Technical Report

TOMOPT: Muon Tomography Experiment Optimization

Maxime Lagrange,[1,2] Giles C. Strong,[1,3] Anna Bordignon,[4] Florian Bury,[5] Tommaso Dorigo,[1,3] Andrea Giammanco,[1,2] Mariam Heikal,[6] Max Lamparth,[1,7] Federico Nardi,[1,8,9] Aitor Orio,[10,11] Pietro Vischia,[1,12] and Haitham Zaraket[1,13,14]

[1]*MODE Collaboration, https://mode-collaboration.github.io/*

[2]*Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, Louvain la Neuve, Belgium*

[3]*Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Italy*

[4]*Department of Statistical Sciences, University of Padova, Italy*

[5]*University of Bristol, UK*

[6]*American University of Beirut, Beirut, Lebanon*

[7]*Physik-Department, Technische Universität München, Germany*

[8]*Université Clermont-Auvergne, Clermont-Ferrand, France*

[9]*Department of Physics and Astronomy, University of Padova, Italy*

[10]*Muon Tomography Systems S.L., Bilbao, Spain*

[11]*Universidad de Cantabria (UC), Santander, Spain*

[12]*Universidad de Oviedo and ICTEA, Oviedo, Spain*

[13]*Multi-Disciplinary Physics Laboratory, Optics and Fiber Optics Group, Faculty of Sciences, Lebanese University, Lebanon*

[14]*Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France*

Corresponding author: Maxime Lagrange
Email address: maxime.lagrange@uclouvain.be

Abstract

The TOMOPT software is a tool to optimize the geometrical layout and specifications of detectors designed for muon scattering tomography. Based on differentiable programming techniques, TOMOPT consists in a modular pipeline that models all the aspects of a muon tomography task, from the generation and interaction of cosmic ray muons with a parameterized detector and passive material, to the inference of the volume properties. This enables the optimization of the detector parameters via gradient descent, to suggest optimal detector configurations and specifications. This optimization is subjected to various external constraints such as cost, logistics, and material identification efficiency.

Keywords: machine learning, artificial intelligence, computer science
DOI: 10.31526/JAIS.2024.492

1. INTRODUCTION

Muon scattering tomography (MST) is an imaging technique that uses the natural flux of muons produced by cosmic rays interacting with the upper atmosphere, and the deflection of those muons in materials due to multiple Coulomb scattering [1]. Many applications are being developed [2] in various fields, which include the identification of nuclear material in waste drums [3], border control [4], nuclear warhead verification in the context of nonproliferation treaties [5], or monitoring and preventive maintenance of industrial facilities [6].

Whatever the use case, MST detector conception aims at reaching the best performance while respecting specific constraints, whether in terms of budget, logistics (e.g., not any size or shape can be allowed on site), or data acquisition time. While standard detector design approaches revolve around optimizing these proxy objectives rather than the final measurement, we propose the use of differentiable programming tools to provide analytic gradients along which to shift design parameters to best improve end-goal performance.

2. TOMOPT

The simulation and optimization procedures developed for this study are embedded in a Python-based software referred to as TOMOPT, which is still in development at the time these lines are written and will be made public in the near future. TOMOPT's ambition is to provide a modular and fully-differentiable pipeline with tools and methods required for the optimization of a muon

scattering tomography detector. The first section of this document will describe TOMOPT's structure and different modules. For an extensive description of TOMOPT, the reader is invited to see [7].

The optimization is performed as an iterative gradient-descent algorithm [8], where each detector parameter x is updated at each step (epoch) by computing the loss function gradient $\nabla_x \mathcal{L}(x^k)$ reweighed by a learning rate η, as shown in equation (1) for the step $k+1$.

$$x^{k+1} = x^k - \eta \nabla_x \mathcal{L}\left(x^k\right). \tag{1}$$

The role of the loss function $\mathcal{L}(x^k)$ is to mathematically express the idea of an optimal detector, in terms of both performance and cost. Because gradient computation requires the loss function to be differentiable with respect to the detector parameters, the whole inference chain is implemented in a differentiable way, utilizing the PYTORCH framework [9]. TOMOPT applies the same workflow at each epoch, as presented in Figure 1:

