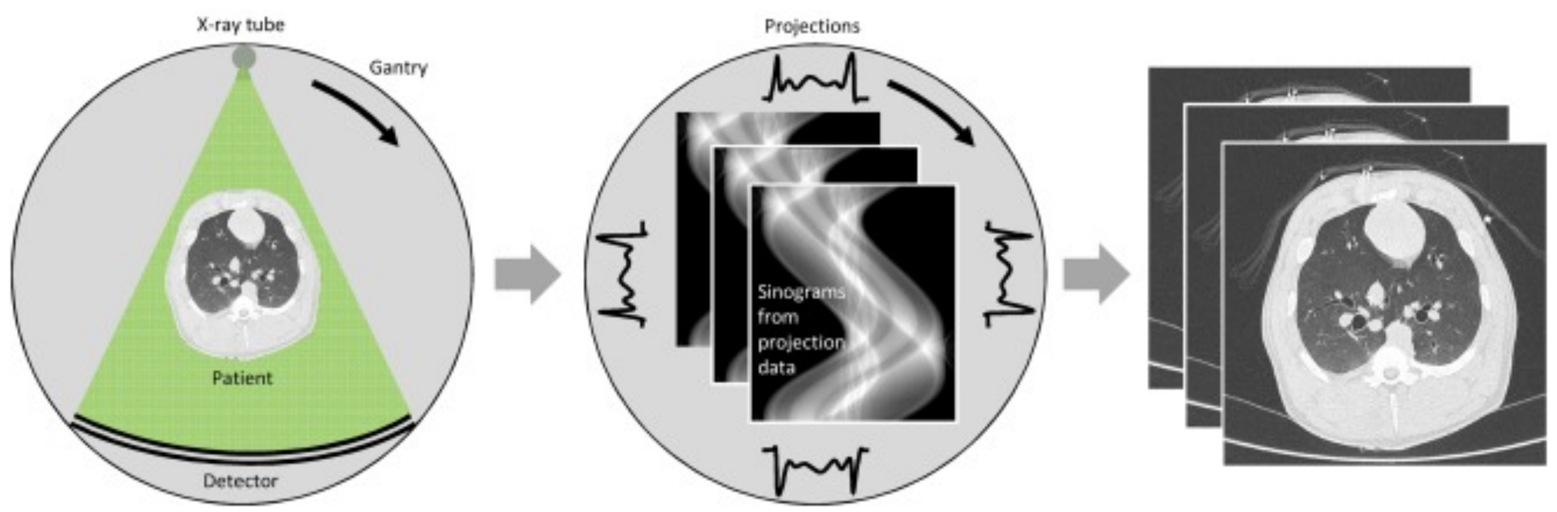
$\mathbf{y}$

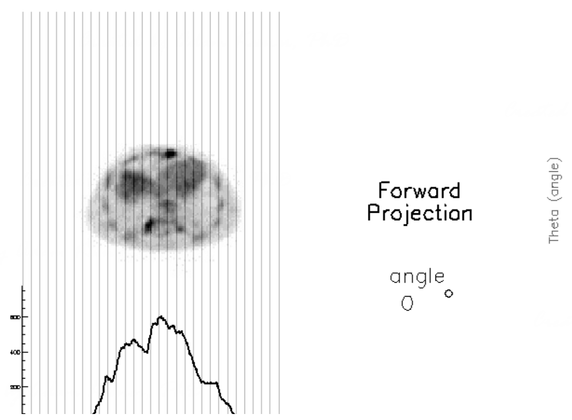
$\hat{\mathbf{x}}$



Unified Framework Across Imaging Modalities

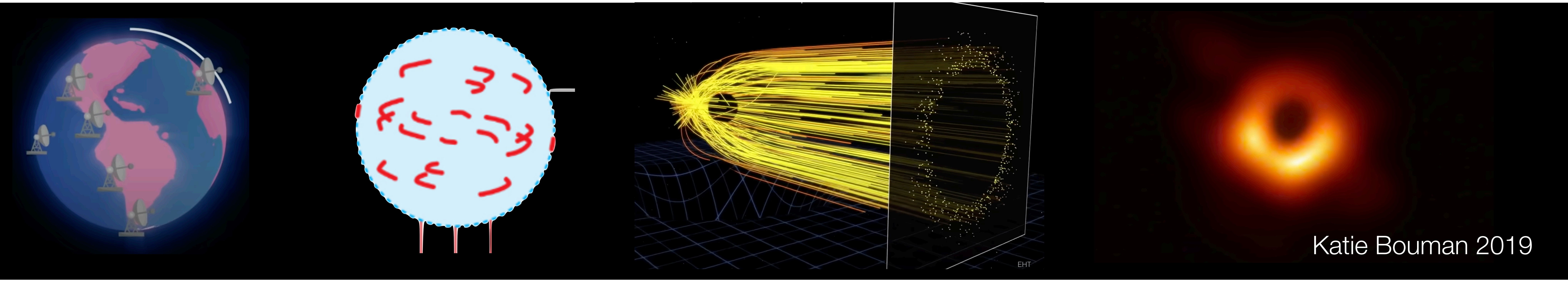
Computed Tomography
Filter-back projection





Sinogram

Event Horizon Telescope
Black Hole Imaging



Katie Bouman 2019

Single Particle Analysis
TEM
Cryo EM

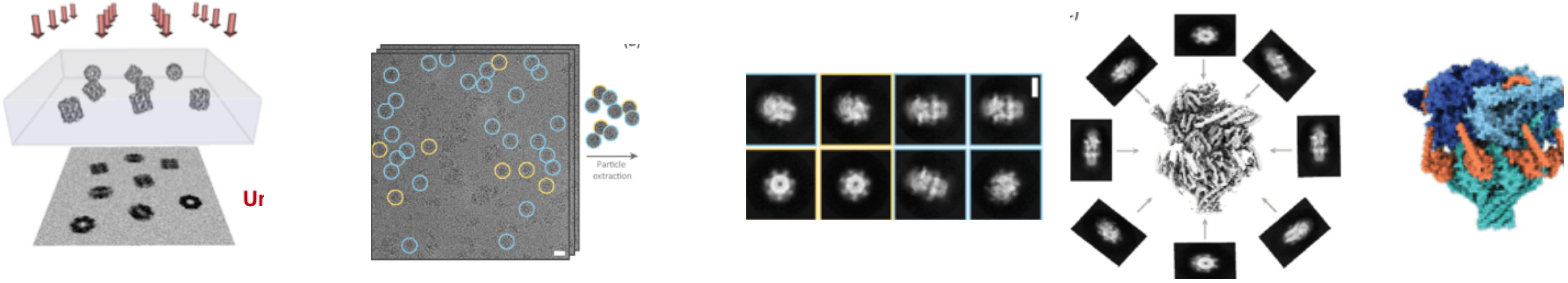
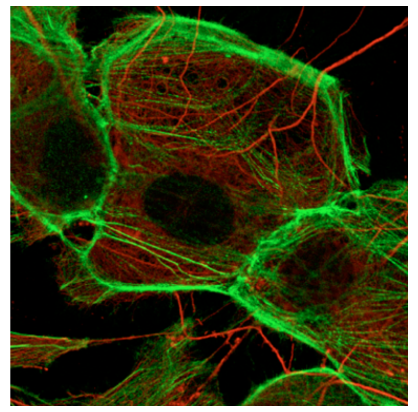
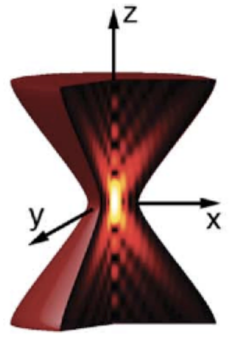
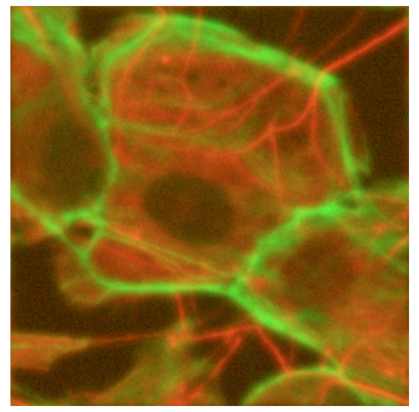


