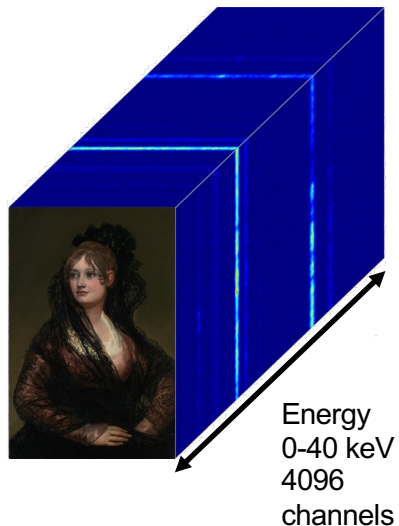


Model-Based Deep Learning for Inverse Problems in Imaging

Pier Luigi Dragotti, Imperial College London

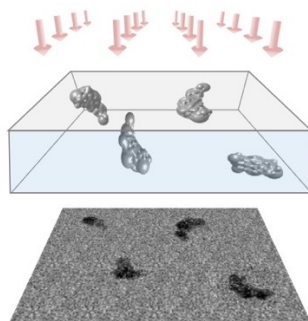
3 February 2025

What do I do for a living: Inverse Imaging Problems

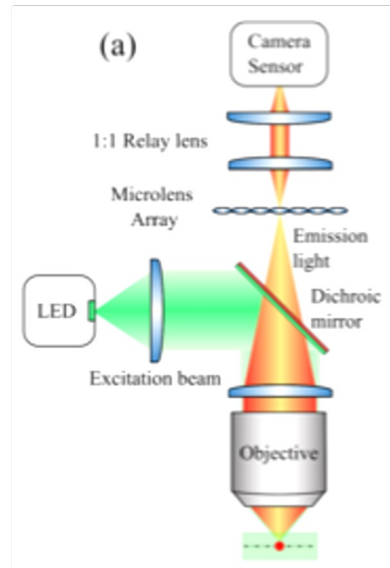


Technical study of
Old Masters
paintings

Image restoration problems



Cryo-EM for structural biology



Light field microscopy for
neuroscience

Plato: models, priors

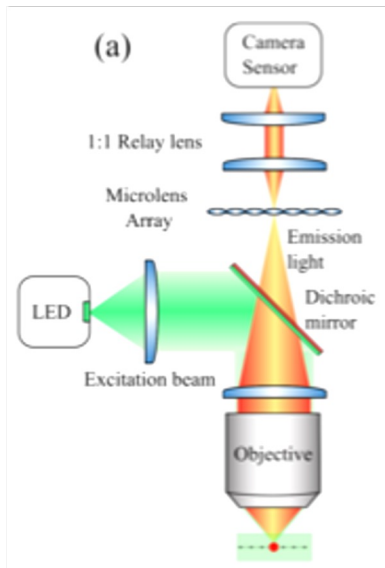
- Inverse problems involve reconstructing unknown physical quantities from indirect measurements.
- The growing complexity of modern imaging workflows calls for a more holistic approach to inverse problems where sensing, physics and computation are analyzed jointly
- **Key in inverse problem** is to find the right balance between **data** and **priors**



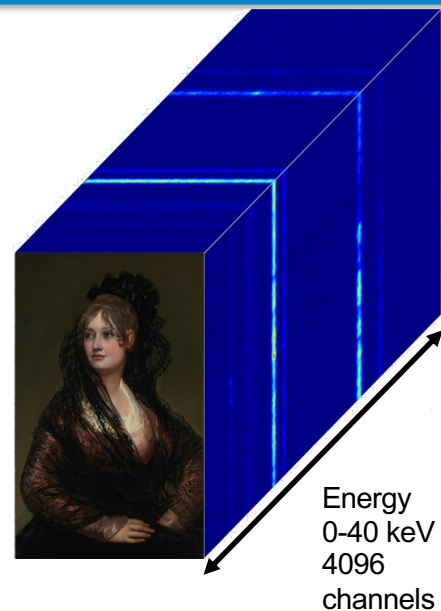
Aristotle: data



Image restoration problems:
Invertible neural networks and
diffusion models



Light field microscopy for
neuroscience



Technical study of Old Masters
paintings



Image restoration problems:
Invertible neural networks and
diffusion models

- $\hat{x} = \min_x \|H(x) - y\|^2 + \lambda \rho(x)$
consistency term prior
- $\hat{x} = \min_{x,v} \|H(x) - y\|^2 + \lambda \rho(v) \quad \text{s.t.} \quad x = v$
- Turn the constraint into a penalty: $\hat{x} = \min_{x,v} \|H(x) - y\|^2 + \lambda \rho(v) + \beta \|x - v\|^2$
- Solve by alternating between x and v
 - Consistency step: $\hat{x} = \min_x \|H(x) - y\|^2 + \beta \|x - v\|^2$
 - A denoiser: $\hat{v} = \min_v \rho(v) + \beta \|x - v\|^2$

Use Deep Learning for
denoising

- $\hat{x} = \min_x \|H(x) - y\|^2 + \lambda \rho(x)$

consistency term

prior

- $\hat{x} = \min_{x,v} \|H(x) - y\|^2 + \lambda \rho(v) \quad \text{s.t.} \quad x = v$

- Turn the constraint into a penalty: $\hat{x} = \min_{x,v} \|H(x) - y\|^2 + \lambda \rho(v) + \beta \|x - v\|^2$

- Solve by alternating between x and v

- Consistency step: $\hat{x} = \min_x \|H(x) - y\|^2 + \beta \|x - v\|^2$

Use INN to impose
consistency

- A denoiser: $\hat{v} = \min_v \rho(v) + \beta \|x - v\|^2$

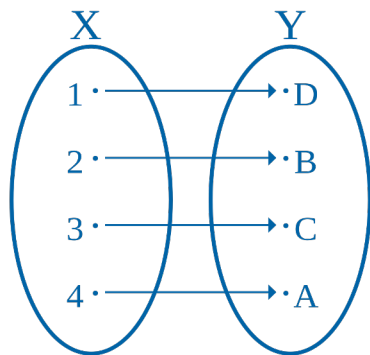
Use Diffusion Models
to impose the prior

Invertible Neural Networks are bijective function approximators with a forward mapping

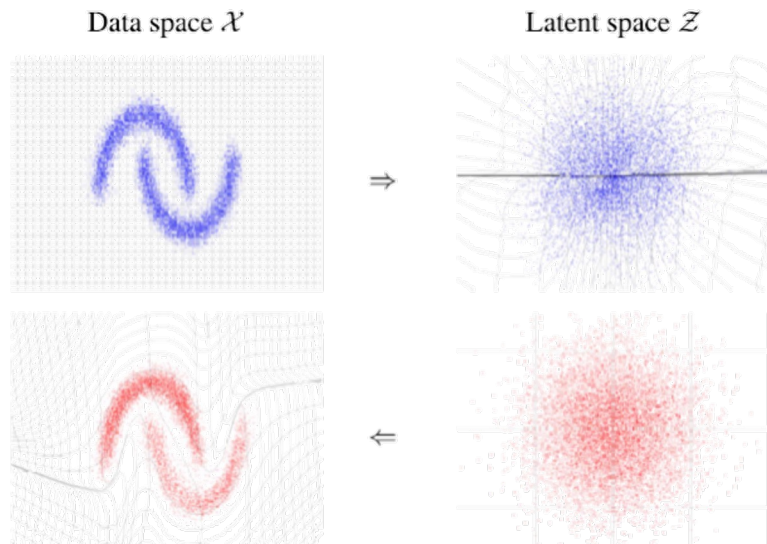
$$F_{\theta}: \mathbb{R}^d \rightarrow \mathbb{R}^l$$
$$x \mapsto z$$

and inverse mapping

$$F_{\theta}^{-1}: \mathbb{R}^l \rightarrow \mathbb{R}^d$$
$$z \mapsto x$$

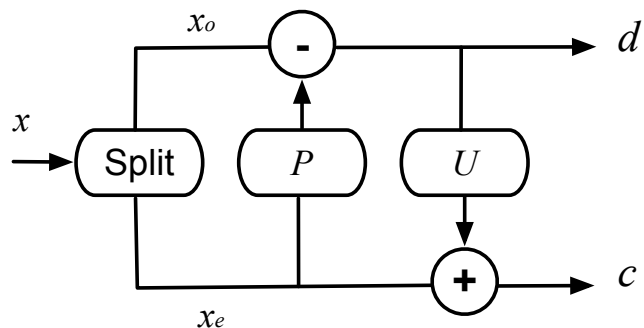


A bijective function (or
invertible function)



How to Achieve Invertibility?

Invertible via lifting scheme

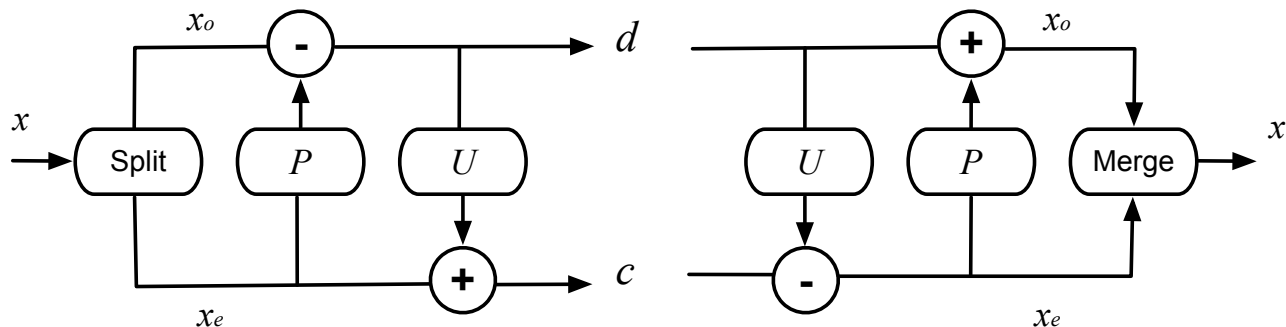


$$\text{Split} \rightarrow \begin{cases} d = x_o - P(x_e) \\ c = x_e + U(d) \end{cases}$$

Forward pass

How to Achieve Invertibility?

Invertible via lifting scheme

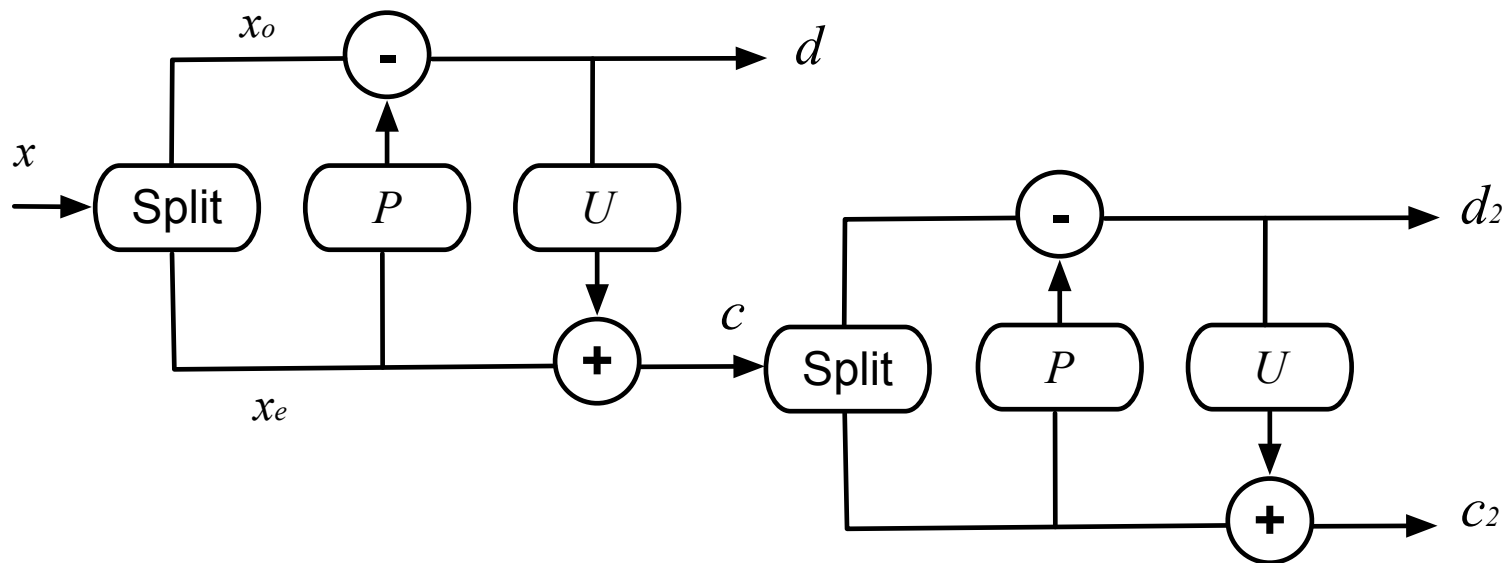


$$\text{Split} \rightarrow \begin{cases} d = x_o - P(x_e) \\ c = x_e + U(d) \end{cases}$$

