



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna | Austria

Dissertation

Track reconstruction in the CMS experiment for the High Luminosity LHC

ausgeführt zum Zwecke der Erlangung des akademischen Grades einer Doktorin der technischen
Wissenschaften unter der Leitung von

UNIV.DOZ. DR. RUDOLF FRÜHWIRTH

am Institut für Hochenergiephysik (HEPHY)
der Österreichischen Akademie der Wissenschaften (ÖAW)

eingereicht an der
TECHNISCHEN UNIVERSITÄT WIEN
FAKULTÄT FÜR PHYSIK
von

DOTT.SSA MAG. ERICA BRONDOLIN
Matrikelnummer: 01428274

CERN-THESIS-2017-328
17/01/2018



Wien, am 14. Dezember 2017

*A tutti i quindici membri
della mia famiglia.*

Abstract

Tracking is one of the crucial parts in the event reconstruction because of its importance in the estimation of particle momenta, their identification, and in the estimation of decay vertices. This task is very challenging at the **LHC**, given the hundreds or even thousands of particles generated in each bunch crossing.

Track reconstruction in **CMS** was developed following an iterative philosophy. It uses an adaptation of the combinatorial **Kalman Filter** (**KF**) algorithm to allow pattern recognition and track fitting to occur in the same framework. For $t\bar{t}$ events under typical Run-1 pileup conditions, the average track reconstruction efficiency for charged particles with transverse momenta of $p_T > 0.9$ GeV is 94% for $|\eta| < 0.9$ and 85% for $0.9 < |\eta| < 2.5$.

During **Long Shutdown 1**, some developments were made in different aspects of tracking. In particular, I implemented the **Deterministic Annealing Filter** (**DAF**) algorithm to protect track reconstruction against wrong hit assignments in noisy environments or in high track density environments. The **DAF** algorithm showed a reduction of mis-identified tracks in high- p_T jets, even more evident inside their core, without any loss in terms of efficiency.

During Run 2, which started in 2015 and will continue until the end of 2018, the machine operates at a centre-of-mass energy of 13 TeV with a bunch crossing separation of 25 ns and targeting an instantaneous luminosity of up to $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. During data taking, a possible way to monitor the consistency between data and **MC** is to measure the muon tracking efficiency using the tag and probe. I found the consistency to be within a few per-mill using $Z \rightarrow \mu^+ \mu^-$ events, and within a few per-cent for the $J/\psi \rightarrow \mu^+ \mu^-$ sample. Moreover, I performed a study in the context of the $H \rightarrow \tau^+ \tau^-$ analysis simulating the Run 2 conditions, with the aim of finding a good signal-versus-background classifier. For this purpose I implemented and optimized a multi-layer perceptron neural network using the NeuroBayes package. I compared its performance to other **Multivariate Analysis** (**MVA**) approaches as well as to a cut-based approach, which is the current technique. The network performs better than the cut-based approach, but it does not deliver the best results among the different **MVAs** analyzed. This is very likely due to the fact that the NeuroBayes neural network requires a larger dataset in order to get a significant improvement in performance.

Between 2024 to 2026, the **Long Shutdown 3** is scheduled where the main preparation of the accelerator and the experiments will take place for the High Luminosity phase of the LHC (**HL-LHC**). The **HL-LHC** will provide an unprecedented instantaneous luminosity of $5 - 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. In order to face this challenge, the **CMS** experiment has decided to install a new silicon-based tracker. In the work described here I laid the foundation for track reconstruction in **CMS** at the **HL-LHC**. In a first step, I adapted the existing tracking algorithms to the new tracker. I evaluated the performance of tracking and vertexing for different event types and detector geometries and I found them to be similar to the current ones despite the harsher pileup conditions at **HL-LHC**. In a second step, I introduced a new type of hits, so-called vector hits, in the outer part of the new tracker, with the aim of answering the question whether they bring tangible benefits in the pattern recognition step at the **HL-LHC**. To this end, I extended the **KF** to make use of the full information contained in the vector hits, and I added a new iteration in the Phase-2 tracking, in which seeds are built using only the vector hits. The results show a reduction of about an order of magnitude of the number of mis-identified tracks for high pileup environments and a significant improvement in the reconstruction of tracks coming from displaced vertices, such as the decay products of the K_S^0 .

Kurzfassung

Die Rekonstruktion von geladenen Spuren ist ein wichtiger Teil der Ereignisanalyse, da sie die Voraussetzung zur Impulsschätzung und zum Auffinden von Kollisions- und Zerfallsvertices darstellt. Am **LHC** ist dies eine schwierige Aufgabe, vor allem wegen der Vielzahl von Teilchen, viele hundert bis einige tausend, die bei einer Kollision von zwei Protonpaketen im **LHC** erzeugt werden.

Während der ersten LHC-Periode zwischen 2010 und 2012 wurde eine schrittweise Methode der Spurrekonstruktion entwickelt und implementiert. Sie verwendet eine Version des kombinatorischen Kalman-Filters zum gleichzeitigen Auffinden und Schätzen der Trajektorien von geladenen Teilchen. Mit $t\bar{t}$ -Ereignissen und typischen Pileup-Verhältnissen betrug die Rekonstruktionseffizienz in dieser Periode für Teilchen mit Transversalimpulsen über 0.9 GeV 94% im Bereich $|\eta| < 0.9$, und 85% im Bereich $0.9 < |\eta| < 2.5$.

Während der Schließperiode **Long Shutdown 1** in den Jahren 2013 und 2014 wurden verschiedene Teile der Spurrekonstruktion weiterentwickelt. Ich implementierte den **Deterministic Annealing Filter (DAF)**-Algorithmus, um die Spurrekonstruktion unempfindlich gegen falsche Zuordnung von Messungen, speziell in Bereichen hoher Teilchendichte, zu machen. Das Resultat war eine Reduktion von falschen Spuren in Jets mit hohem p_T , und hier wiederum speziell im Zentrum der Jets, ohne Verlust an Effizienz.

In Periode 2, die Anfang 2015 begann und bis Ende 2018 dauern wird, läuft der Beschleuniger mit einer Schwerpunktsenergie von 13 TeV und einer Kollisionsfrequenz von 40 MHz. Die Luminosität soll bis zu $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ erreichen. Während der Datennahme kann die Übereinstimmung zwischen echten und Monte-Carlo-Daten laufend durch die Messung der Myoneffizienz kontrolliert werden. Dazu wird die sogenannte “tag-and-probe”-Methode verwendet. Mit Ereignissen vom Typ $Z \rightarrow \mu^+ \mu^-$ ergab sich aus meinen Rechnungen eine Übereinstimmung von einigen Promille, während sie mit Ereignissen vom Typ $J/\psi \rightarrow \mu^+ \mu^-$ einige Prozent betrug.

Ich leistete auch einen Beitrag zur Analyse von Zerfällen des Higgsbosons in zwei τ -Leptonen. Das Ziel war die Entwicklung eines Klassifizierungsalgorithmus, der das Signal möglichst gut vom Untergrund abtrennt. Zu diesem Zweck implementierte und optimierte ich ein neuronales Netzwerk mit einer verborgenen Lage, unter Verwendung des Programmpaketes NeuroBayes. Der Vergleich mit anderen Methoden der Klassifikation ergab, dass das neuronale Netz eine deutliche Verbesserung im Vergleich zur herkömmlichen Methode brachte. Es war jedoch nicht der beste multivariate Klassifikator, vermutlich wegen unzureichender Menge an Trainingsdaten.

Zwischen 2024 und 2026, ist eine weitere Schließperiode **Long Shutdown 3** geplant. Diese dient dem Ausbau des Beschleunigers und der Experimente für die Hochluminositätsphase des LHC (**HL-LHC**), in der die Luminosität auf den beispiellosen Wert von $5 - 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ ansteigen wird. Um diese Herausforderung zu meistern, hat sich das **CMS**-Experiment entschlossen, einen neuen Spurdetektor zu bauen. In dieser Arbeit wird der Grundstein der Spurrekonstruktion in diesem neuen Gerät gelegt.