Image Formation Operators

Convolution







Simplified example of matrix representation

A1	A2	A3
B1	B2	B3
C1	C2	C3

	1	
1	1	1
	1	

Toeplitz matrix

A1
A2
A3
B1
B2
B3
C1
C2
C3

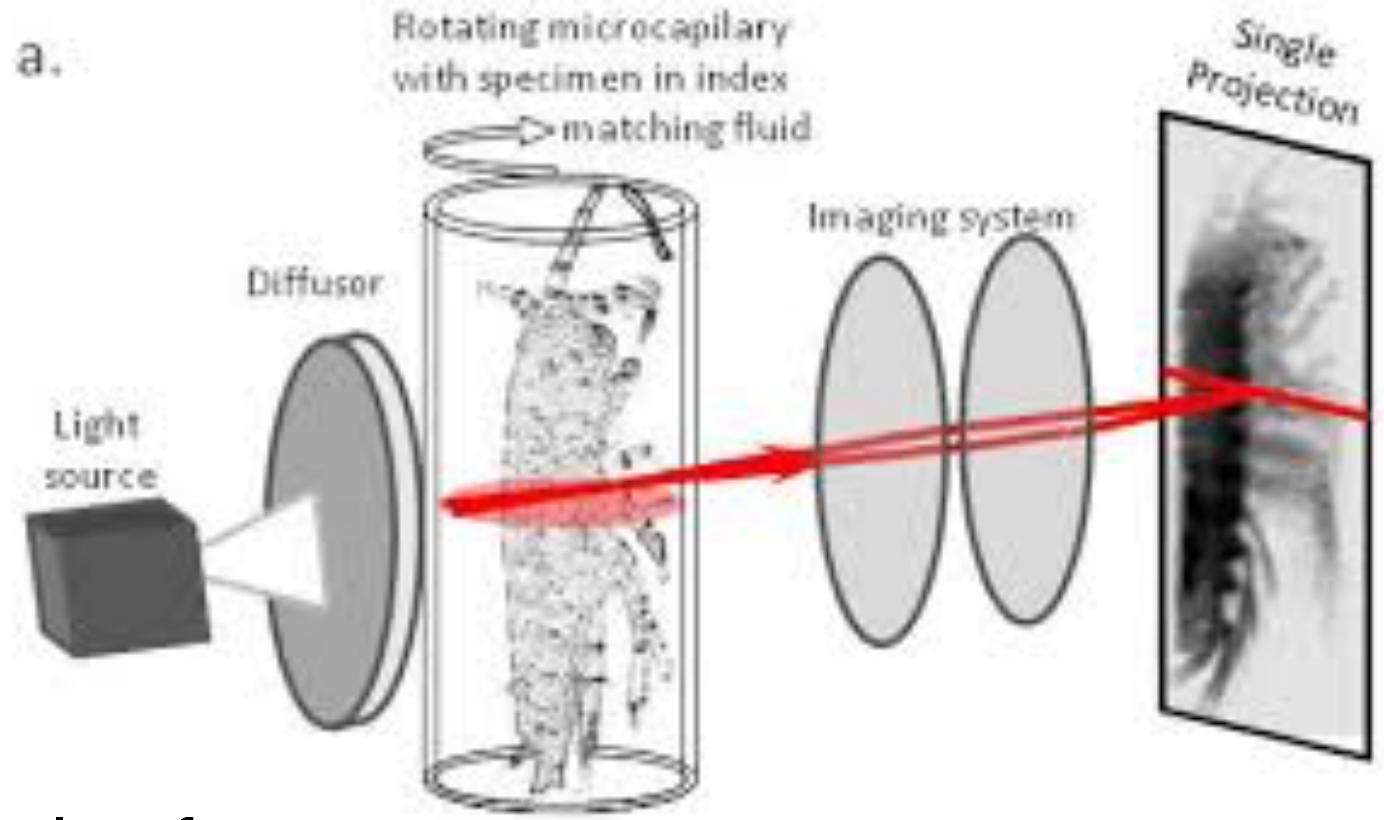
=

1	1		1					
1	1	1		1				
	1	1	1		1			
1		1	1	1		1		
	1		1	1	1		1	
		1		1	1	1		1
			1		1	1	1	
				1		1	1	1
					1		1	1

A1
A2
A3
B1
B2
B3
C1
C2
C3

- Usually the matrix is very large
- The operator is a circular convolution

Projection



Simplified example of matrix representation

a11	a12
a21	a22

m11
m12

a11	a12
a21	a22

m21	m22

m11
m12
m21
m22

=

1	1		
		1	1
1		1	
	1		1

a11
a12
a21
a22

- Usually the matrix is very large
- The operator is the Radon transform

Image Formation in Microscopy

3D Deconvolution

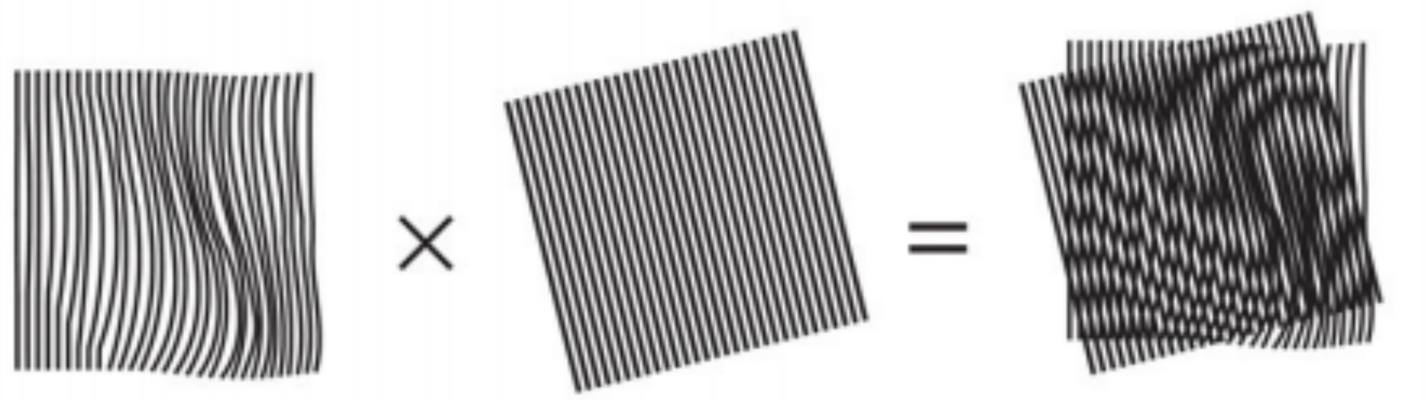
$\mathbf{H}$ is a circulant matrix (PSF)

Structured Illumination Microscopy

SIM

$\mathbf{H} = \mathbf{C} \circ \mathbf{M}$

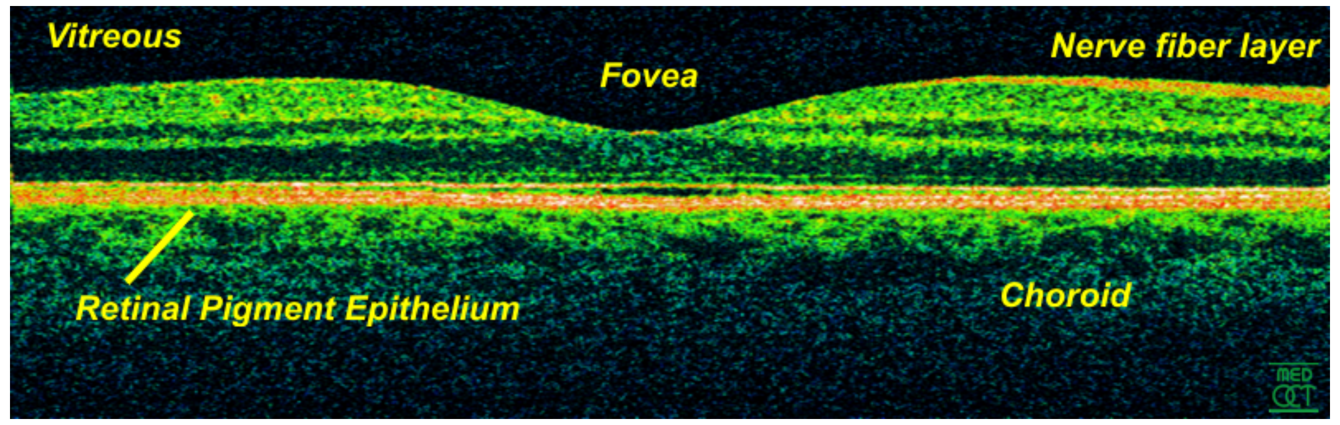
- $\mathbf{C}$ is a convolution (PSF)



