particles

IMPACT FACTOR 1.7

CITESCORE 3.2

Article

Machine Learning Approach to Shield Optimization at Muon Collider

Luca Castelli

Special Issue

Selected Papers from the 4th MODE Workshop on Differentiable Programming for Experiment Design

Edited by
Prof. Dr. Tommaso Dorigo, Dr. Roberto Ruiz de Austri Bazan, Prof. Dr. Jose Salt and
Dr. Pietro Vischia

MDPI

https://doi.org/10.3390/particles8010025 particles

MDPI

Article

Machine Learning Approach to Shield Optimization at Muon Collider

Luca Castelli [1,2]

[1] Dipartimento di Fisica, Sapienza Università di Roma, 00185 Rome, Italy; luca.castelli@uniroma1.it
[2] INFN Sezione di Roma, 00185 Rome, Italy

Abstract: Muon collisions are considered a promising means for exploring the energy frontier, leading to a detailed study of the possible feasibility challenges. Beam intensities of the order of 10^{12} muons per bunch are needed to achieve the necessary luminosity, generating a high flux of secondary and tertiary particles from muons decay that reach both the machine elements and the detector region. To limit the impact of this background on the physics performance, tungsten shieldings have been studied. A machine learning-based approach to the geometry optimization of these shieldings will be discussed.

Keywords: Muon Collider; MDI; Machine-Learning

check for updates

Academic Editors: Tommaso Dorigo, Roberto Ruiz de Austri Bazan, Jose Salt and Pietro Vischia

Received: 27 January 2025
Revised: 19 February 2025
Accepted: 26 February 2025
Published: 3 March 2025

Citation: Castelli, L. Machine Learning Approach to Shield Optimization at Muon Collider. *Particles* **2025**, *8*, 25. https://doi.org/10.3390/particles8010025

1. Introduction

The Muon Collider is a proposed next-generation facility for high-energy physics [1]. Muons, as fundamental particles in the lepton family, are particularly suited for collider applications due to their relatively large mass, about 200 times heavier than electrons, combined with their point-like nature. This makes muons less prone to radiative energy losses compared to electrons when accelerated to high energies, allowing the building of a circular facility to explore the multi-TeV center-of-mass energy region. However, one of the primary challenges associated with the Muon Collider arises from the limited lifetime, even in a highly boosted regime of muons. They decay in the beam pipe, and the resulting decay products interact with the collider components, generating a flux of photons, neutrons, and charged particles that reach the detector, known as Beam-Induced Background (BIB). Effective mitigation strategies are essential to reduce BIB and ensure the desired detector performance and measurement accuracy. To this purpose, a tungsten cone-shaped shield (nozzle) has been designed inside the detector to absorb the incoming particles and reduce their energy. The original geometry of this object was proposed by MAP [2] for a $\sqrt{s} = 1.5$ TeV machine. However, the current project aims to reach $\sqrt{s} = 10$ TeV with a first stage operating at $\sqrt{s} = 3$ TeV; therefore, optimizations for these center of mass energies are required. This contribution aims to describe the activities undertaken and ongoing concerning the nozzle's geometry optimization for the 3 TeV Muon Collider exploiting machine learning (ML) algorithms.

2. Beam-Induced Background

At $\sqrt{s} = 3$ TeV, the muon decay rate is expected to reach 2.34×10^5 decays per meter per beam per bunch crossing, given 2.2×10^{12} particles per beam. The electrons and positrons produced by these decays interact with the machine elements, generating electromagnetic showers and photo-nuclear reactions. As a result, a substantial flux of secondary particles reaches the detector, estimated to be on the order of 10^8 particles per bunch crossing. These particles—primarily photons, neutrons, electrons, and positrons—are widely distributed in time relative to the bunch crossing. This Beam-Induced Background presents a significant challenge for the detector, particularly for the vertex detector near the Interaction Point (IP), since it causes high occupancy and background noise, preventing precise track reconstruction and degrading overall physics performance.

To mitigate the impact of BIB, a nozzle design based on the MAP configuration has been developed. The nozzle features a cone-shaped tungsten core extending from a few centimeters from the IP to the edge of the detector, forming a 6-m-long structure. Near the IP, the nozzle's angular aperture is 10° to address the critical impact of BIB on the vertex detector, narrowing to 5° at a distance of 1 meter, since the BIB intensity decreases, requiring less shielding material and allowing for improved detector acceptance. To further reduce neutron production from interactions within the tungsten core, the nozzle is coated with borated polyethylene. The detector layout and nozzle geometry are illustrated in Figure 1.