Forward pass

$$\begin{cases} x_o = d + P(x_e) \\ x_e = c - U(d) \end{cases} \rightarrow \text{Merge}$$

Backward pass



Invertible Neural Networks are ideal architectures to address inverse problems

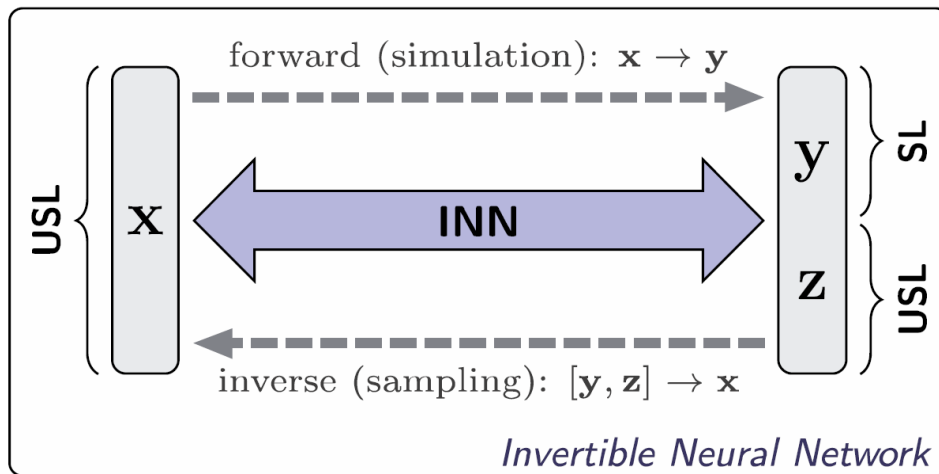
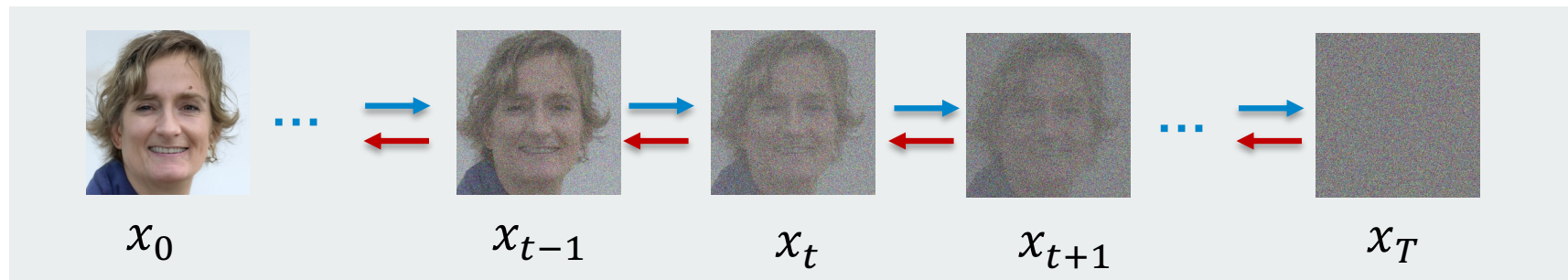


Figure from: Ardizzone, Lynton, Jakob Kruse, Sebastian Wirkert, Daniel Rahner, Eric W. Pellegrini, Ralf S. Klessen, Lena Maier-Hein, Carsten Rother, and Ullrich Köthe. "Analyzing inverse problems with invertible neural networks." in Proc. of *ICLR*, 2019.

Diffusion Models are good for “unconditional” generation of new samples (e.g., Denoising Probabilistic Diffusion Models)



Motivation: Can we use a pretrained “unconditional” diffusion model for inverse problems?

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_{\theta}(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}. \quad \mathbf{x}_{0,t} = \frac{1}{\sqrt{\bar{\alpha}_t}} (\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_{\theta}(\mathbf{x}_t, t))$$

- Diffusion Models are good for “unconditional” generation of new samples (e.g., Denoising Probabilistic Diffusion Models)

- From x_T to x_0 :



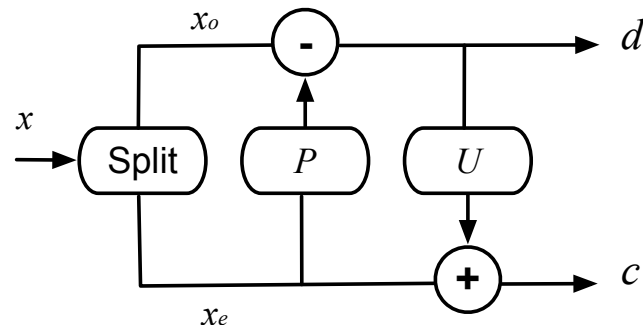
- From $x_{0,T}$ to $x_{0,1}$:

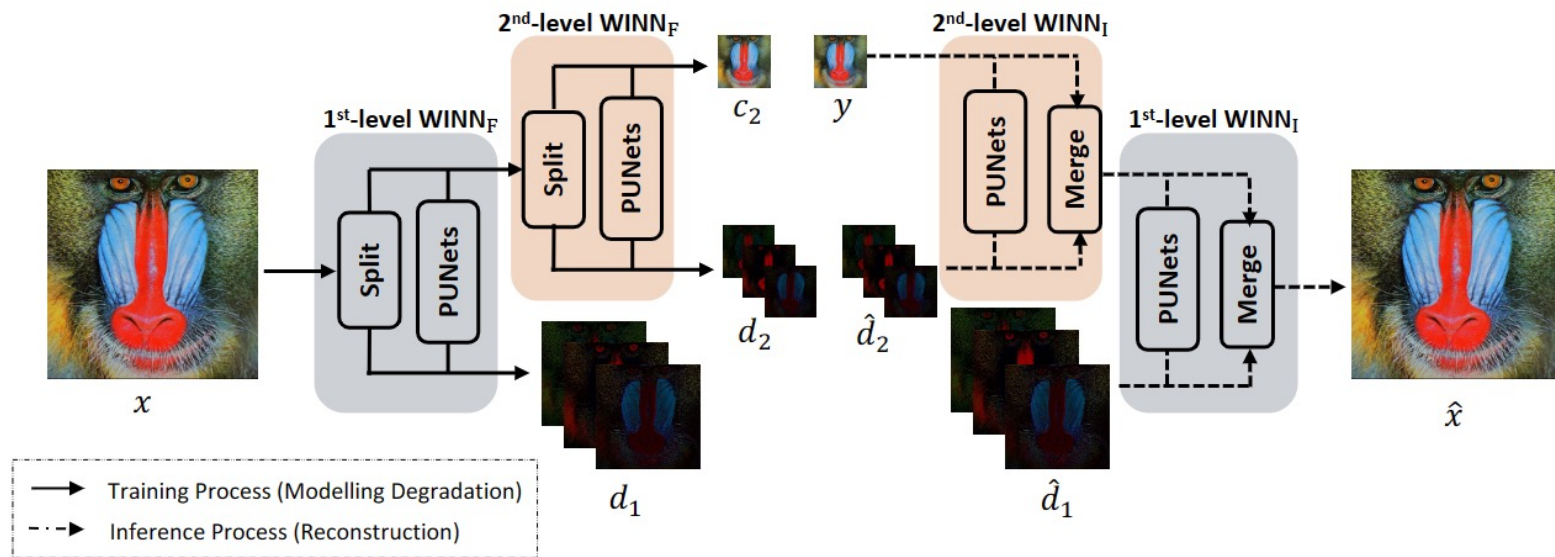


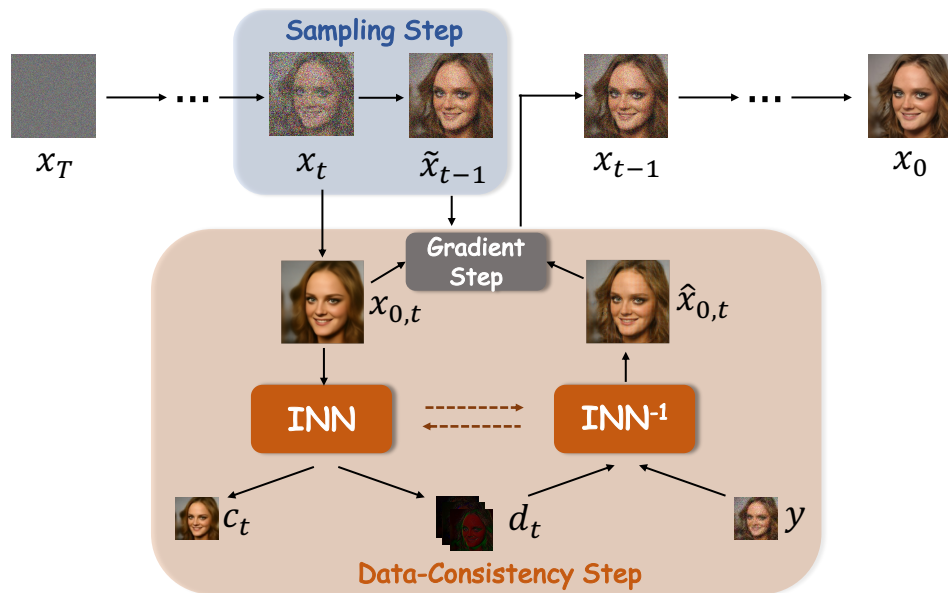
- Given a training set $\{x_i, y_i\}$ which contains N high-quality images and their low-quality counterparts, we learn the forward part of the INN using the following loss:

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{c}^i - \mathbf{y}^i\|_2^2,$$

- Consequently, d models the lost details that need to be recovered with the diffusion model







Algorithm 1 INDigo

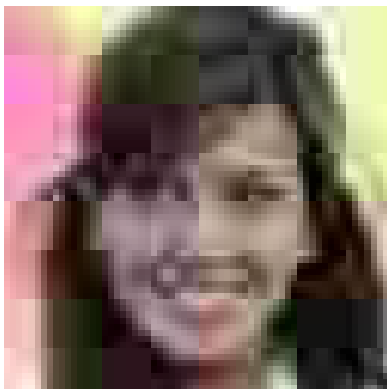
```

1:  $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
2: for  $t = T, \dots, 1$  do
3:    $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  if  $t > 1$ , else  $\mathbf{z} = \mathbf{0}$ 
4:    $\mathbf{x}_{0,t} = \frac{1}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t))$ 
5:    $\tilde{\mathbf{x}}_{t-1} = \frac{\sqrt{\bar{\alpha}_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t}\mathbf{x}_{0,t} + \sigma_t\mathbf{z}$ 
6:    $\mathbf{c}_t, \mathbf{d}_t = f_\phi(\mathbf{x}_{0,t})$ 
7:    $\hat{\mathbf{x}}_{0,t} = f_\phi^{-1}(\mathbf{y}, \mathbf{d}_t)$ 
8:    $\mathbf{x}_{t-1} = \tilde{\mathbf{x}}_{t-1} - \zeta \nabla_{\mathbf{x}_t} \|\hat{\mathbf{x}}_{0,t} - \mathbf{x}_{0,t}\|_2^2$ 
9: end for
10: return  $\mathbf{x}_0$ 

```



Ground Truth



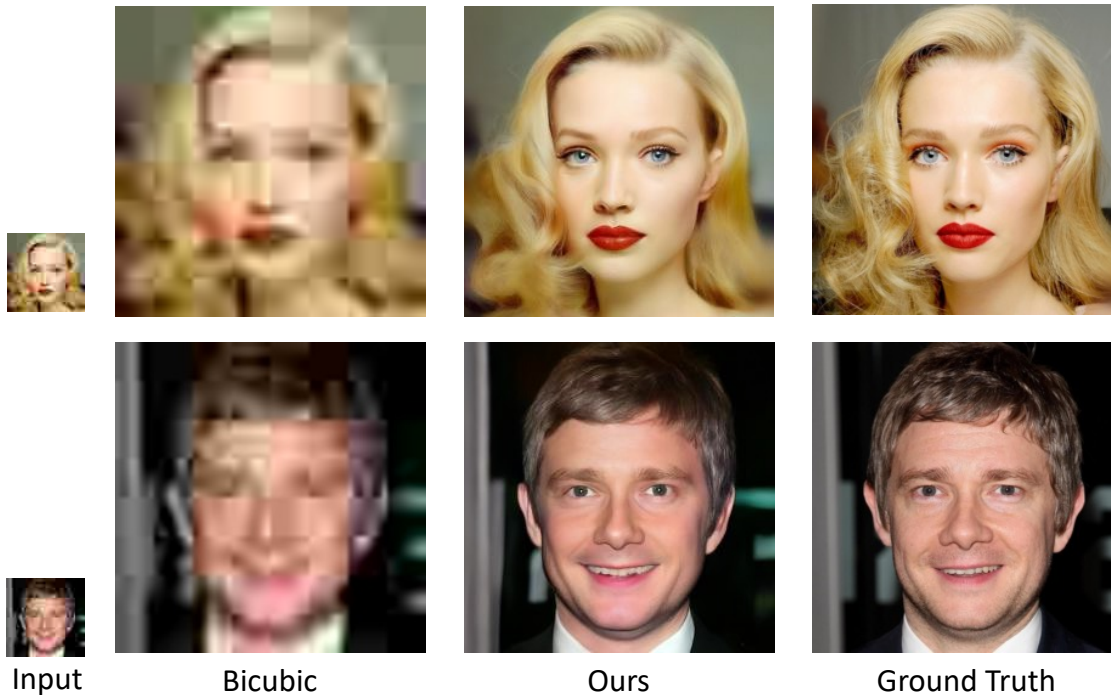
Degraded



Reconstructed

- This approach is simple, flexible and effective
 - No-need to know the degradation process
 - The degradation process can be highly non-linear
 - No need to retrain the diffusion model for every new degradation (just need to train the INN)