Optical Coherence Tomography

OCT

Measures intensity of back-reflected light



Denoising

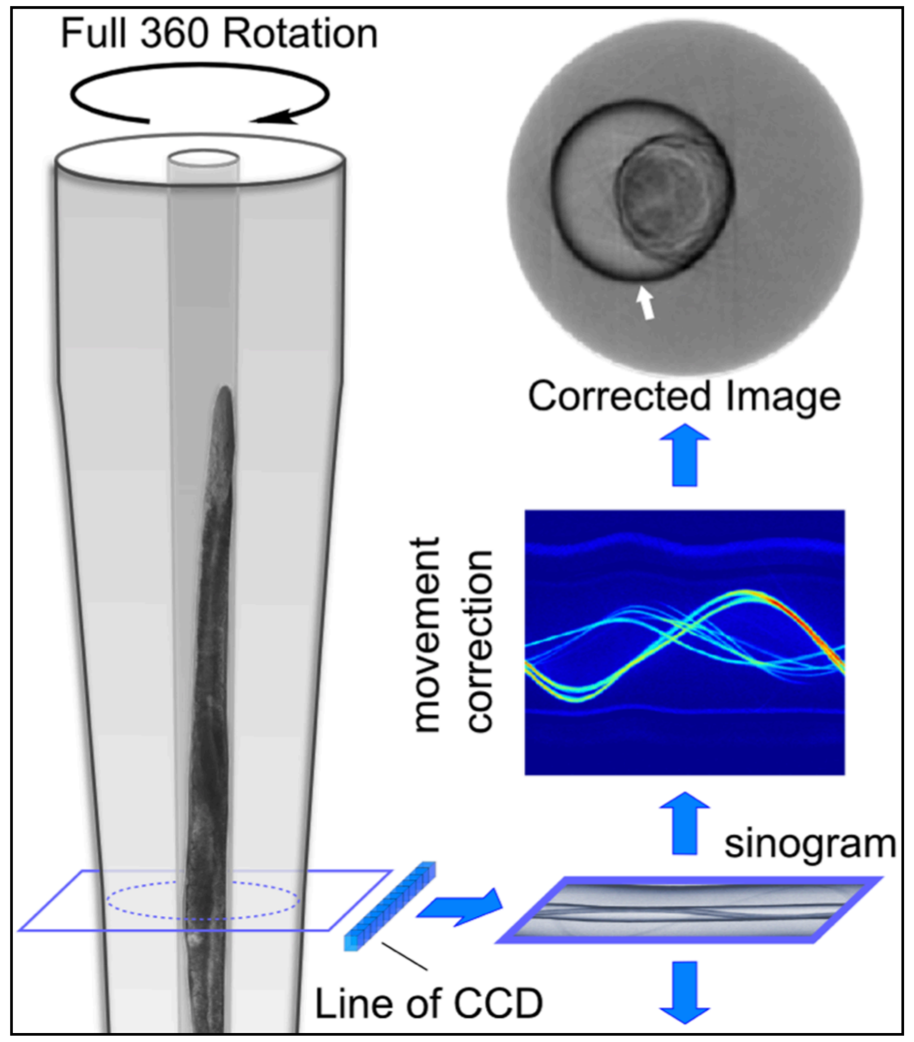
$\mathbf{H}$ is an identity

Optical Projection Tomography

OPT

$\mathbf{H} = \sum \circ \mathbf{R}_\theta$

- $\sum$ is an integration
- $\mathbf{R}_\theta$ is a rotation

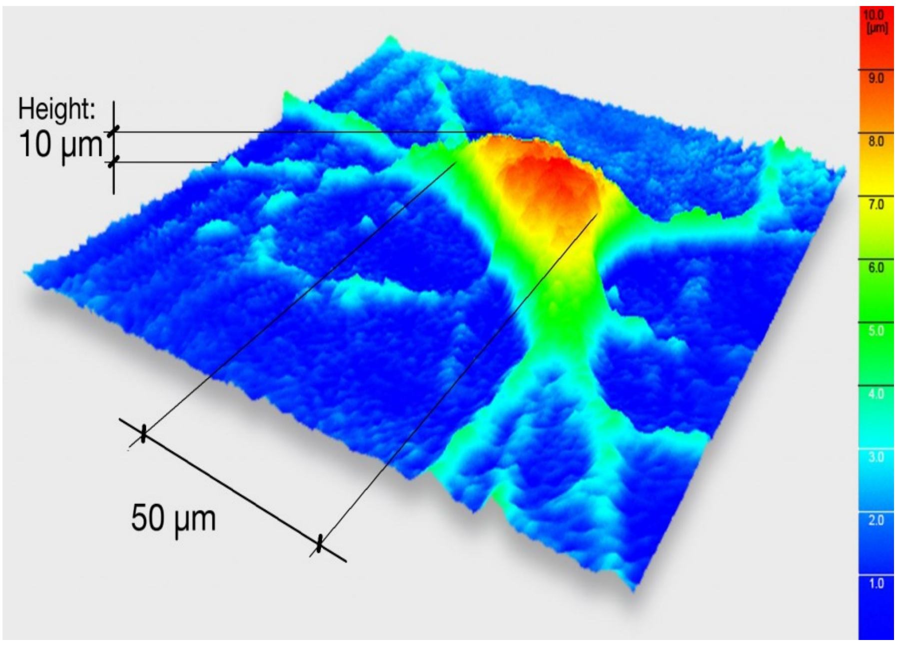


M. Rieckher, PLOS one 2017

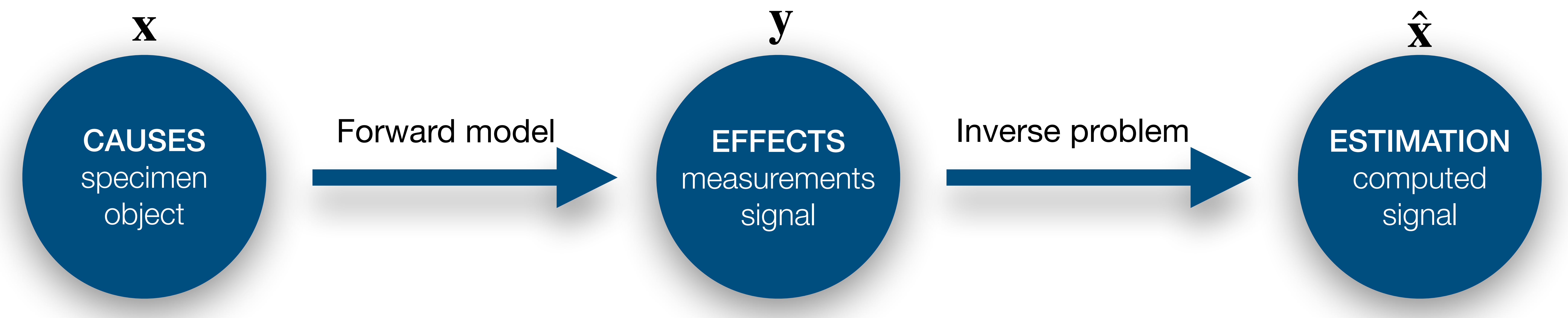
Digital Holographic Microscopy

DHM

- Phase interference registered on a hologram
- Forward model: propagation of coherent light



Inverse Problems in Imaging



Real measures

degraded, noise
partial measurements
non-directly interpretable
non-usable as image

Numerical solving

approximative forward model
estimation noise model
prior on the solution
many unknowns

$\mathbf{H}$ Forward Model

Perfect model $\mathbf{y} = \mathbf{H}\mathbf{x}$

$\mathbf{H}^{-1}$ is never possible

Real model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 \right\}$$ **No regularization**

$$\|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 + \lambda R(\mathbf{x})$$ **With regularization**

Inverse Problem in Practice

$$\tilde{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \mathscr{D}(\mathbf{H}\mathbf{x}, \mathbf{y}) + \lambda \mathscr{R}(\mathbf{x}) \right\}$$

Data Fidelity Term

Objective
Noise model

$\mathscr{D}$

$\mathbf{H}$

Forward Model

Physical process modeling
Fidelity and consistency?
Linear?

Optimizer

$\mathscr{R}$

Regularizer Term

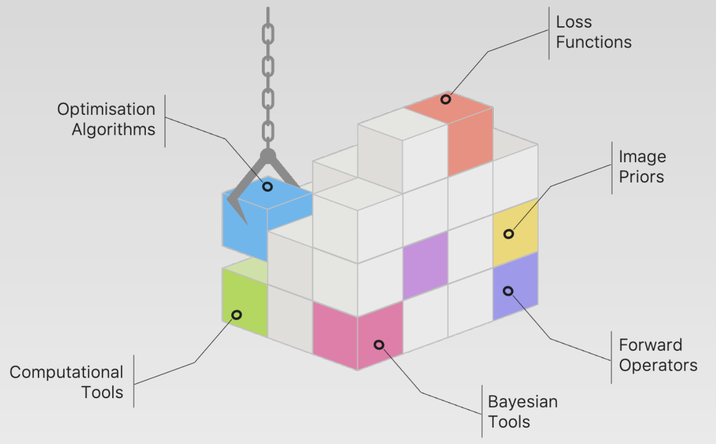
Encourage smooth or piecewise
Promote sparsity
Data-driven

Computational challenges

- Large memory required
- Iterative algorithms are generally slow
- How to stop the iterative algorithms
- How to set up the hyper parameter?

Software toolbox

- Decoupling forward model, optimizer, cost, regularisation
- Unifying the algorithms for any modality
- **EPFL** Pyxu (Python) https://pyxu-org.github.io/
- **EPFL** GlobalBioIm Library (Matlab) https://biomedical-imaging-group.github.io/



Least-square Solution

Without regularization

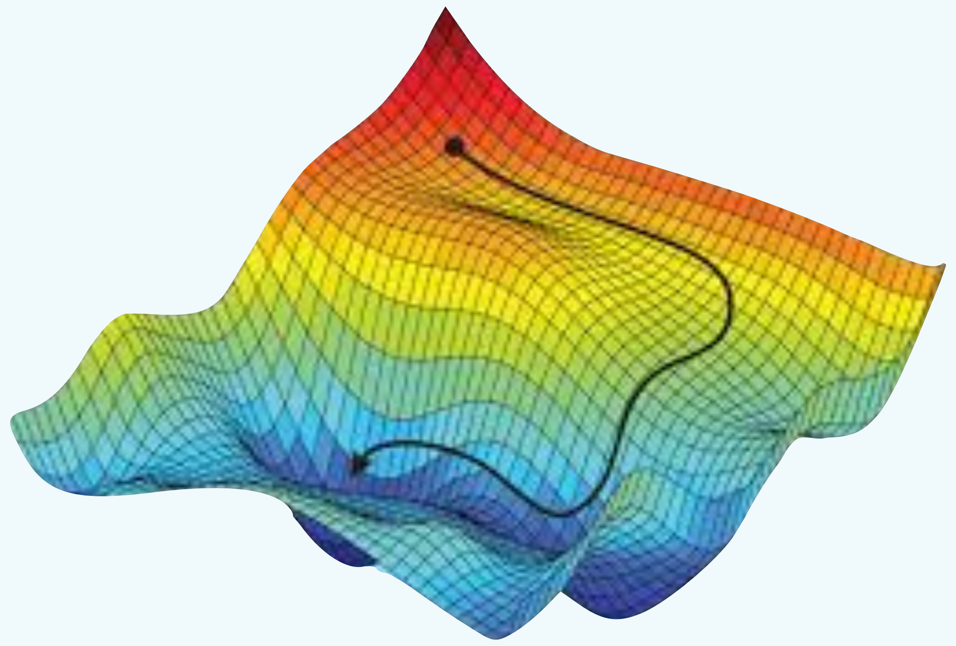
$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 \right\}$$

