

Scintillator Decorrelation for Self-supervised X-ray Radiograph Denoising

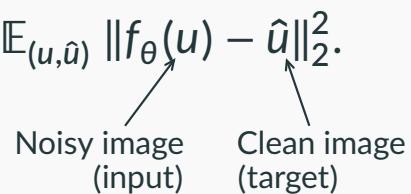
Adriaan Graas Felix Lucka
`adriaan.graas@cwi.nl` `felix.lucka@cwi.nl`

2025-07-31

INTRO: Blind-Spot Denoising



In *supervised denoising*, a network f_θ is trained with

$$\arg \min_{\theta} \mathbb{E}_{(u, \hat{u})} \|f_{\theta}(u) - \hat{u}\|_2^2.$$


Noisy image (input) Clean image (target)

- *Self-supervised denoising* replaces \hat{u} with $\hat{u} + \varepsilon$, where ε is zero-mean noise.

$$\mathbb{E} \|f_\theta(u) - (\hat{u} + \varepsilon)\|_2^2 = \underbrace{\mathbb{E} \|f_\theta(u) - \hat{u}\|_2^2}_{\text{supervised}} + \underbrace{\mathbb{E} 2\langle f_\theta(u) - \hat{u}, \varepsilon \rangle}_{\text{cross-term}} + \cancel{\mathbb{E} \|\varepsilon\|_2^2}$$

- When the cross-term cancels, **self-supervised denoising** → **supervised denoising**.

Noise2Noise

Inputs: image $\hat{u} + \varepsilon_1$

Targets: image $\hat{u} + \varepsilon_2$

Condition: $\varepsilon_1 \perp \varepsilon_2$.

- Loss: $\|f_\theta(\hat{u} + \varepsilon_1) - \hat{u}\|_2^2 + \mathbb{E} 2\langle f(\hat{u} + \varepsilon_1) - \hat{u}, \varepsilon_2 \rangle$. Works ✓.

[1] J. Lehtinen *et al.*, “Noise2Noise: Learning Image Restoration without Clean Data.” 2018.

- *Noise2Noise* was brought to tomographic domain by *Noise2Inverse* [2].
- Two noisy images = two noisy reconstructions each from half of the angles.

Noise2Inverse (simplified)

Inputs: FBP reconstruction $A_{\text{FBP}} \mathbf{y}_{i::2}$ from *odd* angles $i = (1, 3, \dots)$

Targets: FBP reconstruction $A_{\text{FBP}} \mathbf{y}_{i+1::2}$ from *even* angles $i = (0, 2, \dots)$

Condition: Noise uncorrelated between angles.

[2] A. A. Hendriksen, D. M. Pelt, and K. J. Batenburg, “Noise2Inverse: Self-Supervised Deep Convolutional Denoising for Tomography,” 2020.

- *Blind-spot denoisers* (BSNs) split input and target from a single image, via masking of pixels.
- Prominent framework **Noise2Self**, where it is formalized through *J-invariance*.

Noise2Self

Inputs: pixels u_{J^c} — an image with “blind spots”, i.e. J pixels masked out

Targets: pixels u_J — the blind spots J

Condition: $\varepsilon_i \perp \varepsilon_j$ for all pixels (i, j)

$$\bullet \text{ Loss: } \|f_\theta(\hat{u}_{J^c} + \varepsilon_{J^c}) - u_J - \varepsilon_J\|_2^2 + \cancel{\mathbb{E} 2\langle f_\theta(\hat{u}_{J^c} + \varepsilon_{J^c}) - \hat{u}_{J^c}, \varepsilon_J \rangle}$$

- A (small) performance loss due to masking of input pixels.

[3] J. Batson and L. Royer, “Noise2Self: Blind Denoising by Self-Supervision.” 2019.

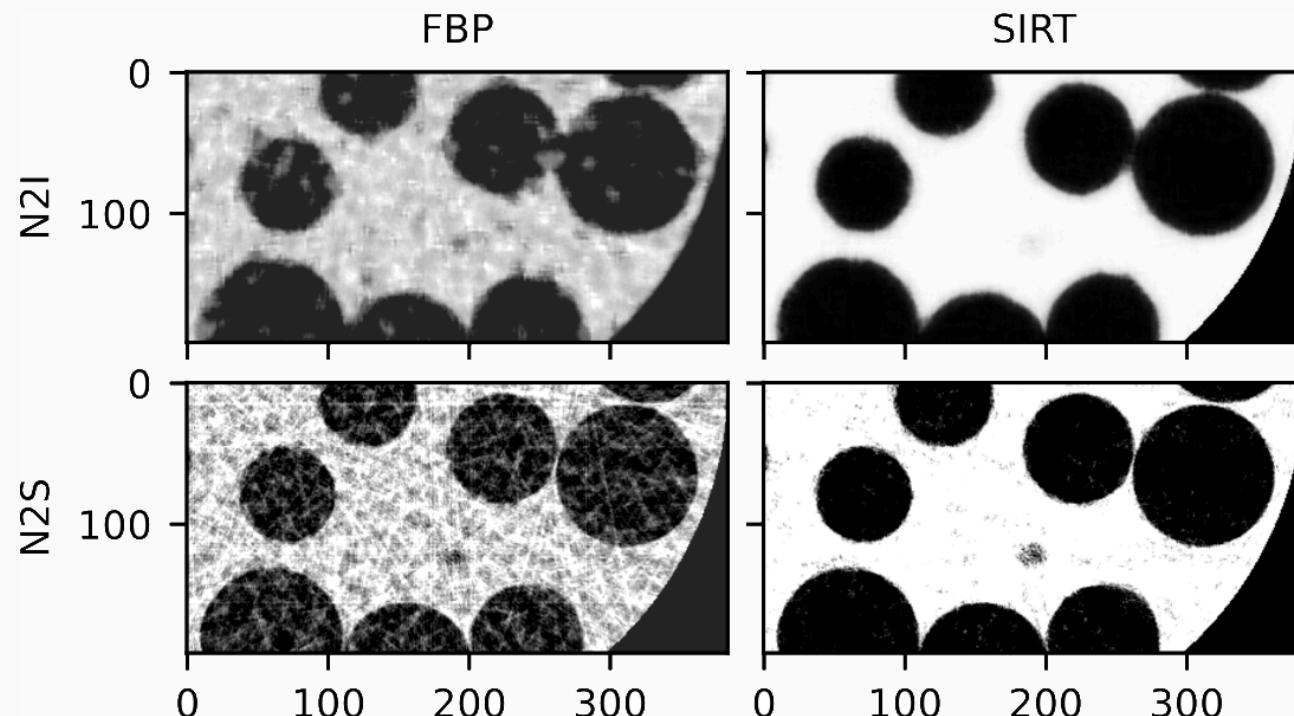
MOTIVATION: Denoising Radiographs for CT



- Denoising *radiographs* (raw projections) for Computed Tomography (CT)? 😊
 - No usage of the X-ray projectors A and A^T .
 - Image features are local in reconstruction.
- Wait... 🤔
 - 💡 Sparse-view / limited-wedge / high-noise:
 - A^T alone gives poor image features.
 - Both A and A^T (unrolled methods)? Not feasible with high-resolution 3D CBCT.
 - 💡 Noise remains local in radiographs → globally correlated in reconstructions:
 - Take loss in projection domain!