Results for non-linear degradation models





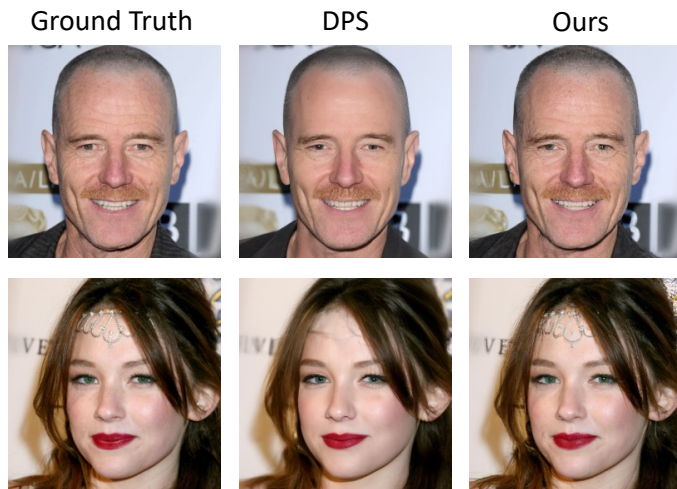
Input

Ours

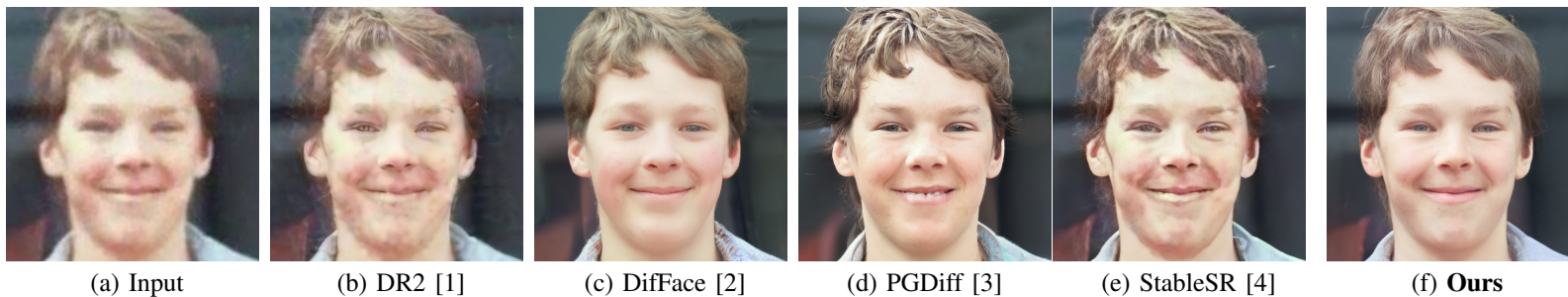
Ground-truth

Results on 4x super-resolution

Method	Noise σ	PSNR $\uparrow$	FID $\downarrow$	LPIPS $\downarrow$	NIQE $\downarrow$
ILVR	0	27.43	44.04	0.2123	5.4689
DDRM	0	28.08	65.80	0.1722	4.4694
DPS	0	26.67	32.44	0.1370	4.4890
Ours	0	28.15	22.33	0.0889	4.1564
ILVR	0.05	26.42	60.27	0.3045	4.6527
DDRM	0.05	27.06	45.90	0.2028	4.8238
DPS	0.05	25.92	31.71	0.1475	4.3743
Ours	0.05	27.16	26.64	0.1215	4.1004
ILVR	0.10	24.60	88.88	0.4833	4.4888
DDRM	0.10	26.16	45.49	0.2273	4.9644
DPS	0.10	24.73	31.66	0.1698	4.2388
Ours	0.10	26.25	28.89	0.1399	3.9659



Results on blind unsupervised deconvolution



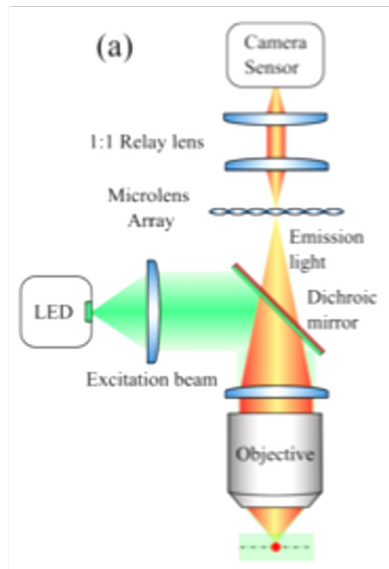
D.You and P.L. Dragotti, "INDIGO+: A Unified INN-Guided Probabilistic Diffusion Algorithm for Blind and Non-Blind Image Restoration", IEEE Journal of Selected Topics in Signal Processing, 2024

- Invertible Neural Networks are an interesting new concept
 - Designing INN and combining them with diffusion models (plug-and-play) leads to more interpretable and simpler architectures
 - Good performance and good **generalization ability**
 - Potential for further developments
-

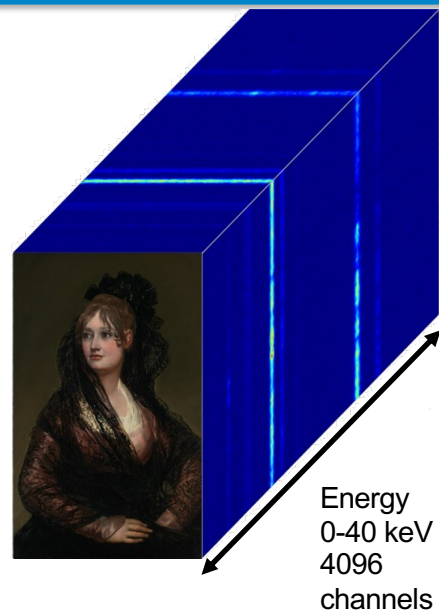
Imaging for scientific
discovery:

- Very complex imaging workflows
- No ground truth data (but each experiment can produce terabytes of data)
- No silver bullet

Part 2: We now get real 😊

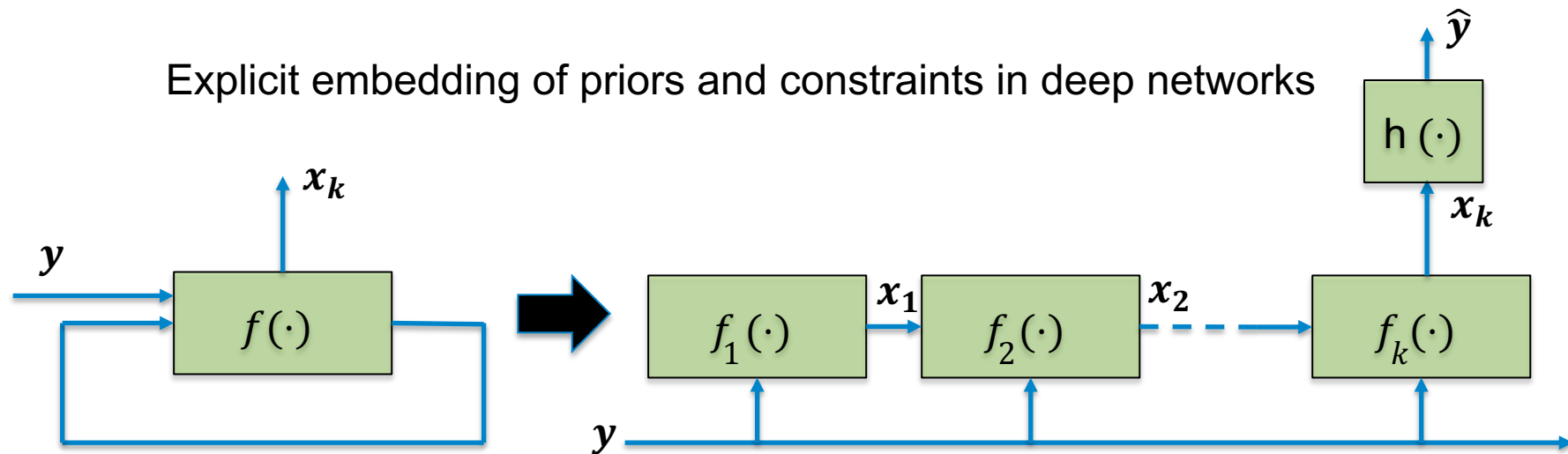


Light field microscopy for
neuroscience



Technical study of Old Masters
paintings

Explicit embedding of priors and constraints in deep networks



Iterative algorithm with y
as input and x as output

Unfolded version of the iterative algorithm with
learnable parameters

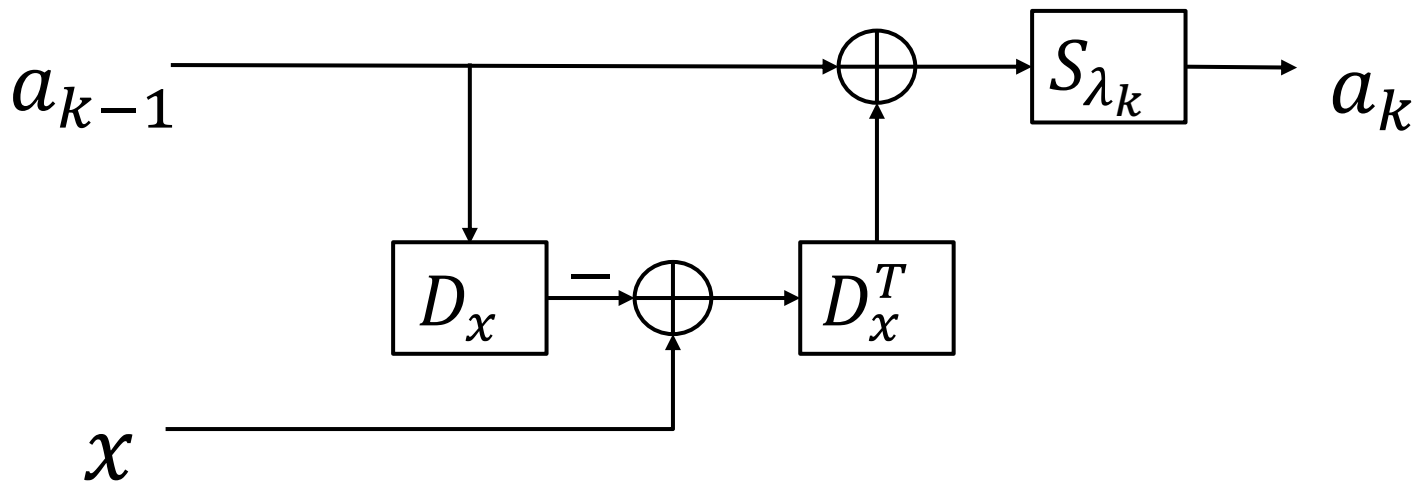
Need to re-synthesize the input, if self-supervised

- The dictionary is usually learned

The diagram illustrates the equation $X = D\alpha$ using matrix visualizations. On the left, matrix X is shown as a vertical column of yellow squares, with a double-headed arrow indicating its height is n . In the center is an equals sign. To the right of the equals sign is matrix D , represented as a grid of blue squares. Three vertical columns in D are highlighted: red, green, and purple. A double-headed arrow below D indicates its width is $m > n$. To the right of D is vector α , shown as a vertical column of gray squares. The red, green, and purple squares in α correspond to the highlighted columns in D , representing the sparse coefficients.

