Inverse Problems and Uncertainty Quantification

Per Christian Hansen
Professor, Villum Investigator
Section for Scientific Computing

Research question

How to make UQ a general and easy-to-use tool for inverse problems

What it's like to do research

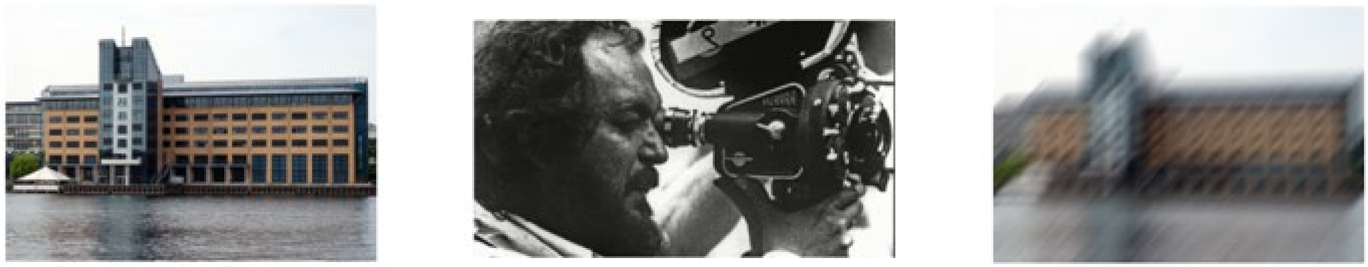
Heian Shrine, Kyoto

DTU Compute
Department of Applied Mathematics and Computer Science

Inverse Problem: Image Deblurring

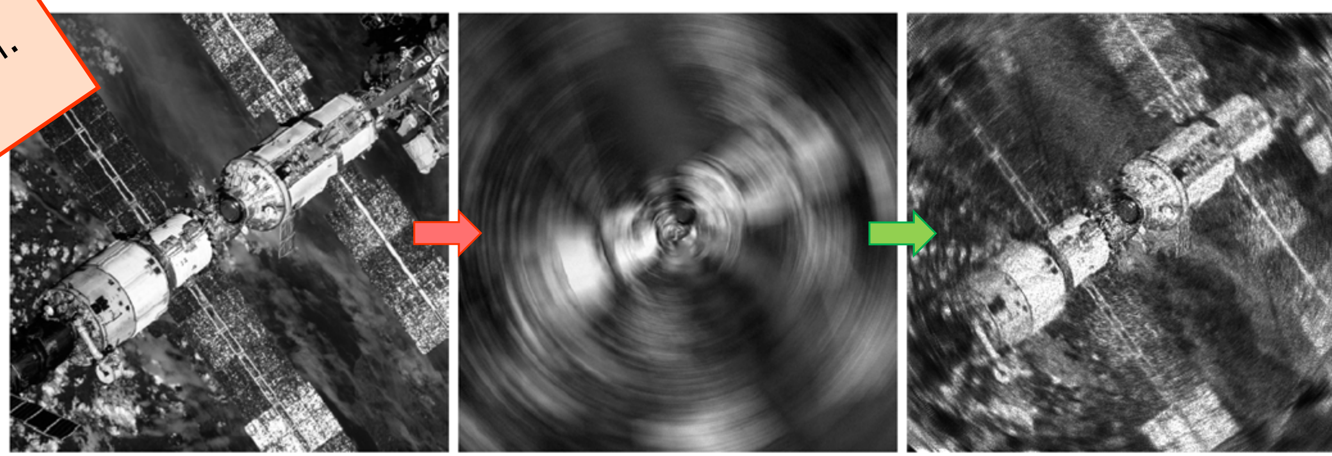
Mathematical model
- Point spread function.

Camera blur.



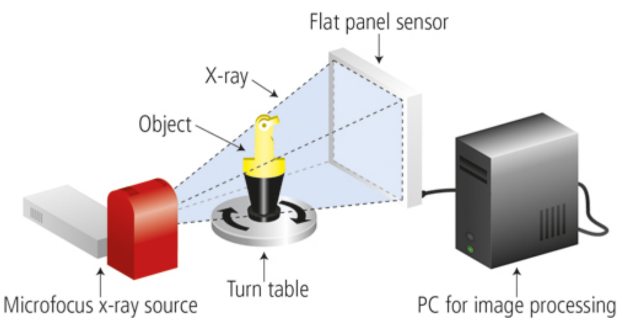
Rotational blur.

Mathematical model
- Center of rotation.
- Rotation angle.



Inverse Problem: X-Ray CT

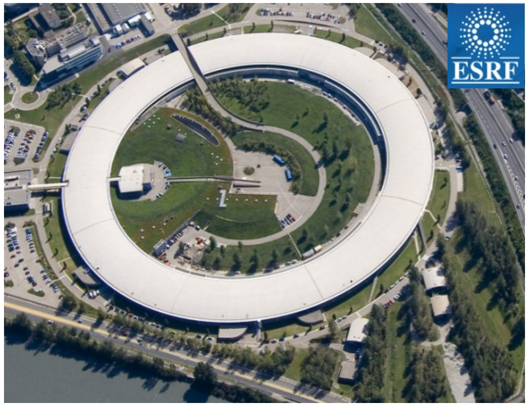
Image reconstruction from measurements of X-ray attenuation in an object..