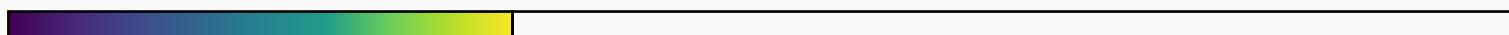
- Empirical evidence, let's compare:
 - **Volume denoising:** FBP or SIRT → then N2I (*Noise2Inverse*)
 - **Radiograph denoising:** N2S (*Noise2Self*) → then FBP or SIRT
 - Data set:
 - Simulated radiographs of foam [4] + Poisson noise.
 - Sparse-view reconstruction with 32 angles.
-

[4] D. M. Pelt, A. A. Hendriksen, and K. J. Batenburg, “Foam-like phantoms for comparing tomography algorithms,” Jan. 2022.



- **Highest MSE:** N2S → SIRT (bottom right)
- Noise2Inverse cannot handle global noise after SIRT (top right)

THE PROBLEM: Correlated Noise



- Many detectors use *scintillator crystals*, e.g. Caesium-Iodine (CsI).
- X-ray interaction lead to a *photon avalanche*.
 - One-to-many effect: 100s to 1000s of visible-light photons per X-ray photon.



Figure 2: A large-area Caesium-Iodine scintillator detector:
Teledyne DALSA Xineos-3131

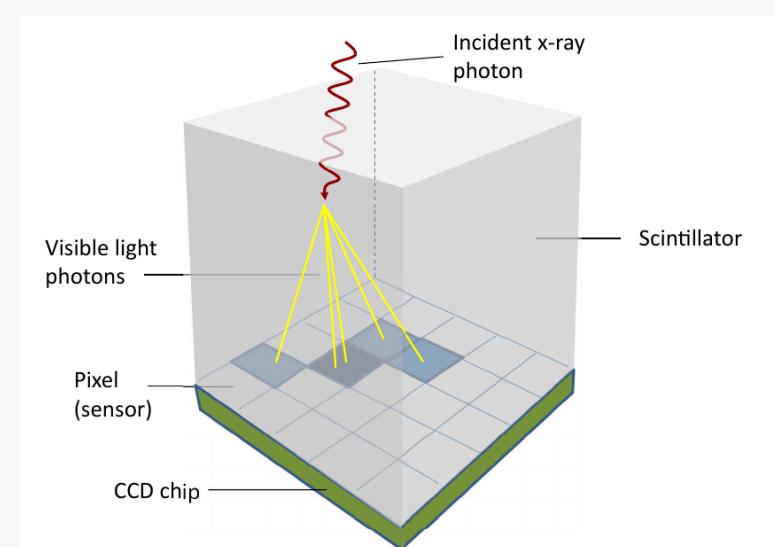


Figure 3: Principle of X-ray detection. Figure reproduced
from Just Enough Physics (Chapter 4 in [5]).

[5] P. C. Hansen, J. S. Jørgensen, and W. R. Lionheart, Eds., *Computed Tomography: Algorithms, Insight, and Just Enough Theory*. 2021.

- Recall cross-term. When noise correlates pixelwise,

$$\mathbb{E}[\varepsilon_J \varepsilon_{J^c}] \neq 0 \Rightarrow \mathbb{E}\langle f_\theta(u_{J^c})_J - \hat{u}_J, \varepsilon_J \rangle \neq 0,$$

it does **not** vanish. Result: **blurry, noisy images**.

- Existing strategies:
 - *AP-BSN* [6], *M-Denoiser* [7]:
 - **Idea:** Subsampling → Denoise subimages → Upsample.
 - **Problem:** Upsampling artifacts & aliasing.
 - *Noisier2Noise* [8]:
 - **Idea:** Train to remove added simulated noise → Inference on original noisy image.
 - **Problem:** overfits to noisy images / noise model required / interpolation invalid.

[7] X. Chong, M. Cheng, W. Fan, Q. Li, and H. Leung, “M-Denoiser: Unsupervised image denoising for real-world optical and electron microscopy data,” 2023.

[6] W. Lee, S. Son, and K. M. Lee, “AP-BSN: Self-Supervised Denoising for Real-World Images via Asymmetric PD and Blind-Spot Network.” 2022.

[8] N. Moran, D. Schmidt, Y. Zhong, and P. Coady, “Noisier2Noise: Learning to Denoise from Unpaired Noisy Data.” 2019.

- For real-world X-ray noise model:

$$I = h \circledast \text{Poisson} \left(\exp \left(- \int_{I_i} \mu(\eta) d\eta \right) \right) + \text{Gaussian} \left(0, \underbrace{\text{diag}(\sigma_1^2, \dots, \sigma_N^2)}_{\text{zero-mean, anisotropic variance}} \right)$$

Beer-Lambert's law

- “The radiograph I is
 1. **signal-dependent** Poisson noise on Beer-Lambert’s law...
 2. ...blurred by a **uniform convolution** “ $h \circledast$ ” of the point-response function (PRF)...
 3. ...further polluted by **additive, independent, anisotropic Gaussian noise.**”

- At sufficient photon counts, this estimates a multi-variate Gaussian:

$$I \sim \mathcal{N}(H\lambda, H\Lambda H^T + \Sigma)$$

- $H \in \mathbb{R}^{N \times N}$ convolution operator, $\lambda \in \mathbb{R}^N$ ground truth image
- $\Lambda \in \mathbb{R}^{N \times N}$ diagonal of ground truths, $\Sigma \in \mathbb{R}^{N \times N}$ diagonal of Gaussian variances