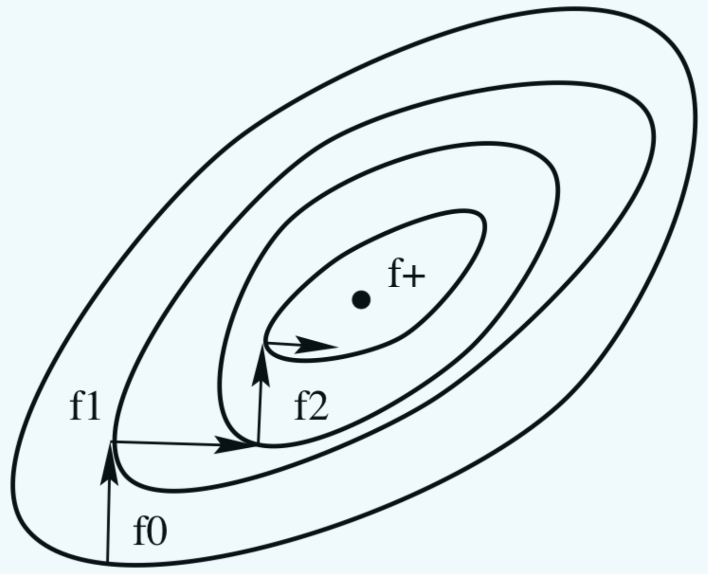
Don't work when it is ill-posed

Iteration of Landweber $\nabla_{\mathbf{x}} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 \right\} = \mathbf{H}^T\mathbf{H}\mathbf{x} - \mathbf{H}^T\mathbf{y}$

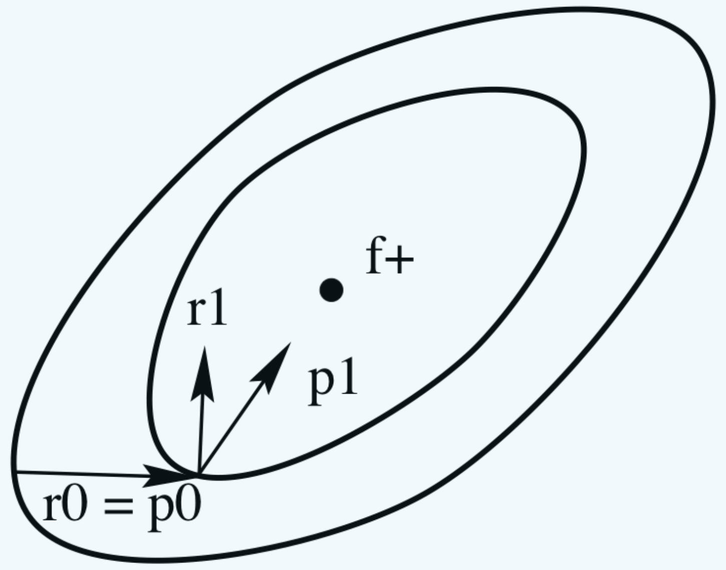
Gradient descent: $\mathbf{r}_k = \mathbf{H}^T(\mathbf{H}\mathbf{x} - \mathbf{y})$

Update rule: $\mathbf{x}_{k+1} = \mathbf{x}_k + \gamma_k \mathbf{r}_k$





Steepest gradient



Conjugate gradient

With regularization

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 + \frac{\lambda}{2} R(\mathbf{x}) \right\}$$

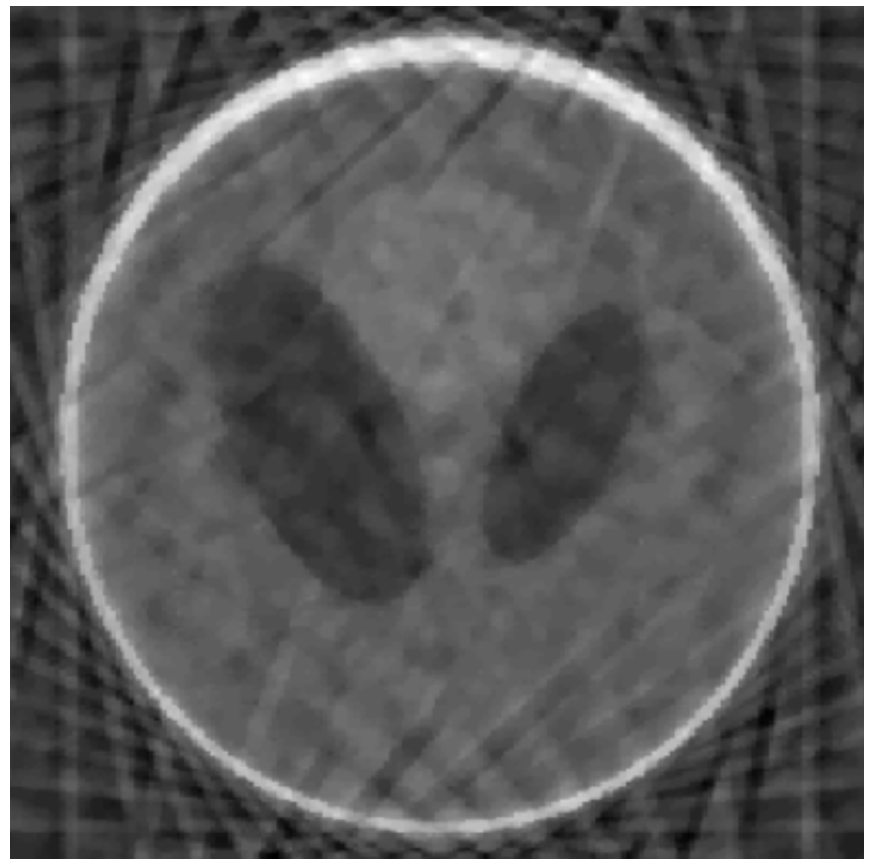
ADMM Alternating Direction Method of Multipliers [Boyd]

Powerful optimization strategy to solve a large problem under a given constraint.

Regularization

$$\tilde{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \mathscr{D}(\mathbf{Hx}, \mathbf{y}) + \lambda \mathscr{R}(\mathbf{x}) \right\}$$