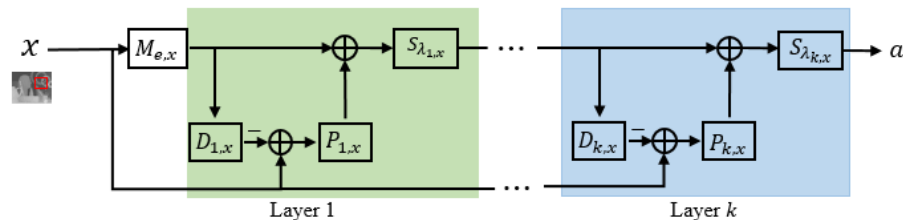
$$\begin{matrix} n \\ \left[\begin{array}{c} \text{yellow squares} \end{array} \right] \\ X \end{matrix} = \begin{matrix} n \\ \left[\begin{array}{c} \text{blue grid with red, green, purple columns} \end{array} \right] \\ m > n \\ D \end{matrix} \begin{matrix} \left[\begin{array}{c} \text{gray column with red, green, purple squares} \end{array} \right] \\ \alpha \end{matrix}$$

- The sparse vector α can be found using ISTA: $\alpha_k = S_{\lambda_k}(\alpha_{k-1} + D_x^T(x - D_x \alpha_{k-1}))$



□ Solving by ISTA algorithm through unfolding:

$$a_k = S_{\lambda_k}(a_{k-1} + D_x^T(x - D_x a_{k-1}))$$

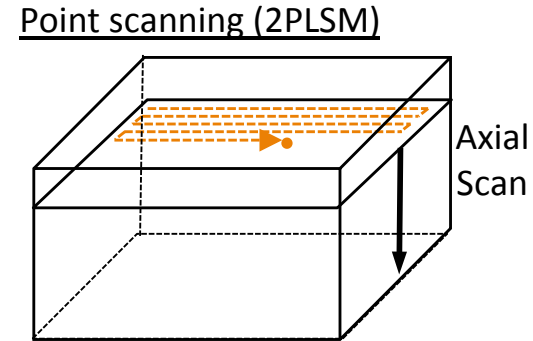


- Gregor Karol and LeCunYann, “Learning fast approximations of sparse coding”, Proceedings of the 27th International Conference on International Conference on Machine Learning, 2010
- Y. Eldar et al, “Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing”, IEEE Signal Processing Magazine, 2021

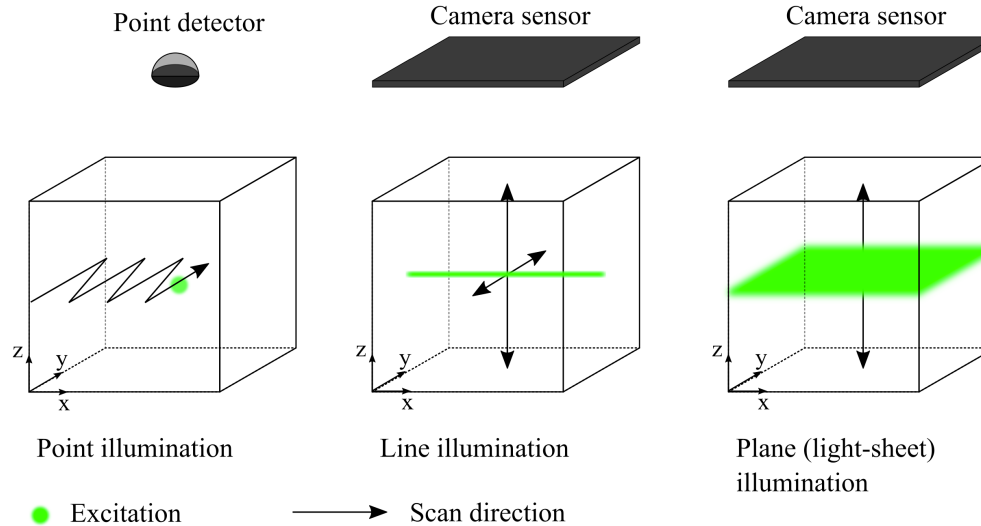
Two-Photon Microscopy for Neuroscience

- Goal of Neuroscience: to study how information is processed in the brain
 - Neurons communicate through pulses called Action Potentials (AP)
 - Need to measure in-vivo the activity of large populations of neurons at cellular level resolution
 - Two-photon microscopy combined with right indicators is the most promising technology to achieve that
-

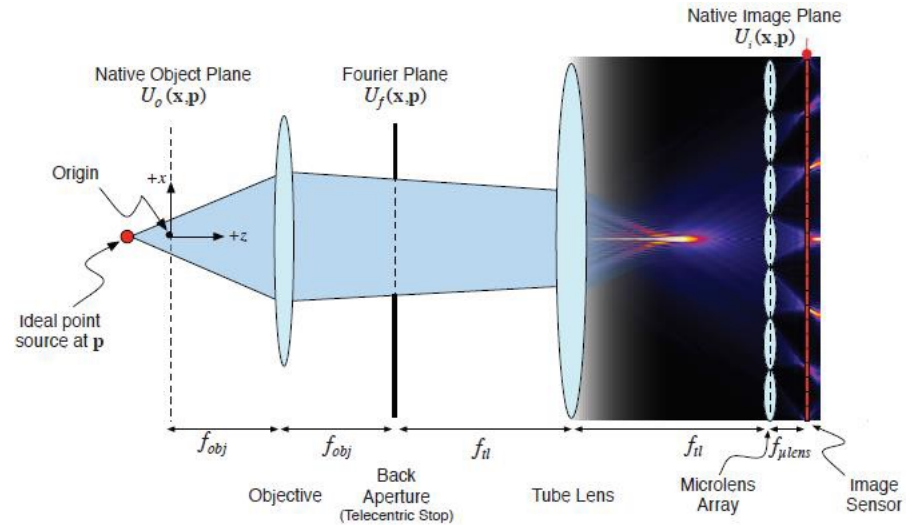
- Fluorescent sensors within tissues
- Highly localized laser excites fluorescence from sensors
- Photons emitted from tissue are collected
- Focal spot sequentially scanned across samples to form image
- Two-photon microscopes in raster scan modality can go deep in the tissue but are **slow**

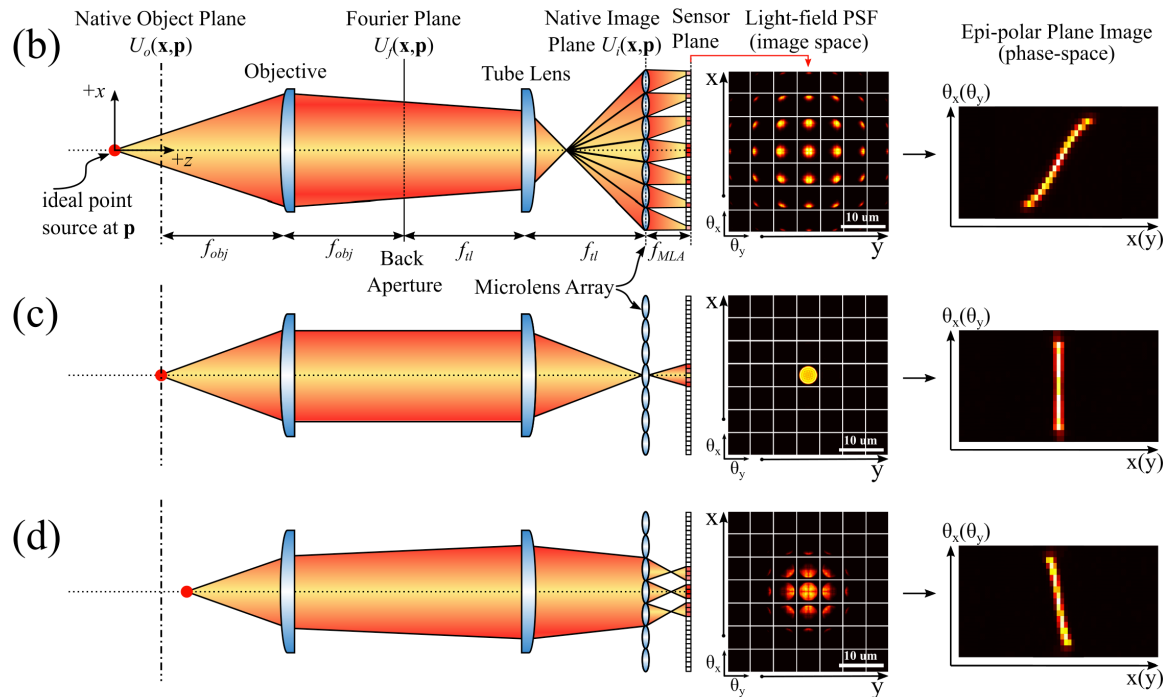
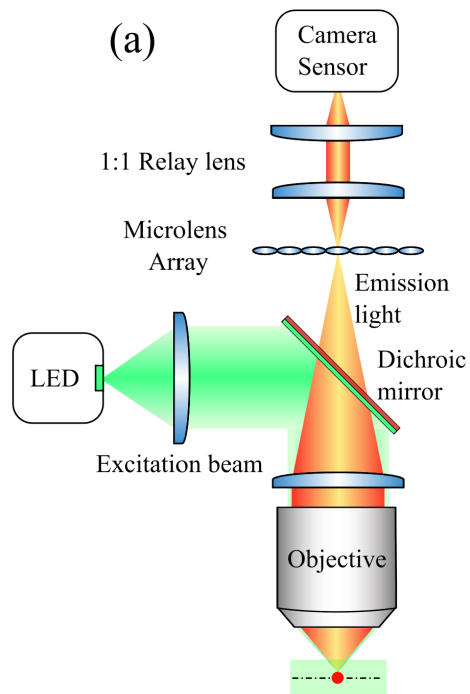


- In order to speed up acquisition one can change the illumination strategy
- This mitigates the issue but does not fix it
- Issue with scattering

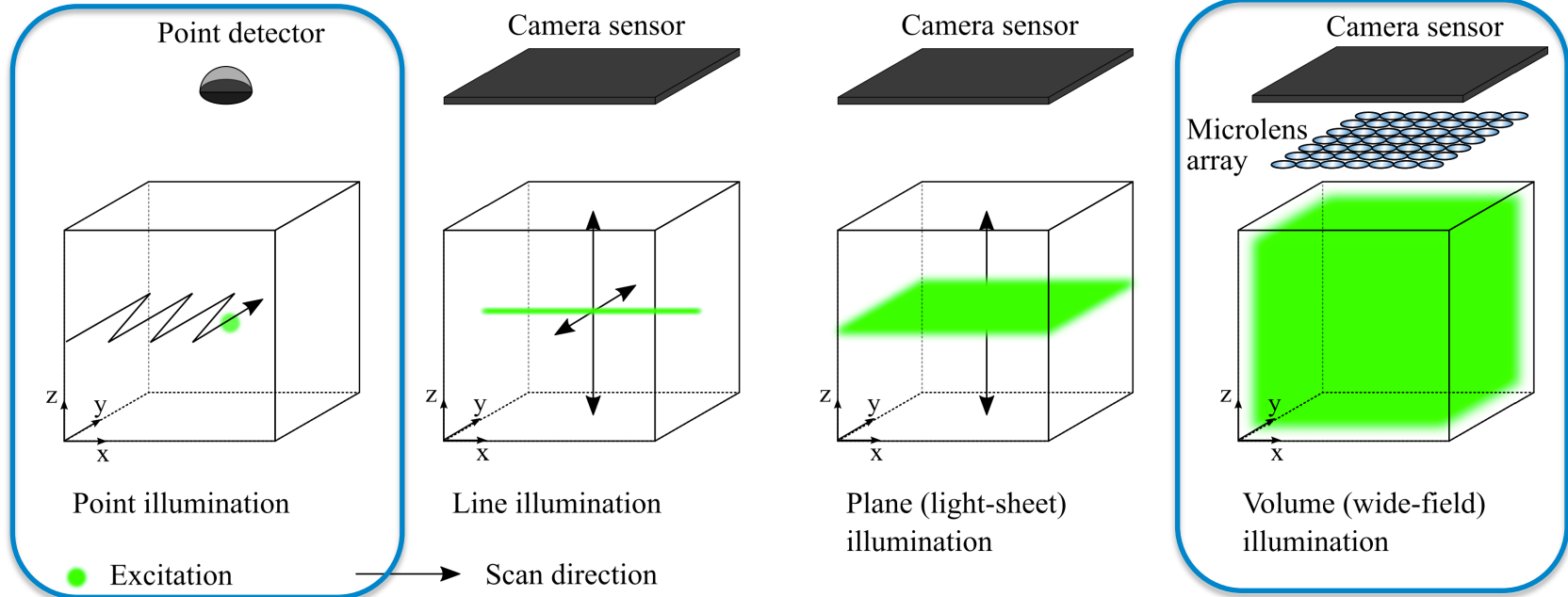


Light-Field Microscopy (LFM) is a high-speed imaging technique that uses a simple modification of a standard microscope to capture a 3D image of an entire volume in a single camera snapshot