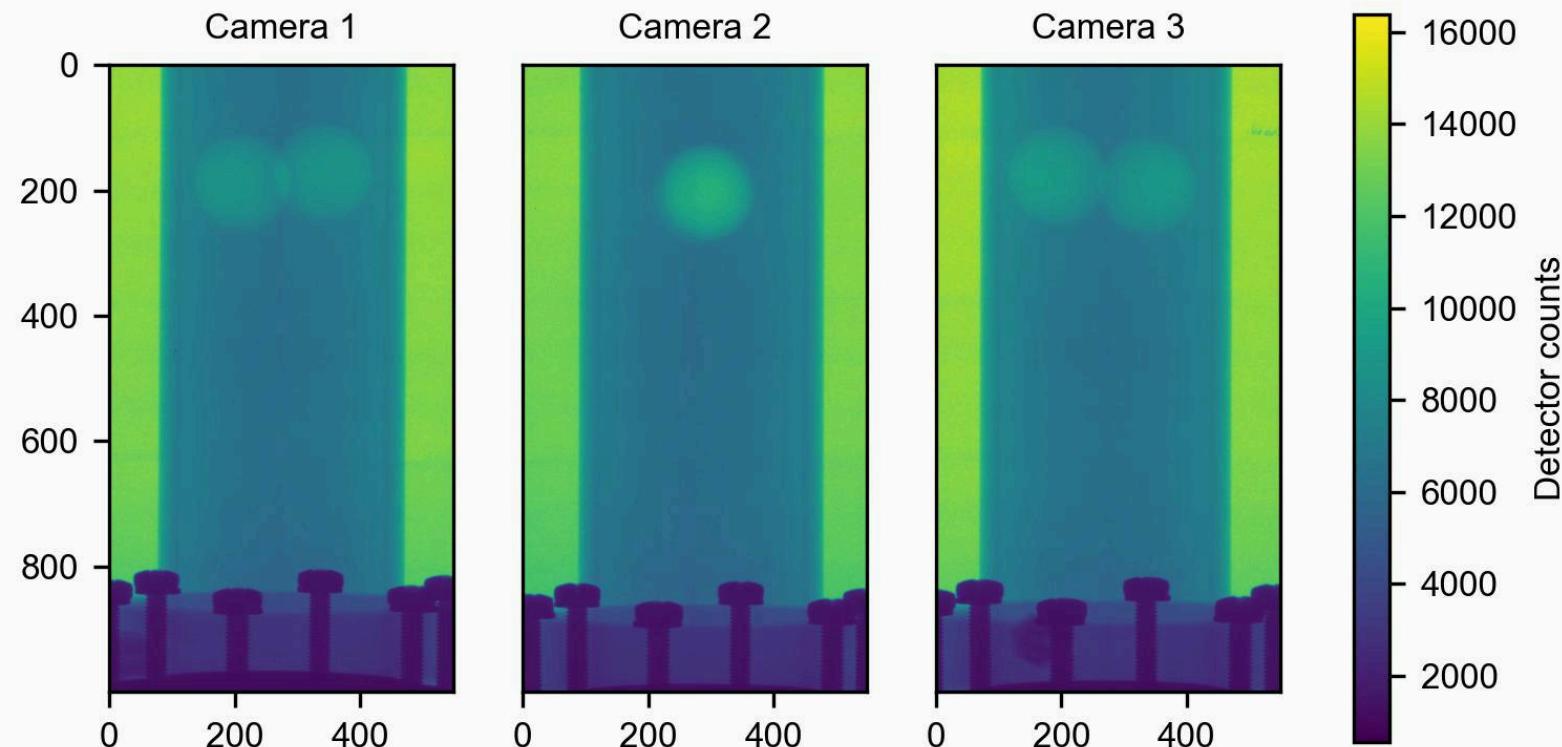


Figure 4: Radiographs of phantoms measured with three Teledyne DALSA Xineos-3131 detectors.
Contains: background, metal screws, polystyrene balls, PMMA column with particles. [9]

[9] A. Graas, E. Wagner, and F. Lucka, "Fluidized-bed-phantom radiographs from three Caesium-Iodine DALSA Xineos-3131 detectors for empirical PRF estimation." Zenodo, 2025.