Ground-truth



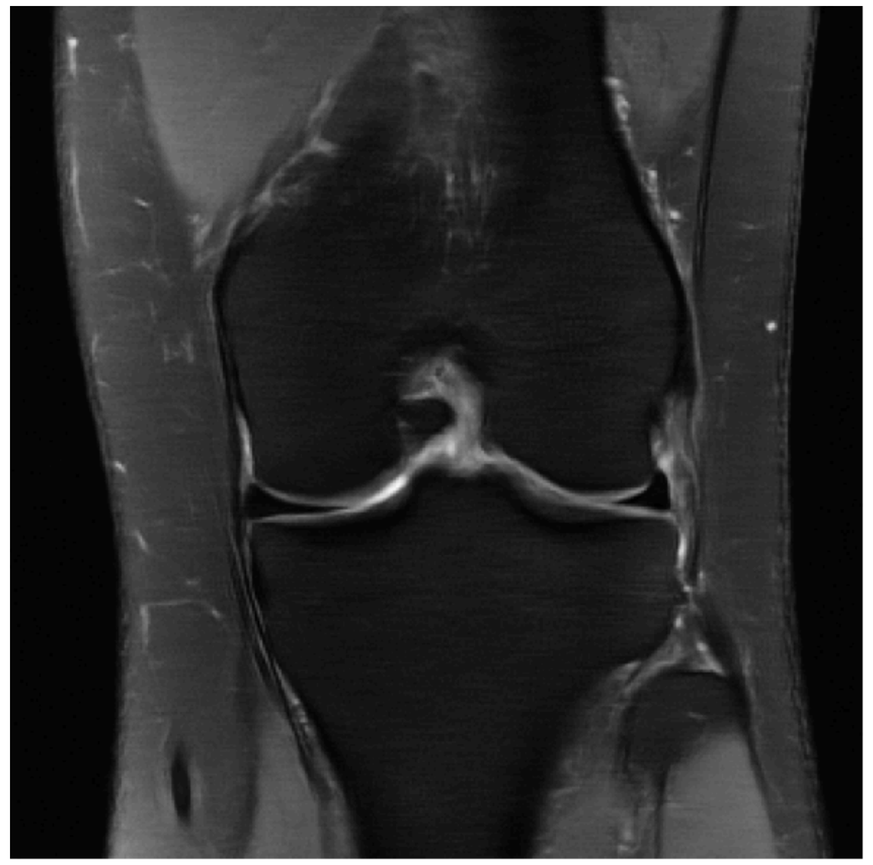
Non regularized

Ground-truth

Non regularized

Regularized solutions

Classic: Image Prior



Learning: Data-Driven

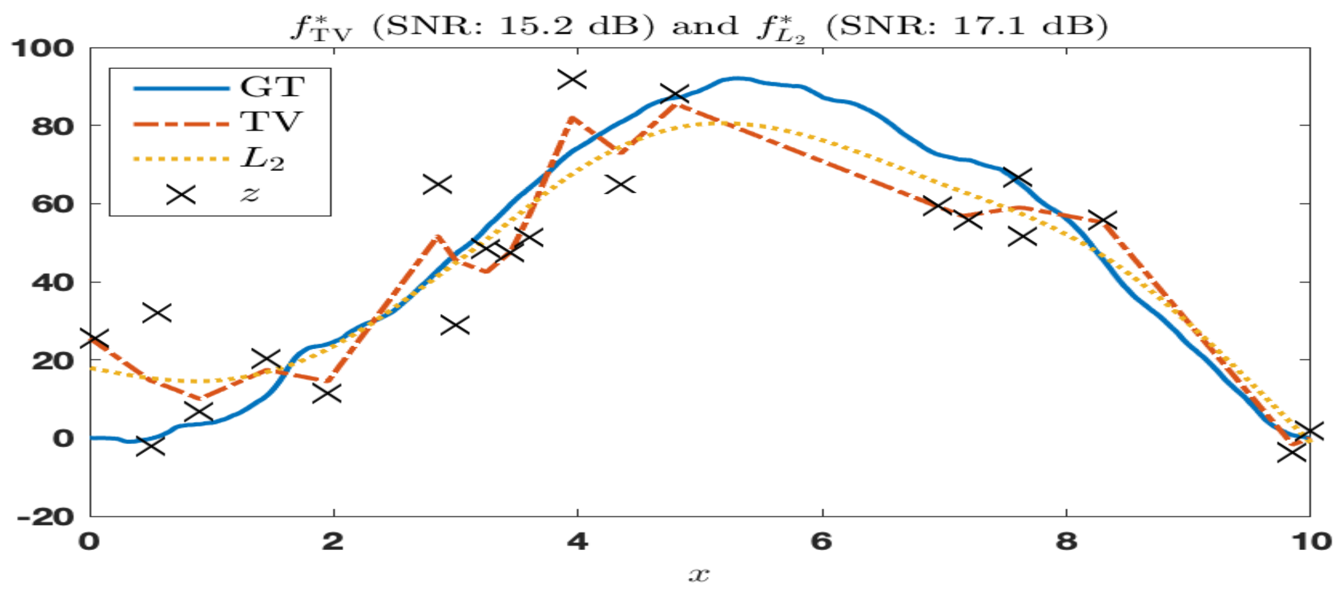
Regularization

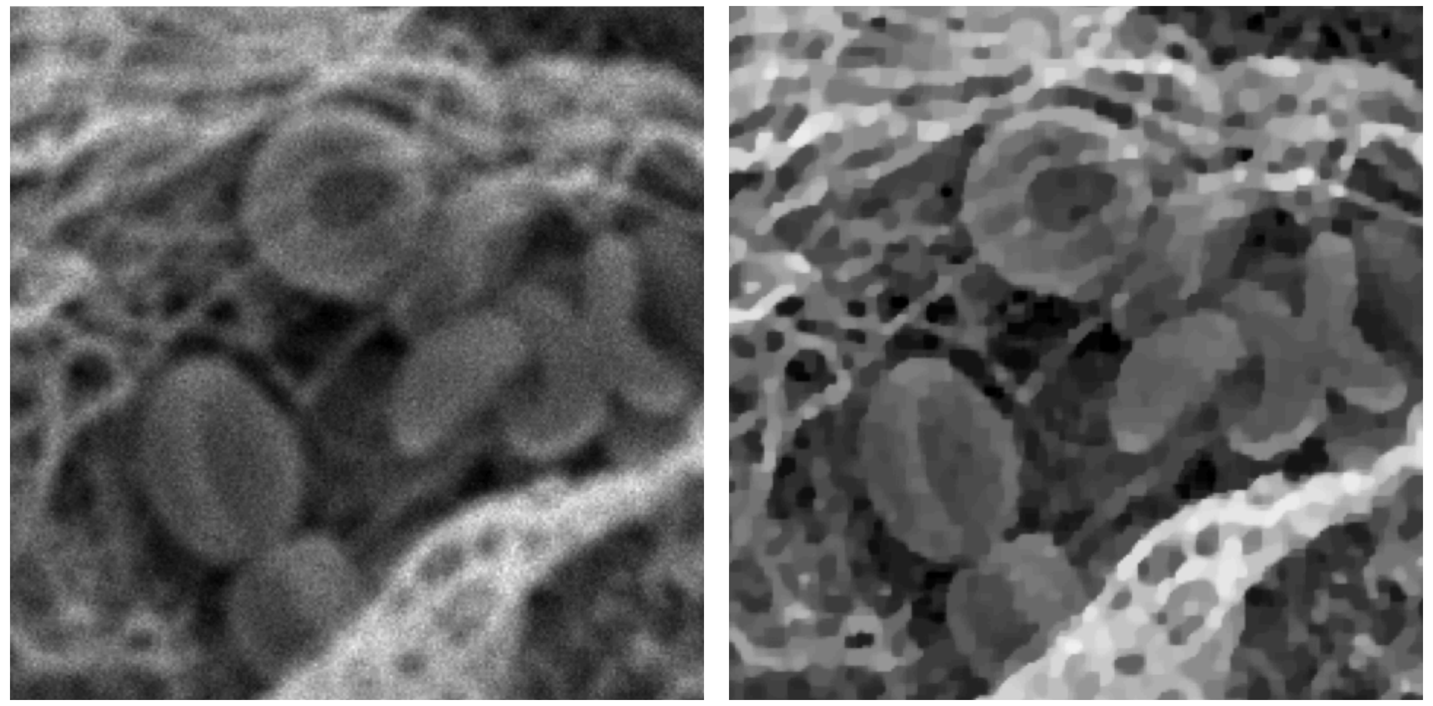
Tikhonov (L2)

