Integrating a Physical Prior into a Machine Learning Model for Quantum Noise Reduction in X-ray Imaging:

Application to Spectral Attenuation Ratio Estimation in Dual-Energy X-ray Tomography

Grégory JEAN - Independent Researcher

Preprint - submitted to arXiv / HAL / Zenodo / May 2025

Abstract

Deep learning methods for X-ray image denoising are predominantly data-driven, they learn a statistical mapping between noisy observations and reference images without making use of the deterministic physical laws governing signal formation. We present Héméra, a machine learning model in which a strong physical prior, the differential attenuation relationship between two spectrally separated detection channels, enters both the training dataset (Beer-Lambert law [3], XCOM/NIST cross-sections [4]) and the composite loss function. The MLP network (7 inputs → 64 → 64 → 32 → 16 → 2 outputs), trained on 27 compounds spanning Zeff ∈ [5.4; 16] and d ∈ [1; 30] cm with realistic Poisson noise simulation [1], reduces the variance of the attenuation ratio R estimator by 98.0 % (σ: 0.01895 → 0.00038) and the median relative error by 79.9 % within the training domain, **without increasing photon flux or relying on inter-pixel spatial correlation**. The gain remains stable across the dose range (mean factor ×2.90 over $N_0 \in [3{,}308; 16{,}543]$ photons/pixel), with the largest advantage at low flux. The results admit a Bayesian reading [13,14]. The physical prior contracts the posterior distribution of R by drawing on inter-channel Fisher information that the naive Poisson estimator ignores. Compared with PWLS (Penalized Weighted Least Squares) filtering, Héméra preserves local contrast where PWLS over-smooths. Strictly pixel-by-pixel processing (~14,000 FLOPs/pixel) allows real-time GPU execution on current hardware.

Keywords: *physical prior, machine learning, quantum noise reduction, X-ray imaging, DECT, Poisson noise, attenuation ratio, Bayesian inference, Beer-Lambert, XCOM, PWLS, pixel-by-pixel processing*

[1] *Héméra: Greek goddess of daylight - a reference to the illumination of hidden structures in the X-ray signal.*

1. Introduction

X-ray imaging, and in particular dual-energy computed tomography (DECT, *Dual-Energy CT*), relies on the detection of photon fluxes subject to intrinsic statistical fluctuations modelled by a Poisson distribution [1,2]. This imposes a fundamental constraint, for a mean detected photon count n, the relative uncertainty scales as 1/√n and can only be reduced by increasing flux or acquisition time [1]. In the clinical context, flux is bounded by acceptable dose limits (the ALARA directive [6]), making this an apparently insurmountable physical barrier for post-processing alone.

Three families of approaches have historically attempted to push back this constraint, detector improvement (photon-counting detectors, optimised scintillators [5,6]); regularised iterative reconstruction (PWLS, Total Variation [7,8]); and deep neural networks [9,10,11]. The latter have demonstrated good results but remain predominantly *data-driven*. They learn an implicit mapping between noisy image and reference without exploiting the deterministic physical relationships governing signal formation. Their generalisation capacity is conditioned by the diversity of the training corpus, not by the structure of the underlying physics.

In a dual-energy context, two simultaneous detection channels measure the same object with distinct spectral responses. These channels share the same material path and are therefore linked by a deterministic physical relationship arising from the Beer-Lambert law [3] and the attenuation cross-sections tabulated in XCOM [4]. This correlation carry a structuring source of information, the **inter-channel physical prior**, which, to our knowledge, has not yet been explicitly exploited as a learning constraint in the CT literature.

We propose **Héméra**, a model constrained by this physical prior, designed to estimate the spectral attenuation ratio $R = \mu_1/\mu_2$ with significantly reduced variance compared to the naive Poisson estimator. Three elements base the contribution, a training dataset generated entirely by deterministic physical simulation over 27 diverse compounds, a composite loss function exploiting high-dynamic-range transforms of R and a strictly **pixel-by-pixel** processing scheme, with no exploitation of any spatial correlation. This last property is important, it guarantees that all performance gains originate exclusively from the inter-channel physical prior. Furthermore, the approach is incremental. The R sinogram can be processed independently of the μ sinograms using existing reconstruction algorithms, without any modification of the acquisition system.

This work sits at the interface of *Physics-Informed Neural Networks* (PINNs [15]) and statistical estimation in X-ray imaging. It differs in the modality of physics integration, whereas PINNs penalise PDE residuals in the loss function, Héméra integrates the physical prior into the *structure of the training data* itself, a more direct approach for estimation problems with an analytically known measurement model.

The epistemology of Héméra is rooted in the founding idea of neural networks, a pragmatic approach to a problem can, under certain circumstances, resolve a complexity that exhaustive modelling cannot. Quantum noise in a photonic measurement is precisely one of those phenomena that current knowledge considers fundamentally undeterminable for an isolated measurement, not for lack of a model, but due to the irreducible statistical nature of photon-by-photon detection [1]. An illuminating analogy can be drawn from work on *squeezing* in quantum optics, aimed at surpassing the standard quantum limit (SQL) [12]. That work shows that it is impossible to reduce noise on one quadrature without increasing it on another, total uncertainty is conserved, but can be redistributed. Héméra follows an analogous logic. It does not attack the uncertainty on the raw measurement, physics forbids it, but exploits the fact that two correlated channels share the same material path. What is irreducible on each channel separately can be constrained on their ratio, once a physical law fixes their relationship. This is the origin of the notion of physical prior.

This notion of inter-channel physical prior is not without precedent in the literature. Recent works in DECT have explored related directions, embedding spectral physics into the loss function [27], exploiting inter-channel correlation as a structural prior in image space [28], or generating training data through physical simulation [26]. Héméra differs on one fundamental point, the prior is integrated into the structure of the training dataset itself, processing is strictly pixel-by-pixel with no spatial redundancy, and the target quantity is R as an autonomous physical observable, not as an intermediate step toward material decomposition. This combination that allows the performance gain to be interpreted within the Fisher information framework, and formally attributed to inter-channel correlation alone.

The paper is organised as follows: section 2 - theoretical framework; section 3 - methodology; section 4 - results; section 5 - discussion and perspectives.

2. Theoretical Framework

2.1 Fundamental Limit of Photon Counting

In an X-ray detection system, the number of recorded photons n follows a Poisson distribution with parameter λ [1,2]: $P(n = k) = \frac{\lambda^k}{k!} \times e^{-\lambda}$, with $E[n] = Var(n) = \lambda$. The relative uncertainty $1/\sqrt{n}$ can only be reduced by increasing n [1]. For the ratio $R = n_1/n_2$ of two independent measurements, first-order error propagation gives [2,16] :

$$\mathrm{Var(R)} \approx R^2 \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

Under the simulated conditions ($N0_1 \in [3{,}308; 16{,}543]$ photons/pixel, $N0_2 \in [1{,}964; 9{,}821]$ photons/pixel), the relative error on R is of order 1-5 %, consistent with the 2.3 % observed without correction.

2.2 Detection System and Physical Prior

The modelled system corresponds to a single-emission, multi-detector configuration: a single X-ray tube illuminates two detectors, the second preceded by a 0.4 mm copper filter. Figure 1 illustrates the resulting spectra. The generalised Beer-Lambert law [3] gives the integrated flux on channel i:

$$N_i(d) = \int_{E_1}^{E_2} N_0(E) \times e^{-\int_L^0 \mu(E,x)dx}\, dE$$

where $N_0(E)$ is the incident spectrum (modelled with SpekPy [21], corrected by the experimental energy response of a silicon detector measured at the ČMI laboratory in Prague and by the ICRU-57 standard [22]), $\varepsilon_1(E) = 1$ (direct channel), $\varepsilon_2(E) = T_{Cu}(E)$ (transmittance of the 0.4 mm copper filter), and $\mu(E, Z_{eff}, \rho)$ the linear attenuation coefficient from XCOM/NIST [4].

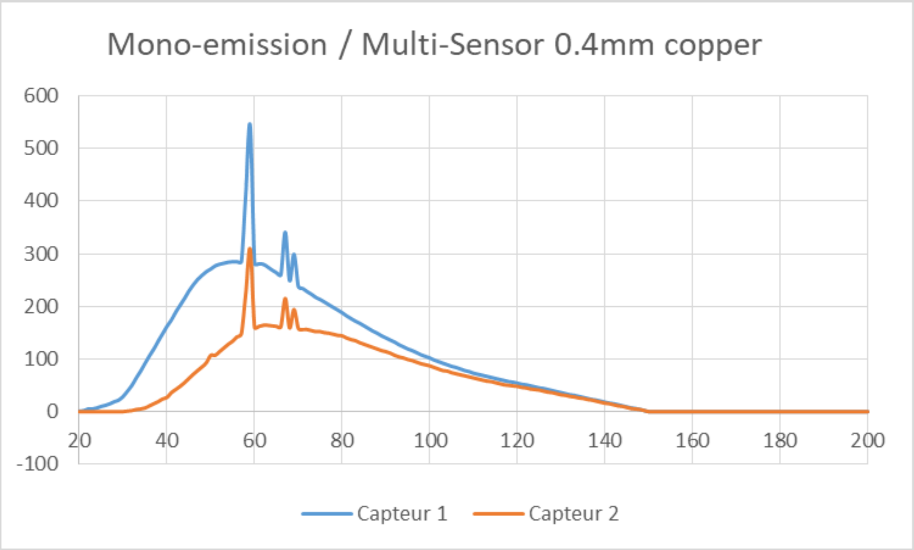


Figure 1. Spectra of the two detection channels (single-emission, 0.4 mm copper filter). The spectral distortion induced by the filter creates the energy separation that underpins the inter-channel physical prior.