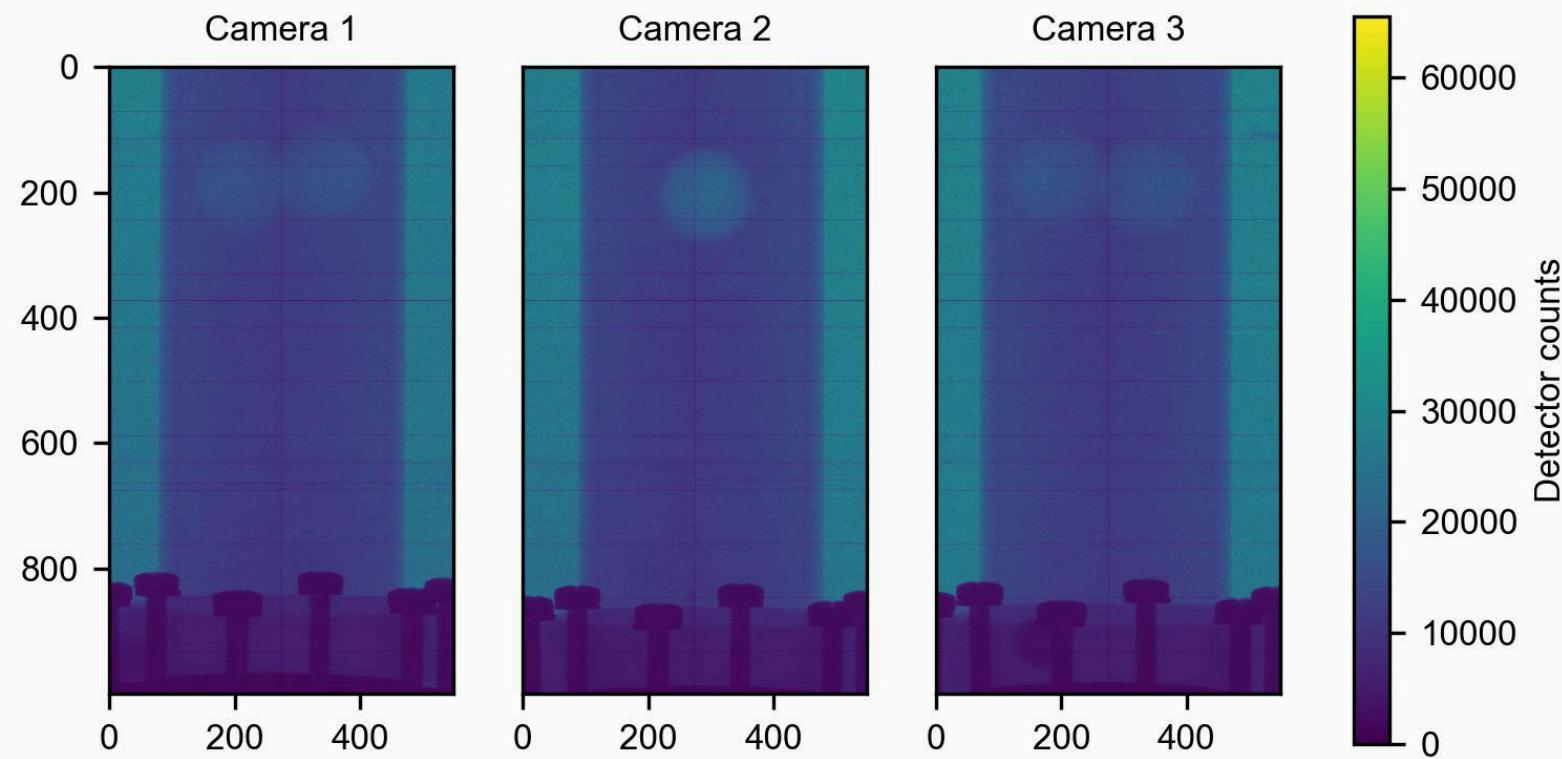


Figure 5: Pixelwise **empirical variance** of the radiographs using 1000 samples.
Signal-dependence: the variance resembles the mean.

METHOD: PRF Estimation & Direct Deconvolution



1.
**Estimate
correlation**
↑

2.
Direct
deconvolution

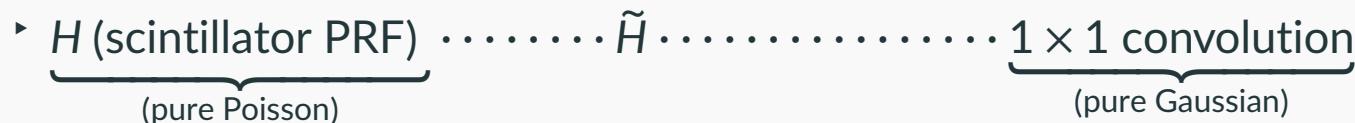
3.
Noise2Self
training

- Recall $I \sim \mathcal{N}(H\lambda, H\Lambda H^T + \Sigma)$. **Assumption:** Covariance is well-approximated by auto-correlation (double convolution),

$$\text{find a } \tilde{H} \text{ s.t. } H\Lambda H^T + \Sigma \approx \tilde{H}\Lambda\tilde{H}^T,$$

with \tilde{H} a convolution operator.

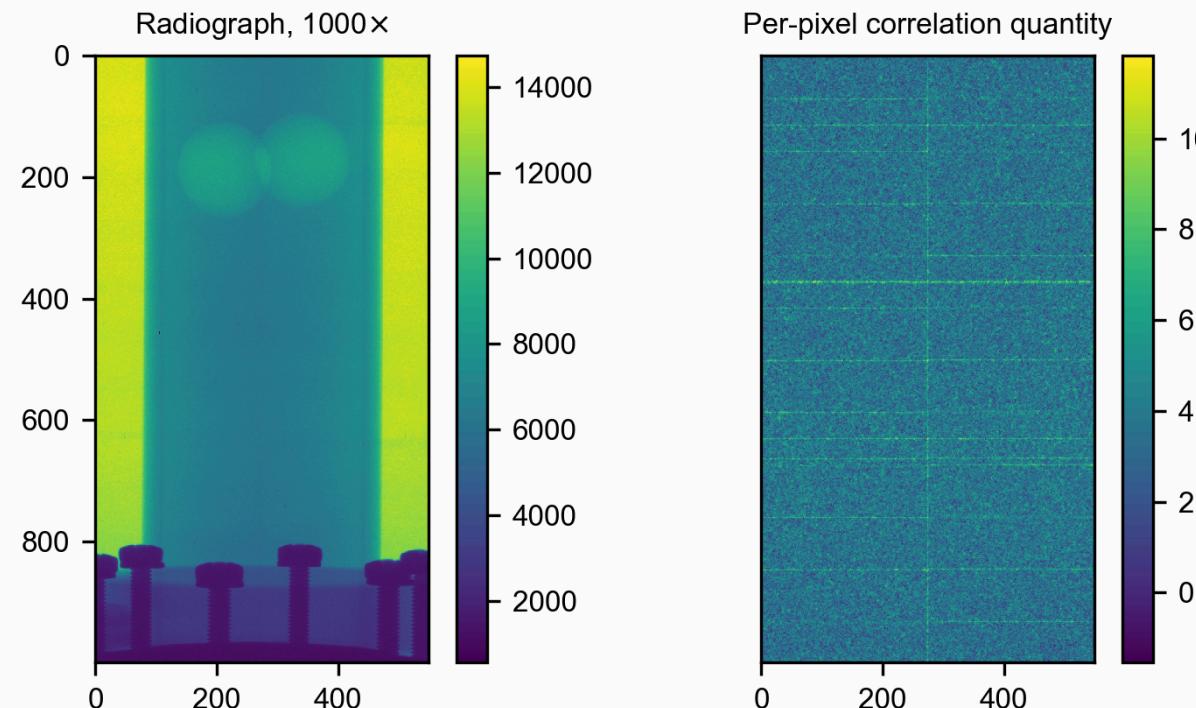
- Perfect \tilde{H} does not exist! However, it balances:



- Assumption holds when:
 - Gaussian noise is small and not too anisotropic.
 - The PRF is uniform; limited effect of energy-dependence (Swank noise) and incident X-ray angle.