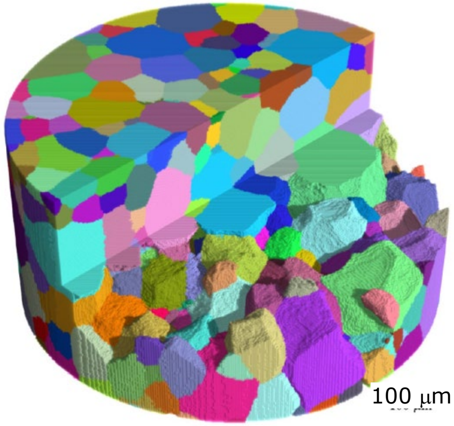


Medical imaging

Matematical model
- Physics of X-ray attenuation.
- Spectrum of X-ray source.
- Specification of the geometry.

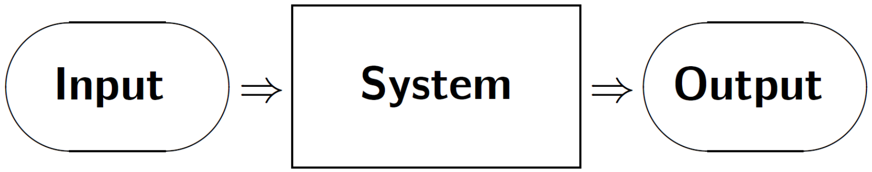
Materials science



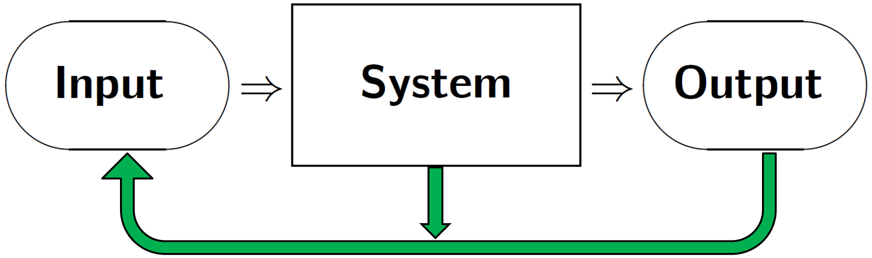


So What is an Inverse Problem?

In a forward problem, we use a mathematical model to compute the output from a “system” given the input.



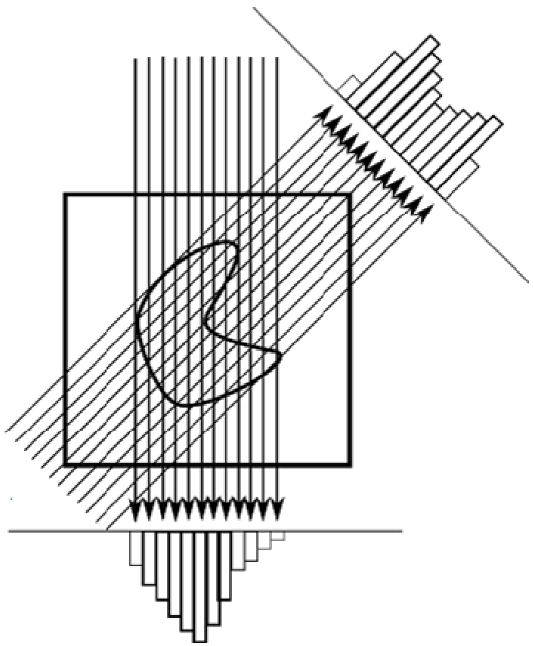
In an inverse problem we compute/estimate a quantity that is not directly observable, using indirect measurements and the forward model.



Solving CT Problems, the Algebraic Way

The Principle

Send X-rays through the object at different angles, and measure the attenuation.



Lambert-Beer law → attenuation of an X-ray through the object f is a line integral:

$$b_i = \int_{\text{ray}_i} f(\xi_1, \xi_2)\, d\ell \ ,$$

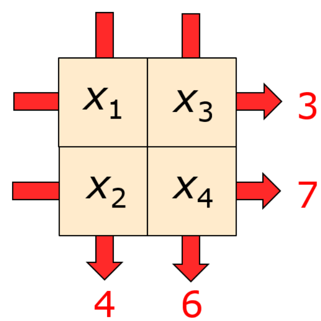
$$f = \text{attenuation coef.}$$

A discrete version:

$$\boxed{A\,x = b}$$

$A \sim$ measurement geometry,

$x \sim$ reconstruction, $\quad b \sim$ data.



$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 4 \\ 6 \end{pmatrix}$$

Just solve for x; how hard is that?

Large-Scale Problems

How to solve **large-scale problems** $A\,x = b$ efficiently?

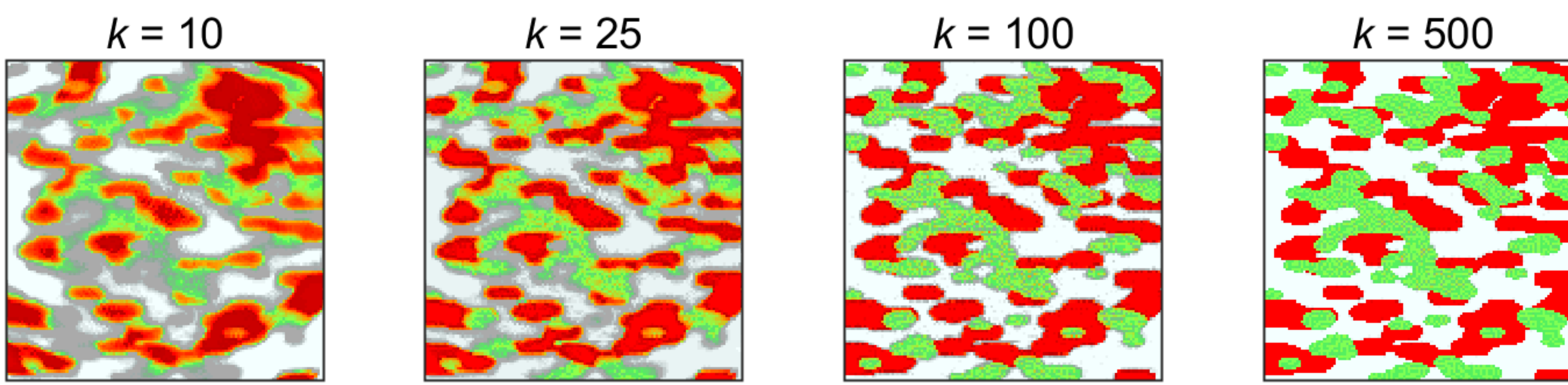
$\rightarrow$ Use iterative methods that produce increasing better reconstructions.

