UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics

Final Dissertation

Study of b- and c- jets identification for Higgs coupling measurement at the Muon Collider

Thesis supervisor

Prof. Donatella Lucchesi

Thesis co-supervisor

Dr. Lorenzo Sestini

Candidate

Giacomo Da Molin

Academic Year 2020/2021

Contents

1. Introduction

The Muon Collider is an innovative project ideated to push the research in the high energy frontier of particle physics. This proposed machine would exploit the collision of μ^+ and μ^- to perform studies at the TeV scale with several key advantages compared to other experiments. One of the biggest opportunities provided by this environment is the relatively large number of Higgs bosons which would arise from these collisions. Such feature would allow precise measurements of all the Higgs boson field parameters, such as the Yukawa coupling as well as the trilinear and quadrilinear Higgs self-couplings, which may provide a link to New Physics (NP) if deviations from the Standard Model (SM) are going to be found.
In order to achieve such a result, one must reconstruct the Higgs decay final products in the challenging background of the Muon Collider, given by the decay products of the μ constituting the beams and their interactions with parts of the accelerator. In the Higgs sector it is of pivotal importance to be able to reconstruct and identify b-jets and c-jets.
The aim of this thesis is to study the properties of the Muon Collider experimental environment at $E_{cm} = 3$ TeV, with a particular care for the vertices and jets reconstruction, and to use a detailed simulation to identify the H$\rightarrow b\bar{b}$, finding out with which statistical precision the cross section can be measured. The study will be performed by selecting b-jets with a secondary vertex, signature of a b-hadron decay, inside them. Both the beam induced background and the physical background are included in this work. After this, a study on the most discriminant jets observables will be performed, in order to identify which features can be used to distinguish b-,c- and light jets. This information, combined with machine learning techniques, will be pivotal to study the H$\rightarrow c\bar{c}$ in the future. One of these possible techniques is the use of a Deep Neural Network (DNN) for jet flavour identification, which, using the observables studied before, will be presented in the last chapter.
This work will be structured as follows:

- **Chapter 2) Higgs Physics:** In this chapter the Standard Model Higgs sector will be presented, along a brief description of its most important parameters. This will be followed by some of the measurements performed so far on said parameters at LHC, showing also which parameters must be known more accurately or have to be explored to investigate possible New Physics. Finally the expected precision on Higgs parameters at future experiments will be shown and compared with the Muon Collider.

- **Chapter 3) The Muon Collider:** In this chapter the possibilities presented by the Muon Collider will be explored. An overview of the advantages provided by the use of μ as colliding particles will be presented. The main technical challenges of building the Muon Collider will be listed: a small summary will be given on the accelerator side and on the radiation hazards, while the unique beam induced background and its consequences on the physics reconstruction will be explored in more detail.

- **Chapter 4) Detector description:** In this chapter the detector used in this work will be described, presenting the parameters of each sub-detector. The strategies employed to reduce the effects on the beam-induced background at detector level will be discussed afterwards.

- **Chapter 5) Physics objects reconstruction:** In this chapter we are going to describe the simulation workflow and the algorithms used for the event reconstruction, with particular care for the jets and vertices algorithms. Finally, these configurations will be used to determine the

secondary vertex tagging efficiency and mistag on both b-jets and c-jets samples.

- **Chapter 6) H$\rightarrow b\bar{b}$ identification and cross section measurement:** In this chapter we are going to characterize the observables of $H \rightarrow b\bar{b}$ events, focusing on jets and secondary vertices. The $H \rightarrow b\bar{b}$ invariant mass distribution will be reconstructed using pairs of tagged jets. From this distribution, in the following section, the backgrounds will be analized and compared with the signal. Finally from the total dijet invariant mass distribution, the expected precision on the $H \rightarrow b\bar{b}$ cross section will be extracted.
- **Chapter 7) c-jets identification and prospects on H$\rightarrow c\bar{c}$:** In this chapter we are first going to apply the technique used for the H$\rightarrow b\bar{b}$ to the H$\rightarrow c\bar{c}$. We are going to see the limitations of tagging c-jets with secondary vertices and how this, combined with the low σ_{Hcc} cross section, limits the possibility to measure accurately the cross section. We are then going to look to other methods to perform c-tagging, and in order to do this, we are first going to study the observables characterizing b-jets, c-jets and jets from light quarks. This study can be used as starting point to the development of machine learning techniques to perform jet flavour identification. A first attempt at the development of a Deep Neural Network tasked with flavour-tagging the jet is going to be shown in the last section.

2. Higgs Physics

As the goal of this thesis is to study with a detailed simulation the cross section of the $H \to b\bar{b}$ and investigate techniques for future studies of the $H \to c\bar{c}$, it is essential to describe what we expect the Higgs boson is and what we measured of it so far. This chapter will first provide a small introduction to the role of the Higgs in the SM and which parameters are used to describe its interactions. Then, in the second section, we show the Higgs properties the physics community was able to measure at LHC. Finally, in the last section, we will present a list of proposed particle physics experiments and compare their performances in this pivotal sector, focusing on the b-coupling and the Higgs self interactions.

2.1 Higgs interactions in the SM

The Higgs boson is a fundamental piece of the Standard Model (SM), the theory that as of today better explains the fundamental interactions of nature. This particle is generated as a consequence of spontaneous symmetry breaking (SSB) of the Higgs potential, which is essential to the SM as it allows massive bosons in the theory without breaking the gauge symmetries. It is predicted to be an electrically neutral scalar (spin 0 and CP even) particle. At the end of this procedure, the SM Lagrangian contains both Higgs interactions with the other massive vector bosons, Z and W, as well as the H self-interactions:

$$\mathscr{L}_{H-bos} = -\frac{M_H^2}{2}H^2 - \lambda v H^3 - \frac{\lambda}{4}H^4 - \left(1 - \frac{H}{v}\right)^2 \left[M_W^2 W^+W^- + \frac{M_Z^2}{2}Z^2\right] \tag{2.1}$$

$$with \qquad M_H^2 = 2\lambda v^2 \qquad M_W^2 = \frac{g^2 v^2}{4} \qquad M_Z^2 = \frac{v^2(g^2+g^2)}{4}$$

This sector is described by 4 parameters, 2 related to the strength of the coupling of the symmetries broken by the SSB (g for the $SU(2)_L$ and g' for the $U(1)_Y$) and 2 by the Higgs potential itself (the quartic parameter λ and the vacuum expectation value of the Higgs potential v).
The Higgs appears in another part of the SM: the Yukawa term of the Lagrangian. This gives masses to the SM fermions and has the following form in the physical (mass) basis:

$$\mathscr{L}_{Yukawa} = -\sum_i m_i \left(1 + \frac{H}{v}\right)\bar{\psi}_i\psi_i \qquad with \qquad m_i = y_i\frac{v}{\sqrt{2}} \tag{2.2}$$

where the i index runs over all massive fermions.
All mass terms in the SM (seen in equations 2.1 and 2.2) appear only thanks to the Higgs and its interactions determine the value of the masses. In fact the masses of both fermions and bosons are proportional the vev parameter v times a combination of some couplings: the Yukawa couplings y_i for the fermions, the respective couplings to the fermions for Z and W and the square root of the quartic parameter λ for the Higgs.
The quartic parameter not only fixes the Higgs mass, but also is the couplings of other interactions provided by this theory: the Higgs self interactions in the first terms of equation 2.1. This allows some theoretical predictions in this crucial sector: if one measures M_H, it can fix the value of λ and so the trilinear and quadrilinear couplings, which can then be measured. For this reason, this measure is of great importance to test one of the main mechanism of the SM, but also for New Physics (NP) theories.

2.2 Higgs measurements at LHC

The Higgs boson was the last discovered piece of the SM, whose evidence was found only in 2012 by CMS [1] and ATLAS [2] at LHC, a Multi-TeV proton-proton collider in Geneva, Switzerland. In p-p collisions at 7 and 8 TeV in the hadron center of mass frame, the main channels producing an Higgs are strongly dependent by the parton distributions functions (PDF) of the proton. In fact, protons are composite particles made up by partons, each present inside the proton with a fraction of its momenta x with a probability determined by the PDF. In LHC, the actual colliding particles are not the protons but their "pieces": quarks and gluons.
The production channels are represented by the Feynman diagrams in figure 1.

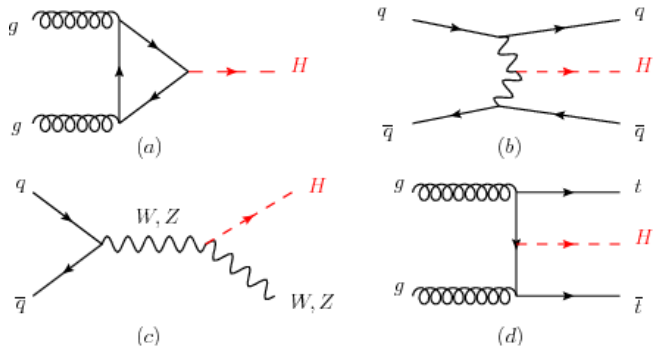


FIG. 1. The most important Higgs production channels at LHC: (a) gluon fusion, where the triangular loop is dominated by top quarks; (b) vector boson fusion; (c) Higgs-strahlung; (d) $\bar{t}t$ associate production. Taken from [3].

Due to the shape of PDF, the main Higgs production channels have mostly light or even massless initial states, and the Higgs is emitted by an heavy intermediate particle, be it a weak vector boson ((b),(c) in fig. 1) or an heavy quark ((a),(d) in fig. 1), with the gluon fusion being the dominant process at this $\sqrt{s}$.
The biggest decay branching ratio is in the heaviest particle compatible with the phase space: for a Higgs of mass $m_H \sim 125$ GeV, as measured by [1] and [2], this is the b-quark, since it is not possible to produce on-shell W^+W^-, ZZ or $\bar{t}t$. However the production of two vector bosons, one of which is off-shell and denoted with *, is still the second larger channel for the W and of order of some % for the Z. Then, before the Higgs direct decay to other heavy particles such as $\tau^+\tau^-$ and $\bar{c}c$, loop diagrams dominated by intermediate top quarks produce a decay channel in 2 gluons (the time reversal of the gluon fusion showed in fig. 1). A summary of the most important branching ratios is given in the table in figure 2.

Decay mode	Branching ratio
$H \to b\bar{b}$	56.1 %
$H \to WW^*$	23.1 %
$H \to gg$	8.5 %
$H \to \tau^+\tau^-$	6.2 %
$H \to c\bar{c}$	2.8 %
$H \to ZZ^*$	2.9 %
$H \to \gamma\gamma$	0.23 %
$H \to Z\gamma$	0.16 %
$H \to \mu^+\mu^-$	0.021 %
Γ_H	4.2 MeV

FIG. 2. Higgs boson decay branching ratios and total width for $m_H = 126GeV$, taken from [4].

In truth [1] and [2] did not use the bottom jet decay mode to measure m_H, but the rarer $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$, since the high energy leptons and photons give a clear signal. Their combined results of m_H are shown in figure 3. The $ZZ^* \to 4\ell$ channel was also used to study the J^P of the Higgs, to verify through its kinematic variables if it is indeed a scalar particle. Results from both ATLAS and CMS [5] discard the 0^- and 2^+ hypothesis with CL of order 99%, confirming, since $J = 1$ states are forbidden by the P symmetry in the $H \to \gamma\gamma$ and $H \to gg$, that indeed the H is a scalar particle as predicted by the SM.

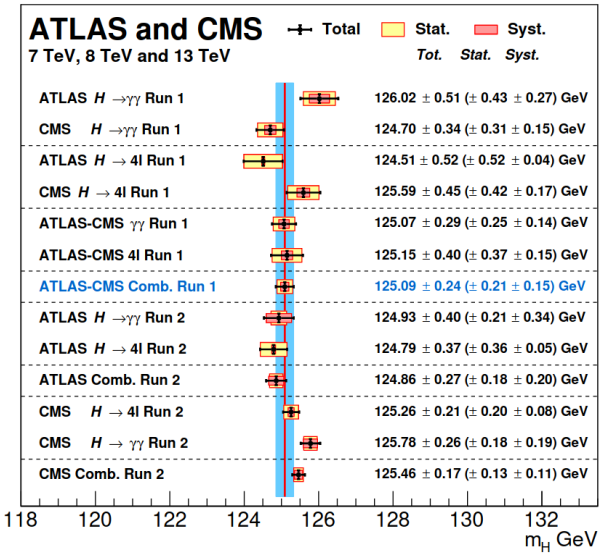


FIG. 3. Summary of the CMS and ATLAS mass measurements in the $\gamma\gamma$ and ZZ channels in Run 1 and Run 2, from [5].

Measurements on other channels have been performed to study the Higgs couplings and see if they are consistent with the SM. Some of the combined results of ATLAS and CMS are showed in figures 4 and 5. Although one can see that the values are compatible with the SM expectations, there is not enough precision to be able to rule out NP, especially for other channels not showed because their uncertainty is still not significant enough, such as $H \to \mu\mu$, $H \to ZZ$ from H produced via Higgstrahlung and $H \to bb$ with H produced via gluon gluon fusion. The production mechanism is reconstructed by the kinematics of the rest of the event: in case of a vector boson fusion, two high p_T jets (originated from the quarks in the final state) with an high invariant mass are required; for the Higgs-strahlung mode, still one expects either more high p_T jets but with lower invariant mass or at least an high p_T lepton not used to reconstruct the Higgs; events with two leptons and two b-jets not from the H are signature of a $\bar{t}t$ associate production; events without these signatures are assumed coming from the dominating gluon gluon fusion.

Instead the Higgs self-couplings have not yet been measured. The SM shows a clear prediction (equation 2.1): the trilinear coupling expected value is λv and the quadrilinear λ. This parameters are known: v is measured in the muon decay, and from it and the measure of M_H, one can extract λ. Any deviation from these predictions will be a clear signal of NP. Yet despite being so interesting, this measure has not been performed yet: the cross section, even at LHC with $\sqrt{s} = 13$ TeV is prohibitively small.
This is one of the main reasons new colliders are needed: not only they could improve the precision on the measurement of the couplings with H and other SM particles, but they would allow to test the Higgs self-interactions. Many Dark Matter models predict interactions between it and the Higgs (the so-called Higgs Portal), and require better knowledge of the Higgs parameters to confirm or dismiss many of such theories, pushing forward the research in the fundamental interactions. Some proposal of future colliders will be shown in the next section, as well as their predicted precision for measurements in this sector, with particular focus on the machine studied in this work: the Muon Collider.

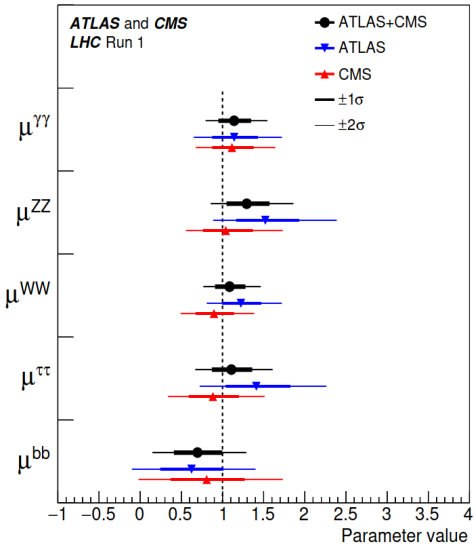


FIG. 4. Best fit for the decay signal strenght $\mu = {}^{BR(H\to b\bar{b})}/_{BR(H\to b\bar{b})_{SM}}$ from combined data of CMS and ATLAS. Taken from [6].

2.3 Prospects of Higgs measurements at future colliders

Other machines have been proposed by the scientific community to proceed the study of important processes in this sector.
One of these has already been approved: it is the High Luminosity Large Hadron Collider (HL-LHC). It is an upgrade of LHC, which aims to study proton-proton collisions at $\sqrt{s} = 14$ TeV and to improve the luminosity of its predecessor by a factor 10. This new statistics will provide an enhancement in the precision of measurement of several Higgs parameters. However the magnitude of these improvements will strongly depend on the attenuation of the theoretical uncertainties on the protons PDF.
The other most prominent proposed machines are:

- High-Energy LHC (HE-LHC) [8], proton-proton collider with $\sqrt{s} = 27$ TeV which could be built in the LHC tunnel if the R&D for the 16 T magnets (needed for FCC) is successful. The main production channel is the gluon-gluon fusion and the main advantage of this candidate is the possibility to study the Higgs production at high transverse momenta, where NP may lie.

- Future Circular Collider (FCC) [9], which could be built as an hadron collider (FCC-hh), a circular electron-positron collider (FCC-ee) or as a $e - h$ machine (FCC-eh). Each would provide several interesting characteristics: FCC-hh can reach unprecedented energy in the C.M. up to $\sqrt{s} = 100$ TeV, which not only can push the energy frontier but also measure the trilinear and quadrilinear Higgs couplings, since the cross section of these processes is much bigger than at LHC; FCC-ee can exploit the cleaner experimental environment to perform precision measurements that allow to estimate the Higgs mass corrections due to loops, giving information on eventual Beyond the Standard Model (BSM) particles; FCC-eh with a 60 GeV e^- and 50 TeV proton can provide measurements of the Higgs self-interactions as well as probing extended Higgs sectors with some dark matter (DM) candidates.

- Circular Electron Positron Collider (CEPC) [10], which can be run as an Higgs factory at $\sqrt{s} = 240$ GeV looking for $e^+e^- \to$ZH and exploiting the clean experimental environment.

- International Linear Collider (ILC) [11], a linear e^+e^- collider able to easily change its energy in the center of mass. The best results in the Higgs sector are expected a $\sqrt{s} = 250$ GeV, exploiting the recoiling Z to tag the Higgs. This experiment would be sensible also to invisible or exotic Higgs decays. Other advantages include the possibility of exploiting polarized beams to produce new observables, as well as the capability to switch to $\sqrt{s} = 500$ GeV to access to the Higgs self couplings and top-Higgs interactions.

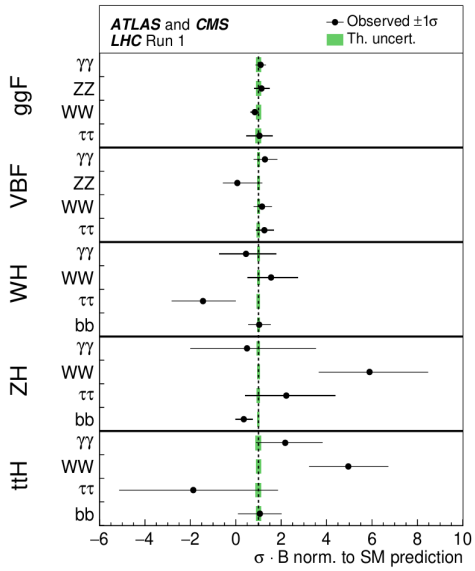


FIG. 5. Best fit value for $i \xrightarrow{(a)} H \xrightarrow{(b)} f$ over the SM predictions from [7]. The leftmost column indicates the Higgs production mechanism (a), while the second one refers to the Higgs decay products (b). As stated before, only the channel with enough significance are presented. The black bars are the 1σ experimental errors, while the green area indicates the theoretical 1σ uncertainty around the expected value, equal to 1.

- Compact Linear Collider (CLIC) [12], forseen to collide polarized e^- with e^+ at $\sqrt{s} = 380$ GeV, 1.5 TeV and 3 TeV. The first energy allows precision studies in the SM Higgs physics and H→invisible, while the higher energies allow to probe the Higgs self-interactions, the $Ht\bar{t}$ channel and look for Higgs exotic decay channels.
- Large Hadron electron Collider (LHeC) [13], in which an electron beam of 60 or 140 GeV, produced with a linear or circular accelerator, would collide with an high energy proton beam. Measurements can be performed using high precision Deep Inelastic Scattering (DIS) with a much bigger phase space available and higher luminosity than its ep predecessor, HERA.

A summary table of the main parameters of these machines is presented in fig. 6.

To better describe the potential of each of these machines, it is convenient to use the "kappa" framework. If the width of the Higgs is negligible, these "kappa" k_i describe the ratio of the coupling of the Higgs to the the particle i compared to the prevision of the SM. In practice, be i the initial state producing the H and f the decay products of the boson

$$(\sigma \cdot BR)(i \to H \to f) = \frac{\sigma_i \cdot \Gamma_f}{\Gamma_H} \tag{2.3}$$

is sent into

$$(\sigma \cdot BR)(i \to H \to f) = \frac{\sigma_i k_i^2 \cdot \Gamma_f^{SM} k_f^2}{\Gamma_H k_H^2} \tag{2.4}$$

where k_H^2 accounts for width corrections of the SM Higgs width due to deviations from the SM-predicted couplings. If the SM is respected, all k are expected to be 1. The projected precision of the measurements is depending also from the fact that H might be decaying into either invisible or exotic particles. For this reason, the predictions are divided in different scenarios, whether H can decay into invisible products or in still unknown (but visible) particles. If the Higgs does not decay in any of those, the scenario is called "kappa-0" and the prediction for all the above cited machines are presented in figure 7. Several e^+e^- colliders and the combined results of FCC seem to be able to probe k_b below the 1% precision, as well as measuring k_c with good accuracy.

A similar study has been performed [15] on the expected uncertainty on the k_i provided by a 10 TeV Muon Collider, and the results are shown in figure 8. These estimations prove that the Muon Collider can provide excellent insight in the Higgs physics.

Collider	Type	$\sqrt{s}$	$\mathscr{P}$ [%] $[e^-/e^+]$	N(Det.)	$\mathscr{L}_{\text{inst}}$ $[10^{34}]$ cm^{-2}s^{-1}	$\mathscr{L}$ [ab^{-1}]	Time [years]
HL-LHC	pp	14 TeV	-	2	5	6.0	12
HE-LHC	pp	27 TeV	-	2	16	15.0	20
FCC-hh$^{(*)}$	pp	100 TeV	-	2	30	30.0	25
FCC-ee	ee	M_Z	0/0	2	100/200	150	4
		$2M_W$	0/0	2	25	10	1-2
		240 GeV	0/0	2	7	5	3
		$2m_{top}$	0/0	2	0.8/1.4	1.5	5 (+1)
ILC	ee	250 GeV	±80/±30	1	1.35/2.7	2.0	11.5
		350 GeV	±80/±30	1	1.6	0.2	1
		500 GeV	±80/±30	1	1.8/3.6	4.0	8.5 (+1)
		1000 GeV	±80/±20	1	3.6/7.2	8.0	8.5 (+1-2)
CEPC	ee	M_Z	0/0	2	17/32	16	2
		$2M_W$	0/0	2	10	2.6	1
		240 GeV	0/0	2	3	5.6	7
CLIC	ee	380 GeV	±80/0	1	1.5	1.0	8
		1.5 TeV	±80/0	1	3.7	2.5	7
		3.0 TeV	±80/0	1	6.0	5.0	8 (+4)
LHeC	ep	1.3 TeV	-	1	0.8	1.0	15
HE-LHeC	ep	1.8 TeV	-	1	1.5	2.0	20
FCC-eh	ep	3.5 TeV	-	1	1.5	2.0	25

FIG. 6. Characteristics of main future collider proposals. $\mathscr{P}$ describe the polarization degree of e^- and e^+ respectively, N is the projected number of collision points where detectors would be placed, $\mathscr{L}_{inst}$ is the instantaneous luminosity, $\mathscr{L}$ the integrated luminosity accumulated in the time reported afterward (the years within brackets refer to the duration of shutdowns, needed for example to change $\sqrt{s}$). Taken from [14].

kappa-0	HL-LHC	LHeC	HE-LHC		ILC			CLIC			CEPC	FCC-ee		FCC-ee/eh/hh
			S2	S2′	250	500	1000	380	1500	3000		240	365	
κ_W [%]	1.7	0.75	1.4	0.98	1.8	0.29	0.24	0.86	0.16	0.11	1.3	1.3	0.43	0.14
κ_Z [%]	1.5	1.2	1.3	0.9	0.29	0.23	0.22	0.5	0.26	0.23	0.14	0.20	0.17	0.12
κ_g [%]	2.3	3.6	1.9	1.2	2.3	0.97	0.66	2.5	1.3	0.9	1.5	1.7	1.0	0.49
κ_γ [%]	1.9	7.6	1.6	1.2	6.7	3.4	1.9	98⋆	5.0	2.2	3.7	4.7	3.9	0.29
$\kappa_{Z\gamma}$ [%]	10.	–	5.7	3.8	99⋆	86⋆	85⋆	120⋆	15	6.9	8.2	81⋆	75⋆	0.69
κ_c [%]	–	4.1	–	–	2.5	1.3	0.9	4.3	1.8	1.4	2.2	1.8	1.3	0.95
κ_t [%]	3.3	–	2.8	1.7	–	6.9	1.6	–	–	2.7	–	–	–	1.0
κ_b [%]	3.6	2.1	3.2	2.3	1.8	0.58	0.48	1.9	0.46	0.37	1.2	1.3	0.67	0.43
κ_μ [%]	4.6	–	2.5	1.7	15	9.4	6.2	320⋆	13	5.8	8.9	10	8.9	0.41
κ_τ [%]	1.9	3.3	1.5	1.1	1.9	0.70	0.57	3.0	1.3	0.88	1.3	1.4	0.73	0.44

FIG. 7. Expected uncertainty on the k in the various colliders, obtained by a fit on the projected measurements [14] in the kappa-0 scenario. The row under ILC, CLIC and FCC-ee specify the $\sqrt{s}$, while the one below HE-LHC express two different scenarios: S2 if the uncertainty on the luminosity is reduced to 1%, as foreseen for HL-LHC, S2' if both theoretical and modelling uncertainty are halved compared to S2. If an experiment lacks the sensitivity to measure a parameter, it's value is fixed to the SM expectation and in the table a - is placed. If the value is let free in the fit due to lack of documentation, a star is placed near that value. The last column corresponds to the combined performance of FCC-ee$_{240}$, FCC-ee$_{365}$, FCC-eh and FCC-hh.

A study on the Muon Collider expected precision on the b-coupling has been performed [16], with a detailed simulation at $\sqrt{s}$ =1.5 TeV and conservative scalings at higher energies in the CM. Assuming the parameters reported in figure 9.a, the expected Muon Collider performance can be compared with the most similar machine: CLIC, the circular leptonic collider with the highest $\sqrt{s}$. This has been done [16] with a detailed simulations and the results are shown in fig. 9.b. Despite the fact that for the Muon Collider conservative assumption were made and the CLIC results at higher energy exploit

	10 TeV Muon Collider
κ_W	0.06
κ_Z	0.23
κ_g	0.15
κ_γ	0.64
$\kappa_{Z\gamma}$	1.0
κ_c	0.89
κ_t	6.0
κ_b	0.16
κ_μ	2.0
κ_τ	0.31

FIG. 8. Expected uncertainty in % in the 10 TeV Muon Collider, in a kappa-0 scenario, with an integrated luminosity of 10 ab^{-1}. These results are extracted with several caveats: scaling from the detailed simulation at 1.5 TeV are used; sensitivities are not optimized; the physical and the beam induced backgrounds and background are not considered (the latter is justified by the fact that at higher $\sqrt{s}$ the BIB is expected to become more manageable, more details in chapter 3.3); off-shell H measurements (which the Muon Collider can provide) are not used. Still the estimations can be compared to those of fig. 7 and show that indeed the Muon Collider is potentially a great Higgs factory.

also the measures previously performed at $\sqrt{s}$, the results are very similar: the Muon Collider is a good machine to measure the coupling of the Higgs to the b quark.

A necessary step forward in particle physics will be given by measurements of the Higgs shape potential, which could confirm the SM expectations or give crucial information for BSM theories. This is done by measuring the trilinear or quadrilinear Higgs self interactions, given by vertices with respectively 3 or 4 Higgs. The SM predictions are immediate from equation 2.1: the trilinear coupling is expected λv, while the quadrilinear is $\frac{\lambda}{4}$. As stated in the previous section, no test was possible at LHC or previous machines because the cross-sections of such processes were too small to be probed. The leap forward will hopefully come from the next generations machines. For the trilinear coupling, more easily accessible than the quadrilinear, several studies have been performed, and we will report here the results of some previsions on leptonic machines, combined with HL-LHC data, [17] and FCC-hh [18]. The former are shown in figure **10**, where the best performance is obtained by ILC and CLIC with a significance of order 20%, whereas the latter, reported in figure **11**, has an expected uncertainty $\delta_{k_\lambda} = \frac{\delta_u}{\frac{du}{dk_\lambda}|_{SM}}$ of 3.4-7.8% depending on the systematic error assumptions, where u is the ratio $\frac{\sigma}{\sigma_{SM}}$ and the channel with the biggest sensitivity is $HH \to \bar{b}b\gamma\gamma$.

For the Muon Collider, an estimation of said coupling in the kappa-0 scenario has been performed in [19]. Although only one final state channel is considered ($HH \to \bar{b}b\bar{b}b$), the precision obtained on δ_{k_3} is greater than that of the previous colliders.

To provide predictions for the quadrilinear coupling, one must also assume the value of λ_3. Several hypothesis and scenarios can be foreseen, but for the results we are about to cite, we are going to compare estimations obtained assuming that λ_3 has no deviations from the SM predictions. A study on ILC and CLIC [20] shows that ILC is expected to have an higher sensitivity and it foresees two data-taking scenarios:

- 4 ab^{-1} at $\sqrt{s} = 500$ GeV + 2.5 ab^{-1} at $\sqrt{s} = 1$ TeV, where the expected precision on deviations from SM is ± 25 at 68% CL;
- 4 ab^{-1} at $\sqrt{s} = 500$ GeV + 8 ab^{-1} at $\sqrt{s} = 1$ TeV, where the expected precision on deviations from SM is ± 20 at 68% CL.

Another study [21] is focused on hadronic colliders at $\sqrt{s} = 27$ TeV (which could be implemented

$\sqrt{s}$ [TeV]	A [%]	ϵ [%]	$\mathcal{L}$ [$cm^{-2}s^{-1}$]	$\mathcal{L}_{int}$ [ab^{-1}]	σ [fb]	N	B	$\frac{\Delta\sigma}{\sigma}$ [%]	$\frac{\Delta g_{Hbb}}{g_{Hbb}}$ [%]
1.5	35	15	$1.25 \cdot 10^{34}$	0.5	203	5500	6700	2.0	1.9
3.0	37	15	$4.4 \cdot 10^{34}$	1.3	324	33000	7700	0.60	1.0
10	39	16	$2 \cdot 10^{35}$	8.0	549	270000	4400	0.20	0.91

(a)

	$\sqrt{s}$ [TeV]	$\mathcal{L}_{int}$ [ab^{-1}]	$\frac{\Delta g_{Hbb}}{g_{Hbb}}$ [%]
Muon Collider	1.5	0.5	1.9
	3.0	1.3	1.0
	10	8.0	0.91
CLIC	0.35	0.5	3.0
	1.4	+1.5	1.0
	3.0	+2.0	0.9

(b)

FIG. 9. (a): H→bb acceptance in the fiducial region (A), selection efficiency (ϵ) in the fiducial region, integrated luminosity $\mathcal{L}_{int}$ as a product of the instantaneous luminosity $\mathcal{L}$ and the data taking time $t = 4 \cdot 10^7$s, number of signal ($N = \mathcal{L}_{int} \cdot A \cdot \epsilon \cdot \sigma$) and background ($B$) events , predicted uncertainty on the cross section ($\frac{\Delta\sigma}{\sigma}$) and couplings ($\frac{\Delta g_{Hbb}}{g_{Hbb}}$). The estimations at higher $\sqrt(s)$ are obtained with a conservative scaling, not accounting that the BIB is predicted to become more tame at higher $\sqrt(s)$. (b) Comparison of Muon Collider projected uncertainties with CLIC's ones. It is important to state that CLIC results at higher energy exploit also the measures previously performed at $\sqrt{s}$, while the Muon Collider estimations are all independent.

by HE-LHC) and $\sqrt{s} = 100$ TeV (aim of FCC-hh) for which data are expected to be taken, with an integrated luminosity of, respectively, 15 ab^{-1} and 30 ab^{-1}. These estimations use both the two Higgs productions channels as well as the 3H. From a global fit in the differential measurements of $pp \to HH$ and an inclusive measurement of $pp \to HHH$, the expected 95% CL bounds on the ratio $k_4 = \frac{\lambda_4}{\lambda_{4SM}}$ are [-21,27] for $\sqrt{s} = 27$ TeV and [-5,12] for $\sqrt{s} = 100$ TeV. Finally, a study [22] has been performed on the Muon Collider as well. The results on the deviations from the SM expectation δ_4, presented in fig. 13, used some simplifying assumptions, but are much better than the ones of the others machines.

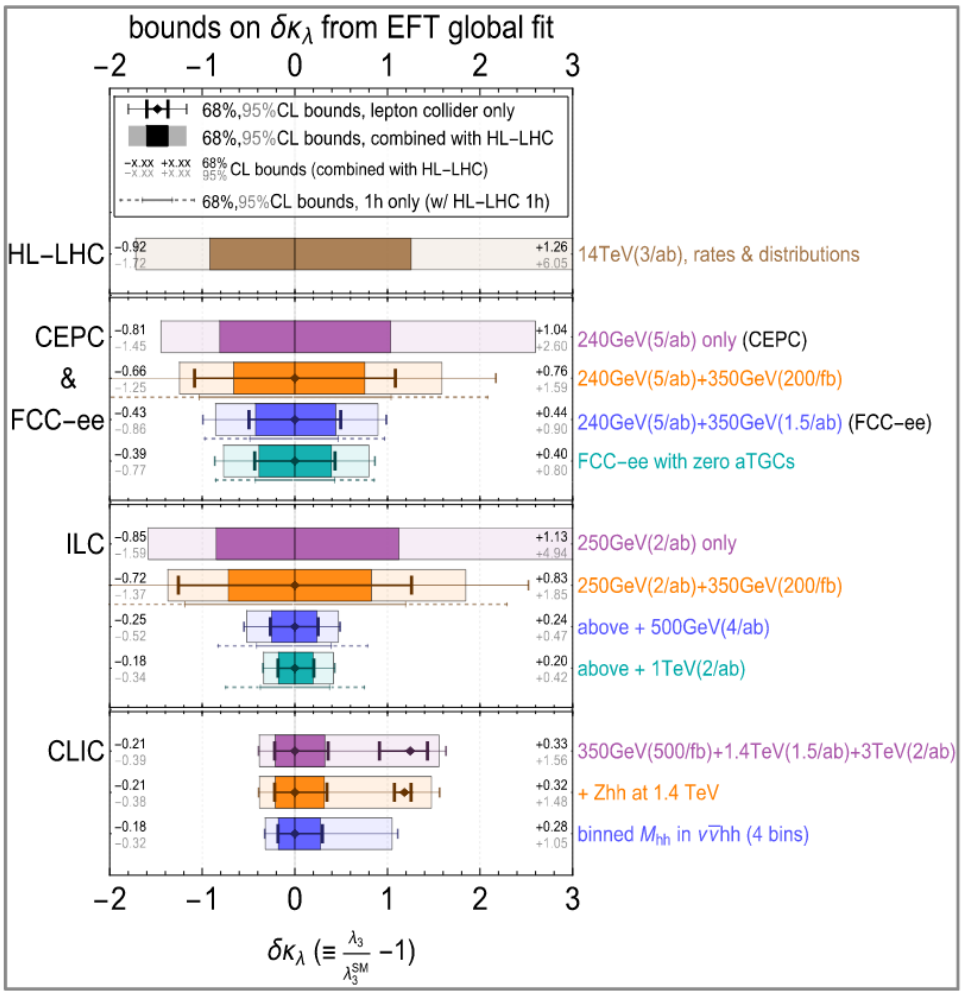


FIG. 10. The main Higgs mechanism productions are the Higgs-strahlung (more sensitive to $\delta_{k_\lambda} < 0$ and the vector boson fusion, more sensitive to $\delta_{k_\lambda} > 0$. Low energy runs can use put some bounds on this parameter through loop corrections to single Higgs production, while high energy can extract it directly from binned fit on invariant mass of the two Higgs bosons. The final bounds are obtained through an EFT global fit. Taken from [17].

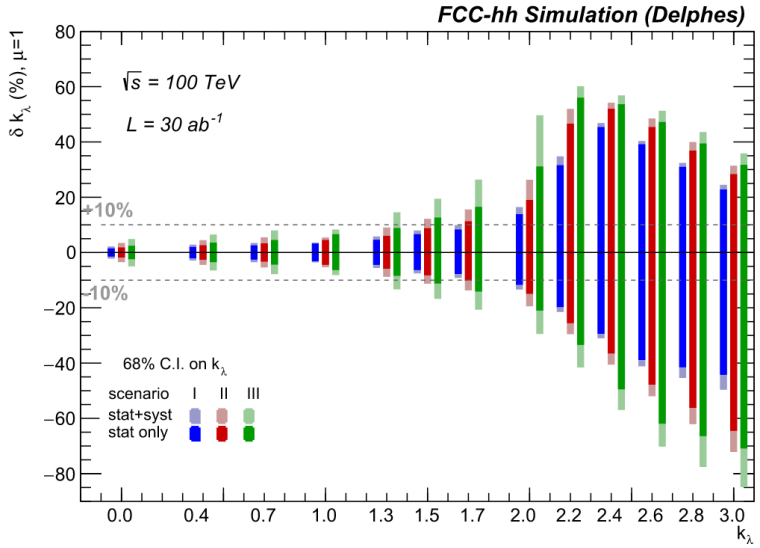


FIG. 11. 68% C.L. on k_λ at FCC-hh after a data collection corresponding to an integrated luminosity of $\mathscr{L}_{int} = 30ab^{-1}$ in function of each scenario and values of k_λ. The channel used in these estimations are $\bar{b}b\gamma\gamma$, $\bar{b}b\bar{b}b$ and $\bar{b}b\tau^+\tau^-$, where machine learning techniques have been used to separate signal and background. The dominant uncertainty in all cases is the systematical one. Taken from [18].