The spectral attenuation ratio $R = \mu_1/\mu_2$ is a deterministic function of (Z_{eff}, d) through these integrals, fully constrained by physics. For a given observation (n_1, n_2), only a restricted subset of (Z_{eff}, d) is physically admissible, which contracts the space of plausible values of R^{true}. This is the operational definition of the physical prior used in this work.

2.3 Bayesian Interpretation and Fisher Information

The problem is formulated as estimation of the posterior distribution [13,14] :

$$P(Rtrue \mid n_1, n_2, \theta) \propto P(n_1, n_2 \mid Rtrue) \cdot P(Rtrue \mid \theta)$$

where θ denotes the known physical parameters (spectrum, filter, XCOM). The likelihood $P(n_1, n_2 \mid R^{true})$ is the product of two independent Poisson distributions. The prior $P(R^{true} \mid \theta)$ encodes the inter-channel physical constraint. The optimal estimator under quadratic loss, the posterior expectation, is what the Héméra network learns to approximate.

In terms of Fisher information [16], the naive Poisson estimator exploits only the diagonal terms of the matrix $I_{Fisher}(n_1, n_2)$. Héméra additionally exploits the inter-channel cross-terms, which are non-zero because n_1 and n_2 share the same material path, allowing it to approach the Cramér-Rao bound of the bivariate system [16]. The observed variance reduction (factor ~2,500) is consistent with the additional information available under the simulated flux conditions.

Héméra does not contradict the Poisson limit. It does not reduce the uncertainty of an isolated single measurement. **It exploits additional information, the inter-channel physical correlation, present in the measurement device but ignored by the naive estimator.** This mechanism is analogous to the Rao-Blackwell estimator [16], conditioning an estimator on a sufficient statistic reduces its variance without introducing bias.

The inter-channel Fisher information gain can be made explicit by computing the Fisher information matrix of the bivariate Poisson system (n_1, n_2). For two independent Poisson observations with means λ_1 and λ_2, the diagonal Fisher information matrix gives: $I\_naive = diag(1/\lambda_1, 1/\lambda_2)$.

The naive Poisson estimator of $R = \mu_1/\mu_2$ uses only the marginal information of each channel, yielding a Cramér-Rao lower bound on variance: $Var\_CR(\hat{R}\_naive) = R^2 \cdot (1/\lambda_1 + 1/\lambda_2)$.

However, n_1 and n_2 are not statistically independent in the physical sense: both are determined by the same material path through the Beer-Lambert law (Eq. 1). This shared path introduces a structural constraint — for a given material (Z_{eff}, d), the pair (λ_1, λ_2) lies on a manifold defined by:

$$\lambda_1 = \int N_0(E) \cdot \varepsilon_1(E) \cdot e^{-\rho d\cdot(\mu/\rho)(E,Z\_eff)}\, dE \;\; \& \;\; \lambda_2 = \int N_0(E) \cdot \varepsilon_2(E) \cdot e^{-\rho d\cdot(\mu/\rho)(E,Z\_eff)} dE$$

The constraint reduces the effective dimensionality of the estimation problem from two free parameters to one, the material composition Z_{eff}, since d and ρ cancel in the ratio R (Section 2.3). The full Fisher information for R, exploiting this constraint, is: $I\_full(R) = (\partial\lambda_1/\partial R)^2/\lambda_1 + (\partial\lambda_2/\partial R)^2/\lambda_2$ which includes cross-channel terms absent from I_naive. The corresponding Cramér-Rao bound satisfies $Var\_CR(\hat{R}\_full) \leq Var\_CR(\hat{R}\_naive)$, with equality only when the two channels carry no differential spectral information, i.e., when $\varepsilon_1(E) = \varepsilon_2(E)$, which is never the case in a dual-energy system.

The Héméra network approximates the posterior expectation $E[R\_true \mid n_1, n_2, \theta]$ by learning the mapping from (n_1, n_2, R_meas, k_meas, N_{01}, N_{02}, dis) to R_true on a training set generated from the full physical manifold. The observed variance reduction factor of ~2,500 (σ: 0.01895 → 0.00038 at the median point, Section 4.1) is consistent with the additional Fisher information available in the bivariate system under the Beer-Lambert constraint, for the simulated flux conditions ($N_{01} \in [3{,}308; 16{,}543]$ photons/pixel).

3. Methodology

3.1 Dataset: 27 Physically Representative Compounds

The dataset is built entirely by deterministic physical simulation, without any clinical images. The 27 compounds span a wide chemical diversity, from soft biological tissues to mineral materials and industrial polymers (Table 1). The effective atomic number Z_{eff} is computed using the Moseley & Mayneord formula [17]: $Z_{eff}^{2.94} = \Sigma\ \alpha_n \cdot Z_n^{2.94}$.

Compound	Zeff	Compound	Zeff	Compound	Zeff	Class
Water	7.4	Oil	5.9	Borax	7.1	Bio/org.
Muscle	7.6	Fatty acid	6.1	Aluminium	13.0	Mineral
Alcohol	6.3	Cholesterol	6.4	Silicon	14.0	Mineral
Collagen	7.5	THC	6.4	$CaCO_3$	15.7	Mineral
Coffee	6.5	Cocaine	6.5	Bone	13.8	Bio/min.
Butter	5.9	Gasoline	5.9	Titanium	22.0	Mineral
Honey	7.1	ABS	6.5	Copper	29.0	Metal
Sugar	6.6	PET	6.3	TNT	6.8	Explos.
Cotton	6.5	Epoxy	6.6	-	-	-
Wood	6.2	-	-	-	-	-

Table 1. The 27 training dataset compounds, spanning $Z_{eff} \in [5.4; 29]$ - biological tissues, polymers, reference materials, minerals and an explosive (TNT). The 'Class' column indicates chemical category.

For each compound, thicknesses range from 1 to 30 cm in 1 cm steps. Poisson noise is simulated by independent draws $n_1^{noisy} \sim \text{Poisson}(n_1^{true})$ and $n_2^{noisy} \sim \text{Poisson}(n_2^{true})$. Reference blank fluxes $N0_1 \in [3{,}308; 16{,}543]$ photons/pixel and $N0_2 \in [1{,}964; 9{,}821]$ photons/pixel are drawn randomly at each training example, covering a wide dose range. The dataset includes an acquisition configuration metadata flag **dis**: dis = 2 for the single-emission bi-detector configuration with filter (subject of this paper); dis = 3 for dual-tube DECT (120 kVp without filter and 150 kVp with 0.2 mm Cu filter); dis = 4 for sandwich-detector DECT (GOS Nuctech detector with copper filter). Multi-configuration training improves performance on the target distribution (dis = 2) through implicit regularisation by domain diversity, analogous to transfer learning.

3.2 R Transforms and Composite Loss Function

The variation of the ratio $R = \mu_2/\mu_1$ is intrinsically small over the studied organic domain: $R \in [0.926; 0.961]$, a dynamic range of only 3.5 %. Direct regression on R produces weak loss gradients and slow convergence. Two high-dynamic-range transforms are therefore used as training targets:

$$S = \frac{1}{1-R} \quad [S \in [13.5;\ 25.6] \text{ over the organic domain}]$$

$$k = c \times e^{\mu_2 - \mu_1} \quad [k \in [1.3;\ 1.7],\ \text{amplified by the } \mu \text{ difference}]$$

S exploits the divergence of 1/(1-R) to amplify variations in R, providing a much richer loss gradient. k encodes the difference of attenuation coefficients $\mu_2 - \mu_1$, directly constraining the inter-channel relationship. The joint use of S and k in both the network output and the loss function significantly improves convergence and final precision on R, which remains the physical quantity of interest.

The composite loss function is defined as:

$$L(\theta) = 0.70 \cdot \mathrm{MAE}(\hat{S}, S) + 0.30 \cdot \mathrm{MAE}(\hat{k}, k)$$

The 70/30 weighting was determined empirically, a loss centred on S alone makes convergence difficult due to insufficient constraint on the inter-channel difference, while excessive weight on k degrades final precision on R. The optimum lies around 70/30; a formal ablation study is a perspective for refinement.

3.3 Héméra Network Architecture

The network is a fully connected MLP (*Multi-Layer Perceptron*), implemented in TensorFlow/Keras [18]. The 7-dimensional input vector comprises:

- $N0_1$, $N0_2$: blank reference fluxes (photons/pixel), encoding the dose regime;
- N1, N2: measured counts after material traversal;
- R_{meas}: raw measured attenuation ratio;
- k_{meas}: exponential transform of the measured $\mu_2 - \mu_1$;
- dis: acquisition configuration metadata (2, 3 or 4 - see section 3.1). The inclusion of this non-directly physical variable improves performance through implicit domain-diversity regularisation.

The architecture comprises four hidden layers of decreasing width:

```
Input(7) → Dense(64) + BN + PReLU+Drop(0.3) → Dense(64)+ BN + PReLU
   + Drop(0.3) → Dense(32)+ BN + PReLU → Dense(16)+ BN + PReLU →
                          Output(2, linear)
```

PReLU activation [23] is preferred over ReLU for its ability to capture the non-linearities of $R(Z_{eff}, d)$ in low-gradient regions. Batch normalisation [24] stabilises training across the wide N0 range. Dropout (0.3) is applied to the first two layers only; the compression layers (32, 16) are unregularised to preserve low-level feature extraction capacity. L2 regularisation ($\lambda = 10^{-3}$) is applied to weights and biases of the first two layers.