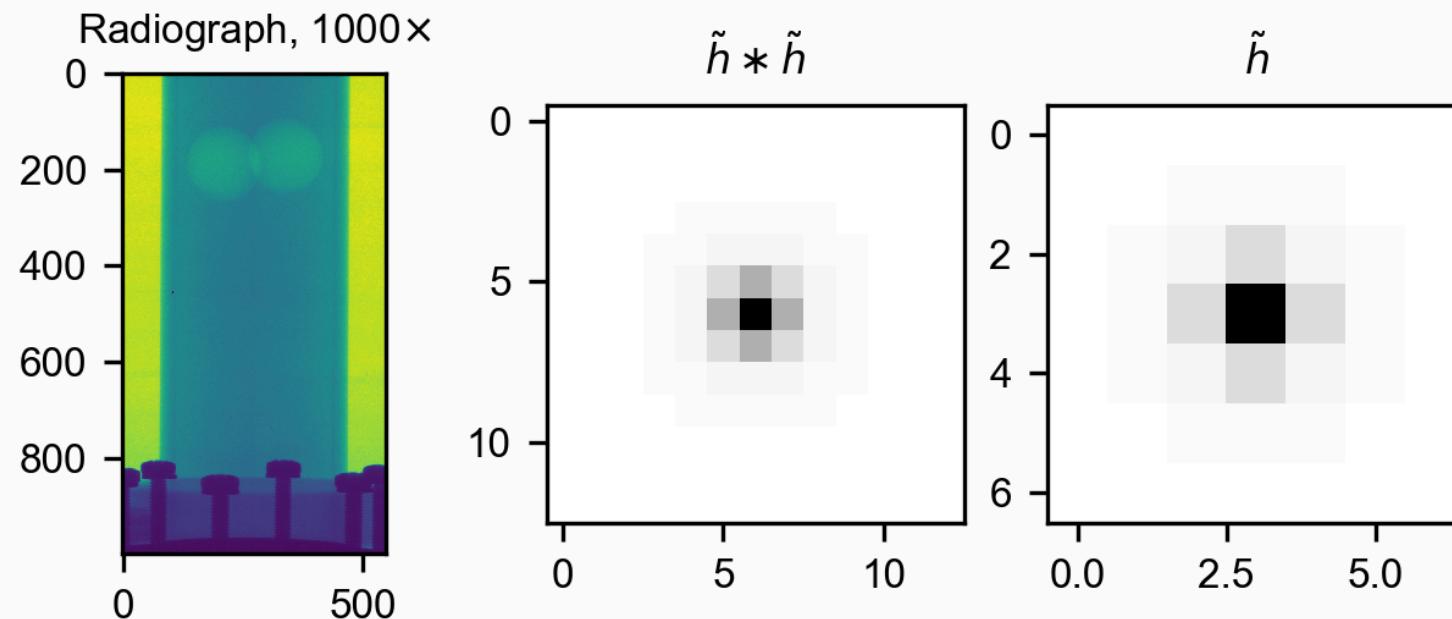
- Assumption valid in practice? Check correlations in static data.

- Per-pixel summed Pearson correlation with neighbors, $\sum_{j \neq i} \frac{\text{Cov}(I_i, I_j)}{\sqrt{\text{Var}(I_i) \text{Var}(I_j)}}$,



- For this set-up (cone beam, source voltage 120 kVp, Xineos-3131, ...),
 - **correlation is spatially homogeneous: about correlation ~350%**
 - regardless of metal objects, background intensity
 - **signal-dependent noise, but no strong signal-dependent correlation**
 - scintillator blur uniform, Gaussian noise not too anisotropic ✓

- How to compute \tilde{H} ?
 - Since convolution is uniform: estimate for an average pixel.
 - Acquire multiple noisy realizations from **static data / static detector region**, preferably **not background radiation**.
- Algorithm:
 1. Estimate $\tilde{h} \circledast \tilde{h}$ (auto-correlation kernel) via sampling.
 2. Set-up an artificially small covariance matrix \tilde{C} .
 3. Solve $\tilde{H} = \sqrt{\tilde{h}\tilde{h}^T} = \sqrt{\tilde{C}}$ using matrix factorization.
 4. Extract a row/column of \tilde{H} : the kernel \tilde{h} .



METHOD: PRF Estimation & Direct Deconvolution



1.
Estimate
correlation
2.
Direct
deconvolution
↑
3.
Noise2Self
training

- Perform a Fourier-based deconvolution (a.k.a. an *inverse filter*):

$$\tilde{H}^+ I := \mathcal{F}^{-1} \left(\frac{\mathcal{F}(I)}{|\mathcal{F}(\tilde{h})|^2} \right).$$

- Regularization, e.g. a Wiener filter, should not be used
 - Inverse filter amplifies noise (but, that is okay, we want decorrelation).
- Compared to iterative deconvolution (e.g. ADMM, Richardson-Lucy)
 - No scary “ringing” artifacts, these are typical for sharpness criteria.
 - Computationally efficient due to FFT.