$\sqrt{s}$/TeV	$\mathcal{L}$/ab^{-1}	$\sigma(\ell^+\ell^- \to hh(4b)\nu\bar{\nu})$/ab	$N_{\rm SM}$ events	68% C.L. $\delta\kappa_3$
3	5	$132 \cdot \left[1 + (3.85\, C_H - 0.87\, C_6)\, v^2 + \left(26.8\, C_H^2 + 0.74\, C_6^2 - 5.52\, C_H C_6\right) v^4\right]$	172	[-8.5, 9.9] %
10	10	$239 \cdot \left[1 + (7.25\, C_H - 0.80\, C_6)\, v^2 + \left(196\, C_H^2 + 0.71\, C_6^2 - 8.40\, C_H C_6\right) v^4\right]$	621	[-4.9, 5.3] %
14	20	$257 \cdot \left[1 + (8.43\, C_H - 0.79\, C_6)\, v^2 + \left(300\, C_H^2 + 0.68\, C_6^2 - 9.28\, C_H C_6\right) v^4\right]$	1336	[-3.4, 3.6] %
30	90	$271 \cdot \left[1 + (12.8\, C_H - 0.79\, C_6)\, v^2 + \left(1389\, C_H^2 + 0.78\, C_6^2 - 13.8\, C_H C_6\right) v^4\right.$	6341	[-1.6, 1.6] %

FIG. 12. Cross sections and uncertainty on k_3 at several lepton collider energies. The cross sections include also effects from new physics couplings parametrized by C_H and C_6. The number of events predicted by the SM is reported after considering a 26% signal selection efficiency, and including only b-quarks with $p_T > 10$ GeV and $10° < \theta < 170°$ to take into account the space occupied by the nozzles. Although this angular requirements cuts significantly the rate at higher energies, this does not spoil the measurement because the sources of background would be boosted along the longitudinal direction too, meaning that the optimal region in which to measure is the central one.

$\sqrt{s}$ (TeV)	Lumi (ab^{-1})	Constraints on δ_4 (with $\delta_3 = 0$) x-sec only, acceptance cuts		
		1 σ	2 σ	3 σ
6	12	$[-0.50, 0.70]$	$[-0.74, 0.95]$	$[-0.93, 1.15]$
10	20	$[-0.37, 0.54]$	$[-0.55, 0.72]$	$[-0.69, 0.85]$
14	33	$[-0.28, 0.43]$	$[-0.42, 0.58]$	$[-0.52, 0.68]$
30	100	$[-0.15, 0.30]$	$[-0.24, 0.38]$	$[-0.30, 0.45]$
3	100	$[-0.34, 0.64]$	$[-0.53, 0.82]$	$[-0.67, 0.97]$

FIG. 13. Expected sensitivity at deviation of λ_4 from the SM in function of the energy in the center of mass $\sqrt{s}$ at a muon collider. The integrated luminosities are computed predicting a data-taking time of 10 years. As for the acceptance cuts, $p_T > 20GeV$ and $|\eta| < 3$ are required on the Higgs products. Taken from [22]

3. The Muon Collider

In this chapter the Muon Collider proposed machine, as well as its capabilities and challenges, will be presented. First we will give a brief introduction to its main characteristics, as well as what advantages it provides with respect to hadronic and electronic colliders. Then the main challenges of building a Muon Collider will be shown: the second section of this chapter will focus on the technological challenges involved in creating a muon accelerator and which schemes the scientific community is studying to overcome them; on the other hand the third section will focus on the effects of the decays of μ used in the beams, on how this requires a study to avoid radiation hazards and on the effect these decays have on the detector, through the Muon Collider unique Beam Induced Background, as well as showing that a carefully designed Machine Detector Interface can dampen this background.

3.1 Main characteristics of the Muon Collider

The Muon Collider is quite unique between the proposed future machines. The signature that distinguish it from its peers is the use of muon beams, which implies several advantages and disadvantages. In fact compared to the e^+e^- used so far by lepton colliders, muons are heavier and so are much less subject to bremsstrahlung emission, making it easier to reach high energy and removing the power loss introduced by this effect. Also, being heavier, they require a small radius of the circular accelerator to reach the same energy of their light leptons' counterparts, which reduces the expenses needed to build the machine. Finally, as shown in figure 14 from [23], the Muon Collider becomes more efficient (i.e. has an higher number of useful collisions, called luminosity, per consumption of energy) as the E_{CM} increases, making it more and more convenient the higher the energy it is run, crowning it the "greener" lepton collider machine proposed so far.

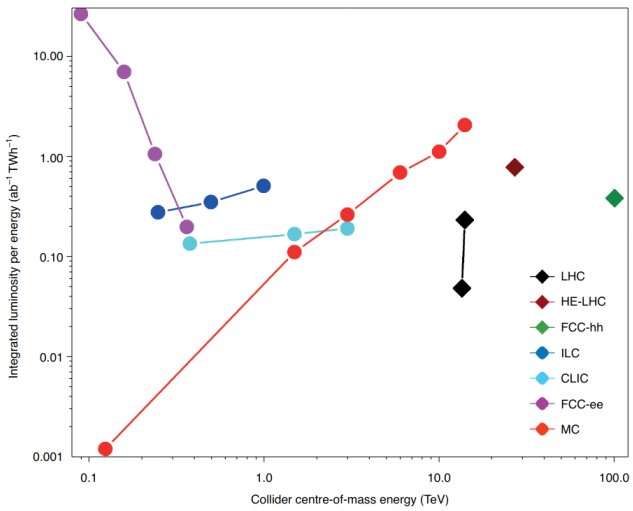


FIG. 14. Energy efficiency comparison in function of the the energy in the center of mass of several present and future colliders. It can be noticed that the higher the E_{CM}, the more efficient the Muon Collider becomes.

Compared to hadronic colliders, the advantage of the Muon Collider is the fact that the collision happens between muons which are elementary particles, and so all of the energy given in the acceleration

is used for interesting physics in the collision. On the other hand, hadrons are composite particles, and in high energy collisions only some parts of them, the so-called partons are interacting, which means that the energy that actually is available in the collision is just the fraction held by the partons, while the rest does not contribute. For these reasons, as shown in [15], the center of mass energy in which a muon collision has the same cross section of a process compared to an hadron collider, is much lower than the one needed for a proton collision. In short the Muon Collider can combine the high precision typical of e^+e^- colliders with the high $\sqrt{s}$ of the hadronic colliders.

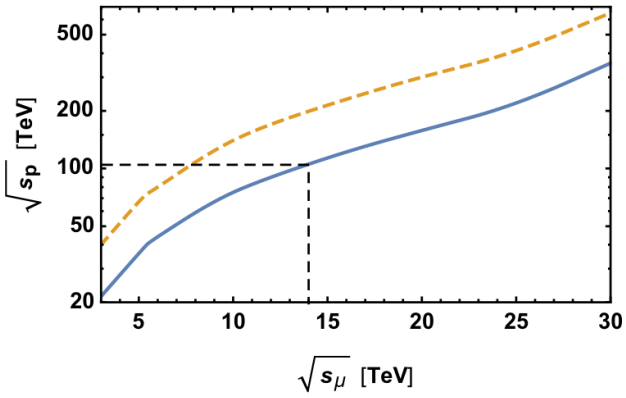


FIG. 15. Center of mass energy in which a proton collider cross section equals that of a muon collider. The blue line is the the cross section of a colored process, where hadron colliders get an increase to the the amplitude due to their constituents being charged under the string interactions. If the amplitude is comparable for the muons and partons of the hadrons, i.e. no colored object are produced, a muon collider reaches the same cross section at lower $\sqrt{s_\mu}$, as shown by the orange broken line.

As stated in the Introduction, the Muon Collider is a very good Higgs factory, as shown by the cross sections represented in figure [17]. At 3 TeV, the main Higgs production channel are the vector boson fusion (where the WW, figure [16].a, dominates over the ZZ, figure [16].b) and the Higgstrahlung, represented in figure , figure [16].c. Assuming an integrated luminosity of 1 ab^{-1}, at 3 TeV $\sim$ 500k Higgs are expected to be produced and this high statistic may allow to measure the most important Higgs parameters with great precision.

As shown in the previous chapter, the Muon Collider can measure the most important Higgs couplings, both with other particles as well as with itself, with unprecedent precision in the same data-taking time. The advantages of building such an experiment are now self-evident.
The main disadvantages of the Muon Collider are the technical difficulties needed to be overcome to have an intense enough muon beam and the background formed by the beam decay in the detectors. We're going to describe both of them in detail in the following sections.

3.2 Overview of the facility proposal

One of the biggest challenges is indeed creating muon beams with sufficient intensity. In fact they must provide an adequate amount of collisions from which physics is studied. This challenge arises because muons, unlike electrons and protons:

1. cannot be extracted from materials but **must be produced as tertiary particles through reactions**
2. are **not stable** since they decay with a lifetime $\tau = 2.2 \cdot 10^{-6}$ s in the rest frame.

So the Muon Collider facility must not only be able to produce muons, but also to uniform their energy and direction after their creation (this process is called cooling) and accelerate them very quickly to lose as less of them as possible.
The Muon Accelerator Program (MAP) collaboration [24] has performed deep studies on how such a

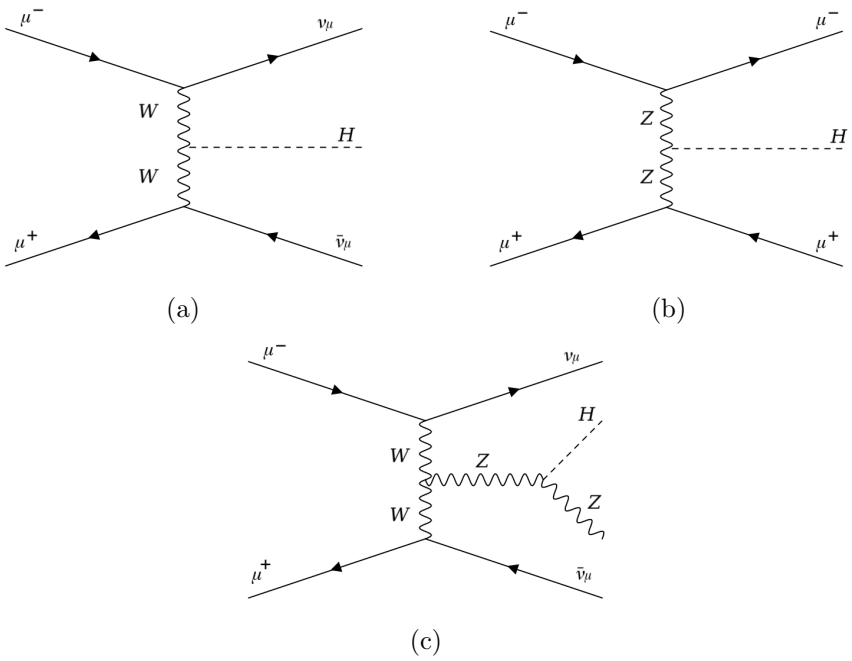


FIG. 16. Most important diagrams of the Higgs production processes at the Muon Collider. Respectively: (a) is the dominant diagram of the WW fusion; (b) the dominant diagram of the ZZ fusion; (c) the dominant (at 3 TeV) diagram of Higgstrahlung.

facility should be built. Muons can be produced as ternary particles, starting from an high intensity (order of MW) proton beam on a heavy material target, which collisions would produce a huge number of pions, which would then decay into muons. These muons are emitted in a large phase space and need to be collimated before injected in the accelerators rings. This is one delicate point: after the muons are accelerated their lifetime in the laboratory gets dilated by the the Lorentz boost, but at this point they are still relatively slow and so are decaying more rapidly in our frame; so the cooling must be done as quickly as possible to lose less muons. After this 6-D cooling is performed , the muons are quickly accelerated and sent into the collider rings. An overview of this facility project is showed in fig. **18**.

This facility operations can be divided in several stages:

- **The production and cooling stage**
  As previously stated muons cannot be extracted directly but must first be produced. To do this protons are accelerated in a super-conducting LINAC to reach 4 GeV, accumulated in a accumulator rings, constrained in 2 ns bunches and sent toward an heavy material target. This target is immersed in a strong solenoidal magnetic field (20 T in the target) directed along the longitudinal direction, which guides the pions produced by the interactions of beam and target in decay channels. This pions produce muons in their decay which are then bunched in Radio Frequency (RF) cavities, where also the higher energy bunches are decelerated while the late low energy ones are accelerated, and sent to the cooling stage. There currently 4 types of cooling techniques used in accelerators: synchrotron-radiation cooling, laser cooling, stochastic cooling and electron cooling. But none of them are suitable for this design: the first works only for low mass particles such as e^+e^-; the second is limited to ions and atomic beams; the third and fourth are too slow (between seconds and hours depending on the beam frequency spread and electron density available). So the proposed technique is an innovative one: ionization cooling. Ionization cooling uses two alternating kind of sections. The first is the transverse ionization cooling, which aim is to subtract the energy of the particles and then re-add some of the momenta but only in the longitudinal direction. It does this by making the muons interact with an absorber via ionization, where they lose momenta along all 3 directions, and then accelerated along the longitudinal one by some RF cavities. In this way one reduces the transverse momenta without diminishing too

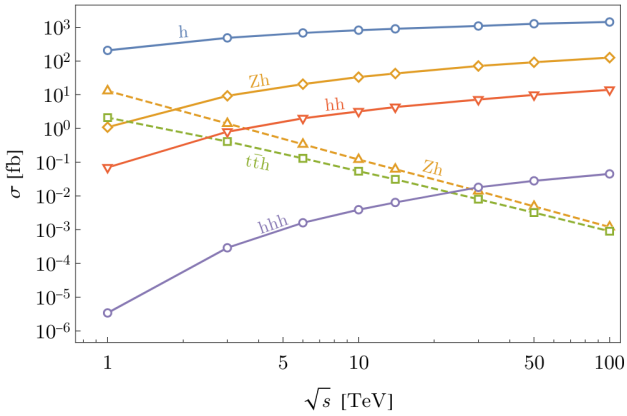


FIG. 17. Higgs production cross sections at a Muon Collider as a function of the energy in the center of mass $\sqrt{s}$. The solid lines represent Vector Boson Fusion processes, while the dashed lines indicate an annihilation process. For reference, assuming an integrated luminosity of 1 ab^{-1}, at 3 TeV $\sim$ 500k Higgs are expected to be produced. Taken from [15].

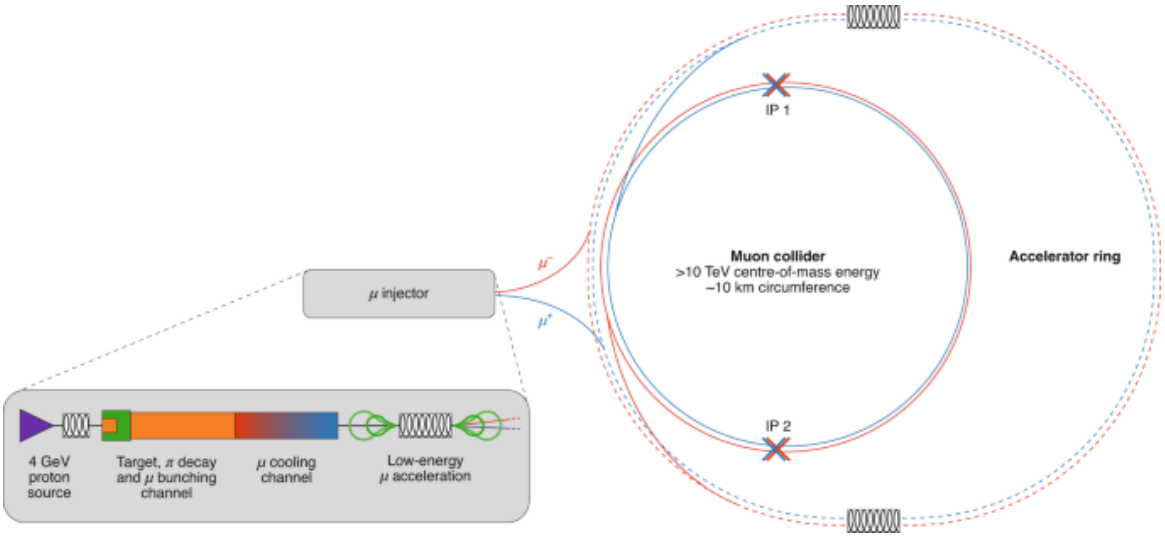


FIG. 18. Scheme of the muon facility proposed by the MAP collaboration.

much the longitudinal one: as a result the angular spread of the beam is reduced and this is obtained without slowing the beam too much and allowing an higher number of muons decays. A scheme of this process is shown in fig. 19.

The second section exploits the "emittance exchange" to reduce the longitudinal emittance. The beam passes through an absorber in a way that the more energetic components interact with longer lengths of material, while the low energy one travel a smaller distance. So the former lose more energy than the latter and so the energy distribution becomes more uniform. An example of this is shown in figure 20.

Alternating the two sections allows to reduce both the tranverse emittance as well as the longitudinal. The transverse ionization cooling has been tested by the MICE collaboration: their results agreed with the simulation. After including the emittance exchange sections, the muon beam 6-D emittance can be reduced by 5 orders of magnitude. After this quick cooling, the muon beam is ready to be sent to the first acceleration stage.

- **The acceleration stage**
  Due to the fact that muons decay, the acceleration needs to be as fast as possible. This can be achieved in stages by using RF cavities: However they are expensive and so ideally one would like to exploit each RF more than once: so small circumference of acceleration stages

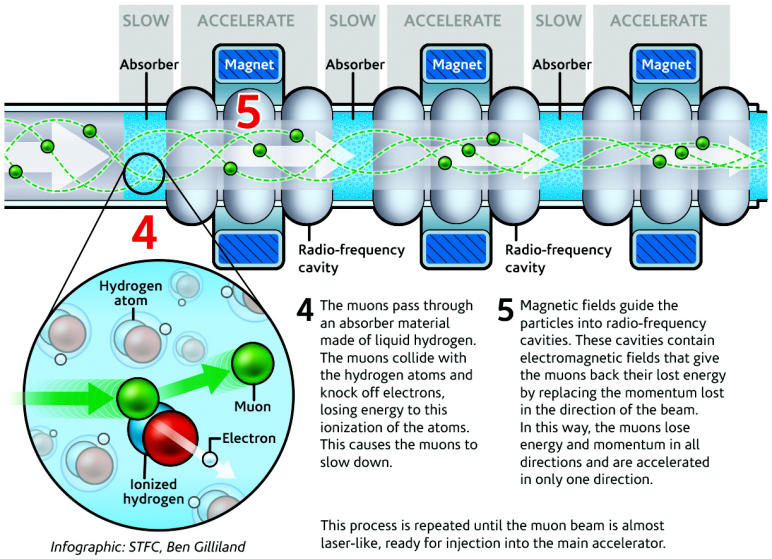


FIG. 19. Scheme of the ionization cooling used by the MICE collaboration. Figure from [25].

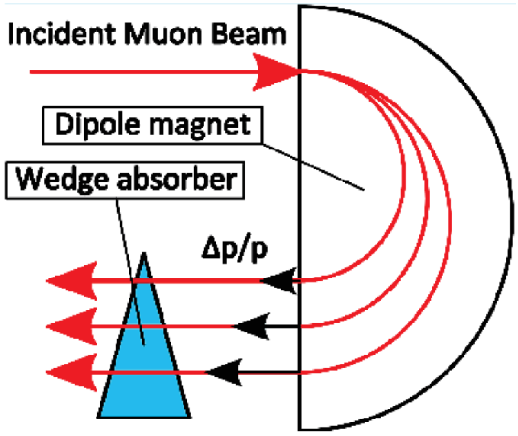


FIG. 20. Emitted exchange technique with a wedge absorber. The particles with higher momenta are less deflected by the magnetic field and pass through the wider parts of the absorber, interacting more with the material, while the slower particles lose less energy since they pass near the tip. A similar result is obtained without a wedge but by filling the magnet with gas absorber: once again particles with higher energy have less curvature and travel more in the gas. Image taken from [26].

are preferred. Several types of accelerator are currently under scrutiny: in the LINAC each RF cavity is used only once and so they are inefficient; Recirculating Linear accelerators (RLA) are multi-pass devices where the particles go in separate passes depending on their energy, but to di this the geometry must be carefully tuned, along with the need of beam focusing; Fixed Field alternating gradient Accelerator(FFA) are made by rings where the magnetic field is constant in time and exploiting strong focusing a single beam-line is used for many energies; Rapid Cycling Synchrotrons (RCS) are pulsed synchrotrons where magnetic fields are proportional to the beam momentum. This last kind is preferred because it allows higher RF frequencies and more turns in the accelerator, however it is also possible to look for a hybrid solution. At the moment no final choice has been made, but this stage will be deeply studied by the International Muon Collaboration.

- **The collider stage**
  Accelerated muon beams are sent to the detector, where physics data collection is performed. More particles arrive, higher statistics is available for the experiments and this is described by the luminosity L. In head-on collisions, with beams of opposite charge and of Gaussian transverse

distribution, the luminosity is given by:

$$L = \frac{f_{coll}}{4\pi} \frac{N^+ N^-}{\sigma_x^* \sigma_y^*} \tag{3.5}$$

To reach a luminosity of order $10^{35} cm^{-2} s^{-1}$, one needs to have an high collision frequency (f_{coll}) achievable by using small collider rings (currently limited by the magnitude of the dipole magnetic field available), high bunch populations (N^+N^-) which require fast acceleration lest many of the muon decay before reaching the interaction point (IP), and small transverse bunch area at the IP ($4\pi\sigma_x^*\sigma_y^*$). This last requirement is proportional on two beam parameters: the emittance ϵ, which describe the phase space area of the bunch, and the parameter β which describe the extremes of each particle oscillations around the trajectory defined by the rings magnet. Naively we should want to minimize both of them around at the IP position, however since a small β is obtained by an high gradient quadrupole magnet, this function has a parabolic behavior. For this reason, if β at the IP is minimized for the center of the bunch, the β of the head and tail of said bunch are much larger and this brings a net loss of luminosity. This is called "hourglass" effect, and to avoid it one must set a β no lower than the bunch longitudinal length σ_l. Moreover since the energy spread of the muons in the beam is quite large in the MAP scheme, the rings needs to be able to keep all particles and provide at least small momentum compaction factor. This rings also need to work with a very large spread of beam energies: from 1.5 TeV (to reach $E_{CM} = 3$ TeV) to 7 TeV (to reach $E_{CM} = 14$ TeV). Finally the particles produced from the muons decay provide several other challenges: they can interact with the accelerators magnets, heating them and making them lose superconducting capabilities; the background suppression in the detector requires a carefully built machine detector interface, as well as several mitigation strategies; their neutrinos can form radiation hazards far away from the facility. The last two points will be discussed in more detail in the next section.

Another idea to produce muons as non tertiary particles has been conceived some years ago: produce them from a e^+e^- collisions. This idea consists in sending a positron to a target which contains electrons, tuning their energy so to produce the muons at threshold (to do this the e^+ must be accelerated at 42 GeV). Hence the muon produced have a tiny phase space region available, which implies they have low emittance (have quite uniform directions and energies) and need very little cooling: in this scheme they can be accelerated much earlier and much less of them are lost due to decay before getting boosted. This idea is called LEMMA (Low Emittance Muon Accelerator) and the scheme of the proposed facility, compared with the one which extract muons from pions, is showed in figure 21.
However this scheme has its own share of technical problems. First and foremost, at such low energy the cross section of $e^+e^- \to \mu^+\mu^-$ is very small and so the muon production rate is very low. Which, combined with the fact that at the moment we don't have a very intense positron source, implies too little muons are produced. Also the target material choice is critical, since it must be dense to increase the production rate and to stand the heating and the mechanical stress caused by the beam, but also reduce multiple scattering interactions as much as possible. R&D around the world is working on this proposal, since the more powerful positron source and the target construction are synergic with other future e^+e^- colliders.

3.3 Beam induced background characterization

One of the biggest obstacles to the physics which can be studied at the Muon Collider is the Beam Induced Background, in short BIB. It originates from the decays of the muon happening while they travelling through the accelerator. This not only implies a loss in luminosity since muons are lost, but also the emission of other high energy particles emerging as its decay products: $e\nu\bar{\nu}$. For 750 GeV beams, the muon decay length is $4.7 \cdot 10^6$, and assuming a bunch intensity of $2 \cdot 10^{12}$ particles, are expected on average $4.3 \cdot 10^5 \mu$ decays/m. It is important to notice that the higher the energy of the beams, the bigger the Lorentz boost γ becomes and so muon decay decreases.

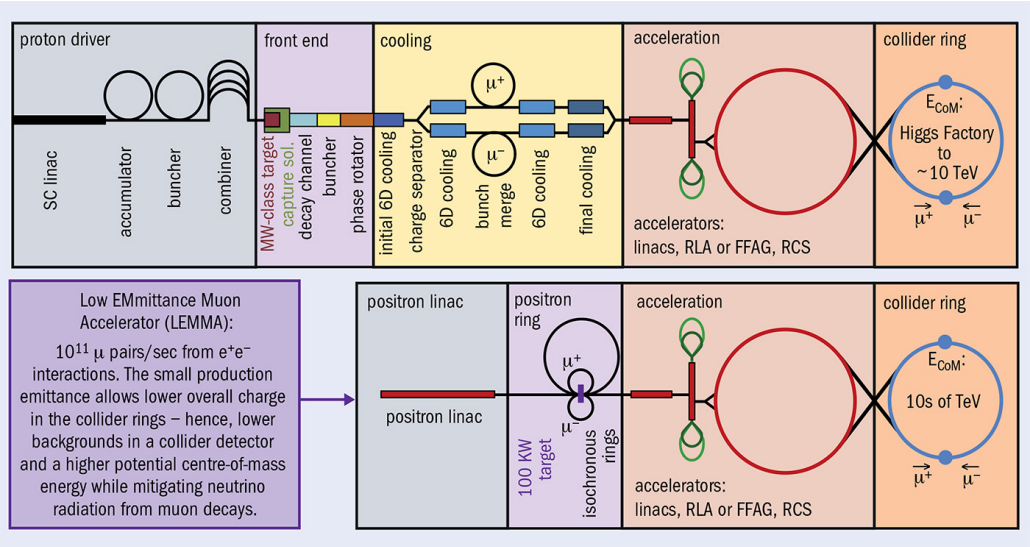


FIG. 21. Scheme of a muon facility producing muons from the π decay (upper row) versus scheme of a muon facility exploiting the muon production from a e^+e^- pair at dimuon energy threshold (lower row).

The μ decay in electrons (or positrons) and pairs of neutrinos. This two types of decay products have different effects: the former is the one creating background in the detector, which will be tackled in this section, while the latter two can exit from the machine and create radiation hazards, and this will be discussed in the next section. So for the physics studies itself, the most troublesome part of the BIB originates from the electron (or positron) of the muon (or antimuon) decay. In fact the high energy light leptons not only emit photons, but also both them and the γ interact with the accelerator material producing a huge number of other particles. In particular in this thesis we will work with the simulated BIB at E_{CM}=1.5 TeV obtained with MARS15.
Such choice does not compromise the studies we will do at E_{CM}=3 TeV because the higher the energy of the beams, the smaller (less muon have time to decay) and more focused along the beamline the BIB is expected to be. So although we cannot use the BIB at E_{CM}=3 TeV because it is currently under study, we are using a configuration with higher background and it is expected that the results of an experiment with the BIB at the correct energy will be more precise than the ones obtained in this thesis, which can then be considered conservative estimations.
The interactions of the electrons with the materials of the accelerator produces several kind of particles:

- electrons and γ, from μ decay, synchrotron emission and interactions with accelerator components
- neutrons, from the interactions with accelerator components
- charged hadrons, from the interactions with accelerator components
- Bethe-Heitler muons and antimuons, coming from pair production

Several studies on the characteristics of such particles have been performed both by using MARS15 and FLUKA [27]. In particular [28] simulated with MARS all the produced particles and their interactions for a single beam: this has been done within 25 meters from the IP for all types of particles except the Bethe-Heite muons, which have not negligible number up until 100 m from the IP. The multiplicity of production of each species can be seen in the figure 22.

To mitigate the BIB, the Machine Detector Interface has been optimized by the MAP collaboration, also by introducing another key component of the machine: in the detector two nozzles are placed along the beam-pipe to absorb as much background as possible. This components, combined with an optimization of the magnets layout and iron and concrete shielding before the detectors, limits the detector acceptance in the forward and backward region, but suppress the number of BIB particles entering the detector by a factor 500 and drastically reduce their energy spectra. The MDI can be

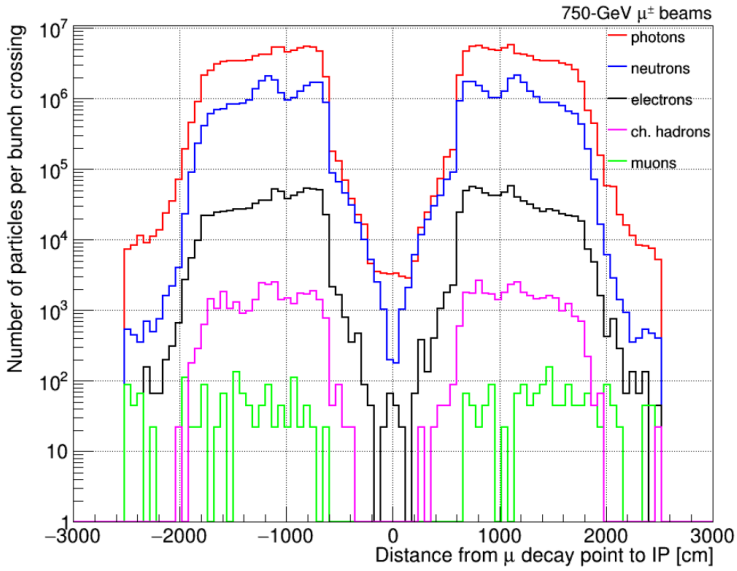


FIG. 22. Decay point of μ producing that particle and multiplicity of the species composing the BIB in 0.75 TeV beams. This picture shows only the particles that arrive to the detector after the MDI mitigation strategy explained in the previous section.

seen in figure 24
The geometry of the nozzles (in fig. 23) is the results of many optimization studies [29] in order to stop as much BIB as possible without limiting too much the detector acceptance. Its closest point to the IP is at Z=6mm, R=1mm: at smaller values of Z only a thin slice of Beryllium is installed, since form this region will originate the signal particles. The nozzle then proceeds with an inclination angle in the Z,R plane of 10° up until 1 meter from the origin where it becomes less step switching to a 5° degrees angle in the Z,R plane. The nozzles have cylindrical symmetry with respect to the z-axis. It's innermost part is made of tungsten, which interacts with the BIB γ producing neutrons, which are then stopped by the Borated Polyethylene (BCH_2).

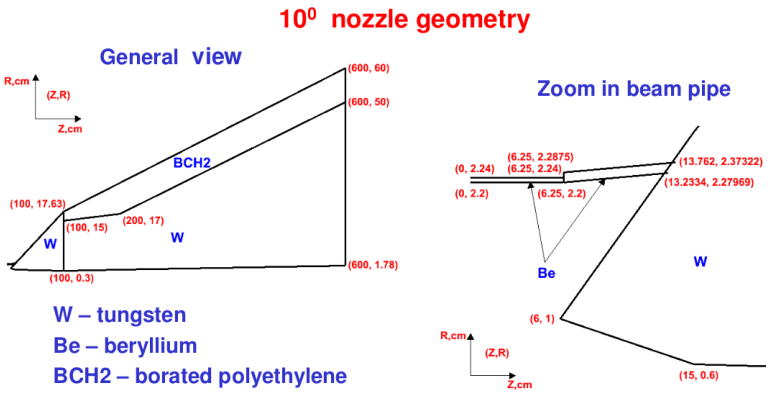


FIG. 23. Nozzle geometric characteristics.

Even after the optimization of the MDI, lots of BIB particles reach the detectors. By studying their characteristics one can implement other ways to mitigate their effect. In particular the most useful ones are:

- Entry time in the detector (fig. 25): the decay products lose time and speed in their interaction with the accelerator materials, so a big fraction of them arrive to the detector delayed.
- Momenta (fig. 26): BIB particles lose energy in their interactions, so they have a soft energy and momenta spectrum.
- Entry point in the detector and flight direction: most BIB particles do not enter in the detector

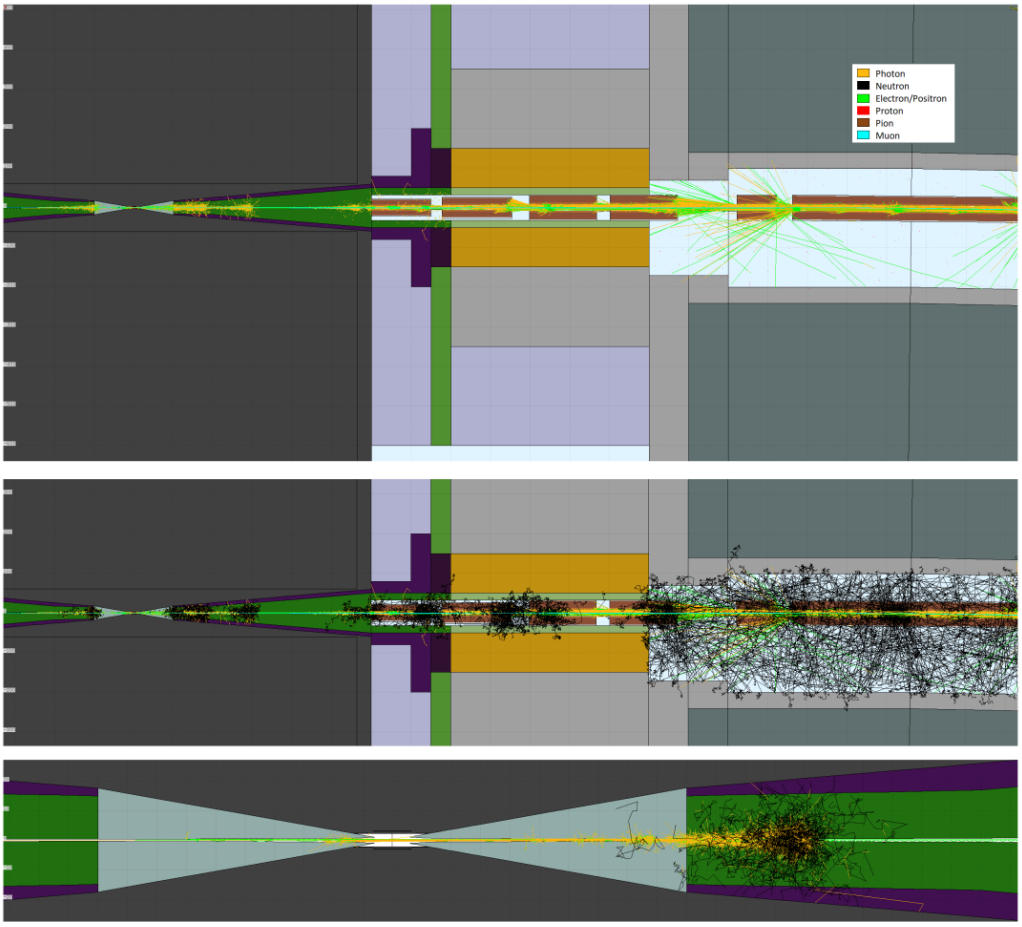


FIG. 24. Depiction of the MDI region and the BIB tracks from some μ decays. On the left the dark grey represents the detector and the cones extending inside it are the nozzles (the central tungsten part is coloured in green, while the outer sides, in BCH_2, is pictured in violet), while the magnets along the beamline are illustrated in brown. On the upper picture, neutrons are removed, while on the middle one they are kept and represented with black tracks. Finally in the lower image a close up in the part of the nozzle closer to the IP. Image taken from [27].

near the IP and they are mostly directed toward the beam longitudinal direction.

Due to the nature of the BIB, the occupancy in the detectors grows the further we get closer to the beams. For this reason, mitigation strategies must be applied at detector level: these will be studied in detail in chapter 4.2, after the detector is presented.

3.4 Neutrino radiation hazard

The neutrinos produced by the decay of the μ along the beams are not seen by the detector and so do not provide a background during physics experiments, but they can exit from the accelerator complex, travel through the ground and interact with it. The secondary particles produced from these processes can emerge from the surface and produce radiation hazards. The magnitude of the dose produced from them, how it scales with energy and some mitigation strategies have been studied [28] both with analytic computations and MARS15 [30] simulations. The neutrinos from decays are emitted mostly in a cone with axis the beam tangent direction and of spread given approximately by the $\frac{1}{\gamma}$ of the parent muon. If the some part of accelerator was linear, all the parts where the beam has the same direction would focus their dose in the same angular region, but this is mitigated by the fact that the accelerator is circular. This also implies that the neutrino radiation is concentrated on the machine plane. In a circular accelerator, the dose scales approximately as E^3, E being the muon energy: a factor E due to the energy deposit, one to the spot size ($\frac{1}{\gamma}$) and the last one to the neutrino cross section. An idea to mitigate the dose is to made the neutrino be emitted more spread out by

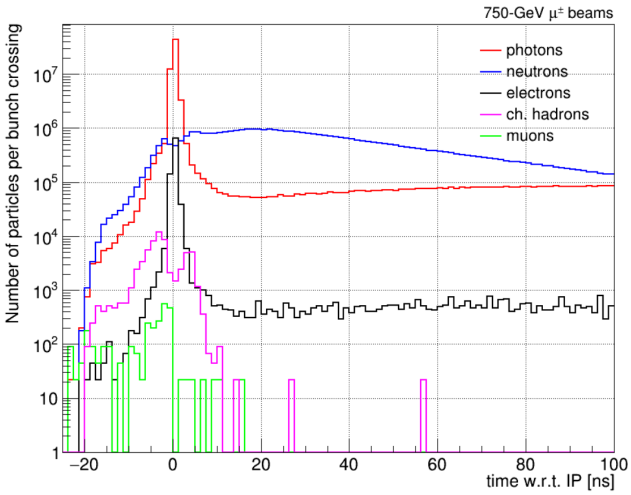


FIG. 25. BIB particles arrival time at the detector w.r.t. the bunch crossing.

