







Status of $\sqrt{s} = 3 TeV$ MDI studies

L. Castelli on behalf of IMCC



Outline

• 3 *TeV* MDI

- MAP design
- FLUKA simulation
- Forward Muon Study
 - Goals
 - Simulation and results

Machine Learning for Nozzle Optimization

- Low statistics approach
- XGBoost Results
- Metric problem
- Gradient Descent approach
- Final results



3 TeV MDI

•MAP nozzle design:

1) 10° closest to the IP

2) 5° starting from $z = 100 \ cm$

MAP design[1] with mixed function FF quadrupoles (Cyan)





BIB simulation with FLUKA



- Generated one beam of μ^+ decays within **55** *m* from the Interaction Point
- Energy threshold for particles production fixed at 100 keV
- Particles which arrives to the nozzles are scored
- Propagation through the Nozzles
- Particles who exit the nozzle and enters the detector area are scored
- $\sim 1.6\%$ of one BIB event (i.e. bunch crossing) considering

only 1 beam \rightarrow 4 days per simulation



Detecting Forward Muons

- Instrumenting the nozzle:
 - Small detector
 - High dose from BIB
- Analysis approach:
 - Three scoring layers implemented in FLUKA
 - Simulation of Forward Muons and BIB
 - Identification of Forward Muons candidate



- The goal is to **evaluate**:
 - % forward muon tagged
 - # fake forward muon from BIB



Detecting Forward Muons

- $\mu^+\mu^- \rightarrow ZZ + \mu^+\mu^- \rightarrow H + \mu^+\mu^- \rightarrow$ $W^+W^- + \mu^+\mu^-$
- Readout window ±100 ps w.r.t. bunch crossing
- Rough tracking of muons in layers (100% efficiency)
- No fake muons from BIB reconstructed
- Energy Measurement achievable only by instrumenting the cave (?)







Nozzle Geometry Optimization

Goal:

- Reduced the BIB flux entering the detector area
- Maximizing the detector acceptance

Approaches:

- Manual tuning with high statistics simulation
- Many low statistics simulation to train
 Machine Learning algorithms
- Bayesian optimization iterating *medium* statistics simulation





Machine Learning Approach

Data preparation

- Addressed 9 parameters to define nozzle geometry
- Performed $2.0 \cdot 10^4$ low statistic (*0.02%* of b.c.) FLUKA simulation
- Trained XGBoost Regressor (80% of simulation):
 - Features → geometrical parameters
 - Target → BIB flux entering the detector (RW applied)





Machine Learning Result

- Model Evaluation
 - $\Delta[\%] = \frac{Flux_{true} Flux_{predicted}}{Flux_{true}} * 100$
 - $Flux_{true} \rightarrow$ Flux obtained from FLUKA simulation
 - Flux_{predicted} → Flux computed by XGBoost Regressor on test dataset (20%)
 - Gaussian fit of Δ distribution results in: $\overline{\Delta} = -0.01\%$, $\sigma = 7.05\%$

Optimization

- Generated a pseudo-dataset with $O(10^6)$ different configurations with trained XGBoost
- Manually identified an optimal configuration \rightarrow **Done**





Optimized Geometry

- Main features:
 - Base radius reduced
 - Nozzle body further reduced starting at 450 cm from the IP
 - Borated polyethylene coat moved under a layer of tungsten
 - Tip moved few millimeters further from the IP
- Beam-Induced Background:
 - Reduced photon and e^+/e^- flux
 - Reduced occupancy in the tracking system
 - Increased neutron flux





The Metric Problem

- Manually → human-driven consideration
- Advance ML allow to analytically determine the best solution:

• y = f(X) with

- $y \rightarrow \text{target variable (BIB flux)}$
- $X \rightarrow$ input features (**Geometrical parameters**)
- $f \rightarrow ML \mod (XBGBoost regressor)$
- $X_{best} \to f'(X_{best}) = 0$
- In technical consideration skipped, because beyond the scope of this presentation ...
- Using BIB flux as target variable, this method is not applyable the algorithm would just find the largest nozzle within parameters space





$$\mathcal{M} = (a+P) \cdot \frac{\Delta flux}{flux_{ref}} + b \cdot \frac{\Delta \theta_{nozzle}}{\theta_{ref}} + c \cdot \frac{\Delta V}{V_{ref}}$$

Observables:

Parameters:

- $\frac{\Delta f lux}{f lux_{ref}}$ \rightarrow BIB flux relative difference from optimized nozzle configuration
- $\frac{\Delta \theta_{nozzle}}{\theta_{ref}}$ \rightarrow Relative variation of the nozzle tip angle,

which determines the detector acceptance

• $\frac{\Delta V}{V_{ref}} \rightarrow$ Relative variation of the overall nozzle volume

• $a = 1 \rightarrow$ reference parameter

•
$$P\left(\frac{\Delta flux}{flux_{ref}}\right) \rightarrow \text{Penalty for «large» flux}$$

- $b = 3.7 \cdot 0.9 \rightarrow$ Quantify acceptance gain relative to performance degradation due to BIB, with a correction
- c = 0.3 → Small award for lighter nozzle (could be changed considering nozzle engineering)



Gradient Descent Approach

- $\mathcal{M} = f(X)$ with
 - *X* → input features (**Geometrical parameters**)
 - $f \rightarrow ML \mod (XBGBoost regressor)$
 - $X_{best} \to f'(X_{best}) = 0$
- SGBoost is not differentiable, not possible to compute f'
- XGBoost used to produce a pseudo-dataset (X, flux) $\rightarrow (X, \mathcal{M}(X, flux))$
- Trained a Sinusoidal Representation Network (SiReN) $\rightarrow \mathcal{M}$ = f(X)
- Using **Stochastic Gradient Descent** $f'(X_{best}) = 0$ has been computed







Final Nozzle Design

- SGD results, Nozzle very similar to Design XXI
- M = -0.02





14



Conclusion

Forward muons

- can be tagged unless they end up in the beam pipe
- measuring momentum is challenging

Nozzle Design

- Optimal design has been achive
- Definition of «Optimal», i.e. metric \mathcal{M} , can be changed, but the ML pipeline is ready to run

Further Study

- Measuring forward muons momentum
- Improve XGBoost as some overtraining might be present





MInternational UON Collider Collaboration

Thank you for the attention



References

[1] Y. Alexahin, E. Gianfelice-Wendt, A 3-TeV MUON COLLIDER LATTICE DESIGN, <u>Insiperhep.net</u>

- [2] P. Li, Z. Liu, K. Lyu, HIGGS WIDTH AND COUPLINGS AT HIGH ENERGY MUON COLLIDERS WITH FORWARD MUON DETECTION, <u>arxiv.org</u>
- [3] M. Ruhdorfer, E. Salvioni, A. Wulzer, INVISIBLE HIGGS FROM FORWARD MUONS AT A MUON COLLIDER, <u>arxiv.org</u>
- [4] MODE Collaboration, <u>mode.github</u>
- [5] A. Baranov et al., OPTIMIZING THE ACTIVE MUON SHIELD FOR THE SHIP EXPERIMENT AT CERN, <u>SHIP optimization</u>
- [6] Z. Liu, HIGGS WIDTH AND COUPLINGS AT HIGH ENERGY MUON COLLIDERS WITH FORWARD MUON DETECTION, <u>indico.cern</u>





Muon decay position



19



High Statistics Approach

- Lessons learned:
 - The Beam Pipe cannot be touched
 - Is Boreth layer really effective?
 - Tried to put the Boreth inside the nozzle











Measuring Forward Muons Energy

- Not feasible with track-like detector
- Energy deposit detector in the cavern only way







BIB simulation with FLUKA





22





- 60 layers of 19-mm steel absorber + plastic scintillating tiles;
- \bullet 30x30 mm² cell size;

electromagnetic calorimeter

- 40 layers of 1.9-mm W absorber + silicon pad sensors;
- 5x5 mm² cell granularity;
- ightarrow 22 X₀ + 1 λ₁.

muon detectors

- 7-barrel, 6-endcap RPC layers interleaved in the magnet's iron yoke;
- 30x30 mm² cell size.

Detector



superconducting solenoid (3.57T)

tracking system

- Vertex Detector:
 - double-sensor layers (4 barrel cylinders and 4+4 endcap disks);
 - 25x25 µm² pixel Si sensors.
- Inner Tracker:
 - 3 barrel layers and 7+7 endcap disks;
 - 50 µm x 1 mm macropixel Si sensors.
- Outer Tracker:
 - 3 barrel layers and 4+4 endcap disks;
 - 50 µm x 10 mm microstrip Si sensors.

shielding nozzles

 Tungsten cones + borated polyethylene cladding.