<u>Computer simulation</u>

Image: 128×128.

Data: 360 projection angles in 0°–360°, 181 detector pixels.

k = iteration number



Algebraic Iterative Reconstruction Methods

How to formalize an iterative method for solving $A\,x = b$,
where A = Radon transform = model of the CT scanner.

Landweber iteration with initial $x^0 = 0$:

$$x^k \leftarrow x^{k-1} + \omega\, A^T(b - A\,x^{k-1}) \ .$$

Lots of software is available …

AIR Tools II

Core Imaging Library

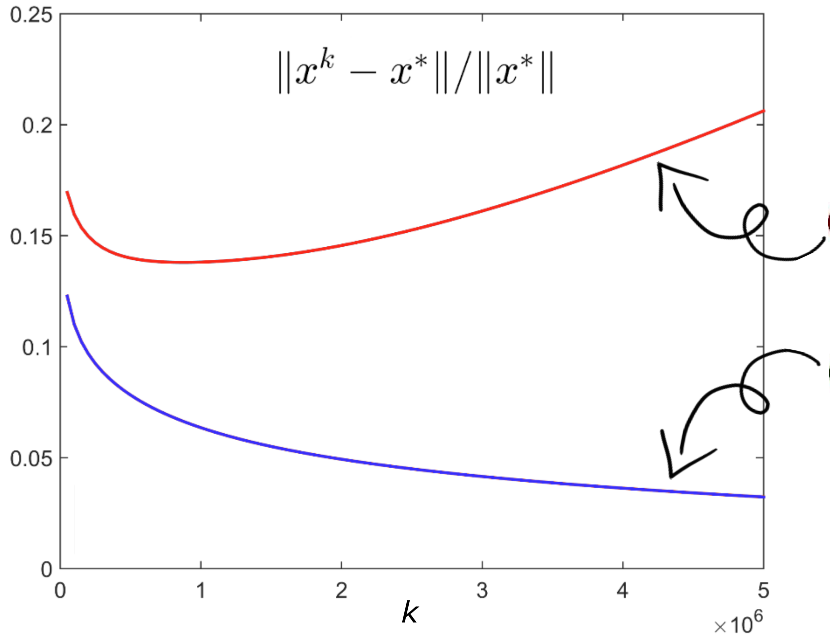
The ASTRA Toolbox

Dealing with an Unmatching Transpose

When the matrix A is too large to store, we perform operations with the Radon transform and its adjoint (the back projection) on a GPU.

The adjoint of A is A^T, but for optimal use of the GPU it is implemented such that it corresponds to a matrix $B \neq A^T$ leading to the iteration:

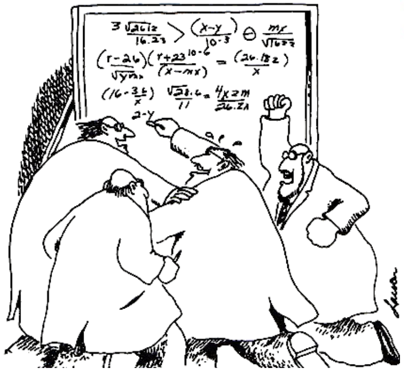
$$x^k \leftarrow x^{k-1} + \omega \, B \, (b - A \, x^{k-1}) \ .$$



Reconstruction errors

- No convergence with unmatched transpose $B \neq A^\mathsf{T}$.
- Convergence with matched transpose A^T.

Fixing the Convergence



1. Ask the software developers to change their implementation of B and/or A?
   $\rightarrow$ Significant loss of comput. efficiency.

2. Use mathematics to *fix* the nonconvergence.

We define the **shifted** version of the iterative algorithm:

$$x^{k+1} = (1 - \alpha\,\omega)\, x^k + \omega\, B\, (b - A\, x^k) \ , \qquad \alpha > 0$$

with just one extra factor $(1 - \alpha\,\omega)$; simple to implement.