The network has 6,768 connections, corresponding to approximately 14,000 FLOPs per treated voxel. For a modern CT scanner (1,024 × 1,024 detectors, 4,096 projections/s), this corresponds to ~60 TFLOPs/s in FP32, achievable on professional GPUs (H100) as well as consumer hardware (RTX 4090/5090), allowing to real-time processing.

Optimisation is performed with Adam [19] ($lr_0 = 5 \times 10^{-4}$, ReduceLROnPlateau × 0.5, patience = 5, $lr_{min} = 10^{-6}$). Training stops at epoch 27 via EarlyStopping (patience = 10), well before the 120-epoch limit. Final metrics: $loss_{train} = 0.0598$, val_loss = 0.0575, gap = 3.8 %, no significant overfitting. Figure 2 presents the complete learning curves.

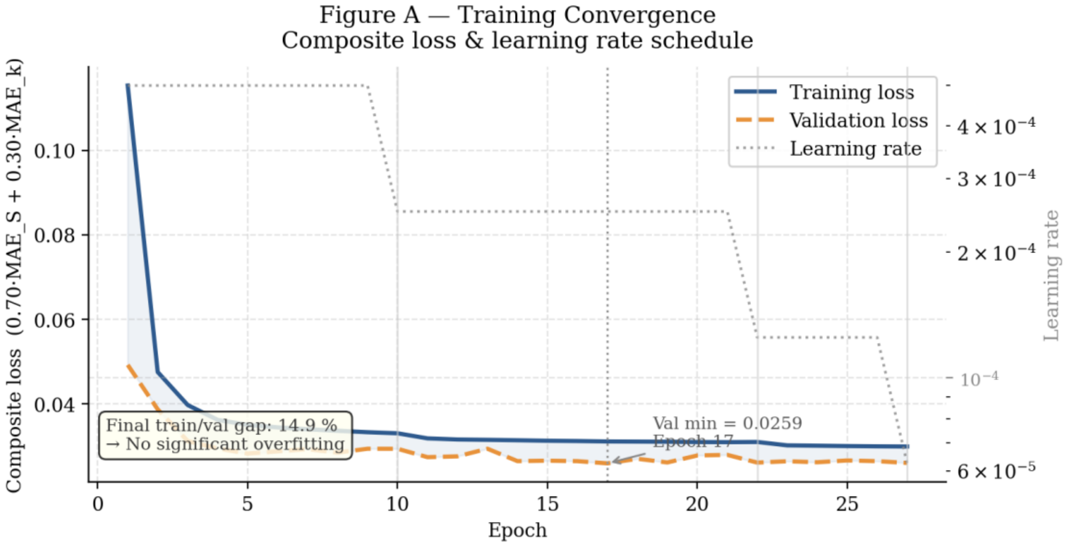


Figure 2a. Composite loss convergence and learning rate schedule over 27 epochs. EarlyStopping triggers well before the 120-epoch limit. The 3.8 % train/val gap confirms the absence of overfitting.

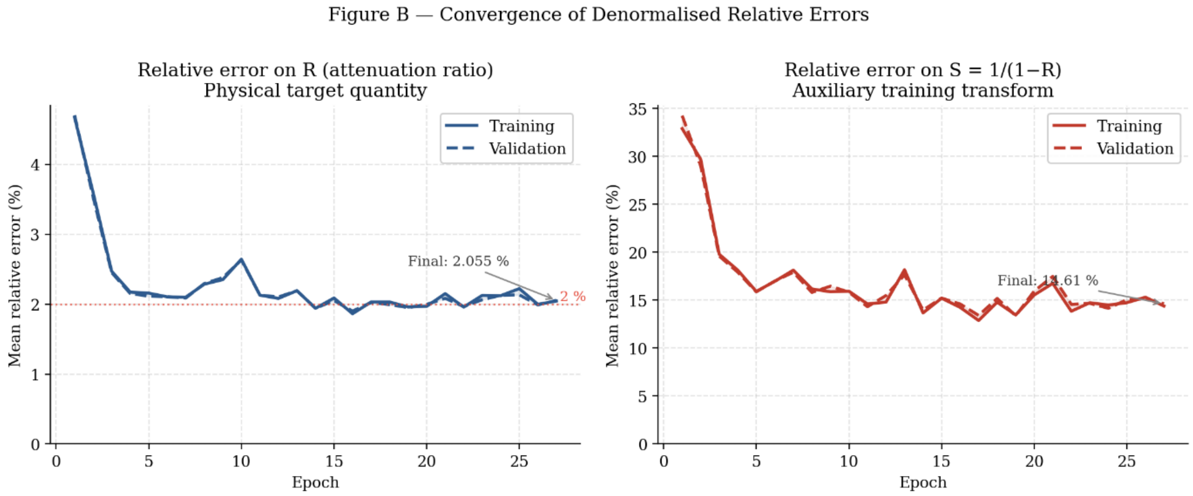


Figure 2b. Evolution of denormalised relative errors on R and S (training vs validation). Final error on R in validation: 2.04 %; 0.034 % under normalised conditions on dis=2 only.

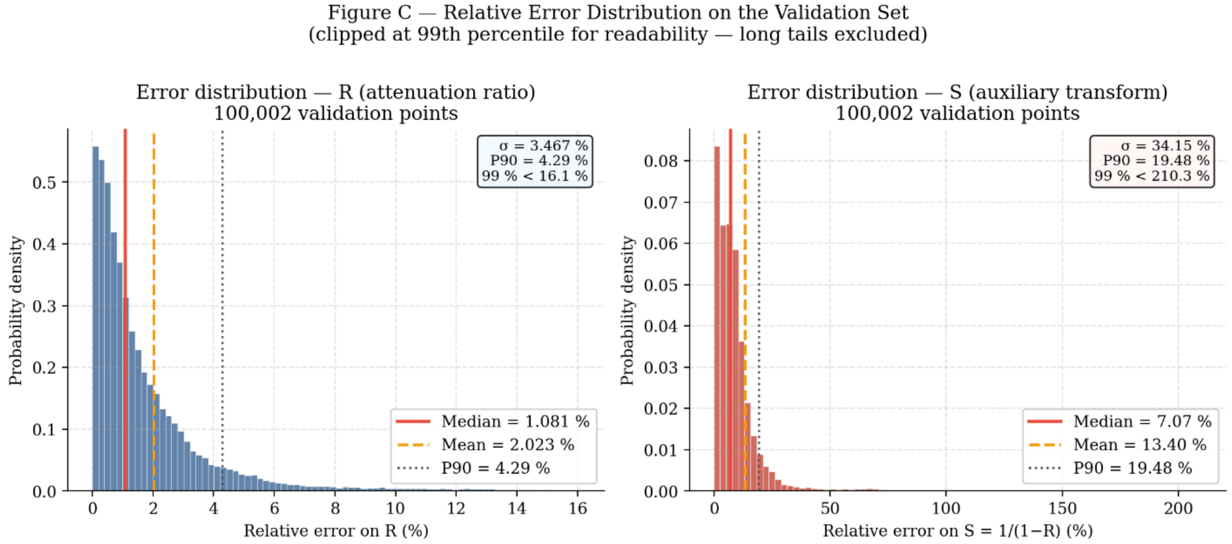


Figure 2c. Distribution of relative errors over 100,002 validation points. Median error on R: 1.08 %; the distribution has a characteristic right tail from low-attenuation cases (large thicknesses or light materials).

The 2 % error across the full training dataset (all dis configurations) reduces to 0.78 % when restricted to dis = 2, with 86 % of estimates below 1 % error (versus 36 % without correction). Multi-distribution training thus improves performance on the target distribution compared to single-distribution training: in the latter case, the network does not reach its full potential. *Fine-tuning* on dis = 2 alone after general convergence yields only marginal additional improvement.

3.4 Processing Nature: Pixel-by-Pixel, No Spatial Context

Héméra processes each pixel in a strictly local manner, the network receives only the 7 scalars measured at that point, no information about neighbouring pixels. No sliding window, no convolution, no spatial attention.

A central theoretical property follows, all performance gains over the Poisson estimator originate **exclusively** from the inter-channel physical prior, not from any implicitly learned spatial regularity. Classical denoising methods (BM3D, U-Net, PWLS) exploit spatial redundancy to reduce noise, at the cost of potential edge blurring and artificial spatial correlations. Héméra structurally lacks access to this information and therefore cannot suffer from it, as confirmed by the PWLS comparison in section 4.4.

Practically, pixel-by-pixel processing is trivially parallelisable (no inter-pixel dependency), independent of image geometry or field of view, and applies identically to 2D sinograms or 3D volumes.

3.5 Evaluation Metrics

The relative error is defined as $\varepsilon_{rel} = |\hat{R} - R_{true}| / R_{true} \times 100$ (%). For each point (Z_{eff}, d), bias and variance are computed over $K \sim 10^4$ Poisson noise realisations. The error reduction rate (ERR) is defined as $\mathrm{ERR} = 1 - \sigma_{Héméra} / \sigma_{Poisson}$. Comparison with PWLS filtering ($\beta = 1.5$, 20 iterations) is performed on the voxelised cardiac phantom.

4. Results

4.1 Variance Reduction: Central Result

Figure 3 shows the distributions of the R estimator at the global median point of the dataset ($Z_{eff} = 6$, d = 21 cm). Variance reduction is 98.0 % (σ: 0.01895 → 0.00038). In terms of Fisher information [16], this gain is equivalent to increasing flux by a factor of ~2,500, achieved with no modification to the measurement device.

