PROCEEDINGS OF SCIENCE

Bias Reduction Using Expectation Maximization in the Optimization of an AI-Assisted Muon Tomography System

Marta de la Puente Santos,[a,*] **Zahraa Zaher,**[b,d] **Maxime Lagrange,**[b,d] **Andrea Giammanco**[b,d] **and Pietro Vischia**[a,c,d]

[a] *Universidad de Oviedo*
[b] *Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université catholique de Louvain*
[c] *ICTEA*
[d] *MODE Collaboration*
E-mail: marta.delapuente.santos@gmail.com, zahraa.zaher@uclouvain.be, maxime.lagrange@uclouvain.be, Andrea.Giammanco@cern.ch, pietro.vischia@cern.ch

Muon scattering tomography is a non-invasive imaging technique that exploits cosmic ray muons to visualise the interior of dense objects. We implemented an Expectation Maximisation (EM) algorithm within a modern muography reconstruction framework and compared its performance with established methods (PoCA, BCA, ASR). Using simulated data, EM produced reconstructions with clearer density contrast in the XY projection and demonstrated potential for bias reduction. However, the method exhibited systematic deviations in high-density regions, poor performance in XZ and YZ projections, and high computational cost. These findings highlight both the promise of EM for enhanced imaging and the need for further optimisation, including improved parameter tuning, noise mitigation, and integration with real experimental data.

Fifth MODE Workshop on Differentiable Programming for Experiment Design (MODE2025)
8-13 June 2025
Kolymbari, Crete, Greece

*Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0) All rights for text and data mining, AI training, and similar technologies for commercial purposes, are reserved. ISSN 1824-8039 . Published by SISSA Medialab.

PoS(MODE2025)014

1. Introduction

Muon scattering tomography (MST) enables non-destructive imaging of dense or shielded materials by analysing the deviation of cosmic ray muons as they traverse matter. These muons are secondary particles generated in the upper atmosphere, primarily through the decay of charged pions and kaons produced in cosmic ray interactions with atomic nuclei [1].

At sea level, muons have a mean energy of approximately 4–5 GeV [1, 2]. Their relatively high energy and weak interactions with matter allow them to penetrate several metres of dense material. When traversing a medium, muons undergo multiple Coulomb scattering, mainly through interactions with atomic nuclei [3, 4]. By measuring deflection angles and reconstructing muon trajectories upstream and downstream of the volume of interest, one can infer the internal scattering density of the material, which correlates with its atomic number and structure.

The goal of MST is to reconstruct a three-dimensional (3D) map of scattering density within a discretised volume. Several methods have been proposed, ranging from simple geometric approaches such as the point-of-closest-approach (PoCA) algorithm [4], to statistical methods including the Binned Clustered Algorithm (BCA) [5], Angle Statistics Reconstruction (ASR) [6], and expectation maximisation (EM) [4, 7].

The EM algorithm provides a probabilistic framework that iteratively refines the estimate of scattering density using measured muon trajectories and scattering angles. This work presents the integration of EM into a full simulation and reconstruction pipeline (`TomOpt`, [8]) and compares it with other methods in terms of bias, resolution, and computational efficiency.

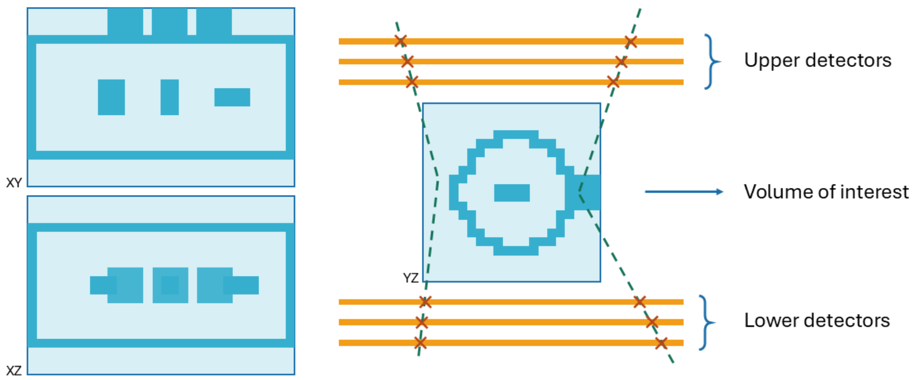


Figure 1: Schematic representation of the muon-tomography system. The upper and lower tracking detectors (orange) record the trajectories of cosmic ray muons (green dashed lines), enabling the reconstruction of their scattering as they traverse the volume of interest (blue). The internal structure of the scanned object is shown through its three principal projections (XY, XZ, YZ).