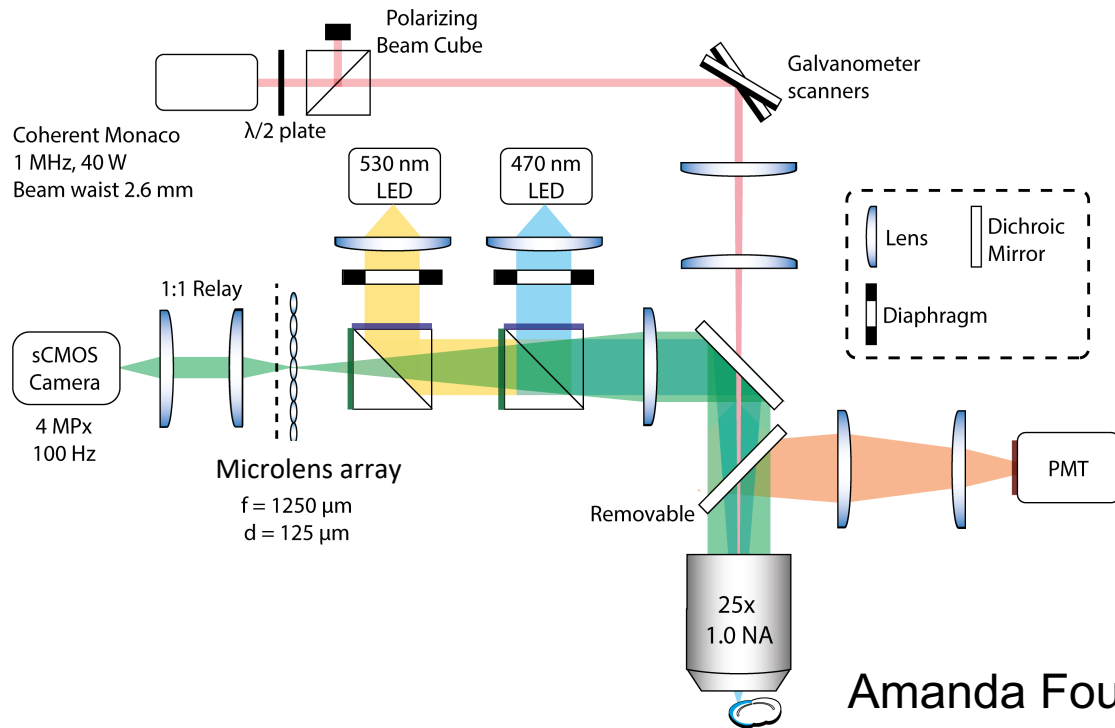


Light-field Microscopy and Illumination Strategies



Key insight: use the 2P microscope for high-resolution structural information and the LFM for monitoring the activity of neurons.

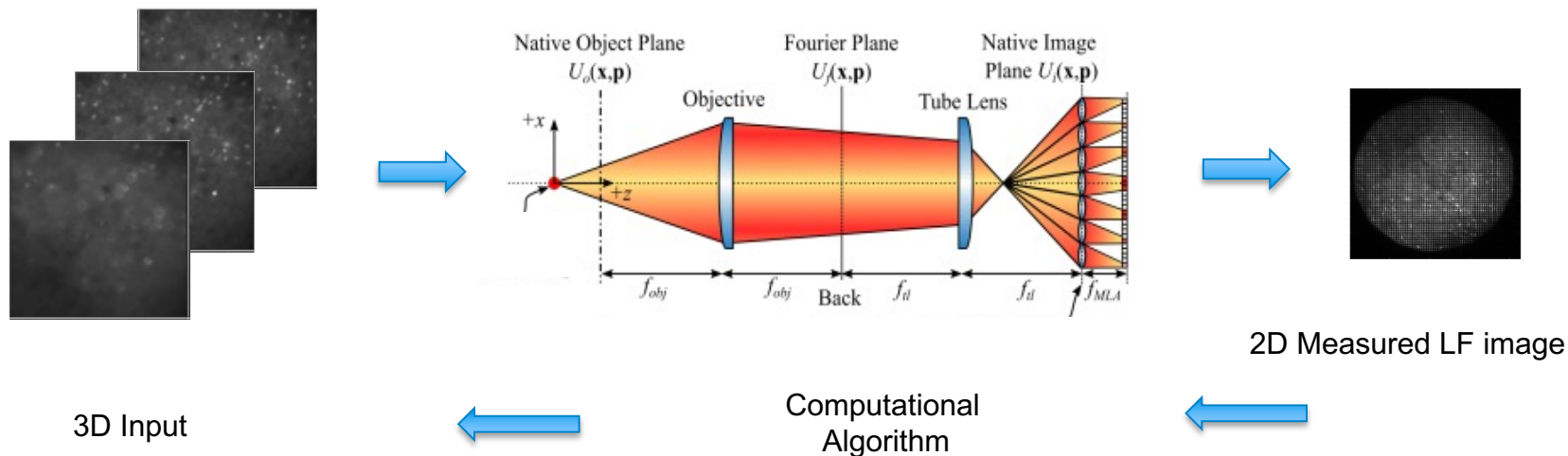
Our Solution: Scattering-robust structural volumes + high-bandwidth, scanless functional volumes



Amanda Foust



Challenge: given a sequence of lightfields (2-D signals), need to reconstruct a sequence of volumes (3-D+t)

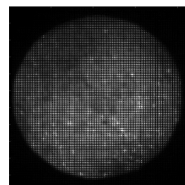


- **Challenges**

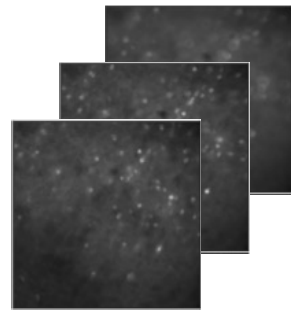
- Scattering induces blur, making inversion more challenging
- Lack of ground-truth data for learning

- **Opportunities**

- Forward model structured and linear
- Data is **sparse** (neurons fire rarely and are localized in space)
- Occlusion can be ignored

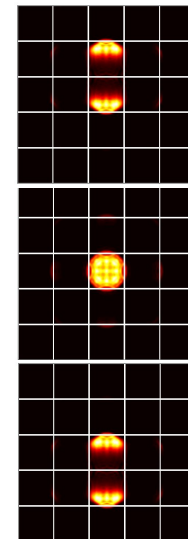
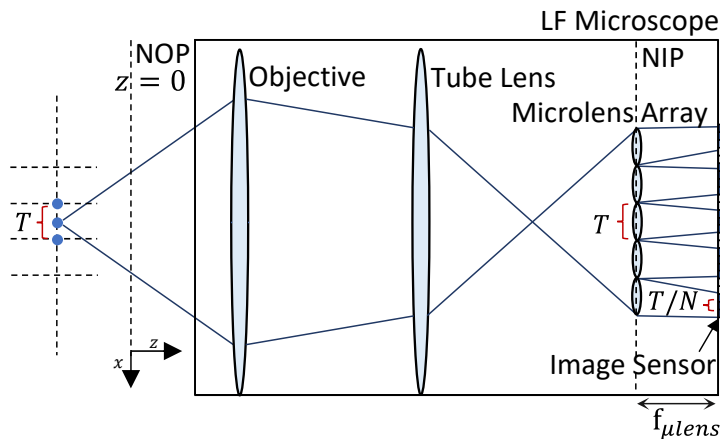


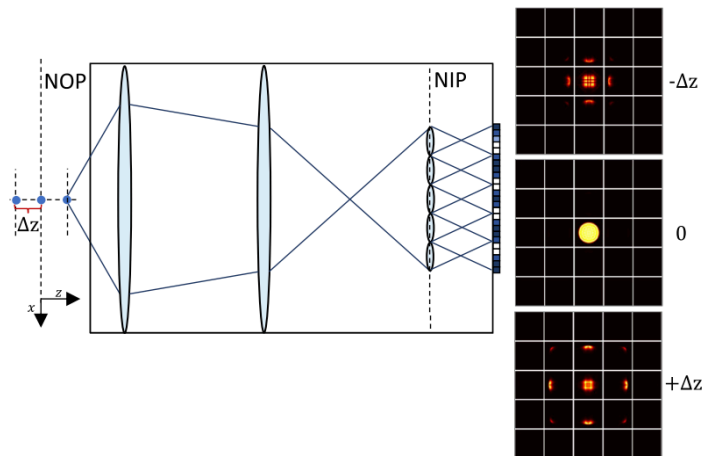
2-D LF



Volume

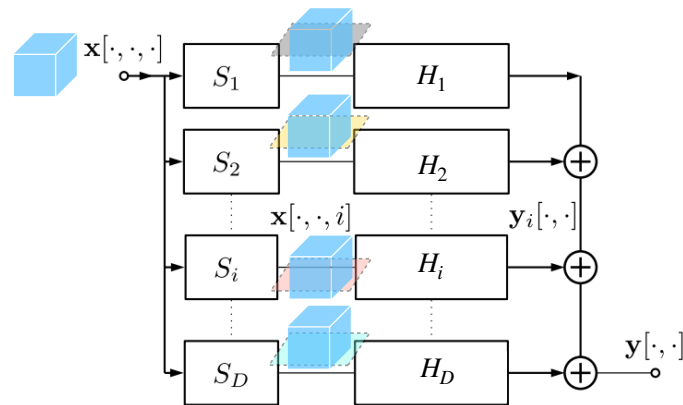
- Forward model is linear which means $\mathbf{y} = \mathbf{H}\mathbf{x}$
 - $\mathbf{H}$ is estimated using wave-optics
 - For each depth, $\mathbf{H}$ is block-circulant (periodically shift invariant) and can be modelled with a filter-bank
 - The entire forward model can be modelled using a linear convolutional network with known parameters (given by the wave-optics model)





- The discrete LFM system: $\mathbf{y} = \mathbf{H}\mathbf{x}$

$$\mathbf{H} = \sum_{i=1}^D \mathbf{H}_i \mathbf{S}_i$$



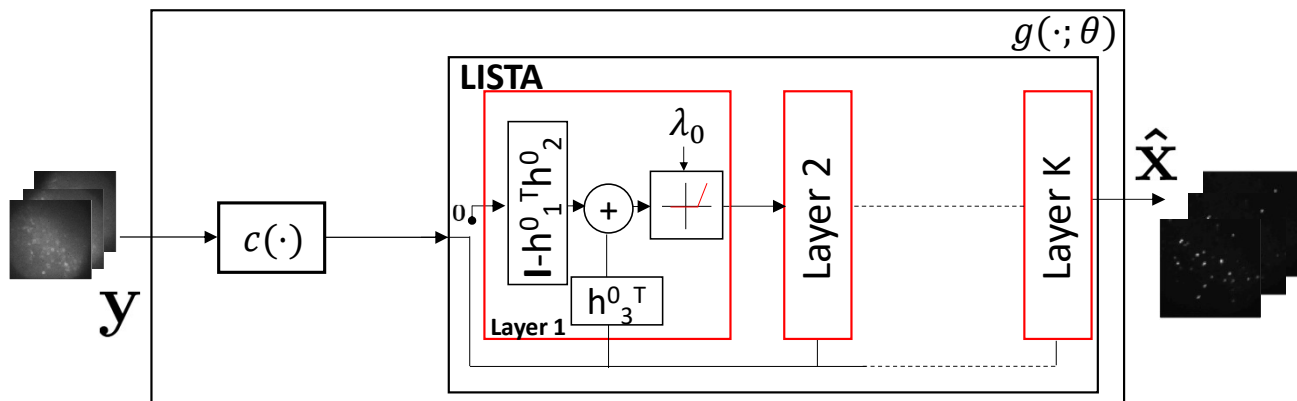
- Data is **sparse** (neurons fire rarely and are localized in space)
- Solve $\min_x (\|y - Hx\|^2 + \|x\|_1)$ s.t $x \geq 0$
- This leads to the following iteration:

$$x_{k+1} = \text{ReLU}(x_k - H^T H x_k + H^T y + \lambda)$$

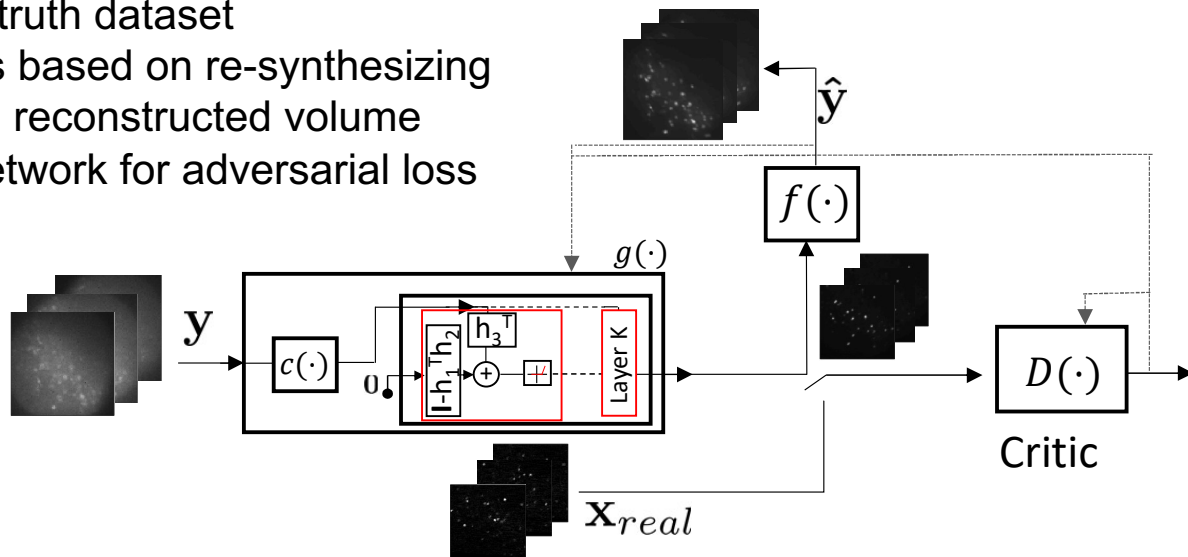
- Approach: Convert the iteration in a deep neural network using the unfolding technique