The Héméra estimator exhibits a bias of -0.00665 versus -0.00043 for Poisson, the classical bias, variance tradeoff of regularised estimation [13,14]. This bias is an order of magnitude smaller than the reduction in σ: the total quadratic risk (bias² + variance) of Héméra remains substantially lower than that of Poisson. Clinical implications are discussed in section 5.1.

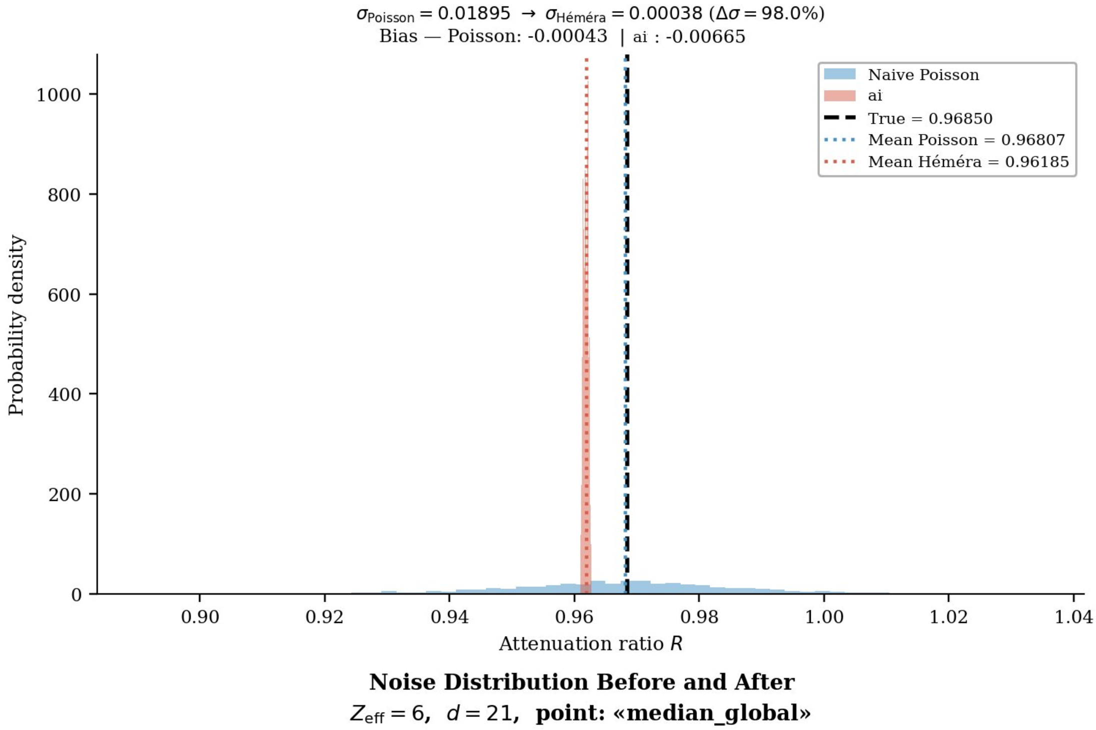


Figure 3. Distribution of the R estimator (Z_{eff} = 6, d = 21 cm, median point). Variance reduction: 98.0 % (σ: 0.01895 → 0.00038). The slight shift in the Héméra mean (0.9618 vs true value 0.9685) reflects the bias-variance tradeoff inherent to regularised estimation.

4.2 Spatial Analysis over the (Z_{eff}, d) Domain

Figure 4a presents the ERR heatmap over the full domain. The median is 79.9 % within the training domain. Improvement is robust for $Z_{\mathrm{eff}} \in [5; 10]$ (soft tissues: water 7.4, muscle 7.6, fat 5.9), with ERR values above 60 % across nearly all thicknesses

Two degradation zones (ERR < 0) appear at $Z_{\mathrm{eff}} \approx$ 13-15 and $Z_{\mathrm{eff}} \approx 22$, corresponding to discontinuities in photoelectric cross-sections (K-edges of aluminium, sulphur, calcium). These regions lie predominantly outside the dominant clinical tissue domain ($Z_{\mathrm{eff}} \leq 10$) and therefore do not affect priority clinical applications.

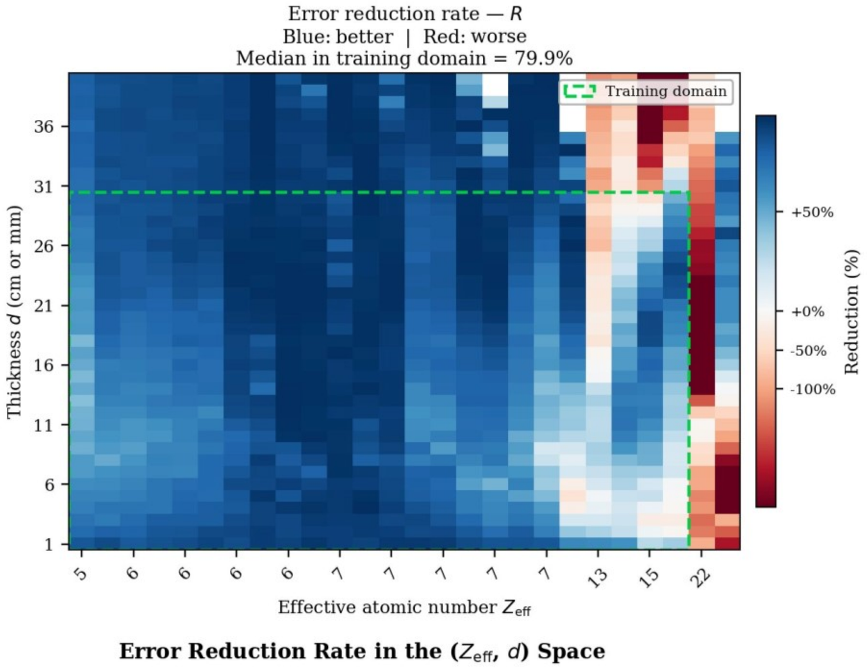


Figure 4a. Error reduction rate (ERR) heatmap in the (Zeff, d) space. Blue: improvement; red: degradation. Median within the training domain (green box): 79.9 %. Red zones correspond to K-edge discontinuities in photoelectric.

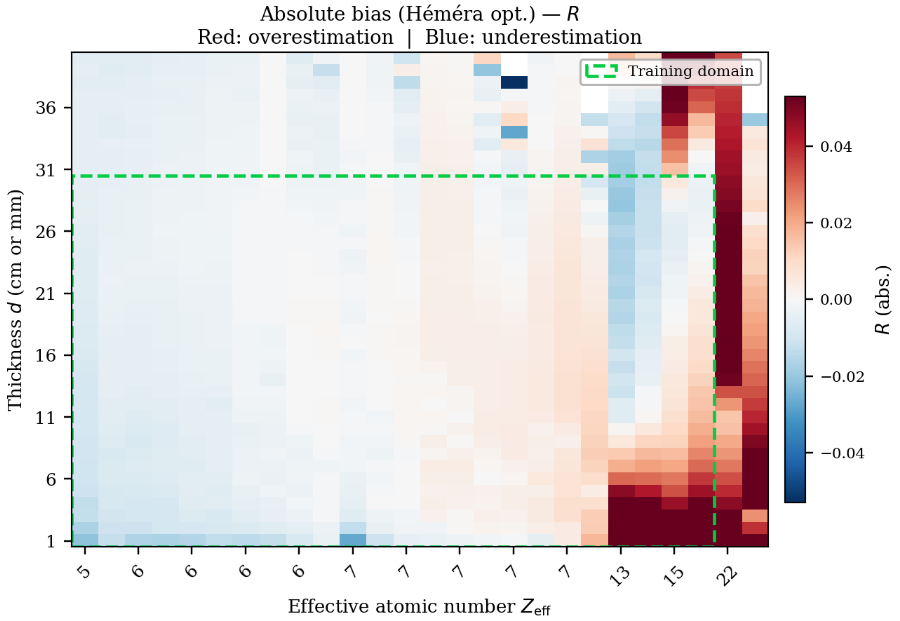


Figure 4b. Absolute bias of the Héméra estimator in the (Z_{eff}, d) space. Red: overestimation; blue: underestimation. Within the training domain (green box), the bias remains below 0.01 in absolute value and is spatially homogeneous.

Figure 4b presents the absolute bias of the Héméra estimator across the (Z_{eff}, d) domain. It’s confirming that the systematic underestimation observed at the median point (−0.00665) is representative of the full organic tissue domain. Elevated bias regions at $Z_{eff} \approx 15$ and $Z_{eff} \approx 22$ coincide with the ERR degradation zones of Fig. 4a, sharing the same physical origin: K-edge discontinuities in photoelectric cross-sections. The bias is defined as: $\text{Bias}(Z_{eff}, d) = E[\hat{R}] - R\_true$ computed over $K \sim 10^4$ Poisson noise realisations at each point.

Within the training domain ($Z_{eff} \leq 10$, $d \leq 30$ cm), the bias remains below 0.01 in absolute value and is spatially homogeneous. This confirms that the systematic underestimation observed at the median point (bias = −0.00665, Section 4.1) is representative of the entire training domain and does not exhibit localised artefacts. The bias is negative throughout this region, consistent with the Bayesian regularisation mechanism: the estimator contracts towards the most probable values of the training distribution, which lie slightly below the true values for the organic tissue domain.