The time of arrival of BIB particles at the detector surface is highly dispersed relative to the bunch crossing. A time-based rejection strategy is employed, using a readout window from -1 ns to $+15$ ns around the bunch crossing to significantly reduce off-time background events. The composition of the BIB before and after time discrimination is detailed in Table 1, where the first column shows the initial particle flux and the second column the flux remaining after time rejection. Additionally, Figure 2 displays the full-time distribution and energy spectra for each particle species based on the MAP configuration.

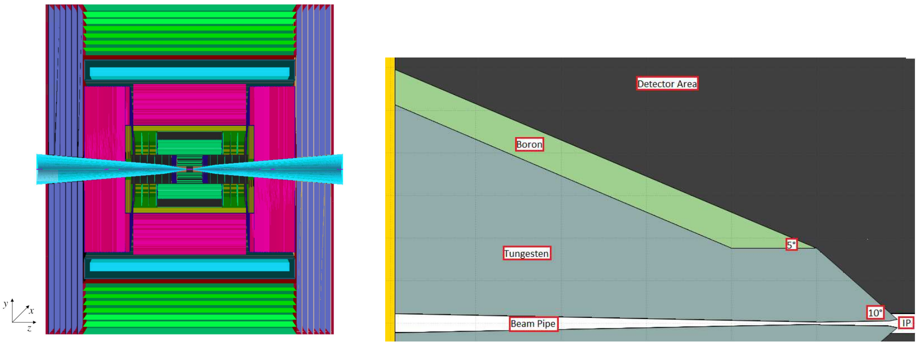

Figure 1. (**Left**) The detector used for $\sqrt{s} = 3$ TeV Muon Collider studies [3]. The nozzles are shown in cyan. (**Right**) Section of left nozzle.

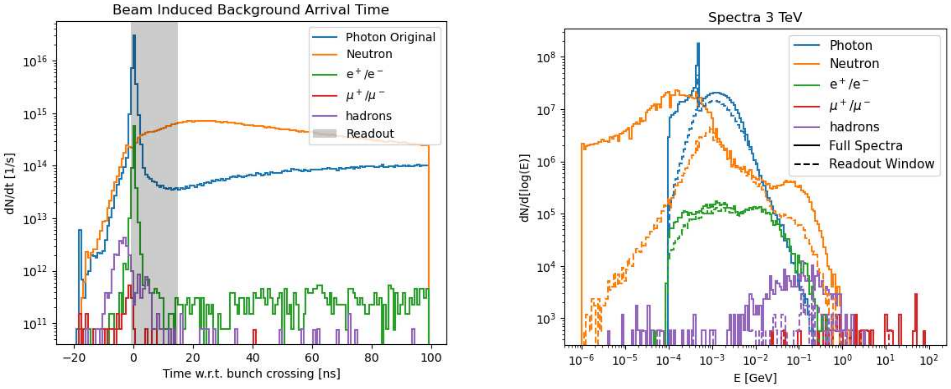

Figure 2. (**Left**) Arrival times of the main BIB components at the detector relative to the bunch crossing. The light gray band between -1 and 15 ns indicates the assumed readout window for the detector. (**Right**) Energy spectra of BIB particles reaching the detector. The solid line represents all particles, while the dashed line includes only those within the readout window [3].

Table 1. Particle species per Bunch Crossing contributing to the Beam-Induced Background at $\sqrt{s} = 3$ TeV with the MAP. The first column considers the whole flux entering the detector. The second column considers the application on a Readout Window of $[-1, 15]$ ns with respect to bunch crossing.