2. Reconstruction Techniques

In muon tomography, estimating scattering density requires reconstructing how muons are deflected as they pass through the target. Since muons are minimally ionising particles, their trajectories can be measured upstream and downstream of the volume of interest (see Fig. 1). Several methods exist for inferring the scattering distribution from these data. Here we compare

PoS(MODE2025)014

three widely used approaches—PoCA, BCA, and ASR—with the Expectation Maximisation (EM) algorithm, discussed in the next section.

PoCA: Point of Closest Approach

PoCA approximates the scattering by assuming that each muon undergoes a single dominant deflection. The scattering site is taken as the point where the incoming and outgoing tracks are closest. The volume is discretised into voxels, and for each voxel the number of PoCA points it contains is recorded. This yields a coarse but efficient map, though accuracy is limited when scattering is distributed or multiple deflections occur.

BCA: Binned Clustered Algorithm

BCA extends PoCA by emphasising muons with large scattering angles, which are more likely to signal high-Z materials. After computing PoCA points, they are binned into voxels, and the $n_{\max}$ points with the largest angles are selected. A voxel score is then obtained from an angle-weighted distance metric, with a cutoff applied to suppress background from small-angle scatters. This improves localisation of dense regions, though intensity inconsistencies can remain.

ASR: Angle Statistics Reconstruction

ASR distributes each muon's scattering contribution along the portion of its trajectory common to both incoming and outgoing tracks. The scattering angle is projected only into those voxels, each receiving a weighted share. The resulting maps are smoother and capture transverse (XY) variations more faithfully than PoCA or BCA, especially for extended targets. However, because contributions are spread along the path, resolution degrades in depth (XZ and YZ) where smearing becomes significant.

3. Expectation Maximisation Algorithm

The Expectation Maximisation (EM) algorithm offers a robust statistical framework for estimating the scattering density λ_j across a discretised three-dimensional volume in muon tomography, where j denotes the voxel index. Unlike deterministic algorithms such as PoCA or ASR, EM accounts for the uncertainty inherent in multiple Coulomb scattering by modelling the scattering process as a latent-variable problem. This implementation was developed in full as the core contribution of the author's undergraduate thesis [10].

Iterative Structure

Each iteration of the EM algorithm consists of two alternating steps:

- **E-step (Expectation):** For each muon and voxel, calculate the expected contribution to scattering based on the current estimate of voxel density.
- **M-step (Maximisation):** Update the voxel-wise scattering density λ_j to maximise the likelihood of the observed scattering angles across the dataset.

PoS(MODE2025)014

This procedure is repeated iteratively, refining the estimate of λ_j until convergence. The overall workflow of the EM algorithm, from the input data to the voxel-density update, is summarised schematically in Fig. 2.

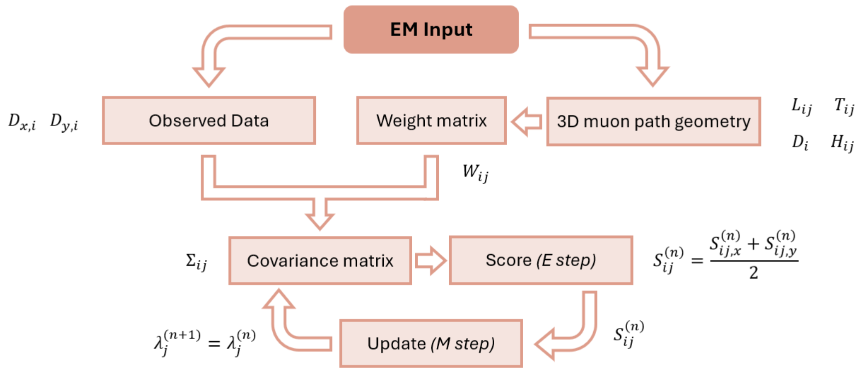


Figure 2: Simplified flowchart of the EM algorithm applied to muon tomography. The diagram shows the iterative cycle between the observed data D_i, the hidden voxel-level scattering information H_{ij} (related to the 3D muon path geometry), covariance estimation, score evaluation, and voxel-density update.