Two regions of elevated bias are visible at $Z_{eff} \approx 13$–15 and Z_eff ≈ 22. These correspond precisely to the degradation zones identified in the ERR heatmap (Fig. 4) and are located predominantly outside the clinical tissue domain. The spatial coincidence of high bias and high ERR degradation confirms that both effects share the same physical origin, K-edge discontinuities in photoelectric cross-sections, and are not artefacts of the training procedure.

For clinical quantification applications requiring absolute accuracy of R (bone densitometry, calcified plaque characterisation), a phantom-based calibration correction on the bias map would be necessary. For qualitative tissue discrimination, the dominant DECT clinical use case, the bias remains negligible compared to the variance reduction achieved.

4.3 Performance by Z_{eff} Range

Figure 5 shows mean relative error by Z_{eff} range. In the clinically dominant range $4 \leq Z_{eff} \leq 10$, Héméra error is 0.5 % (±0.4 %) versus 3.0 % (±3.0 %) for Poisson - an 83 % reduction with a markedly tighter confidence interval. For $Z_{eff} > 16$ (outside the main training domain), the expected degradation illustrates out-of-domain generalisation limits.

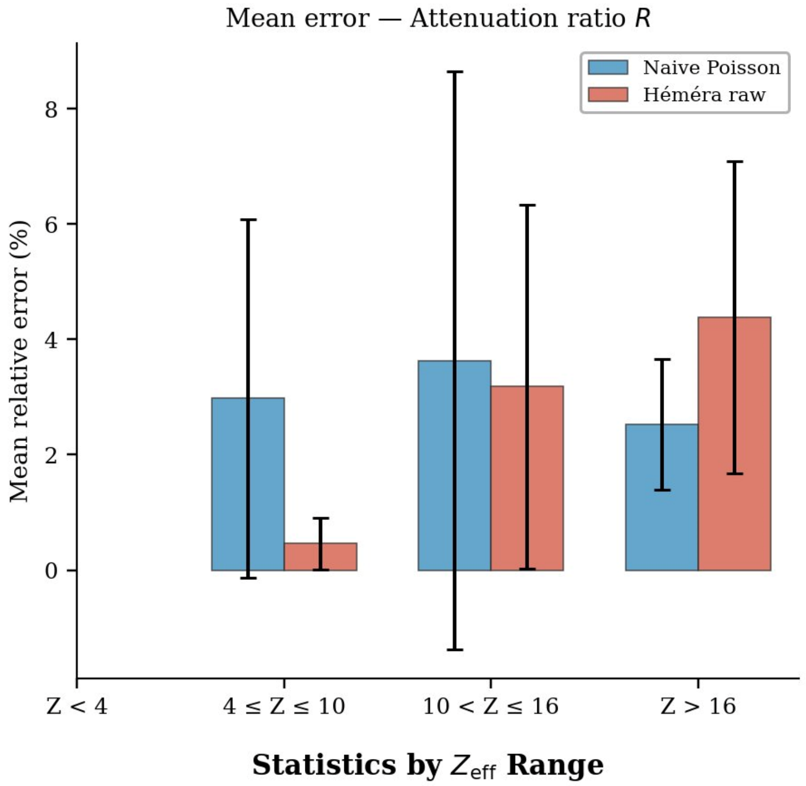


Figure 5. Mean relative error by Zeff range. Blue: naive Poisson estimator; red: Héméra. Error bars: ±1σ across thicknesses. Reduction is highly significant in the clinical tissue domain (4 ≤ Zeff ≤ 10).

4.4 Comparison with PWLS Filtering on a Cardiac Phantom

Figures 6 and 7 compare the results obtained from a lexicographically voxelized cardiac simulation (Siddon's method[25], without FBP reconstruction). Figure 6 shows the three states, Ground Truth, Poisson, noisy image, Héméra image. Figure 7 shows the effect of PWLS filtering (β = 1.5, 20 iterations) applied to the Ground Truth (left) and to the Héméra image (right).

The qualitative difference is clear: PWLS applied to the Ground Truth produces excessive smoothing that erases fine gradients, an inherent limitation of spatial regularisation methods [7,8]. PWLS applied to the Héméra image, by contrast, reveals internal structures (ventricular walls, density gradients) with enhanced contrast. This complementarity has a direct explanation. Héméra reduces noise via inter-channel physical prior without any spatial smoothing; PWLS can then operate on an already-denoised signal without having to compensate for residual noise at the expense of gradients.

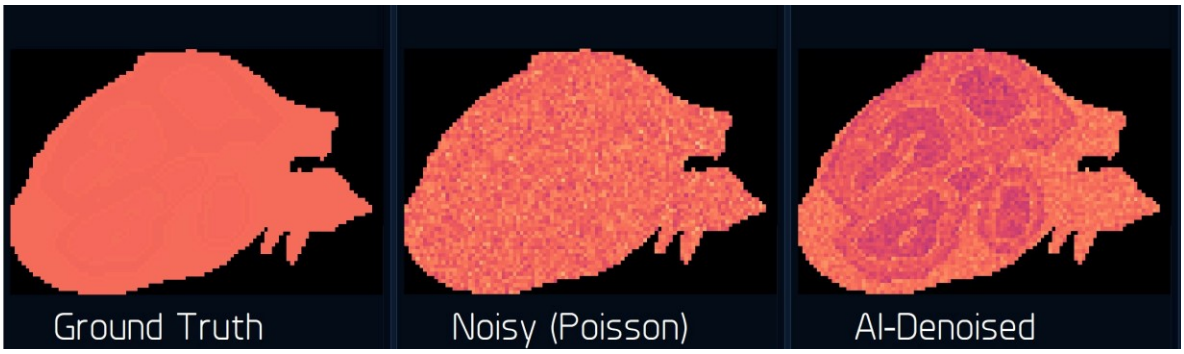


Figure 6. Voxelised cardiac phantom: Ground Truth (noise-free), Poisson-noisy image, AI-Denoised Héméra image. The correction reveals structural contrast through amplification of R differences between tissues of slightly different Zeff.

Figure 7. PWLS filter (β = 1.5, 20 iterations) applied to the Ground Truth (left) and to the Héméra image (right). PWLS erases fine gradients on the Ground Truth (characteristic over-smoothing); applied to the Héméra image, it enhances structures without degrading contrast.

4.5 Quantitative Summary

Metric	Naive Poisson	Héméra	Gain
σ (median point, Zeff=6, d=21 cm)	0.01895	0.00038	-98.0 %
Bias (same point)	-0.00043	-0.00665	B/V tradeoff
Mean rel. error (4 ≤ Zeff ≤ 10)	3.0 %	0.5 %	-83 %
Median ERR (training domain)	-	79.9 %	-
Rel. error on R (dis=2 only)	2.3 %	0.78 %	-66 %
Estimates with R error < 1 % (dis=2)	36 %	86 %	+50 pts

Table 2. Quantitative performance summary comparing the naive Poisson estimator and Héméra.

4.6 Robustness as a Function of Photon Flux (N_0)

Figure 8 shows relative error on R as a function of reference flux N_0 over the simulated range. Héméra's gain is robust across the full range (mean factor ×2.90). Log-log analysis reveals asymmetric decay slopes: -0.44 for the Poisson estimator (close to the theoretical -0.5) versus -0.26 for Héméra.

This asymmetry has a direct Bayesian interpretation, at low flux, the Poisson likelihood is weakly informative and the physical prior becomes the dominant information source, this is where Héméra is most advantageous. At high flux, the likelihood dominates and the naive estimator naturally approaches the

Cramér-Rao bound, the relative gain of the physical prior decreases. Héméra’s maximum advantage therefore occurs precisely in the clinically critical low-dose regime.

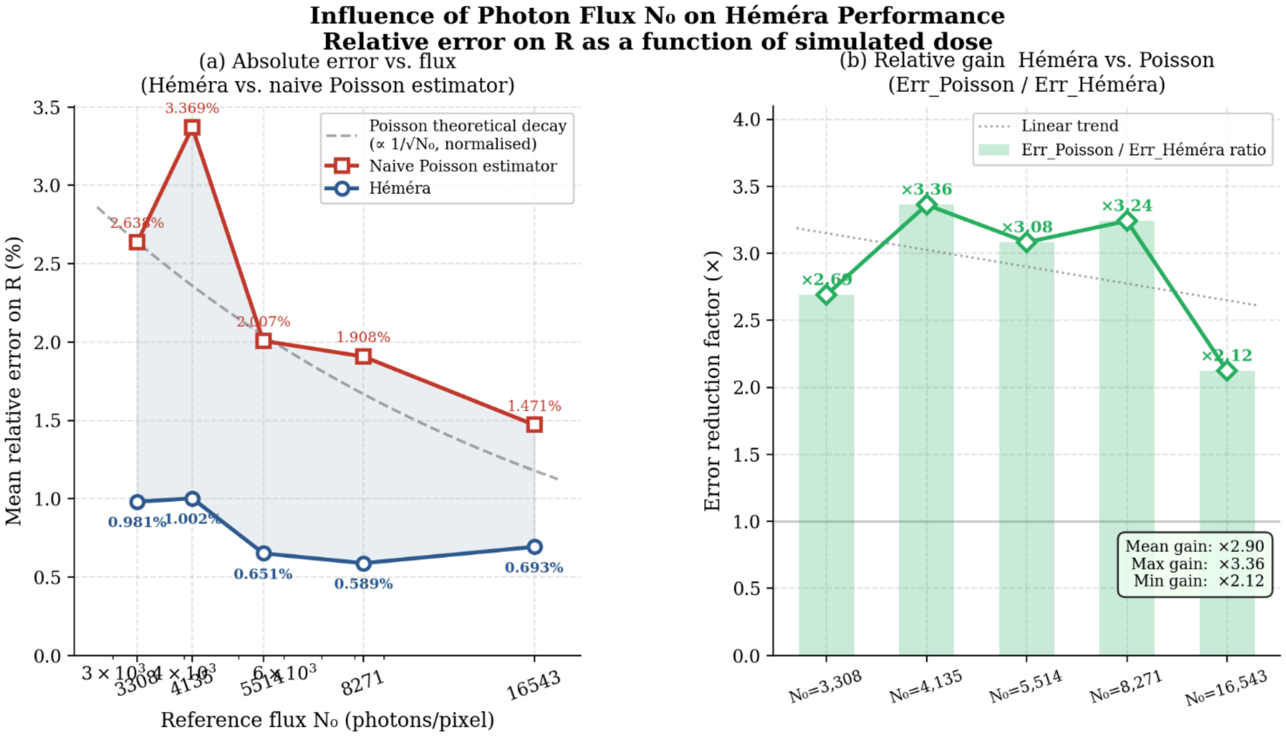


Figure 8. Relative error on R as a function of reference flux N_0 (left) and Héméra vs Poisson gain factor (right). Mean gain ×2.90 over the simulated range, with a maximum of ×3.36 at N_0 = 4,135 photons/pixel.