Particle Species	**Flux**	**Flux − R.W.**
Photons	8.48×10^7	4.22×10^7
Neutrons	7.35×10^7	7.18×10^6
Electrons/Positrons	7.00×10^5	5.75×10^5
Hadrons	2.43×10^4	4.91×10^3
Muons	2.42×10^3	6.76×10^2

The simulation of the BIB, which led to the plots shown in Figure 2, has been achieved by implementing the design in Figure 1 (right) together with the beam pipe and all the machine elements in FLUKA [4,5], exploiting a two-step simulation process:

- **Step 1**: Muon decays along the beam pipe are simulated, and the interaction with the environment is propagated up to the detector area. There, the incoming particles are killed in the simulation, and their kinematic information is stored.
- **Step 2**: Particles arriving on the detector surface are propagated inside the nozzles. The flux surviving the interactions inside the nozzles and reaching the detector is stored as the Beam-Induced Background.

This approach allows saving a significant amount of time since it is not necessary to repeat the first step each time the nozzles geometry changes, considering that only the second step of the simulation takes on average 4 days for a 1.6% of the expected decays for a bunch crossing, considering only one beam.

3. Nozzle Geometry Optimization

As mentioned in the Introduction, the nozzle geometry was originally designed for a $\sqrt{s} = 1.5$ TeV machine. Therefore, it must be adapted to the energy stages currently envisioned for the Muon Collider, namely 3 and 10 TeV in the center of mass. The primary objective of this adaptation is to minimize the Beam-Induced Background in the detector, thereby mitigating its impact on event reconstruction. Additionally, reducing the nozzle size offers several advantages, particularly cost-wise. Given that tungsten has a density of approximately 19 g/cm^3, depending on the alloy, even a slight reduction in nozzle volume can lead to a substantial decrease in material requirements. This reduction also lowers the overall weight that needs to be supported, simplifying the design of the nozzle's supporting structure. Finally, decreasing the nozzle volume creates more space for detectors, thereby enhancing the acceptance of the entire experiment. This work discusses the optimization achieved and the studies ongoing for the $\sqrt{s} = 3$ TeV stage.

3.1. FLUKA Simulation

The first step in the optimization process focused on identifying how the geometry influences the resulting Beam-Induced Background. To this end, several FLUKA simulations were conducted, varying three key geometrical parameters: the angle of the nozzle tip, θ_{nozzle}, the position along the beam axis where the aperture changes, z_{change}, and the base radius, r_{base}. Figure 3 illustrates these parameters.

The simulations used a sample size corresponding to 1.6% of the expected decays, approximately 2×10^5 muons (high statistics). This sample size ensures a negligible statistical error in the flux of particles entering the detector surface. Figure 4 on the right plots presents the composition of the BIB for two nozzle designs compared to the original configuration. For photons, neutrons, and e^+/e^-, the statistical error is negligible.

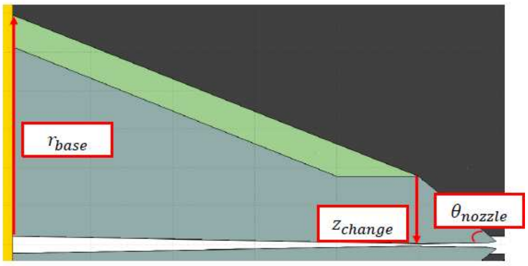

Figure 3. Geometrical parameters changed during the first studies on the Beam-Induced Background.

The results indicated that careful optimization of the last meter of the nozzle is critical, whereas material savings can be achieved further from the IP, as can be seen in Figure 4. Two different designs are proposed. On the top, a tentative with a larger base radius and a tinier tip with respect to the original design results in an increase of about a factor in the Beam-Induced Background (BIB). Conversely, the bottom plots show a reduction of photon and e^+/e^- of the same factor achieved by increasing the amount of material in the tip but reducing the base radius. However, due to the sensitivity of the geometry—where even small variations result in significant increases in BIB—no definitive guidelines for modifications could be established.

