Reconstructing Tracker Material with Information Field Theory

Alexander Pesch Berrocal, Johannes Erdmann, Jan Lukas Späh

Physics Institute III A, RWTH Aachen University
alexander.pesch.berrocal@rwth-aachen.de

ML4Jets September 15, 2026

Gefördert durch:
Bundesministerium für Forschung, Technologie und Raumfahrt

Physics Institute III A | RWTH AACHEN UNIVERSITY

Overview

Goal

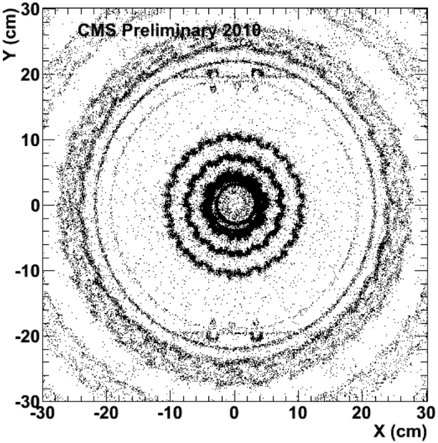
Infer the material distribution in the tracker volume from collision data.

Why this matters

A better material description can reduce material-related systematic uncertainties and improve tracking and object reconstruction.

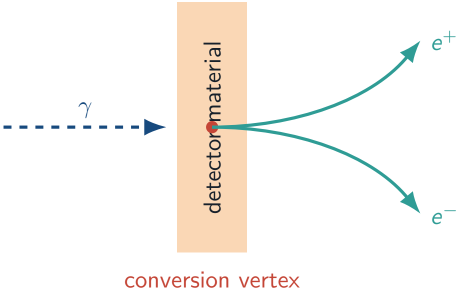
Key observables

- photon-conversion vertices $\Leftrightarrow$ radiation length X_0 — focus of this talk;
- nuclear-interaction vertices $\Leftrightarrow$ interaction length λ_I — future extension.



Source: CMS Collaboration, CMS-PAS-TRK-10-003, Fig. 9; cds.cern.ch/record/1279138.

Why photon conversions trace radiation length

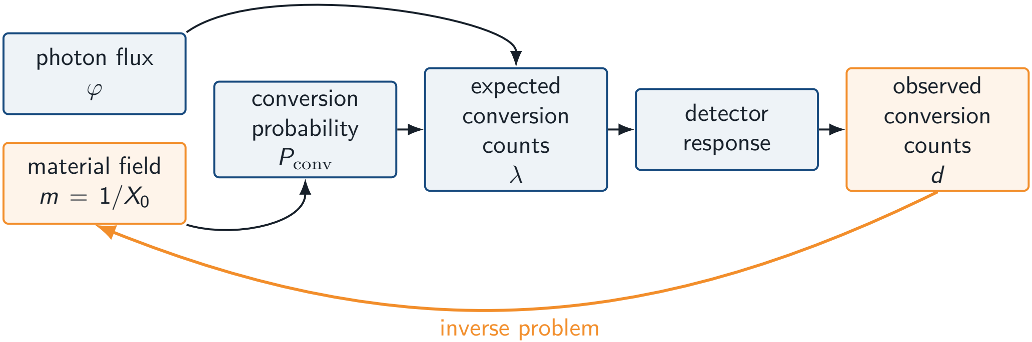


For a high-energy photon crossing material budget x/X_0,

$$P_{\text{conv}} \simeq 1 - \exp\left[-\frac{7}{9}\frac{x}{X_0}\right].$$

- more $x/X_0 \Rightarrow$ higher P_{conv};
- reconstructed conversion vertices can probe the underlying material distribution.

From conversion vertices back to material



Reconstruction task

We observe noisy counts of conversion vertices. How can we infer possible material fields that explain the data?

Information Field Theory: Bayesian inference for fields

Infer a joint probability distribution over material and photon-flux fields.

$$\underbrace{P(m,\varphi \mid d,c)}_{\text{posterior}} \propto \underbrace{P(d,c \mid m,\varphi)}_{\text{likelihood}} \underbrace{P(m)P(\varphi)}_{\text{field priors}}$$

d: conversion counts, c: photon counts

Joint likelihood $P(d,c \mid m,\varphi)$

How probable is the observed data given the fields?