Conditions for convergence, with λ_j = eigenvalues of $B\,A$:

$$0 < \omega < 2\, \frac{\mathrm{Re}\,\lambda_j + \alpha}{|\lambda_j|^2 + \alpha\,(\alpha + 2\,\mathrm{Re}\,\lambda_j)} \qquad \text{and}$$

$$\mathrm{Re}\,\lambda_j + \alpha > 0 \ .$$

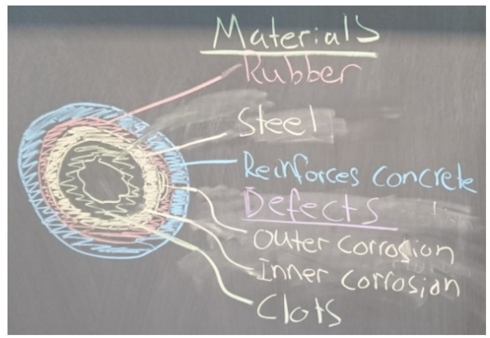
Choose the shift α *just large enough!*

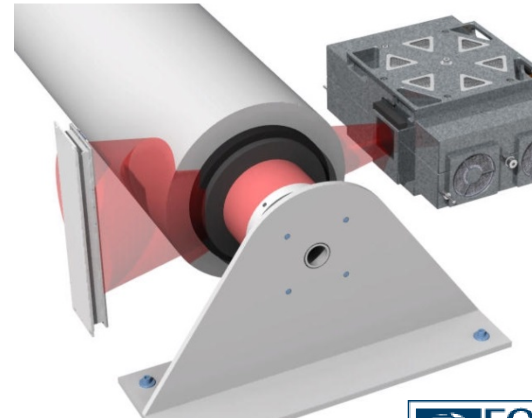
Dong, H, Hochstenbach, Riis; SISC, 2019.

Towards the Villum Project

Use X-ray scanning to compute cross-sectional images of oil pipes on the seabed.

Detect *defects*, *cracks*, etc. in the pipe.

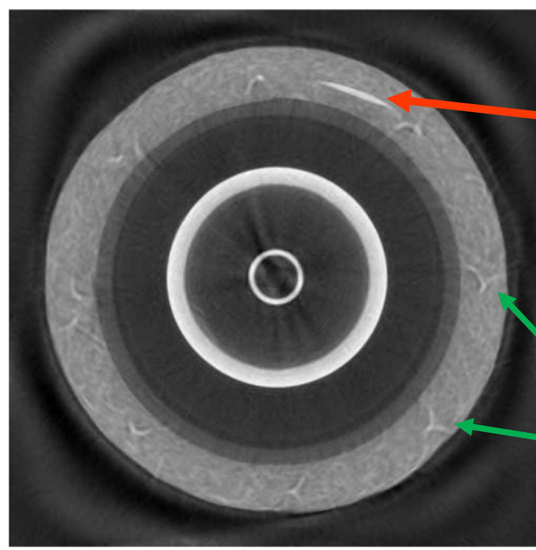




FORCE TECHNOLOGY

<u>Required in the math. model</u>

- Strength of the X-ray source.
- Specification of the geometry.
- Structure of the oil pipe.



Defect!
How much can we trust the size and the location?

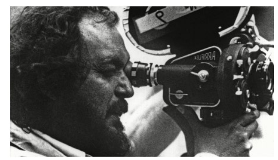
Reinforcing bars

<u>C</u>omputational <u>U</u>ncertainty <u>Q</u>uanti-fication for <u>I</u>nverse Problems

Inverse problem: compute hidden features from external data.

Data:
blurred image

Model of
blurring



Reconstruction
w/ edge prior

The problems are hampered by:

- measurement errors in the data,
- errors/uncertainties in the mathematical model,
- uncertainties in our prior knowledge about the solution.

Uncertainty Quantification (UQ) is the study of the impact of all forms of error and uncertainty in the data and models, through the posterior obtained via Bayes' rule.

CUQI Sampling the posterior is computationally challenging and calls for hierarchical prior modeling, model reduction, and many other "tools."

Example: Archeology as an Inverse Problem

What did buildings of former times look like?

Data: whatever ruins are left.

Prior: everything we know about the culture, building styles, aesthetics, etc.

Model: a temple that is worn down by the elements over 2000 years.

Example: Reconstruction of Viking Halls

Very limited data: traces of the sturdy timbers that the hall was built from show as dark patches in the light natural subsoil.

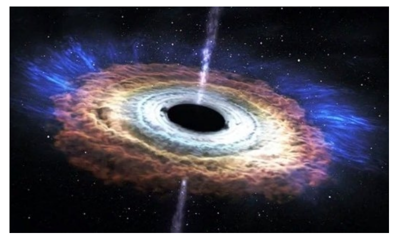
There might be many possible solutions!

The UQ Approach to Inverse Problems

Uncertainty Quantification (UQ) is based on Bayesian statistics.

Instead of producing a single solution (i.e., $x = A^{-1} b$) we obtain the *distribution* (the posterior) of all possible solutions.

Black hole?



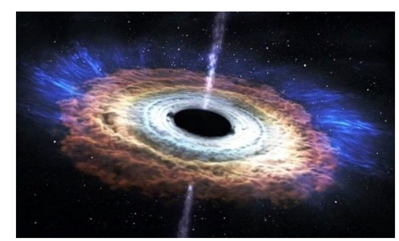
Classical computation methods produce a single image; but can we trust it?

Uncertainty Quantification (UQ) reveals the reliability of all possible reconstructions.

12%

17%

63%



8%

The Lowdown

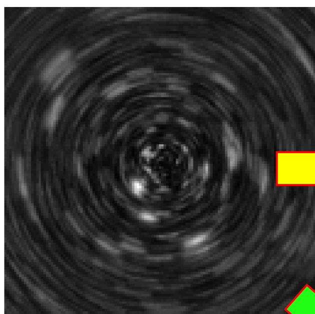
Simple case: $A\,x = b$ where the unknown x is a random vector.