Mathematical Model

The EM algorithm relies on an approximate description of muon trajectories through the volume. A common approach is the straight-line approximation, where a single trajectory connects the entry and exit points of the muon. Here we adopt a more accurate variant based on the PoCA approximation, modelling the trajectory as two connected segments—from the entry point to the PoCA, and from the PoCA to the exit. This better reflects the localisation of the dominant scattering event while remaining tractable.

For each muon i, the distribution of its observed scattering among traversed voxels is described by a weight matrix W_{ij}. Voxels where the muon travels a longer distance (L_{ij}) and has a greater remaining path length (T_{ij}) receive larger weights, reflecting their higher probability of contributing to the deflection. Thus, each muon assigns fractions of its scattering to the voxels along its path, and these weighted contributions enter the iterative EM update.

A uniform radiation length L_{rad} is used only to initialise the voxel densities λ_j before the EM iterations.[1] The voxel density λ_j at iteration $n+1$ is updated as

$$\lambda_j^{(n+1)} = \frac{1}{2M_j} \sum_{i:L_{ij}\neq 0} S_{ij}^{(n)} ,$$

where M_j is the number of muons intersecting voxel j, and $S_{ij}^{(n)}$ the fraction of muon i's scattering attributed to voxel j at iteration n. In the E-step, the expectations are computed using the voxel weights W_{ij} and the covariance matrix Σ_{ij} of scattering angles and displacements, estimating how much each voxel contributes to the observed deflection.

In the M-step, the voxel densities λ_j are updated by maximising the expected log-likelihood of the hidden variables H_{ij}, conditioned on the measured data D_i and the current estimates. Iterating

[1] L_{rad} is the characteristic scale of electromagnetic interactions in a material and determines its scattering power.

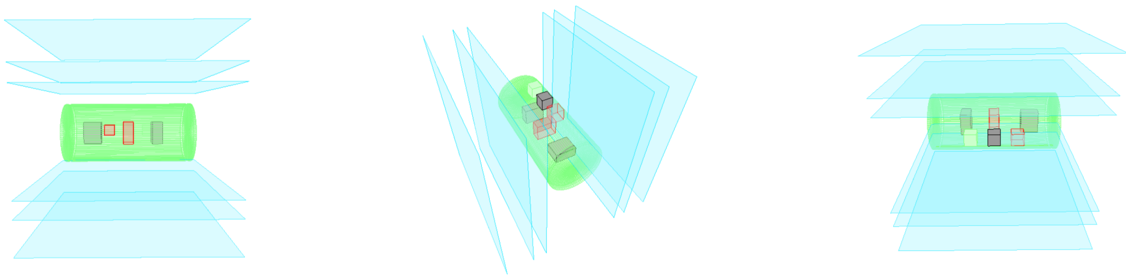
PoS(MODE2025)014

this process progressively refines the scattering-density map, yielding an increasingly accurate reconstruction of the object's internal structure. Further details of S_{ij}, W_{ij}, and Σ_{ij} are given in [4, 10].

4. Simulation Setup

The scattering data used in this study were generated with the GEANT4 framework [9]. The simulated geometry (see Fig. 3) consisted of an iron barrel containing several high-density blocks placed inside, as well as additional blocks located outside the barrel. The full specification of the geometry, including positions, dimensions and materials, is available in the public repository at `muograph/data/iron_barrel/specs.txt`.

A total of $n_\mu = 9584$ cosmic-ray muons were simulated with realistic angular and momentum distributions at sea level. Their trajectories were propagated through the detector planes and the test object, and all scattering interactions were recorded. These simulated tracks constitute the input to the reconstruction algorithms discussed in the following sections.



(a) Front view **(b)** Oblique view **(c)** Tilted view

Figure 3: Visualisations of the simulated geometry in GEANT4 from different perspectives. The setup includes an iron barrel (green) with internal blocks as well as external blocks placed above it. The tracking detector planes are shown in blue.

PoS(MODE2025)014

5. Results

To compare the performance of the different reconstruction algorithms (PoCA, BCA, ASR, and EM), identical event-selection criteria were applied. Only muons with a scattering angle larger than $\Delta\theta > \pi/180$ were retained, ensuring that the analysis focuses on significantly deflected tracks. This cut reduces the available statistics but enhances sensitivity to the high-Z materials of interest, while also keeping the computational requirements of the EM algorithm manageable. All reconstruction algorithms were evaluated using a common voxel grid with a spatial resolution of 30 mm per voxel.