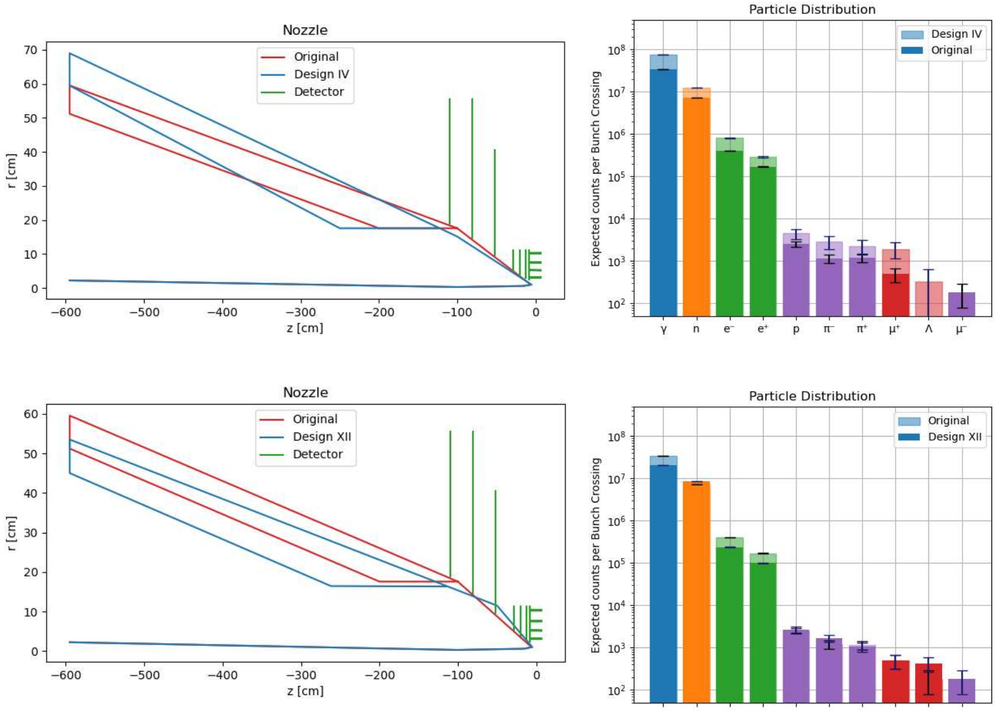

Figure 4. (**Left**) Nozzle design compared to the original one. The plots also include the tracking system as a reference. (**Right**) BIB particle distribution after the application of the Read-Out window. Reducing the size of the tip and increasing the nozzle base radius resulted in the increase of BIB (**top**). Increasing the size of the tip while reducing the base radius reduced the BIB (**bottom**). The latter design cannot be considered for optimization since it overlaps with the detector.

Additionally, an assessment of the borated polyethylene coating was made after analyzing the neutron fluence distribution shown in Figure 5. Observations revealed that the highest concentration of particles does not occur on the external surface. Consequently, it was proposed to relocate the coating internally, placing a layer of tungsten on the outside. This configuration would help block photons emitted during neutron capture in the boron.

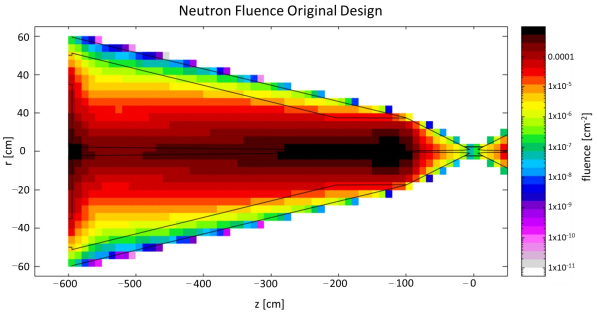

Figure 5. Neutron fluence in the nozzle. The highest concentration of particle is located in the internal part.

3.2. Machine Learning Studies

Manual optimization achieved by iterating FLUKA simulations is not the most effective strategy for finalizing the nozzle design due to the extensive computational time required for a single simulation. To address this challenge, an alternative approach leveraging Machine Learning (ML) algorithms has been adopted.

3.2.1. Data Preparation

Building on the insights from the studies summarized in the previous section, a set of geometrical parameters has been identified to redefine the nozzle geometry. Specifically, the borated polyethylene has been incorporated inside the tungsten, and a step has been introduced to reduce material usage in areas where it is not required. The defined parameters are listed below, and the new nozzle design is illustrated in Figure 6:

- r_{base}: Radius of the nozzle at the end of the final focusing stage.
- r_{boron}: Outer radius of the borated polyethylene section of the nozzle, expressed as a fraction of r_{base}.
- z_{step}: Position along the beam axis where the nozzle size is reduced.
- r_{step}: Magnitude of the step, defined as a fraction of the nozzle's radial size at position z_{step}.
- z_{change}: Position along the beam axis where the angular aperture of the nozzle changes.
- θ_{tip}: Angular aperture at the tip of the nozzle.
- z_{tip}: Position along the beam axis of the nozzle tip.
- r_{tip}: Radial coordinate of the nozzle tip.