- Per-pixel correlation maps **before and after deconvolution.**

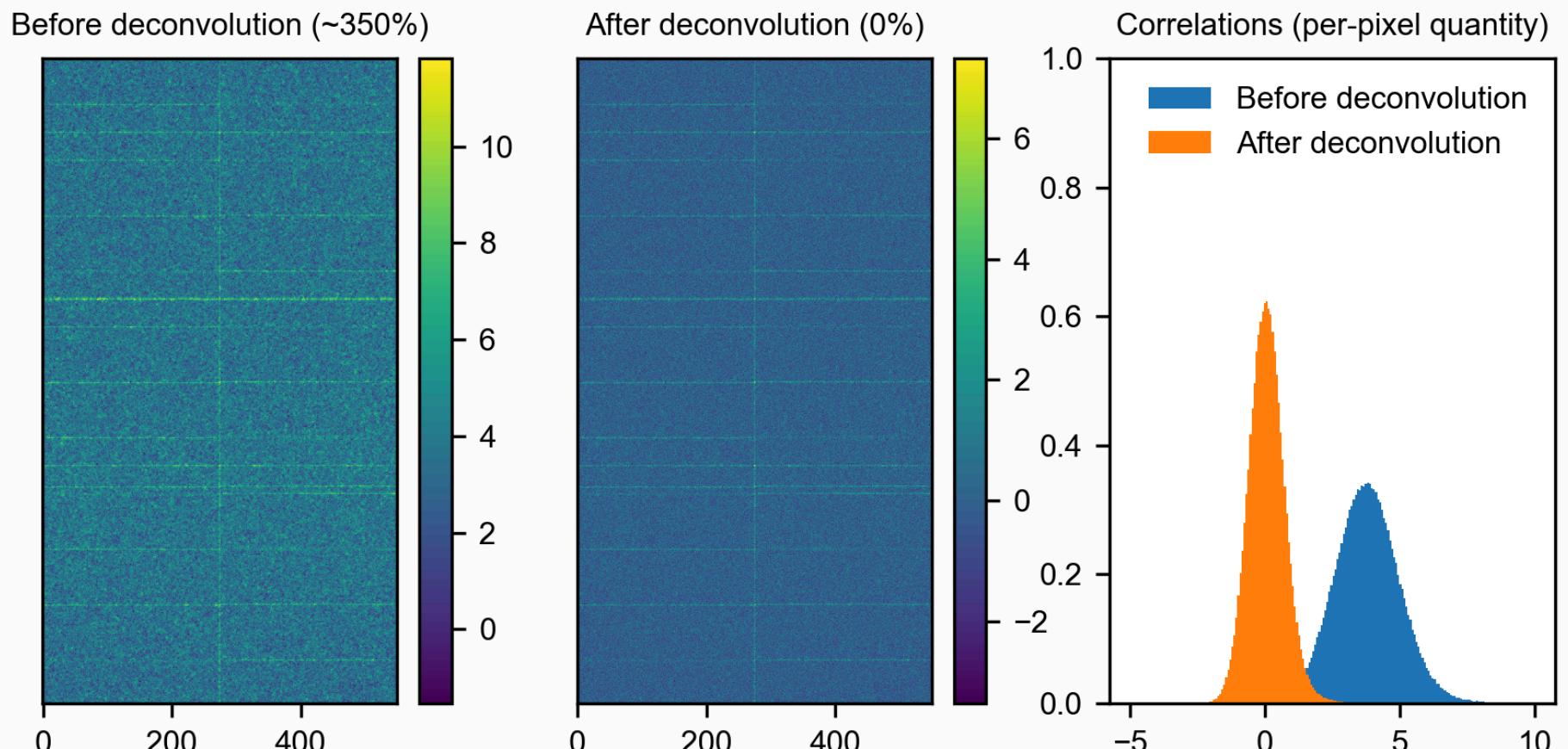


Figure 8: Deconvolution \Rightarrow Shift in correlation distribution

METHOD: PRF Estimation & Direct Deconvolution



1.
Estimate
correlation

2.
Direct
deconvolution

3.
Noise2Self
training
↑

- Noise2Self loss, f_θ a U-Net, in log-domain:

$$\arg \min_{\theta} \mathbb{E} \|\log f_\theta(\textcolor{blue}{u}) - \log \textcolor{red}{u}'\|_2^2.$$

- Here,
 - Input: $\textcolor{blue}{u} = (\tilde{H}^+ I)_{J^c}$
 - Target: $\textcolor{red}{u}' = (\tilde{H}^+ I)_J$

- Result: deconvolution achieves similar accuracy as uncorrelated

 - Training data set: $1000 \times$ same phantom radiograph.

 - Comparisons:** raw data (**left**), simulated uncorrelated data (**right**).

 - Same accuracy (MSE) and same “neighborhood-dependency” $p_i = \mathbb{E}_\varepsilon \sum_j \frac{\text{Cov}(f_\theta(\hat{u} + \varepsilon)_i, (\hat{u} + \varepsilon)_j)}{\sqrt{\text{Var}(\varepsilon_i)} \sqrt{\text{Var}(\varepsilon_j)}}$

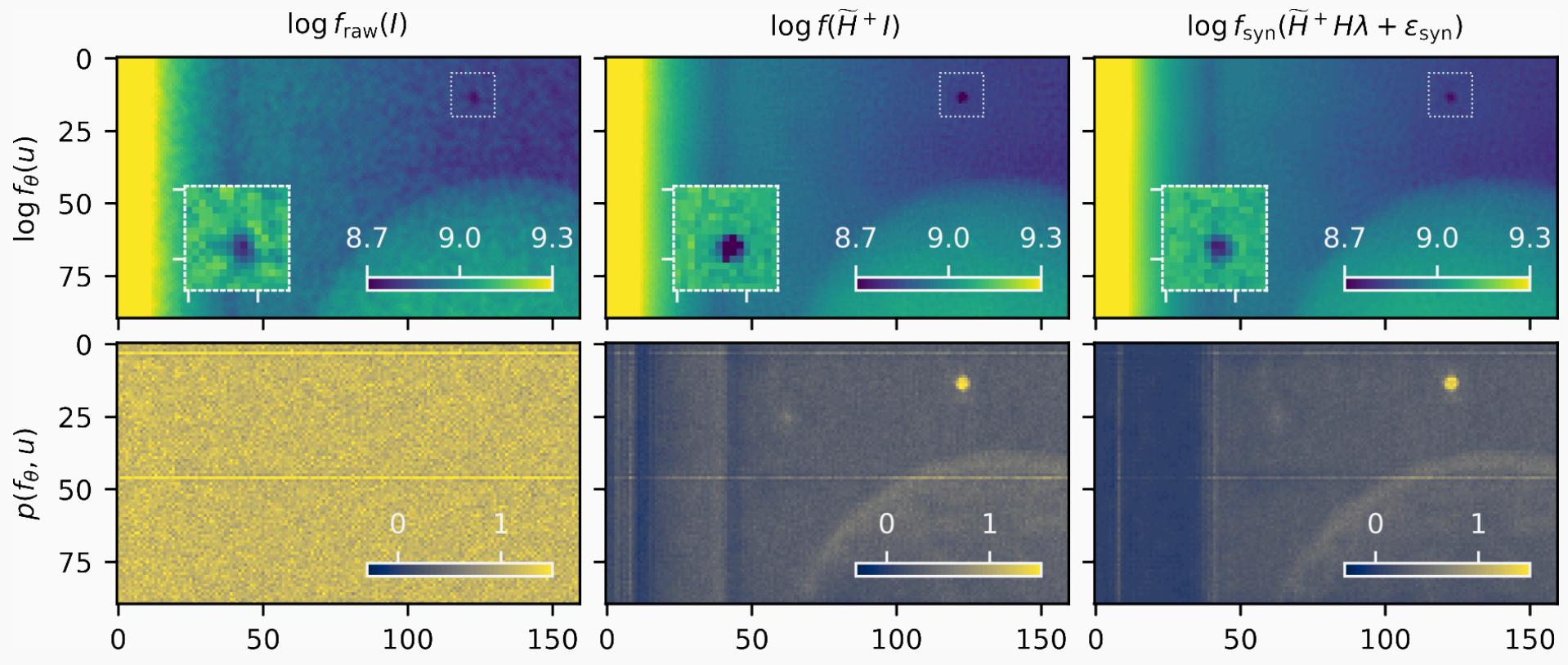


Figure 9: Raw vs. Real-world vs. Uncorrelated

- Watch out for overfitting to residual correlations

- Dead-pixel inpainting causes artifacts
- Occur in overfitting stage during training (stop in time?)