Im ersten Schritt adaptierte ich die existierenden Algorithmen an den neuen Detektor. Ich studierte die Effizienz und allgemeine Leistungsfähigkeit von Spur- und Vertexrekonstruktion für verschiedene Detektorgeometrien und Ereignistypen. Sie stellten sich als mindestens ebenso gut als die aktuellen Werte heraus, trotz wesentlich höherem Pileup. Im zweiten Schritt definierte ich neuartige Typen von Messungen, sogenannte “vector hits”, und untersuchte ob durch diese die Mustererkennung am **HL-LHC** verbessert werden kann. Dazu erweiterte ich den **KF**, um die

volle Information der vector hits auszunützen. Des weiteren entwickelte ich eine neue Stufe der Spurrekonstruktion, in der Spuren lediglich unter Benützung von vector hits gesucht und gefunden werden. Als Resultat zeigte sich, dass bei gleichbleibender Effizienz die Anzahl von falschen Spuren um fast eine Größenordnung gesenkt werden kann. Auch zeigte sich eine signifikante Verbesserung der Rekonstruktion von Spuren, die, wie zum Beispiel die Zerfallsprodukte von K_S^0 -Mesonen, außerhalb der Wechselwirkungszone erzeugt werden.

Contents

Abstract	I
Kurzfassung	III
Contents	V
1 Theoretical overview	1
1.1 The Standard Model	1
1.1.1 Fundamental particles and their interactions	2
1.1.2 Electroweak theory	4
1.1.3 Spontaneous symmetry breaking	6
1.2 Higgs boson discovery at the LHC	8
1.2.1 Higgs production mechanisms at LHC	8
1.2.2 Higgs decay channels and discovery at LHC	9
1.2.3 Evidence for an SM Higgs boson decaying into a pair of τ leptons	11
1.3 Beyond the Standard Model	13
2 The CMS detector at the LHC	15
2.1 The Large Hadron Collider	15
2.1.1 The injection chain	17
2.1.2 Operational periods of the LHC	17
2.1.3 LHC operations	19
2.1.4 Experiments at the LHC	20
2.2 The CMS detector	23
2.2.1 The tracking system	25
2.2.2 The electromagnetic calorimeter	27
2.2.3 The hadron calorimeter	30
2.2.4 The superconducting solenoid	31
2.2.5 The muon system	31
2.2.6 Trigger and data acquisition	34
2.2.7 Reconstruction software	34
2.3 Physics object reconstruction	36
2.3.1 Muons	37
2.3.2 Electrons	37
2.3.3 Particle Flow reconstruction	39
2.3.4 Jets and missing transverse energy	40
2.3.5 Hadronic τ lepton decays	41
3 Track reconstruction in CMS	45
3.1 Theoretical and experimental context	45
3.1.1 The track model	46
3.1.2 The track fit	48
3.1.3 Specific choices in the CMS experiment	55
3.2 Track reconstruction during Run 1	56

3.2.1	Local reconstruction	57
3.2.2	Iterative tracking	58
3.2.3	Performance	61
3.3	Developments towards Run 2	69
3.3.1	Physics- and timing-oriented developments	69
3.3.2	Deterministic Annealing Filter results	71
3.4	Run-2 results using data and simulation	72
3.4.1	Validation of simulation using 2016 data	73
3.4.2	Simulation results for the Pixel Tracker upgrade	81
4	Multivariate Analysis of $H \rightarrow \tau^+\tau^-$	87
4.1	Signal-versus-background classifiers	87
4.1.1	Cut-based approach	88
4.1.2	Multivariate approach	88
4.1.3	Quantification of performance	94
4.2	Performance of NeuroBayes neural network $H \rightarrow \tau^+\tau^-$	95
4.2.1	13 TeV simulation samples	95
4.2.2	Input variables	95
4.2.3	Optimization of the network	99
4.2.4	Result and comparison with other classification approaches	103
5	The CMS detector at the High Luminosity LHC	105
5.1	The High Luminosity LHC	105
5.1.1	Introduction	105
5.1.2	Theoretical motivations for the HL-LHC	106
5.1.3	The machine	107
5.1.4	Experiments at the HL-LHC	109
5.2	The upgrade of the CMS experiment	111
5.2.1	Overview of the upgrade plan	111
5.2.2	The CMS detector at the HL-LHC	113
5.3	The Phase-2 upgrade of the CMS tracker	115
5.3.1	Requirements at the HL-LHC	116
5.3.2	The Inner Tracker	118
5.3.3	The Outer Tracker	119
5.3.4	Services	122
5.3.5	The Level-1 Track Trigger	123
6	Track reconstruction in Phase-2 CMS	127
6.1	Track reconstruction during Phase 2	127
6.1.1	Local reconstruction	128
6.1.2	Iterative tracking	130
6.1.3	Performance	131
6.1.4	Possible developments	137
6.2	Vector Hits	139
6.2.1	Local reconstruction	139
6.2.2	Global reconstruction	150
6.2.3	Seeding in the Outer Tracker	151
6.3	Results	151
7	Summary and outlook	157
	Bibliography	159
	Acronyms	170
	Thanks	171

CHAPTER 1

Theoretical overview

*To Laura Bassi, first woman
university professor. (1731)*

In this chapter a theoretical overview of elementary particle physics is presented. [Section 1.1](#) will give an introduction to the [Standard Model](#), which summarizes the current knowledge of elementary particles and their interactions. In the [Standard Model](#), the symmetry breaking mechanism of the [Electroweak](#) theory is introduced to explain how the fundamental particles acquire mass. This mechanism is originating spontaneously via the Brout-Englert-Higgs theory, which leads to the existence of a new particle, the Higgs boson. Its properties, such as production and decay mode at the [Large Hadron Collider](#), are presented in the [Section 1.2](#) as well as its discovery in 2012. Particular emphasis is given to the challenges and latest results of the τ -pair decay channel. The last part of this chapter, [Section 1.3](#), addresses briefly the main open questions in elementary particle physics.

In this thesis, natural units will be used if not specified otherwise, i.e. $\hbar = c = 1$.

A glossary explaining important acronyms is provided at the end of the thesis.

1.1 The Standard Model

The greek philosopher Democritus and his teacher Leucippus can be considered the fathers of the idea that all matter is made of not further divisible particles, called *atoms*. Today, this idea is explored by elementary particle physics. The actual beginning of elementary particle physics can be placed in 1897 with the discovery by J. J. Thomson [\[1\]](#) that cathode rays were actually made of negatively charged particles, what we call today *electrons* and that these are an essential constituent of atoms. He imagined that the electrons were suspended in a heavy, positively charged paste. This model was rejected by E. Rutherford's famous scattering experiment in 1909, which investigated the atomic structure in more detail by firing a beam of α -particles (ionized helium atoms) against a sheet of gold foil [\[2\]](#). Rutherford showed with his experiment that the entire positive charge of the atom is located at the center, occupying only a small fraction of the entire volume. He named the nucleus of the lightest atom, hydrogen, *proton*. The discovery of the *neutron* by J. Chadwick [\[3\]](#) in 1932 concludes the so-called *classical period* of elementary particle physics, where the picture of the atom and therefore all matter seemed to be complete: The atom is neutral, composed of a nucleus with protons and neutrons and surrounded by electrons.

However, in the so-called *middle period* (1934–1960) many particles of different masses, charges and spins were discovered, which did not fit in this simplified model of the atom [\[4\]](#). Thanks to a remarkable interplay between theoretical predictions and experimental evidence, the fundamental theory which describes elementary particles and their interactions, the so-called [Standard Model](#)

Table 1.1: Mediator bosons carrying the four type of interactions with their boson mass and the relative magnitude.
 *Not observed yet.

Interaction	Gauge Boson	Mass (GeV)	Charge	Relative strength
electromagnetic	γ	0	electrical	10^{-2}
strong	gluons	0	colour (r,g,b)	1
weak	$W^{\pm}$	80.38 [5]	weak isospin	10^{-7}
	Z^0	91.18 [5]		
gravitation	graviton*	0		10^{-39}

(SM) of particle physics, was formulated in the 1970s and since then it has been confirmed with great precision by a vast amount of experimental data.

1.1.1 Fundamental particles and their interactions

In the SM, every elementary particle is identified by its quantum numbers, a representation of the particle interaction properties. According to this model, fundamental particles are grouped into two categories: *fermions* and *bosons*. The former have half-integral spin and obey Fermi-Dirac statistics, while the latter have integral spin and obey Bose-Einstein statistics. In the SM, all matter is made of two kinds of half-integer spin fermions: *leptons* and *quarks*. They interact mainly via three fundamental forces: the electromagnetic, the weak and the strong interaction. The first two can be unified in the **Electroweak (EWK)** force. At the typical mass scales, the strength of the fourth fundamental force, gravitation, is negligible compared to that of the other interactions.