$$\mathscr{R}(\mathbf{x}) = \| \nabla \mathbf{x} \|^2$$

$$\mathscr{R}(f) = \| \mathbf{x} \|^2$$

Favorize smooth solutions



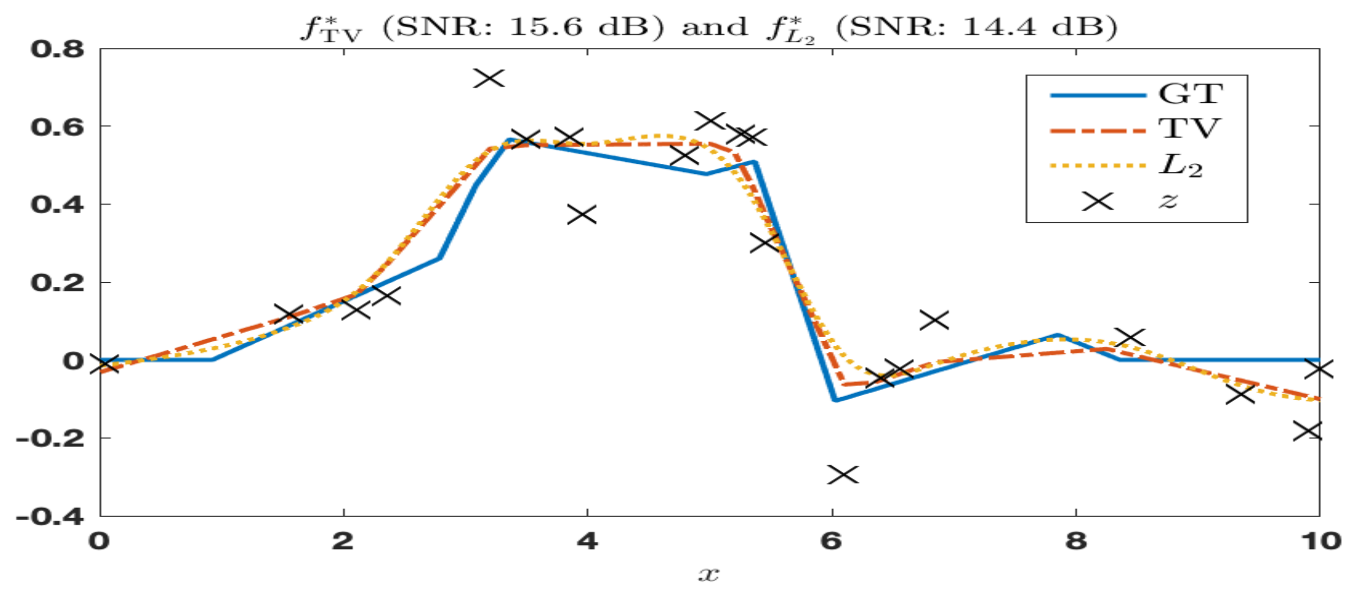


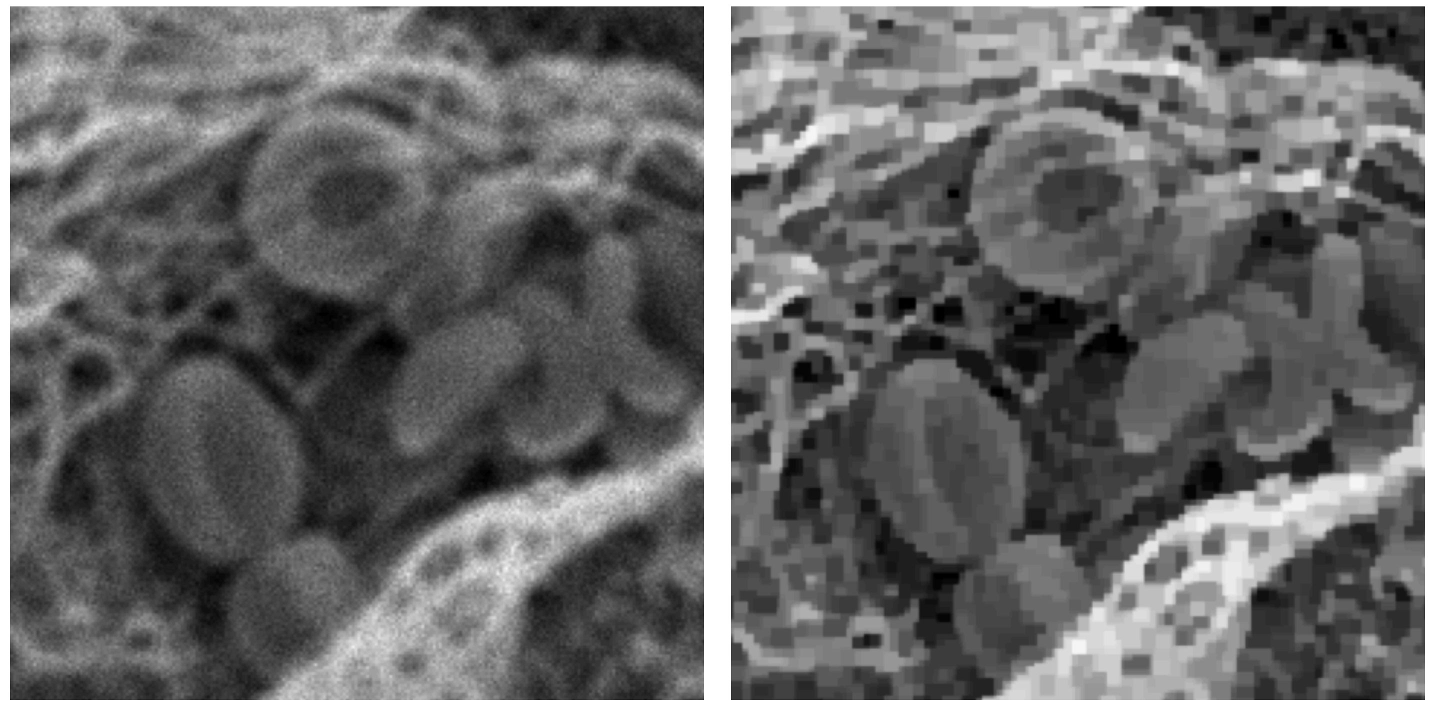
Total Variation (L1)

$$\mathscr{R}(\mathbf{x}) = \| \nabla \mathbf{x} \|_1$$

$$\mathscr{R}(\mathbf{x}) = | \nabla_x \mathbf{x} | + | \nabla_y \mathbf{x} |$$

Favorize piecewise constant solutions





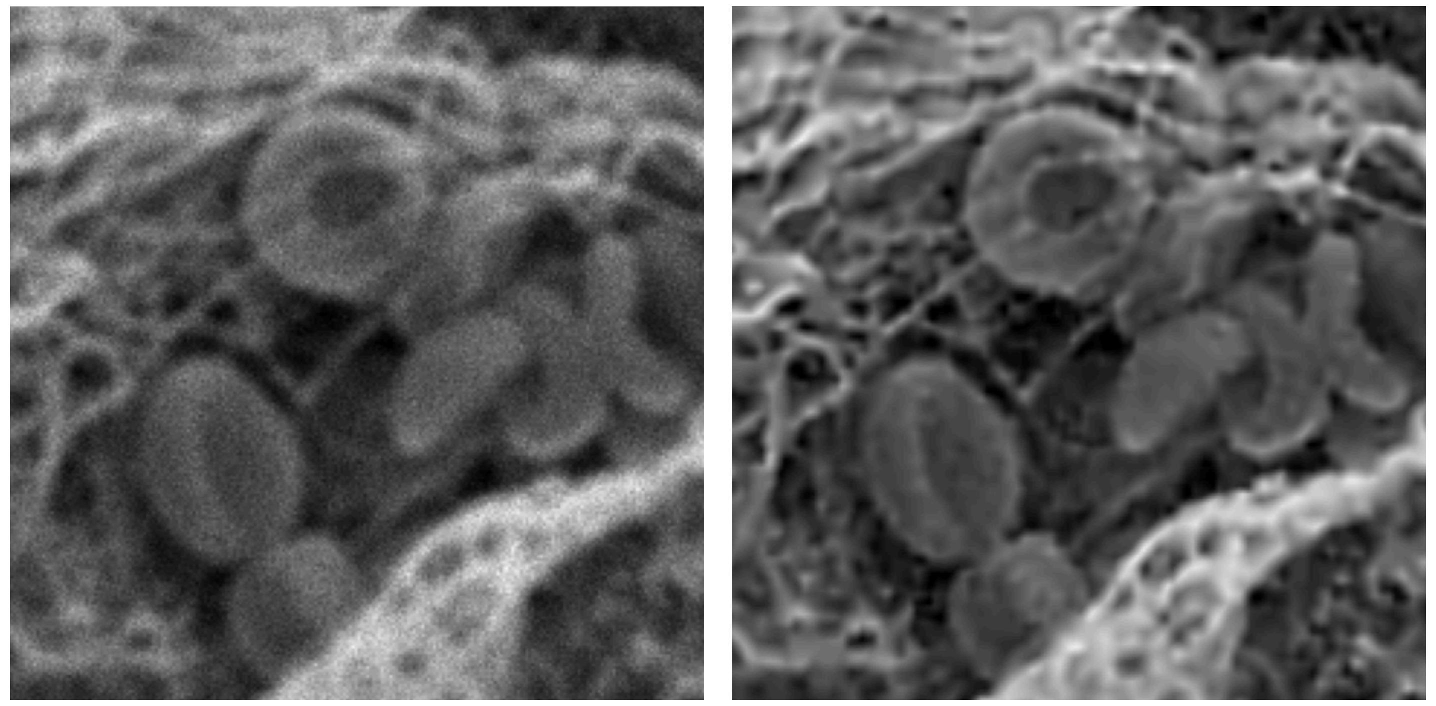
Hessian Schatten-Norm

$$H = \begin{bmatrix} \frac{\partial f^2}{\partial x} & \frac{\partial f^2}{\partial x \partial y} \\ \frac{\partial f^2}{\partial x \partial y} & \frac{\partial f^2}{\partial y} \end{bmatrix}$$

Favorize thin structure solutions

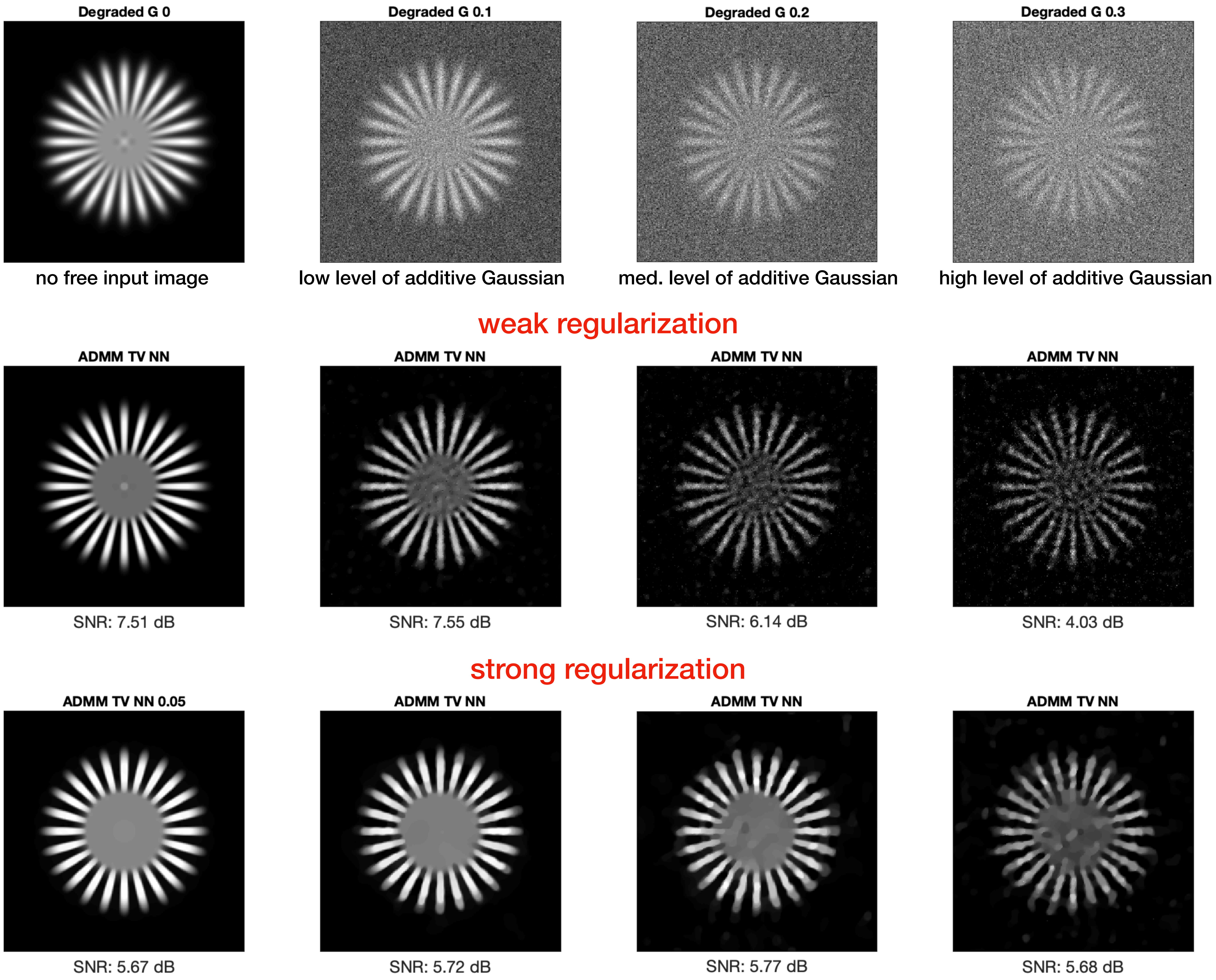
$$T = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$R(f) = \| f \|_{\mathcal{S}} = \mathrm{Tr}(T)$$