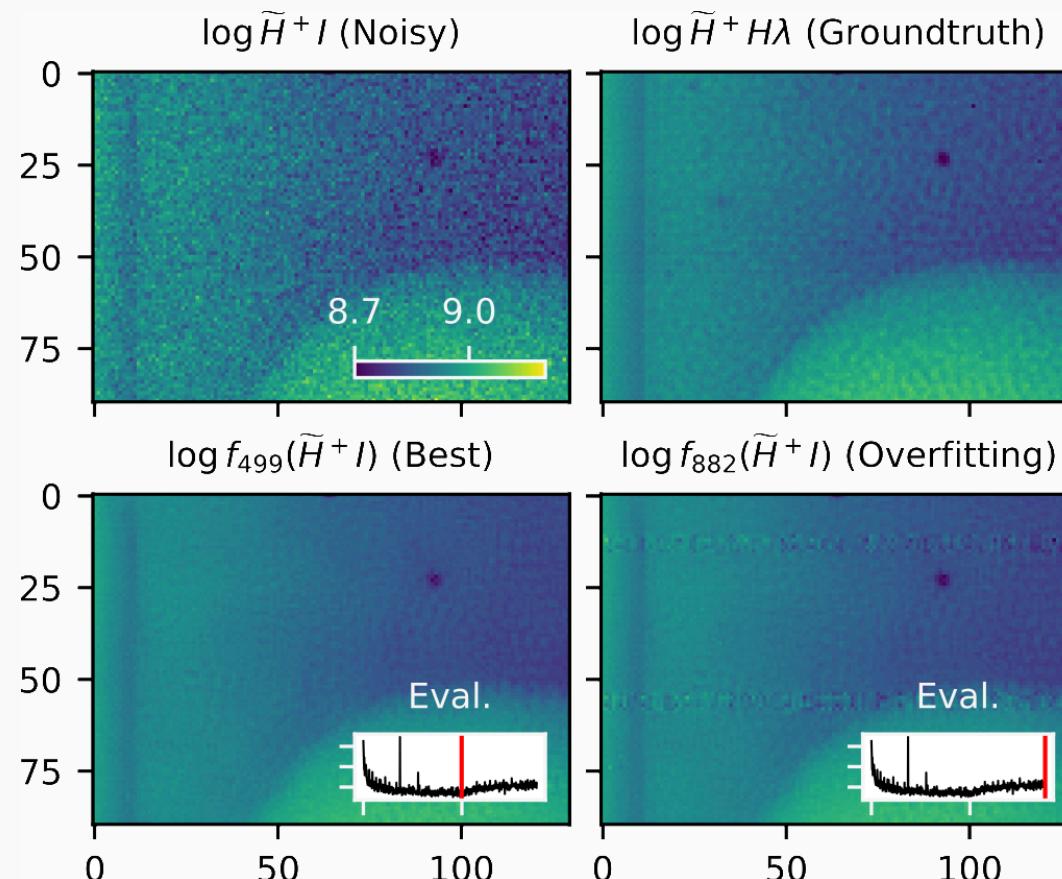


Figure 10: Blurry pixel rows after overfitting, in lower-right plot.

EXPERIMENTAL RESULTS: Fluidized Beds



- Fluidized beds: dynamic mixture of particles and gas, like a fluid.
- **No groundtruths or paired noisy images + sparse-view situation (section: Motivation).**

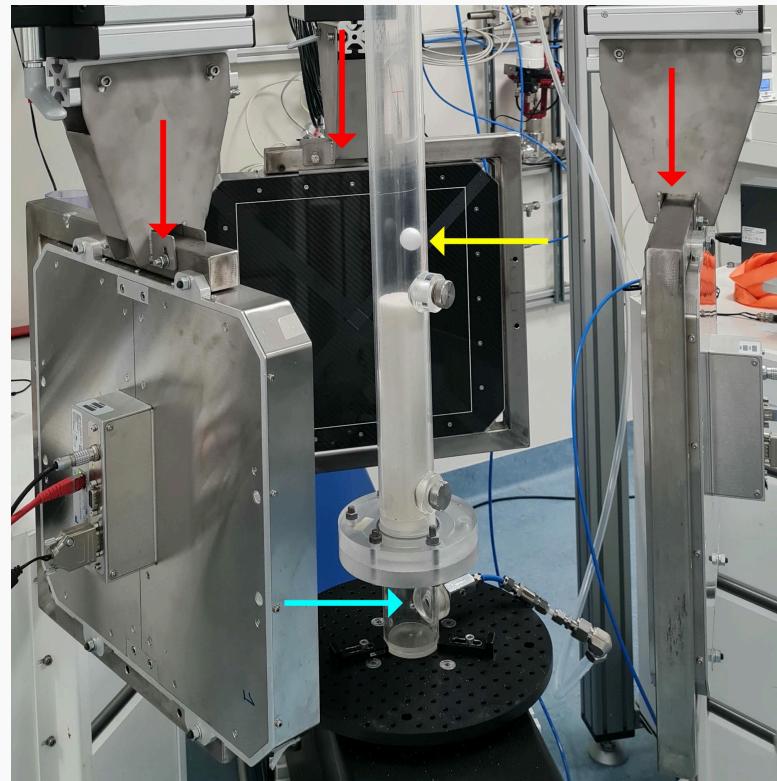
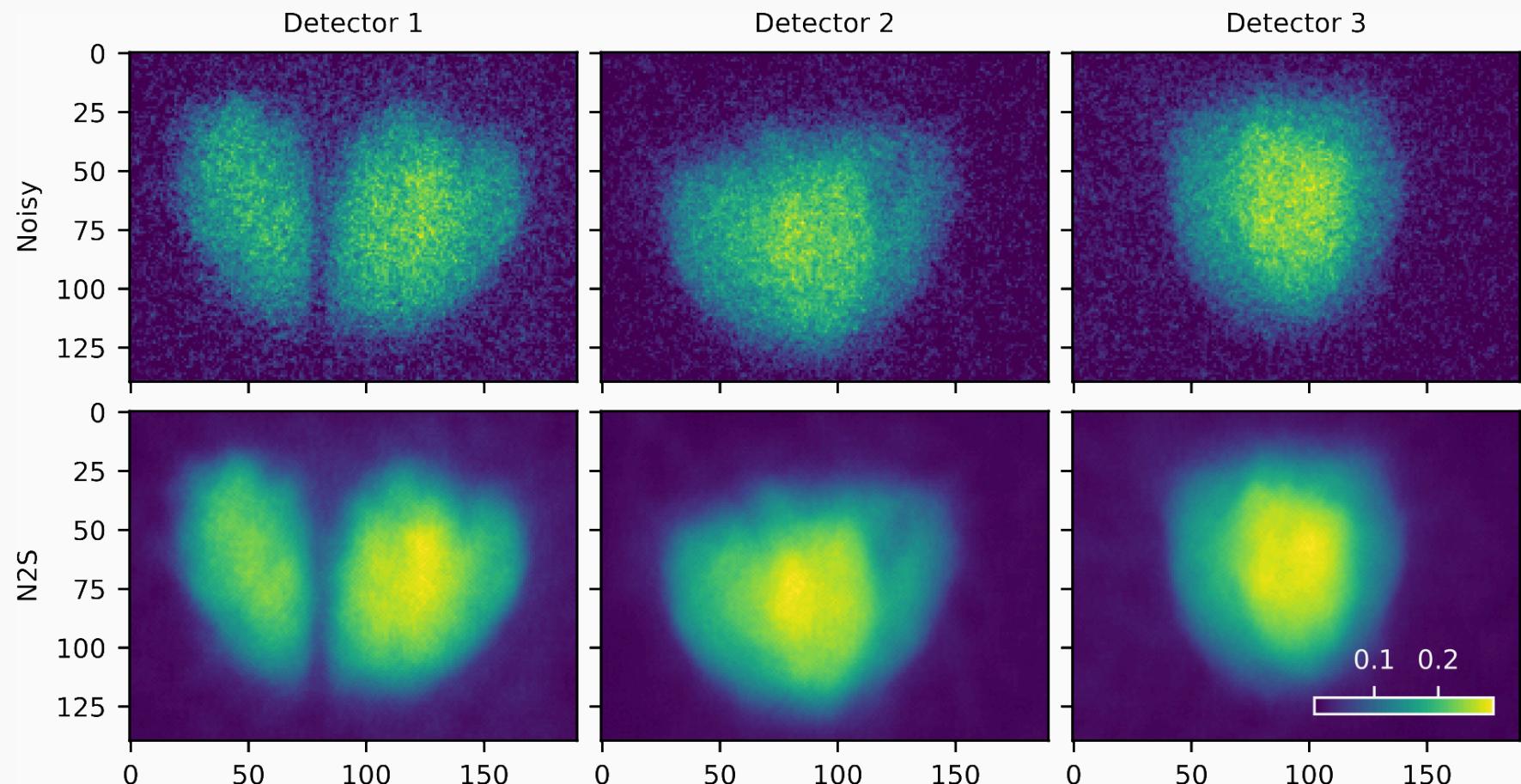