The interactions between quarks and leptons are carried by the exchange of a gauge boson:

- The *photon* (γ) mediates the electromagnetic interaction. It is a massless particle with no electric charge. **Quantum Electrodynamics (QED)** describes this force. It acts only between electrically charged particles and is responsible for almost all microscopic phenomena outside the nuclei, such as the bound states of electrons with nuclei and the intermolecular forces in liquids and solids.
- There are eight electrically neutral and massless *gluons* (g) that mediate the strong interaction. The charge is called colour and can take three values, commonly labeled *red*, *green* and *blue*, and the respective anticolours. Gluons themselves carry colour charge, which allows them to interact with each other. This short range force is responsible for the phenomenon known as *quark confinement*, which for example keeps quarks bound inside the proton. **Quantum Chromodynamics (QCD)** describes this interaction.
- Three massive bosons (W^- , W^+ and Z^0) carry the weak interaction, which is responsible, for example, for radioactive decays. More details about this interaction and the associated charge are given in **Section 1.1.2**.
- The *graviton* is the hypothetical boson which carries the gravitational interaction. The graviton has not been experimentally observed yet, but it is expected to be massless and a spin-2 boson. The effect of this force becomes relevant only in the presence of large masses, such as planets or stars.

The list of interactions with their corresponding bosons is given in **Table 1.1**.

Six leptons and six quarks are the fundamental fermions of the SM. They are grouped in three different *families*, also called *generations*, with masses increasing from one family to the other. Each fermion has an anti-particle which has opposite quantum numbers. In the leptonic case, the family consists of one (negatively) charged lepton and a neutrino that only carries weak charge. The electron (e) is the lightest charged lepton with a mass of approximately 511 keV, while muon (μ) and tau (τ) are heavier with masses of $\simeq 105$ MeV and $\simeq 1.7$ GeV, respectively [5]. In the SM, neutrinos (ν) are treated as massless, although the observation of neutrino oscillations indicates that they must have non-zero mass [6]. The current value of the experimental upper bound is

$m_{\nu_e} < 2 \text{ eV}$ [5]. The six quarks are labeled up (u), down (d), charm (c), strange (s), top (t) and bottom (b). Each family contains an up- and a down-type quark. Historically, these names refer to the third isospin component, which takes values $+\frac{1}{2}$ or $-\frac{1}{2}$, respectively. Unlike leptons, quarks have a not integer electric charge. The up-type quarks have an electric charge of $\frac{2}{3}|e|$, while the down-type quarks of $-\frac{1}{3}|e|$, where $|e|$ is the absolute value of the electron charge. Similar to the leptons, the quark masses increase throughout the families from the $\simeq 2.2 \text{ MeV}$ mass of the up-quark, up to the $\simeq 173 \text{ GeV}$ mass of the top-quark [5]. All stable matter consists of first generation fermions: atoms consist of electrons, protons and neutrons, the latter two being composed of up- and down-quarks. Fermions belonging to the other two generations and their compounds are unstable and always decay into lighter particles.

All known fermions interact among themselves through the **EWK** force described in **Section 1.1.2**. Quarks carry also the color charge, and therefore they take part in the strong interaction. They do not exist as free particles in nature, but are confined in bound states, called *hadrons*. Experimentally, it has been observed that all hadrons appear to be colorless, an example are *mesons*, which are bound states of $q\bar{q}$ pairs, or *baryons*, which are $qqq/\bar{q}\bar{q}\bar{q}$ aggregates. The list of fermions and their masses is reported in **Table 1.2**.

Table 1.2: Table of fundamental fermions with respective masses and electric charge (Q) [5].

	1 st gen.		2 nd gen.		3 rd gen.		Q
Leptons	ν_e	$< 2.2 \text{ eV}$	ν_μ	$< 0.17 \text{ MeV}$	ν_τ	$< 15.5 \text{ MeV}$	0
	e	511 keV	μ	105.7 MeV	τ	1.777 GeV	-1
Quarks	u	$\simeq 2.3 \text{ MeV}$	c	$\simeq 1.275 \text{ GeV}$	t	$\simeq 173.07 \text{ GeV}$	2/3
	d	$\simeq 4.8 \text{ MeV}$	s	$\simeq 95 \text{ MeV}$	b	$\simeq 4.18 \text{ GeV}$	-1/3

In the **SM**, the fundamental particles are represented as *quantum fields*. The Lagrange formalism is used to describe their propagation and their interactions [4]. In particular, interactions in the **SM** are described by *gauge fields*. They are created from the postulation of local gauge invariance under certain local gauge transformations which define a symmetry. These transformations generate thus the corresponding symmetry group. The gauge symmetry invariance requirement brings the conservation of charges, via Noether's theorem, and it allows for the introduction of new fields and interactions in the theory. For each generator of the symmetry group, a *vector-boson field* is introduced in order to re-establish the local gauge invariance. This can be illustrated using the following example.

A free fermion of mass m is represented by a 4-component Dirac spinor ψ . Its Lagrangian, with the associated equation of motion, holds

$$L_{Dirac}(x) = i\bar{\psi}(x)\gamma^\mu\partial_\mu\psi(x) - m\bar{\psi}(x)\psi(x) \quad \rightarrow \quad i(\gamma^\mu\partial_\mu - m)\psi(x) = 0 \quad (1.1)$$

where γ_μ are the Dirac matrices and $\bar{\psi} = \psi\gamma_0$ is the adjoint spinor. If a local phase transformation of the form $\psi'(x) = \exp(ief(x))\psi(x)$ is performed, it leads to

$$L_{Dirac} \rightarrow L_{Dirac}' = L_{Dirac} + e\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu(x) \quad (1.2)$$

To preserve the Dirac's equation invariance, a new field $A_\mu(x)$ can be introduced and the usual derivative replaced with the so-called *covariant derivative*

$$D_\mu\psi(x) = [\partial_\mu + ieA_\mu(x)]\psi(x) \quad (1.3)$$

where e is the charge of the fermion. Under gauge transformations, the derivative and the field become

$$\begin{cases} D_\mu\psi(x) & \rightarrow D'_\mu\psi(x) = \exp(-ief(x))D_\mu\psi(x) \\ A_\mu(x) & \rightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu f(x) \end{cases} \quad (1.4)$$

With the modified derivative the Lagrangian 1.2 becomes:

$$L_{Dirac}(x) = i\bar{\psi}(x)\gamma_\mu\partial_\mu\psi(x) - m\bar{\psi}(x)\psi(x) + e\bar{\psi}(x)\gamma_\mu A_\mu\psi(x) \quad (1.5)$$

This Lagrangian can be identified with part of the Lagrangian density of the **QED**: the first two terms are the kinetic terms for the fermion field and its mass term, the third term is the interaction between the new field and the fermion field via the charge e . In the complete Lagrangian a kinetic term of the new field with the form $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ must be also included. The invariance under the transformation in Equation 1.1 corresponds to a symmetry under a U(1) gauge transformation and the new field, A_μ , can be identified with the photon field. This boson must be massless to conserve the symmetry.

For further convenience, the Dirac spinors can be rewritten using the Weyl spinor representation which divides the Dirac spinor ψ into its left-handed and right-handed components, ψ_L and ψ_R , respectively

$$\psi = \psi_L + \psi_R = \begin{pmatrix} \psi'_L \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi'_R \end{pmatrix} \quad (1.6)$$

where

$$\begin{cases} \psi_L = P_L\psi = \frac{1}{2}(1 - \gamma_5)\psi \\ \psi_R = P_R\psi = \frac{1}{2}(1 + \gamma_5)\psi \end{cases} \quad (1.7)$$

$P_{L/R}$ are the *chirality projection* operators. The left or right-handedness of spinors is called *chirality*.

Similarly to **QED**, the formalism can be generalized to the other interactions. The **SM** is described by a $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry. The Lagrangian density of the **QCD** can be obtained by requiring symmetry under local $SU(3)_C$ gauge transformations. Each quark appears in three different color states, C , thus belonging to a $SU(3)_C$ triplet. The $SU(3)$ has eight generators, which are associated to the eight previously mentioned gluons. The Lagrangian density of the **EWK** is obtained using the local $SU(2)_L \otimes U(1)_Y$ gauge transformations. It has four generators: The long-range electromagnetic interaction is mediated by the massless photon, while the short-range weak force carriers are the massive W^+ , W^- and Z^0 bosons. Among the three observed vector bosons of the **EWK** interaction, only the photon is massless, while W and Z bosons are massive, indicating that the gauge symmetry is broken. In order to give them mass, a spontaneous symmetry breaking mechanism is introduced. It predicts the existence of another fundamental boson which must have spin 0. This boson is commonly referred to as the Higgs boson, named after Peter Higgs, who was one of the first to predict its existence [7, 8, 9].