(1) *Initialization*: The user provides an initial detector configuration to be optimized. This step is only performed once, prior to the first epoch.
(2) *Muon generation*: Cosmic muons position, direction, and energy are sampled from models found in the literature, as explained in Section 2.1.
(3) *Muon propagation and detection*: Muons are propagated through the volume of interest and their position is recorded by the detection planes, referred to as *panels*, as described in Section 2.2.
(4) *Volume inference*: Muons tracks are reconstructed and used by the volume inference method. The latter being task-specific, the user is free to implement its own inference method. TOMOPT provides an in-build method for X_0 inference, further described in Section 2.3.
(5) *Loss function computation*: Given the detector parameters x^k, the loss function $\mathcal{L}(x^k)$ is computed. Again, the choice of the loss function is free and has to be adapted to the use-case.
(6) *Gradient-descent*: Given the loss function $\mathcal{L}(x^k)$, detector parameter gradients $\nabla_x \mathcal{L}(x^k)$ are computed.
(7) *Detector parameter update*: The updated detector parameters x^{k+1} are computed using equation (1).

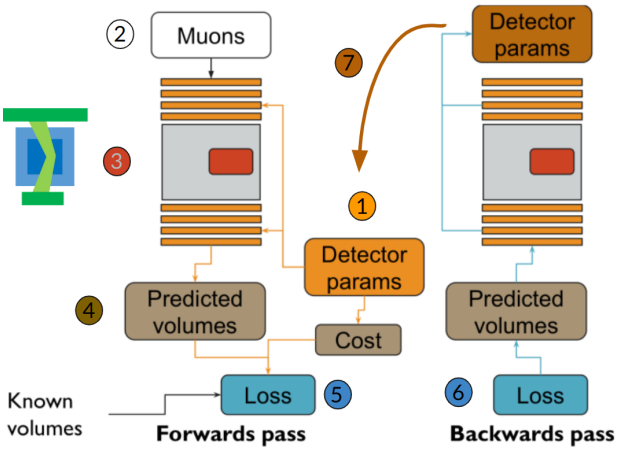

FIGURE 1: Schematic view of TOMOPT workflow within a single epoch, reproduced from [10].

2.1. Passive Volume and Muon Transport

Muons are generated by sampling cosmic muon flux models, namely Guan et al. [11] or Shukla et al. [12]. TOMOPT then simulates the passage of muons through the volume of interest (VOI) and their interaction with matter. The VOI is modeled as a 3D voxelized volume, where each voxel bears the information of the material density and radiation length. An example of VOI is presented in Figure 2. TOMOPT has its own database of radiation length for various materials but the user is also free to add custom radiation lengths and densities. Being in its first phase of development, only multiple Coulomb scattering is taken into account during muon propagation. It is done through the implementation of a step-by-step algorithm where muon displacement and deflection are sampled from the scattering model recommended by the Particle Data Group (PDG) [13].

2.2. Detector Modeling and Hit Recording

Detectors are modeled as infinitely thin horizontal planes and are parameterized by a position x, y, z and spans d_x, d_y for a total of 5 learnable parameters per panel which will get updated at each epoch. Additionally, each panel has a spatial resolution $\sigma_{\max}$ and efficiency $\epsilon_{\max}$ which remain constant throughout the optimization. Muon's positions (hits) are recorded in x, y, z and the effect