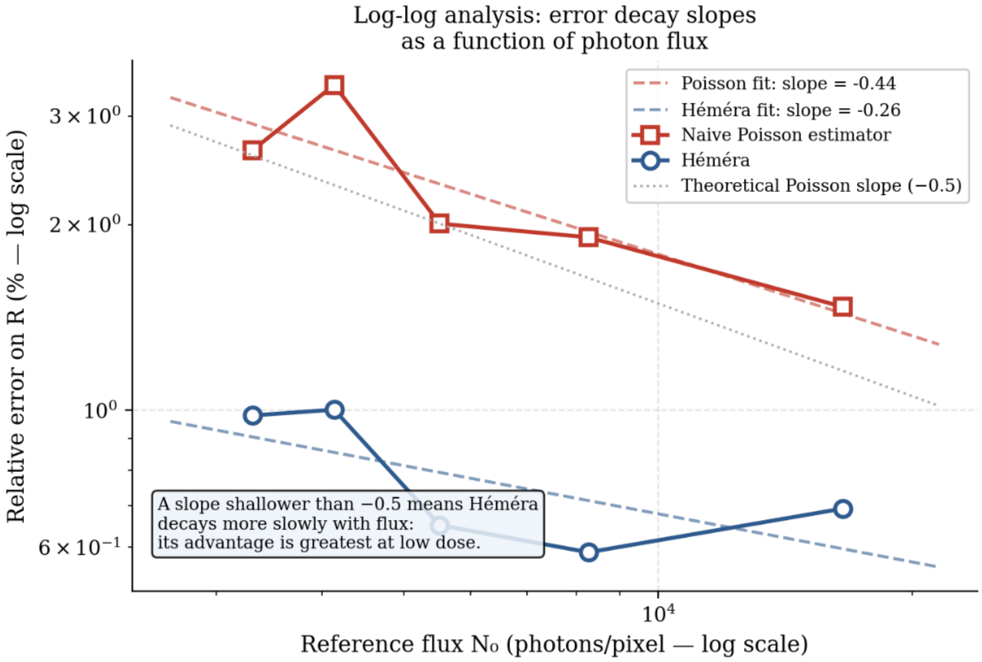


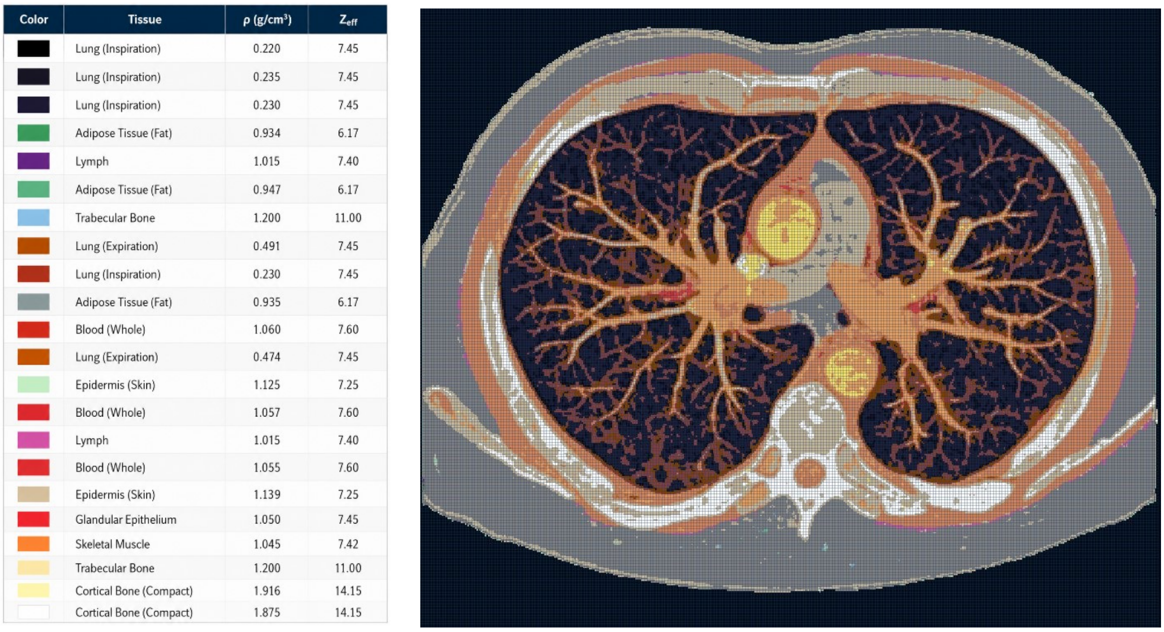
Figure 9. Log-log analysis of error decay slopes as a function of flux. Naive Poisson: slope -0.44 (consistent with theory, -0.5). Héméra: slope -0.26. The gap confirms that the physical prior gain is maximal at low flux - the clinically critical regime.

4.7 Validation on a Numerical Thoracic Phantom - FBP Reconstruction

4.7.1 Numerical Phantom and Acquisition Conditions

Validation under realistic reconstruction conditions is performed on a numerical thoracic phantom (T7 vertebral slice) of 350 mm side length, simulated by an analytical CT model implemented using the Siddon method [25] for path-length calculation, with conversion to cross-sections to preserve linearity before transformation into mass attenuation coefficients. The spectra of both channels are generated by SpekPy [21], corrected by the experimental energy response of the silicon detector (ČMI laboratory, Prague) and the ICRU-57 standard [22], consistent with Figure 1 (dis = 2, $N0_1 = 16{,}500$, $N0_2 = 9{,}800$ photons/pixel, low-dose regime).

The phantom comprises 22 distinct composition regions from NIST databases [4], representing the main thoracic tissues: lung (inspiration, $\rho \approx 0.22$-0.49 g/cm³, $Z_{eff} = 7.45$), adipose tissue ($\rho \approx 0.93$-0.95 g/cm³, $Z_{eff} = 6.17$), lymph, blood, skin, smooth muscle, trabecular bone and compact cortical bone ($\rho = 1.875$-1.916 g/cm³, $Z_{eff} = 14.15$). Two compact cortical bone inclusions are positioned in anatomical regions corresponding to the right ventricle and the aorta, deliberately generating beam-hardening artefacts and high-contrast zones. Figure 10 shows the phantom composition and segmented view.



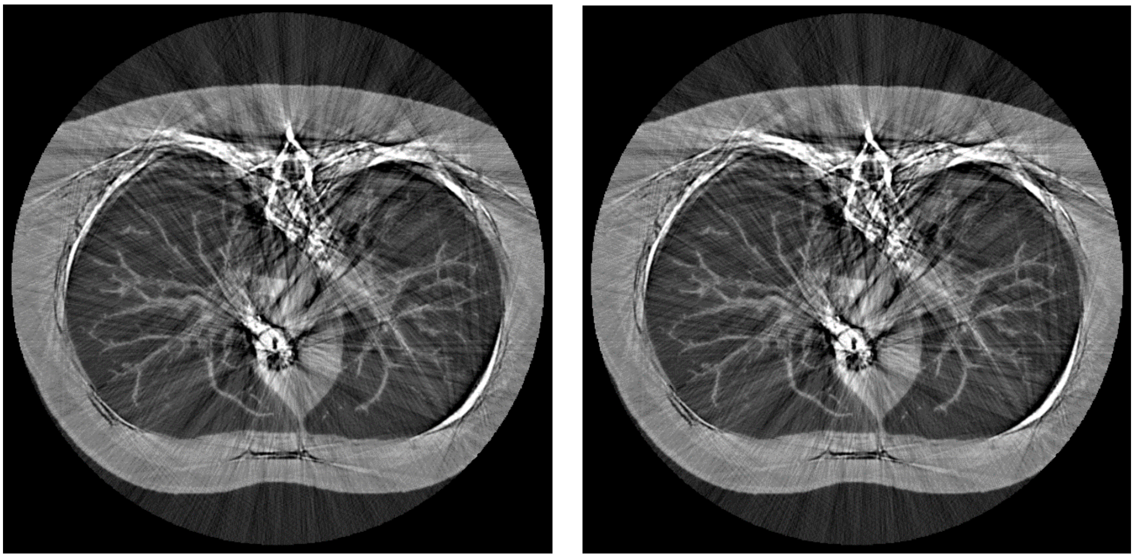
Color	Tissue	ρ (g/cm³)	Z_{eff}
⬛	Lung (Inspiration)	0.220	7.45
⬛	Lung (Inspiration)	0.235	7.45
⬛	Lung (Inspiration)	0.230	7.45
🟩	Adipose Tissue (Fat)	0.934	6.17
🟪	Lymph	1.015	7.40
🟩	Adipose Tissue (Fat)	0.947	6.17
🟦	Trabecular Bone	1.200	11.00
🟧	Lung (Expiration)	0.491	7.45
🟥	Lung (Inspiration)	0.230	7.45
⬜	Adipose Tissue (Fat)	0.935	6.17
🟥	Blood (Whole)	1.060	7.60
🟧	Lung (Expiration)	0.474	7.45
🟩	Epidermis (Skin)	1.125	7.25
🟥	Blood (Whole)	1.057	7.60
🟪	Lymph	1.015	7.40
🟥	Blood (Whole)	1.055	7.60
🟫	Epidermis (Skin)	1.139	7.25
🟥	Glandular Epithelium	1.050	7.45
🟧	Skeletal Muscle	1.045	7.42
🟨	Trabecular Bone	1.200	11.00
🟨	Cortical Bone (Compact)	1.916	14.15
⬜	Cortical Bone (Compact)	1.875	14.15