1.1.2 Electroweak theory

One of the most important steps in the history of the **SM** was the unification of the electromagnetic and the weak interaction in the **EWK** of Glashow, Salam and Weinberg in 1979 [10]. For a long time, only weak charged current interactions, in which the charge of the leptons or quarks changes sign, were known. For this reason, at least two exchange particles were predicted, with charge $+1$ and -1 , the already-mentioned W^+ and W^- . If these particles have spin 1, the interaction can be described by a combination of a vector (V) and an axial-vector (A) operator, and the strength of each contribution by the coefficients c_V and c_A . If the interaction conserves parity, i.e. couples equally to left- and right-handed particles (defined in Equation 1.6), then it can only be either purely vectorial ($c_A = 0$) or purely axial-vectorial ($c_V = 0$). For example, in the case of electromagnetism, it is a pure vectorial interaction. If both coefficients have the same absolute value, parity is *maximally violated* and it is referred to as the $V - A$ -theory. Experimental evidence proves that the charged gauge bosons couple only to the left-handed components ψ_L of the fermion fields, and that the coupling strength is the same for all fermions. The $SU(2)_L$ doublets are

$$f_L = \left\{ \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \right\}. \quad (1.8)$$

This is different from neutral currents, which do not change the electric charge of the participating fermions and couple to right-handed components too. These are $SU(2)_R$ singlets

$$f_R = \{e_R, \mu_R, \tau_R, u_R, d_R, c_R, s_R, t_R, b_R\} \quad (1.9)$$

The first observation of the neutral weak current was in 1973 with the Gargamelle bubble chamber at CERN [11]. In January 1983 the W boson was observed by the UA1 experiment and five months later also the production of a neutral vector boson, Z^0 , was experimentally shown [12, 13]. A new quantum number, the *weak isospin* I , is associated to each doublet. Neutrinos and up-like quarks possess a third positive weak isospin component $I_3^W = +1/2$, whereas the rest of leptons and quarks exhibit a negative I_3^W . Right-handed fermions are weak isospin singlets, $I = I_3 = 0$, since they do not participate in weak charged current interactions.

$W^\pm$ bosons must have $I = 1$ and $I_3 = \pm 1$ to allow transitions between left-handed charged leptons and neutrinos or up- and down-type quarks. To explain the transitions between different generations, the electroweak eigenstates of down-type quarks are interpreted not as the actual quark mass eigenstates (d, s, b) but mixtures of those, labelled d', s' and b', according to the unitary Cabibbo-Kobayashi-Maskawa matrix [14, 15]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.10)$$

The diagonal elements describe the transitions within the same generation. Given that these elements are close to unity, the transitions between different families are therefore strongly suppressed.

The **EWK** foresees also another boson which mediates transitions that do not change the fermion flavour, just like the neutral currents, and therefore has $I_3 = 0$ and does not change the 3-component. Moreover, this fourth field has to couple to fermions in the same way as $W^\pm$. Experimentally, two such bosons are observed: the photon and the Z boson. The two theories can be unified by expressing the observed bosons as mixtures of the two bosons with $I_3 = 0$. The Gell-Mann and Nishijima relation defines a new quantum number Y , called *hypercharge*, which relates electric and weak charge in the following way:

$$Y = Q/e - I_3^W \quad (1.11)$$

where e is the charge of the electron. The symmetry group of the electroweak interaction is $SU(2)_L \otimes U(1)_Y$. As shown in [Section 1.1.1](#), the Lagrangian is invariant under gauge transformations only if we introduce the covariant derivative, which in this case is given by

$$D_\mu = (\partial_\mu - ig \frac{\tau_i W_{\mu i}}{2} - ig' Y B_\mu) \quad (1.12)$$

where g and g' are the coupling constants related to $SU(2)_L$ and $U(1)_Y$ and the four fields are $W_{i\mu}$ and B_μ . For left-handed fermions, τ_i are the Pauli-matrices with $i = 1, 2, 3$. For right-handed fermions, $\tau_i = 0$, for $i = 1, 2, 3$. The generator of the hypercharge group is $Y/2$.

With this, the relations for the observable vector bosons are expressed as:

$$W^\pm_\mu(x) = \frac{1}{\sqrt{2}} [W_{1\mu}(x) \mp i W_{2\mu}(x)] \quad (1.13)$$

and $W_{3\mu}$ and B_μ as a linear combinations of two different Hermitian fields (A_μ and Z_μ)

$$\begin{pmatrix} W_{3\mu} \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \quad (1.14)$$

where θ_W is the *weak mixing angle*, measured as $\sin^2 \theta_W = 0.23129(5)$ [16].

It is important to notice that the free-boson Lagrangian must include the following kinetic terms

$$-\frac{1}{4} W_{i,\mu\nu} W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (1.15)$$

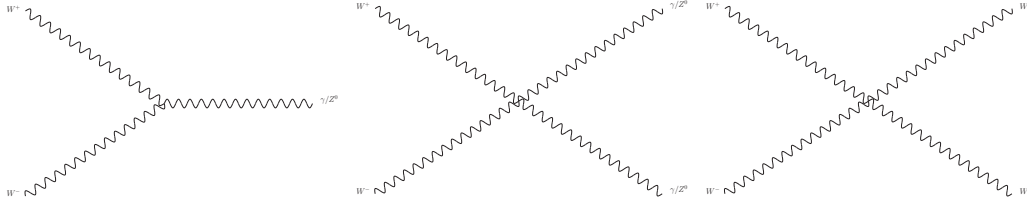


Figure 1.1: Feynman diagrams representing the interactions among mediator bosons (Z^0 , $W^\pm$ and γ^0).

where, given the $SU(2)$ algebra,

$$\begin{cases} W_i^\mu \nu = \partial^\mu W_i^\nu - \partial^\nu W_i^\mu - ig\epsilon_{ijk} W_j^\mu W_k^\nu \\ B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu. \end{cases} \quad (1.16)$$

These terms, caused by the fact that $W_i^\mu(x)$ carry weak electric charge, generate the interactions of the gauge bosons among themselves (Figure 1.1). Writing the covariant derivative for left- and right-handed fermions (ψ_L and ψ_R) explicitly, the Lagrangian reads

$$L = \psi_L^\dagger \gamma^\mu (i\partial_\mu + g \frac{\tau_i W_{\mu i}}{2} + g' Y B_\mu) \psi_L + \psi_R^\dagger \gamma^\mu (i\partial_\mu + ig' Y B_\mu) \psi_R - \frac{1}{4} W_{i,\mu\nu} W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \quad (1.17)$$

Up to this point, fermions and gauge bosons have been considered to be massless. A possible solution to this problem is the spontaneous symmetry breaking mechanism.

1.1.3 Spontaneous symmetry breaking

In **EWK**, the spontaneous symmetry breaking mechanism is applied as following [7, 8, 9]: The basic idea is to choose a particular state among the different minimum-energy states, which is not invariant under the symmetry and which is characterized by a non-vanishing quantity in the vacuum. A generic continuous symmetry which is spontaneously broken is described by the Nambu and Goldstone Theorem [17], which states that a new massless scalar particle appears in the spectrum of possible excitations for each broken generator of the symmetry. This new particle is commonly referred to as *Goldstone boson*.

A simple extension to the **EWK** is obtained by introducing the mass terms of two bosons using complex scalar fields ϕ^+ and ϕ^0 , respectively. These two fields form an isospin doublet with no colour charge:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \quad (I = \frac{1}{2}, Y = 1). \quad (1.18)$$

The Lagrangian, manifestly invariant under $SU(2)_L \times U(1)_Y$, is given by

$$L_{\text{Higgs}} = (\mathcal{D}^\mu \Phi)^\dagger (\mathcal{D}_\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2, \quad \mu^2, \lambda \in \mathbb{R} \quad (1.19)$$

where the latter two terms represent the potential of the field, $V(\Phi)$.

The first term vanishes for constant Φ . It follows that the minimum value for L_{Higgs} corresponds to the value of Φ minimizing $V(\Phi)$. The shape of the potential $V(\Phi)$ depends on the choice of the parameters μ and λ . Firstly, it can be noted that λ has to be greater than 0 to ensure vacuum stability. Secondly, when $\mu^2 < 0$ is chosen, $V(\Phi)$ has a local maximum at the origin and degenerate minima at

$$\Phi_0 = -\frac{\mu^2}{2\lambda} = \frac{1}{2}v^2 \quad \text{where} \quad v = \sqrt{-\frac{\mu^2}{\lambda}}, \quad (1.20)$$

with v the vacuum expectation value of the field Φ .