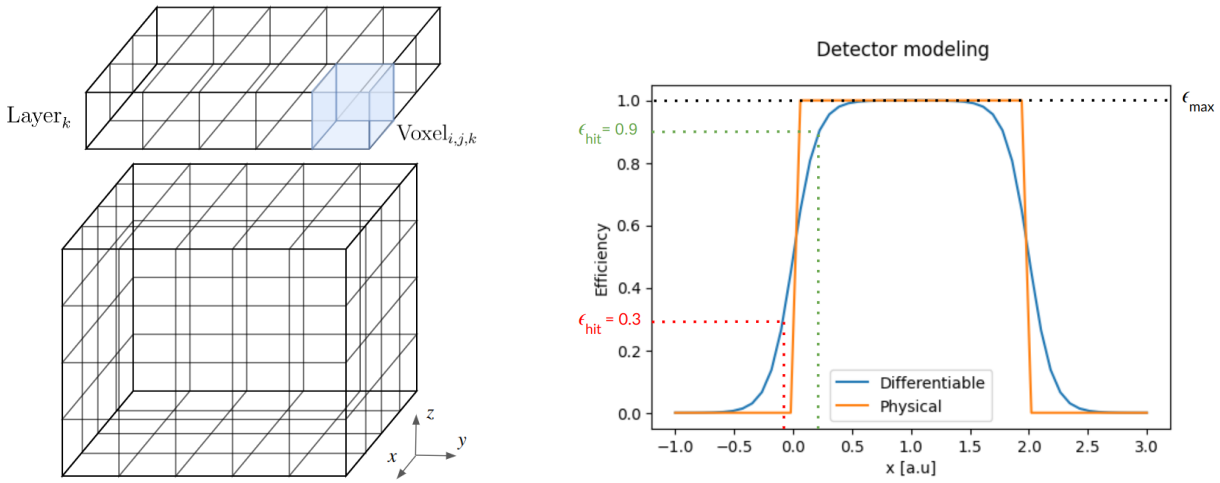

FIGURE 2: Example of voxelized volume of interest (left). Example of detector panel modeling $\mathcal{M}$ with sigmoid function used during optimization (blue) and realistic detector panel modeling (orange) (right).

of spatial resolution is simulated by applying a Gaussian smearing with mean $\mu = 0$ and scaled by the spatial resolution σ_{xy}, as shown in equation (2).

$$xy_{\text{rec}} = xy_{\text{true}} + \mathcal{G}\left(\mu = 0, \sigma = \sigma_{xy}\right), \quad z_{\text{rec}} = z_{\text{true}}, \tag{2}$$

where, in reality, a muon would be either detected or not, TomOpt has to take a different approach in order to preserve the differentiability of the inference chain with respect to the panel parameters. Hits are always recorded by the panels at their current z position, even if the muon is located outside of the panel's active area. However, the effective hit efficiency ϵ_{hit} and resolution σ_{xy} will decrease the further away the hits get from the center of the panel. The model $\mathcal{M}$ used in TomOpt is a pair double-sigmoid function in x and y, as shown in Figure 2. This differentiability trick provides a detection performance close to the actual detector for hits within the panel, and a smooth transition to zero efficiency and resolution outside the panel. In this way, each hit and muon trajectory will be differentiable with respect to all parameters of the panel. As the inference accounts for the efficiencies of the recorded hits, muons outside of the panels will still contribute but way less than a muon detected within the panel's active area.

2.3. *Volume Inference*

Once muon hits have been recorded, they can be combined to reconstruct the incoming and outgoing muon tracks. They are computed analytically through likelihood minimization and are reweighted as a function of the xy resolution of the hits σ_{xy} and their efficiencies ϵ_{hit}. As a consequence, tracks from hits within the active detector region are likely to be properly reconstructed whereas tracks from hits outside the active detector region will be poorly reconstructed and will bear a large uncertainty. Once tracking is performed, various quantities such as zenith, azimuthal and scattering angles can be computed, which will then be used by the inference method.

Volume inference is implemented in such a way that the user is free to choose his own inference method, as long as it respects the differentiability of the loss function concerning detector parameters. The default volume inference method implemented in TomOpt, is based on the Point of Closest Approach [14] and accounts for POCA point uncertainty computed via automatic differentiation. Once the POCA points are computed, they are used to compute an estimate of the radiation length X_0 [m] of a voxel with height δz, thus converting muon-wise data into voxel-wise data.

$$\theta_0^{\text{rms}} = \frac{\theta_{\text{tot.}}^{\text{rms}}}{\sqrt{2}}, \tag{3}$$

$$\theta_0^{\text{rms}} = \frac{0.0136\,\text{GeV}}{p^{\text{rms}}}\sqrt{\frac{\delta z}{X_0 \cos\left(\bar{\theta}^{\text{rms}}\right)}}, \quad \bar{\theta}^{\text{rms}} = \frac{\theta_{\text{in}}^{\text{rms}} + \theta_{\text{out}}^{\text{rms}}}{2}, \tag{4}$$

$$X_0 = \left(\frac{0.0136\,\text{GeV}}{p^{\text{rms}}}\right)^2 \frac{\delta z}{\cos\left(\bar{\theta}^{\text{rms}}\right)} \frac{2}{\theta_{\text{tot.}}^{\text{rms}}}, \tag{5}$$