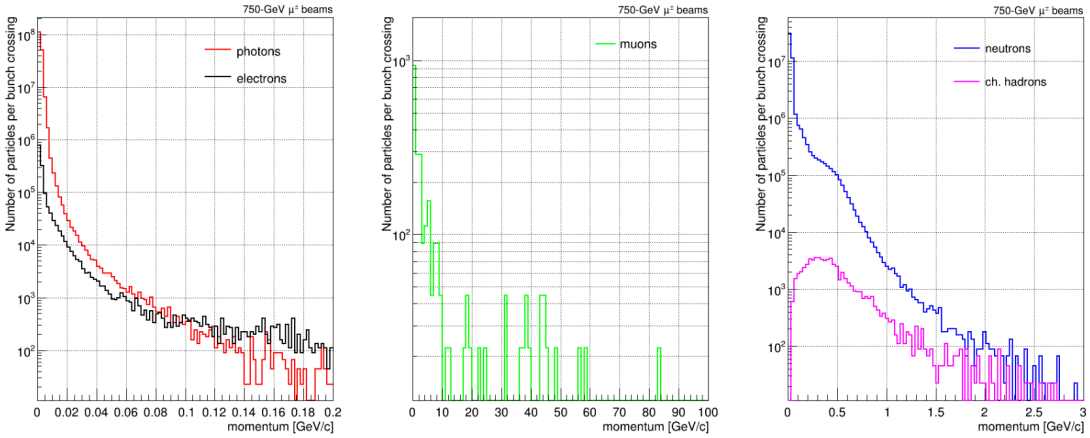


FIG. 26. BIB particles momenta spectra.

introducing some vertical wobbling on the beam exploiting the magnets (see fig. **27**): by periodically shifting the wobbling over time one can reduce the dose by one order of magnitude (fig. **28**). These studies show that with careful planning in the construction of the accelerator, radiation can be kept under control.

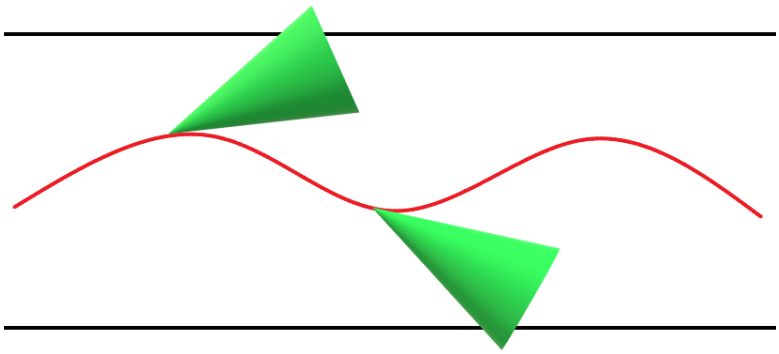


FIG. 27. Schematic visualisation of the wobbling: alternating the vertical direction of the beam (red line) changes the direction of the cone where ν are emitted (represented in green) . If no wobbling is used, the cone of emitted neutrinos along several points of the beam will be directed in the same direction and produce a bigger equivalent dose in a small area.

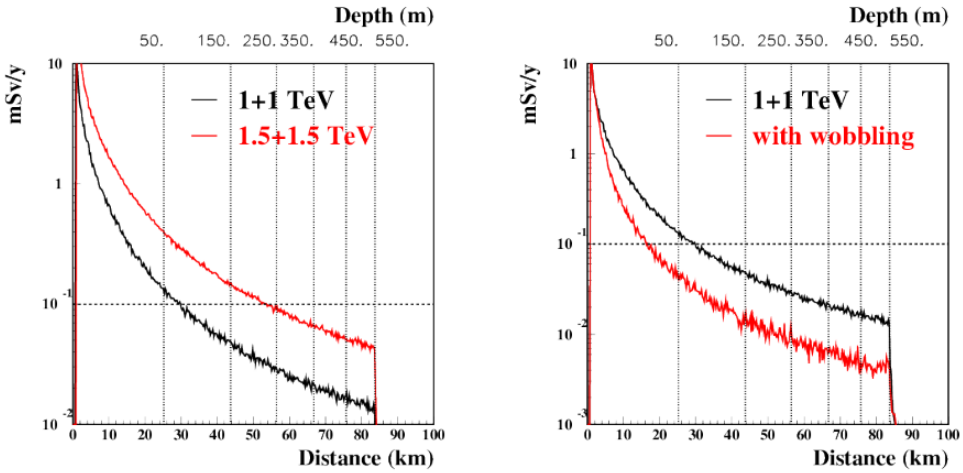


FIG. 28. Left: Ambient Dose equivalent for 1 TeV and 1.5 TeV beams as a function of the distance at which the ν reach the Earth surface (which depend on the depth at which the accelerator ring is built). Right: Ambient Dose equivalent for 1 TeV beams with and without wobbling. In both pictures an horizontal dashed line shows the desired radiation goal of $10^{-1}\,\frac{mSv}{y}$

4. Detector description

In this chapter, the detector used in our simulation to extract the cross section of the H$\rightarrow b\bar{b}$ and characterize the jets will be described. After a brief overview of the whole detector, all subdetectors will be described with all their most important parameters, starting from the tracking system, then proceeding with the calorimeters and lastly concluding the muon detector. This description will be integrated by the contents of the third section, showing the strategies employed on each subdetector to reduce the noise caused by the beam induced background.

4.1 Detector overview

The Muon Collider detector used as starting basis the proposed detector by CLIC, adapting some components to better adaapt to its unique beam induced background. As such, it is resembling a 4π detector, except for its unique nozzles described in chapter 3.3. The overall detector has a cylindrical shape with axis along the beam direction, having a radius of 6.45 meters and being 11.276 meters long with an internal magnetic field of B=3.57 T, while on the magnet return yoke the field is B_{yoke}=-1.34 T. Starting from the innermost subdetector and going outward, one finds:

- the Tracking system: composed by the barrel and endcap layers of the Vertex Detector (VXD), of the Inner Tracker (IT) and of the Outer Tracker (OT);
- the Calorimetry, made by an electromagnetic calorimeter (ECAL) and an hadronic calorimeter (HCAL)
- the solenoid and muon detector: after the superconducting solenoid an iron return yoke with some layers of Resistive Plate Chambers (RPC) detect the passage of the muon escaping from our system.

In the rest of the thesis we will describe everything with the following coordinates: the origin is set at the interaction point, with the z-axis directed along the μ^- beam direction, the y-axis parallel with gravity and x-axis orthogonal to the previous ones. A scheme of the full detector is shown in figure **29**.

4.1.1 Tracking detectors

The moment of charged particles travelling in the detector is measured by the tracking detectors. In fact, the charged particle will have helix-like trajectories due to the magnetic field and, during this motion, they will pass through pixels of the tracking detectors. Doing so, they lose energy in them and leave marks of their passage, forming hits, which can then be used to reconstruct their trajectory and measure the momenta.
Each of the tracking detectors has several layers arranged in two different regions exploiting the cylindrical symmetry of the system: the barrel region and the endcap region. The barrel layers are shaped like the lateral surface of a cylinder, they are coaxial with the beam direction and are centered along z at the interaction point. They are used to track particles in the central region of the detector. On the other hand, endcap layers are annulus placed perpendicularly to the beam direction. They are disposed along z after the end of the barrel layers symmetrically for both z<0 and z>0 regions. They are used to reconstruct the tracks of the particles emitted in the forward or backward direction.

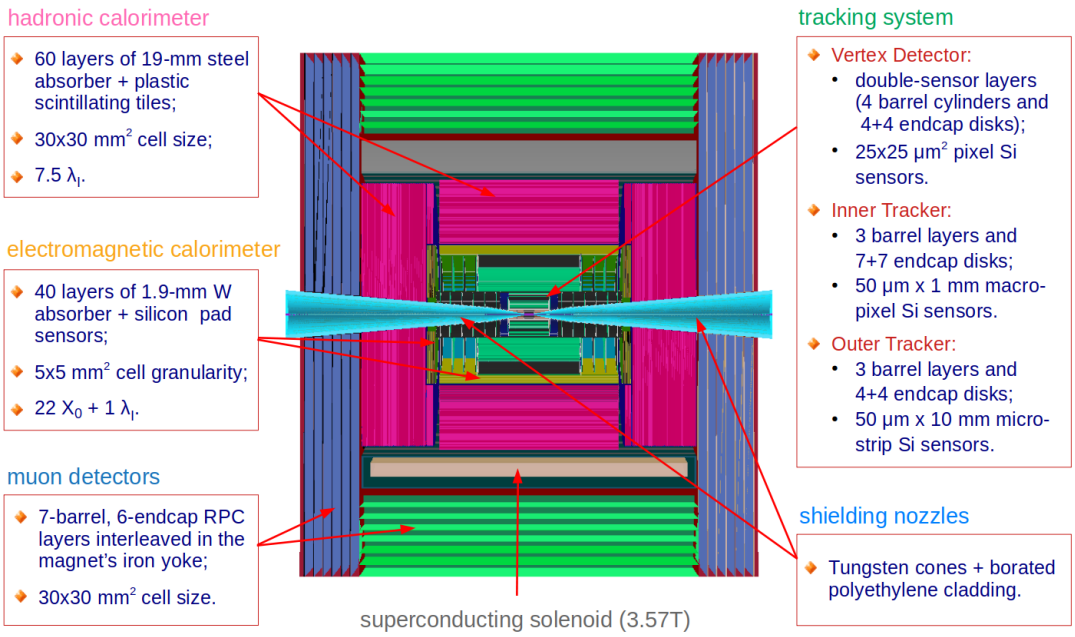


FIG. 29. Schematic drawing of the detector used in these simulations with a brief description of its components. Image taken from [31].

The specific dimensions of each layer depend on which detector they are in as well as how they are placed in the detector itself. A schematic view of the overall system is presented in figure **30**.

The resolution of each detector is taken from results of test beams of similar silicon devices. This information is employed to decide the smearing of each hit in the tracking system in our simulation. The time windows of sensitivity of each detector will be reported during the discussion of the BIB mitigation strategies at detector level in the next section 4.2.

Vertex Detector

The Vertex Detector is placed very close to the interaction point in order to be able to reconstruct precisely the tracks positions near their origin and so reconstruct secondary vertices. This is very important because secondary vertices are signatures of decay of b- and c-hadrons and are one of the main tools we will use to tag heavy flavour jets. The barrel subdetectors are made of 50 μm thick silicon rectangular modules of 25 μm x 25 μm pixels. They allow a time resolution of 30 ps. Each of these modules is double layered, i.e. it has one module of pixel on the IP side of the layer and one toward the outer side, separated by a gap of 2 mm. The double layers are extremely useful because, as will be seen in the next section 4.2, the BIB has the highest occupancy in the VXD detector and to reduce the multiplicity of hits the tracking algorithm has to deal with, information on the direction between the two hits in the same module is exploited: it is required that the second hit in the DL is not too displaced from the direction determined by the line passing from the IP and the first hit. In the assignment of hits to each tracks, the 2 hits are counted separately. This helps mitigate a huge part of the BIB, since unlike signal particles, it does not originate from the IP. The geometric parameters of each of the 4 shells making up the barrel are shown in the following table **1**.

Shell	Radial distance from IP [cm]	Length along z [cm]	Number of modules
1	3.0	13.0	16
2	5.1	13.0	15
3	7.4	13.0	21
4	10.2	13.0	29

TAB. 1. Geometric parameters of the layers of the VXD in the barrel region.

The assumed spacial resolution is 5 μm x 5 μm.

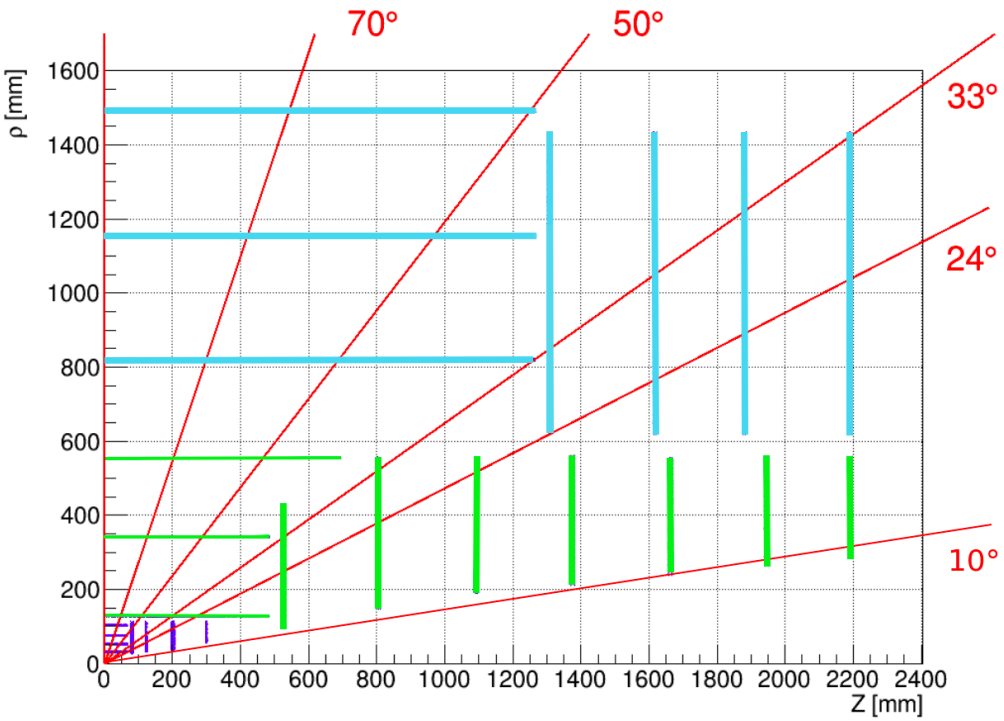


FIG. 30. Section of the tracking detectors in the forward region. The back region is symmetric. Layers have different colors depending on the subdetectors: violet for the VXD, green for the inner tracker and light blue for the outer tracker. The horizontal lines are a section of the barrel layers, while the vertical ones represent the endcap. The double layers in the VXD detector are not visible in this picture. The part at low ρ without any layer is occupied by the nozzle.

The endcap modules are trapezoidal and arranged as "petals" to form the disks. As for the barrel, double layers (separated by the 2mm gap) and time cuts are used to reduce the huge occupancy due to the BIB. The silicon pixels are of the same kind of the barrel: squared pixels of 25μm x 25μm with a thickness of 50 μm and 30 ps of time resolution. The layers are disposed symmetrically on $z < 0$ and $z > 0$, 4 for each side, so we can just use $|z|$ as a parameter to describe the layers placement in table 2.

Disk	Position along beamline $\|z\|$ [cm]	Inner Radius [cm]	Outer Radius [cm]	Number of modules
1	8.0	2.5	11.2	16
2	12.0	3.1	11.2	16
3	20.0	3.8	11.2	16
4	28.0	5.3	11.2	16

TAB. 2. Geometric parameters of the layers of the VXD in the endcap region.

The spacial resolution expected from the system is 5 μm x 5 μm, as for the barrel.

Inner Tracker

The Inner tracker exploits for both barrel and endcap silicon macropixels of size 50 μm $\times$ 1 mm. For the barrel layers the long side is parallel to the beam axis, while for the endcap ones it is parallel with the radial direction. They are also thicker (100 μm) and have a slightly worse temporal resolution of 60 ps. Also they don't use double layers, unlike the VXD. The looser time resolution, bigger pixel size and absence of the double layers are possible because the occupancy of the BIB in this region is much lower than the VXD. The barrel is made up by 3 layers with rectangular modules, which geometric characteristics are reported in table 3 .

The endcap regions are still symmetrical for positive and negative z, each presenting 7 layers made by rectangular modules, whose geometry is illustrated in table 4.

Shell	Radial distance from IP [cm]	Length along z [cm]	Number of modules
1	12.7	96.32	28
2	34.0	96.32	76
3	55.4	138.46	124

TAB. 3. Geometric parameters of the layers of the IT in the barrel region.

Disk	Position along beamline $\|z\|$ [cm]	Inner Radius [cm]	Outer Radius [cm]	Number of modules
1	52.4	9.5	42.7	26
2	80.8	14.7	55.8	26
3	109.3	19.0	55.6	26
4	137.7	21.2	56.1	26
5	166.1	23.7	55.7	26
6	194.6	26.4	55.4	26
7	219.0	28.4	55.8	26

TAB. 4. Geometric parameters of the layers of the IT in the endcap region.

For both endcap and barrel it is expected a position resolution of 7 μm along the smaller direction of the macro pixel and 90 μm along the longer one.

Outer Tracker

The Outer tracker uses silicon microstrips instead of pixels: they are 50 μm long, 10 mm wide and 100 μm thick. The direction of the longer sides are the one already reported for the IT. The time and the spacial resolution are respectively 60 ps and 7 μm x 90 μm , as the IT. No double layers are used. The barrel region is made up by 3 layers of rectangular modules with the geometry described in the following table 5.

Shell	Radial distance from IP [cm]	Length along z [cm]	Number of modules
1	81.9	252.84	184
2	115.3	252.84	256
3	148.6	252.84	328

TAB. 5. Geometric parameters of the layers of the OT in the barrel region.

The endcap region is composed by 4 disk for each sign of z, as shown in table 6, made by rectangular modules.

Disk	Position along beamline $\|z\|$ [cm]	Inner Radius [cm]	Outer Radius [cm]	Number of modules
1	131.0	61.7	143.0	48
2	161.7	61.7	143.0	48
3	188.3	61.7	143.0	48
4	219.0	61.7	143.0	48

TAB. 6. Geometric parameters of the layers of the OT in the endcap region.

4.1.2 Calorimeters

Calorimetry is used to measure the energy of a particle by their energy loss in the calorimeter material. Calorimeters are essential because they allow to measure also neutral particles, which are not seen by the tracking system. As inherited from the CLIC design, the Muon Collider uses an electromagnetic calorimeter (ECAL) followed by an hadronic calorimeter (HCAL). What follows is the configurations that have been used during this thesis.

The ECAL is composed by a dodecagonal barrel and two endcaps. Each is made of 40 layers of tungsten absorber (each 1.9 mm thick) intermitted by silicon pads sensor layers. This corresponds to

22 radiation lenghts X_0 or 1 interaction length λ_l. Each of its cells is 5mm x 5mm, and this granularity allows for better BIB rejection: BIB has a soft energy spectra and so loses most of its energy in the first layers. Geometrically the barrel has an inner radius of $r_{inn} = 1500$mm and an outer radius of $r_{out} = 1702$ mm. Along Z it is $l = 4420$ mm long. Instead, the endcap radius is $r = 1700$ mm, and, longitudinally, its starting point is $z_{min} = 2307$ mm and ends in $z_{max} = 2509$ mm.

In future an alternative ECAL could be CRILIN [32] (CRystal calorimeter with Longitudinal Information), semi-homogeneous calorimeter where Cherenkov light is emitted from particles passing through the PbF_2 absorber and read by SiPM. A prototype of the cells of CRILIN is currently being tested in LNF ("Laboratori Nazionali Frascati").

The HCAL follows the same symmetry. It uses 60 layers of Steel absorber 19 mm thick alternated with plastic scintillator tiles of dimensions 30mm x 30mm. This corresponds to 7 interaction length λ_l. The HCAL is shielded by ECAL, but still some mitigation strategies must be applied: the signal showers go deeper inside the calorimeter than the background ones and so longitudinal shower analysis will be key in reducing the BIB, as well as time cuts (the same of the ECAL) and an energy threshold of 0.25 MeV is set on each pixel. The barrel is designed with an inner radius $r_{inn} = 1740$mm, an outer radius $r_{out} = 3300$mm and a longitudinal lenght of $r_{min} = 4420$mm. The endcap starts at $z_{min} = 2539$ mm and ends in $z_{max} = 4129$ mm with a radius in the transverse plane of $r = 3246$ mm. Drawings of the calorimeters are shown in figures **31** and **32**: the first one shows both calorimeters with the left an event of H→bb without the BIB, while the second one shows only the ECAL with the information used by the PandoraPFA algorithm (chapter 5.2).

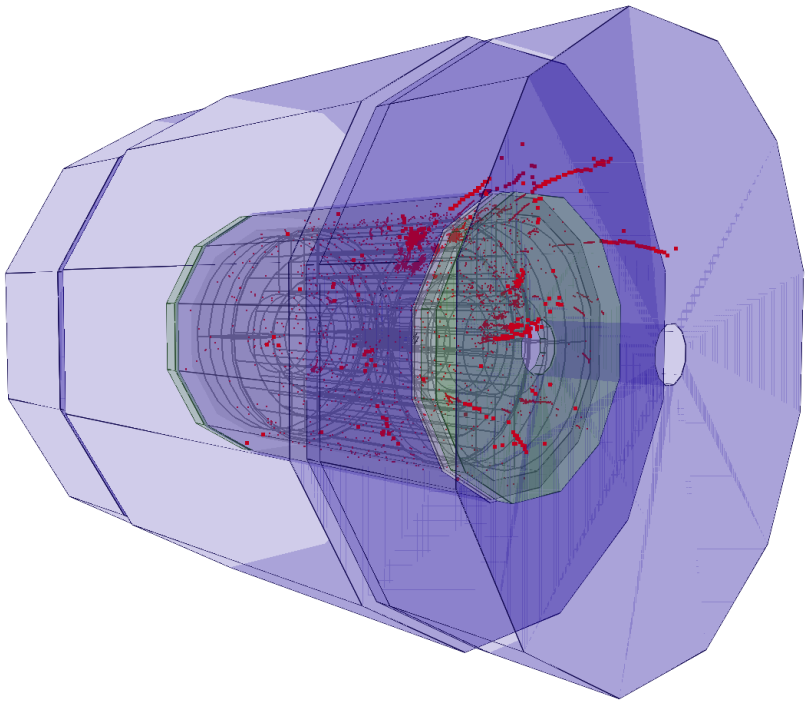


FIG. 31. Image of the calorimeters used in the simulation, with superimposed in red the hits in the calorimeters by an event of H→bb without the BIB. The ECAL barrel is represented in light grey, while the ECAL endcaps are shown in green and the HCAL in violet. The central grey lines are mechanical supports. Image made with [33].

4.1.3 Muon detectors

The muon detectors are installed in the iron return yoke of the solenoid. As the name suggests, they are used to identify and measure the momenta of muons, and they can do this thanks to their external position in the overall detector. In fact any other particle coming from the detector would stop in the hadronic calorimeter or before, while only muons are penetrating enough to reach the muon detector. This provide their identification, while their momenta is measured thanks to the trajectory

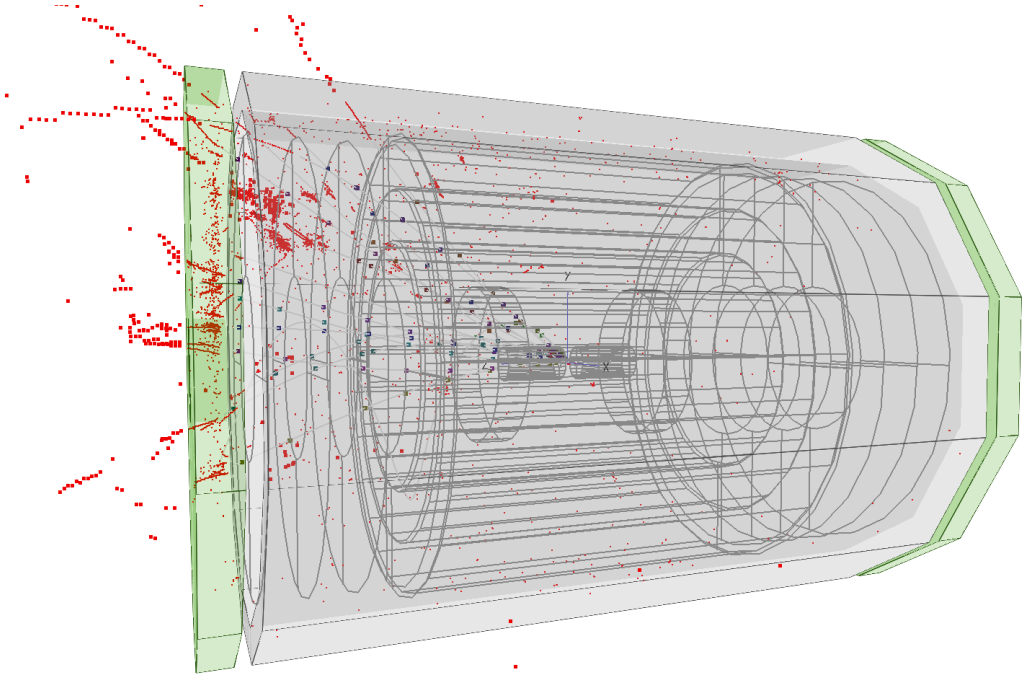


FIG. 32. Representation of the ECAL barrel (light grey) and endcaps (green) with the calorimeter hits (red) and the tracker hits (colored in dark blue, dark green or dark purple) associated to a reconstructed particle by PandoraPFA (chapter 5.2). One can notice that the hits in the ECAL are smaller than the ones outside it (where the HCAL would be) due to the bigger granularity of the former. The central grey lines are mechanical supports. Image made with [33].

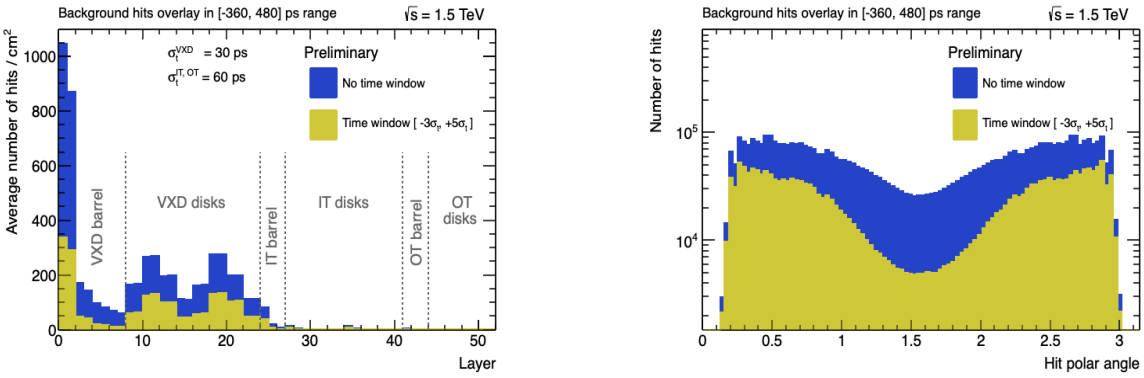
deflection in the magnetic field. The detector itself is made by layers of Glass Resistive Plate Chamber (RPC), which give an electric signal when the particle passes, signaling their position. The geometry is different for barrel and endcaps:

- **Barrel**: extends from a radius of 4461 to 6450 mm, has $|z|$ <4179 mm, with 7 layers of sensors, each with an area of 1942 mm^2. They are made of 30x30 mm^2 cells and have a total of $2.2{\cdot}10^6$ channels.
- **Endcap**: extends from a radius of 446 to 6450 mm, has 4179< $|z|$ <5700 mm, with 6 layers of sensors, each with an area of 1547 mm^2. They are made of 30x30 mm^2 cells and have a total of $1.7{\cdot}10^6$ channels.

4.2 Beam induced background mitigation strategies

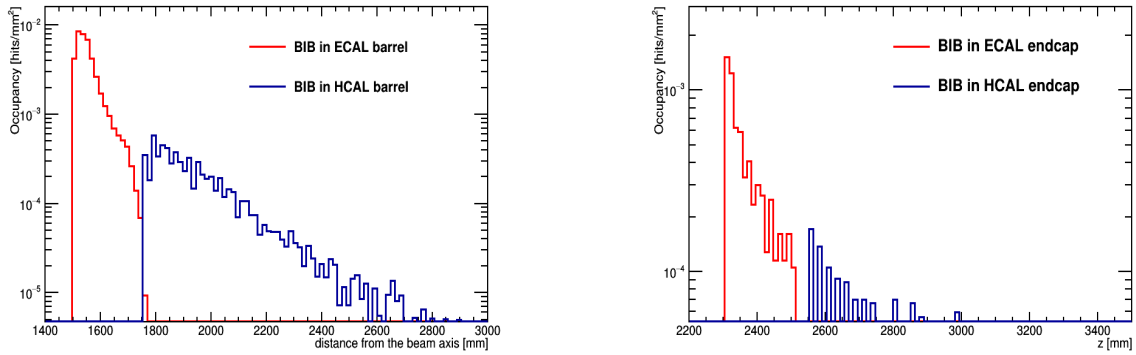
Now that the detector has been described, it is necessary to see the effects of the BIB on each of its parts. As stated in chapter 3.3, the detector occupancy grows the closer one gets to the beampipe: this is shown in figures **33**, **34** and **35**. In these figures it is evident not only that the BIB is leaving much more hits in the tracking system (and in particular in the Vertex Detector) than in the outermost subdetectors, but also that even inside each subdetector the occupancy drastically change: for example the ECAL first layers, have an higher occupancy than the deeper layers or the HCAL, which are de-facto shielded by the first layers. It follows that each subdetector has different requirements depending on the type of signal it is tasked to collect and on the level of background. This is reflected in different mitigation strategies, implemented at detector level:

- **Tracking detectors**: timing cuts are employed to reject out of time background. As of now, the cuts are the same for barrel and endcap and are shown in table 7. Even after this requirement, the vertex detector has still a very high occupancy. Another strategy is needed in order to reduce the hits of the BIB particles and this is done by exploiting their entering position in the detector and their direction thanks to the Double Layer Filter. Each layer of the vertex detectors is actually made by two active parts, forming double layers. In this way, the particle



(a) (b)

FIG. 33. (a) Occupancy of the layers of the tracking system. The first layers of the barrel part of the vertex detector are very close to the beam pipe and have the highest occupancy. Due to the direction of the BIB, also the endcap disks of the vertex detector are very much affected by the BIB. The geometrical characteristics of each layer are given in Chapter 4.1.1. (b) Number of hits as a function of the angle with the μ^- beam direction. The central region is less affected by the BIB. Image taken from [31].



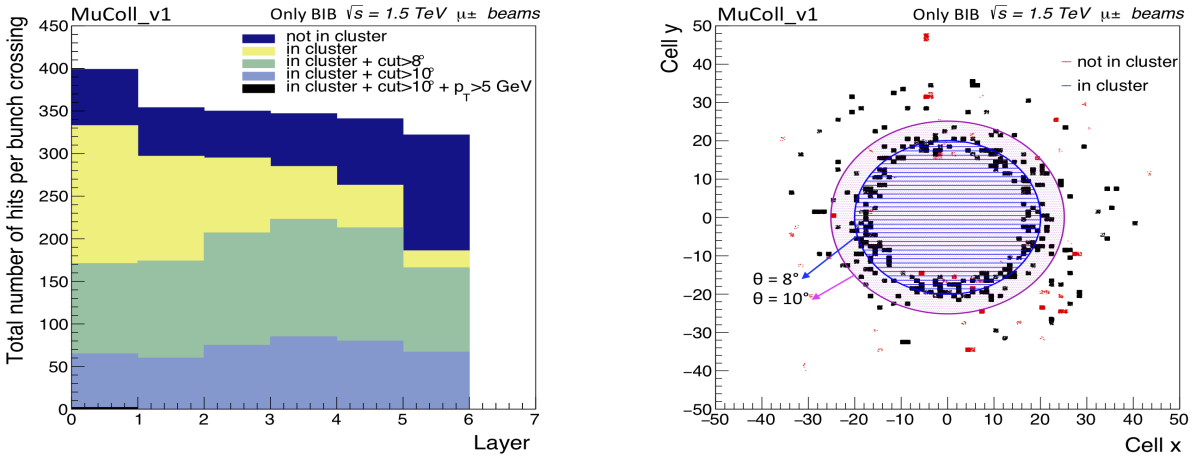
(a) (b)

FIG. 34. (a) Occupancy in the barrel region of the calorimeters. Since the BIB is mostly made of soft particles, the deepest layers of the HCAL are barely reached by any BIB-originated showers. (b) Occupancy in the barrel region of the calorimeters. The same consideration of (a) holds here. Image taken from [31].

Detector	Start of the window [ns]	End of the window [ns]
Vertex detector	−0.18	0.24
Inner tracker	−0.36	0.48
Outer tracker	−0.36	0.48

TAB. 7. Acceptance time windows w.r.t. bunch time for the different tracking detectors.

position is known in two very close points and this allows to deduce the particle travel direction. If a particle is coming from the signal, it will come from the IP direction, leaving two hits in the double layer that will point to that direction. However if it's coming from a decay of the beam, it will come from completely another direction. This is shown in figure 36.a. In practice the Double Layer Filter puts two requirements on the two consecutive hits (if there are no close consecutive hits, the single hit is discarded), one in the longitudinal-radial plane and one on the transverse plane. On the longitudinal plane, it requires the outermost hit to lie within a cone of angle $\Delta\theta$ and axis pointing to the IP. If there is no second hit in this area, the first hit is discarded. A similar requirement is applied on the transverse plane: signal particles have higher p_T and so are less deflected by the B. One can discard much of the BIB by putting an upper bound on the displacement in the ϕ direction, called ΔU, between the two hits. If no second



(a) (b)

FIG. 35. (a) Number of hits per bunch crossing in the layers of the endcap of the muon detector. The number of hit in the barrel region is negligible. Several cuts on the angle (the nozzle has an angular aperture of 10°) and p_T. (b) BIB hit distribution in the first layer of the endcap of the muon detectors. The positions are given in cm. Image taken from [31].

hit is found within ΔU, the first hit is discarded. In tables 8 and 9 the settings used in this simulation are reported as a function of the angle in the longitudinal plane $\Delta\theta_{max}$ and of the displacement ΔU_{max} in the transverse plane.

Barrel layers	ΔU_{max} [mm]	$\Delta\theta_{max}$ [mrad]
0 and 1	0.55	0.3
2 and 3	0.55	0.2
4 and 5	0.5	0.15
6 and 7	0.4	0.12

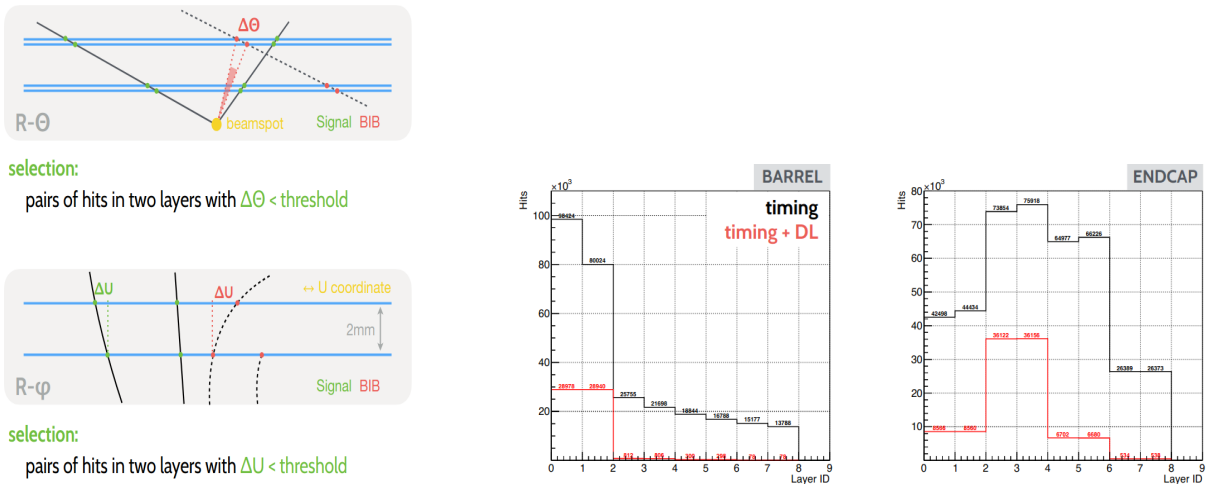
TAB. 8. Double layer filter parameters for the Vertex detector sensors in the barrel region.

Endcap layers	ΔU_{max} [mm]	$\Delta\theta_{max}$ [mrad]
0 and 1	0.7	0.11
2 and 3	0.7	0.09
4 and 5	0.4	0.06
6 and 7	0.3	0.042

TAB. 9. Double layer filter parameters for the Vertex detector sensors in the endcap region.

This discards a lot of hits and greatly improves the CPU time needed to reconstruct an event with BIB overlayed, as shown in figure 36.b. However this may cause problems to displaced tracks, coming from SV. This will be studied in detail in chapter 5.3. Still, during the tracking, the multiplicity of remaining hits is too big for any tracking algorithm to work with in acceptable time. For this reason, the regional tracking technique is employed: detail will be given in chapter 5.2.