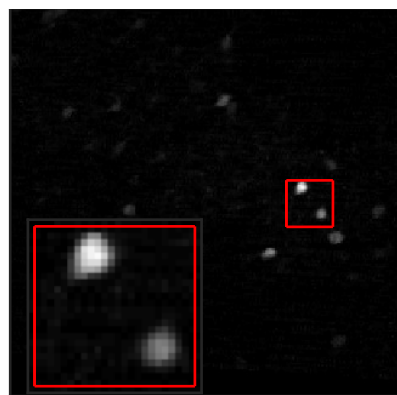
- Convert the iteration in a deep neural network using the unfolding technique

$$x^{k+1} = \text{ReLU}(x^k - H^T H x^k + H^T y + \lambda)$$

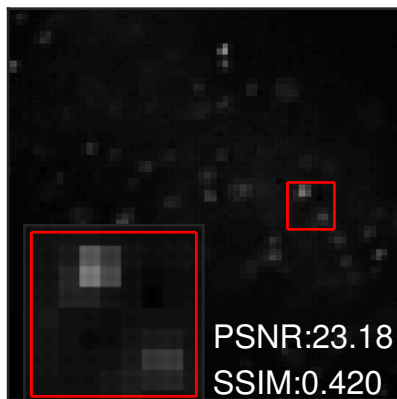


- Training, in this context, is difficult due to lack of ground-truth data
- Our approach: semi supervised learning
 - Small ground truth dataset
 - Light-field loss based on re-synthesizing light-field from reconstructed volume
 - Adversarial network for adversarial loss

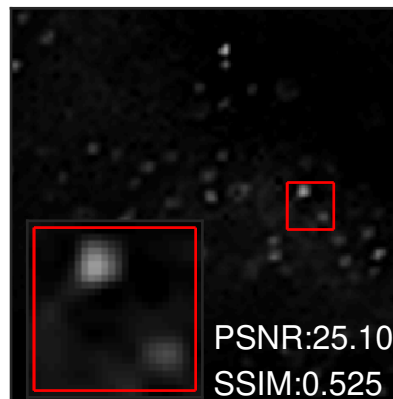




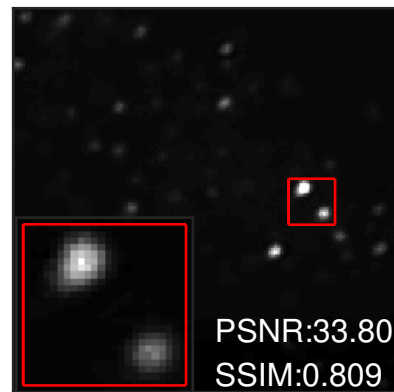
Ground-truth



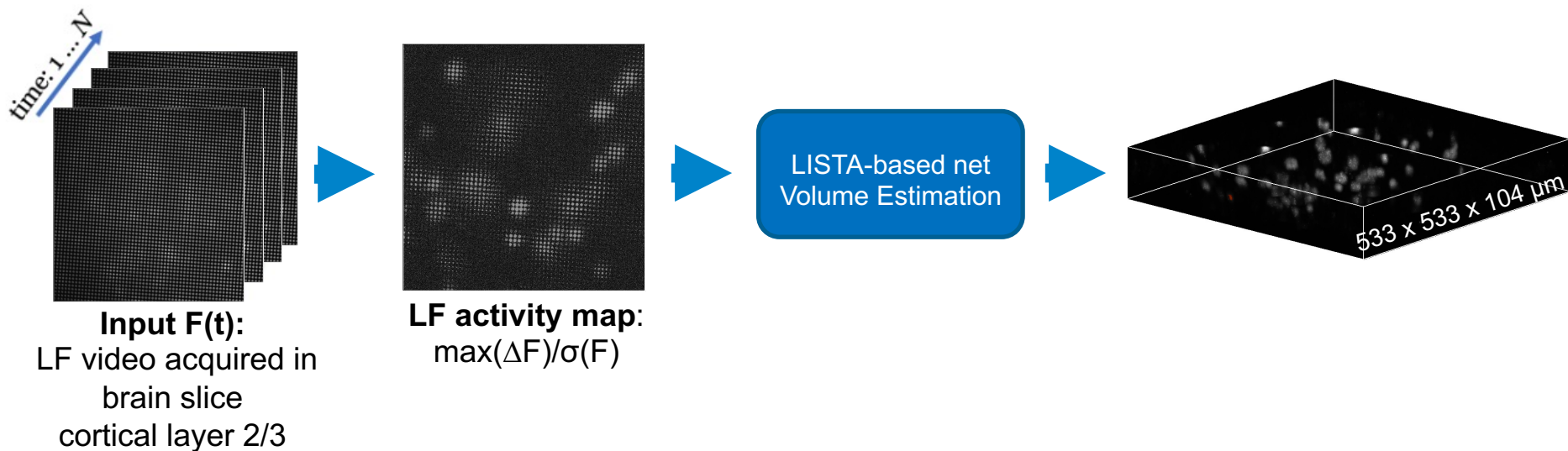
ISRA



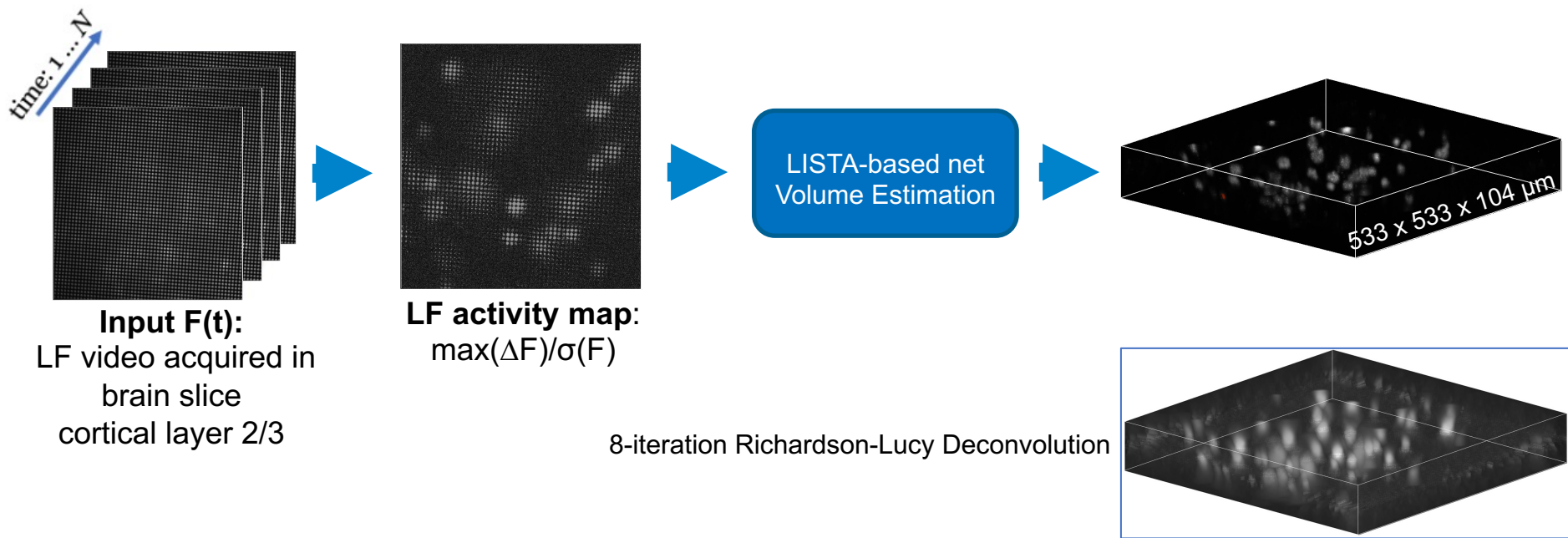
ADMM



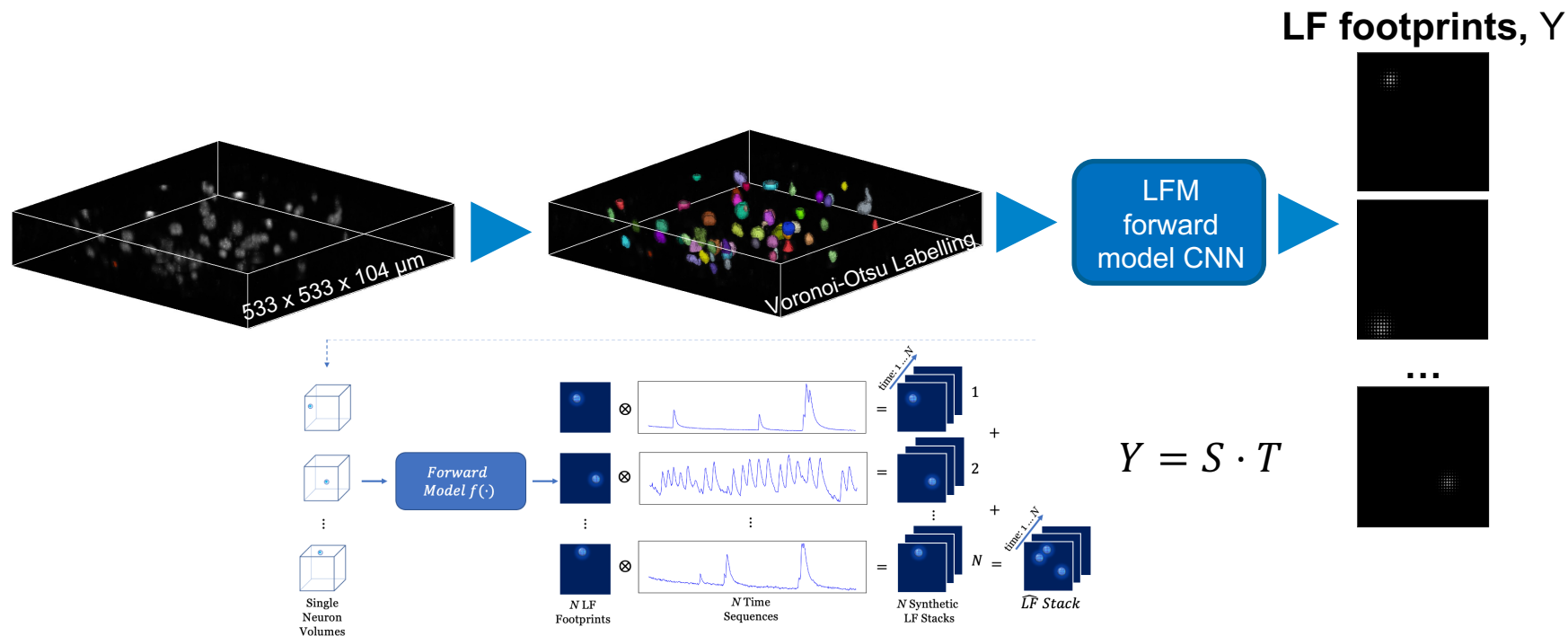
Fast volumetric jGCaMP8f time-series extraction