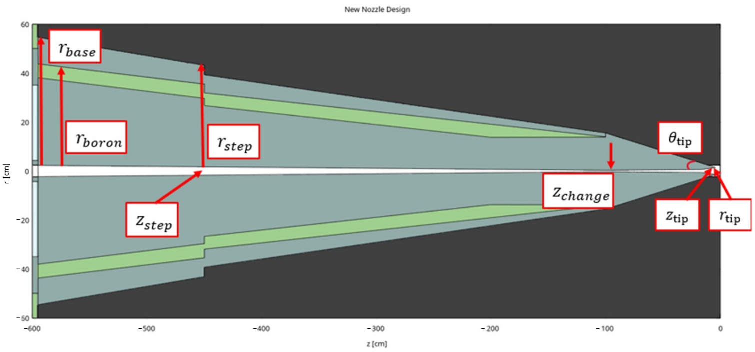

Figure 6. Geometry layout of the nozzle considered for the Machine-Learning studies. In figure are reported the parameters described in text section.

The first step in this study was to construct a dataset for training the ML algorithms (model details in Section 3.2.2). Given the objective of significantly reducing computational time, the simulated decays in each simulation were limited to only 0.02% of a bunch crossing (low statistics). This approach enabled the creation of a dataset comprising 1.3×10^4 different geometries and the related BIB flux simulated with FLUKA within a reasonable timeframe while maintaining adequate statistical accuracy. 75% of the dataset has been used to train the models, and the remaining 25% formed the test sample to evaluate the performance (more in Section 3.2.2).

Figure 7 compares the photon energy spectrum of the BIB obtained using the original nozzle design with 1.6% decays simulated (blue) to that of a design studied with ML algorithms and 0.02% decays simulated (black). Both simulations have been performed with FLUKA. The difference in the error bar size is noticeable primarily in the higher-energy region of the spectrum, whereas the relative error in the most densely populated energy range remains around 10^{-4}.

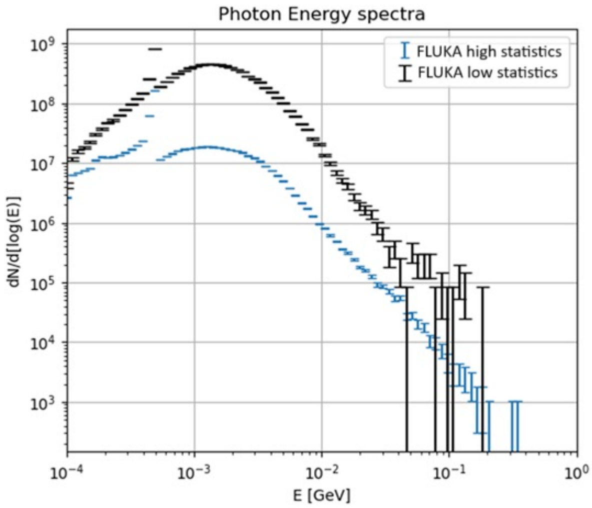

Figure 7. Photon energy spectrum obtained with original nozzle design and 1.6% of expected decay simulated (blue) compared to one of the low statistics simulations used for the ML studies (black). Both spectra have been obtained with FLUKA simulation. The absolute number of photons differs because different configurations are considered.

3.2.2. Model Evaluation

The second step of this study aimed to develop a model capable of parametrizing the particle flux entering the detector surface as a function of the geometrical parameters listed above. The integrated flux alone is not an ideal metric, as different particle types affect various detector components differently [1], and their flux does not vary coherently with changes in nozzle geometry. However, evaluating the effect of each geometry on individual sub-detectors would significantly increase the simulation time, counteracting the primary goal of this study.