where $\theta_{\text{tot.}}^{\text{rms}}$ is the root-mean-square (RMS) of the measured scattering angles. In equation (4), $\theta_{\text{in}}^{\text{rms}}$ and $\theta_{\text{out}}^{\text{rms}}$, respectively, refer to the incoming and outgoing muon zenith angles. Where in most POCA algorithms, a given POCA point only contributes to the voxel it is located in, TomOpt extends the POCA points to other voxels using the POCA point uncertainty. A POCA point with a large uncertainty will be extended to other voxels, whereas a POCA with low uncertainty will only contribute to a single voxel.

2.4. Demonstration

The following demonstration was made using 2 sets of 3 detection panels each. The initial configuration is purposely chosen to be suboptimal: the panels' span does not fully cover the volume of interest. Also, panels that are close to each other grant a poor angular resolution, which drives the precision and quality of the inference. A single passive volume was used at each step, but the user is free to optimize the detector using as many passive volumes as he wants. The passive volume used was a block of sand containing a smaller block of lead. The initial efficiency and spatial resolution values are $\epsilon_{\max} = 0.9$ and $\sigma_{\max} = 1\,\text{mm}$. The loss function used is the Mean Squared Error MSE $= \frac{1}{N}\sum_{i=1}^{N}(X_{0,\text{true}} - X_{0,\text{pred}})^2$ computed over every voxel X_0 prediction. The optimization ran for 30 epochs and used 2000 muons per epoch, for a total computing time of 8 min on an 11th Gen Intel(R) Core(TM) i5 CPU.

The initial detector configuration and optimized geometry are presented in Figure 3. Detector panels' spans have been extended such that they now cover the whole passive volume. Throughout optimization, the upper and lower panels' z positions were, respectively, forced to remain within $[1.00, 0.80]$ m and $[0.00, 0.20]$ m. After optimization, outer panels are placed as far as possible from each other, which grants an almost optimal angular resolution. The most optimal configuration would be the one where middle panels are placed as close as possible to the volume of interest. After further investigations, it appeared that the gain in angular resolution provided by such configuration remains smaller than the loss gradient fluctuations from one epoch to another.

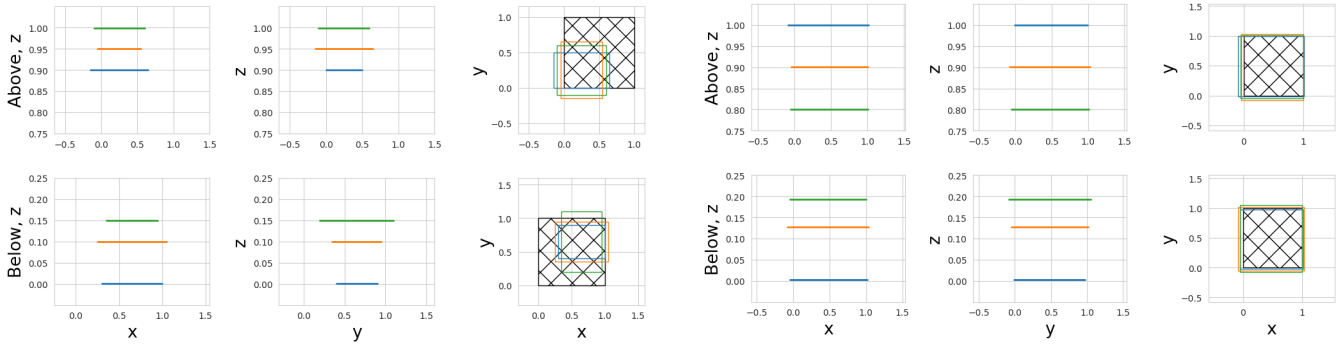

FIGURE 3: Initial detector configuration (left) and optimized detector configuration (right). Detector panels are represented as colored lines and squares. The volume of interest horizontal profile is represented by the black hatched squares.