Assuming $\mu^2 < 0$, the choice of a particular ground state for Φ realizes the symmetry breaking. Expanding the Higgs field $\Phi(x)$ around the chosen minimum

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (1.21)$$

considering a small excitation

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.22)$$

and inserting it into the Lagrangian yields to:

$$\begin{aligned} L_{\text{Higgs}} = & \frac{1}{2} \partial_\mu H \partial^\mu H - \frac{1}{2} 2v^2 \lambda H^2 - \frac{1}{3!} 6v\lambda H^3 - \frac{1}{4!} 6\lambda H^4 \\ & + \frac{1}{2} \frac{v^2 g^2}{4} W_\mu^{-\dagger} W^{-\mu} + \frac{1}{2} \frac{v^2 g^2}{4} W_\mu^{+\dagger} W^{+\mu} \\ & + \frac{1}{2} \frac{v^2 (g^2 + g'^2)}{4} \left(\frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} \right)^2 + 0 \cdot \left(\frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}} \right)^2 \\ & + \frac{1}{4} (2vH + H^2) \left[g^2 W_\mu^- W^{+\mu} + \frac{1}{2} (g^2 + g'^2) \left(\frac{W_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} \right)^2 \right] \end{aligned} \quad (1.23)$$

The first line represents the kinetic term for the Higgs boson, its mass term ($m_H^2 = 2v^2\lambda$) and the Higgs boson self-interaction terms. The second line is the description of the charged $W^\pm$ bosons, obtained with the substitution seen in Equation 1.13, which are now massive ($m_W^2 = \frac{1}{4}v^2g^2$). In the third line, the massive Z^0 boson ($m_Z^2 = \frac{1}{4}v^2(g^2 + g'^2)$) and the massless γ^0 boson can be easily identified with a substitution¹ equivalent to Equation 1.14

$$\begin{cases} Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} \\ A_\mu = \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}} \end{cases} \quad (1.24)$$

It can be noted that the mixing of W_μ^3 and B_μ yields the following relation between the Z^0 and $W^\pm$ boson masses:

$$m_Z = \frac{m_W}{\cos \theta_W} \quad (1.25)$$

Finally, the last line of the Equation 1.23 shows the couplings of the weak gauge bosons to the Higgs boson. From the masses of $W^\pm$ and Z^0 , an approximate branching ratio of the Higgs boson into a pair of vector bosons can be estimated, which corresponds to $\simeq 2.7$ in first approximation.

All SM predictions can be tested with a measurement of three parameters: v and λ described in Equation 1.21 and θ_W mentioned in Equation 1.14. The parameter v can be estimated from precise muon lifetime measurements, i.e. $v = (\sqrt{2}G_F)^{-1/2} \simeq 246 \text{ GeV}$ [5]. The weak mixing angle θ_W has also been already measured with good precision as mentioned in Section 1.1.2. With regards to the λ parameter, it defines the Higgs mass and cannot be predicted by the theory. The discovered Higgs particle has a mass of roughly 125 GeV, thus fixing the value of the last unknown parameter.

¹The numerical factor $\sqrt{g^2 + g'^2}$ has been introduced in order to normalize the combinations of gauge fields $gW_\mu^3 - g'B_\mu$ and $g'W_\mu^3 + gB_\mu$.

To include the fermions masses in the Lagrangian the Yukawa couplings to the Higgs-field is included with the following term

$$L_{\text{Yukawa}} = -g_{d_{ij}} \bar{q}_{L_i} \Phi_C d_{R_j} - g_{u_{ij}} \bar{q}_{L_i} \Phi u_{R_j} - g_{l_{ij}} \bar{l}_{L_i} \Phi e_{R_j} + h.c. \quad (1.26)$$

where

$$\Phi_C = -i\tau_2 \Phi^* \quad (1.27)$$

and q_L (l_L) and u_R , d_R (e_R) being the quark (lepton) SU(2) doublets and singlets in the Lists 1.8 and 1.9. The mass of a fermion f is therefore proportional to the coupling g_ℓ

$$m_f = \frac{g_f v}{\sqrt{2}} \quad (1.28)$$

In conclusion, the Higgs direct coupling to fermions and to bosons is indeed expected to be proportional to the mass in the case of fermions and to the mass squared for bosons.

1.2 Higgs boson discovery at the LHC

The **LHC** is a proton-proton accelerator located at **CERN** in Geneva. The main goal of the **LHC** is to better understand the **SM** and to give an answer to the still open questions in elementary particle physics. Four main experiments have been installed for this purpose. Two of them, **ATLAS** and **CMS**, have been built in particular with the primary scientific goal of discovering the **SM** Higgs boson. The discovery of a Higgs-boson like particle with a mass of about 125 GeV has been published in June 2012 by both experiments simultaneously [18, 19]. An extensive description of the accelerator and the main experiments is given in **Chapter 2**.

1.2.1 Higgs production mechanisms at LHC

The four main processes to produce an **SM** Higgs boson at the **LHC** are represented in **Figure 1.2** as **leading-order (LO)** Feynman diagrams. Their characteristics are listed here along with the level of known accuracy, such as **next-to-leading-order (NLO)** or **next-to-next-to-leading-order (NNLO)** etc.:

- **Gluon Fusion (ggH or gg→H)** is the dominant production process over the entire range of m_H accessible to the **LHC**. The process involves a fusion of two gluons via an intermediate particle loop. As already discussed in **Section 1.1.3**, the Higgs couplings to fermions are proportional to the square of their masses. It follows that the main contribution in the loop comes from the top quark. The inclusive cross section of this process is known at **NNLO QCD** and **NLO EWK** accuracy [20].
- **Vector Boson Fusion (VBF) (qqH or qq→qqH)** is the second largest production mechanism at the **LHC**. Its cross section is about one order of magnitude below the one for the gluon fusion. The Higgs boson is created by an exchange of a W, a Z or a γ boson between two quarks. The final state includes two heavily boosted jets, coming from the hadronization of the quarks which radiate the bosons. These jets are characterized by a large separation in pseudorapidity as well as a large invariant mass and this allows the identification of Higgs events produced via **VBF**. The cross section of this process is known at **NNLO QCD** and **NLO EWK** accuracy [20].
- **Higgs-strahlung (q $\bar{q}'$ →WH and q $\bar{q}'$ →ZH)** is characterized by quark-antiquark annihilation which leads to a virtual boson decaying into a pair formed by an **SM** Higgs and a W or Z boson. If the centre-of-mass energy permits to produce an on-shell associated boson, as in the case of the **LHC**, then its decay products can be used to tag more easily the event. This process, as in the case of **VBF**, is directly sensitive to the Higgs couplings to vector bosons. The cross section of the process is known at **NNLO QCD** and **NLO EWK** accuracy [20].

- In the $t\bar{t}$ associated production ($gg \rightarrow t\bar{t}H$ and $q\bar{q}' \rightarrow t\bar{t}H$) a pair of gluons or quarks annihilates producing two top quarks, which radiate a Higgs boson. Although its cross section is several orders of magnitude smaller than the gluon fusion one, this process is particularly important because it is directly sensitive to the coupling between the Higgs boson and the top quark. Its cross section is known at NLO QCD accuracy [20].

The total production cross section at different centre-of-mass energies for each production mode is shown in Figure 1.3(a) [21].