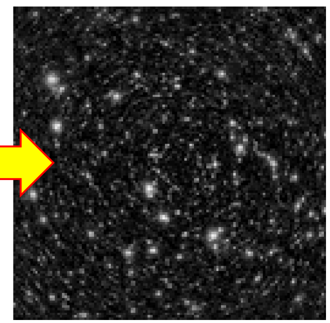
Bayes rule/law/theorem defines the *posterior* for x:

$$p(x|b) \propto p(b|x)\, p(x)\ .$$

Here, $p(b|x)$ is the data's likelihood and $p(x)$ is the prior for the solution.

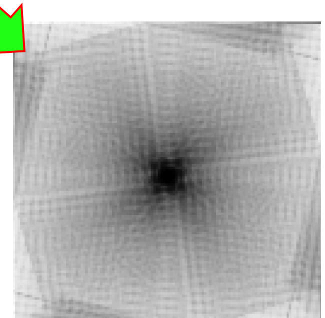
Measured blurred image *b*





Classical solution *x*

MAP estimate



UQ shows the uncertainty (variance) in each pixel; white = high uncertainty.

CUQI

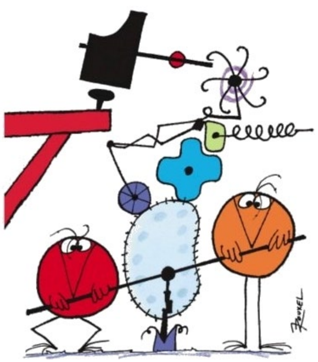
Vision: Computational UQ becomes an essential part of solving inverse problems in science and engineering.

Ingredients

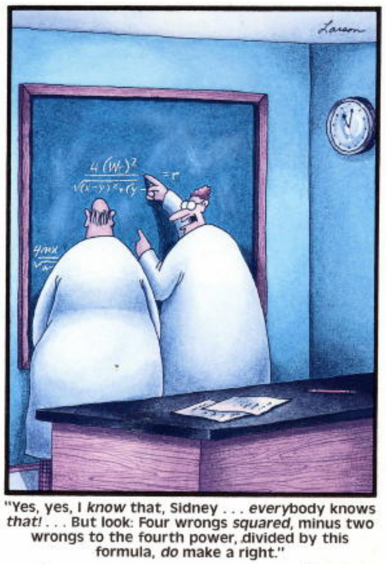
- Develop formulations of inverse problems that incorporate all uncertainties in the data, the models, the assumptions, the computations, etc.
- Develop mathematical & statistical methods and algorithms suited for practical applications.
- Create a modeling framework and a computational platform for non-experts.

VILLUM FONDEN

Comput. ”engine”



Mathematics



The Computational Aspect

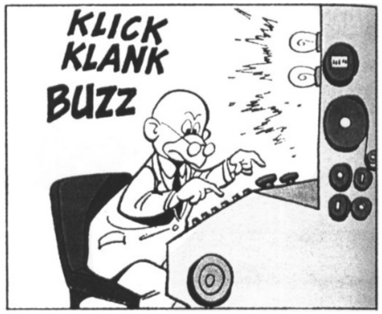
Philosophy

- Hide mathematics, statistics and scientific computing from non-expert users.
- Give expert users full control of the UQ methods and computations.
- All users can focusing on their modeling of the inverse problem.

Case: UQ for edge-preserving reconstruction

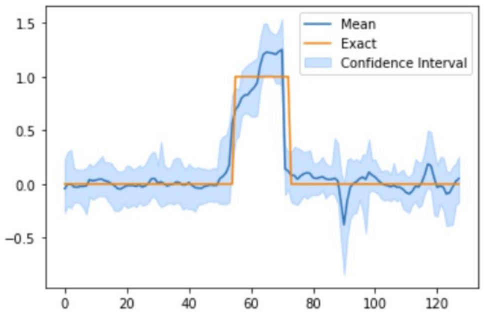
$$\min_x \|A\,x - b\|_2 \quad \text{s.t.} \quad \|\nabla x\|_1 \leq \delta$$

➢ Today: 500 lines of code

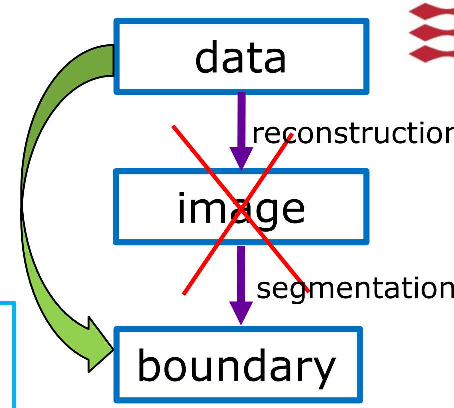


Actual CUQYpy code

```
model = LinearModel(A)
noise = Gaussian(np.zeros(m),0.05)
prior = Cauchy_diff(np.zeros(n),0.05,'neumann')
IP    = Type1(data,model,noise,prior) #data=model(prior)+noise
IP.UQ()
```