3. MUON MOMENTUM MEASUREMENT

The in-build X_0 inference previously described in Section 2.3 assumes that the muon momentum p is known with a 100% precision (equation (5)), which is quite unrealistic in the context of the MST experiment. The following section describes how a momentum measurement module based on muon deflection measurement in a known material can be used for momentum inference [15, 16]. The performance of the machine learning based inference method, similar to [17], will be discussed and compared to the standard regression approach.

3.1. Muon Momentum Measurement Module

The muon momentum measurement module in the study consists of an alternation of 10 detector panels and 10 cm thick scattering material as illustrated in Figure 4. The latter is made of lead, a material chosen for its high scattering power (e.g., 4 GeV muons traversing 10 cm of lead have a scattering angle and displacement distribution with respective standard deviation $\sigma_{\Delta\theta} = 1.16^\circ$ and $\sigma_{\Delta x} = 0.08\,\text{cm}$). Detector panels are separated by a 10 cm gap from each other and record the true muon's x and y position. The effect of spatial resolution on the reconstructed hits is then simulated by applying a Gaussian smearing with the desired spatial resolution σ.

3.2. Tracks Reconstruction

Once muon hits have been recorded, they are used to reconstruct muon tracks throughout the apparatus. Each two consecutive position measurements constitute a track, which is then used to compute scattering and deflection angles. Additionally, a fit of all reconstructed hits is performed. Variables relevant to the momentum inference are the following, as represented in Figure 4:

(i) Deflection angle $\theta_{\text{deflection}}$: the angle between 2 consecutive tracks. These values bear the measurement of muon displacement dx: $\theta_{\text{deflection}} = \tan(dx/d_r)$ where $d_r = d/\cos(\theta)$ with θ being the muon zenith angle and d being the lead slab thickness.

(ii) Scattering angle $\Delta\theta$: the angle between muon trajectory before and after interacting with the lead block. The root mean square θ_0^{RMS} of the $\Delta\theta$ distribution is computed as $\theta_0^{\text{RMS}} = \sqrt{\frac{1}{N}\sum_{i}^{N=4}\Delta\theta^2}$.

(iii) Fit residuals r: The residuals of the linear track fit performed over all recorded hits.

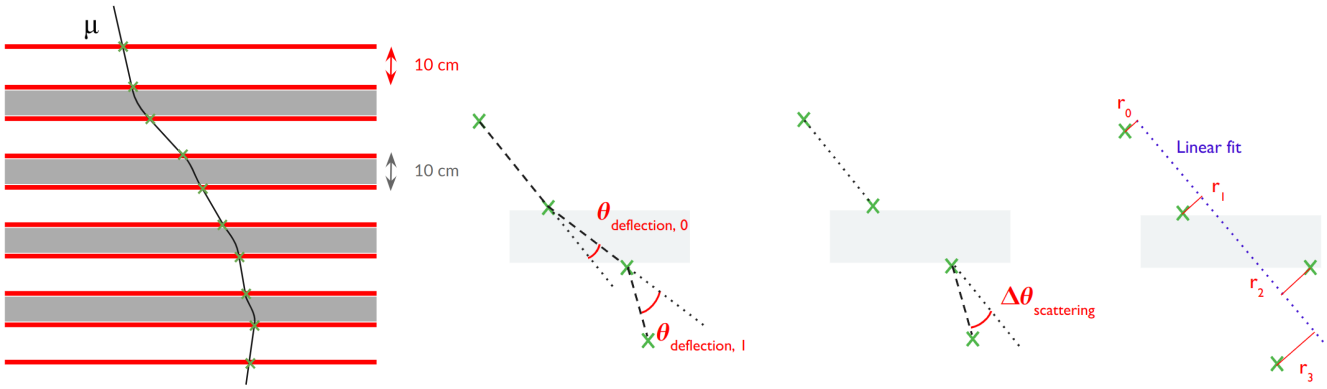

FIGURE 4: Muon momentum measurement module (left). Variables used for momentum inference (right). From left to right, deflection angle $\theta_{\text{deflection}}$, scattering angle $\Delta\theta_{\text{scattering}}$, and linear fit residuals r. Green crosses and the grey rectangle, respectively, represent the reconstructed hits and the lead slab.