To address this challenge, Deep Neural Networks [6], Random Forest [7], and XGBoost [8] regressors were evaluated. Data transformation techniques were also explored, given that the parameters span different phase spaces and small variations in some parameters can lead to large changes in flux. This behavior can negatively affect the performance of ML algorithms. The trained ML model predicts the particle flux as a function of the nozzle's geometrical parameter:

$$Flux_{pred} = f_{\mathrm{ML}}(r_{base}, r_{boron}, z_{step}, r_{step}, z_{change}, \theta_{tip}, z_{tip}, r_{tip}). \tag{1}$$

Here, f_{ML} represents the trained ML model, which has learned the mapping between the nozzle parameters and the resulting flux. The performance of the model is then assessed by comparing the predictions for each parameter configuration in the test dataset with the respective FLUKA-simulated flux, $Flux_{sim}$:

$$\Delta[\%] = \frac{Flux_{sim} - Flux_{pred}}{Flux_{sim}} \times 100 \tag{2}$$

Among the tested models, XGBoost paired with the Standard Scaler [9] applied to the dataset features (i.e., the parameters) produced a Δ distribution peaked closest to 0% with the minimum width. The iper-parameters of the model are summarized in Table 2. The performance was quantified by fitting the Δ distribution on the test dataset with a Gaussian function, yielding an expected value of $\overline{\Delta} = -0.12\%$ with a standard deviation of $\sigma = 5.24\%$, as shown in Figure 8. This demonstrated that the XGBoost model could reliably serve as a surrogate for the FLUKA simulation while requiring only 200 ms for training and evaluation, making it highly efficient.

Table 2. Summary of the hyperparameters used for the XGBoost model that achieved the best performance in predicting the BIB flux based on the geometrical parameters.

Iper-Parameter	Value
Loss function	MSE
Estimators	100
Max Depth	6
Learning rate	0.3

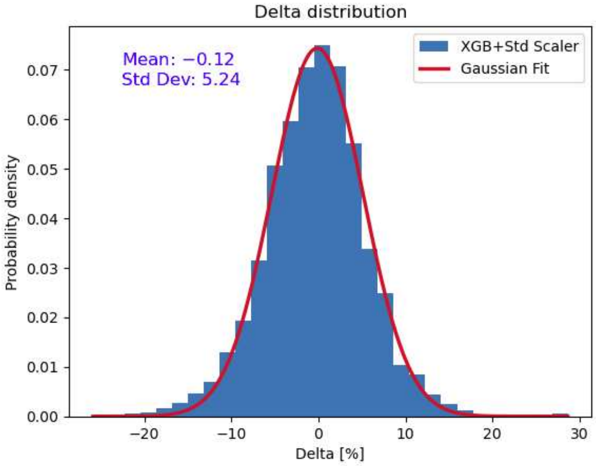

Figure 8. Gaussian fit of the Δ distribution achieved with the XGBoost model.

3.2.3. Results

Using the trained model, a pseudo-dataset was generated to identify parameter sets that simultaneously minimized the Beam-Induced Background and the total volume of the nozzle. The optimized geometry, compared to the original design (Figure 9), was subsequently tested using high-statistics FLUKA simulations.

The new design reduced the photon flux by 34% and the e^+/e^- flux by 27% at the cost of an 86% increase in the neutron flux after the application of the readout window. This trade-off is considered favorable because the primary challenge for event reconstruction is the occupancy in the tracking system, which is primarily affected by the e^+/e^- flux and is improved with the optimized nozzle. Furthermore, the hadronic calorimeter, which is impacted by the neutron flux, is located farther from the interaction point, and the neutron flux is distributed over a larger surface area, mitigating its overall impact. Figure 10 shows the energy spectra of the Beam-Induced Background obtained from high-statistics FLUKA simulations, comparing the original nozzle geometry with the optimized design.

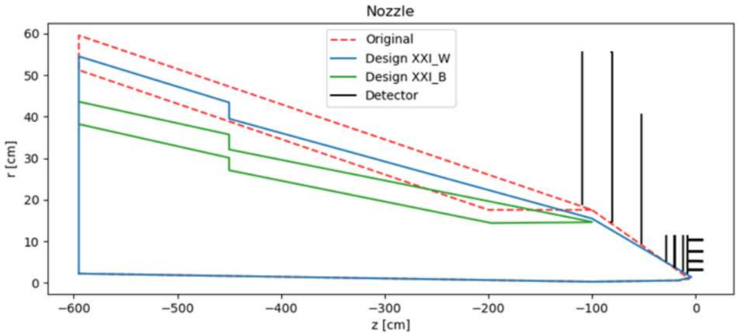

Figure 9. Comparison of the optimized nozzle geometry (blue) featuring the repositioned borated polyethylene (green) with the original design (dashed red). The vertex and tracker detector layers are also depicted (black).