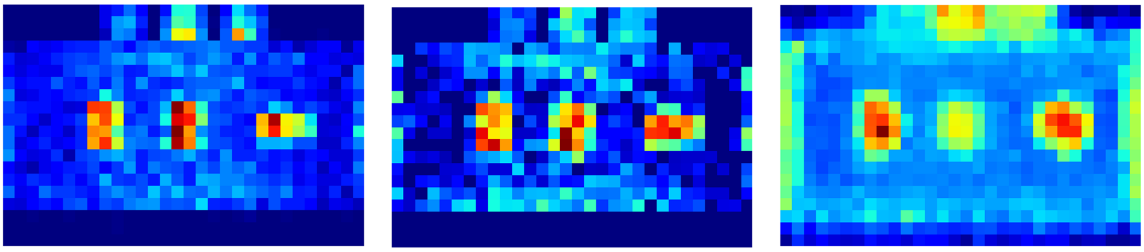
5.1 Reference algorithms (PoCA, BCA, ASR)

The reference algorithms were evaluated first, focusing on the XY plane (see Fig. 4).

- PoCA provided a basic but quick overview of the internal structure, with visible high-scattering regions but no absolute density calibration.

- BCA improved discrimination of highly scattered muons but produced intensity inconsistencies, especially in the upper cubes.
- ASR spread scattering along the full muon path, providing smoother maps and better edge visibility, at the expense of some spatial resolution.

These results set a baseline for evaluating EM, which was the focus of this work.



(a) PoCA (b) BCA (c) ASR

Figure 4: XY projections of the reconstructed scores obtained with the three reference algorithms: (a) PoCA, (b) BCA, and (c) ASR. These maps correspond to the same object shown in Fig. 1. PoCA provides a quick but coarse overview, BCA enhances high-scattering regions though with some intensity inconsistencies, and ASR yields smoother maps with better edge visibility.

5.2 EM algorithm behaviour

For EM, voxelised scattering density maps were reconstructed (voxel size 30 mm, 100 iterations) to study the convergence from an initial uniform estimate toward the internal structure.

5.2.1 XY projection

Fig. 5 shows how the XY projection evolves from a nearly featureless distribution at the first iteration to one where the three inner blocks are clearly visible after 100 iterations. High-Z regions appear with stronger intensity and the noise level is reduced, illustrating the algorithm's ability to progressively refine voxel contributions.

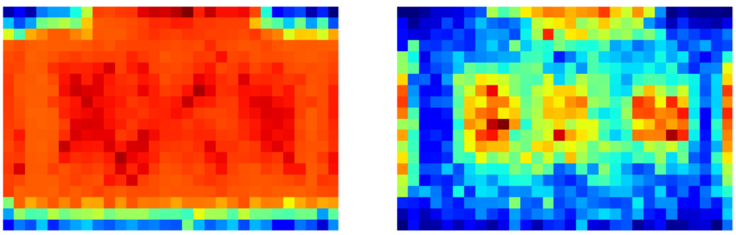


Figure 5: XY projection of the reconstructed scattering density using EM with $L_{\mathrm{rad}} = 70$ mm. Left: first iteration. Right: iteration 100, where the inner blocks and density contrast become visible.

5.2.2 XZ projection

As shown in Fig. 6, the XZ projection fails to recover the expected horizontal band structure. Even after convergence the map remains dominated by diffuse noise, likely due to limited muon statistics at shallow angles, which are essential to constrain this plane.

PoS(MODE2025)014

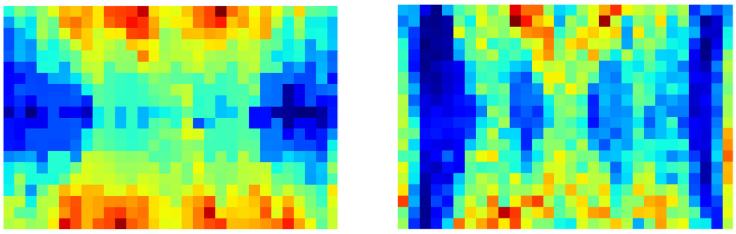


Figure 6: XZ projection of the reconstructed scattering density. Left: first iteration. Right: iteration 100, where the expected band structure remains poorly defined.

5.2.3 YZ projection

In the YZ projection (Fig. 7), the circular container and internal square feature are not well reproduced after 100 iterations. The map is dominated by blurred, noise-like patterns, suggesting insufficient geometric constraints in this orientation due to the near-vertical muon flux.