Priors $P(m)$ and $P(\varphi)$

Which fields are physically plausible before seeing the data?

IFT framework · NIFTy software implementation

IFT: Bayesian inference for fields.

NIFTy: computes an approximate posterior using variational inference.

Posterior samples provide material and photon-flux estimates with uncertainties.

NIFTy → G. Edenhofer et al. arXiv:2402.16683

Developing the method with known detector material

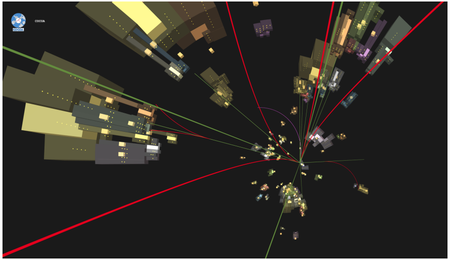
- **COCOA:** Pythia8 event generation + Geant4 particle transport in a configurable detector.
- Custom ATLAS-like geometry.
- Minimum-bias events provide photons with broad angular coverage and large statistics.

Idealized reconstruction

100% reconstruction efficiency, no fakes, perfect vertex resolution.

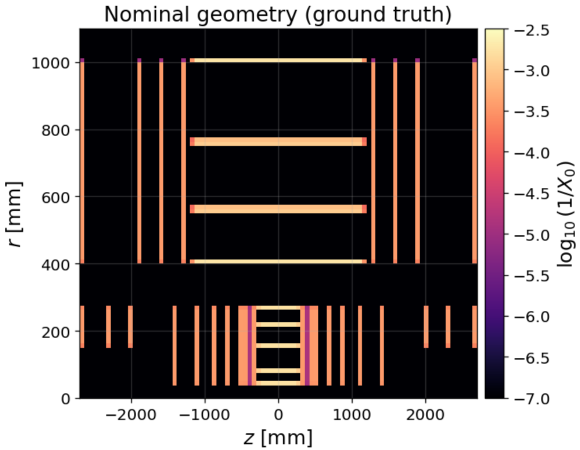
COCOA → F. A. Di Bello et al. arXiv:2303.02101

simulated collision events	500,000
conversion vertices used	655,411
phase-space cuts	$E_\gamma \geq 500$ MeV
	$\lvert\eta\rvert < 3$



Example minimum-bias event with multiple photon conversions (red curves=e^+/e^-)

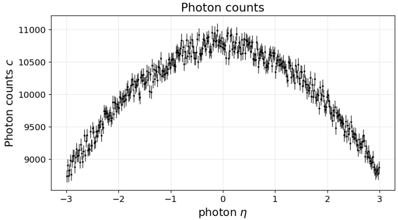
Ground truth: an idealized layered detector

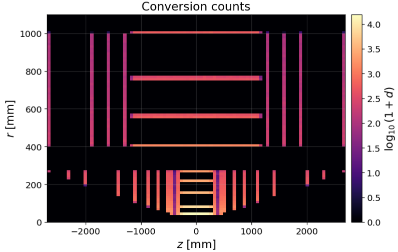


- ϕ symmetry $\Rightarrow$ 2D study
- 88×90 bins in (r, z)
- $0 < r < 1100$ mm
- $|z| < 2700$ mm
- thin barrel layers and endcap discs

The known Geant4 geometry is used as ground truth.