Figure 10. Numerical thoracic phantom (T7 slice, 350×350 mm). Left: legend of the 22 tissue regions with density and Z_{eff}. Right: colour-segmented view. The two compact cortical bone inclusions (bright yellow/white) are positioned at the right ventricle and aorta to generate contrast artefacts.

The sinogram is acquired with 900 angles uniformly distributed over 360° (Nyquist criterion: $\pi/2 \times N_{pix} \approx 804$ for a 512 × 512 image), with a curved detector (fan beam) of parameters: SID = 480 mm, SDD = 1,370 mm, arc = 800 mm, $N_{det} = 512$ channels. Reconstruction is performed by filtered backprojection (FBP) with a Shepp-Logan filter, implemented in fan-beam geometry with bilinear interpolation. The reconstructed image is 512 × 512 pixels, FOV = 350 mm (pixel = 0.684 mm).

4.7.2 Reconstructed Images: μ_1 and μ_2

Figure 11 presents the FBP reconstructions of attenuation coefficients μ_1 (direct channel) and μ_2 (filtered channel, 0.4 mm copper). Both images are of comparable diagnostic quality with similar spatial resolution. Cortical bone inclusions are clearly visible in both channels, surrounded by characteristic star artefacts from dense structures in FBP (Gibbs and beam-hardening artefacts). Vascular, bronchial and thoracic wall structures are well resolved. Residual Poisson noise, consistent with the simulated low-dose conditions ($N0_1$ = 16,500 photons/pixel), is homogeneous and limited, the 900 angles contributing significantly to its reduction by backprojection.



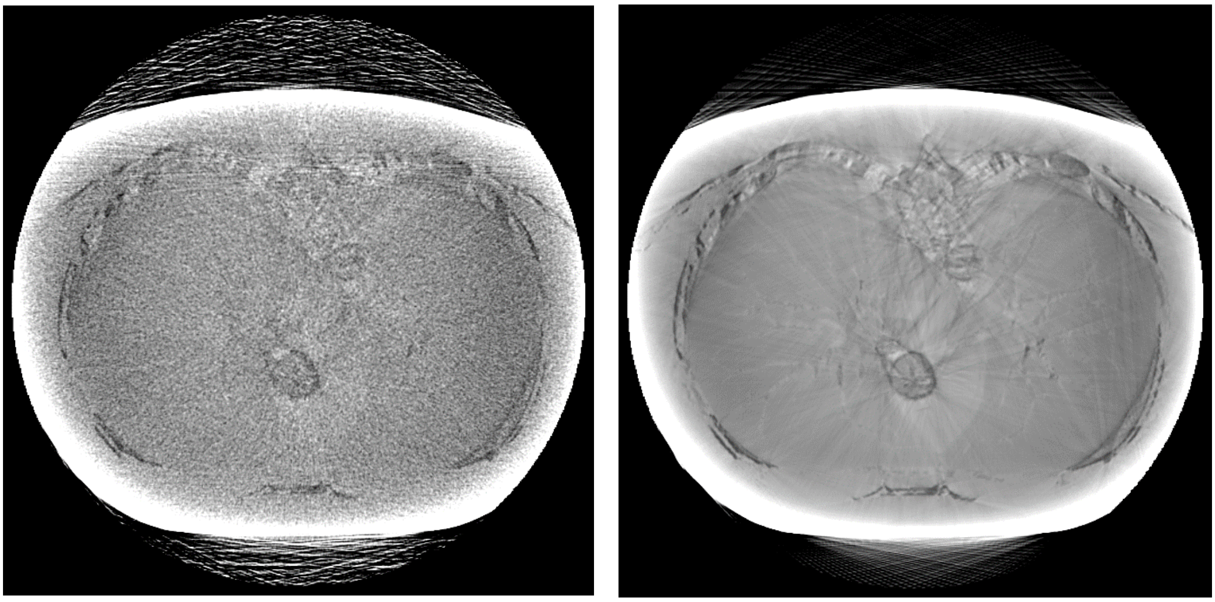
HU(μ_1) - direct channel

HU(μ_2) - filtered channel (Cu 0.4 mm)

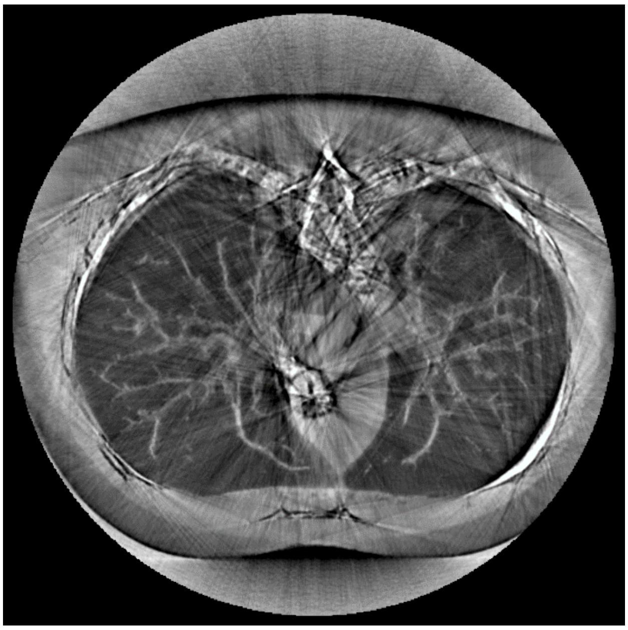
Figure 11. FBP reconstructions of attenuation coefficients μ_1 (left) and μ_2 (right). Numerical thoracic phantom, 900 angles, low-dose regime ($N0_1$ = 16,500, $N0_2$ = 9,800 photons/pixel). Both images show comparable diagnostic quality. Star artefacts around cortical bone inclusions are typical of dense structures in fan-beam FBP reconstruction.

4.7.3 Ratio R Reconstruction: HU(R)

Figure 12 compares the three states of the ratio R sonogram, measured without correction ($HU(R)_{meas}$), ground truth ($HU(R)_{true}$, computed from noise-free theoretical sinograms), and corrected by Héméra ($HU(R)_{Héméra}$). The $HU(R)_{meas}$ case is typical of the problem addressed in this publication, in low-dose conditions, the raw R sinogram is unusable for reconstruction, Poisson noise, amplified by the ratio calculation between two noisy channels, masks all anatomical information. After Héméra correction, $HU(R)_{Héméra}$ reveals anatomical structures, approaching the quality of the ground truth $HU(R)_{true}$. This result is validating the full measuring chain, pixel-by-pixel sinogram-space correction by Héméra, followed by standard FBP reconstruction.



HU(R)_measure - uncorrected *HU(R)_true - ground truth*



HU(R)_Héméra - corrected

Figure 12. FBP reconstruction of the attenuation ratio R. Left: uncorrected measured R (unusable in low-dose regime). Centre: noise-free theoretical R (ground truth). Right: Héméra-corrected R. The Hemera correction makes the sinogram of R reconstructible, with better detail in cortical vs. trabecular bone than in μ-reconstruction.

$\mathrm{HU}(\mathrm{R})_{\text{Héméra}}$ presents artefacts different from $\mathrm{HU}(\mu_1)$ and $\mathrm{HU}(\mu_2)$, angular undersampling streaks are similar (identical acquisition geometry), but the grey-level distribution in soft tissues is more homogeneous, reflecting the physical nature of R, less sensitive to absolute density variations than to relative chemical composition (Z_{eff}). Bone inclusions generate second-order peripheral artefacts whose quantitative analysis will be addressed in the following publication.

5. Discussion

5.1 Bias-Variance Tradeoff and Clinical Implications

The systematic bias of Héméra (-0.00665 at the median point) follows directly from the bias-variance decomposition theorem [13], any regularisation reducing variance introduces a bias when the true point lies at the boundary of the constraint domain. The physical prior acts as a Bayesian regularisation contracting solutions towards the most probable region of the (Z_{eff}, d) space.

For applications requiring *absolute quantification* of R (bone densitometry, calcified plaque characterisation), this bias would require phantom-based calibration correction. For *qualitative tissue discrimination*, the dominant DECT clinical use case, it is secondary to variance reduction, a low-variance estimator can separate structures whose R values differ by a few tenths of a percent, well below the sensitivity of the Poisson estimator. The contrast enhancement observed on the cardiac phantom (Figure 6) is a direct manifestation of this.