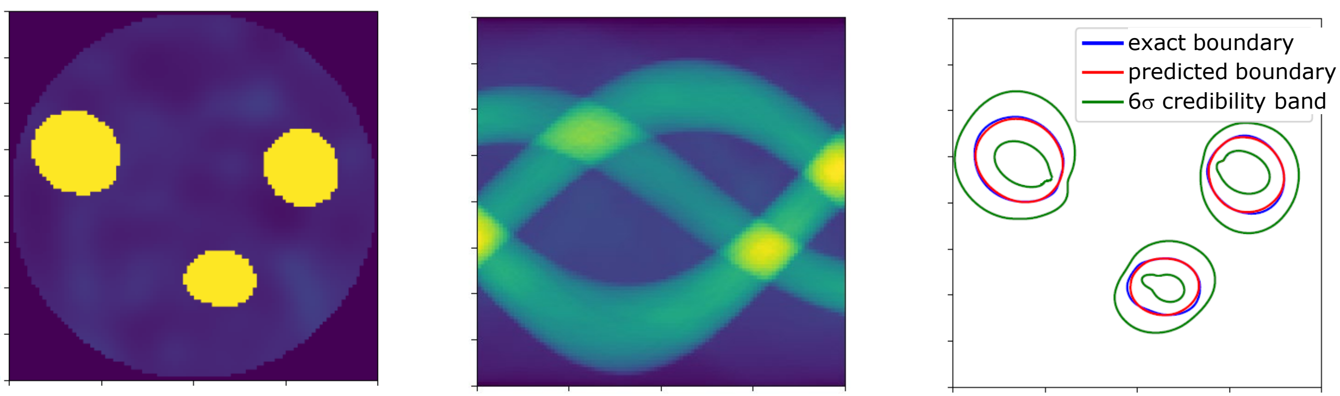
Case: Goal-Oriented CUQI



Reconstruct the desired quantity directly from data, and perform UQ on this quantity.

Example in X-ray imaging: Find inclusion boundaries without a classical two-stage process & perform UQ on the boundaries.

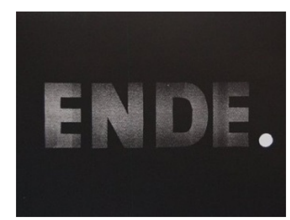
ground truth | data | boundaries w/ UQ



- *2D → 1D* computational problem, no pixels, no error accumulation.
- Represent the inclusion boundaries as *random-field functions*.
- Assign a hyper-parameter that controls the boundary's *regularity*.
- Perform UQ by assigning *probabilities* to the functions and their regularity.

Thanks for your attention

Any questions or uncertainties?



Appendix: Fixing the Convergence

We introduce a scaling factor slightly smaller than one:

$$x^k \leftarrow (1 - \alpha\,\omega)\, x^{k-1} + \omega\, B\, (b - A\, x^{k-1}) \ .$$

Dong, Hansen, Hochstenbach, Riis (2019)

Let λ_j denote those eigenvalues of BA that are different from $-\alpha$.

Then the Shifted BA Iteration converges to a fixed point if and only if α and ω satisfy

$$0 < \omega < 2\, \frac{\mathrm{Re}\,\lambda_j + \alpha}{|\lambda_j|^2 + \alpha\,(\alpha + 2\,\mathrm{Re}\,\lambda_j)} \qquad \text{and} \qquad \mathrm{Re}\,\lambda_j + \alpha > 0 \ .$$

The fixed point x^*_α satisfies

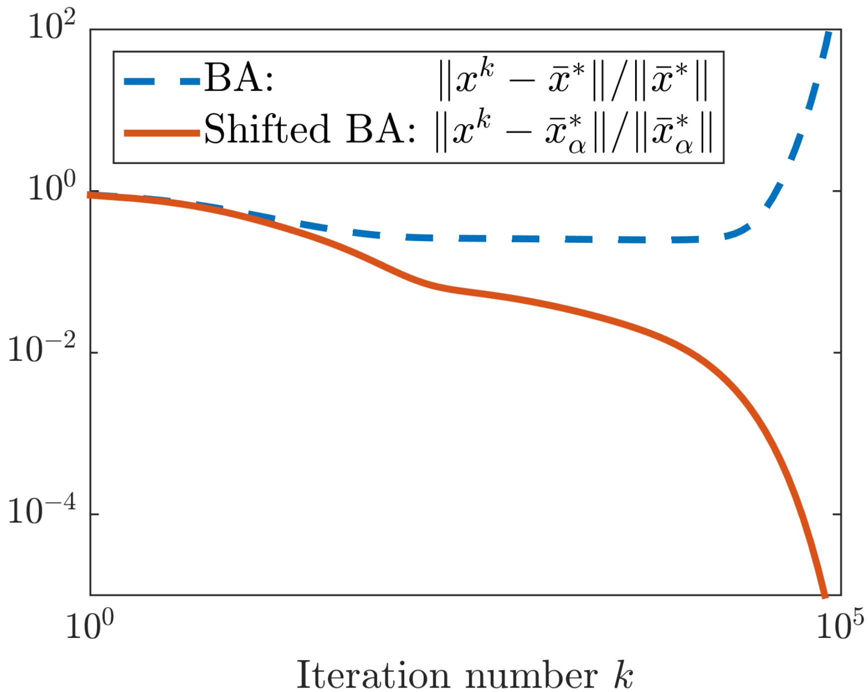
$$(BA + \alpha\, I)\, x^*_\alpha = B\, b \ .$$

$$\bar{x} - \bar{x}^*_\alpha = \alpha\, (BA + \alpha\, I)^{-1} \bar{x} \ .$$

Convergence to a slightly perturbed solution.

Appendix: Nonconvergence → Convergence

Image: 128×128. Data: 90 projection angles in 0°–180°, 80 detector pixels.
Both *A* and *B* are from the GPU-version of the ASTRA toolbox.



The Landweber iteration diverges from $\bar{x}^* = (BA)^{-1} B\, b$.

The shifted iteration converges to fixed point $\bar{x}^*_\alpha = (BA + \alpha I)^{-1} B\, b$.

Appendix: Gaussian Likelihood & Prior

Model: $b = A\,\bar{x} + e$ with $A \in \mathbb{R}^{m \times n}$ fixed and $e = \mathcal{N}(0, \sigma^2 I)$.