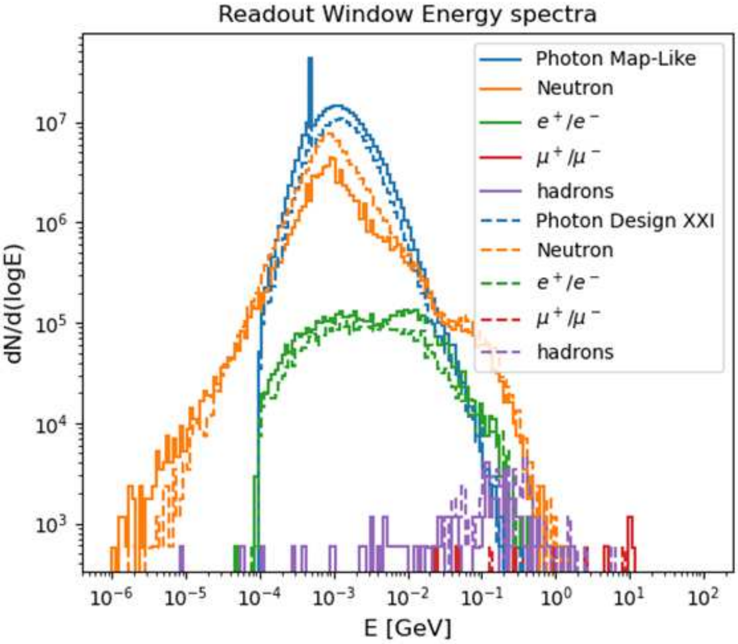

Figure 10. Energy spectrum of the Beam-Induced Background at $\sqrt{s} = 3$ TeV obtained with high-statics FLUKA simulation. The plot compares the optimized geometry reported in Figure 9 (dashed) to the original design (solid). The results are obtained after the application of a readout window of $[-1, +15]$ ns with respect to the bunch crossing.

4. Further Developments

Although an initial optimization of the nozzle geometry has been achieved, more advanced ML methods may enhance the design further. The approach described in the previous section has its limitations, as the pseudo-dataset must be manually analyzed to identify an optimal configuration. To address this constraint, two parallel approaches are currently under investigation, in collaboration with MODE [10].

The first approach involves implementing a Bayesian Optimization Loop [11], which iteratively performs FLUKA simulations based on informed predictions of optimal parameters. The second approach leverages a Deep Neural Network (DNN) to approximate the XGBoost model and, by extension, the FLUKA simulations. The goal of this approach is to obtain a differentiable algorithm (the DNN), enabling the application of the Gradient Descent method [12] to analytically determine the configuration that minimizes the chosen metric.

Although the Bayesian optimization method is still under study, the DNN-based approach has already been tested. However, it did not yield an improved algorithm because the metric used in the study was the integrated Beam-Induced Background flux in the detector. The optimization process led to the selection of the largest possible geometry within the given parameter space, which was suboptimal.

To effectively exploit this method, a more nuanced metric M is required, defined as:

$$M = a \cdot \gamma + b \cdot n + c \cdot e + d \cdot V \tag{3}$$

Here:

- γ, n, e: the total fluxes of photons, neutrons, and electrons, respectively, in the detector;
- V: the total volume of the nozzle;
- a, b, c: parameters reflecting the respective impacts of photons, neutrons, and electrons on the sub-detectors;
- d: a parameter related to the nozzle's total volume.

Ongoing studies aim to evaluate the parameters a, b, c, and d to accurately reflect the trade-offs between particle fluxes and nozzle volume.

5. Conclusions

This paper presented the optimization process of the nozzle's geometry by leveraging a trained XGBoost model to parametrize the Beam-Induced Background flux in the detector as a function of the shielding's geometrical parameters. This approach enabled the exploration of different geometrical configurations without relying on FLUKA simulations, which require several days per run.

A significant improvement was achieved over the original design, simultaneously reducing the BIB flux in the detector and decreasing the nozzle size. This enhancement increases detector acceptance while also lowering material costs.