Regularization

Input



$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 + \frac{\lambda}{2} R(\mathbf{x}) \right\}$$

Tune the regularization to the noise level

Regularization

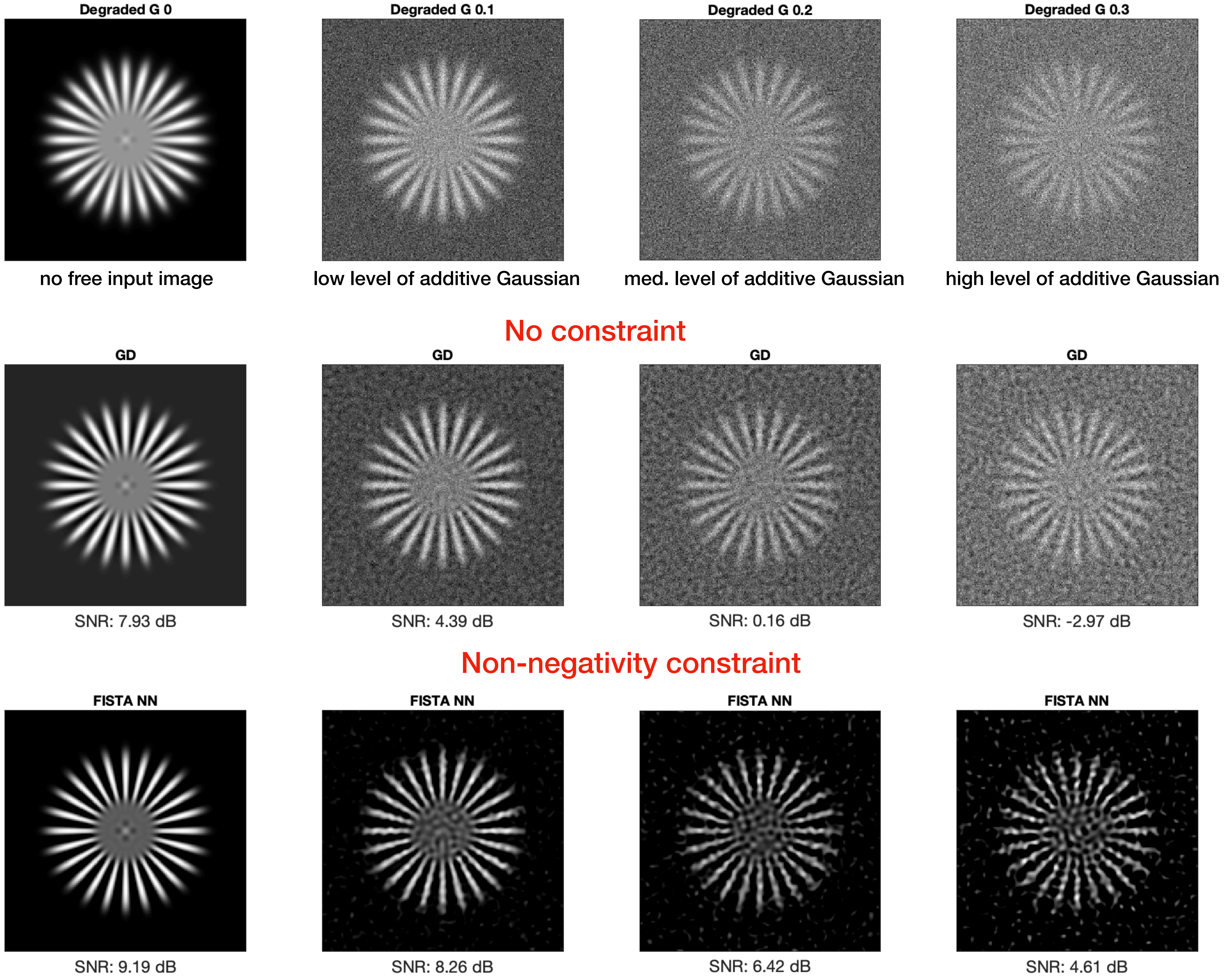
Implicit regularization

Non-negativity (positivity)

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmin}} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 + \lambda R(\mathbf{x}) + i_{\geq 0}(\mathbf{x}) \right\}$$

Early stop

- Stop the iterations before convergence
- Tricky regularization: complicated to controlled



Solving Inverse Problem with Deep Learning

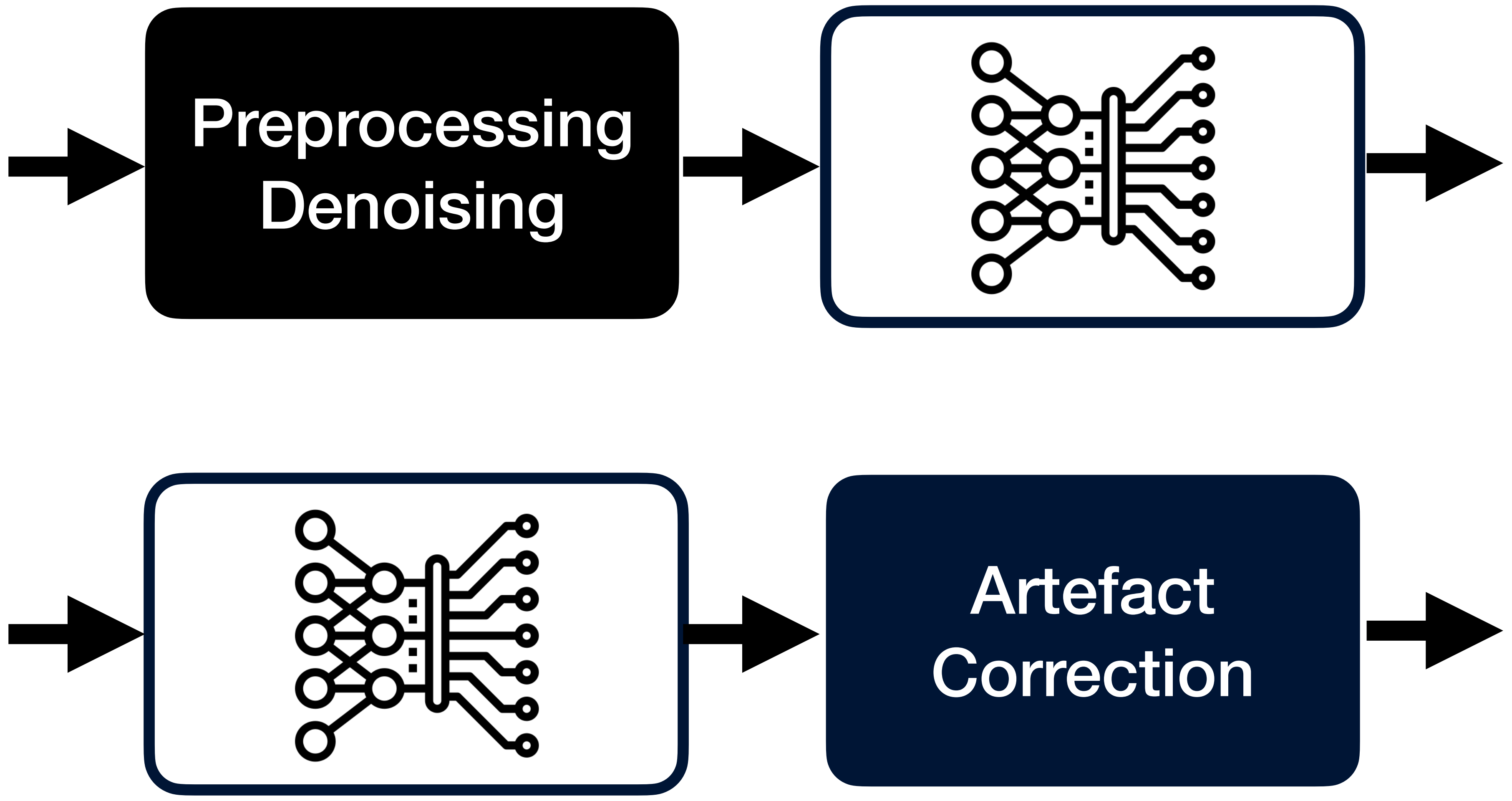
Motivation

- DL can learn complex priors from data
- Handles noise, incomplete data, nonlinearities
- Once it is trained ➜ fast reconstruction (non iterative)

Learn

- PSF
- Hyper-parameter
- Image transform

Image-to-image



End-to-end Mapping

Train a neural network to directly map measurements to images (e.g. Unet)

No ground-truth data
Don't rely on the physical model

Physics-Informed Learning

Plug-and-Plug Priors (PnP)

Plug in a learned denoiser in the optimizer
[Venkatakrishnan 2013, Hurault 2022, Goujon 2024]