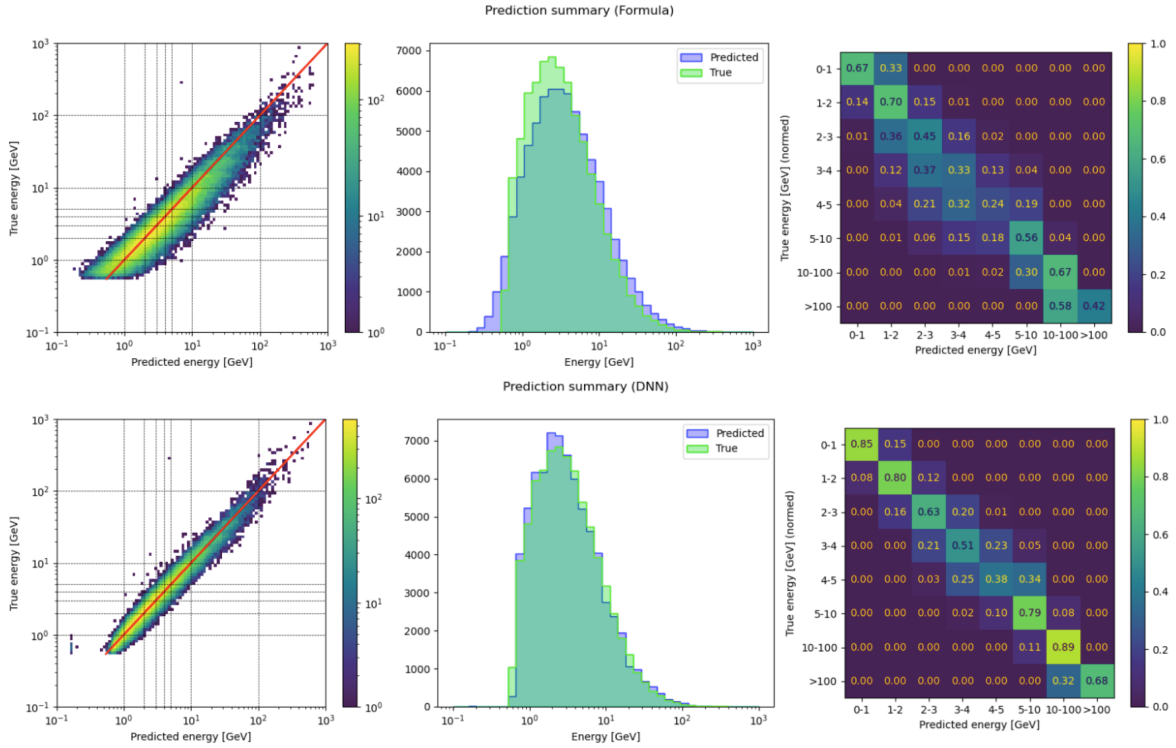

FIGURE 5: Standard regression (up) and DNN regression (down) for a perfect detector spatial resolution. From left to right are shown the predicted muon energy as a function of its true energy, the predicted and true muon energy distributions and the predicted energy confusion matrix.

3.3. *Inference*

3.3.1. *Standard Regression*

The scattering angle variable $\Delta\theta$ can be used to infer the muon momentum, using equation (6). Having 4 passive layers allows us to measure 4 values of $\Delta\theta$, which are used to compute θ_0^{RMS}. When detector panels have a perfect spatial resolution, this inference method performs reasonably well and almost recovers the cosmic muon momentum distribution, as shown in Figure 5. When a detector spatial resolution res $= 1$ mm is simulated, reconstructed hits become $x_{\text{rec}} = x_{\text{true}} + \mathcal{G}(\mu = 0, \sigma = \text{res})$. This resolution is propagated to the angles used in the inference, which implicitly creates an upper limit in the momentum prediction. While in the ideal case a high-energy muon creates an almost linear track with scattering angles close to zero, in a realistic setup, the angle measurement is affected by the angular resolution which creates the illusion of a lower energy muon scatterings. This effect is clearly visible in Figure 6, where it is impossible to predict muon momentum beyond 10 GeV. This naive approach is compared to

a machine-learning-based regression method in the following section.

$$p = \frac{13.6\,\text{MeV}}{\theta_0^{\text{RMS}}}\sqrt{\frac{d}{X_0\cos(\theta)}}. \tag{6}$$