Photon flux must be inferred too





The ambiguity

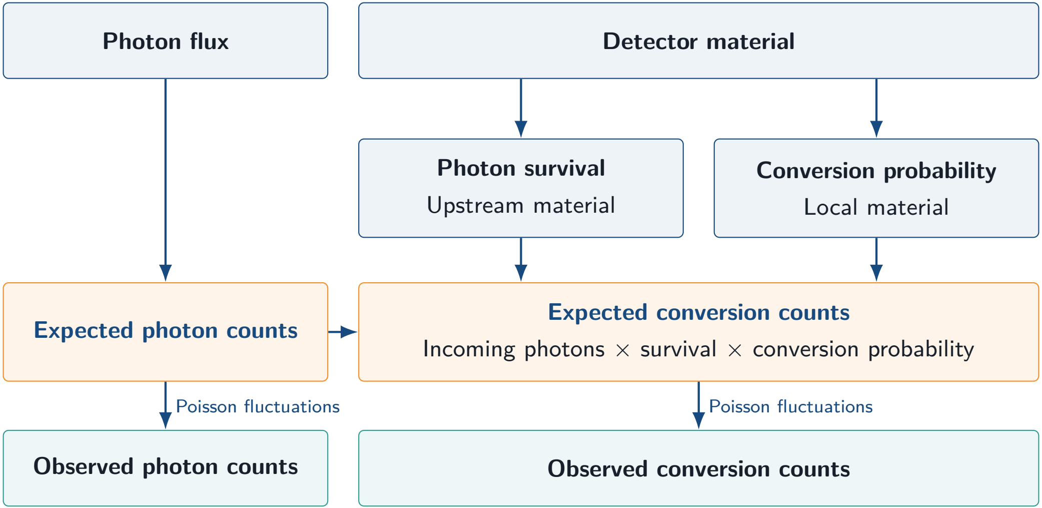
More conversion vertices can mean more material budget, or simply more incident photons.

Joint inference

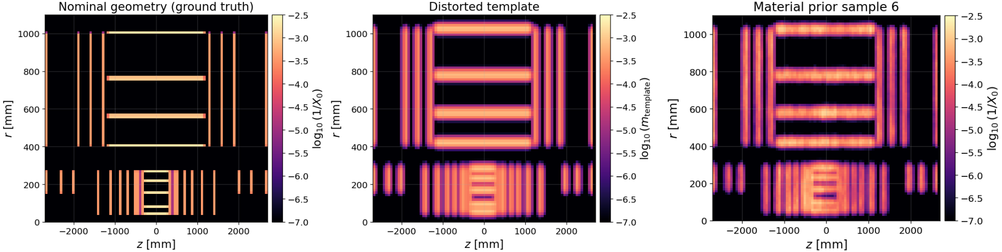
Photon counts c constrain the photon flux versus η.

We infer the flux and material fields jointly, propagating flux uncertainty into the material reconstruction.