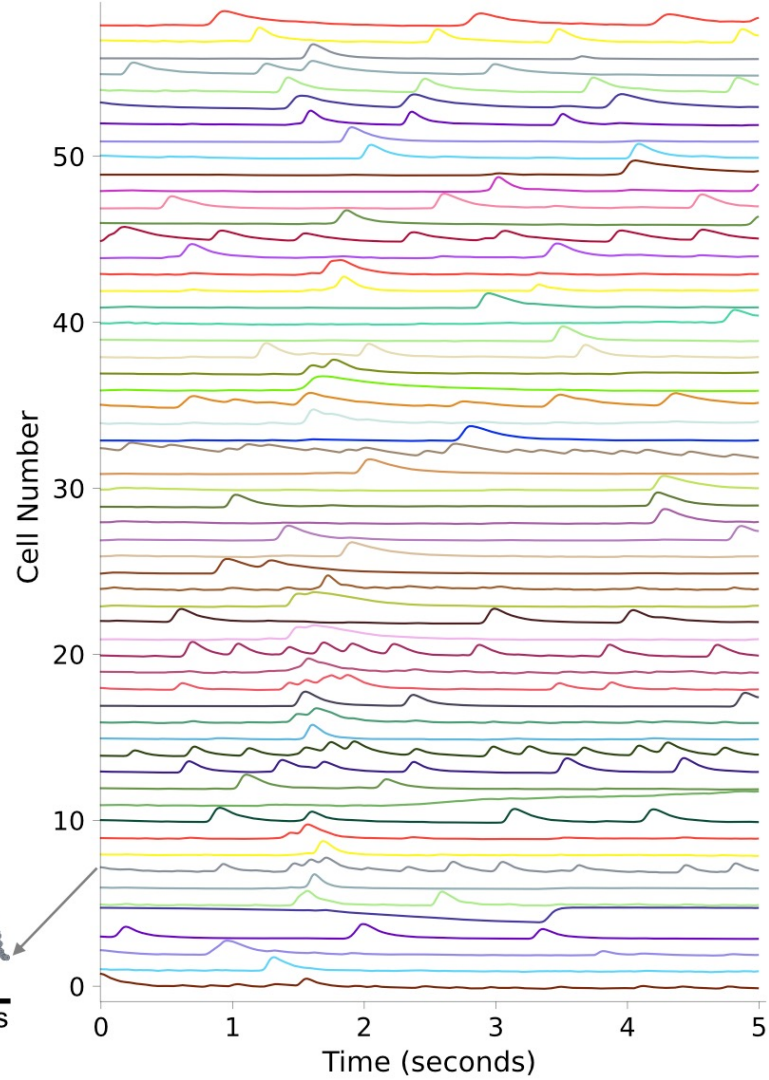
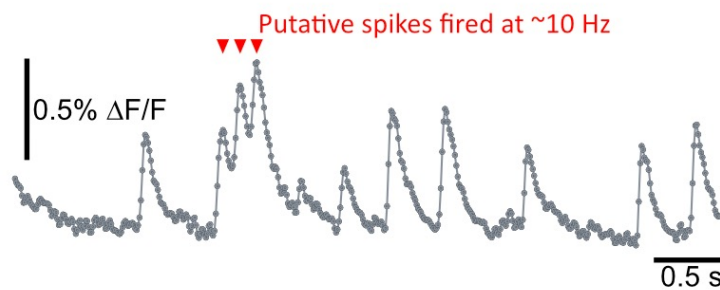
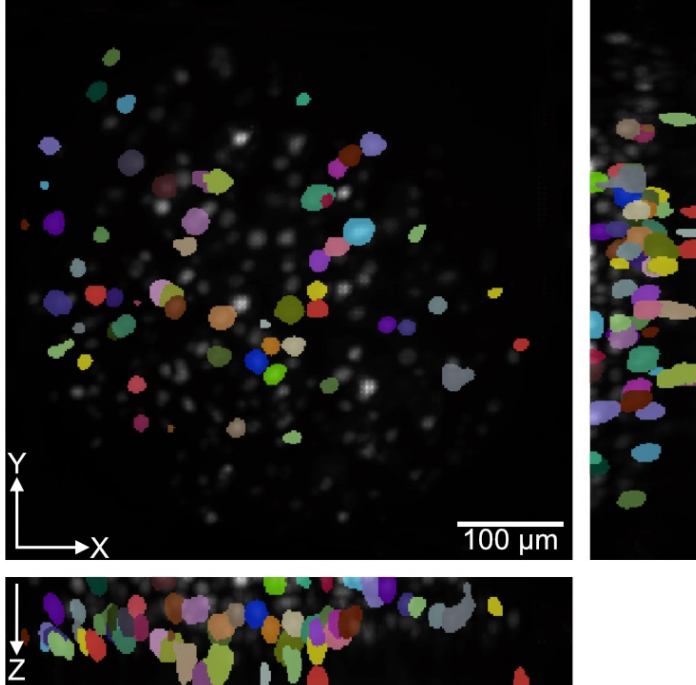


Fast volumetric jGCaMP8f time-series extraction

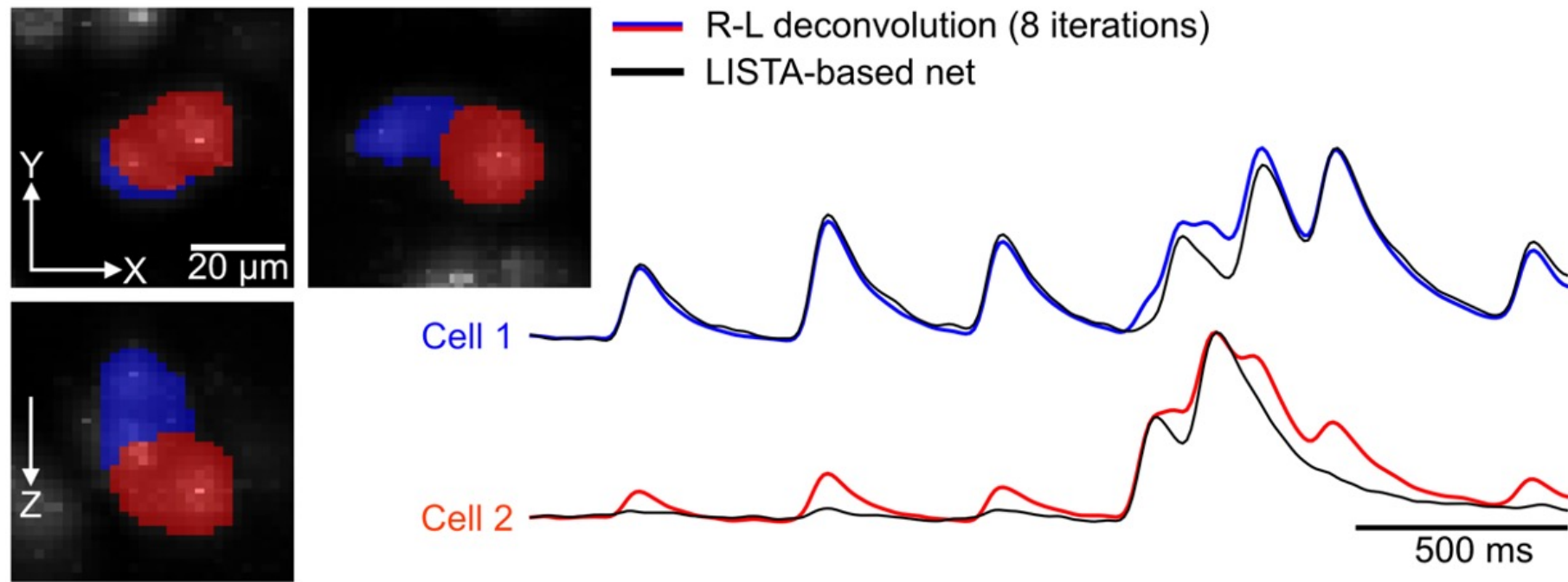


Fast volumetric jGCaMP8f time-series extraction





LISTA-based net decreases crosstalk between neighbouring neurons



- Goal: we want to separate the two x-ray images
- Approach:
 - Use the visible RGB image as side information (x-ray visible similar to RGB image)
 - Exclusion loss: the “contours” of the two x-ray images should be as different as possible



Visible



X-ray

$$\begin{aligned}
 x_1 &= \sum_{k=1}^K \Xi_k * z_{1,k}, & x_2 &= \sum_{k=1}^K \Xi_k * z_{2,k}, \\
 r_{1,s} &= \sum_{k=1}^K \Omega_{k,s} * z_{1,k}, & x &= \sum_{k=1}^K \Xi_k * (z_{1,k} + z_{2,k}),
 \end{aligned}$$

- The visible image and the two separated X-ray images have a sparse representation in proper dictionaries
- RGB image and visible X-ray share the same sparse representation
- The two X-rays x_1, x_2 share the same dictionary
- The measured X-ray is $x = x_1 + x_2$



Visible



X-ray

- Given the reconstructed X-ray images x_1, x_2 , we expect that their edges are as different as possible we therefore add an “exclusion term” in the optimization

$$\begin{aligned}
 \min_{\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_{1,k}, \mathbf{z}_{2,k}} & \quad \|\mathbf{x} - \Psi * \mathbf{y}_1 - \Psi * \mathbf{y}_2\|_F^2 \\
 & + \tau_1 \|\mathbf{y}_1 - \sum_{k=1}^K \Theta_k * \mathbf{z}_{1,k}\|_F^2 \\
 & + \tau_2 \|\mathbf{y}_2 - \sum_{k=1}^K \Theta_k * \mathbf{z}_{2,k}\|_F^2 \\
 & + \gamma \sum_{s=1}^3 \|\mathbf{r}_{1,s} - \Phi_s * \mathbf{y}_1\|_F^2 \\
 & + \lambda_1 \sum_{k=1}^K \|\mathbf{z}_{1,k}\|_1 + \lambda_2 \sum_{k=1}^K \|\mathbf{z}_{2,k}\|_1 \\
 & + \sum_{i=1}^I \mu_i \|(\mathbf{W}_i * \mathbf{y}_1) \odot (\mathbf{W}_i * \mathbf{y}_2)\|_1,
 \end{aligned}$$

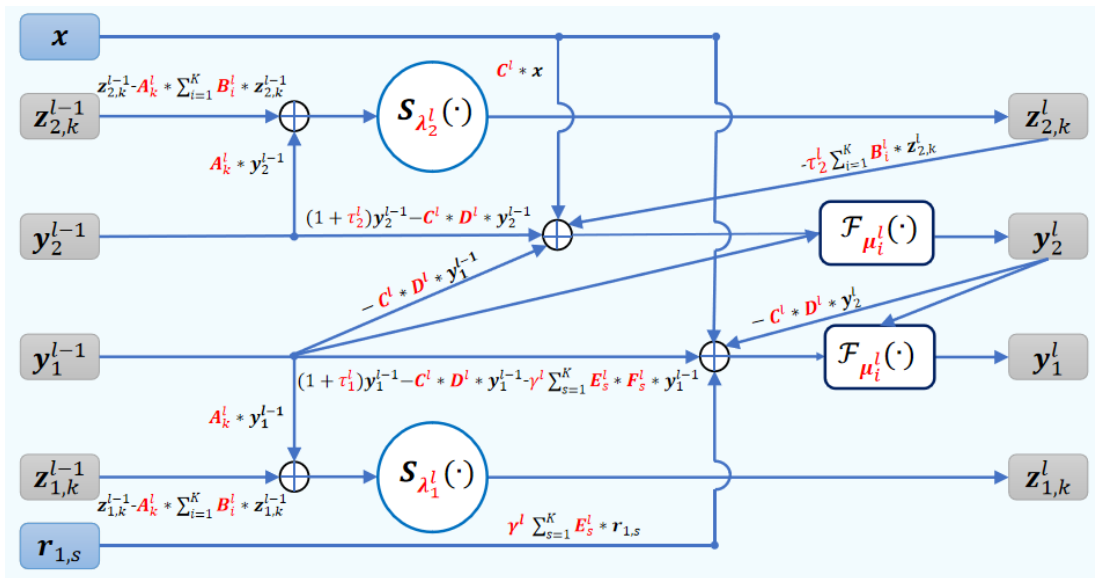


Visible



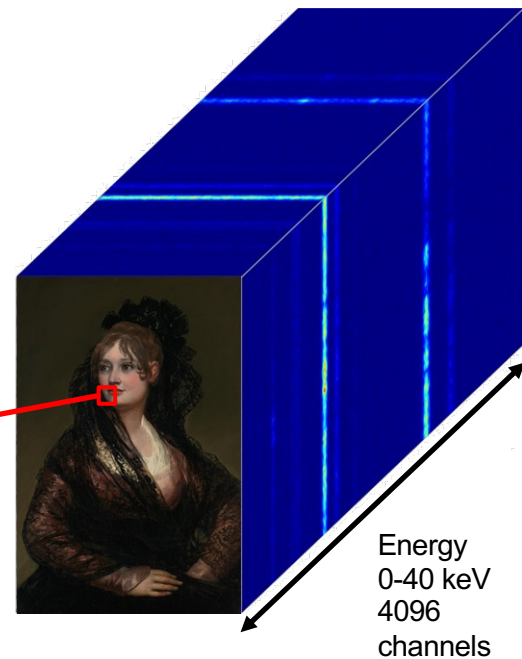
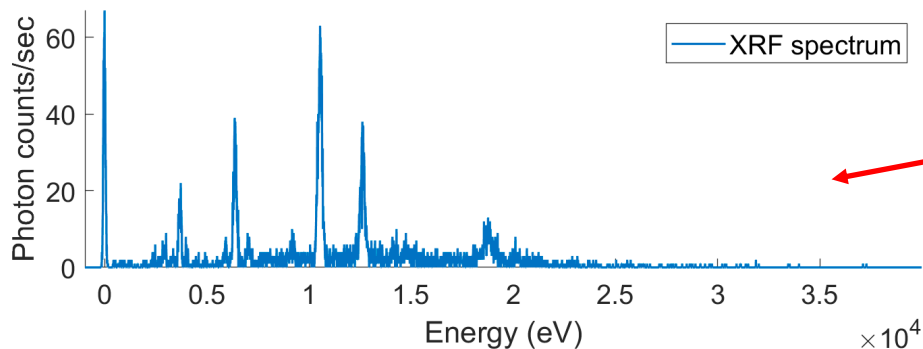
X-ray