- **Calorimeters**: an energy subtraction algorithm is employed to remove the average BIB contributions to the measured energy. First, through simulations, the average energy deposit of the BIB E_{BIB} and its fluctuations σ_{BIB} in regions of longitudinal coordinate in the calorimeter and θ is found. Then this value is subtracted from the actual energy measured by the calorimeter with signal+BIB, yielding an estimation of the energy deposited by only the signal $E_{SIG} = E_{TOT} - E_{BIB}$. Also for an hit to be selected, an energy $E_{TOT} > E_{BIB} + 2\sigma_{BIB}$ is required. Time cuts can eliminate out of time BIB particles. For both ECAL and HCAL and



(a) (b)

FIG. 36. (a) Schematic working of the DL filter. ΔU cuts on the transverse plane, while $\Delta\theta$ does so in the longitudinal-radial one. (b) Number of hits in the barrel and vertex detectors with only time cuts and with time cuts + DLfilter. The latter reduces the number of hits by an order of magnitude.

all of their regions the time window is [-0.25,0.25] ns. Lastly, the shower profile is different for signal and BIB, as showed in 37, and this will be exploited in the future by using ML techniques on the shower shape to further reject the background.

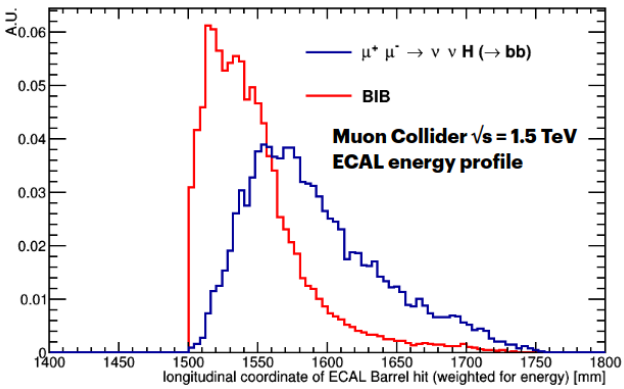


FIG. 37. Signal and BIB shower depth (weighted for energy loss) in the Barrel of ECAL. Signal particles are more energetic and lose more energy deeper in ECAL: this can be used to separate them from BIB-originated showers. Image taken from [34].

- **Muon detector**: this detector has not a huge occupancy and so less mitigation strategies are needed. Once again a time window acceptance of [-0.25,0.25] ns is used. Further requirements in the endcap can be imposed on the angles by requiring all hist to have $\theta > 10°$ and on the pT of the particle, such as $p_T > 5$ GeV.

5. Physics objects reconstruction

In this chapter, the structure of the simulation used in this thesis will be described. First, its workflow and the software used in it will be presented. Then we are going to proceed with the event reconstruction, following with a comprehensive description of the physical objects reconstruction. In particular, the focus is going to be put in the choice of parameters of the objects most crucial in our H$\rightarrow b\bar{b}$ and H$\rightarrow c\bar{c}$ study: jets and secondary vertices reconstruction. Finally the performance of tagging a jet from an heavy flavour quark with a secondary vertex will be evaluated.

5.1 Software and simulation

The following studies have been performed with a detailed simulation of the signal and BIB interactions with the complete detector. This simulation has been performed with the ILCSoftware [35] developed for ILC. It uses the following packages: DD4hep [36], which provide the tools to describe the geometry of the detector and the interface to the package handling the particles interactions; GEANT4 [37], simulating the interactions of the particles passing through the detector materials; the MARLIN [38] software environment The workflow, showed in figure **38**, is similar to what CLIC used, with a key exception given by the overlay of the BIB particles.
So, first and foremost, the signal is generated with an external program such as WHIZARD or PYTHIA8. On the other hand, the particles of the BIB are provided by the combination of FLUKA and FlukaLineBuilder, which computes all products emerging from the beam interactions within 25 meters from the IP, except for the Bethe-Heithe muons which are computed from 100 meters from the IP. Since the BIB multiplicity is huge, some strategies are employed to save RAM and CPU time: all out-of-time particles (time after bunch crossing of arrival at MDI $>$ 25ns) are discarded; for very soft particles detector response is not simulated; for relatively soft neutrons a simplified and less CPU consuming model of interaction with matter is used.
Then the detector response is simulated with GEANT4 separately for the signal and the BIB and superimposed in the overlay phase, producing the simulated hits SimHits.
After this step the hits are converted in experimental signals in the digitization steps. There two classes of hits, each with its own digitization: CalorimeterHit (signals from ECAL, HCAL but also from the muon detector) in which all E_{loss} of the various hits in a cell is integrated in the readout time window, saving the ID of the cell, the total energy collected and the signal arrival time; TrackerHit (from the tracking detectors), in which one saves the ID of the pixel or microstrip where the hit has been recorded, a more precise timing information as well as the energy collected. For the former the number of cells is limited and so the hits are within the time window are merged and their signal fully simulated while for the latter, which have a huge number of channels, two types of digitization are possible: a quicker but less precise one, where a 4-D spread in time e position is applied on the hits and a slower but much more accurate, which fully simulate the sensor and the read-out response. At the moment, the first kind of digitization is used, since it is simpler and much faster in CPU. However this method does not account for pile-up effects or electronic noise and so the full digitization is currently under development, which will also allow to use clustering to reject BIB tracks, greatly improving the performance in the tracking detector.
Now with the reconstructed hits, physics objects such as jets and tracks can be built. Finally the information of the subdetectors are combined by the particle flow algorithms to form the reconstructed

particles. From them, higher level analysis is performed with external tools.

Despite the RAM and CPU saving strategies, simulating the BIB and its interactions is still an heavy task, which requires a lot of time and resources. For this reason, one generates a pool of BIB only events, and then, when one wants to perform a detailed simulation, he or she superimposes the SimHits of the signal to the SimHits of the BIB. The pool of BIB events used in this study is composed by about 30 events. Still, the reconstruction of an event with the BIB turned on is a slow process, and this computational issue is going to limit the size of our samples and force us to use a limited statistic, as we are going to see in the last chapter.

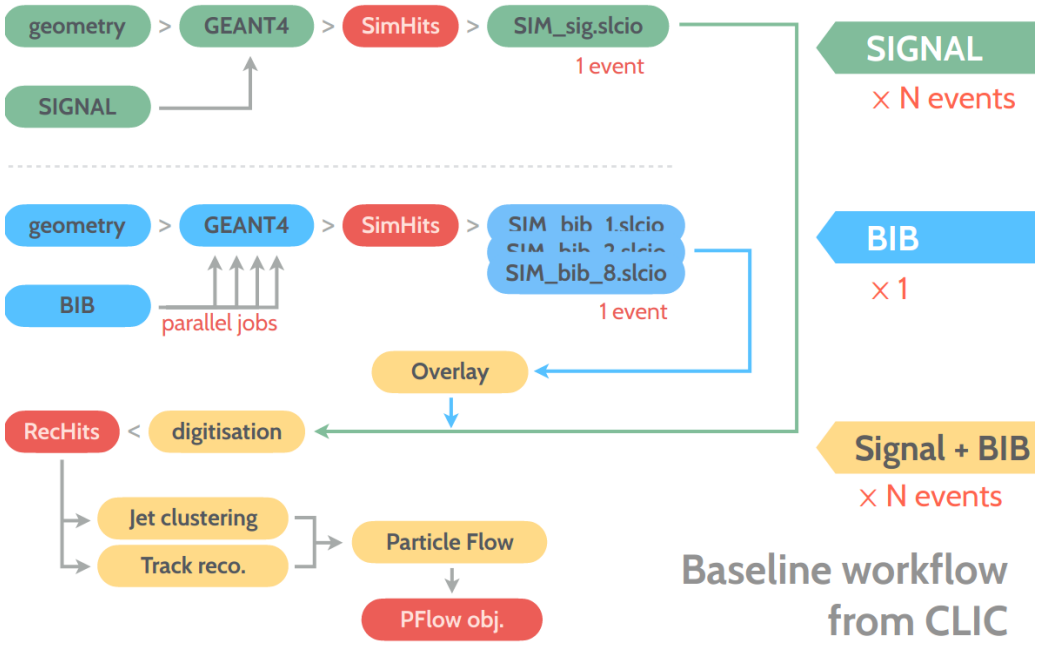


FIG. 38. Simulation workflow. The various steps are explained in the text. Taken from [39].

Jets and vertices, required for higher level analysis (such as looking for b-jets), exploit independent algorithms which will be described in detail in the following sections. This allows to optimize the cuts on the particle flow objects to best suit each reconstruction.
For example, the vertices reconstruction depends mostly from the vertex detector of the tracking system, where the BIB occupancy is the highest. So, this algorithm requires looser cuts, in order to not lose tracks and consequently, secondary vertices. This may improve when the digitization which fully simulates the detector and read-out response will be implemented, since this will allow the clustering, a key item to suppress the BIB in the tracking system,

In this work the beamspot smearing has always been kept turned off. Its effects are going to be studied when the full digitization will be operative.

5.2 Algorithms for physic objects reconstruction

The **event reconstruction** must account for the BIB hits multiplicity, which is higher closer to the beam pipe. For this reason, the event reconstruction starts by looking for energy clusters deposited in the calorimeters. While this is signature of a jet, also a lepton, for example a μ emerging from a semileptonic decay of a b-hadron, will leave a signal in the calorimeters. In this phase, any cluster of energy remaining after the BIB energy subtraction in the calorimeters will be considered as a candidate jet.
In the tracking detector, only hits deposited within $R = \sqrt{\eta^2 + \phi^2} < 0.7$ of the candidate jet axis are used to build tracks: this strategy, motivated by physics, greatly reduces the multiplicity of hits due to the BIB and saves a lot of CPU time. The tracks found in this way are combined with the calorimeter clusters to form reconstructed particles, which are used to build higher level objects such

as jets and vertices. In the following sections we are going to describe each of these objects.

5.2.1 Track reconstruction

The track reconstruction is performed in 2 steps:

1. the use of pattern recognition to identify hits belonging to a tracks, performed with the Conformal Tracking;
2. the fit of the chosen tracks with the Kalman Filter.

Conformal Tracking:
The Conformal Tracking [40] combines the the conformal mapping and the cellular automaton track finding method. The conformal mapping transforms the coordinates in the transverse plane (x, y) in a new set of coordinates (u, v). A charged particle moves in the detector with an helix like trajectory whose projection on the transverse plane is a circle of equation:

$$(x-a)^2 + (y-b)^2 = r^2$$

In the conformal coordinate system, if the circle is passing through the origin such that $r^2 = a^2 + b^2$, the following change of coordinates can be applied

$$u = \frac{x}{x^2+y^2} \qquad v = \frac{y}{x^2+y^2}$$

in which space the equation of motion now becomes that of a straight line

$$v = -\frac{a}{b}u + \frac{1}{2b}$$

which is much easier to find than a circle. Another characteristic of the trajectory in the conformal space is that the hits in the innermost part of the detector are mapped to the outermost part of the conformal space and viceversa: the radial ordering is inverted. If in (u, v) trajectories are straight lines, one can just look for tracks by selecting aligned hits looking at the angular hot distribution in the conformal space. However this method becomes less effective when particles are not following circular trajectories in the (x, y) space, for example due to multiple scattering interactions, or when the circular trajectory is not passing through the origin, as it is the case for displaced tracks coming from secondary vertices. To take this effects into account, the pattern recognition in the (u, v) coordinates is performed with the cellular automaton technique.
The cellular automaton technique define cells as the aims to build cellular tracks by linking the various cells. It works with 2 algorithms which are called recursively in specific regions: the **building** and the **extension** of cellular tracks.

The building algorithm aims to construct the cellular tracks.

1. It sorts all hits in decreasing conformal radius, and, starting to the firsts in the lists (the one closer to the IP in the (x, y) space), it looks for the closest hits in the conformal polar angle and builds seed cells (figure **39**) by linking the first hit with each further hit in a window of $\Delta\theta_{neighbours}$. These seed cells are kept if such hits are not already part of a track (built with a previous iteration of this algorithm), the second hit has a smaller conformal radius, the length of the cell is below l_{max} and the hits have a slope on the longitudinal plane $\Delta z < z_{max}$.
2. The remaining seed cells are extrapolated in the conformal R direction by a parameter $\Delta R_{neighbours}$, adding the closest hits in the polar angle window by building another cell. This new cells are kept if they form an angle $\alpha < \alpha_{max}$ with the previous cell. If new hits are successfully added, the extrapolation is repeated. Candidate cellular tracks are formed (figure **39**).

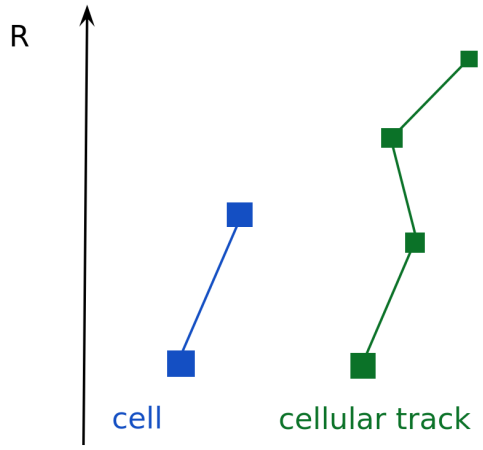


FIG. 39. Pictorial representation of a cell (blue) and a cellular track (green). They are built starting from the hit with the lowest R in the (x, y) space.

3. Each cell stores its starting hit, its last hit and a weight proportional to the number of seeds. Starting from the highest weight track, a fit with a linear regression is performed to find $\chi^2_{u,v}$ and p_T. The information on the z direction is obtained with a fit on the helix in (s, z) where s is the coordinate along the helix. From this, a $\chi^2_{s,z}$ is computed. The candidate cellular track is kept if both $\chi^2_{u,v}$ and $\chi^2_{s,z}$ are below a certain threshold. To avoid using spurious hits, each of them is removed, recomputing the normalized χ^2. Finally, cellular tracks sharing more than hit (clones) are removed choosing the one with more hits, or if candidates have the same number of hits, by comparing the χ^2s. Hits are now marked as used.

The extension algorithm extends the previously created cellular tracks to other regions of the detector. Two different procedures are employed depending on the p_T of the previously found cellular tracks. First will use the first for all tracks satisfying the p_T criteria, and then all the remaining hits are used for the second procedure

- if $p_T > p_{T,cut}$, then the algorithm is similar to the points 2) and 3) of the building one, with the only difference that the extension goes layer by layer and select which hit to add using the χ^2.
- if $p_T < p_{T,cut}$, no nearest hit search is employed. All hits in the same hemisphere w.r.t. the PV are considered and used as seeds hits. The algorithm then works as 2), 3) of building, but a quadratic term is added to the fit of the cellular candidate.

In our simulation, there are 3 calls to this algorithm, shown in table 10.

Algorithm	Region	α [rad]	$\Delta\theta_{neighbours}$ [rad]	Δz	χ^2	Nhits	l_{max} [mm^{-1}]	$p_{T,cut}$ [GeV]
Building	VXD (B)	0.025	0.025	5.0	100	4	0.015	0.5
Building	VXD (B+E)	0.025	0.025	10.0	100	4	0.015	0.5
Extension	(I+O) Tracker (B+E)	0.05	0.05	10.0	2000	4	0.02	0.5

TAB. 10. Steps of the tracking algorithm and respective parameters, described in the text. "I" and "O" describe the Inner and Outer trackers respectively, while "B" and "E" stand for Barrel and Endcap regions of the detector previously mentioned. The $\Delta R_{neighbour}$ parameter is set as 75 % of l_{max}.

This setup can be optimized by using more steps, in particular with a building on the trackers to recover displaced tracks, but due the high occupancy in the detector this is not possible without severely slowing down the simulation. The full digitization of tracker hits could help in this department by removing much of the occupancy and allowing for a more complete tracking.

Track Fit
Neglecting the energy loss and possible multiple scatterings, a charged particle in magnetic field is expected to have an helical trajectory, which can be fitted with the parameters described in table 11.

Parameter	Definition	Description
D_0	$y_0 cos(\phi) - x_0 sin(\phi)$	Signed transverse impact parameter w.r.t. the interaction point
Z_0	z_0	Longitudinal impact parameter w.r.t the interaction point
$tan(\lambda)$	pz/p_T	The "dip" angle: the angle of the helix w.r.t. the $(x.y)$ plane
Ω	$c \cdot B \cdot q/p_T$	Signed curvature of the track with charge q in the B magnetic field
ϕ	$\arctan(p_y/p_x)$	Angle in the azimuthal plane at (x_0, y_0, z_0)

TAB. 11. Formula and description of each parameter used to fit a track. The quantities with the 0 subscript are referred to the point of closest approach of the track trajectory with the interaction point.

However, to account for the presence of deviations from this trajectory, we use the Kalman Filter [41]. Overall, using the previously described parametrization, the fit is performed in 2 steps:

1. The **Pre-fit**: using only the first, the middle and the last hit. This give a starting value of the parameters for the second step.
2. The proper **Kalman Filter**: starting from the previous parameters and the innermost hit in the detector, all other hits are added to the track, updating every time the parameters accounting for multiple scattering and the energy loss.

Tracking Efficiency:
The efficiency of track reconstruction has been evaluated on a sample of 1000 events of $H \rightarrow b\bar{b}$ at $\sqrt{s} = 3$ TeV including the BIB. As described before (and illustrated in figure 43), only particles inside the cones are reconstructed. However in the efficiency computation, it is not possible to use all jets found in an event, as this may lead to look for tracks into fake jets. To remove them, jets are built using Monte Carlo truth information particles from the physic signal as input to the jet clustering (see chapter 5.2.4). So only reconstructed jets with axis with $\Delta R < 0.5$ w.r.t. the axis of a Monte Carlo truth jet are considered in this procedure. The tracking efficiency is defined as

$$Eff_{trk} = \frac{Number\ of\ truth\ particles\ in\ a\ jet\ matched\ with\ a\ track}{Number\ of\ truth\ particles\ in\ a\ jet}$$

and we will look for all stable, charged truth particles with $\Delta R = \sqrt{\phi^2 + \eta^2} < 0.5$ w.r.t. the jet axis. The matching of a Monte Carlo truth particle with a track is performed by minimizing a χ^2, defined as

$$\chi^2 = \frac{(\Omega - \Omega_{MC})^2}{\sigma_\Omega^2} + \frac{(tan(\lambda) - tan(\lambda)_{MC})^2}{\sigma_{tan(\lambda)}^2} + \frac{(\phi - \phi_{MC})^2}{\sigma_\phi^2} \tag{5.6}$$

where $(\Omega, tan(\lambda), \phi)$ are the track parameters with $(\sigma_\Omega, \sigma_{tan(\lambda)}, \sigma_\phi)$ their respective errors, while $(\Omega_{MC}, tan(\lambda)_{MC}, \phi_{MC})$ are the the same quantities reconstructed from the momentum of the Monte Carlo truth particle. It must be noticed that this χ^2 is not normalized to the number of degrees of freedom. We consider a Monte Carlo truth particle as "matched with a track" if the non-normalized χ^2 is < 30. The tracking efficiency as a function of the p_T of the Monte Carlo truth particle is reported in figure 40.

The efficiency is around 0.55 % for Monte Carlo truth particle of $p_T < 40$ GeV and 0.35 % for $p_T > 40$ GeV.

5.2.2 PandoraPFA

In order to achieve highest precision on the invariant masses of bosons decaying in jets and reconstruct the various physics processes, the particle flow algorithm PandoraPFA [42] is used to extract the energy information from the subdetector that is expected to provide it with highest precision. So the energy

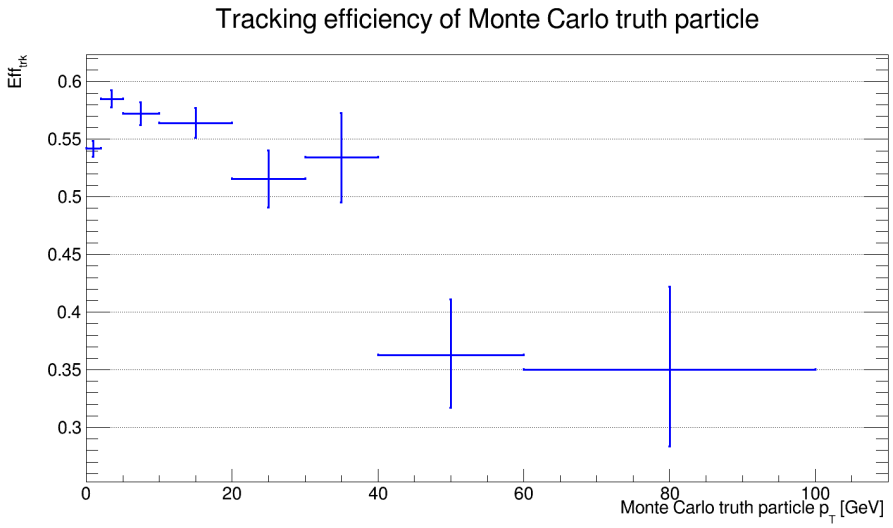


FIG. 40. Tracking efficiency versus transverse momentum of Monte Carlo truth particle within jets.

of the charged tracks is obtained as the module of the momenta of the tracks, while the energy of neutrals is given by the clusters in the calorimeters not associated to charged tracks. This moves the limiting factor to the resolution from the calorimeters performance to the confusion in the association of tracks and clusters, which can be improved with high granularity and the precision of the pattern recognition algorithms in Pandora.

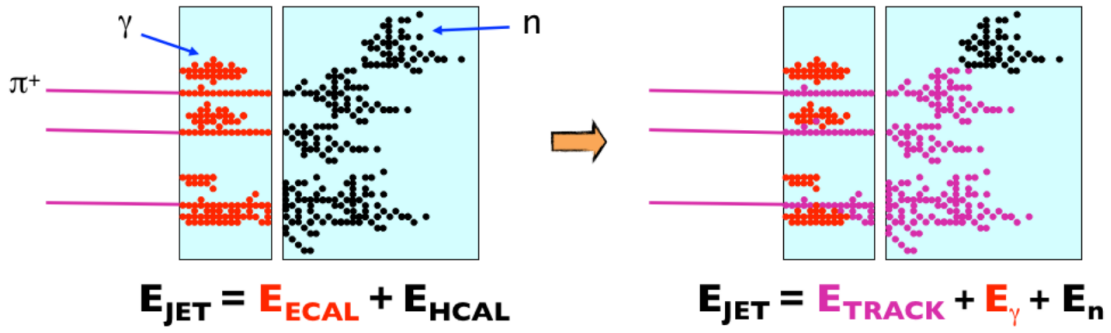


FIG. 41. On the left: energy of a jet computed with standard calorimetry, so as the sum of energy lost in the ECAL (first light blue rectangle) and in the HCAL (second light blue rectangle). On the right, the particle flow approach for the same particles: the energy of the charged tracks is better estimated by the tracking system, while clusters in the ECAL not assigned to tracks form γ candidates and clusters in the HCAL neutrons candidates. Image taken from [42].

Pandora uses several algorithms to perform operations that can be divided in steps. In the step 1 of the jet reconstruction, only calorimeter clusters are given to Pandora and so in that case all steps regarding tracks are skipped.

1. **Track selection and topology:** tracks are projected up to the ECAL. The topology of the track is deduced as kinks and decays from neutral particles are reconstructed. Both track parameters and the topology are given as input to the following steps.
2. **Calorimeter hits selection and ordering:** Position and energy of the digitized hits in the calorimeters are, in order:
   - selected: kept if their energy is bigger than a set fraction (0.5 for the ECal and 0.3 for the HCAL) of E_{loss} by a minimum ionizing particle (MIP);
   - calibrated: the energy of the hits is corrected to account for the invisible energy (of course, the corrections are different for HCAL and ECAL);

- tested with isolation requirements: if the signal was deposited by an high energy particle, other hits must be present nearby the first one. Hits that do not satisfy this criteria are defined as isolated hits. They are mostly due to neutrons produced in the hadronic showers and are treated separately in the following steps.
- tested for MIP identification: the hits with energy deposit below a certain multiple of E_{loss} of a MIP are associated to such a particle. To satisfy this requirement, also nearby cells must not have more than one hit above this threshold.
- subjected to hit ordering: hits within the same layers are ordered by energy

3. **Clustering:** starting from the first layer of the ECAL, clusters are formed using hits in a cone with axis defined by the momenta of the incoming track (if no track is compatible with that hits, the momenta direction is assumed to be along the line connecting the IP and this first hits). Hits within this cones are associated to the clusters, unmatched hits are used as seeds for new clusters.

4. **Topological cluster merging:** different clusters of hits in the calorimeter from 3 are merged depending on topological signatures, which are (note, in the following list, tracks are intended as series of consecutive calorimeter hits, not as the objects reconstructed in tracking detectors): i) looping tracks, ii) track segments with gaps, iii) tracks segments pointing to hadronic showers, iv) track-like neutral clusters pointing to an hadronic shower, v) backscattered track-like clusters originated from hadronic showers, vi) neutral clusters lying in the direction of a charged cluster, vii) neutral cluster in proximity to a charged cluster, viii) cone association, ix) recovery of photons overlapping with other clusters. A picture of these different topologies is shown in figure 42.

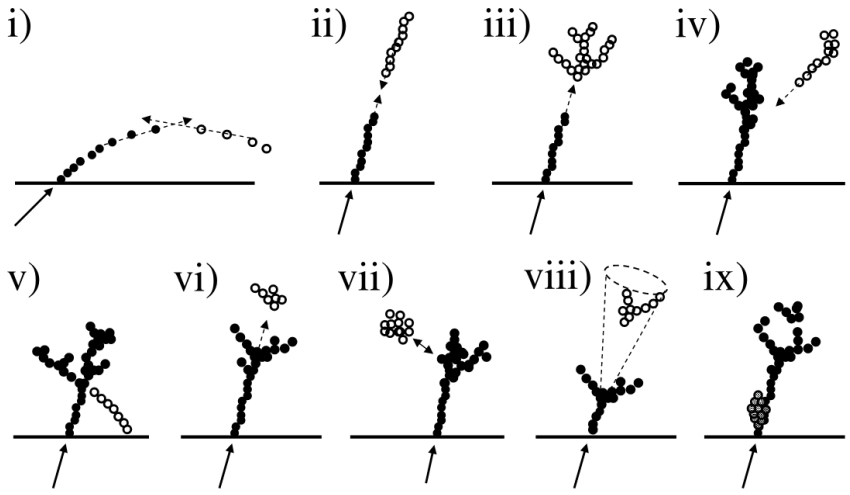


FIG. 42. Picture of the different type of cluster merging performed by the step 4 of PandoraPFA. Each topology is described in the text. Image from [43].

5. **Statistical reclustering:** If a discrepancy is found between the momenta of a track and its associated energy in a cluster, steps 3) and 4) are rerun with different parameters.

6. **Photon recovery and identification:** a scan in the longitudinal and transverse shape of the showers in the ECAL is used to recover clusters compatible with photons.

7. **Fragment removal:** neutral clusters not compatible with γ candidates are selected and, if consistency checks allow it, merged with charged hadronic showers, from which they may have originated.

8. **Formation of particle flow objects:** Reconstructed particles objects are finally built, where their energy is obtained from the $|p|$ of the track if it is a charged particle, while for the neutrals

the energy is obtained from the measurements of the calorimeters. For the latter also scale corrections to the final energy are applied, depending if the reconstructed particle is a photon or a neutral hadron.

Different parameters have been used in the Jet reconstruction algorithm and in the vertex one. In general, the requirements on the latter were much looser to keep a sufficient amount of tracks to build the vertices.

5.2.3 Jet clustering

After the PandoraPFA algorithm produces the reconstructed particles, they are given as input to the k_T algorithm [44] to perform clustering and obtain a jet. Let us call y_i the rapidity of the particle i, p_{Ti} its transverse momenta, ϕ_i its azimuthal angle. The k_T works as follows:

1. A list of all the available particles is created
2. Every possible pair of particles in the list is selected and used to compute the k_T distance:

$$d_{ij} = min(p_{Ti}^2, p_{Tj}^2)\frac{\Delta R_{ij}^2}{R^2}$$

   $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the angular distance of the two particle (i, j) directions while R is defined as $R = \sqrt{y^2 + \phi^2}$ and expresses the aperture of the cone. All results presented in this thesis are obtained with R=0.5.

3. the distance between each particle and the beam is calculated

$$d_{iB} = p_{Ti}^2$$

4. The minimum of all d_{ij}, d_{iB} is found. If the minimum is a d_{ij}, particles i and j are merged into one by summing their 4-momenta. This allows to compute the jet 4-momenta with the E-recombination scheme by summing the components of the 4-momenta of all the particles composing the jet $(E_{jet}, p_{jet}) = (\sum_i E_i, \sum_i p_i)$. If the minimum distance is instead a d_{iB}, particle i is considered the final jet and is removed from the list of all particle.
5. This process is reiterated from step 2 until there are no more particles in the list.

5.2.4 Jet reconstruction

The jet reconstruction is performed in 3 steps, illustrated in figure **43**:

1. first, one looks for energy deposits in the calorimeters remaining after the BIB energy deposit subtraction. They are given to the PandoraPFA algorithm (see 5.2.2), which is used to find hit clusters and transform them in reconstructed particles built from only calorimetry information. In turn, these particles are fed to the k_T algorithm (see 5.2.4) with a radius $R = \sqrt{\eta^2 + \phi^2} = 0.5$, where η is the pseudorapidity and ϕ the angle in radians in the transverse plane of the beam. After the k_T algorithm we have jet candidates built only with the calorimeters information.
2. all hits in the tracking detectors within R=0.7 of the axis of the candidate jet from the previous step are collected and used to form tracks.
3. the clusters in the calorimeters and the tracks found in the precious step are combined by the PandoraPFA algorithm to reconstruct particles and these information are used to recompute the jet with the k_T algorithm with R=0.5.

Jet four-momentum correction due to the BIB

When the BIB is overlaid to the signal, reconstructing the jets becomes more challenging. The BIB energy subtraction removes energy of the jets, and so a calibration is needed. By comparing the true energy of jets generated with PYTHIA8 with the reconstructed energy, a calibration function on the

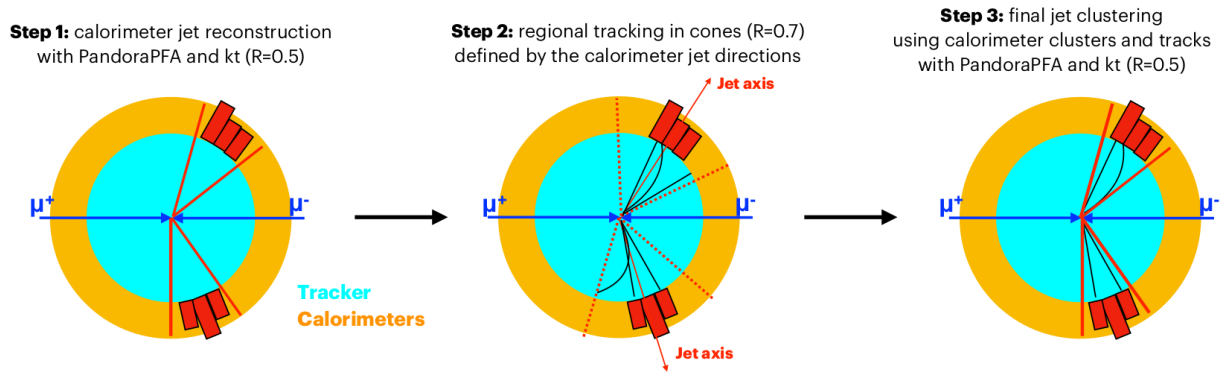


FIG. 43. The 3 steps of the jet reconstruction, previously explained in the text.

four-momenta of the jets is computed. It is parametrized as a polynomial up to the third power in the p_T of the jet before the correction, whose coefficient are different in 5 regions of pseudorapidity [0-0.5, 0.5-1, 1-1.5, 1.5-2, 2-2.5]. The function used in the pseudorapidity range [0.5,1] is plotted in 44.

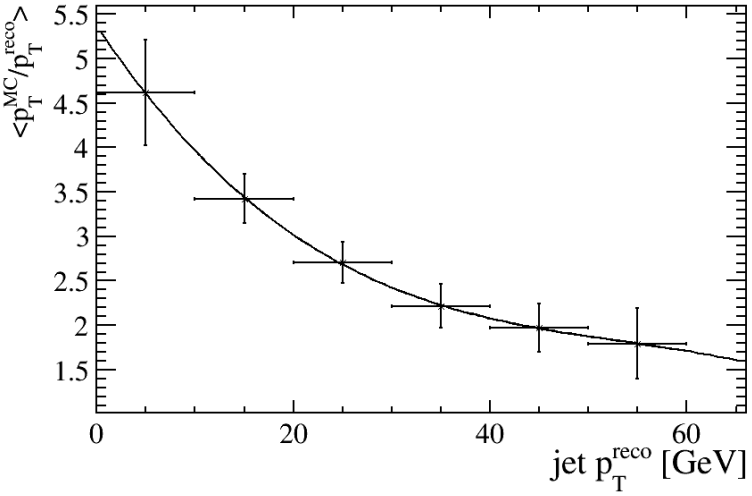


FIG. 44. Four-momentum calibration curve of jets n the pseudorapidity range [0.5,1]. The p_T of the reconstructed jet is placed on the x-axis, while the y-axis lable is the correction factor, computed as the mean of the ratio of the p_T of the Monte Carlo jets and the p_T of the corresponding reconstructed jet. Monte Carlo jets and the matching procedure are described in chapter 5.3.

The performance of the b-jets reconstruction obtained from an inclusive bb sample generated with PYTHIA8 [45] + BIB at 1.5 TeV is showed in figure 45. b-jets are identified with the matching procedure described in detail in chapter 5.3. Despite the use of non-optimized algorithms, the results are promising.

5.2.5 Vertexing algorithm

The vertexing algorithm is optimized depending on the type of vertex that has to be reconstructed. The procedure is inspired by [47], but is then modified to adapt it to and our environment.

- **Primary vertex finding:** First a selection of the tracks to be used is performed. All tracks with $D_0 < 0.1$ mm and $Z_0 < 0.1$ mm w.r.t. the center of the constrain region (0,0,0) are kept and used to fit a first PV candidate. Also, to remove the number of tracks coming from the BIB and displaced vertices, each track is required to have at least 4 hits associated in the vertex detector (Barrel+Endcap).

  Then the so called tear-down algorithm begins. A parameter analyzing how well a track fits into a vertex $\chi^2_{threshold} = 10$ is chosen: this states how close a track must pass near a Vertex

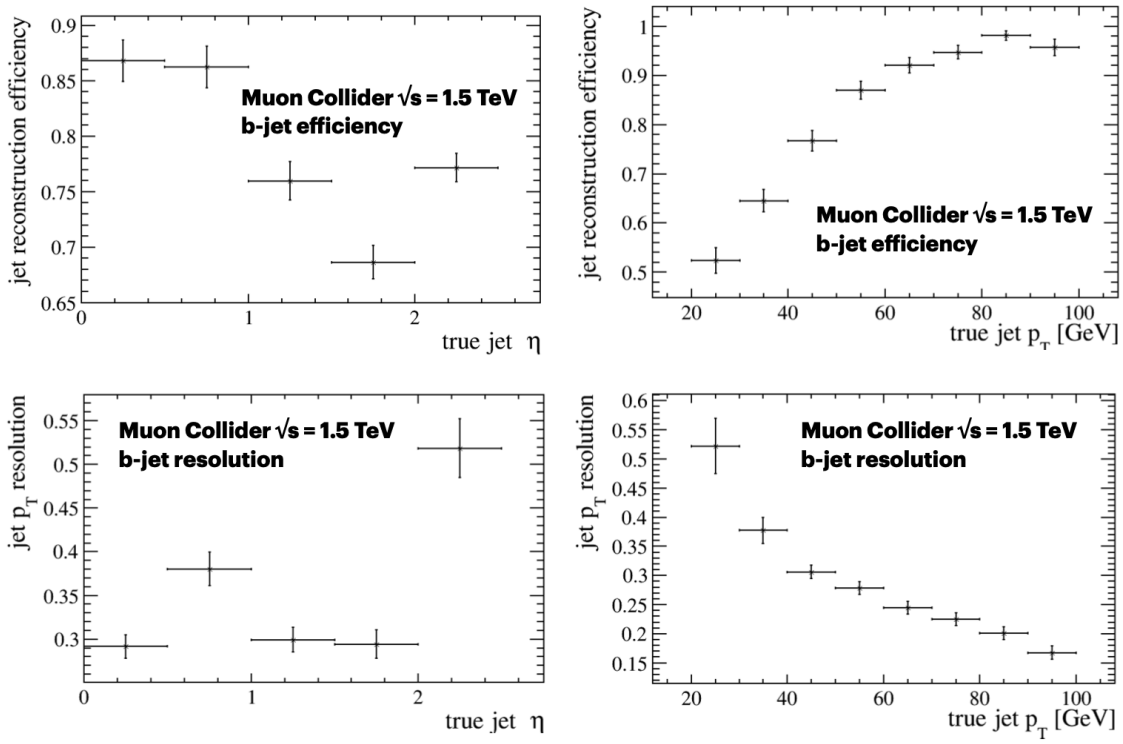


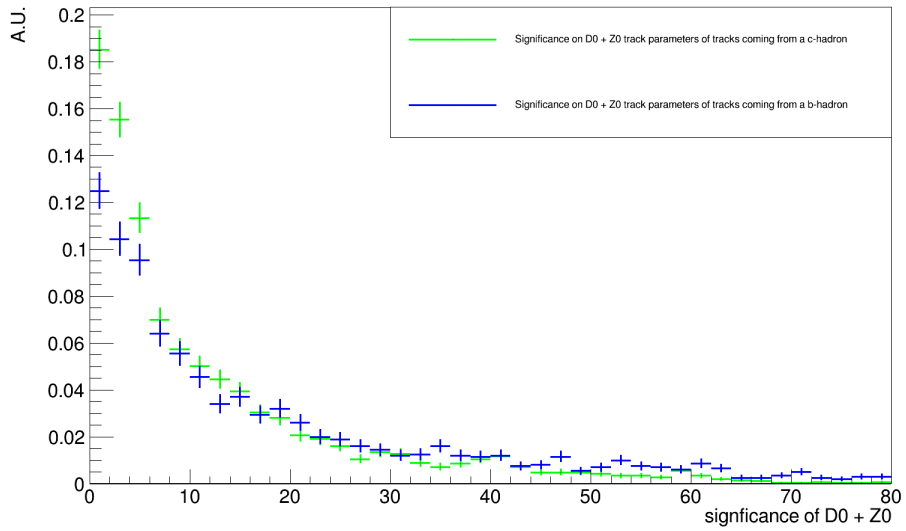
FIG. 45. Upper row: b-jet reconstruction efficiency as a function of the true jet pseudorapidity η and p_T. Lower row: b-jet reconstruction efficiency as a function of the true jet pseudorapidity η and p_T. The results are better at high p_T, due to a better separation from the BIB (which leaves soft energy deposits in the calorimeter), and low η, since the nozzles can obstruct jets with higher pseudorapidities. Taken from [46].

to be considered coming from it. Then the track with the highest χ^2 (if $\chi^2 > \chi^2_{threshold}$) w.r.t. the vertex is removed from the set of tracks and a new candidate PV is obtained fitting the remaining tracks. Then this procedure is repeated with the new highest χ^2 track, and so on until no track has a χ^2 bigger than the $\chi^2_{threshold}$. The fit of the remaining tracks determines the Primary Vertex position.