Forward model: from material and flux to counts



A deliberately imperfect starting template

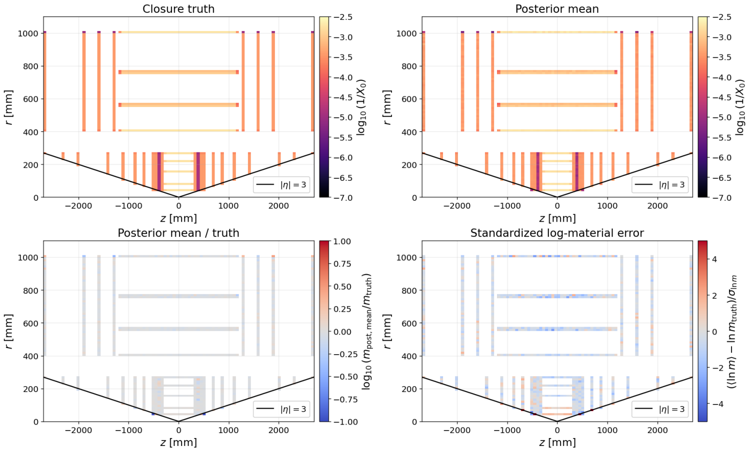


Template: shifted, axis-scaled and blurred

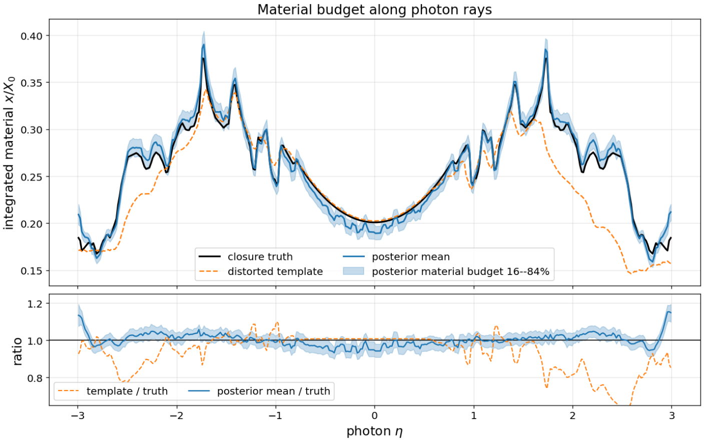
$$\Delta r = +12 \text{ mm}, \ \Delta z = -30 \text{ mm}$$

We roughly know where material should be, the question is where it is precisely.

Reconstruction recovers material distribution



Ray-integrated material closes within posterior uncertainty



Ratio means

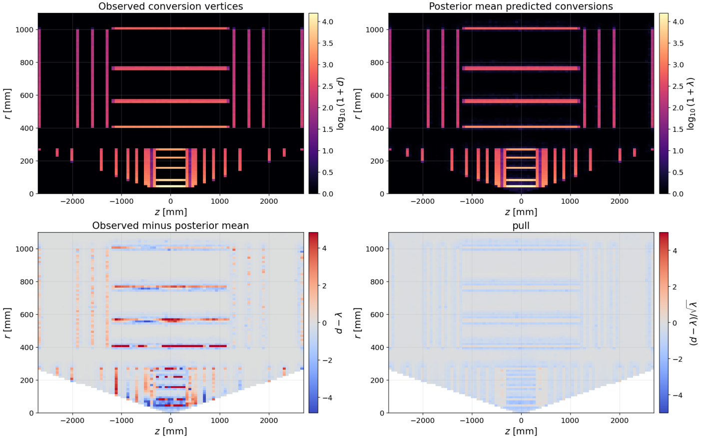
$$\frac{B_{\mathrm{template}}}{B_{\mathrm{truth}}} = 0.937,$$

$$\frac{B_{\mathrm{post,mean}}}{B_{\mathrm{truth}}} = 1.003$$

central 68% interval $[0.974, 1.032]$.

The inferred field restores the integrated material budget.

Posterior-predictive check



Count closure

Observed conversions:

655,411

Posterior mean predicted:

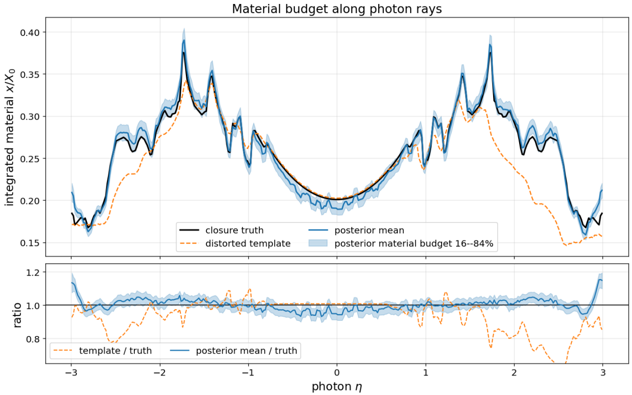
655,828

Total-count difference (prediction vs. observed):

+0.064%

Pull reveals prior and forward-model dependence.