The pdf for b, given x and σ (known as the *likelihood*):

$$p(b|x,\sigma) = \left(\frac{1}{2\pi\sigma^2} \right)^{m/2} \exp\left(-\frac{1}{2\sigma^2} \, \|A\,x - b\|_2^2 \right) .$$

The unknown x is a random vector. Assume a Gaussian prior $x \sim \mathcal{N}(0, \delta^{-1} I)$; this yields the prior

$$p(x|\delta) = \left(\frac{\delta}{2\pi} \right)^{n/2} \exp\left(-\frac{\delta}{2} \, \|x\|_2^2 \right) .$$

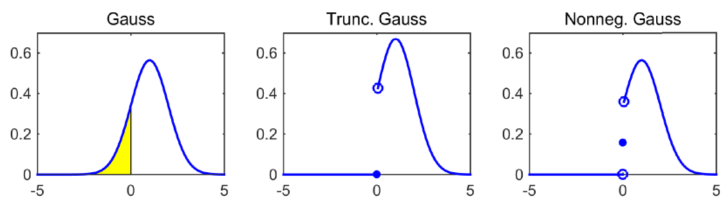
Bayes rule/law/theorem defines the *posterior* for x:

$$\begin{aligned} p(x|b,\sigma,\delta) &= \frac{p(b|x,\sigma)\, p(x|\delta)}{p(b|\sigma,\delta)} \propto p(b|x,\sigma)\, p(x|\delta) \\ &\propto \text{const} \cdot \exp\left(-\frac{1}{2\sigma^2} \, \|A\,x - b\|_2^2 \right) \cdot \exp\left(-\frac{\delta}{2} \, \|x\|_2^2 \right) \\ &\propto \exp\left(-\|A\,x - b\|_2^2 - \alpha \, \|x\|_2^2 \right) , \qquad \alpha = \delta\,\sigma^2 . \end{aligned}$$

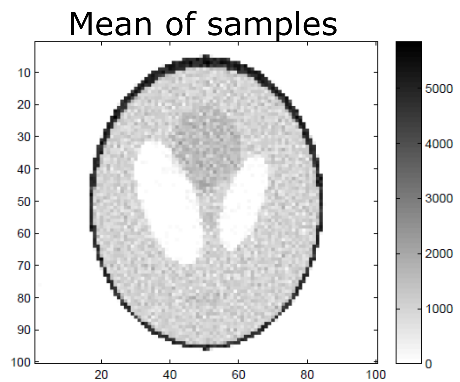
Appendix: UQ with Non-Negative Prior

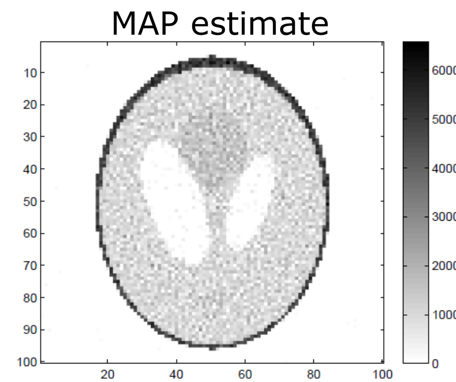
If the prior or likelihood is non-Gaussian, we must **sample** the posterior: we generate <u>many</u> random instances of the regularized solution with the specified likelihood and prior.

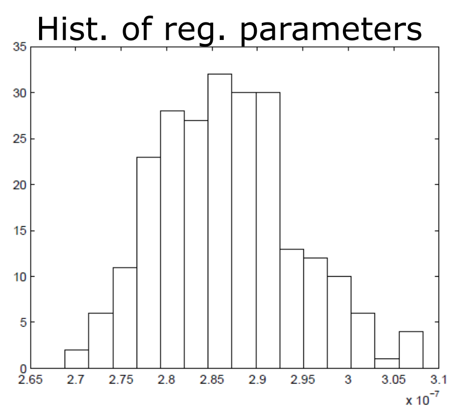
Bardsley, Hansen, *MCMC Algorithms for Non-negativity Constrained Inverse Problems*, 2019.

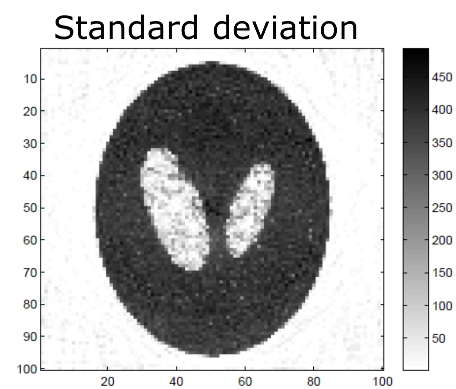


We have an analytical expression for the prior, but no analytical expression for the posterior.