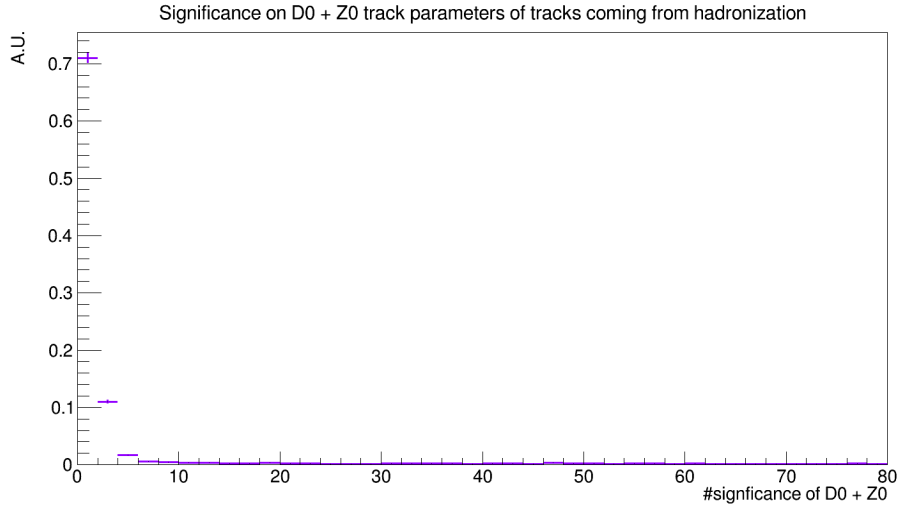
- **Secondary vertices finding:** Tracks not used to build the PV are selected. The requirements on the parameters listed in the following are used with the aim to pick tracks displaced enough to be coming from a secondary vertex and reject the ones coming from the BIB. The choice of the parameters of this algorithm has been performed by looking at the characteristics of tracks matched with Monte Carlo truth particles coming from a b-hadron. The matching has been performed using the handmade χ^2 defined in equation 5.6, in the tracking efficiency section.
  To further remove tracks possibly coming from a PV (erroneously removed by the teardown algorithm of the PV), a cut on the significance of Z_0 and D_0, $\sqrt{(\frac{\sigma_{Z0}}{Z_0})^2 + (\frac{\sigma_{D0}}{D_0})^2}$, is set. By requiring that each track is at least 2 sigma away from the PV position, we remove 70 % of the hadronization tracks, as can be seen looking at the distributions in figure **46**.

  However setting a maximum displacement is required to remove displaced tracks from the BIB. The secondary vertices of b- and c-hadrons travel only a small distance before decaying, so almost no signal track is expected to have Z_0 and D_0 above 5 mm in each case. As can be seen for Z_0 by the plots in figure **47**, most of the signal particles have Z_0 below 5 mm; on the other hand, the BIB tracks longitudinal impact parameter distribution is much wider, having the tail of Z_0 up to 300 mm: only a small fraction of the tracks fits in the canvas. This is to be expected, since BIB particles may originate from other points than the IP of the beamline.

  To further remove the BIB, each track is required to have at least 4 valid (i.e. surviving the Double Layer Filter) hits in the vertex detector (Barrel+Endcap). The selection is performed using the hits in the VXD detector only due to the fact that tracks are built from hits in the



(a)



(b)

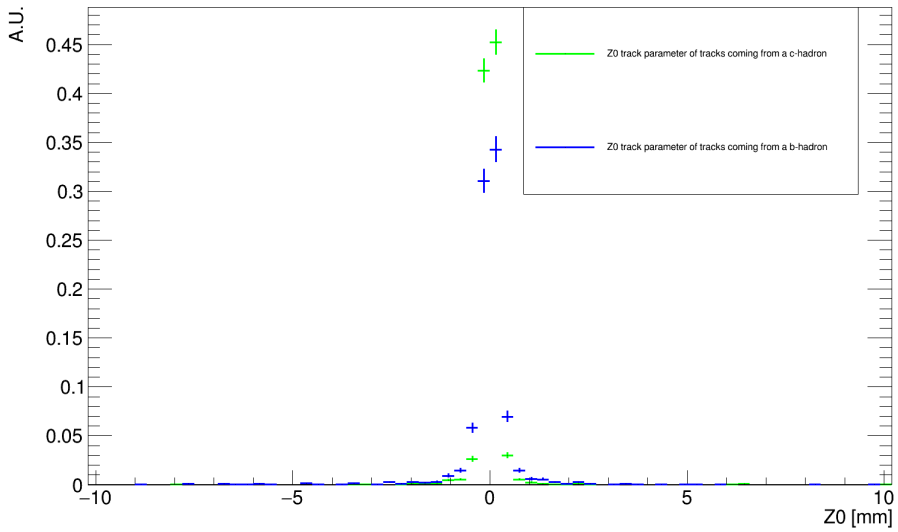
FIG. 46. Distributions normalized to 1 of the combined significance of Z_0 and D_0 $\sqrt{(\frac{\sigma_{Z0}}{Z_0})^2 + (\frac{\sigma_{D0}}{D_0})^2}$ w.r.t. the PV of a track matched with a particle daughter of a b- or c-hadron (blue and green respectively) (a) or of a track matched with a particle produced during the hadronization (violet) (b).

VXD detector and then extended to the trackers, as explained in the tracking section (5.2.1). This requirement removes 50 % of the BIB particles while keeping the majority of the signal tracks, as can be seen in figure **48**. Also, we can exploit the fact that the BIB has a soft momenta spectra (figure **26**), and by requiring $p_T > 0.8$ GeV we can discard $\sim 80\%$ of the BIB tracks and lose a small amount of signal particles, as shown in **49**.

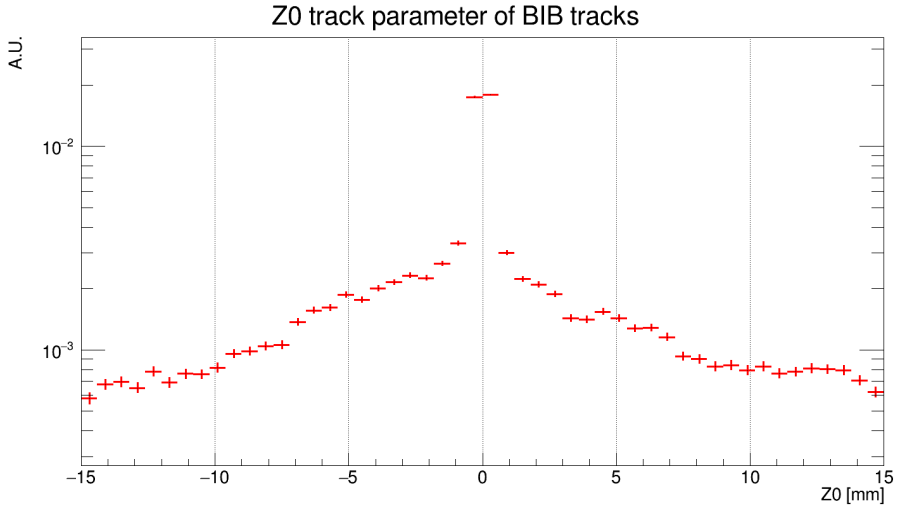
Finally a maximum for the errors of Z_0 and D_0 is fixed to 1 mm, to reject decays from long lived particles and BIB. The distributions presented in figure **50** show that almost no signal track is cut by these requirements, while tracks coming from the BIB have a distribution that extends over 2 mm for both parameters.

The tracks surviving all these requests are put in a list and used in the following steps. From every pair of seed tracks in such list, starting vertex candidate are built from the intersection of circular projection of helices of the tracks. The Z position of this candidate is determined scanning along the helices. From this starting point a χ^2 fit with Minuit is performed, and the vertex is kept if some criteria are met:

– the invariant mass of the SV must be below a threshold, $m_{SV} < 10$ GeV;



(a)

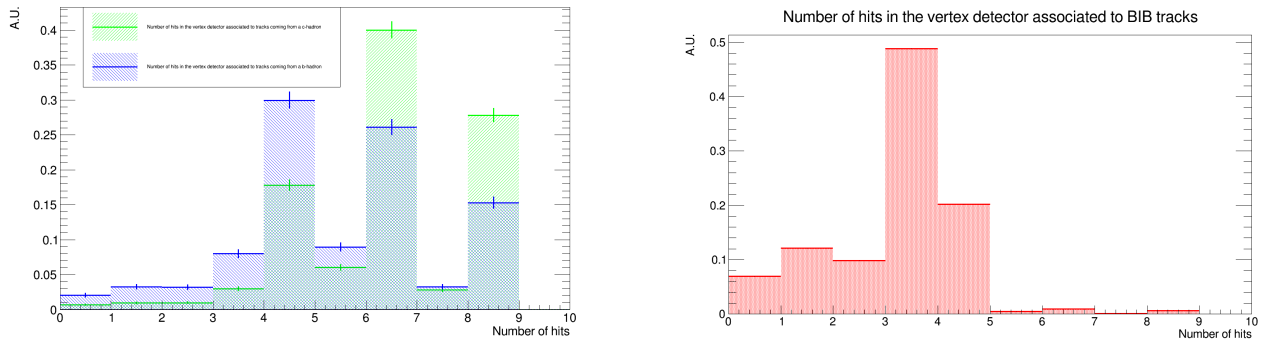


(b)

FIG. 47. Distributions normalized to 1 of Z_0 and D_0 w.r.t. the PV of tracks matched with particles daughters of a b- or c-hadron (blue and green respectively) (a) and from the BIB (red) (b).

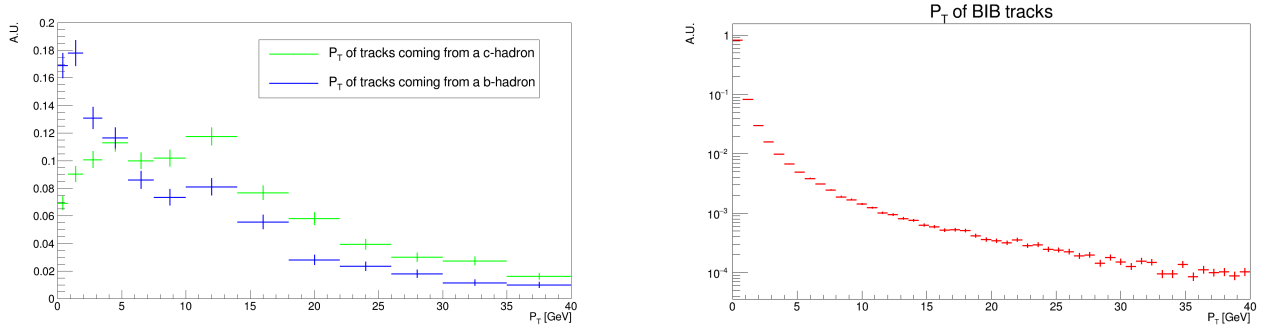
- the invariant mass of the track pair must be smaller than the energy of each track
- the track pair must not be compatible with coming from the decay of a neutral long lived particles V^0, such as a K^0
- the vertex must lie in the same side of the sum of the track momentum w.r.t. the PV
- the χ^2 of the track w.r.t. the vertex position must be below a certain threshold $\chi^2_{thr} = 5$

Additional tracks are added to a 2-tracks vertex if the track's χ^2 w.r.t. the vertex is below $\chi^2_{thr} = 5$, after all the others criteria listed above (except the V^0 check) are satisfied. Finally one eliminates overlapping tracks, i.e. tracks compatible with more than one vertex. Starting from the vertices with the highest number of tracks N (which are the ones less likely to appear from the combinatorial background), all tracks are assigned to the vertex with the lowest $\chi^2 = \sum \chi^2_{trk}$ and removed from all the other vertices. Such vertex is saved and considered final. This process is repeated until no vertex with N tracks remains, and then is used for vertices with N-1 tracks. The algorithm ends when no more vertices are available. Finally a check whether the SV is inside the nozzle is performed, discarding the vertex if this is true.



(a) (b)

FIG. 48. Distributions normalized to 1 of number of hits in the Vertex detector of tracks matched with Monte Carlo truth particles coming from a b- or c-hadron (blue and green respectively) (a) and from the BIB (red) (b). The peaks at 4,6 and 8 in picture (a) are due to the fact that the hits of Double Layer are counted separately. Odd number of hits are present if two hits are assigned to a track in the same layer



(a) (b)

FIG. 49. Distributions normalized to 1 of p_T of tracks matched with Monte Carlo truth particles coming from a b- or c-hadron (blue and green respectively) (a) and from the BIB (red) (b). The cut in p_T rejects the first bin of each histogram: so it removes only a small fraction of signal particles, while discarding $\sim 80\%$ of the BIB tracks.

5.3 SV-tagging performance on b- and c-jets

The jets coming from heavy flavour quarks can be identified by looking inside them for Secondary Vertices (SV), created from the products of the decays of mesons originating from the initial quark. On the other hand, jets produced by lighter quarks or gluons are much less likely to have SVs, and so the presence or absence of it provides a good discriminant to separate jets from light flavour quarks and from b and c quarks. In this section we are going to evaluate the SV-tagging performance on inclusive bb, cc and qq samples.
Before diving in the performance tests, it is necessary to define some useful quantities, as well as describing some correction factors used in the analysis.
First, to see how well we are recognizing b- and c-jets, we need to know which jets of the event are coming from them and which are of other types, such as jets coming from light (u,d,s) quarks. This is done using Monte Carlo information with the procedure called **matching**.

1. One looks at all the Monte Carlo truth quarks produced by the physics process before the hadronization begins, and for each of them the PDG identity code and momenta direction is saved.

2. All the Monte Carlo truth particles are given as input to the jet clustering processor, which builds PandoraPFA (see chapter 5.2.2) from them and using the kT algorithm (chapter 5.2.4) with identical parameters used for the reconstructed particles, builds the so-called Monte Carlo truth jets (MCjets).

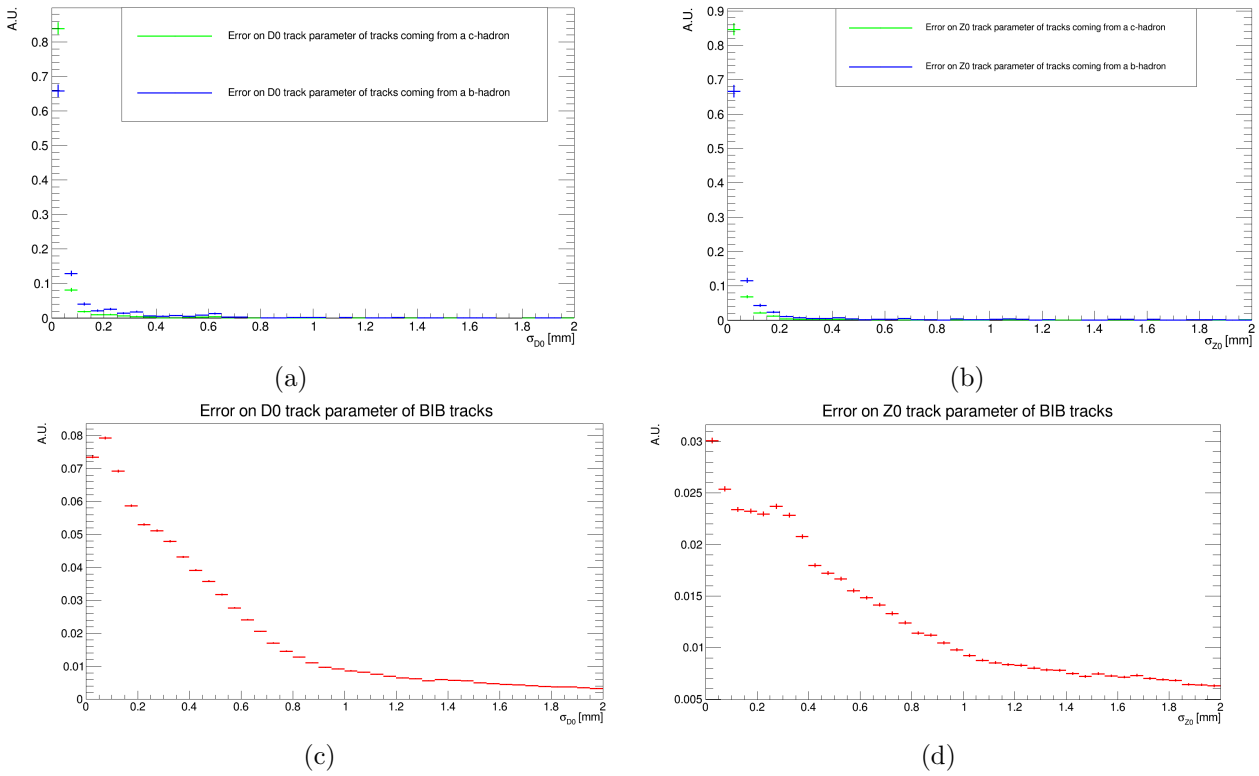


FIG. 50. Distributions normalized to 1 of p_T of σ_{Z0} and σ_{D0} tracks matched with Monte Carlo truth particles coming from a b- or c-hadron (blue and green respectively) ((a),(b)) and from the BIB (red) ((c),(d)).

3. One can classify the MCjets by looking at the flavour (from the PDG identity code) of the quarks found in step 1 inside the MCjets: since the jets are built with $R = \sqrt{\phi^2 + \eta^2}$, this means that the axis of the MCjets must have $\Delta R < 0.5$ w.r.t. the quark momenta direction. If more MCjets contain the same quark, only the MCjet with axis closer to the quark is considered matched. If a MCjet contains more than a single quark, this MCjet is considered a fat jet.

4. Jets from reconstructed particles are built as defined in 5.2.4. If one of them contains ($\Delta R < 0.5$) the axis of a matched MCjet, then it is considered a matched jet and it is assigned teh same flavour of the MCjet. If more jets match the same MCjet, only the one with smallest ΔR is considered matched.

A pictorial representation of this process is shown in figure 51. At the end of this procedure we have classified jets depending on their expected flavour quark origin. However two considerations are needed. First, in this chapter fat jets will be considered coming from the heaviest flavour quark they are matched to, however in the future, if a Deep Neural Network (DNN) is going to be used to identify the jet flavour, this types of jets will be discarded as their substructure is vastly different from that of a normal jet and this may reduce the network performance. Secondly, not all quarks have an MCjet within $\Delta R < 0.5$; however being quark they must hadronize: so it is likely their MCjet is present but has $\Delta R > 0.5$ w.r.t. the quark direction before the hadronization. This last fact will lead to an overrate of the mistag, as these jets will not be considered jets from heavy quarks even if they in fact are and as such will contain SVs.

We can now define other two crucial quantities in our reconstruction: the **SV-tagging efficiency** of a jet

$$Eff_{SV-tag} = \frac{Number\ of\ jets\ tagged\ AND\ matched\ to\ b,c\ truth\ quarks}{Number\ of\ jets\ matched\ to\ b,c\ truth\ quarks} \tag{5.7}$$

and the **mistag**

$$Mistag_{SV-tag} = \frac{Number\ of\ jets\ tagged\ AND\ NOTmatched\ to\ b,c\ truth\ quarks}{Number\ of\ jets\ NOTmatched\ to\ b,c\ truth\ quarks} \tag{5.8}$$

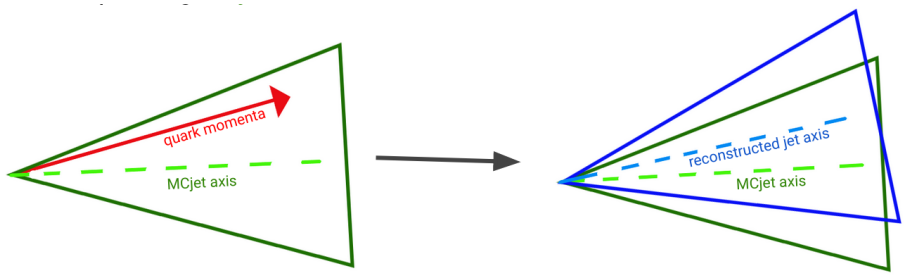


FIG. 51. Pictorial representation of the matching procedure explained in the text: first the quark momenta direction is used to give a flavour to the Monte Carlo jets, then these Monte Carlo jets are used to match the reconstructed jets.

The former estimates how many heavy flavour jets we can recognize, while the latter expresses how many light jets may contain a SV and so be misidentified with an heavy one.
All the previous definitions hold for all samples we are going to analyze.

In this analysis the following samples have been used:

- without BIB: 10000 inclusive $b\bar{b}$, 10000 $c\bar{c}$ and 1500 $q\bar{q}$ events generated in the $\sum |p_T|$ range [120,160] GeV at leading order (L.O.) with PYTHIA8.
- with BIB: 1500 inclusive $b\bar{b}$, 1600 $c\bar{c}$ and 1500 $q\bar{q}$ events generated in the $\sum |p_T|$ range [120,160] GeV at leading order (L.O.) with PYTHIA8.

The samples without BIB have been used to estimate the effect of the Double Layer filter on the displaced tracks coming from a SV We could not use samples with BIB to estimate the magnitude of this effect: it takes too long to run an event with BIB without this filter, and so to estimate how much this affects our results, we have to use these samples, which also allow higher statistic. We reconstruct the samples without BIB two times: one with DL filter turned on and one without it. Apart for the DL filter, in each configuration we will be using the same tracking configuration employed for the BIB, with the only difference being the fact that in the former the tracking is performed in the full detector, instead of the regional strategy employed to reduce the BIB occupancy in the latter category.
In both configurations, the number of tagged jets will be saved in bins of θ and p_T of the jet. Since the events and the analysis are the same, the only differences that appear must be due to the DL. A ratio of the tagged jets found without double layer filter over the number of the tagged jets with such algorithm activated is computed for each bin of the previously mentioned variables, and this provides a correction factor that accounts for the effect of the DL.
This procedure will be done separately for jet tagged and matched with an heavy Monte Carlo quark (which are the numerator of our tagging efficiency) and for jet tagged but not matched (forming the numerator of the mistag). The same procedure will be applied separately for b and c jets.
The correction factors can be seen in the appendix 7.3, together with the invariant mass distribution of the $H \to b\bar{b}$ and the position of the SVs before and after turning on the DL: these last distributions were used as cross-checks. From this point onwards, all efficiencies presented will have already been corrected with this procedure.

The vertex tagging efficiency has been evaluated on the samples with BIB overlaid on.

The SV-tagging efficiency for **b-jets** is presented in figure 52, as a function of the p_T and θ angle of the jet. At low values of the angle the tagging efficiency is lower due to the presence of the nozzle and to the fact that the BIB combinatorial is high in this region, making more difficult to correctly reconstruct track and be able to form SVs. In the central region, so at high values of θ, BIB occupancy diminishes and in fact the tagging efficiency is higher. Also it can be noticed that at high p_T the efficiency increases: the SVs become more boosted and more easily separated from the PV.

The SV-tagging efficiency for **c-jets** is presented in figure 53, as a function of the same variables. As for the b-jets, the tagging efficiency increases in the central region (due to the presence of less BIB)

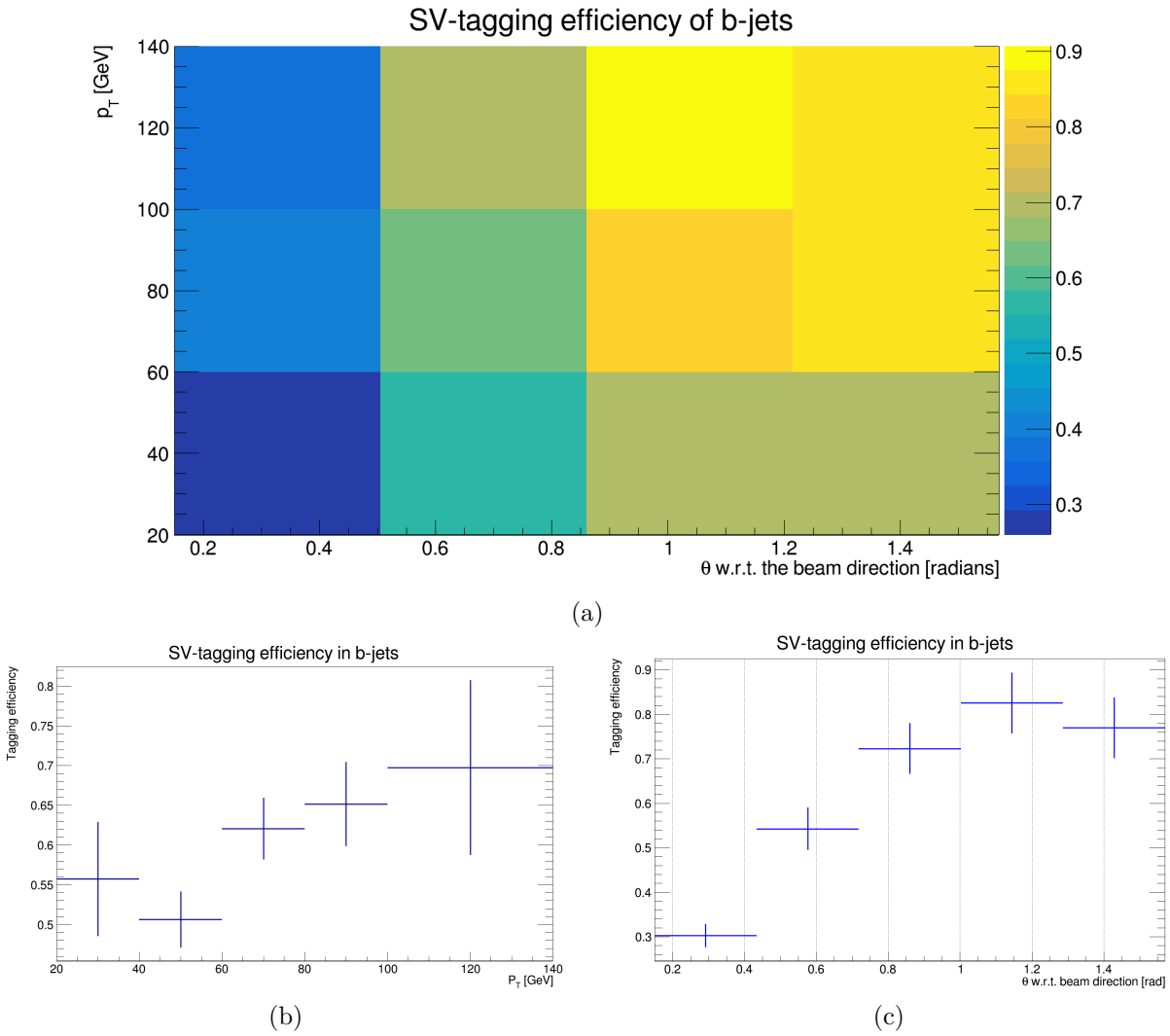


FIG. 52. SV tagging efficiency for b-jets as a function of p_T and θ angle of the jet (a). The efficiency is higher in the central region where the BIB occupancy is smaller, and at high p_T, since the SV become more displaced. Below are reported the SV tagging efficiencies for b-jets as a function of p_T only (b) and θ only (c) with thinner binning.

and at high p_T. However the magnitude of the c-jet tagging efficiency is lower than the b-jet one: the c-hadrons travel a smaller distance before decaying and are more likely to be mistaken with the PV; moreover, on average a b-hadron has more tracks than a c-hadron, and so it is more likely to be reconstructed.

The mistag is shown in figure 54 as a function of p_T and θ angle of the jet. The binning is larger due to the lower statistics, especially at high p_T. The mistag is slightly higher in the forward/backward region, due to the presence of more BIB tracks which may intertwine and form fake SVs. Finally, high p_T jets coming from high p_T gluons are more likely to produce heavy flavour quarks during the hadronization and as such have a SV coming from them.

The Mistag of the light jets sample has been confronted with the one of the $b\bar{b}$ and $c\bar{c}$ sample, which show a similar distribution, although with a different overall magnitude (is bigger by 0.5 % in both $b\bar{b}$ and $c\bar{c}$). Some contributions to this deviation might be:

- light jets from the heavy flavour samples are mostly from almost collinear gluons emitted by b and c quarks during the hadronization and due to their vicinity or partial overlap with heavy flavour jets, they may contain the SV of the heavy quark;
- Sometimes the matching procedure fails (for example when no MCjet is found near the Monte

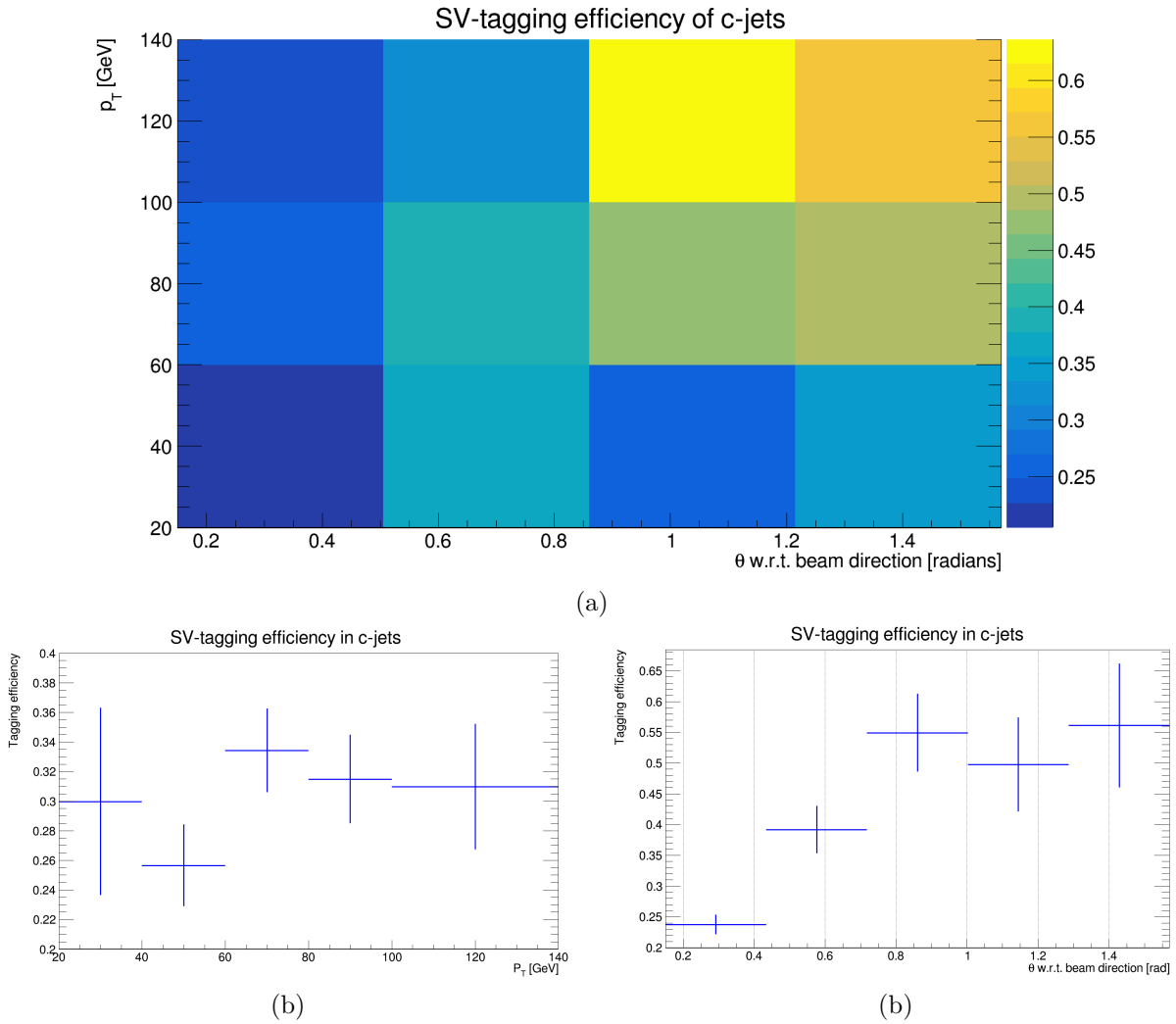


FIG. 53. SV tagging efficiency for c-jets as a function of p_T and θ angle of the jet. The same considerations for the b-jets tagging efficiency [52] hold; however the magnitude of the tagging efficiency for c-jets is lower due to the considerations written in the text. Below are reported the SV tagging efficiencies for c-jets as a function of p_T only (b) and θ only (c) with thinner binning.

Carlo truth quark). However the SV from the b- or c-hadron decay can still be reconstructed and may be found inside a inside a jet erroneously not matched.

The magnitude of each effect is still to be studied.
The values of differential SV-tagging efficiency measured in the inclusive $b\bar{b}$ and $c\bar{c}$ samples is found to be compatible with the same quantities computed, respectively, in the H$\rightarrow b\bar{b}$ and H$\rightarrow c\bar{c}$ samples.

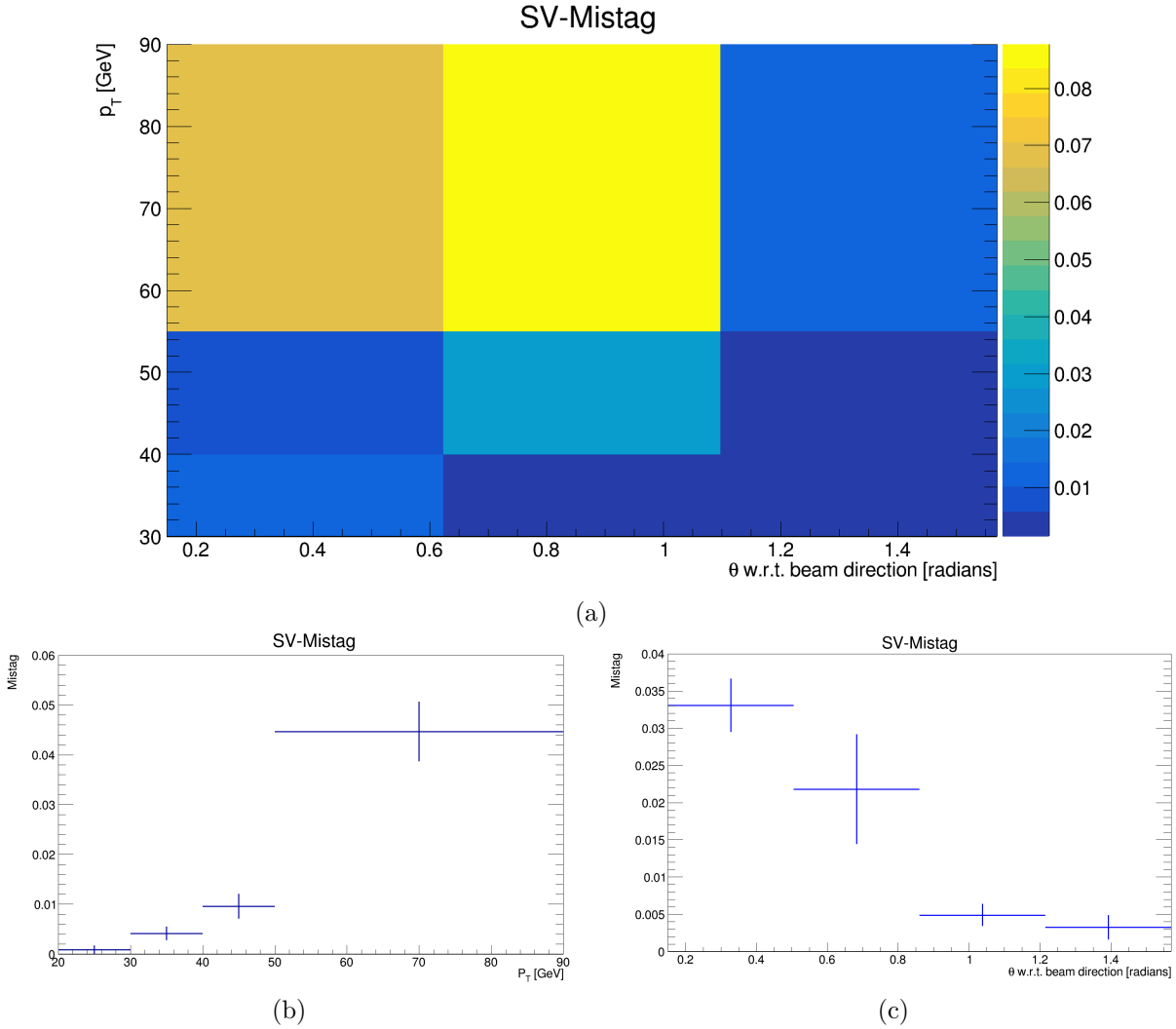


FIG. 54. Mistag as a function of p_T and θ angle the jet, extracted from the $q\bar{q}$ sample. The Mistag is higher in the forward/backward region due to the BIB higher occupancy and at high p_T, since energetic gluons are more likely to produce heavy flavour quark during the hadronization, which will produce SVs. Below are reported the Mistag as a function of p_T only (b) and θ only (c) with thinner binning.