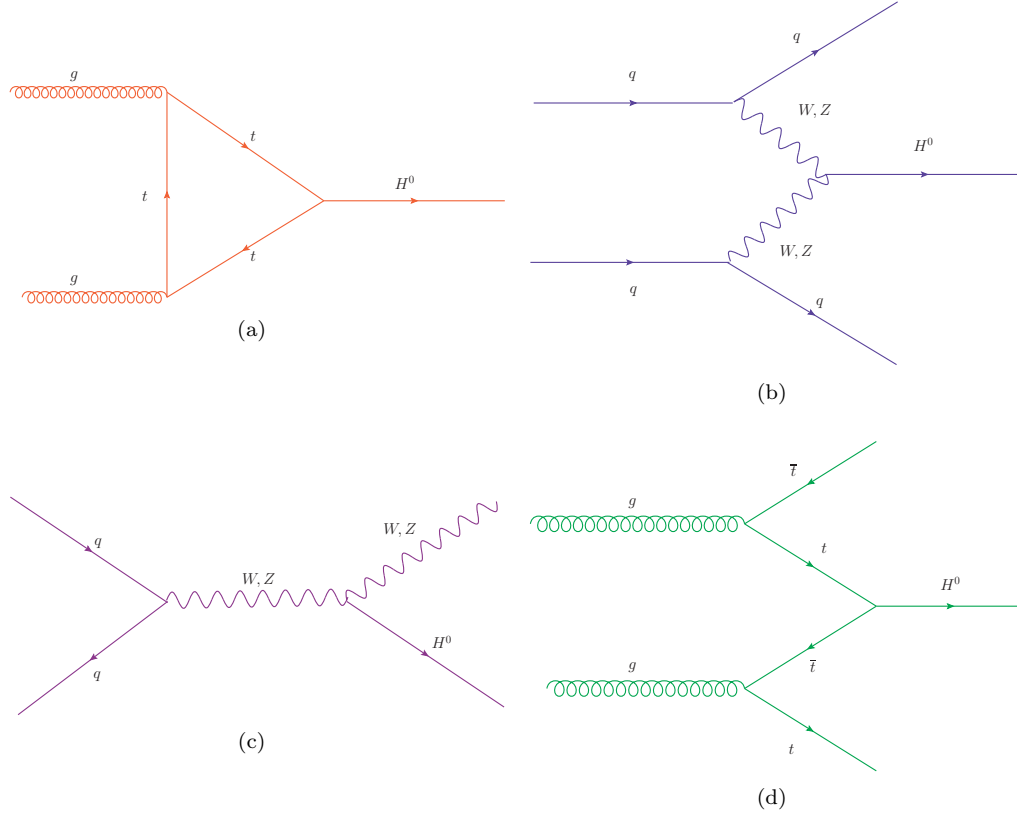


Figure 1.2: Feynman diagrams for the most important LO production processes of the SM Higgs boson: gluon fusion (a), vector boson fusion (b), Higgs-strahlung (c), and $t\bar{t}$ associated production (d).

1.2.2 Higgs decay channels and discovery at LHC

For a better understanding and interpretation of the experimental results in the context of the SM Higgs boson analysis, different decay channels can be exploited. A Higgs boson mass of about 125 GeV provides an excellent opportunity to explore the Higgs couplings to many SM particles as shown in Figure 1.3(b), where the predicted branching ratios for the most relevant decay modes of the Higgs boson as a function of m_H are presented. In Table 1.3 the values for $m_H = 125$ GeV are listed. In particular, the dominant decay modes are $H \rightarrow b\bar{b}$, $H \rightarrow W^+W^-$ and $H \rightarrow gg$, followed by $H \rightarrow c\bar{c}$, $H \rightarrow \tau^+\tau^-$ and $H \rightarrow ZZ$. With much smaller rates the Higgs decays $H \rightarrow \gamma\gamma$, $H \rightarrow \gamma Z$ and $H \rightarrow \mu^+\mu^-$ follow. The $H \rightarrow gg$, $H \rightarrow \gamma\gamma$, and $H \rightarrow \gamma Z$ decays are particularly important to test the validity of the SM, because they are loop induced and provide indirect information on the Higgs couplings to WW , ZZ and $t\bar{t}$ for different combinations.

The total width of the Higgs boson resonance increases rapidly with the Higgs mass due to the opening of new channels and the associated phase space increases. For a 125 GeV SM Higgs boson, the total width is $\Gamma_H = 4.07 \times 10^{-3} \text{ GeV}^{+4.0\%}_{-3.9\%}$ [5].

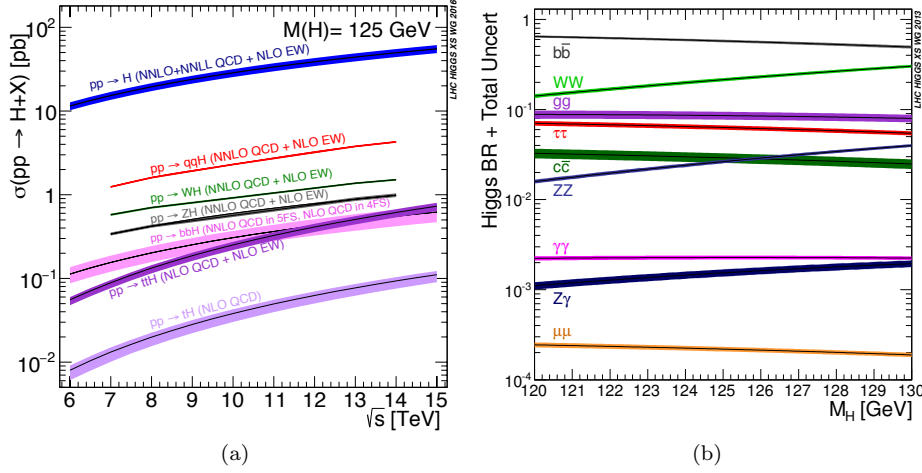


Figure 1.3: (a) **Standard Model** Higgs boson production cross sections as a function of the center of mass energy for pp collisions for the different production mechanisms [5].
 (b) The branching ratios for the main decays of the **SM** Higgs boson near $m_H = 125$ GeV [5].

Table 1.3: The predicted branching ratios and the relative uncertainty for a **SM** Higgs boson with $m_H = 125$ GeV [5].

Decay Channel	Branching ratio	Rel. uncertainty
$H \rightarrow \gamma\gamma$	2.27×10^{-3}	+5.0% -4.9%
$H \rightarrow ZZ$	2.62×10^{-2}	+4.3% -4.1%
$H \rightarrow W^+W^-$	2.14×10^{-1}	+4.3% -4.2%
$H \rightarrow \tau^+\tau^-$	6.27×10^{-2}	+5.7% -5.7%
$H \rightarrow b\bar{b}$	5.48×10^{-1}	+3.2% -3.3%
$H \rightarrow \gamma Z$	1.53×10^{-3}	+9.0% -8.9%
$H \rightarrow \mu^+\mu^-$	2.18×10^{-4}	+6.0% -5.9%

Among all the possible decays, **ATLAS** and **CMS** focus their observation on the five decay channels $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^{(*)} \rightarrow l^+l^-l^+l^-$, $H \rightarrow WW^{(*)} \rightarrow l\nu_l l\nu_l$, $H \rightarrow \tau^+\tau^-$ and $H \rightarrow b\bar{b}$. The different characteristics of these decay channels are the following:

- $H \rightarrow \gamma\gamma$: Because of the loop process, this decay channel has a very low branching ratio compared to the other channels. Despite this, the decay into two photons has the advantage that the signal signature, two isolated photons, can be well identified experimentally over a large background. The background comes from the **SM** production of prompt photon² pairs or from the misidentification of jets or electrons. The most critical aspect in this analysis is achieving the narrowest possible peak in the diphoton invariant mass distribution. Since the **SM** Higgs total width is lower than 1 GeV for a 125 GeV Higgs mass, the energy resolution of the electromagnetic calorimeter dominates the resonance.
- $H \rightarrow ZZ$: This is one of the “golden modes” for the discovery of the Higgs boson, in particular in the fully leptonic final state $H \rightarrow ZZ^{(*)} \rightarrow l^+l^-l^+l^-$ with $l^+l^- = e^+e^-$ or $\mu^+\mu^-$.

²Prompt photons are photons produced promptly in the collision, before the quarks and gluons have had time to form hadrons.

This search is performed by looking for a clean signature with four isolated leptons in the final state, producing a narrow invariant-mass peak above a small continuous background. The backgrounds for this channel are the irreducible **SM** production of ZZ , and the $t\bar{t}$ and $Zb\bar{b}$ processes, that can be suppressed by requirements on the lepton isolation, transverse momentum, event vertex and invariant mass. For example, the invariant mass of a lepton pair must be approximately around the mass of the Z^0 particle.

- $H \rightarrow W^+W^-$: In the low mass region the search in this channel is performed in the fully leptonic mode $H \rightarrow WW^{(*)} \rightarrow l\nu_l l\nu_l$, with $l\nu_l = e\nu_e, \mu\nu_\mu$ or $\tau\nu_\tau$. The signature for this channel is two isolated and opposite-charge leptons with same or different flavor, with significant missing momentum due to undetected neutrinos. To enhance the sensitivity to a **SM** Higgs boson signal, a counting analysis is performed in each final state and category using a selection optimized for each m_H hypothesis considered. The dominant ones are non-resonant WW , $t\bar{t}$ and $W + \text{jets}$, reduced by particular selections applied on lepton momentum and isolation.
- $H \rightarrow b\bar{b}$: This channel is the one with the largest branching ratio of the five search modes. The main challenges of this analysis are that the signal is overwhelmed by the background, mainly **QCD** multijet production of b quarks, and that the mass resolution achieved is worse compared to other channels, since the Higgs decays into jets.
- $H \rightarrow \tau^+\tau^-$: This channel has similar characteristics as the $H \rightarrow b\bar{b}$ one, including a large branching ratio compared to the other channels as well as the background challenges. It is discussed in more detail in [Section 1.2.3](#).