- The sparsity model and the exclusion constraint leads to an iterative optimization method which leads to a network through unfolding





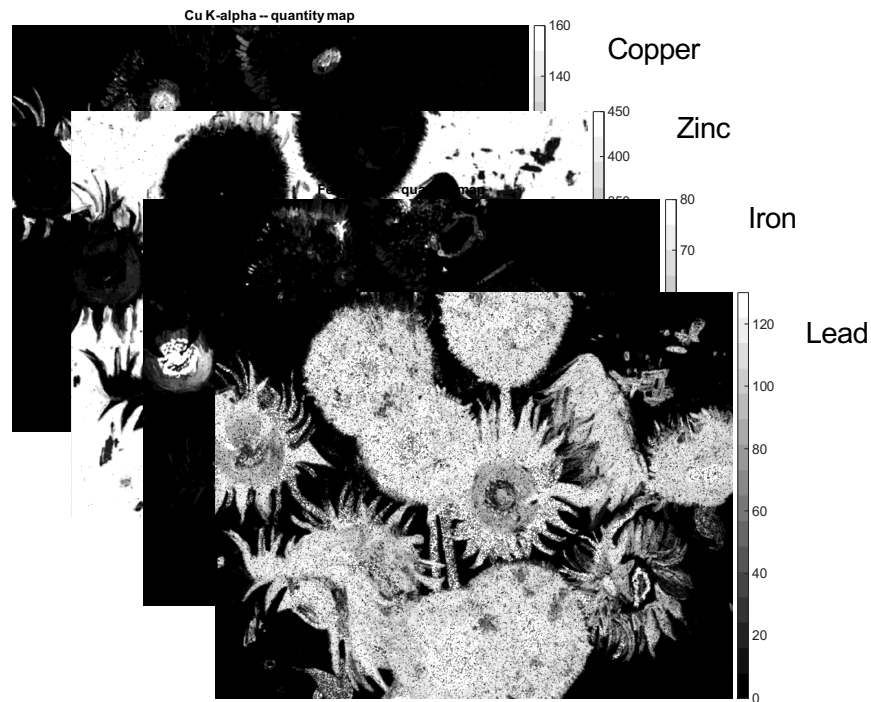
- Macro X-ray provides volumetric data and the locations of the pulses in the energy direction are related to the chemical elements present in the painting.
- This potentially allows us to create maps that show the distribution of different chemical elements



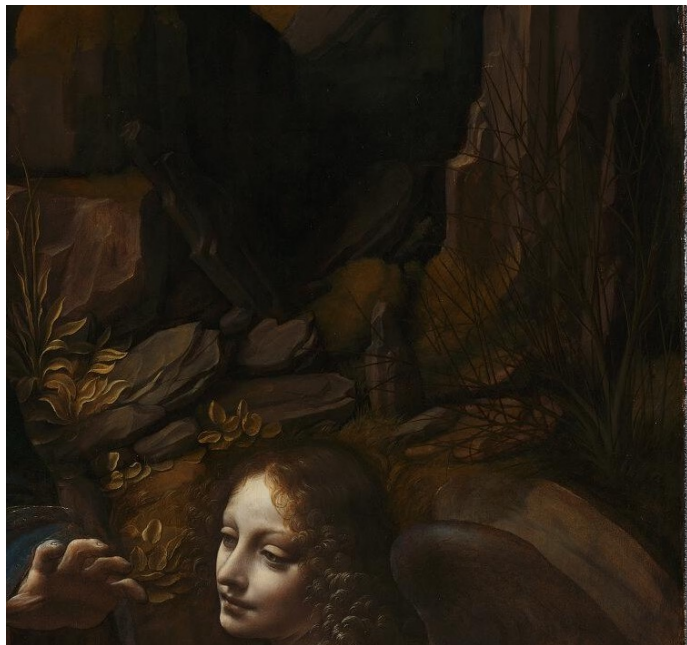


Our XRF
Deconvolution
Algorithm

Extraction of Elemental Maps

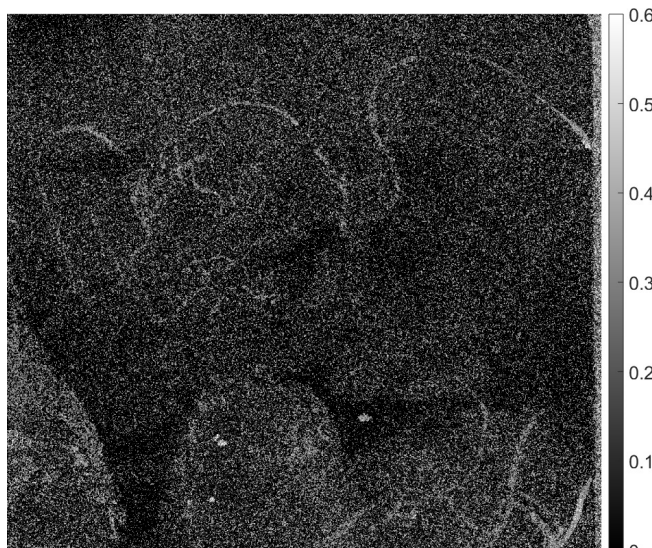
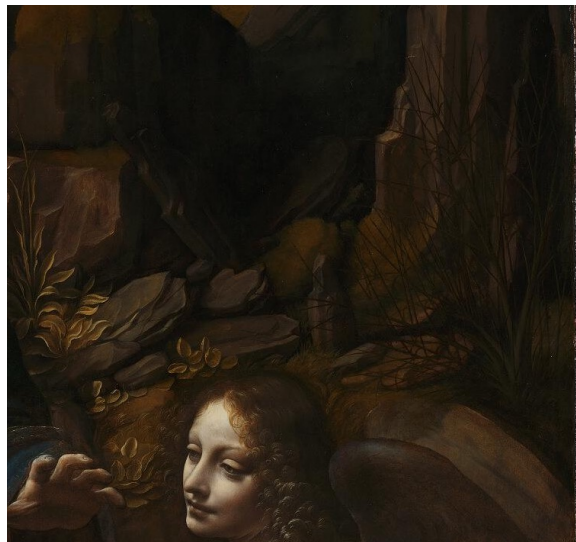


Leonardo da Vinci's "The Virgin of the Rocks"

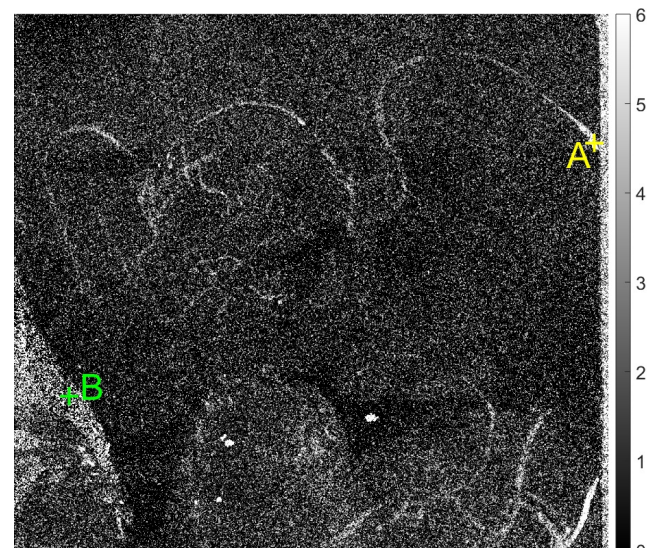


Highlighted is the region of an XRF dataset collected on the painting with an M6 Bruker JETSTREAM instrument (30 W Rh anode at 50 kV and 600 μA , 60 mm² Si drift detector, and data collected with 350 μm beam and pixel size and 10 ms dwell time).

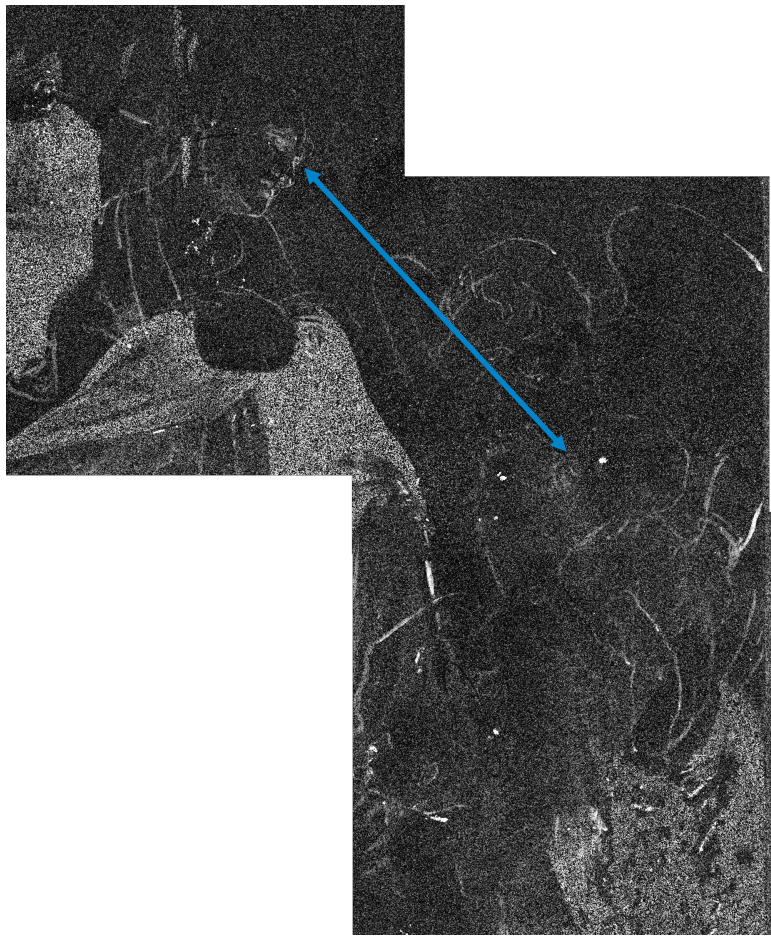
Zinc (Zn) distribution maps



Zn confidence map



Zn quantity map



- In imaging problems:
 - operating at the interface between maths/physics and computation is essential
 - Cross fertilization between model-based approaches and deep learning is fruitful
 - Some computational approaches are transferable
- Inverse imaging problems:
 - are fun 😊
 - and inter-disciplinary

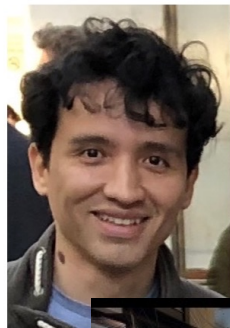
A special thank to:



Di You



Herman Verinaz



Kate Zhao



Pingfan Song



Peter Quicke

Carmel Howe

Amanda Foust



Junjie Huang

Consortium involving: UCL, ICL, Duke and
National Gallery

Thank you!

- Wavelet-inspired INN and Diffusion Models:
 - J. Huang and P.L. Dragotti, “WINNet: Wavelet-inspired Invertible Network for Image Denoising”, IEEE Transactions on Image Processing, 2022
 - D. You, A. Floros and P.L. Dragotti, “INDigo: An INN-Guided Probabilistic Diffusion Algorithm for Inverse Problems”, IEEE MMSP conference, September 2023, arxiv:2306.02949
 - D.You and P.L. Dragotti, “INDIGO+: A Unified INN-Guided Probabilistic Diffusion Algorithm for Blind and Non-Blind Image Restoration”, IEEE Journal of Selected Topics in Signal Processing, 2024
- Light-field Microscopy:
 - H. Verinaz et al. "Physics-based Deep Learning for Imaging Neuronal Activity via Two-photon and Light-field Microscopy", IEEE Trans. on Computational Imaging, 2023.
- Art Investigation
 - W. Pu, J. Huang et al., “Mixed X-Ray Image Separation for Artworks with Concealed Designs”, IEEE Transactions on Image Processing, 2022
 - S. Yan, J.-J. Huang, N. Daly, C. Higgitt, and P. L. Dragotti, “When de Prony Met Leonardo: An Automatic Algorithm for Chemical Element Extraction in Macro X-ray Fluorescence Data”, IEEE Transactions on Computational Imaging, vol.7, 2021.