6. H$\rightarrow b\bar{b}$ identification and cross section measurement

Now that the detector and algorithms have been described in detail, it is possible to proceed with the main objectives of this thesis: the estimation of the statistic uncertainty of $\sigma_{H\rightarrow b\bar{b}}$ and the study of jets characteristics for a future $\sigma_{H\rightarrow c\bar{c}}$ measurement. In this chapter we will see to the former.
In the first section, we are going to look at the main characteristics of an H$\rightarrow b\bar{b}$ sample, focusing on the jets. The b-jets coming from the Higgs decay will be identified by looking inside them for secondary vertices. With this type of flavour jet tagging, we are going to reconstruct the peak of the Higgs mass. In the second section, the spectrum of the dijet reconstructed invariant mass from tagged jets will be used to extract the uncertainty on $\sigma_{H\rightarrow b\bar{b}}$, taking into account the effects of the main sources of background.

6.1 H$\rightarrow b\bar{b}$ reconstruction

In this analysis 2000 events of H$\rightarrow b\bar{b}$ were generated with PYTHIA8 at leading order, requiring:

- both quarks to have a pseudorapidity $|\eta| < 3$,
- both quarks to have $p_T > 5$ GeV.

These cuts, motivated by the detector acceptance , will be imposed on every following sample. With these requirements, the cross section of the H$\rightarrow b\bar{b}$ is found to be $\sigma_{Hbb} = 324$ fb. 2000 events provide a limited statistics, however as described in section 5.2, the event reconstruction with the BIB is a slow process and reconstructing big samples requires a lot of time.
On the reconstruction side, to reject background, cuts are imposed on all samples:

- the p_T of a jet is required to be > 20 GeV;
- the module of the pseudorapidity $|\eta|$ of a jet must be below 2.5;
- the invariant mass reconstructed from 2 tagged jets M_{jj} is required to lie in the interval [0, 300] GeV.

These cuts on the jets have been chosen looking at the jet reconstruction efficiency which, as shown in figure **45**, is small at low p_T and at high η due to the presence of the nozzles. If a jet does not satisfy these criteria, it is removed. The boundaries of the mass fiducial region will be discussed when we will introduce the physical background in the next section. H$\rightarrow b\bar{b}$ events are characterized by the emission of jets whose distribution can be seen in figure **55** and **56**. We are going to call b-jets (surviving the cuts) the jets matched with a Monte Carlo quark with the procedure explained in chapter 5.3. The jets, coming from high energy quarks, appear to have a p_T spectra mostly higher than 50 GeV, which is good since the jet reconstruction efficiency, shown in figure **45**, is higher at high p_T. The ϕ and p_T distributions are presented in figure .

The forward/backward region is the region where the BIB is most present, however looking at the χ^2 distribution of the secondary vertices in figure **57**, it appears that our track selection in the SV algorithm is able to avoid a large creation of spurious vertices from BIB tracks. In fact not only the

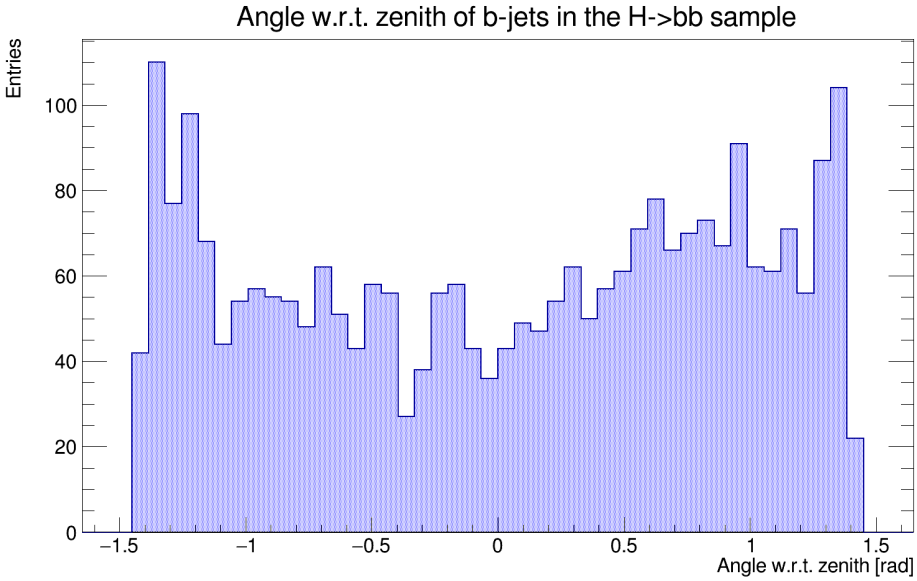
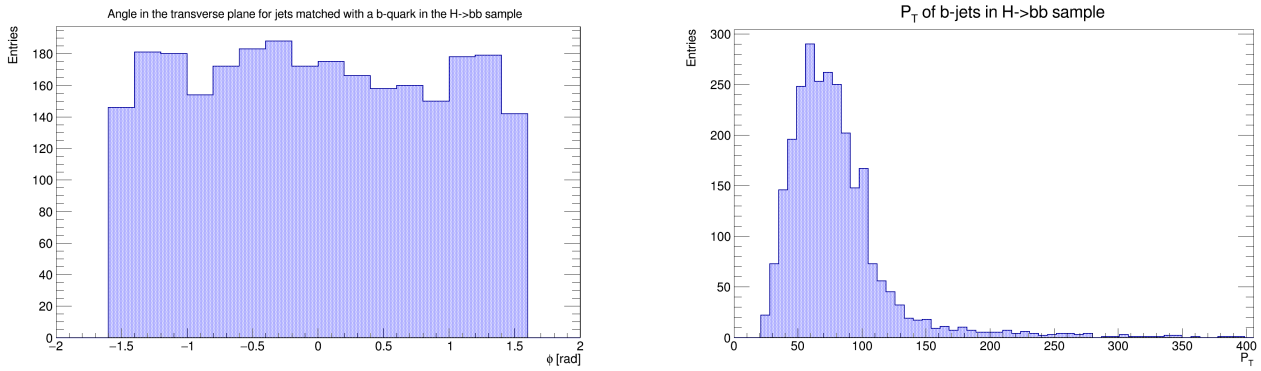


FIG. 55. Angular distribution of jets matched with a b-quark, where the angle has been computed as $\arctan(p_Z/p_T)$.



(a) (b)

FIG. 56. ϕ and p_T distributions for jets matched with a b quark. The ϕ distribution is symmetric as expected. The p_T distribution shows that the requirement for jets to have $p_T > 20$ GeV did not cut many events. The p_T spectra lies in the region where the jet reconstruction efficiency and resolution are better, as shown in the jet clustering performance section (in particular, figure 45).

χ^2 distribution of SVs found inside and outside the cone have the same shape, but also the position distribution of the vertices shows that the vast majority of the SVs outside the cones are found close to the IP as can be checked in figure 58. This can be explained with a geometrical consideration: if due to a small mis-reconstruction of a track the correspective vertex is found shifted w.r.t. to its original position, this is much likely to go outside the cone if said vertex is close close to (0,0,0), as when close to the origin, a small shift in position may imply a big shift in (θ, ϕ). As a cross-check, we also show the p_T distribution of the SVs: once again (figure 58) we found that the distributions of SVs outside the jets have the same distribution of the ones inside, showing that they the vertices outside the cones are likely due to mis-reconstruction of signal tracks.

After having studied separately b-jets and SVs, we can also show how tagged jets are distributed in a same event: the cosine of the angle between the two jets presented in figure 59 shows that tagged jets are mostly close-by. This is confirmed by the fact that the p_T distribution of the Higgs reconstructed from the two tagged jets shows that the Higgs is quite boosted (figure 59). In a limited number of cases, the two b-quark were emitted so collinearly that in the matching assigned to the same Monte Carlo truth jet: in this case only a b-jet is present in the event and with our reconstruction strategy we are not able to reconstruct the Higgs. Although this is an uncommon case ($\sim 2\%$ of the events),

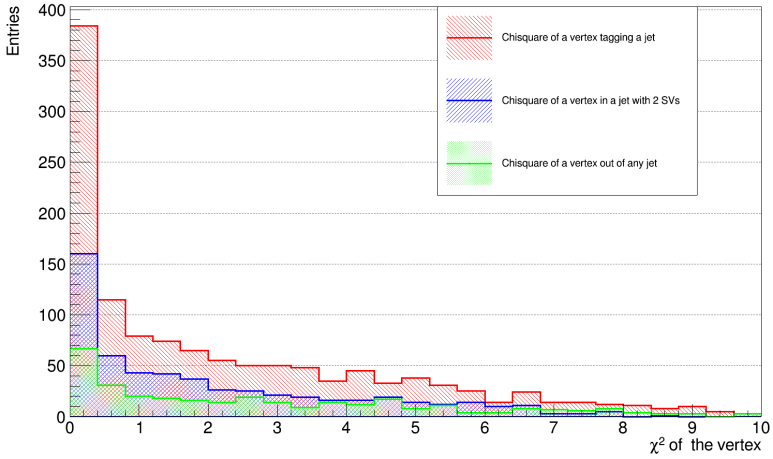
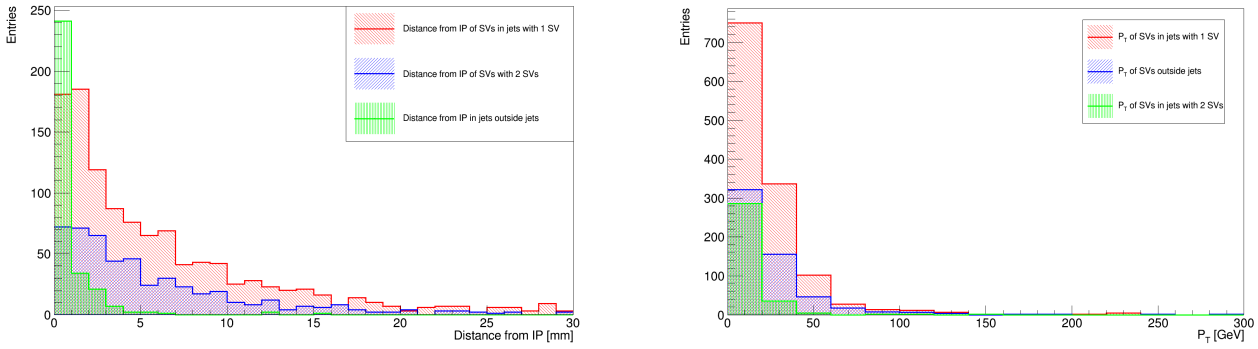


FIG. 57. χ^2 distribution of SVs divided in: SVs tagging a jet (red); SVs in jet tagged by 2 SVs (blue) and SVs outside the jet cones (green). The distributions have the same shape, with a peak at low χ^2 values.



(a) (b)

FIG. 58. Distribution of distance from IP (a) and p_T (b) of SVs divided in: SVs tagging a jet (red); SVs in jet tagged by 2 SVs (blue) and SVs outside the jet cones (green). The three types of SVs have the same p_T distribution while, as explained in the text, the vertices outside of the cones are mostly near the IP.

a different strategy should be thought to be able to reconstruct these type of events.

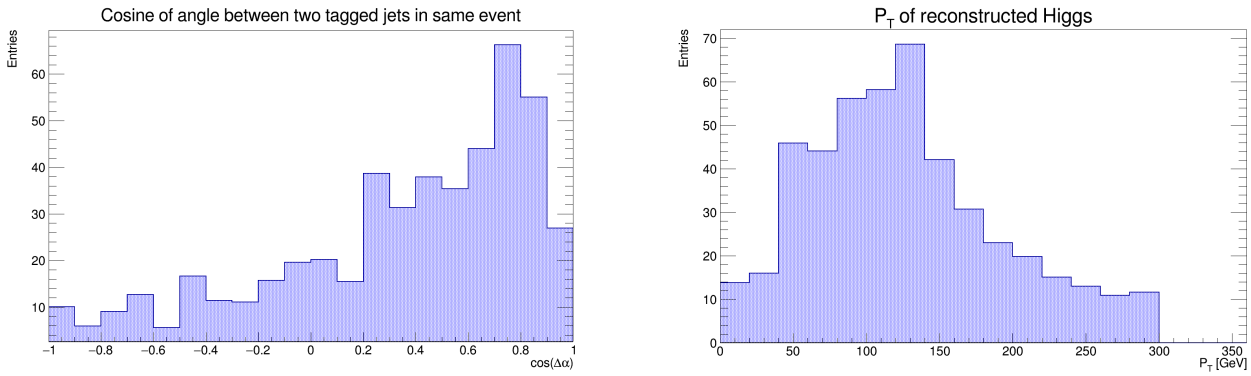
The SV-tagging efficiency and SV-mistag, as defined in equations 5.7 and 5.8, have distributions in perfect agreement with what presented in chapter 5.3.A cross checks on the normalized invariant mass distribution (shown in the appendix) was performed to make sure the Double Layer corrections were not creating any bias.
The SV-tagging efficiency on the H→ $b\bar{b}$ amounts to 0.672±0.019, while the mistag is found to be 0.024±0.002.
If in an event 2 tagged jets are found, their invariant mass is reconstructed. If more than 2 tagged jets are present, the two with highest p_T are selected. Each mass is weighted with the product of the DL corrections associated to each of the two jets used to reconstruct it. We can define the **selection efficiency** as

$$Eff_{rec} = \frac{Number\ of\ Events\ in\ the\ fiducial\ region}{Number\ of\ generated\ Events} \tag{6.9}$$

which tells the fraction of Higgs we are able to reconstruct. The overall reconstruction efficiency amounts to:

$$Eff_{rec} = 0.244 \pm 0.015$$

The invariant mass distribution is shown in figure 60. The Higgs peak is fitted with a double gaussian through the minimization of the χ^2. The double gaussian has been chosen to account for the different resolutions due to jets at different p_T have, as shown in figure 45. The mean of the gaussians is lower than m_H: this is likely due to the fact that some particle, such as particle produced with a sizable



(a) (b)

FIG. 59. Cosine of the angle between two tagged jets in the same events (a) and p_T of the reconstructed H (b), computed from the momenta of the tagged jets. The jets seem to be mostly close by, likely due to the sizable transverse boost the Higgs is shown to have.

transverse momentum component orthogonal to the jet direction or neutrinos arising from b-hadronic semileptonic decays, are not reconstructed in the jets, and as such, their contribution to the invariant mass is not accounted. The parameters of the fit are reported in table **12**. As mentioned at the start of the chapter, due to our computational limitations the statistic is still relatively small and more events would be needed to perform a fit with smaller uncertainty.

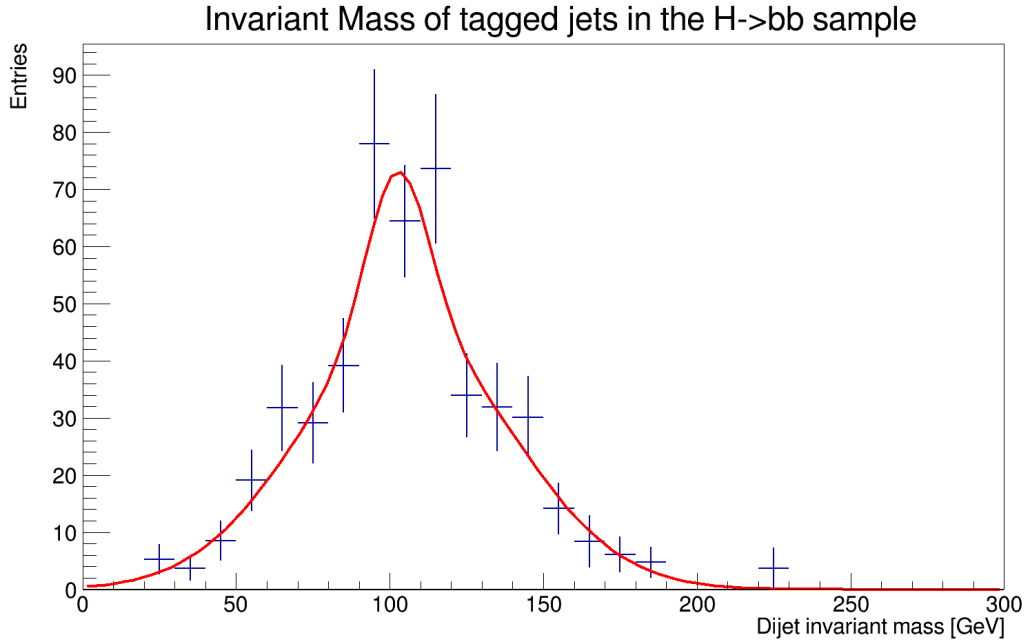


FIG. 60. Dijet invariant mass distribution for tagged jets in the H→ $b\bar{b}$ sample. A fit with a double gaussian has been performed to account for the different resolutions of different p_T jets: the parameters of the fit are reported in table **12**.

6.2 Extraction of $\sigma_{H\to b\bar{b}}$

Before extracting the uncertainty of $\sigma_{H\to b\bar{b}}$, we must discuss which processes constitute a background. Any process with two jets in the final state can, in theory, form a background. However, since the mistag is quite low and the cross section of the inclusive $q\bar{q}$ is not much higher than the heavy flavour inclusive cross section as shown in figure **61**, this background can be considered negligible at first approximation.

The remaining processes to be considered are the ones that have 2 heavy jets in the final state, and so

χ^2 /ndf	Const$_1$ (Entries/ 10 GeV)	σ_1 [GeV]	Mean$_1$ [GeV]
10.8/12	28.5±11.9	10.1 ± 3.1	102.5 ± 4.5
Probability	Const$_2$ (Entries/ 10 GeV)	σ_2 [GeV]	Mean$_2$ [GeV]
0.55	44.8±7.5	34.7 ± 2.7	105.4 ± 2.5

TAB. 12. Fit parameters of the double gaussian function plotted in figure 60. In the first column the χ^2 /ndf and probability of the fit. The parameters of the narrow gaussian are shown in the rest of the first row and are labeled by the subscript 1, while the ones of the wider gaussian are in the second row with subscript 2.

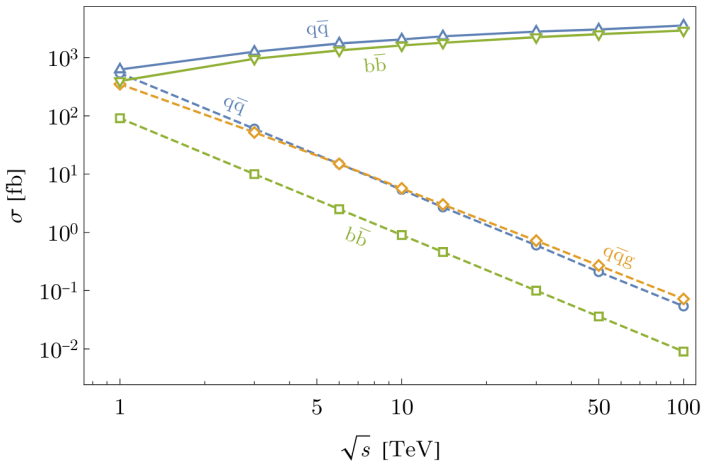


FIG. 61. Inclusive cross sections of di-jets process at the Muon Collider as a function of $\sqrt{s}$. The dashed lines represent annihilation processes, while the continuous line indicate vector boson fusion channels. Taken from [15].

we are going to study the $b\bar{b}$ and $c\bar{c}$ inclusive background. The s-channel with Z/γ^* as propagators is a subdominant component. The reconstructed masses of these events are mostly well above M_{jj} =300 GeV, maximum of the fiducial region we chose in the last section. Important background is instead given by the VBF producing a Z which decays in heavy quarks, whose diagram is shown in figure 62. This process is the dominant one of the 2000 events generated at L.O. with WHIZARD [48], requiring

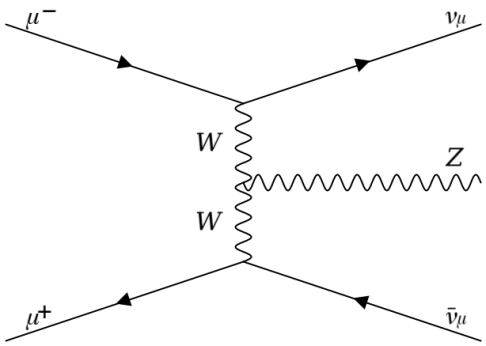


FIG. 62. Feynman diagram of the dominant background process: a WW fusion producing a Z which then decays in $b\bar{b}$ or $c\bar{c}$.

the final state to be $\nu\bar{\nu}b\bar{b}$ or $\nu\bar{\nu}c\bar{c}$. The angular and p_T requirements on the final quarks were the ones described for the H→ $b\bar{b}$. The total cross section of the background is reported in table 13 and compared with the Higgs one. It must be noticed that the cross section of the $b\bar{b}$ (338 fb) background is slightly higher than the one of the $c\bar{c}$ (271 fb).This background unfortunately for us, has kinematic characteristics very similar to the one of the signal. This is shown in figures 63 and 64: the p_T and angular distribution of the jets are nearly identical to the ones of the H→ $b\bar{b}$.

Also the other event variables, such as the angle between the two tagged jets and the $\sum p_T$ of the two

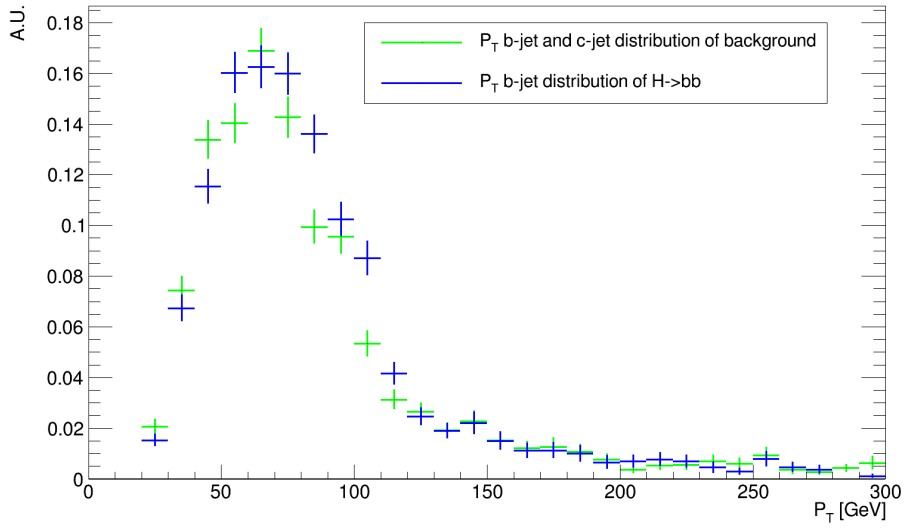


FIG. 63. P_T distributions of jet matched to a b-quark in the H→ $b\bar{b}$ sample (blue) and matched to a b- or c-quark in the background sample (green).

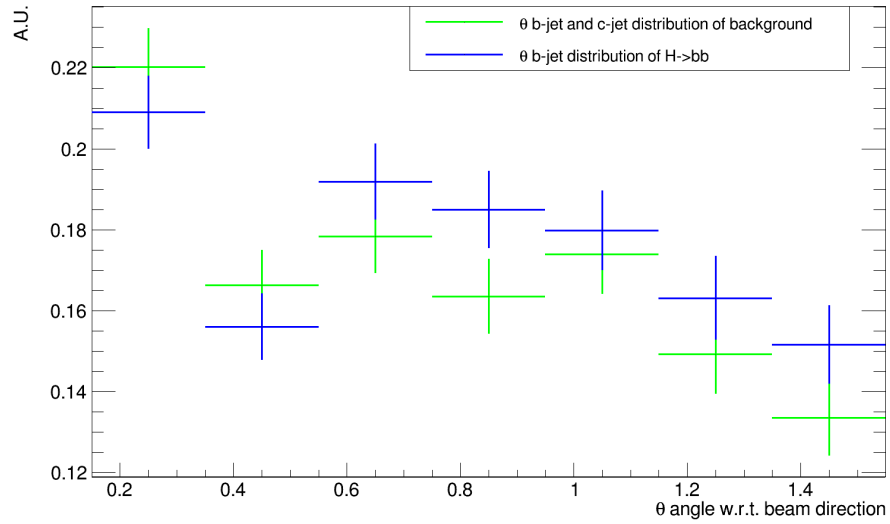


FIG. 64. θ distributions of jet matched to a b-quark in the H→ $b\bar{b}$ sample (blue) and matched to a b- or c-quark in the background sample (green).

tagged jets have a distribution compatible with what presented for the Higgs in figure 59. The total tagging efficiency of this sample is found to be 0.569±0.016, while the mistag is 0.0119±0.0006. The value of the integrated efficiency is smaller compared to the one of the H→ $b\bar{b}$ due to the presence of events with c-jets instead of b-jets, which have a SV less displaced and with less tracks, which implies overall a lower tagging efficiency, as we will see in chapter 7.1. The selection efficiency, that is the ratio of number of events with 2 tagged jets and invariant mass in the fiducial region over the total number of events, amounts to

$$Eff_{rec} = 0.155 \pm 0.011.$$

The number of expected events assuming an integrated luminosity of 1 ab^{-1} (with instantaneous luminosity $\mathscr{L}_{inst} = 1.810^{34}$ cm^{-2} s^{-1} [49] and a data taking period of $t = 5 \cdot 10^7$ s) is computed as $N = \sigma \cdot \mathscr{L}_{int} \cdot Eff_{rec}$ and compared with the one of H→ $b\bar{b}$ in table 13. We can see that the background is expected to be of the same order of magnitude of the signal, but to see how this affects our measurement we must also study how the invariant mass of 2 tagged jets is distributed.

The invariant mass distribution has been fitted with a gaussian in figure 65 and the parameters of the fit are presented in table 14. The background peaks at the Z boson mass, as expected from the dominant diagram. This peak is close to the Higgs one (which, as said in the previous section, is

Process*	σ [fb]	Eff_{rec}	Events with $\mathscr{L}_{int} = 1$ ab^{-1}
H$\to b\bar{b}$	324.0	0.244 $\pm$ 0.015	79.1k
$b\bar{b}$ + $c\bar{c}$ background	610.7	0.155 $\pm$ 0.011	94.8k

TAB. 13. Summary table of cross section σ, selection efficiency Eff_{rec} and number of expected events for the various channels considered in this analysis. The * near the "Process" label is a reminder that such processes are generated with the requirements specified at the start of this chapter.

shifted to masses lower than $m_H = 125$ GeV due to missing particles) as can be seen from figure 66, however the two contributions can still be distinguished.

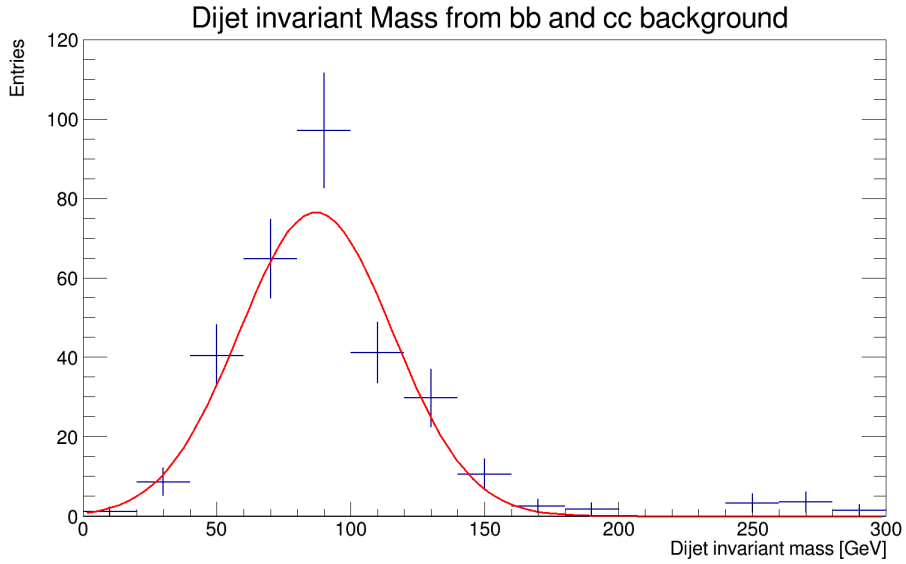


FIG. 65. Dijet invariant mass distribution obtained from tagged jets of the background sample. A peak is present in correspondence to the Z boson mass. This is expected since the main background channel is the WW fusion into a Z boson. More statistics is needed to obtain a more precise distribution and be able to perform a fit with smaller uncertainties.

χ^2 /ndf	Probability	Const (Entries/ 20 GeV)	σ [GeV]	Mean [GeV]
14.6/10	0.15	76.5$\pm$7.9	28.6 $\pm$ 1.9	87.0 $\pm$ 2.5

TAB. 14. Fit parameters of the gaussian function plotted in figure 65. The normalized χ^2 and the errors on the fit parameters are still elevated: more statistics is needed to extract them with smaller uncertainty.

From the fits we performed on the H$\to b\bar{b}$ dijet mass distribution and to the $b\bar{b}$ + $c\bar{c}$ background in this one, we can extract the probability distribution function (PDF) for the signal and the background. The use of the PDF instead of the histograms is chosen to mitigate the effects of the low statistic present in particular in the background sample: in each bin the statistical fluctuations are elevated and this would greatly affect a template distribution created from this histogram.

Each of these PDF is then normalized by the number of expected events of each category, presented in table 13. From these distributions, 2500 pseudo-experiments were performed with Roofit [50]. In figure 67 one of these pseudo-experiments is represented by the black dots, where the total mass spectra is depicted by the blue line, while the red and green distributions indicate the contributions of H$\to b\bar{b}$ and the background respectively.

In each pseudo-experiment, an unbinned extended maximum likelihood fit has been performed, leaving both signal and background yield as free parameters. On each pseudo-experiment, the relative uncertainty on the $\sigma_{H\to b\bar{b}}$ has been extracted from the fit using N_H the normalization parameter of the H$\to b\bar{b}$ distribution, corresponding to the number of signal events. The relative uncertainties of each pseudo-experiment have been plotted in 68 and the relative uncertainty on the cross section has

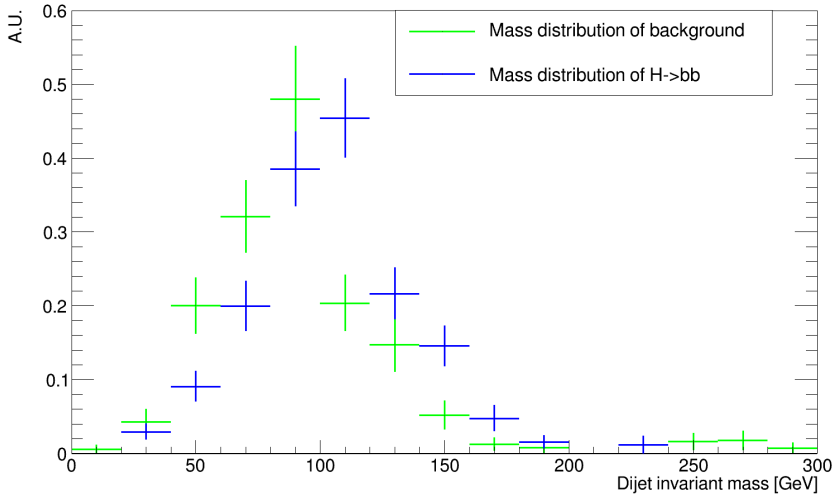


FIG. 66. Normalized distributions of the dijet invariant mass from the H$\to$ $b\bar{b}$ (blue) and the $b\bar{b}+c\bar{c}$ background (green).

been extracted with a gaussian fit. The fit parameters are shown in table 15. Finally, the relative statistical uncertainty on $\sigma_{H\to b\bar{b}}$ with 1 ab^{-1} is found to be:

$$\frac{\Delta\sigma_{H\to b\bar{b}}}{\sigma_{H\to b\bar{b}}} = 0.8\%$$

This is less precise compared to what found by CLIC, which expects a relative uncertainty of 0.3% with 2ab^{-1} (from [4]). However, since the reconstruction algorithms and the analysis procedure are not fully optimized, this result can be further improved and only serves as a starting benchmark.

χ^2 /ndf	Probability	Const Entries	σ [GeV]	Mean [GeV]
27.1/33	0.76	194.3±4.8	$(6.10 \pm 0.09)10^{-5}$	$(8.214 \pm 0.001)10^{-3}$

TAB. 15. Fit parameters of the double gaussian function plotted in figure 65. The normalized χ^2 and the errors on the fit parameters are still elevated: more statistics is needed to extract them with smaller uncertainty.

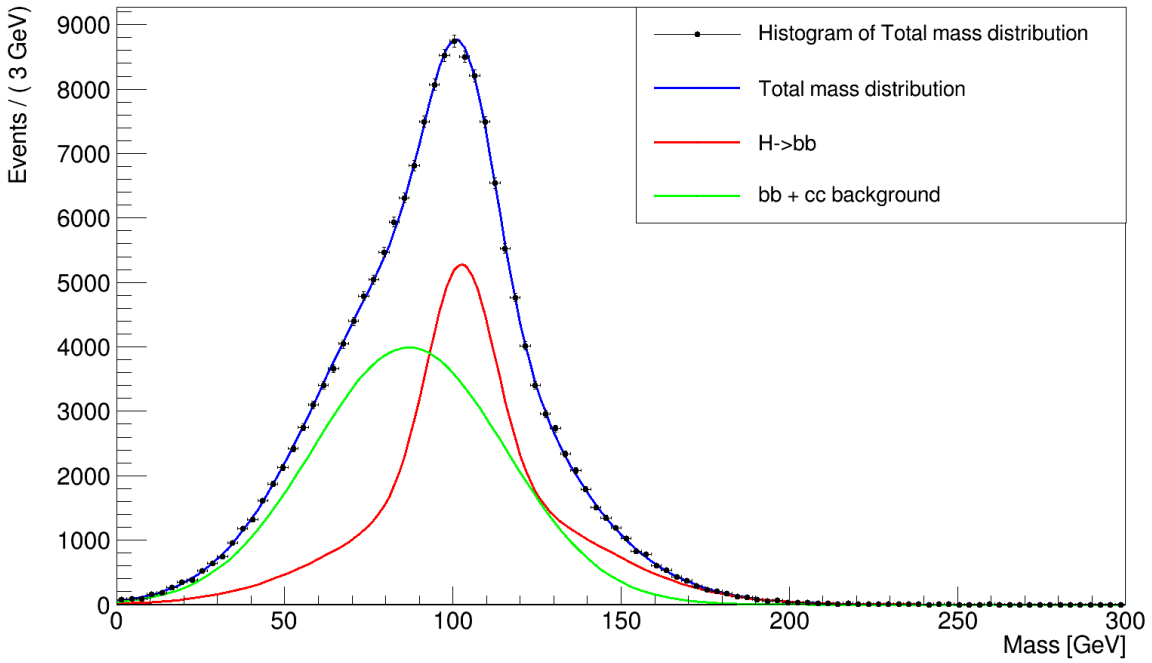


FIG. 67. Distributions of one pseudo-experiment, where the total mass spectra is represented by the blue line, while the red and green distributions indicate the contributions of H$\to b\bar{b}$ and the background respectively.

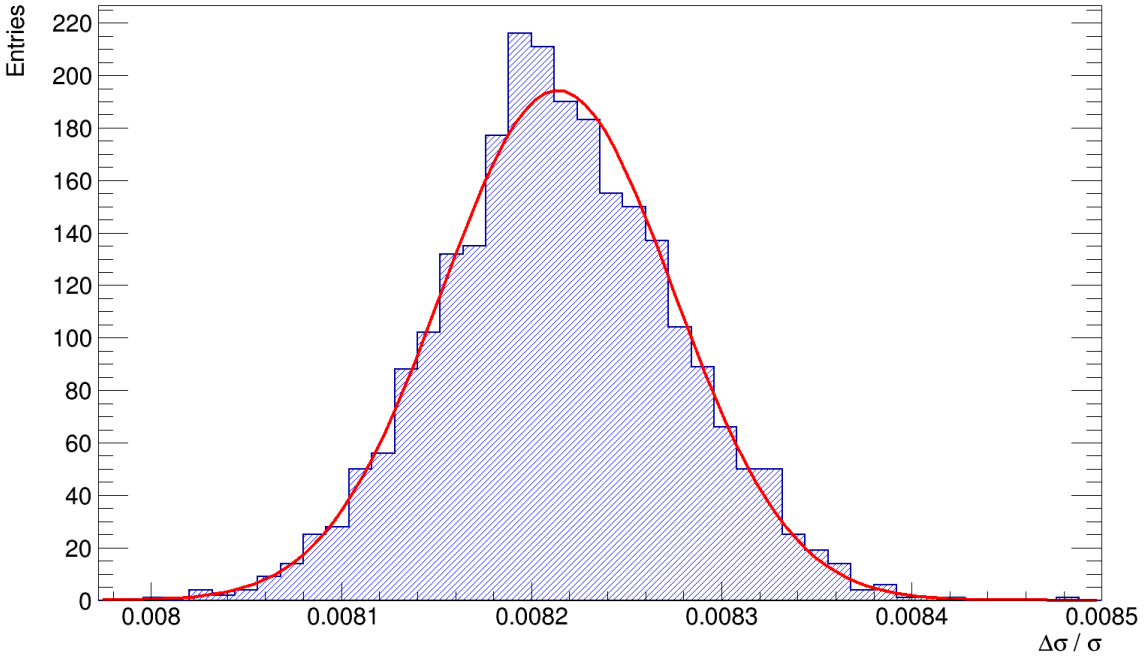


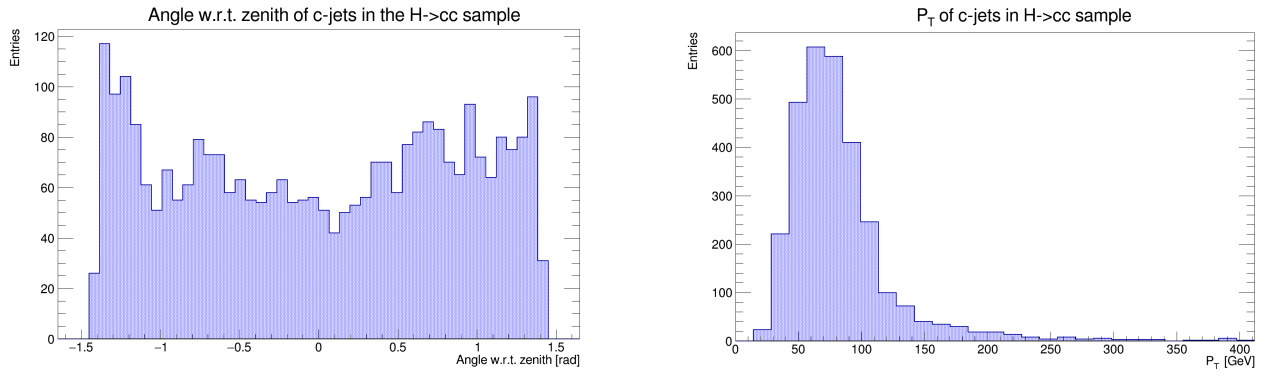
FIG. 68. Distribution of $\Delta\sigma/\sigma$ of the 2500 pseudo-experiments. A gaussian function has been fitted to this distribution and the mean parameter of the fit is used to estimate the final relative uncertainty on the cross section.