The discovery and the study of the **SM** Higgs boson was one of the main goals of the **LHC** project. As already mentioned at the beginning of this section, both **CMS** and **ATLAS** experiments reported in 2012 the observation of a Higgs-like boson with a mass of about 125 GeV. The results for the $H \rightarrow \gamma\gamma$ channel with the **ATLAS** experiment and the $H \rightarrow ZZ$ with **CMS** are shown in [Figure 1.4](#) as examples. The properties of this new resonance have been extensively measured using the full dataset of 5.1 fb^{-1} at 7 TeV and 19.7 fb^{-1} at 8 TeV, collected in 2011 and 2012, and final combination results were published for all decay modes in 2015 [5]. In [Figure 1.5](#) the measurement of the signal strengths can be viewed as the measurements of the cross sections (σ) times the branching fraction (BR) normalized to the **SM** prediction. The couplings of the new boson to the **SM** particles have also been measured extensively, and no significant deviations from the **SM** prediction have been found.

1.2.3 Evidence for an SM Higgs boson decaying into a pair of τ leptons

At hadron colliders, $H \rightarrow b\bar{b}$ and $H \rightarrow \tau^+\tau^-$ are the most promising channels for probing the coupling of the Higgs field to quarks and leptons. The branching fraction is about 55% and about 6% for a Higgs boson with $m_H \simeq 125 \text{ GeV}$, respectively. Nevertheless, the presence of very large backgrounds makes the isolation of a Higgs boson signal in these channels very challenging.

The $H \rightarrow \tau^+\tau^-$ analysis strategy is to reconstruct the invariant mass of the $\tau^+\tau^-$ pair from a kinematic fit of the visible products from the two τ leptons and the missing energy observed in the event. The τ can decay into electrons, muons or hadrons. In total, the τ lepton decays hadronically (τ_h) approximately 64.79% of the time, while 17.82% into electrons (τ_e) and 17.39% into muons (τ_μ). In the hadronic decay, τ_h decays typically into either one or three charged pions or kaons, up to two neutral pions (π^0), and one neutrino (ν_τ). Due to the presence of missing neutrinos in the τ decay, the invariant mass resolution achieved is worse compared to other channels, around 15%. The major sources of background are the irreducible $Z \rightarrow \tau^+\tau^-$ process, and $Z \rightarrow e^+e^-$, $W + \text{jets}$, $t\bar{t}$ and multijet production. The sensitivity of the search is enhanced by requesting a Higgs boson produced via **VBF**, consisting of a $\tau^+\tau^-$ pair with two energetic jets separated by large pseudorapidity. In order to profit from this, **ATLAS** and **CMS** divide the $\tau^+\tau^-$ events into categories based on the number and kinematic properties of additional energetic jets in the event. All six combinations of leptonic ($\tau \rightarrow l\nu$ with $l = e, \mu$) and hadronic ($\tau \rightarrow \text{hadrons}$ ν) tau decays are considered from both experiments. Both experiments report the measurement of the signal

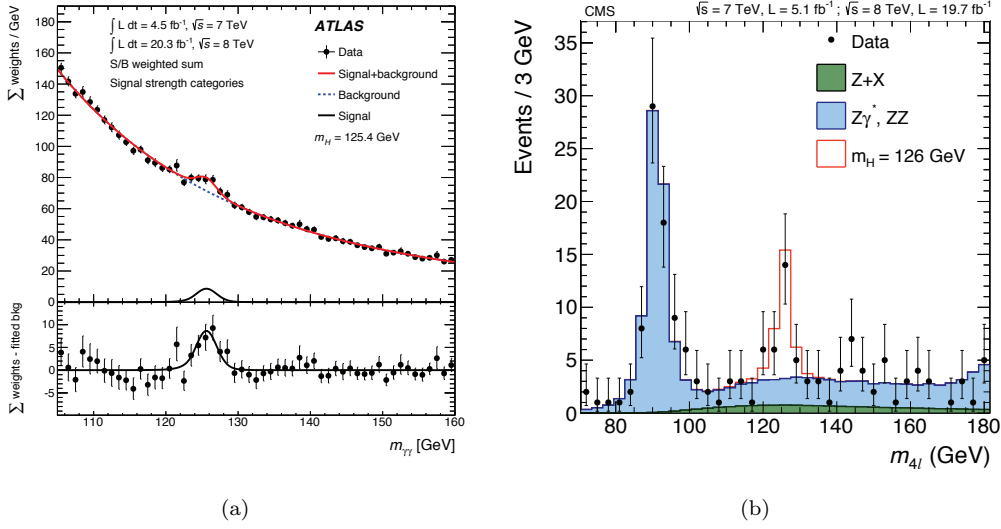


Figure 1.4: Two decay channel discovery examples for the **ATLAS** and **CMS** experiment: the $H \rightarrow \gamma\gamma$ invariant mass distribution observed by **ATLAS** (a) [22] and the $H \rightarrow ZZ^{(*)} \rightarrow l^+l^-l^+l^-$ combined distribution from **CMS** Run 1 data (b) [23].

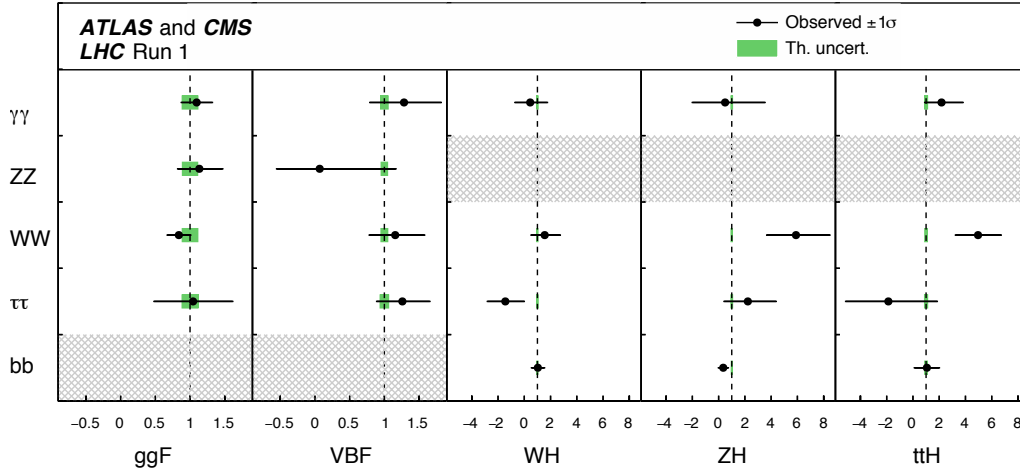


Figure 1.5: Combined measurements of the products $\sigma \cdot BR$ for the five main production and five main decay modes [5].

strength parameter μ , which is the observed product of the Higgs boson production cross section and its branching ratio in units of the corresponding **SM** values:

$$\mu = \frac{(\sigma \cdot BR)_{\text{obs}}}{(\sigma \cdot BR)_{\text{SM}}} \quad (1.29)$$

A value of $\mu = 1$ corresponds to the presence of a **SM** Higgs boson signal and a null value to the absence of the Higgs boson.

ATLAS search

The search for a Higgs boson decaying into a pair of τ leptons by the **ATLAS** experiment is designed to be sensitive to the major production processes of a **SM** Higgs boson, i.e. production via gluon fusion, **VBF** and production with associated W or Z boson. The different signatures produced by different final-states are exploited by defining an event categorization. After preselection and categorization of the $H \rightarrow \tau^+\tau^-$ events, a **Multivariate Analysis (MVA)** technique based on a **Boosted Decision Tree (BDT)** is used to extract the final results. This technique allows to exploit

correlations between final-state observables. It is described in more detail in [Chapter 4](#). As a cross-check, a separate analysis where threshold on kinematic variables are applied is carried out. The results using the entire dataset collected by the [ATLAS](#) experiment during the 2011 and 2012 data-taking periods, corresponding to integrated luminosities of 4.5 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$ and 20.3 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$, is shown in [Figure 1.6\(a\)](#). An excess of events over the expected background from other Standard Model processes is found with an observed (expected) significance of 4.5 (3.4) standard deviations. This excess provides evidence for the direct coupling of the Higgs boson to fermions. The measured signal strength, normalised to the Standard Model expectation, of $\mu = 1.43^{+0.43}_{-0.37}$ is consistent with the predicted Yukawa coupling strength in the [Standard Model](#) [24].