23



BIB characteristics

• By requiring a window of $\pm 100 \ ps$ with respect to the expected time of arrival in the layers

BIB reduced by 5 order of magnitudes





BIB characteristics

BIB particles passing through the layers within the time window (1.4% of b.c)





(a rough) Tracking

Assuming that forward muons are

produced at the IP, a straight line

is the defined for each point in

layer 1

- The line is propagated to layer 2 and 3. If at least 1 particle is present in the expected position
 - ± 1 cm, the particle is tagged as a forward muon





Performance

 Total counts within ±100 ps time window with respect to muons arrival time on layers:

Event	Layer 1	Layer 2	Layer 3
BIB*	$2.5\cdot 10^4$	$2.7\cdot 10^4$	$3.0\cdot10^4$
Z fusion**	3228/6150	3232/6150	3225/6150

 A rough tracking is performed to discard particles that are not coming from IP:

Event	Global Efficiency [%]	Tracking Efficiency [%]
BIB [#]	< 0.28	
Z fusion ^{##}	49.5	99.2

0 particles tracked, estimation on the total bunch crossing computed according to [1] ## Efficiency computed on the total muon generated, i.e. 6150, not on only the ones who pass through the nozzle and the layers



Low Statistic simulation



- Two step: 2% of one beam, one bunch crossing
- Pipeline: 0.025% of one beam, one bunch crossing
- Pipeline nozzles smaller than

28

original (aperture = 20 cm)

• $\sigma = \sqrt{\# particles}$



XGBoost Features importance



29

- -



XGBoost possible overtraining

- 2 observables were simulated with only fixed values
- That might cause overtraining





$$\mathcal{M} = (a+P) \cdot \frac{\Delta f lux}{f lux_{ref}} + b \cdot \frac{\Delta \theta_{nozzle}}{\theta_{ref}} + c \cdot \frac{\Delta V}{V_{ref}}$$

Penalty function:

• $P = p0 \cdot \frac{1}{1 - e^{-X}}$

•
$$X = p1 \cdot \left(\frac{\Delta flux}{flux_{ref}} - p2\right)$$

• Parameters:

- $p0 = 20 \rightarrow \text{penalty weight}$
- p1 = 40 → stepness, i.e. how «fast» is the transition
- $p2 = 0.93 \rightarrow$ relative flux threshold



31





• Parameter $b = 3.7 \cdot 0.9$:

Studied Energy resolution as function of BIB level







• Parameter $b = 3.7 \cdot 0.9$:

- Studied Energy resolution as function of BIB level
- Studied Efficiency gain as function of the nozzle tip angle





$$\mathcal{M} = (a+P) \cdot \frac{\Delta f lux}{f lux_{ref}} + b \cdot \frac{\Delta \theta_{nozzle}}{\theta_{ref}} + c \cdot \frac{\Delta V}{V_{ref}}$$

• Parameter $b = 3.7 \cdot 0.9$:

- Studied Energy resolution as function of BIB level
- Studied Efficiency gain as function of the nozzle tip angle
- Estimated how BIB increase as function of the nozzle tip angle

Normalized Flux vs. Binned Theta Tip Rescaled Average Flux [%] 240 Rescaled Linear Fit (100% at $\theta = 0.174$ rad) 220 0.174 rad] 200 II 180 at 0 : tij 160 J 140 [% of fir 140 120 100 0.12 0.13 0.14 0.15 0.16 0.17 Theta Tip (binned)



$$\mathcal{M} = (a+P) \cdot \frac{\Delta f lux}{f lux_{ref}} + b \cdot \frac{\Delta \theta_{nozzle}}{\theta_{ref}} + c \cdot \frac{\Delta V}{V_{ref}}$$

• Parameter $b = 3.7 \cdot 0.9$:

- Studied Energy resolution as function of BIB level
- Studied Efficiency gain as function of the nozzle tip angle
- Estimated how BIB increase as function of the nozzle tip angle
- Plotted Resolution and Efficiency as function of BIB level, computed the ratio of the derivatives





SiReN performance



36



Incoherent Pair Production

Another source of background due to beam-

beam interaction

- Produced the e^{\pm} pairs with GUINEAPIG
- Products propagated in FLUKA as for two
 Step Simulation
- Reconstruction in the tracking system
- Slightly increase in occupancy (about 5%)