7. c-jets identification and prospects on H→ $c\bar{c}$

In this chapter we will apply the tagging strategy used for the H→ $b\bar{b}$ to the H→ $c\bar{c}$. However we will see this procedure has two main shortcomings: first, it becomes extremely difficult if not outright impossible to distinguish b and c jets; in addition, the c-hadrons travel a distance shorter than b-hadrons before decaying and so their SVs are more difficult to reconstruct. This, combined with the fact that the cross section of H→ $c\bar{c}$ is about 20 times smaller than the one of H→ $b\bar{b}$, assures it is not possible to measure $\sigma_{H\to c\bar{c}}$ with this strategy. For this reason, in the second section of this chapter we are going to study the most significant features of b- and c-jets, comparing them with also the ones from light flavour quarks. The observables will be used with Machine Learning techniques to distinguish the jet flavour and allow to identify H→ $c\bar{c}$ in the future. A first attempt is described in the last section, where some preliminary tests of a Deep Neural Network (DNN) for jet flavour identification are presented.

7.1 H→ $c\bar{c}$ with SV-tagging

Since SVs are going to provide several jet-flavour discriminant variables, here we see in how many c-jets we can find an SV and which resolution we have on the Higgs mass.
A sample of 2000 events of H→ $c\bar{c}$ has been simulated with PYTHIA8 at L.O., with the same acceptance cuts employed for the H→ $b\bar{b}$. As expected from the H decay branching ratios, the cross section is much smaller than the bottom case: $\sigma_{Hcc} = 14.6$ fb.
c-jets have in general similar characteristics of the b-jets of the H→ $c\bar{c}$, as shown in figure 69: many of them are directed in the forward/backward direction and their p_T lay mostly in the region where our jet resolution performances are higher. However some differences are present in the SVs: since



(a) (b)

FIG. 69. Angle respect to zenith and p_T of jets matched with a c-quark in the H→ $c\bar{c}$ sample.

c-hadrons have a smaller lifetime than the b-hadrons, their decay vertex is less displaced. This implies that the SV-tagging efficiency will be lower, since the tracks from their SVs are more likely to be associated with the PV. However, as shown in 5.3, this means that the DL is affecting them less. Another important consideration is that the number of tracks coming from a c-hadron is on average

smaller than the one from a b-hadron vertex: the SV is less likely to be found. Overall we expect a smaller SV-tagging efficiency. The same considerations can be applied also to the mistag: when a b- or c-jet is not recognized by the matching, the b-jet is more likely to be tagged w.r.t. a c-jet; moreover, considering a couple of collinear jets, if they originate from a b-quark they will contain more tracks (see also 7.2) which may more easily give rise to a SV.
The integrated SV-tagging efficiency is 0.381±0.016 and the integrated mistag found in the H→ $c\bar{c}$ is 0.0159 ± 0.0067, both smaller than what found for the H→ $b\bar{b}$. Using the same fiducial region defined for the H→ $b\bar{b}$, the selection efficiency (as defined in equation 6.9) is found to be

$$Eff_{rec} = 0.095 \pm 0.008$$

which is obviously lower than the b case due to the lower SV-tagging efficiency. A summary table of the cross section, selection efficiency and expected events with $\mathscr{L}_{int} = 1$ ab^{-1} is shown in table 16.

Process*	σ [fb]	Eff_{rec}	Events with $\mathscr{L}_{int} = 1$ ab^{-1}
H→ $c\bar{c}$	14.6	0.095 ±0.008	1.4k

TAB. 16. Summary table of cross section σ, selection efficiency Eff_{rec} and number of expected events for the H→ $c\bar{c}$. The * near the "Process" label is a reminder that such processes are generated with the requirements specified at the start of this chapter.

The mass distribution obtained from 2000 events of H→ $c\bar{c}$ is shown in figure 70. Due to the smaller reconstruction efficiency, the statistics is lower than what used for the H→ $b\bar{b}$. However the Higgs peak is clearly present, even if the gaussian fit gives quite a large width of the peak, as can be seen in table 17. Since we cannot distinguish b- and c-jets, it is not possible to measure the H→ $c\bar{c}$ with SV-tagging. However we proved that the Higgs peak is visible and could be measurable if we had a reliable way to tag the jets flavour. Finding a way to do this will be the final objective of the next sections.

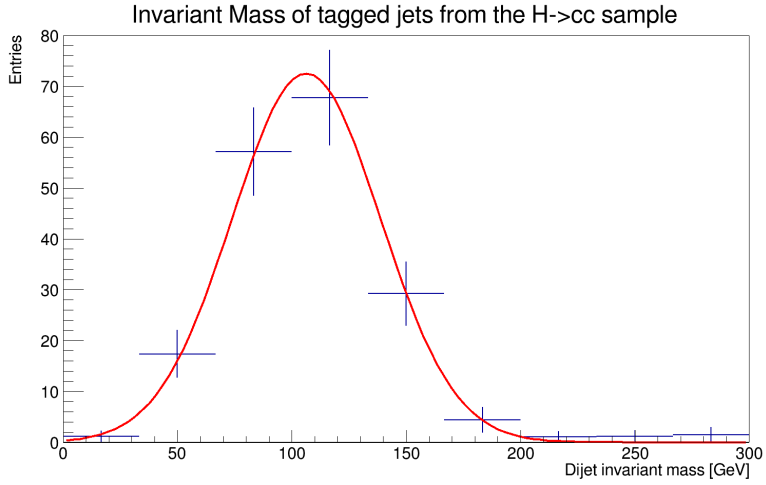


FIG. 70. Invariant dijet mass distribution for the H→ $c\bar{c}$. The gaussian fit shows a resolutions of $\sigma = 32 \pm 2$ GeV. The mean is slightly shifted toward lower values of the invariant mass due to missed particles in the jets.

χ^2 /ndf	Probability	Const [Entries/(33 GeV)]	σ [GeV]	Mean [GeV]
2.9/6	0.82	72.5±7.6	32.5 ± 2.0	106.3 ± 2.8

TAB. 17. Fit parameters of the gaussian function plotted in figure 70. The binning is wide to compensate for the low statistics. As for the H→ $b\bar{b}$, the peak is shifted w.r.t. to the nominal Higgs mass due to particles not reconstructed in the jets and imprecisions in the jet energy correction.

7.2 Study of jets observables for jets identification

For a deep study of jets, the following samples were reconstructed (with the BIB overlayed) and analyzed:

- 4 sets of 1500 events of $b\bar{b}$ inclusive events, generated with PYTHIA8 at L.O., divided in the following ranges of $\sum p_T$ of b-jets: [40-80,80-120,120-160,160-200] GeV
- 4 sets of 1500 events of $q\bar{q}$ inclusive events, with q any quark, generated with PYTHIA8 at L.O., divided in the following ranges of $\sum p_T$ of jets: [40-80,80-120,120-160,160-200] GeV
- 4 sets of 1600 events of $c\bar{c}$ inclusive events, with q any quark, generated with of PYTHIA8 at L.O., divided in the following ranges of $\sum p_T$ of c-jets: [40-80,80-120,120-160,160-200] GeV
- 2000 events of H$\rightarrow b\bar{b}$, 2000 events of H$\rightarrow c\bar{c}$, 2000 events of $b\bar{b}$ background, 2000 events of $c\bar{c}$ background and 2000 events of H$\rightarrow WW \rightarrow c\bar{c}X+ \rightarrow c\bar{c}$, all generated with PYTHIA8 at L.O.
- 2000 events of $b\bar{b}$+$c\bar{c}$ background from the WW fusion, generated with WHIZARD at L.O.

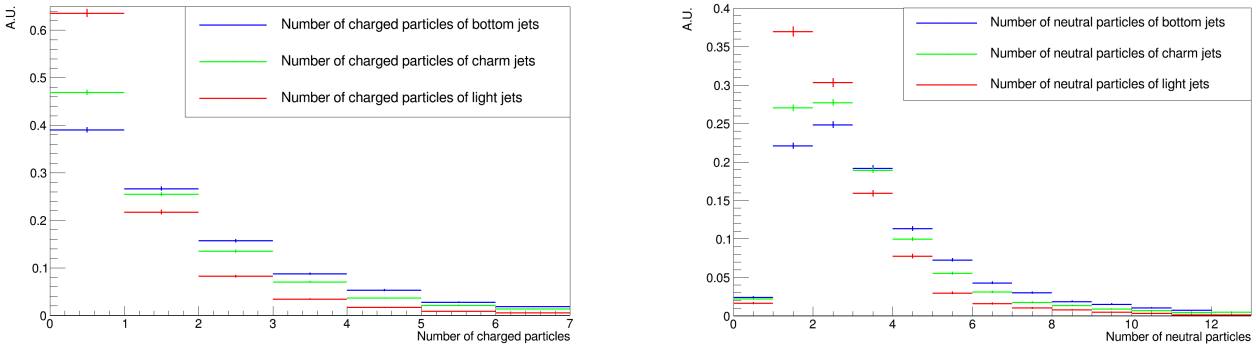
Due to the variety of processes under our consideration, we have to be careful not to choose observables strongly related to the specific process and not to the type of jet: for example, kinematic variables such as p_T and p_Z of the jets were found to be process dependent and were consequently discarded. Ideally we would want to train our machine learning algorithm on only the dijet samples; however the statistics available is very scarce for a machine learning problem (and not all jets in the event can be used for training, but just the ones matched with a Monte Carlo quark or gluon). In fact, reconstructing a sample with the BIB overlayed is an heavy and slow task. Due to this computational limitation, we are going to use every jet at our disposal, supposing many more jets will be available when this strategy can be implemented at its fullest potential. For this reason, this study must be considered more a preliminary estimation than an evaluation of the possible capabilities of this technique at the Muon Collider.
Many observables were studied, however some of them were seen to be not very discriminating and were removed to diminish the number of parameters of the algorithm. An example are the variables related to the presence of μ in the jet, together with the component of their momenta orthogonal to the jet direction and $R = \sqrt{\eta^2 + \phi^2}$ w.r.t. the jet axis. We will divide the observables in several paragraphs depending on what they are describing: the jet globally; the secondary vertex, if present; or the substructure of the jet. In each paragraph we are going to show the normalized distributions of such variables for the different classes of jets: b-jets, c-jets and light jets. The flavour of each jet is identified with the matching procedure described in chapter 5.3.

Global jet variables: Examples of variables describing the jet as a whole are the kinematic variables (such as p_Z, p_T and E) or the ones describing the number of particles inside the jets. However, the kinematics variable were found to be very process dependent: for example many processes with b-jets were mostly directed forward, while this was not true for the lighter jets. Since we want an algorithm that works for every process, we cannot use them. We can however look for kinematic invariants, such as the mass of the jet, which does not strongly depend on the process.
This will be a general rule in the whole chapter: we will always try to use variables invariants under some basic transformations, such as rotations of the jets around its axis and translations of the jet in (η, ϕ). This will be checked by comparing jets of the same flavour between different samples and checking the distribution of their variables has not a completely different shape. After these selections, we remain with only 4 variables for the global sector:

- the mass of the jet;
- the number of charged particles, found from the number of tracks in the jets;
- the number of neutral particles, given by the number of clusters that PandoraPFA could not associate with a track;

- the absolute value of the weighted charge of the jet, computed as the weighted average of the charge of the tracks, where the weights are the ratio of the p orthogonal to the jet axis over the total p orthogonal of charged tracks $wi = {p_{ort,i}}/{p_{ort,tot}}$.

The distributions of the number of charged and neutral particles are shown in figure 71.



(a) (b)

FIG. 71. Normalized distributions of number of charged particles (a) and neutral particles (b) in bottom jets (blue), charm jets (green) and light jets (red). The bottom jets seem to have on average more particles (charged and neutrals) than the others, while the light jets are the less populated.

Secondary vertex variables: Already, without looking at the SV variables, the presence or absence of a SV is a good indicator on whether the jet can be heavy or light. Unfortunately the SV-tagging efficiency here cannot be corrected easily for the effect of the Double Layer: as such it is not very high (slightly more than 50% for b-jets and around 33 % for c-jets) and this will remove in many jets a very powerful discriminant.
If more secondary vertices are present in the jet, the variables we are going to discuss will always refer only to the one with highest p_T. If no SV is found in a jet, all these variables are set to 0.
Guided by general considerations on the b-hadrons and c-hadrons lifetime and decay topology, we chose 9 observables:

- the number of tracks coming from the vertex;
- the absolute value of the sum of the charges of tracks coming from the SV, $|\sum_{trks} q_{trks}|$;
- the ratio of p_T of the SV over the p_T of the jet;
- ΔR between the jet axis and the SV position with respect to the PV (flight direction);
- the projection of the position of the SV along the jet axis;
- the projection of the position of the SV in the plane orthogonal to the jet axis;
- the mass M of the SV;
- the time of flight, computed as $t = {dM}/{cp}$, where d is the distance between SV and PV;
- the corrected mass, defined as the minimum mass consistent with the direction of flight, computed as $M_{corr} = \sqrt{M^2 + p^2(\sin\theta)^2} + p\sin\theta$, where p is the momentum of the SV and θ is the angle between the momentum and the direction of flight of the SV.

Some of the more discriminant variables are plotted in figure 72.

Substructure The substructure of the jet may give important information on the nature of the jet, especially if no SV is found. Initially the substructure information was studied in two blocks: one for the neutral particles and one for the charged ones. The observables were features of the 5 highest p_T charged particles in the jet or the 8 highest p_T neutral particles.

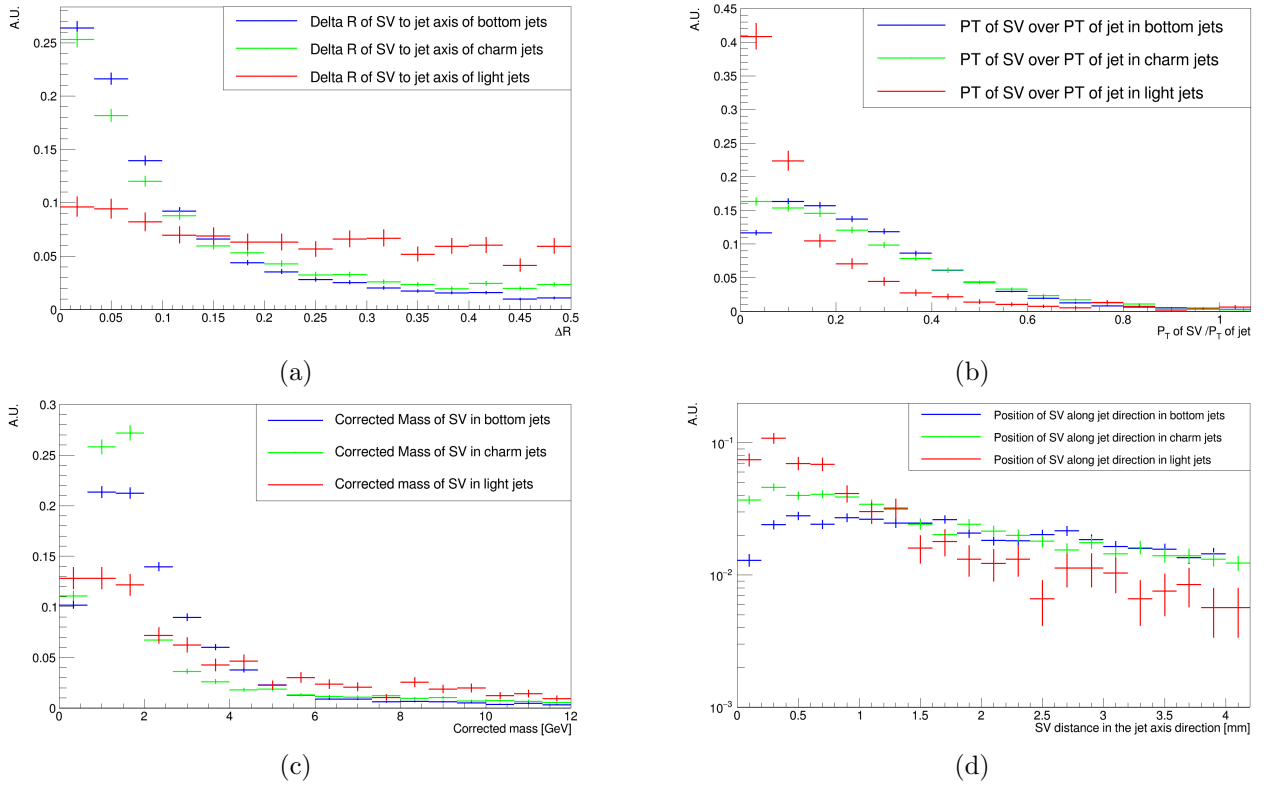
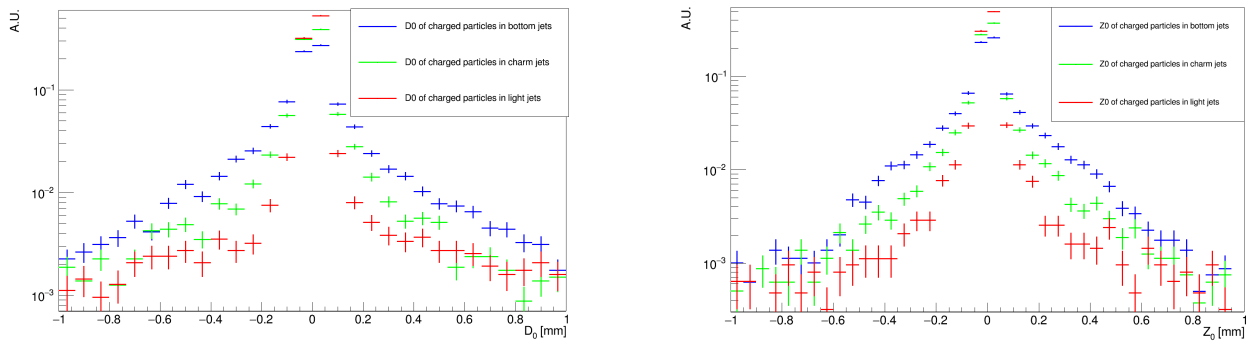


FIG. 72. Normalized distributions in bottom jets (blue), charm jets (green) and light jets (red) of: ΔR between the jet axis and the SV flight direction (a); the ratio of p_T of the SV over the p_T of the jet (b); the corrected mass (c); the flight direction of the SV along the jet axis (d). From (a) we see that the SVs of b- and c- are directed more along the jet axis compared to SVs in light jets. The SVs in the light jets have also a lower p_T (b) w.r.t. to the one of their jet. The corrected mass variable, as expected, peaks (c) for the charm jets around 1.5 GeV. The SV position (d) along the jet direction is smaller of the SVs of light quarks; at higher Z, outside the plot, one can also see that the tail of the bottom jets is higher than the one of the charm. It is likely that the SVs in the light quarks are formed using also tracks from the BIB, as they are close to the IP (where the BIB occupancy is higher), with low p_T, as the BIB has, and their random position inside the jet, showed by the flat ΔR distribution.

For each charged particle initially we used: p component parallel to the jet axis; p component orthogonal to the jet axis; the track parameters D_0, Z_0, Ω, $\tan(\lambda)$; the ratio of the charge of the track over the module of its momentum.
For the neutral particles we had no track related information, so we just used the p component parallel to the jet axis and the p component orthogonal to the jet axis. However during the tests on the DNN described in the next chapter, we noticed that removing or keeping the substructure variables, the output of the DNN had negligible changes: it was not using the substructure information.
So we studied the features of the particles and noticed that many were nearly identical, with only some variables being a little different in the different flavour of jets, such as the D_0 and the Z_0 of the particle with highest p_T, reported in figure 73. It became apparent to us that the DNN, with so little statistics, was not able to build discriminant enough features from the observables.

For this reason we decided to try to build global jet variables combining the information of the substructure. The idea is that if the network is not able to build discriminating features on its own due to the lack of statistics, we can try to feed it more discriminant variables and see if it is able to use them. For example, instead of giving all the D_0 of the charged particles (so 5 parameters), we can give the sum of the absolute value of the D_0 of all the particles. We expect this sum to be bigger for b- and c-jets, due to the contribute in the sum of tracks coming from the SV; this should account for the cases in which the SV is not reconstructed, but one of the tracks coming from it (which should have an higher impact parameter) is. In this way, we are giving only one parameter instead of 5, and



(a) (b)

FIG. 73. Normalized distributions of D_0 (a) and Z_0 (b) of charged particles in bottom jets (blue), charm jets (green) and light jets (red). The bottom jets have more displaced tracks compared to the others, as expected from the lifetime of the B meson, but the difference between charm and light jets is also visible, even if smaller.

since it is built to be more discriminant than each of the singular track parameter, we hope that the DNN is able to use it to perform the classification. When the statistic will be larger, the DNN will be able to extract the most interesting features by itself, and likely it will be able to find combinations better than the one we are going to use in this study. For the moment, however, this isn't the case and as such we will be using some global observables describing the substructure. These variables are

- for the **charged particles**: the average and standard deviation of the orthogonal component of the momentum p_{ort} of each charged particle; the average and standard deviation of the ratio of p_{ort} and p_{par} (the component parallel to the jet axis of the momentum) of each charged particle; the sum and standard deviation of $|Z_0|$; the sum and standard deviation of $|D_0|$; the average of Z_0/D_0; the maximum $|D_0|$ in the jet; the maximum $\sqrt{D_0^2 + Z_0^2}$ in the jet; the minimum, maximum and average value of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ of the tracks with respect to the jet axis;
- for the **neutral particles**: the average and standard deviation of the orthogonal component of the momentum p_{ort} of each neutral particle; the average and standard deviation of the ratio of p_{ort} and p_{par} (the component parallel to the jet axis of the momentum) of each neutral particle.

Overall the substructure is represented by these 14 variable for the charged particles and 4 for the neutrals. The distributions of the sum and the standard deviations of $|Z_0|$, the average of p_{ort}/p_{par} for the neutral particles and the average of R of the charged particles are represented in figure 74.

All these variable are going to be used as input for the Deep Neural Network described in the next section.

7.3 Machine learning techniques for future H→ $c\bar{c}$ extraction

A neural network (NN) is a set of algorithms that, using correlation between features, is able to cluster and classify data. These networks receive as input a huge number of variables describing an event and, after undergoing some computation in its layers, the last layer gives as output the probability that the event is of each of the available classes. A NN is made up by layers. The first layer (called "Visible layer" or "Input layer") receive as input the features of the events and passes its output to the second layer, which uses it to produce another output then passed to the following layer, and so on until the last layer (called "Output layer"). All the layers between the Input and the Output one are called "Hidden layers". Each layer is made up by neurons, which take some numbers as input (which may be the starting variables or the output of other neurons) and give another as output.
A Deep Neural Network (DNN) is a neural network with more than one Hidden Layer. An example of a DNN is shown in figure 75.

The building block of a DNN is the neuron. A neuron computes a weighted sum of all the inputs it receives from the previous layers. This sum, after it is subtracted by a parameter called bias, is then

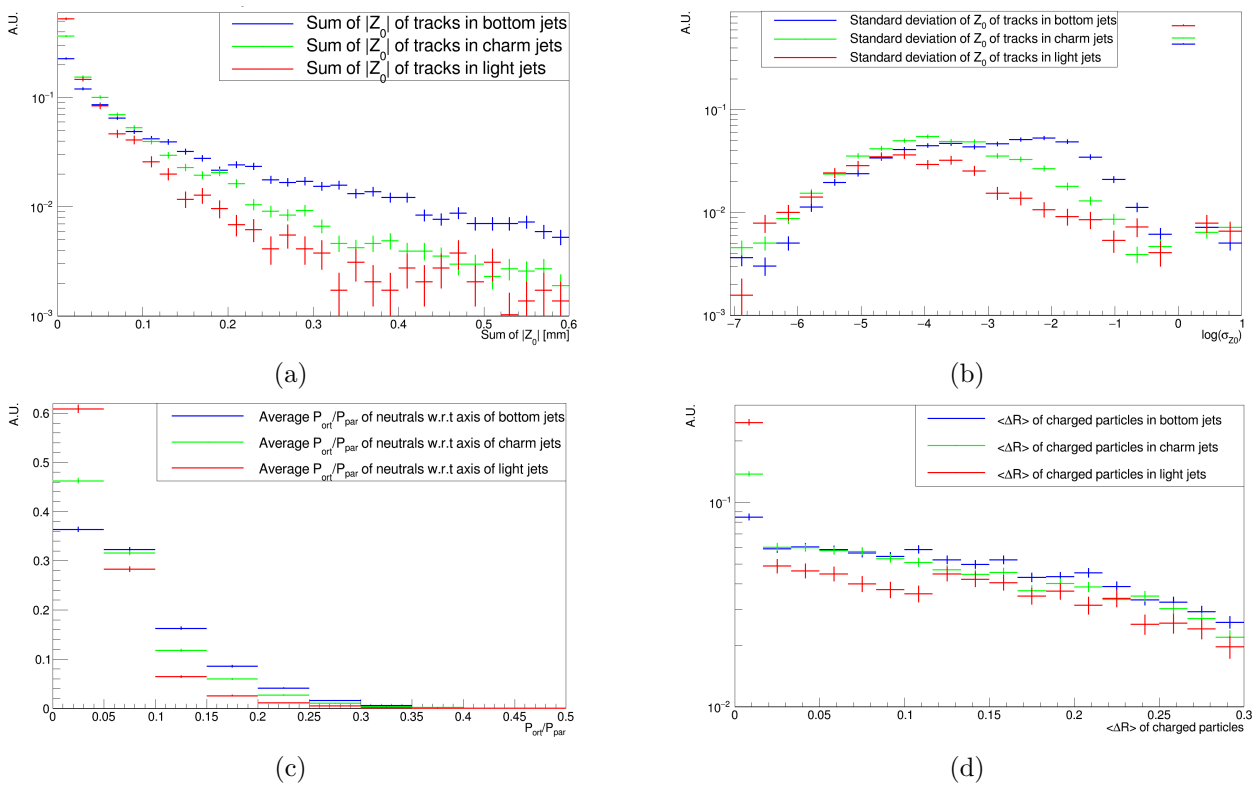


FIG. 74. Normalized distributions in bottom jets (blue), charm jets (green) and light jets (red) of: sum of $|Z_0|$ of the tracks (a); logarithm of standard deviation of Z_0 of the tracks (b), set to 0 if there are less than 2 tracks; p_{ort}/p_{par} for the neutral particles (c), set to 0 if neutral particle is present; average of ΔR of tracks w.r.t the jet axis (d), set to 0 if no charged track is present. The displaced tracks coming from SVs cause the Z_0 of b-jets to have higher absolute value and have bigger deviation ((a),(b)). The decay of the B also produces particle with momenta less aligned with jet axis, as shown in figure ((c),(d)).

fed to an activation function. The value of the activation function is the output of the neuron.

All the weights and the biases used by each neuron are free parameters that the DNN is gonna set in order to be able to correctly classify objects. It does this with the "training" procedure, in which labeled events are shown to the DNN which tries to set the parameters to classify as precisely as possible the labeled data. This is called back-propagation: the weights and biases are changed accordingly to the gradient of a scalar function proportional to how close was the guess to the true event class, called loss-function.
The DNN used in this thesis is inspired by the one [51] proposed by the LHCb INFN group of Padova to tag the jet flavour in LHCb. Their proposal has been adapted to our software environment and undergone a careful selection of variables, which were listed in the previous section, to fit in our limited computational resources.
The DNN is tasked to classify jets in 3 classes: jets coming from a b-quark, jets coming from a c quarks and jets coming from light quarks or gluons. Its scheme is showed in figure **76** and consists of three main blocks. The input variables are divided in two groups, global variables or secondary vertex variables, and each given to one of first two blocks. Several different layouts of the network were tried: for brevity's sake we are just going to describe one of them and the corresponding results.

This first two blocks have first a "Batch normalization" layer, who renormalizes the scale of each variable, and then a small network of dense layers, where the inputs are summed with some weights, as described at the start of the section. For the global variables, there is one dense layers with 14 neurons. On the other hand, the SV block has after the "Batch normalization" just one layer with 8 neurons. The output of these two blocks is used as the inputs of the last one, which proceeds with 4 dense layers of 22, 22, 8 and 3 neurons respectively. The results of last layer are the output probability of the jets being a b-jet, a c-jet or a q-jet. The total number of parameters of the network amounts

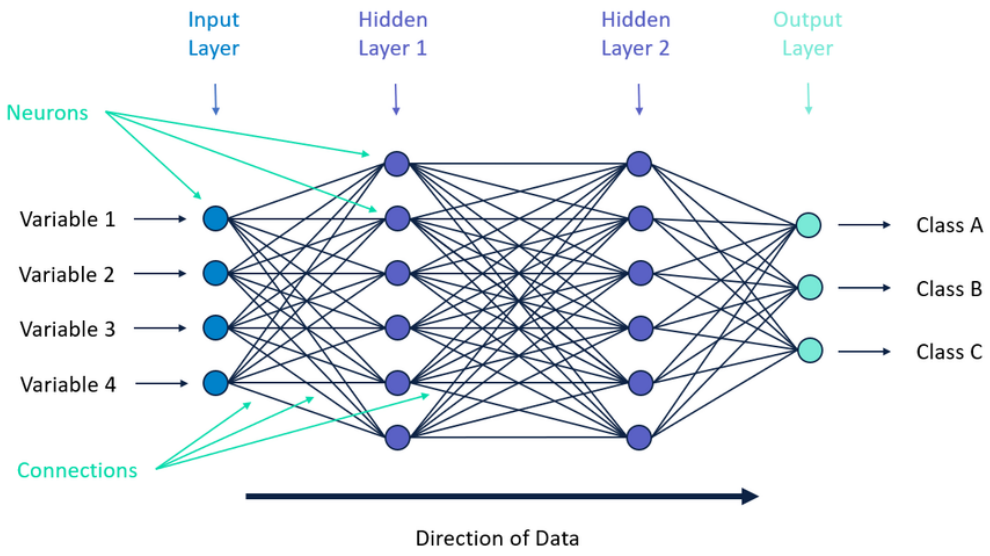


FIG. 75. This picture shows an example of a DNN, which takes 4 variables as input, has 2 hidden layers and 3 output classes. Each of the colored dot is a neuron, and the vertical formations of neurons are the so called layers. Every neuron in each layer takes as input the output of all neurons in the previous layers (or the input variables if the neuron is in the Input Layer) and sends its output to all the neurons in the following layers, as shown by the black lines. In the end one obtains the probability that the event is class A, the probability it is class B and the same for class C.

to 1749, of which 1687 are trainable with the back-propagation.
The activation function of all but the last layer is the rectified linear unit "Relu", which gives as output the weighted sum itself if it is a positive value or gives 0 if the sum is ≤ 0. Each of these layers are then followed by a dropout layer, which deactivates each neuron of the corresponding dense layer with probability 0.1. The dropout layer are used to avoid that the network depends too much on a single neuron. In the last layer instead, the prediction is made with the Softmax activation function. The Softmax is defined as $\sigma(x_j) = e^{x_j}/\sum_i e^{x_i}$ where i are all the possible outcomes. It gives output in the [0,1] range and normalizes their sum to 1, so that its outputs can be treated as probabilities. Another interesting feature is that it highlights the larger inputs while suppressing the smaller ones. Due to these properties, it is often used with the categorical cross-entropy loss-function which we are going to use as it is suited to our multi-class classification problem. The categorical cross-entropy is defined as:

$$CE = -\sum_d \sum_i y_{id} \cdot log(\hat{y}_{id}) + (1 - y_{id}) \cdot log(1 - \hat{y}_{id})$$

where d is an index running on the events, i is an index running on the data classes, y_{id} is the true label of the event d (y_{id}=1 if the event d is of class i and y_{id} =0 if it is of another class), and $\hat{y}_{id}$ is the probability that the network assigns to the event d to be of class i. The minimization of the loss function is made with ADAM optimizer. ADAM is a gradient descent optimizer which uses a running average of the first and second moments of the gradient, respectively m and s, to adaptively change the learning rate. It is defined in figure 77, where θ are the parameters of the function to minimize and t is an index counting the step number. As ADAM parameters, we set the normalization parameter at $\epsilon = 10^{-7}$, the memory lifetime of the first and second moments respectively at $\beta_1 = 0.9$ and $\beta_2 = 0.99$ and the learning rate η at $3 \cdot 10^{-4}$.

Our final jet sample is made up by N_b =12.7k b-jets, N_c =12.7k c-jets and N_q =7.6k light jets. Of these, 18% are going to be used as test samples, while the rest are used for the training. The fact that the samples have different dimensions may affect our results: to eliminate this bias, we re-weight the light jet sample with a factor $w_l = N_b/N_q$ in the loss computation.
The training sample has been divided in mini-batches of 256 jets. Since the statistic in each bunch

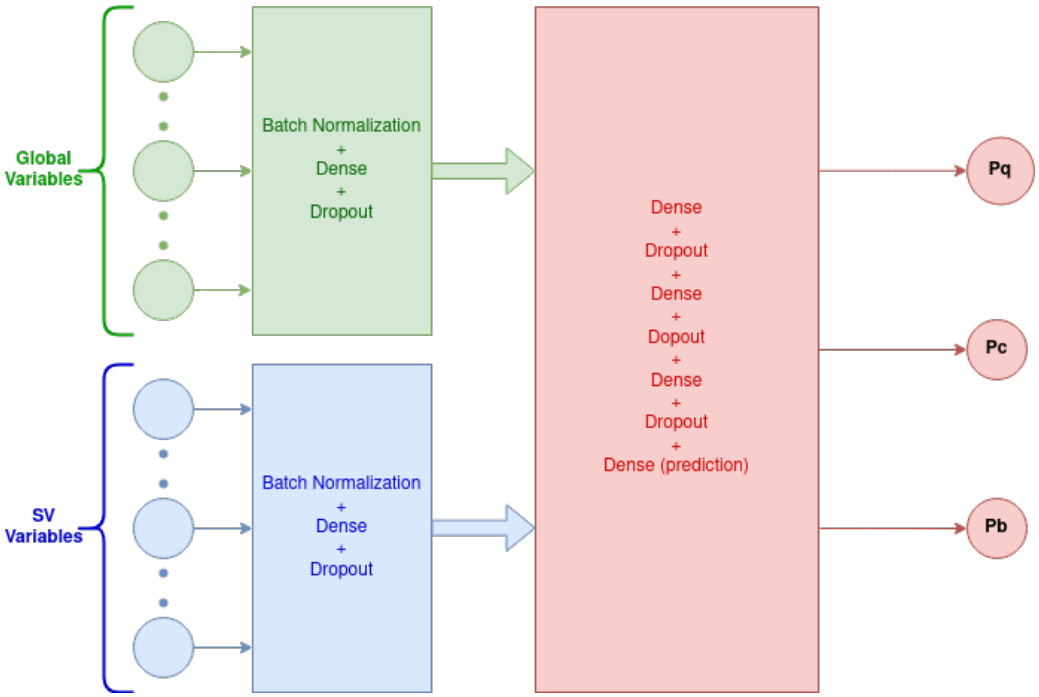


FIG. 76. Scheme of the DNN described in this chapter. The block receiving as input the global jet variables is colored in green, while the one receiving the vertex variables is represented in blue. Their output is the input of the final block, colored in red, which then gives the final predictions on the jet flavour.

$$\begin{aligned} \mathbf{g}_t &= \nabla_\theta E(\boldsymbol{\theta}) \\ \mathbf{m}_t &= \beta_1 \mathbf{m}_{t-1} + (1-\beta_1)\mathbf{g}_t \\ \mathbf{s}_t &= \beta_2 \mathbf{s}_{t-1} + (1-\beta_2)\mathbf{g}_t^2 \end{aligned} \qquad \begin{aligned} \hat{\mathbf{m}}_t &= \frac{\mathbf{m}_t}{1-(\beta_1)^t} \\ \hat{\mathbf{s}}_t &= \frac{\mathbf{s}_t}{1-(\beta_2)^t} \\ \boldsymbol{\theta}_{t+1} &= \boldsymbol{\theta}_t - \eta_t \frac{\hat{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{s}}_t} + \epsilon} \end{aligned}$$

FIG. 77. ADAM optimizer. The symbols are explained in the text. Image taken from [52].

is relatively low, an high batch momentum renormalization [53], chosen at 0.9, is used to smooth statistical fluctuations. The training had a maximum epoch number of 3500, which was not reached because it was set to stop if in a certain amount of epochs, called patience and set to 120, the loss didn't decrease.
We are now going to show the outcome of one training. Some metrics are reported in figure 78. The plot of loss function for the training and the test sample show we don't have an over-training. The confusion matrix is built assigning to each jet the flavour with the highest probability and it tells us that our network is not always successful in distinguishing b vs c or c vs q. The receiver operand characteristic (ROC) curve of the individual flavours tell us from the area under the curve that our network separates light flavour jets from the others, while improvements in the distinctions of the heavy flavours are still needed, as shown by the ROC of b vs c.