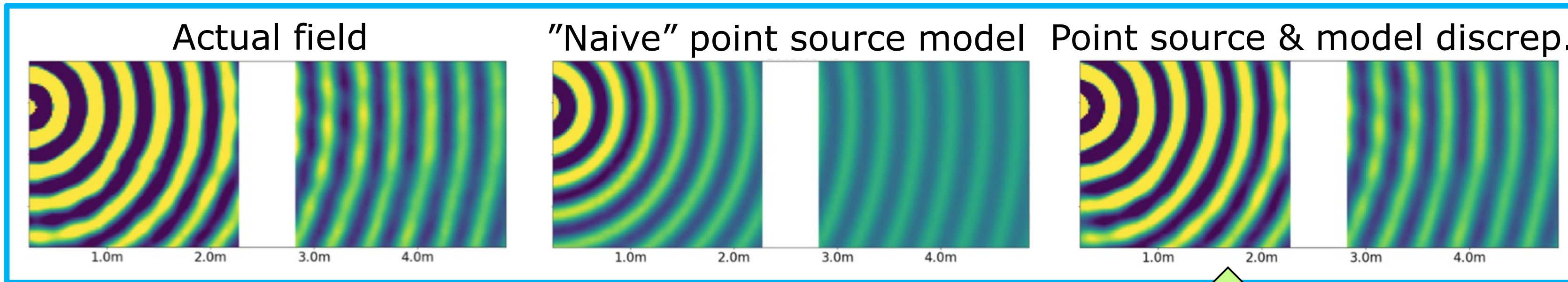


Positron Emission Tomography.

Solutions sampled by a new Poisson Hierarchical Gibbs Sampler.

Appendix: UQ for Model Discrepancies

Dong, Riis, Hansen, *Modeling of sound fields*, joint with DTU Elektro, 2019.



Described by a Gaussian process

Measured data = Physical model + Model discrep-ancy + Data errors

Physical model: Cannot include all possible aspects

Model discrepancy: Accounts for *known unknowns* & *unknown unknowns*

Data errors: Known statistics

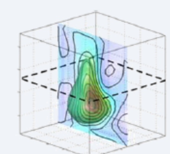
HD-Tomo: High-Definition Tomography

The following examples are from the project **HD-Tomo**, which was funded by an ERC Advanced Research Grant, 2012–17.

Objective: Optimal Use Prior Information

Tomographic imaging allows us to see inside objects. Doctors look for cancer, physicists study microscopic details of materials, security personnel inspect luggage, engineers identify defects in pipes, concrete, etc.

To achieve **high-definition tomography**, sharp images with reliable details, we must use *prior information* = *accumulated knowledge about the object.* **This project: how to do this in an optimal way.**



Outcome: Insight, Framework and Algorithms

We developed *new theory* that provides insight and understanding of the challenges and possibilities of using advanced priors.

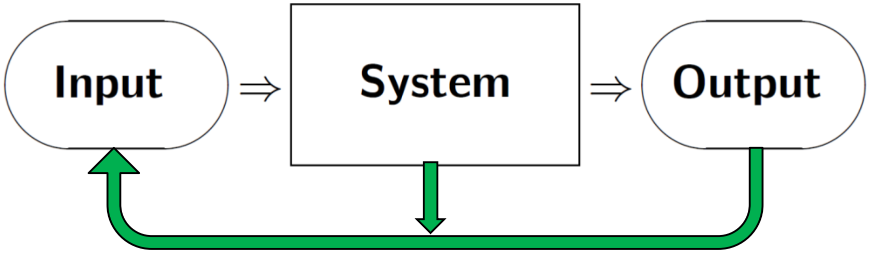
This insight allowed us to develop a *framework for precisely formulated tomographic algorithms* that produce *well-defined results.*

We laid the groundwork for the next generation of algorithms that will further optimize the use of prior information.

The project produced **47** journal papers, **6** proceeding papers, **7** software packages, **25** bachelor/master projects and **3** workshops.

What is an Inverse Problem?

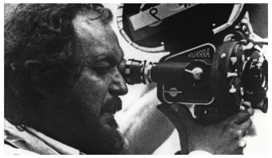
In a **forward problem**, we use a mathematical model to compute the output from a "system" given the input.



In an **inverse problem** we compute/estimate a quantity that is not directly observable, using indirect measurements and the forward model.

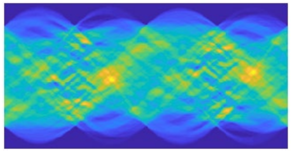
Data:
blurred image

Model of
blurring



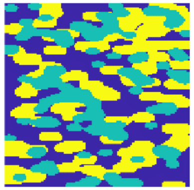
Reconstruction
w/ sharp edges

Data:
sinogram

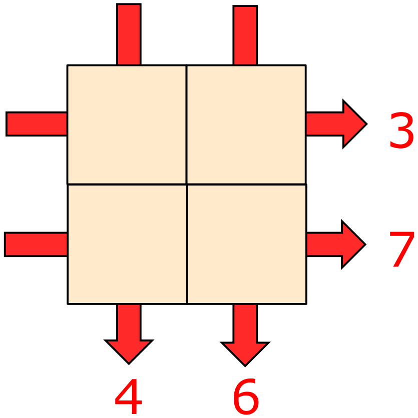


Model of
CT scan

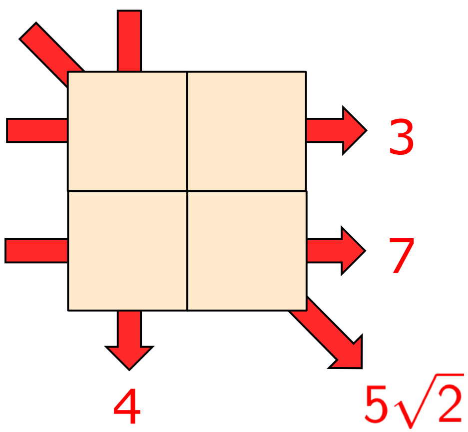
Reconstruction
w/ domains



Analogy: the "Sudoku" Problem – 数独



$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 4 \\ 6 \end{pmatrix}$$



$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ \sqrt{2} & 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 4 \\ 6 \\ 5\sqrt{2} \end{pmatrix}$$

Appendix: Project Overview