Takeaways and path to collision data



What works already

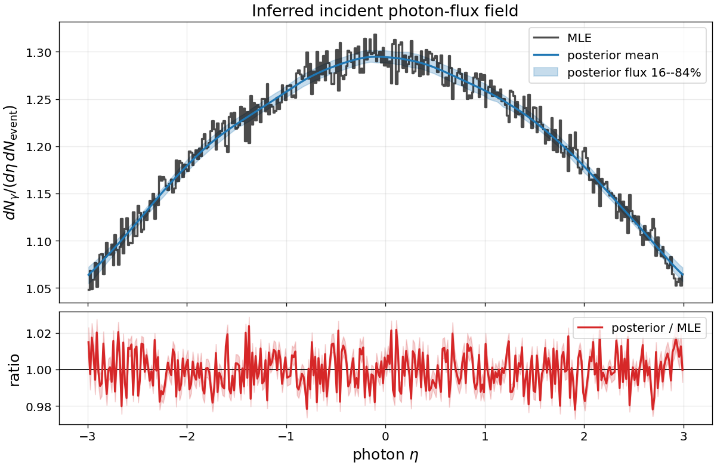
- Conversion vertices enable Bayesian reconstruction of a structured r–z material field.
- Posterior samples provide uncertainties and predictive validation.

Next steps

1. make inference more realistic;
2. use nuclear interactions as an independent λ_I channel.

Backup

Backup: inferred photon-flux field



Source: Independent 500k-event calibration sample; 48 posterior samples.

Backup: forward model from material and flux to counts

Photon flux

φ_a

Material

$$\tau_{ak} = \frac{7}{9}\, m_{q(a,k)}\, \Delta s_{ak}$$

Survive to segment k

$$S_{ak} = \exp\left[-\sum_{l<k} \tau_{al}\right]$$

Convert in segment k

$$p_{ak} = 1 - e^{-\tau_{ak}}$$

Expected photon counts

$$\Phi_a = N_{\mathrm{evts}}\, \Delta\eta\, \varphi_a$$

Expected conversion counts

$$\lambda_{ak} = \Phi_a\, S_{ak}\, p_{ak}$$

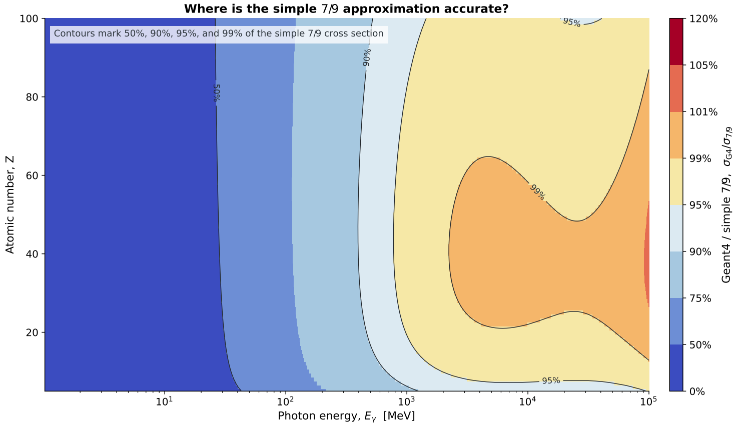
Observed photon counts

$$c_a \sim \mathrm{Poisson}(\Phi_a[\varphi])$$

Sum into pixels $\longrightarrow$ observed conversion counts

$$\lambda_q = \sum_{a,k:\, q(a,k)=q} \lambda_{ak} \qquad d_q \sim \mathrm{Poisson}(\lambda_q[m, \varphi])$$

Backup: why can the 7/9 approximation bias material?



Response mismatch

In much of the relevant energy range,

$$\frac{\sigma_{\mathrm{G4}}}{\sigma_{7/9}} < 1.$$

The 7/9 factor is a simple high-energy approximation. Where the ratio is below one, it overpredicts conversions for fixed material.

$$m_{\mathrm{fit}} \approx m_{\mathrm{true}} \frac{\sigma_{\mathrm{G4}}}{\sigma_{7/9}} < m_{\mathrm{true}}.$$

Bias depends on energy and composition; use an energy-dependent response.

Backup: references

- CMS Collaboration, *Studies of Tracker Material*, CMS-PAS-TRK-10-003 (2010), CDS 1279138.
- T. Enßlin, *Information field theory*, arXiv:1301.2556.
- G. Edenhofer et al., *Re-Envisioning Numerical Information Field Theory (NIFTy.re): A Library for Gaussian Processes and Variational Inference*, arXiv:2402.16683.
- F. A. Di Bello et al., *Configurable calorimeter simulation for AI applications*, arXiv:2303.02101.