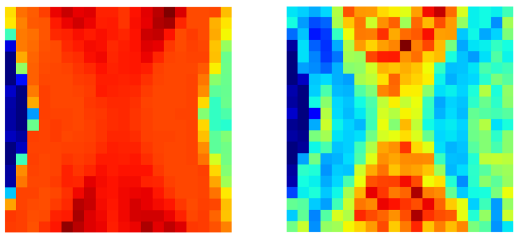


Figure 7: YZ projection of the reconstructed scattering density. Left: first iteration. Right: iteration 100, still dominated by noise-like features.

6. Limitations and Future Work

The EM algorithm developed in this work was able to reconstruct scattering density maps from simulated muon data, but the results did not reach the desired quality. No clear convergence toward physically consistent values was observed, and the reconstructed densities showed significant deviations from expectations. Systematic bias appeared in radiation length estimates, particularly in high-density regions, and the algorithm is computationally demanding, taking about 40 minutes per full run[2]. Statistics are further limited by the angular selection applied, while thresholding is required to suppress extreme outlier voxels that would otherwise dominate the maps. The method also shows strong sensitivity to the initial value of L_{rad}, which affects both stability and convergence.

Future work will focus on applying the algorithm to real data to test robustness under non-ideal conditions. Automatic parameter tuning, including a more reliable initialisation of L_{rad}, will be introduced to improve convergence and reduce bias. Detector spatial and angular resolution will be incorporated to achieve more realistic, noise-resilient reconstructions. Finally, the algorithm will be integrated into the `TomOpt` framework, with the prospect of combining EM with machine-learning techniques to enhance stability and computational performance.

[2] All simulations were carried out on a local workstation running Jupyter Notebook within Visual Studio Code. The code was executed sequentially (no parallelization).

PoS(MODE2025)014

7. Conclusions

This work presented the development and implementation of the Expectation Maximisation algorithm for muon scattering tomography, aimed at reconstructing the scattering density in a discretised three-dimensional volume. The algorithm was built from scratch, incorporating a statistical treatment of angular deviations and lateral displacements, and integrated into a voxel-based reconstruction framework.

Although the reconstruction did not fully recover the expected internal structure, it provided useful insight into the sensitivity of EM to geometry, initial parameters, and muon angular distribution. The study revealed systematic patterns and emphasised the need for robust initialization and sufficient statistics for reliable results.

The modular codebase, validated with realistic GEANT4 simulations, is ready for integration into the `TomOpt` framework. Future work will target performance optimisation, physical regularisation, and hybrid methods combining EM with machine learning. Despite current limitations, EM shows strong potential as a foundation for more advanced algorithms in muon tomography.

References

[1] S. Navas *et al.* (Particle Data Group), "Review of particle physics," *Phys. Rev. D* **110** (2024) 030001.

[2] L. Bonechi, R. D'Alessandro and A. Giammanco, "Atmospheric muons as an imaging tool," *Reviews in Physics* **5** (2020) 100038.

[3] D. E. Groom, N. V. Mokhov and S. I. Striganov, "Muon stopping power and range tables 10 MeV–100 TeV," *Atomic Data and Nuclear Data Tables* **78** (2001) 183–356.

[4] L. J. Schultz *et al.*, "Statistical reconstruction for cosmic ray muon tomography," *IEEE Trans. Image Process.* **16** (2007) 1985–1993.

[5] C. Thomay *et al.*, "A binned clustering algorithm to detect high-Z material using cosmic muons," *J. Instrum.* **8** (2013) P10013.

[6] M. Stapleton *et al.*, "Angle statistics reconstruction: a robust reconstruction algorithm for muon scattering tomography," *J. Instrum.* **9** (2014) P11019.

[7] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Series B (Methodological)* **39** (1977) 1–22.

[8] G. C. Strong *et al.*, "TomOpt: Differential optimisation for task- and constraint-aware design of particle detectors in the context of muon tomography," *Mach. Learn.: Sci. Technol.* **5** (2024) 035002.

[9] S. Agostinelli *et al.*, "Geant4—a simulation toolkit," *Nucl. Instrum. Meth. A* **506** (2003) 250–303.

[10] M. de la Puente Santos, *Bias reduction using Expectation Maximization in the optimization of an AI-assited muon tomography system*, Undergraduate thesis, Universidad de Oviedo (2025).

PoS(MODE2025)014