Figure 11: Set-up with three source-detector pairs at Delft University of Technology [10].

[10] A. Graas *et al.*, “X-ray tomography for fully-3D time-resolved reconstruction of bubbling fluidized beds,” *Powder Technology*, 2024.



- 3D reconstruction using SIRT (with $[0, 1]$ box constraints)

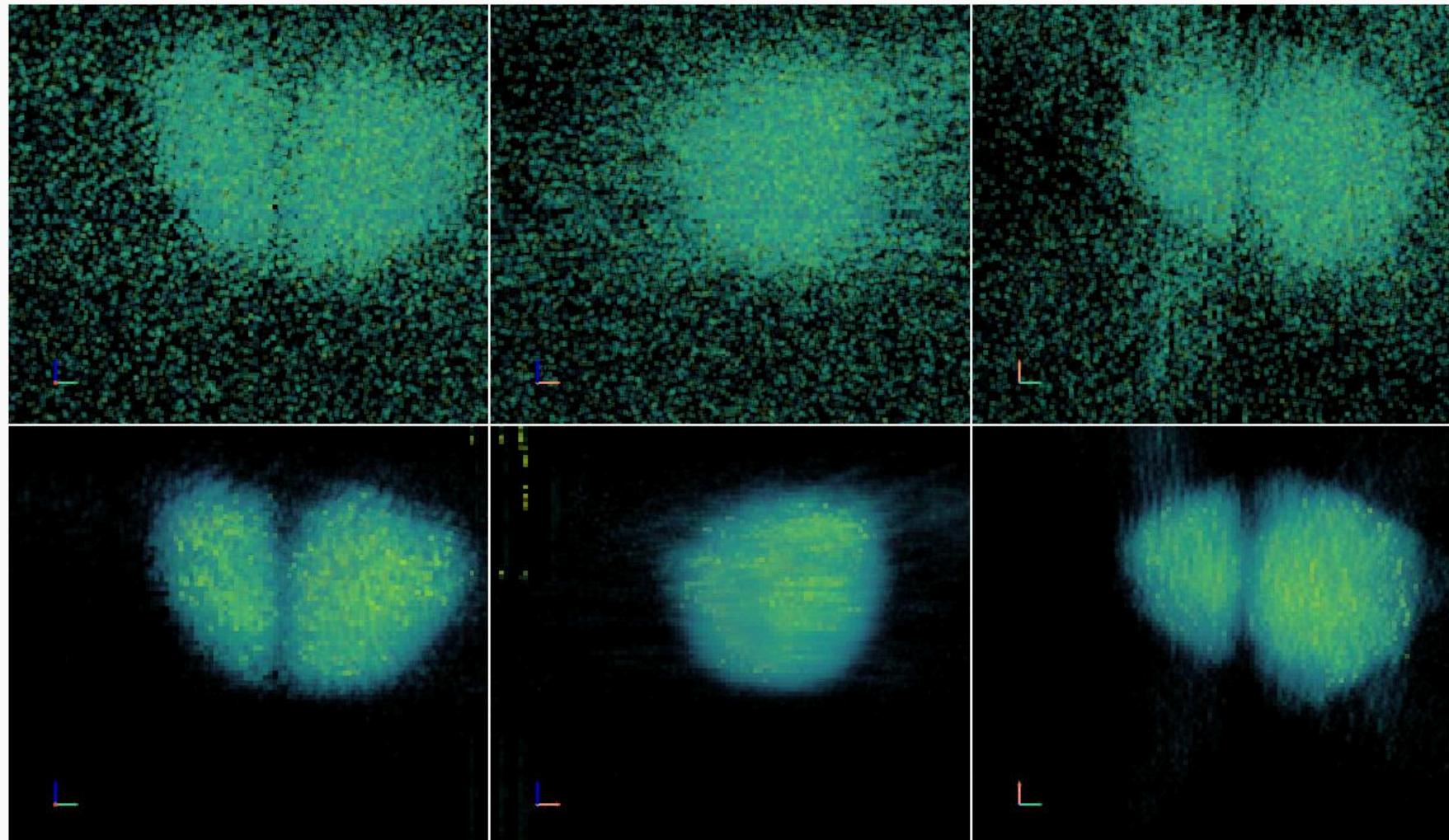


Figure 13: 3D SIRT with and without denoising, “camera view” from three axes.

Discussed

- A simple deconvolution workflow to get Noise2Self to work on radiographs.
- Promising for radiograph denoising, sparse-view CT.

Future work

- Work In Progress: Application to 2DeteCT data set [11], Equivariance2Inverse, by Dirk Schut.
- Extension to non-uniform noise model.
- Mitigate effects of additive Gaussian noise.
- Investigation of BSNs + radiograph loss + self-supervised CT.

 **Code online:** github.com/adriaangraas/scintillatordecorrelator

[11] M. B. Kiss, S. B. Coban, K. J. Batenburg, T. van Leeuwen, and F. Lucka, “2DeteCT - A large 2D expandable, trainable, experimental Computed Tomography dataset for machine learning,” *Scientific Data*, 2023.