Further optimization remains possible. The current approach requires manual evaluation of each configuration since the XGBoost model estimates only the BIB flux and does not account for volume reduction. Ongoing work aims to develop a fully automated pipeline that integrates all relevant factors to identify the optimal geometry based on the metric defined in Equation (3).

In addition to geometry optimization, ongoing studies aim to enhance the nozzle's functionality [13]. Since the nozzles reduce the detector acceptance, particles cannot be observed in the very forward region. To address this, the instrumentation of the nozzle base is under study to enable forward muon tagging, which would enhance the detector's physics reach [14,15].

Funding: This research received no external funding

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Accettura, C.; Adams, D.; Agarwal, R.; Ahdida, C.; Aimè, C.; Amapane, N.; Amorim, D.; Andreetto, P.; Anulli, F.; Appleby, R.; et al. Towards a Muon Collider. *Eur. Phys. J. C* **2023**, *83*, 864. [CrossRef]
2. Alexahin, Y.; Gianfelice-Wendt, E.; Kapin, V. Muon Collider lattice concepts. *J. Instrum.* **2018**, *13*, P11002. [CrossRef]

3. Andreetto, P.; Bartosik, N.; Buonincontri, L.; Calzolari, D.; Candelise, V.; Casarsa, M.; Castelli, L.; Chiesa, M.; Colaleo, A.; Molin, G.D.; et al. Higgs Physics at a $\sqrt{s} = 3$ TeV Muon Collider with detailed detector simulation. *arXiv* **2024**, arXiv:2405.19314. [CrossRef]
4. Battistoni, G.; Boehlen, T.; Cerutti, F.; Chin, P.W.; Esposito, L.S.; Fassò, A.; Ferrari, A.; Lechner, A.; Empl, A.; Mairani, A.; et al. Overview of the FLUKA code. *Ann. Nucl. Energy* **2015**, *82*, 10–18. [CrossRef]
5. Ahdida, C.; Bozzato, D.; Calzolari, D.; Cerutti, F.; Charitonidis, N.; Cimmino, A.; Coronetti, A.; D'Alessandro, G.L.; Donadon Servelle, A.; Esposito, L.S.; et al. New Capabilities of the FLUKA Multi-Purpose Code. *Front. Phys.* **2022**, *9*, 788253. [CrossRef]
6. Yi, H.; Shiyu, S.; Xiusheng, D.; Zhigang, C. A study on Deep Neural Networks framework. In Proceedings of the 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Xi'an, China, 3–5 October 2016; pp. 1519–1522. [CrossRef]
7. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
8. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), San Francisco, CA, USA, 13–17 August 2016; ACM: New York, NY, USA, 2016; pp. 785–794. [CrossRef]
9. de Amorim, L.B.; Cavalcanti, G.D.; Cruz, R.M. The choice of scaling technique matters for classification performance. *Appl. Soft Comput.* **2023**, *133*, 109924. [CrossRef]
10. Baydin, A.G.; Cranmer, K.; de Castro Manzano, P.; Delaere, C.; Derkach, D.; Donini, J.; Dorigo, T.; Giammanco, A.; Kieseler, J.; Layer, L.; et al. Toward Machine Learning Optimization of Experimental Design. *Nucl. Phys. News* **2021**, *31*, 25–28. [CrossRef]
11. Frazier, P.I. A Tutorial on Bayesian Optimization. *arXiv* **2018**, arXiv:1807.02811.
12. Ruder, S. An overview of gradient descent optimization algorithms. *arXiv* **2017**, arXiv:1609.04747.
13. Castelli, L.; Collamati, F.; Calzolari, D.; Lucchesi, D. Status of 3 TeV MDI Studies. Available online: https://agenda.infn.it/event/44387/contributions/250843/attachments/129104/191446/rd_mucoll24.pdf (accessed on 27 January 2025).
14. Li, P.; Liu, Z.; Lyu, K.F. Higgs boson width and couplings at high energy muon colliders with forward muon detection. *Phys. Rev. D* **2024**, *109*, 073009. [CrossRef]
15. Ruhdorfer, M.; Salvioni, E.; Wulzer, A. Invisible Higgs boson decay from forward muons at a muon collider. *Phys. Rev. D* **2023**, *107*, 095038. [CrossRef]