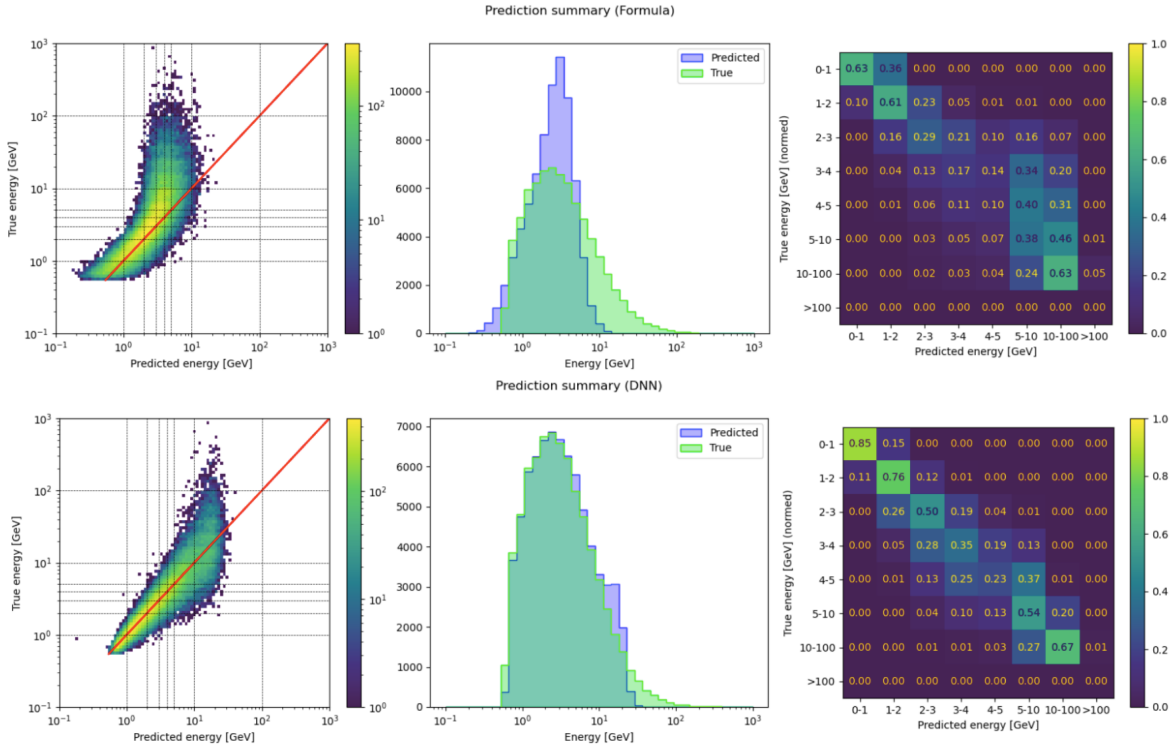

FIGURE 6: Standard regression (up) and DNN regression (down) for a 1 mm detector spatial resolution. From left to right are shown the predicted muon energy as a function of its true energy, the predicted and true muon energy distributions and the predicted energy confusion matrix.

3.3.2. NN-Based Classification

As opposed to the standard regression approach described above, Deep Neural Networks (DNNs) offer more flexibility and allow us to use multiple inputs: 9 deflection angles $\theta_{\text{deflection}}$, 4 scattering angles $\Delta\theta$, 10 fit residuals r, and the squared sum of the fit residuals $r_{\text{sum}} = \sum_{i=1}^{10} r_i^2$. The DNN used in this analysis is made of three 64 neuron layers connected by ReLU activation functions and utilizes the Adam optimizer [18]. The training was performed over 600 k events, and validation with 75 k events. Because of the wide range of momenta, the regression is done on $\log(E)$. Figures 5 and 6 clearly demonstrate that there is more information to extract from the hits than the scattering angles alone, something that the DNN is able to leverage.

4. CONCLUSION

We demonstrated the capabilities of TomOpt in a generic use case of muon scattering tomography. Using a loss function evaluating the performance of the radiation length predictions, we were able to optimize the initial poorly designed detector configuration into one that shows almost optimal performance in terms of angular resolution. In this example, no additional constraints were included in the loss function; however, alternative optimization could be performed to minimize the detector cost. Additionally, as a simplified simulator of muon physics, the TomOpt framework can be exploited for developing and testing advanced inference algorithms.