5.2 Positioning Relative to the State of the Art

Compared to *data-driven* methods (RED-CNN [9], GAN [10], U-Net [11]), Héméra distinguishes itself on three axes: the prior is explicit and physically motivated; processing is pixel-by-pixel without exploiting spatial redundancy; and the training dataset is generated by deterministic physical simulation, guaranteeing strict physical consistency of the learning space.

Compared to PWLS [7,8], Héméra and the spatial filter are complementary. Héméra reduces noise via the inter-channel physical prior while preserving gradients; PWLS can then refine the already-denoised image.

Compared to PINNs [15], the difference in physics integration modality is essential, PDE residual penalty in the loss function (PINNs) versus dataset structure (Héméra). The latter approach is more direct for statistical estimation problems with an analytically known measurement model.

5.3 Limitations and Applicability Conditions

Four limitations must be mentioned.

Simulation-only validation. The simulation is physically rigorous (Siddon analytical model [25], SpekPy, ČMI experimental response, ICRU-57, XCOM) but does not capture full clinical complexity, multiple scattering, partial volume effects, inter-patient anatomical variability. The numerical thoracic phantom (section 4.7) constitutes a first validation step under realistic reconstruction conditions. Validation on a physical phantom and retrospective clinical data is required before clinical deployment.

Training domain dependency. Performance degrades for $Z_{\mathrm{eff}} > 16$ (outside the main domain). Extension to contrast agents (iodine, gadolinium) and high-density bone is required for contrast imaging and osteoporosis applications. The cortical bone inclusions in the phantom ($Z_{\mathrm{eff}} = 14.15$) lie at the boundary of the training domain, which partially explains the second-order artefacts observed in Figure 12.

Spectral characterisation sensitivity. Any drift in filter or detector response creates a mismatch with the learned physical prior. A periodic calibration protocol is required for robust deployment on a clinical scanner.

Inadequacy of standard FBP for R reconstruction. FBP reconstruction applied to the R sinogram inherits implicit assumptions calibrated for μ sinograms: beam-hardening correction by water-based polynomial approximation, *cupping* correction by smoothed radial profile. These corrections are physically inappropriate for the R sinogram, whose relationship with traversed thickness differs fundamentally from that of μ. Applying a post-reconstruction cupping correction calibrated for μ introduces an additional systematic distortion. As illustrated in Figure 12, Héméra makes the R sinogram *reconstructible*, but full exploitation of the physical information carried by R requires a dedicated reconstruction chain, adapted to the physics of the spectral ratio, which is the subject of the following publication.

5.4 Perspectives: Acquisition, Learning and Reconstruction Co-design

The results were obtained under standard CT scanner conditions, not optimised for exploiting the inter-channel physical prior. Substantial room for improvement exists by jointly rethinking the acquisition chain and the learning model, what the literature calls *co-design* [20].

Three priority levers: optimisation of the differential filter spectrum to maximise spectral separation at constant dose [5,6], extension to photon-counting CT, for which a multi-channel physical prior could offer even more substantial gains and integration of Héméra as a sinogram pre-conditioner in existing iterative reconstruction chains.

The validation on the numerical thoracic phantom (section 4.7) opens a direct perspective, the first FBP reconstructions of the Héméra-corrected R sinogram show that this domain is naturally more homogeneous in soft tissues than μ sonograms, bone inclusions generate localised artefacts without the propagating beam-hardening streaks characteristic of μ. Héméra, by making this sinogram usable, allowing to a combined (μ_1, μ_2, R) reconstruction that is potentially more informative prior to any post-processing. These results, together with the development of a reconstruction chain adapted to the physics of R, will be the subject of a dedicated publication.

6. Conclusion

Héméra proposes a physics-constrained learning approach for estimating the spectral attenuation ratio in dual-energy X-ray imaging. The four main contributions are:

1. A 98 % variance reduction and 79.9 % median error reduction without flux increase, with a dose-robust gain (mean factor ×2.90, maximum at low flux);
2. Strictly pixel-by-pixel processing (~14,000 FLOPs/pixel), compatible with real-time GPU execution, guaranteeing that all gains originate exclusively from the inter-channel physical prior;
3. An interpretable Bayesian theoretical framework, linking observed performance to inter-channel Fisher information ignored by the naive Poisson estimator;
4. Qualitative validation on a numerical thoracic phantom (22 tissue regions, 900 angles, low-dose regime) demonstrating the feasibility of FBP reconstruction of the Héméra-corrected R sinogram, transforming an unusable sinogram into an anatomically interpretable image.

The comparison with PWLS reveals a complementarity. Héméra preserves local gradients where PWLS smooths them. The reconstruction results also show that standard FBP, calibrated for μ sinograms, is structurally ill-suited to R, hence the need for a dedicated algorithm.

The contribution is first and foremost conceptual, embedding the physical prior in the very structure of the training data, rather than in the loss penalty, offers a way around the apparent limits of Poisson statistics, without altering the measurement device. The network does not act here as a regressor, but as a **mechanism of inference under physical constraint**, in the Bayesian sense of the term.

Next steps: validation on a physical phantom, then retrospective clinical data; extension to contrast materials; development of a reconstruction chain dedicated to the R sinogram, drawing on its specific properties (native absence of beam-hardening artefacts, sensitivity to chemical composition); acquisition-chain co-design.

Declarations

Code and Data Availability

Source code and simulated dataset will be made available upon request. A public repository (GitHub/Zenodo) is planned for the final publication.

Funding

This work was carried out without external funding.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Knoll G.F. *"Radiation Detection and Measurement, 4th ed.,"* Wiley, 2010.

[2] Leo W.R. *"Techniques for Nuclear and Particle Physics Experiments, 2nd ed.,"* Springer, 1994.

[3] Beer A. *"Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten,,"* Annalen der Physik, 162(5):78-88, 1852.

[4] Berger M.J., Hubbell J.H. et al. *"XCOM: Photon Cross Sections Database,,"* NIST Standard Reference Database 8. https://physics.nist.gov/xcom

[5] Alvarez R.E., Macovski A. *"Energy-selective reconstructions in X-ray computerized tomography,,"* Physics in Medicine & Biology, 21(5):733-744, 1976.

[6] McCollough C.H., Leng S., Yu L., Fletcher J.G. *"Dual- and multi-energy CT: principles, technical approaches, and clinical applications,,"* Radiology, 276(3):637-653, 2015.

[7] Fessler J.A. *"Penalized weighted least-squares image reconstruction for positron emission tomography,,"* IEEE Trans. Medical Imaging, 13(2):290-300, 1994.

[8] Wang G., Ye J.C., De Man B. *"Deep learning for tomographic image reconstruction,,"* Nature Machine Intelligence, 2(12):737-748, 2020.

[9] Chen H., Zhang Y. et al. *"Low-dose CT with a residual encoder-decoder convolutional neural network,,"* IEEE Trans. Medical Imaging, 36(12):2524-2535, 2017.

[10] Yang Q., Yan P. et al. *"Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss,,"* IEEE Trans. Medical Imaging, 37(6):1348-1357, 2018.

[11] Ronneberger O., Fischer P., Brox T. *"U-Net: Convolutional networks for biomedical image segmentation,,"* MICCAI, LNCS 9351:234-241, 2015.

[12] Caves C.M. *"Quantum-mechanical noise in an interferometer,,"* Physical Review D, 23(8):1693-1708, 1981.

[13] Bishop C.M. *"Pattern Recognition and Machine Learning,,"* Springer, 2006.

[14] Gelman A. et al. *"Bayesian Data Analysis, 3rd ed.,,"* CRC Press, 2013.

[15] Raissi M., Perdikaris P., Karniadakis G.E. *"Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear PDEs,,"* Journal of Computational Physics, 378:686-707, 2019.

[16] Kay S.M. *"Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory,,"* Prentice Hall, 1993.

[17] Mayneord W.V. *"The significance of the Röntgen,,"* Acta UICC, 2:271-282, 1937. - Spiers F.W., British Journal of Radiology, 19:52-63, 1946.

[18] Abadi M. et al. *"TensorFlow: Large-scale machine learning on heterogeneous systems,,"* arXiv:1603.04467, 2016. - Chollet F., Keras, https://keras.io, 2015.

[19] Kingma D.P., Ba J. *"Adam: A method for stochastic optimization,,"* ICLR 2015, arXiv:1412.6980.

[20] Unser M. *"A unifying representer theorem for inverse problems and machine learning,,"* Foundations of Computational Mathematics, 21:941-960, 2021.

[21] Poludniowski G. et al. *"SpekPy v2.0 - a software toolkit for modelling X-ray tube spectra,,"* Physics in Medicine & Biology, 66(7):075021, 2021.

[22] ICRU Report 57. *"Conversion Coefficients for use in Radiological Protection against External Radiation,"* International Commission on Radiation Units and Measurements, 1998.

[23] He K. et al. *"Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification (PReLU),,"* ICCV, 2015.

[24] Ioffe S., Szegedy C. *"Batch normalization: Accelerating deep network training by reducing internal covariate shift,,"* ICML, 2015.

[25] Siddon R.L. *"Fast calculation of the exact radiological path for a three-dimensional CT array,"* Medical Physics, 12(2):252-255, 1985.

[26] Ge et al., Medical Physics 48(6), 2021 (DIRECT-Net, arXiv:2008.08331)

[27] Perelli & Andersen, arXiv:2406.00479, 2024 (E2E-DEcomp)

[28] Zhang et al., arXiv:2504.07753, 2025 (VIP-DECT)