The probabilities found by this training are showed in figure 79. Some peaks are present at $p_b \sim 0.2$, $p_c \sim 0.34$, $p_q \sim 0.44$: they may be due to jets which the DNN is not able to classify correctly. This may happen because of jets matched with the wrong flavour or, most importantly, due to the low statistic of the training data for machine learning methods. However, from the tails outside the peak, we can see that the DNN is correctly learning: for example in the p_b plot, the b-jets have many more entries at high p_b than other flavours; in the p_c plot the b-jets have a tail at low p_c while the c-jets are more distributed at values higher than the peak; in the p_q plot, the heavy flavour jets are peaked at 0.
We can conclude that the use of a DNN for jet flavour identification is promising for a future H→ $c\bar{c}$ extraction but more statistics is needed to improve the classification and remove those peaks. A

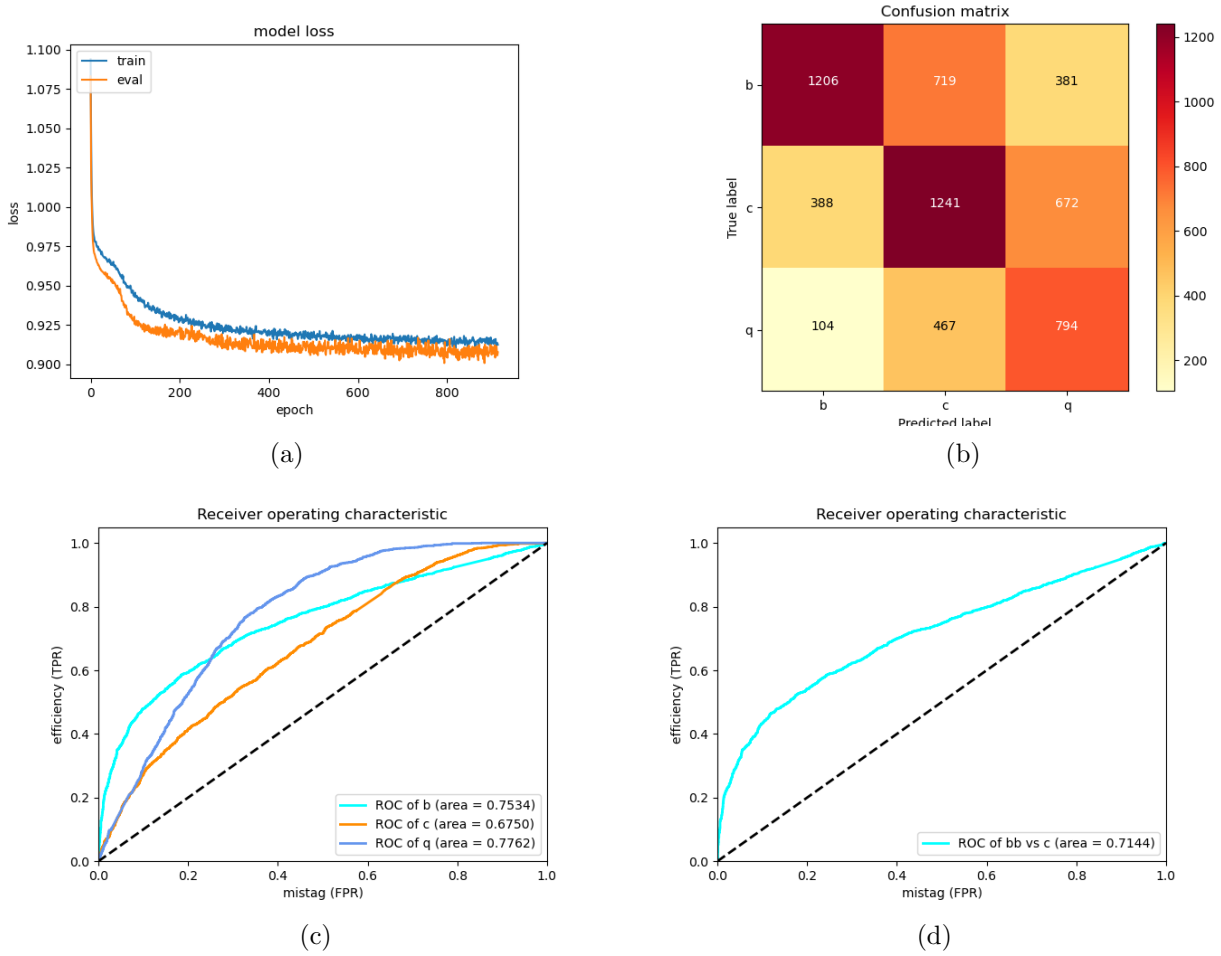
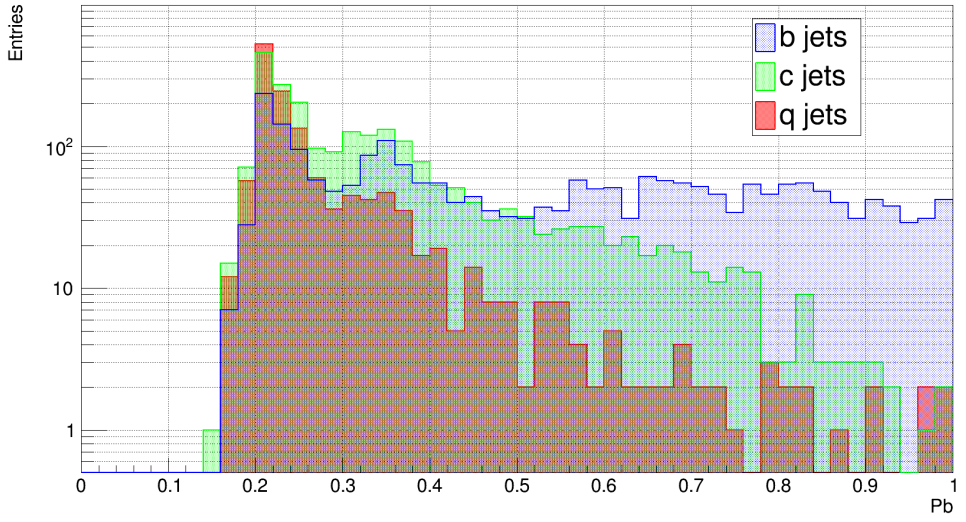
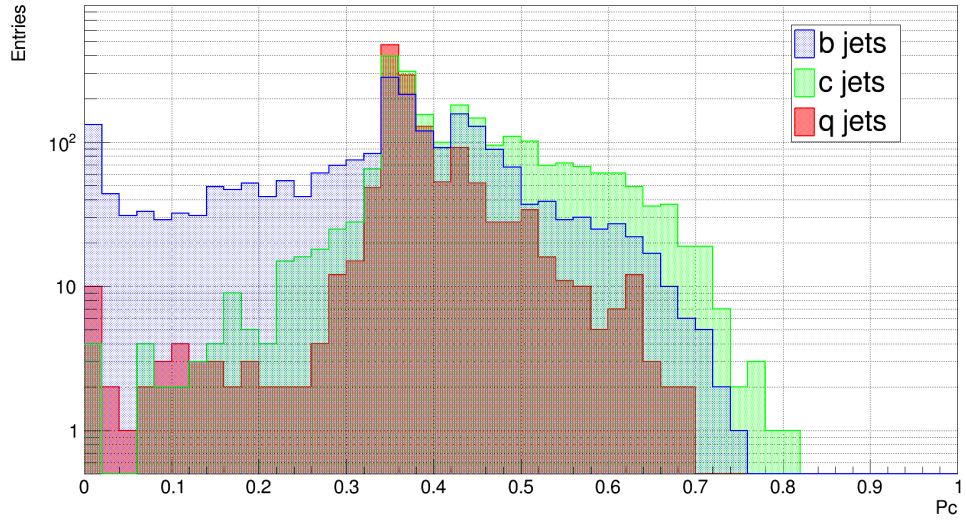


FIG. 78. On the top-left: loss function for the training and test sample as a function of the epochs (a). On the top-right: the confusion matrix, whose rows are the flavour of the jets assigned with the matching and whose columns are the predicted flavour (b). On the bottom-left: the ROC function of each class of jet individually (c). The area under the curve of each is reported in the legend. On the bottom-right: the ROC function of b vs c, whose area under the curve is reported in the legend (d).

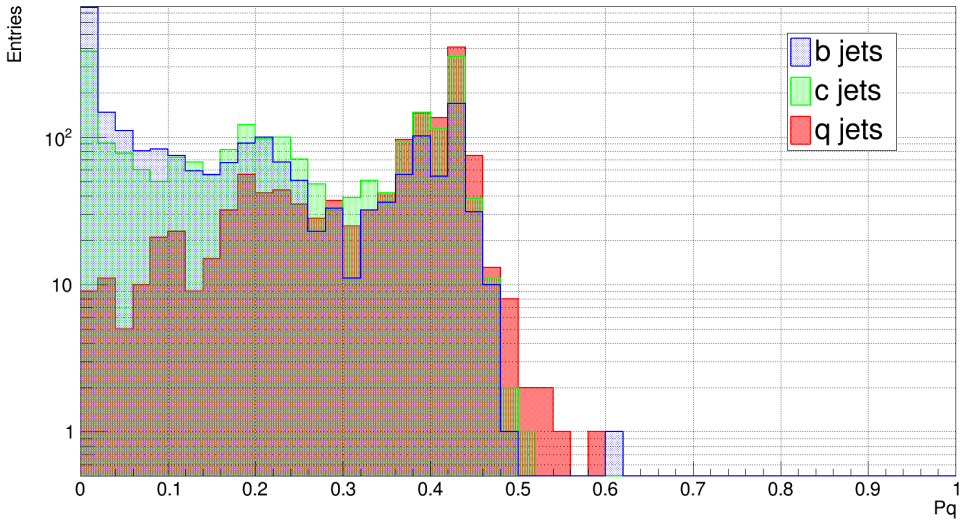
possible strategy to achieve this is the augmentation technique [54]. This method consists in identifying some symmetries our DNN should have and creating copies of the original data that are just shifted by this symmetry: for example we want that our flavour classification does not depend on the direction of the jet; we can then take each jet and create some copies by just shifting it in the (θ, ϕ) plane. In this way, we make sure that the DNN is invariant under the desired symmetry and also increment our training sample.



(a)



(b)



(c)

FIG. 79. DNN probability distribution for the different flavours of jets: at the top, probability of being a b-jet (a); in the middle: probability of being a c-jet (b); at the bottom, probability of being a light jet (c).

8. Conclusions

In this thesis I studied, with a detailed simulation, the preliminary performance of reconstruction of the H$\to b\bar{b}$ at the Muon Collider at $\sqrt{s} = 3$ TeV and estimated the statistical uncertainty on its cross section. I also investigated machine learning techniques for the H$\to c\bar{c}$ identification. In the study I included the effects of the physics background and the beam induced background, showing that the latter can be reduced to manageable levels with a carefully designed machine detector interface and detector. The technologies under study for future accelerators experiments will allow to further improve the detector performance. Algorithms not fully optimized have been used to reconstruct signal and physics background.
By tagging heavy flavour jets with a secondary vertex, I was able to obtain a relative uncertainty on the H$\to b\bar{b}$ cross section of 0.8 % with $\mathscr{L}_{int} = 1$ ab^{-1}, corresponding to 5 years of data taking at the nominal instantaneous luminosity at $\sqrt{s} = 3$ TeV. This is less precise than what CLIC achieved, corresponding to a precision of 0.4 % scaled at our luminosity, however this is just a preliminary result as many improvements are already in progress.

The requirements we put in the vertex processor were optimized to tag both b-jets and c-jets. If the focus is on H$\to b\bar{b}$, cuts could be applied to remove the c-jets to diminish the background; the complete task of jet identification will be performed by using the DNN in the future. The result obtained so far on H$\to b\bar{b}$ represents the starting point. A new algorithm for track and jet reconstruction is under optimization, which will also reduce the computational time. A dedicated analysis strategy is also being studied for the H$\to b\bar{b}$.
I showed that a DNN for jet flavour identification is a promising technique to improve the results of the H$\to b\bar{b}$ and also allows to identify the H$\to c\bar{c}$. However more events are needed for this method to be finalized. One of the possible paths to have more data in a reasonable time is the augmentation, consisting in creating copies of our jets translated in (θ, ϕ) or rotated around their axis, as described in the last chapter.

Further improvements can be achieved in the physics objects reconstruction. Fake jets can be removed looking at their substructure, as they often have a smaller number of particles than real jets and a large fraction of their p_T is coming from a single particle. Machine learning techniques can be used in the calorimeter to distinguish BIB and signal particles deposits from the shower shapes, thus improving the energy resolution of our jets. The full digitization of hits in the tracking system can pave the way to the use of clustering techniques to remove BIB tracks and improve the signal tracking efficiency.

Appendix: Double Layer filter corrections

In this appendix we are going to present the matrices used for the Double Layer filter (DL) corrections. Since b-hadrons and c-hadrons have different lifetimes, the former produce more displaced tracks and are more likely to be removed by the DL. For this reason, the corrections are computed separately for b- and c-jets.

DL Corrections for b-jets:
The correction matrices for the tagging efficiencies (ratio of tagged and matched jets found without DL over the ones with DL) and for the mistag (ratio of tagged and NON-matched jets found without DL over the ones with DL) are shown in figures **80** and **81**. A small θ implies that the jet is almost collinear to the beam direction, while jets emitted in the central region have an high θ.

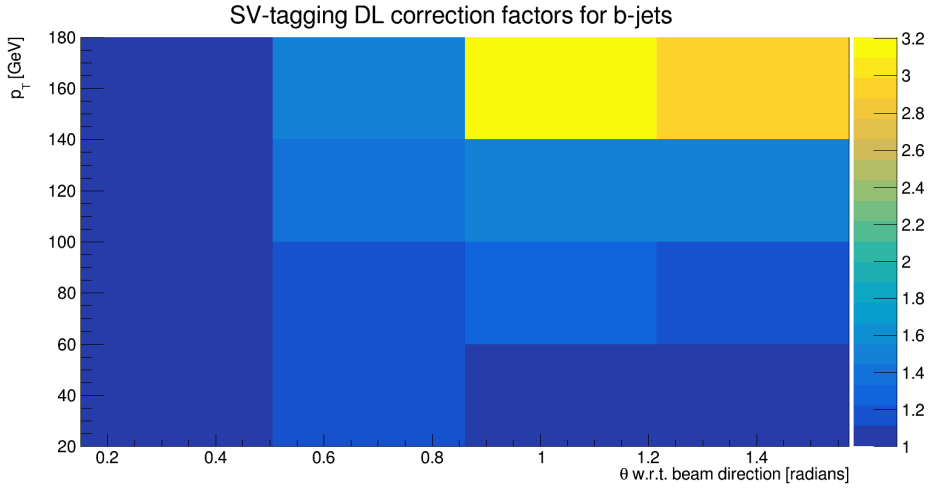


FIG. 80. Correction factors due to the presence of the Double Layer to the tagging efficiency. Each bin is the correction factor which will be applied at jet level, computed as the ratio of tagged and matched jets found without the Double Layer over the one found with the Double layer. The double layer cuts more SVs for jets at high p_T and in the central region.

One can notice that the ratio is bigger (i.e. the double layer cuts more)for high p_T jets and in the central region of the detector. This is to be expected: in high p_T jets the secondary is more boosted and produces tracks more displaced and so more likely to be removed; in the barrel detectors, which are traversed by the particles of more central jets, the parameters of the filter are more strict, due to the closer vicinity of the layers to the IP compared to the endcap detector. The distribution of the SVs position before and after the DL sample can be seen in figure **82**. As a cross check, in figure **83** we can notice that our corrections are not applying any deformation on the reconstructed mass distribution.

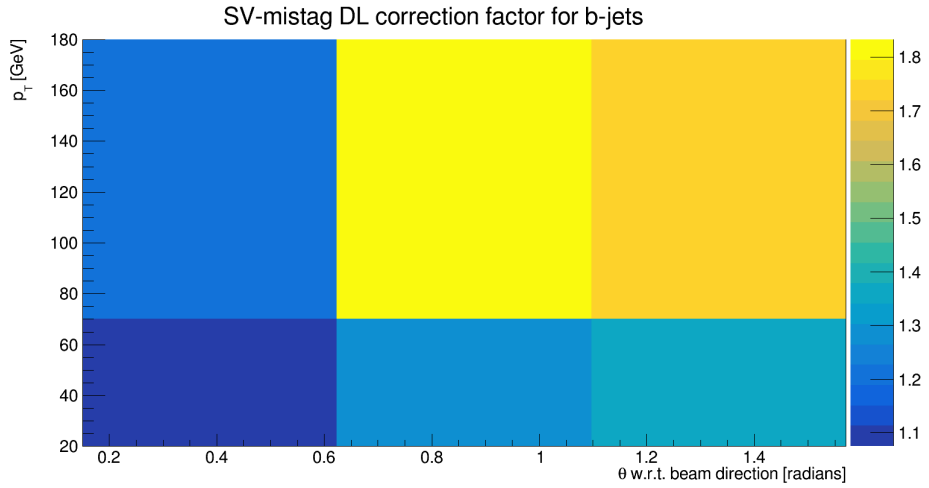


FIG. 81. Correction factors due to the presence of the Double Layer to the tagging efficiency. Each bin is the correction factor which will be applied at jet level, computed as the ratio of tagged and not-matched jets found without the Double Layer over the one found with the Double layer. A wider binning is employed at high p_T since the statistics is lower.

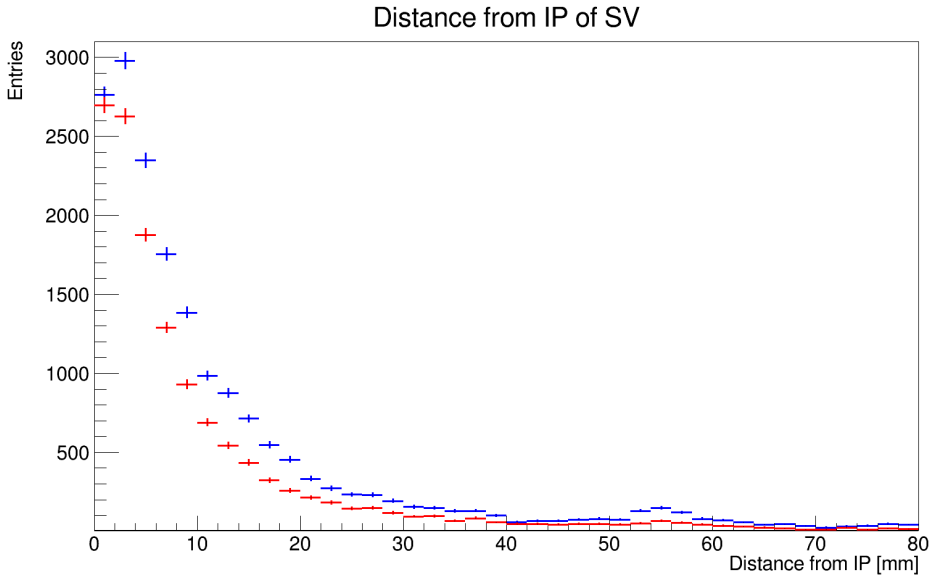


FIG. 82. Distance from the Interaction Point of SVs in the sample with the Double Layer (red) and without (blue). The Double Layer seems to cut mnore at high distances (proportionally to the number of SVs). The small peak at $d \sim 50$ mm is due to the interactions of the particles with the first layer of the barrel of the Vertex detector.

DL Corrections for c-jets:

The correction matrices for the tagging efficiencies and mistag are reported in figures 84 and 85. The effect of the DL filter is much smaller on this sample since D mesons have a smaller lifetime, producing vertices and tracks less displaced. The same cross checks previously showed for the b-jets have been performed for the c-jets.

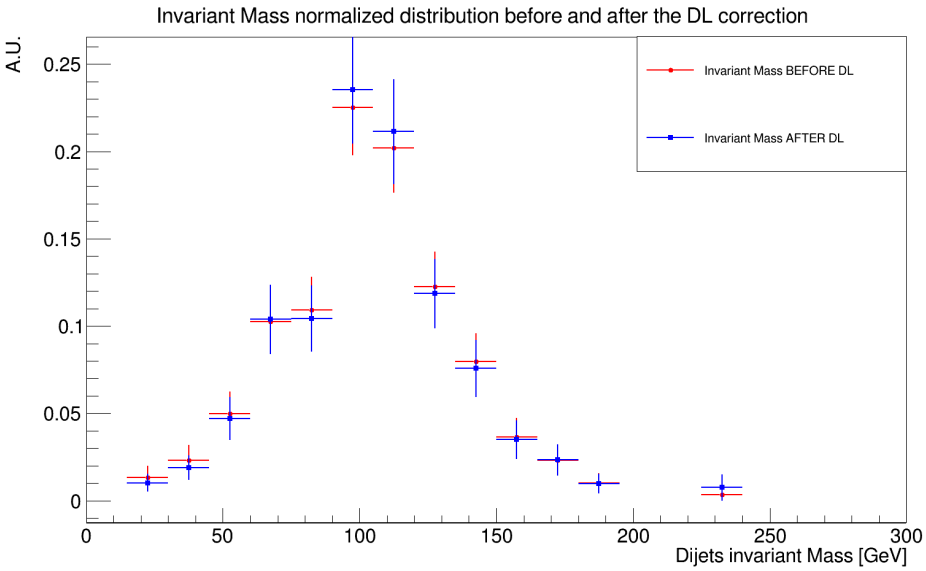


FIG. 83. Normalized H→ $b\bar{b}$ mass distributions before the Double layer filter correction (red) and after it (blue). Our corrections do not misshape the Higgs peak.

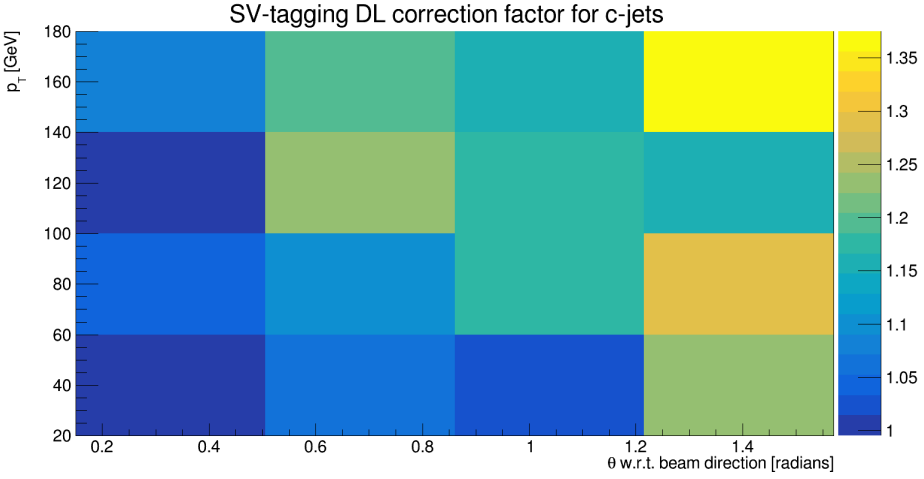


FIG. 84. Correction factors due to the presence of the Double Layer to the tagging efficiency. Each bin is the correction factor which will be applied at jet level, computed as the ratio of tagged and matched jets found without the Double Layer over the one found with the Double layer. C-jets have a smaller tagging efficiency and so the binning has been chosen wider.

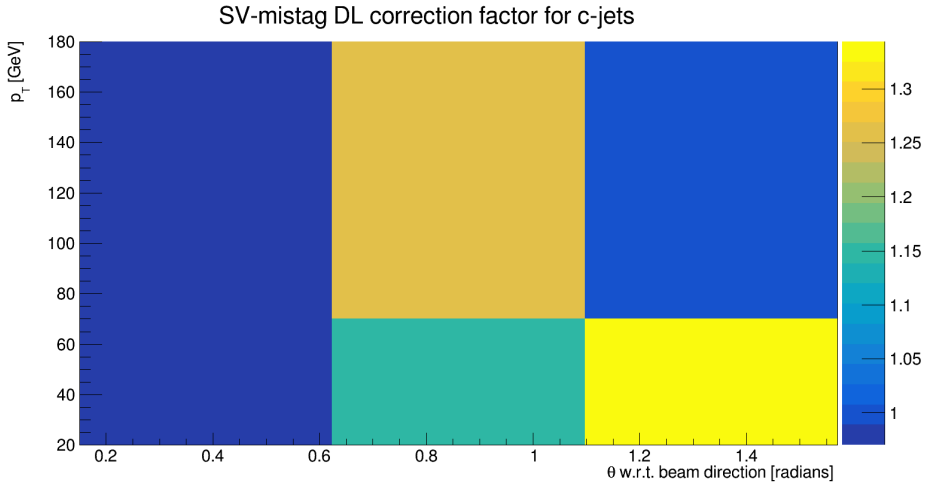


FIG. 85. Correction factors due to the presence of the Double Layer to the tagging efficiency. Each bin is the correction factor which will be applied at jet level, computed as the ratio of tagged and not-matched jets found without the Double Layer over the one found with the Double layer.

Bibliography

[1] S. Chatrchyan, V. Khachatryan, *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012. doi https://doi.org/10.1016/j.physletb.2012.08.021.

[2] G. Aad, T. Abajyan, *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 1–29, 2012. doi: https://doi.org/10.1016/j.physletb.2012.08.020.

[3] M. Y. Hussein, "Higgs Boson Production at the LHC," *arXiv:1703.03952*. https://arxiv.org/pdf/1703.03952.pdf.

[4] H. Abramowicz, A. Abusleme, K. Afanaciev, *et al.*, "Higgs physics at the CLIC electron–positron linear collider," *Eur. Phys. J.*, vol. C 77, p. 475, 2017. doi: https://doi.org/10.1140/epjc/s10052-017-4968-5.

[5] M. Tanabashi and others (Particle Data Group), "11. Status of Higgs Boson Physics," *Phys. Rev.*, vol. D 98, 030001 (2018) and 2019 update.

[6] G. Aad, B. Abbott, and others (The ATLAS Collaboration), "Measurements of the higgs boson production and decay rates and constraints on its couplings from a combined atlas and cms analysis of the lhc pp collision data at $\sqrt{s}$=7 and 8 tev.," *J. High Energ. Phys.*, p. 45, 2016. https://doi.org/10.1007/JHEP08(2016)045.

[7] G. Aad and others (ATLAS collaboration), "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$=7 and 8 TeV," *J. High Energ. Phys*, vol. 45 (2016). doi: https://doi.org/10.1007/JHEP08(2016)045.

[8] A. Abada, M. Abbrescia, AbdusSalam, and others (FCC collaboration), "Higgs Physics at the HL-LHC and HE-LHC," *The European Physical Journal Special Topics*, vol. 228, p. 1109–1382, 2019. doi: https://doi.org/10.1140/epjst/e2019-900088-6.

[9] A. Abada, M. Abbrescia, S. S. AbdusSalam, *et al.*, "FCC Physics Opportunities," *Eur. Phys. J. C 79*, vol. 474, 2019. doi: https://doi.org/10.1140/epjc/s10052-019-6904-3.

[10] C. S. Group, "CEPC Conceptual Design Report: Volume 2 - Physics & Detector," *IHEP-CEPC-DR-2018-02, IHEP-EP-2018-01, IHEP-TH-2018-01*, vol. arXiv:1811.10545v1.

[11] P. Bambade, T. Barklow, and others (ILC collaboration), "The International Linear Collider: A Global Project," *DESY 19-037, FERMILAB-FN-1067-PPD, IFIC/19-10, IRFU-19-10, JLAB-PHY-19-2854, KEK Preprint 2018-92, LAL/RT 19-001, PNNL-SA-142168, SLAC-PUB-17412*, vol. arXiv:1903.01629v3.

[12] Burrows, C. Lasheras, Linssen, Petric, Robson, Schulte, Sicking, and Stapnes, "The Compact Linear Collider -2018 summary report," *CERN-2018-005-M*, vol. Vol 2/2018. doi: https//doi.org/10.23731/CYRM-2018-002.

[13] J. L. A. Fernandez, C. Adolphsen, and others (FCC collaboration), "A Large Hadron Electron Collider at CERN Report on the Physics and Design Concepts for Machine and Detector," *Journal of Physics G: Nuclear and Particle Physics*, vol. 39, p. 075001, jul 2012.

[14] J. de Blas, M. Cepeda, J. D'Hondt, *et al.*, "Higgs Boson studies at future particle colliders," *Journal of High Energy Physics 2020*, vol. 139. doi: https://doi.org/10.1007/JHEP01(2020)139.

[15] H. Al Ali, N. Arkani-Hamed, I. Banta, *et al.*, "The Muon Smasher's Guide," https://arxiv.org/abs/2103.14043.

[16] N. Bartosik, A. Bertolin, L. Buonincontri, *et al.*, "Detector and Physics Performance at a Muon Collider," *Journal of Instrumentation*, vol. 15, pp. P05001–P05001, 2020. doi: https://doi.org/10.1088/1748-0221/15/05/p05001.

[17] D. V. Stefano, "A global view on the Higgs self-coupling at future colliders," p. 34. https://agenda.infn.it/event/16897/attachments/24347/27773/divita-global_view_higgs_selfcoupling-genova.pdf.

[18] M. L. Mangano, G. Ortona, and M. Selvaggi, "Measuring the higgs self-coupling via higgs-pair production at a 100 tev p–p collider," *The European Physical Journal*, vol. 80, p. 1030, 2020. doi: https://doi.org/10.1140/epjc/s10052-020-08595-3.

[19] D. Buttazzo, R. Franceschini, and A. Wulzer, "Two Paths Towards Precision at a Very High Energy Lepton Collider," *CERN-TH-2020-216*.

[20] T. Liu, K.-F. Lyu, J. Ren, and H. X. Zhu, "Probing the quartic Higgs boson self-interaction," *Phys. Rev. D*, vol. 98, p. 093004, Nov 2018. doi: https://link.aps.org/doi/10.1103/PhysRevD.98.093004.

[21] W. Bizoń, U. Haisch, and L. Rottoli, "Constraints on the quartic Higgs self-coupling from double-Higgs production at future hadron colliders," *Journal of High Energy Physics*, vol. 2019. doi: 10.1007/JHEP10(2019)267.

[22] M. Chiesa, F. Maltoni, L. Mantani, B. Mele, and F. Piccinini, "Measuring the quartic Higgs self-coupling at a multi-TeV muon collider," *Journal of High Energy Physics*, vol. 2020. doi: 10.1007/JHEP09(2020)098.

[23] K. R. Long, D. Lucchesi, M. A. Palmer, *et al.*, "Muon colliders to expand frontiers of particle physics," *Nature Physics*, vol. 17, pp. 289–292, 2021. doi:https://doi.org/10.1038/s41567-020-01130-x.

[24] Y. Torum, H. Kirk, A. Bross, *et al.*, "THE US MUON ACCELERATOR PROGRAM," *Proceedings of IPAC'10, Kyoto, Japan.* https://accelconf.web.cern.ch/IPAC10/papers/wepe065.pdf.

[25] "MICE infographic," news.fnal.gov/wp-content/uploads/2020/02/mice-infographic-stfc.jpg.

[26] R. P. Johnson, "Ionization Cooling," *Conf. Proc. C*, vol. 07091010, p. tum2i04, 2007.

[27] F. Collamati, C. Curatolo, D. Lucchesi, A. Mereghetti, N. Mokhov, M. Palmer, and P. Sala, "Advanced assessment of Beam Induced Background at a Muon Collider," https://arxiv.org/pdf/2105.09116v2.pdf.

[28] N. Bartosik, A. Bertolin, M. Casarsa, *et al.*, "Preliminary report of the study of beam-induced background effects at a muon collider," 2019. arXiv:1905.03725v1 [hep-ex].

[29] M. A. Palmer, "An Overview of the US Muon Accelerator Program," in *International Workshop on Beam Cooling and Related Topics*, 6 2013.

[30] N. Mokhov, "MARS15, Version 00," 7 2016. https://www.osti.gov/biblio/1282121.

[31] "Muon Collider INFN confluence site," INFN Muon Collider group, https://confluence.infn.it/display/muoncollider/Muon+Collider+Detector.

[32] A. Cemmi, A. Colangeli, B. D'orsi, I. Di Sarcina, *et al.*, "Radiation study of lead fluoride crystals for the crilin calorimeter," https://arxiv.org/pdf/2107.12307v1.pdf.

[33] H. Hauke and G. Frank, "C event display (ced)," https://ilcsoft.desy.de/CED/current/doc/manual.pdf.

[34] S. Lorenzo, "Report on jet reconstruction at muon collider," https://indico.cern.ch/event/1019298/contributions/4277832/attachments/2209659/3739393/MC_jets_report.pdf, page 16.

[35] "ILCsoftware Github repository," url: https://github.com/iLCSoft/ilcsoftDoc.

[36] M. Frank, F. Gaede, C. Grefe, and P. Mato, "DD4hep: A detector description toolkit for high energy physics experiments," *Journal of Physics: Conference Series*, vol. 513, p. 022010, jun 2014. doi: https://doi.org/10.1088/1742-6596/513/2/022010.

[37] S. Agostinelli, J. Allison, K. Amako, *et al.*, "Geant4—a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003. doi: https://doi.org/10.1016/S0168-9002(03)01368-8.

[38] F. Gaede, "Marlin and LCCD: Software tools for the ILC," *Nucl. Instrum. Meth. A*, vol. 559, pp. 177–180, 2006. doi: 10.1016/j.nima.2005.11.138.

[39] B. Nazar, "Full detector simulation with unprecedented background occupancy at a Muon Collider," p. 5. https://indico.cern.ch/event/948465/contributions/4323702/attachments/2246302/3809598/2021_05_17_bartosik_v1.pdf.

[40] E. Brondolin, E. Leogrande, D. Hynds, F. Gaede, M. Petrič, A. Sailer, and R. Simoniello, "Conformal tracking for all-silicon trackers at future electron–positron colliders," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 956, p. 163304, 2020. doi: https://doi.org/10.1016/j.nima.2019.163304.

[41] F. Keisuke, "Extended kalman filter," (The ACFA-Sim-J Group), https://www-jlc.kek.jp/subg/offl/kaltest/doc/ReferenceManual.pdf.

[42] J. S. Marshall and M. A. Thomson, "The Pandora Particle Flow Algorithm," from Proceedings of CHEF2013 - Calorimetry for the High Energy Frontier, arxiv: https://arxiv.org/abs/1308.4537.

[43] M. Thomson, "Particle flow calorimetry and the PandoraPFA algorithm," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 611, no. 1, pp. 25–40, 2009. doi: https://doi.org/10.1016/j.nima.2009.09.009.

[44] S. Catani, Y. Dokshitzer, M. Seymour, and B. Webber, "Longitudinally-invariant k-clustering algorithms for hadron-hadron collisions," *Nuclear Physics B*, vol. 406, no. 1, pp. 187–224, 1993. doi: https://doi.org/10.1016/0550-3213(93)90166-M.

[45] T. Sjöstrand, S. Mrenna, and P. Skands, "A brief introduction to Pythia8.1," *Computer Physics Communications*, vol. 178, no. 11, pp. 852–867, 2008. doi: https://doi.org/10.1016/j.cpc.2008.01.036.

[46] L. Sestini, A. Gianelle, L. Buonincontri, I. Sarra, and N. Pastore, "Update on jet reconstruction and calorimeter simulation," https://agenda.infn.it/event/26925/contributions/136243/attachments/81218/106303/calo_meeting.pdf, page 6.

[47] T. Suehara and T. Tanabe, "LCFIPlus: A framework for jet analysis in linear collider studies," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 808, pp. 109–116, 2016. doi: https://doi.org/10.1016/j.nima.2015.11.054.

[48] "WHIZARD—simulating multi-particle processes at LHC and ILC," *The European Physical Journal C*, vol. 71, p. 1742, 2011. doi: https://doi.org/10.1140/epjc/s10052-011-1742-y.

[49] D. Schulte, "Muon Collider," https://indico.cern.ch/event/855372/contributions/4456520/attachments/2303915/3919730/neutrino.pdf, page 5.

[50] H. S., "What the new roofit can do for your analysis," *PoS*, vol. ICHEP2020, p. 910, 2021. doi: https://doi.org/10.22323/1.390.0910.

[51] "Identification of beauty and charm quark jets at LHCb," *Journal of Instrumentation*, vol. 10, 2015. The LHCb collaboration, P06013.

[52] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, "A high-bias, low-variance introduction to Machine Learning for physicists," *Physics Reports*, vol. 810, pp. 1–124, 2019. doi: https://doi.org/10.1016/j.physrep.2019.03.001.

[53] H. Yong, J. Huang, D. Meng, X. Hua, and L. Zhang, "Momentum Batch Normalization for Deep Learning with Small Batch Size," in *Computer Vision – ECCV 2020*, pp. 224–240, Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-58610-2_14.

[54] B. M. Dillon, G. Kasieczka, H. Olischlager, *et al.*, "Symmetries, Safety, and Self-Supervision," https://arxiv.org/abs/2108.04253v1.