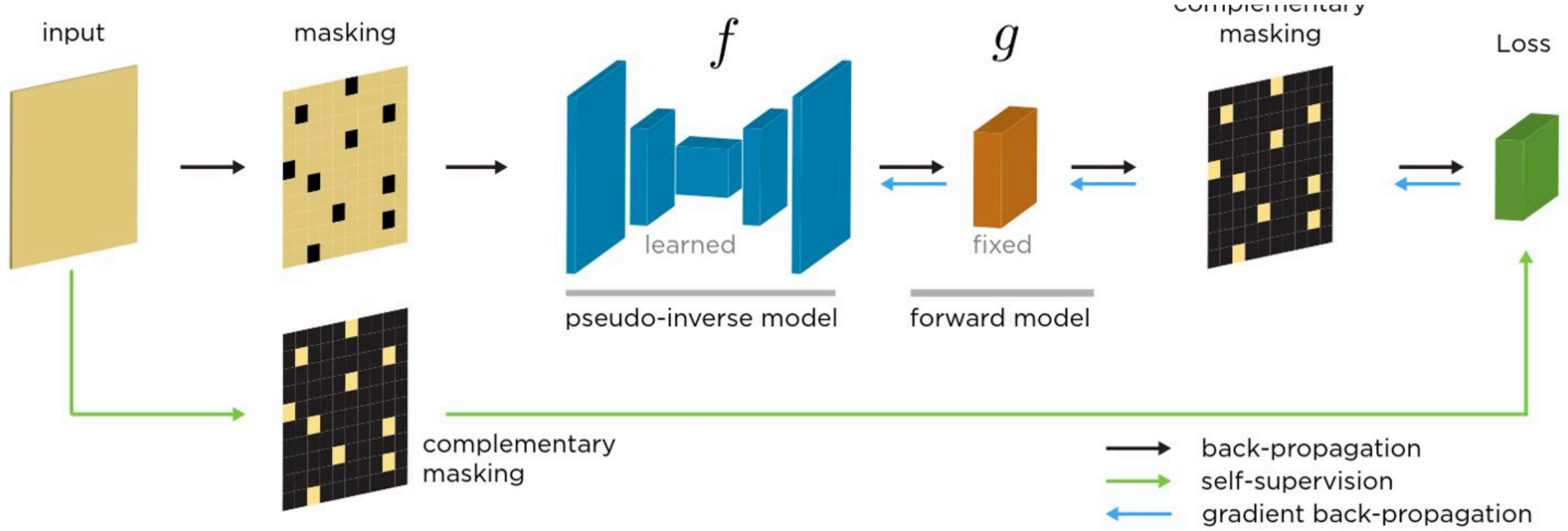
$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \{\mathscr{D}(\mathbf{Hx}, \mathbf{y}) + \lambda \mathscr{R}(\mathbf{x})\}$$

Data fidelity — Regularization

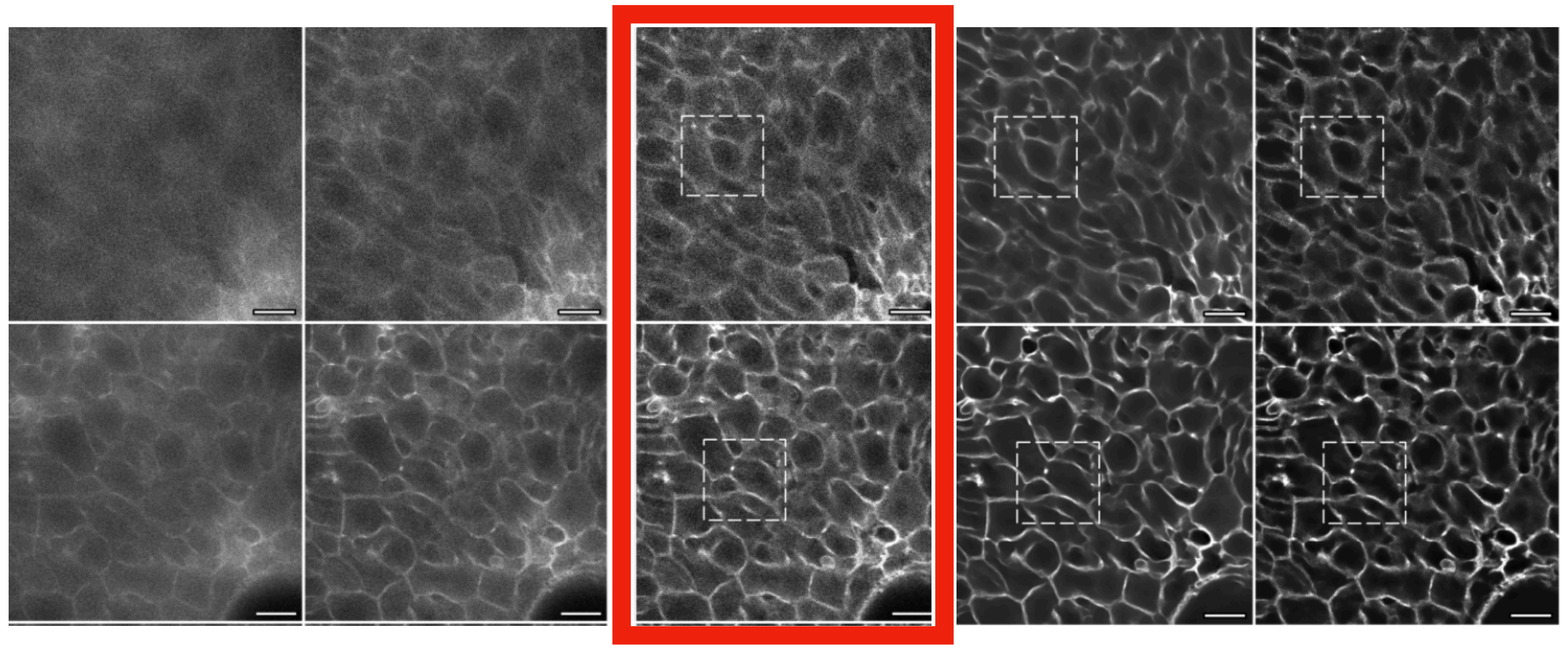
Prior Model
Data-Driven

Self-Supervised Inversion [Kobayashi 2020]

Know forward model - Assumes statistical independence of noise

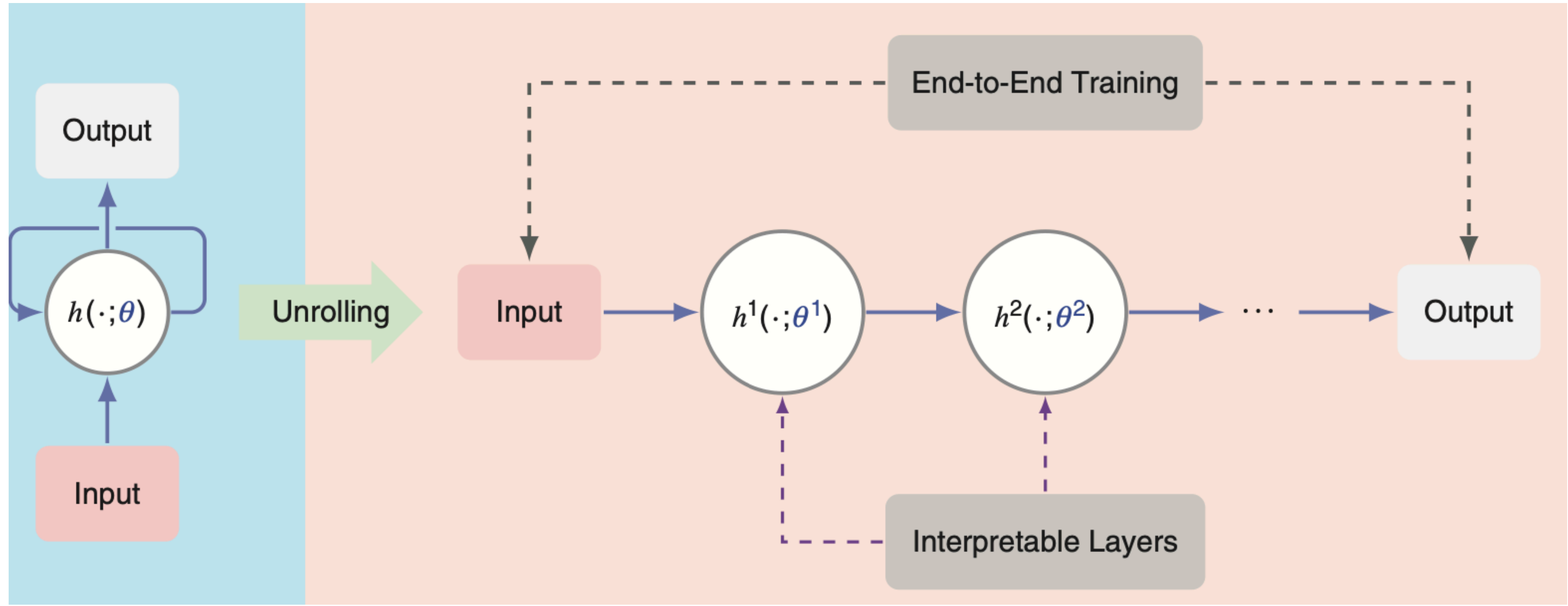


Learn from Degradation [N. Pimpão Martins, 2023]



2x Degraded | 1x Degraded | Raw | 1x Prediction | 2x Prediction

Deep Unrolling / Unfolding [Gregor &LeCun 2010, Monga 2021]



Unroll an iterative algorithm into a NN with each layer corresponding to an iteration.

More interpretable (e.g. ADMM-Net)