We also showed how a muon momentum measurement module can be used to perform momentum classification. In this context, it was shown that the use of Deep Neural Networks can out-perform standard regression approaches, by extracting more relevant information from the detector hits than the scattering angles alone.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

ACKNOWLEDGMENTS

This work was partially supported by the EU Horizon 2020 Research and Innovation Programme under grant agreement No. 101021812 (“SilentBorder”) and by the Fonds de la Recherche Scientifique - FNRS under Grants No. T.0099.19 and J.0070.21.

References

[1] Radiographic imaging with cosmic-ray muons, Borozdin, Konstantin N. and others, Nature 422, 10.1038/422277a (2003).
[2] Atmospheric muons as an imaging tool, L. Bonechi, R. D’Alessandro, A. Giammanco, Reviews in Physics, https://doi.org/10.1016\%2Fj.revip.2020.100038, (2020).
[3] Material identification in nuclear waste drums150 using muon scattering tomography and multivariate analysis, M. Weekes, A. Alrheli, D. Barker, D. Kikola, A. K. Kopp, M. Mhaidra, J. Stowell, L. Thompson, and J. J. Velthuis, Journal of Instrumentation 16 no. 05, (2021) P05007.
[4] Cosmic-Ray Tomography for Border Security, S. Barnes, A. Georgadze, A. Giammanco, M. Kiisk, V. A. Kudryavtsev, M. Lagrange, and O. L. Pinto, Instruments 7 no. 1, (2023) 13.
[5] Horizontal cosmic ray muon radiography for imaging nuclear threats, C. L. Morris, J. Bacon, K. Borozdin, J. Fabritius, H. Miyadera, J. Perry, and T. Sugita, Nucl. Instr. Meth B 330 (2014)42–46.
[6] Non-destructive testing of industrial equipment using muon radiography, P. Martinez Ruiz del Arbol, P. Gomez Garcia, C. Diez Gonzalez, and A. Orio Alonso, Philosophical Transactions of the Royal Society A 377 no. 2137, (2019) 20180054.
[7] TomOpt: Differential optimisation for task- and constraint-aware design of particle detectors in the context of muon tomography, Strong, Giles C. and others, 2309.14027 (2023).
[8] Mémoire sur le problème d’analyse relatif à l’équilibre des plaques élastiques encastrées, Hadamard, J., http://books.google.com.au/books?id=BTEPAAAAIAAJ
[9] PyTorch: An Imperative Style, High-Performance Deep Learning Library, Paszke, Adam and others, Advances in Neural Information Processing Systems 32, (2019).
[10] Toward the end-to-end optimization of particle physics instruments with differentiable programming, T. Dorigo et al., Reviews in Physics, https://doi.org/10.1016/j.revip.2023.100085 (2023).
[11] A parametrization of the cosmic-ray muon flux at sea-level, Guan et al., 1509.06176 (2015).
[12] Energy and angular distributions of atmospheric muons at the Earth, Shukla et al., Int. J. Mod. Phys. A, 1606.06907 (2018).
[13] Particle Data Group Collaboration, R. L. Workman and Others, Review of Particle Physics, PTEP 2022 (2022) 083C01.
[14] Image reconstruction and material Z discrimination via cosmic ray muon radiography, Schultz et al., Nucl. Instrum. Methods Phys. Res. A, 0168-9002 (2004).
[15] Muography of different structures using muon scattering and absorption algorithms, S. Vanini et al., Phil. Trans. R. Soc. A.3772018005120180051 (2018), https://doi.org/10.1098/rsta.2018.0051
[16] A plastic scintillator-based muon tomography system with an integrated muon spectrometer, V. Anghel et al., Nuclear Instruments and Methods in Physics Research 798 (2015), https://doi.org/10.1016/j.nima.2015.06.054
[17] Development of Machine Learning-Assisted Spectra Analyzer for the NEWCUT Muon Spectrometer, L. Oláh et al., Advanced Instrumentation in Science, vol. 2022, Apr. 2022, https://doi.org/10.31526/jais.2022.264
[18] Adam: A Method for Stochastic Optimization, Diederik P. Kingma, Jimmy Ba, 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings, http://arxiv.org/abs/1412.6980