CMS search

The $H \rightarrow \tau^+\tau^-$ analysis performed by the [CMS](#) collaboration is similar to the [ATLAS](#) one. Also in this case, the $\tau^+\tau^-$ pair is produced via [VBF](#) or gluon fusion. Furthermore, the associated vector boson production mode is included by requiring additional leptons that originate from a decaying W or Z boson in the final state. The analysis categorizes events using thresholds on kinematic variables: firstly the event sample is split into three exclusive categories per decay channel, according to the number of jets and then, for each category, a two-dimensional distribution of the variables that maximize the discovery potential are chosen. In the dataset collected between 2015 and 2016 with centre-of-mass energy increased to 13 TeV , an excess of events is observed over the expected background prediction with a significance of 4.9 standard deviations for $m_H = 125 \text{ GeV}$, in comparison to an expected significance of 4.7 standard deviations [25]. The combination of data taken at $\sqrt{s} = 7 \text{ TeV}$ and at $\sqrt{s} = 8 \text{ TeV}$ leads to a total observed significance of 5.9 standard deviations, equal to the expected significance. The measured signal strength is $\mu = 1.09^{+0.27}_{-0.26}$ times the [SM](#) expectation, and therefore is also consistent with the predicted value in the [SM](#). Some results for the 13 TeV dataset and the analysis used in Run 1 with 2011 and 2012 data are shown in [Figure 1.6](#).

1.3 Beyond the Standard Model

The discovery of the Higgs boson was the last missing piece in the fundamental properties of the [SM](#). However, there are many observed phenomena that the [SM](#) provides no explanation for. Some of the open questions are:

- The reason for the existence of exactly three generation of quarks and leptons with vastly different masses.
- The composition of the main part of the matter-energy content of the universe: It has been experimentally shown that the particles described in the [SM](#) account for merely about 20% of the matter in the universe, while the remaining 80% is so-called *Dark Matter* which is not accounted for in the [SM](#). Even when taking *Dark Matter* into account, there remain approximately 70% of the matter-energy-content of the universe, which is referred to as *Dark Energy* [27].
- The matter-antimatter asymmetry in our universe: The [SM](#) predicts that matter and anti-matter are created in the same quantity in each process, no known mechanism in the [SM](#) is sufficient to explain why we live in a universe dominated by matter.
- The gravitational interaction: As already mention in [Section 1.1.1](#), gravitation is not included in the [SM](#), although many attempts to do so have been made. The boson that should carry the gravitational interaction is still only hypothetical [28].
- The mass of neutrino particles and their mixing as demonstrated by neutrino oscillation experiments [5, 6].

Several [Beyond the Standard Model](#) (BSM) models exist that try to address some of these unsolved issues. This provides a large terrain to probe experimentally, and both [ATLAS](#) and [CMS](#) experiments at the [LHC](#) run extensive physics programs to probe as many scenarios as possible.

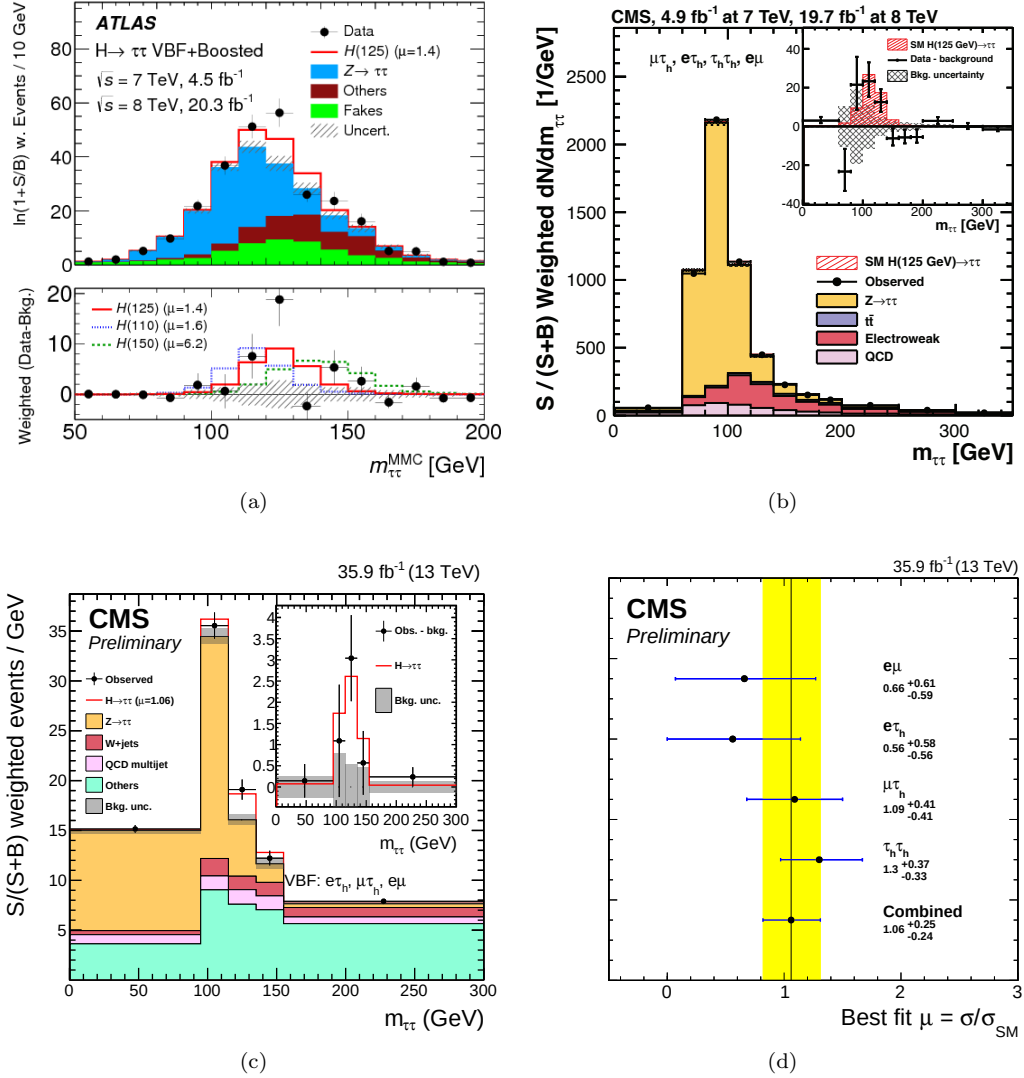


Figure 1.6: (a) Distribution of the reconstructed invariant mass, where events are weighted by $\ln(1+S/B)$ predicted with the BDT for all channels by the **ATLAS** experiment [24]. (b) The observed and predicted invariant mass distributions for all $H \rightarrow \tau^+\tau^-$ subchannels combined by the **CMS** experiment. The inset shows the difference between the observed data and the expected SM background contributions, together with the expected signal distribution for a SM Higgs boson with $m_H = 125$ GeV [26]. (c) Combined observed and predicted $m_{\tau\tau}$ distributions for the VBF category of $\tau_e\tau_h, \tau_\mu\tau_h, \tau_e\tau_\mu$ [25]. (d) Best-fit signal strength per channel, for $m_H = 125$ GeV. The constraints from the global fit are used to extract each of the individual best-fit signal strengths [25].

CHAPTER 2

The CMS detector at the LHC

To Ellen Richards, first woman to attend and become professor at the MIT. (1884)

The **Large Hadron Collider (LHC)** is a machine built to give answers to the still open questions in particle physics described in **Chapter 1**, for example the existence of the Higgs boson or other particles produced by mechanism beyond the **Standard Model**. To achieve this goal, several independent experiments are installed at the **LHC**. A brief overview of the **LHC** accelerator complex and operations is given in **Section 2.1**. In **Section 2.2** the **Compact Muon Solenoid (CMS)** experiment with its sub-detectors and software is presented. Finally, the reconstruction of the high-level physics objects used in **CMS** analyses is described in **Section 2.3**.

2.1 The Large Hadron Collider

The **Large Hadron Collider (LHC)** [29] is a proton-proton and heavy ions accelerator and collider built by the **European Organisation for Nuclear Research (CERN)** in the Geneva region under the border between France and Switzerland. It is located in the existing 26.7 km tunnel, 100 m below ground level, which was constructed between 1984 and 1989 for the **CERN Large Electron-Positron collider (LEP)** machine. The **LHC** is an unprecedented project, not only for the performance planned and achieved, but also in terms of human effort, costs and cutting-edge technologies. Its physics program consists mostly of proton-proton collisions. The heavy ions program is run for about one month of each year, during which ionised lead atoms are made to collide either with each other or with protons.

The main goal of the **LHC** is the expected discovery of the Higgs boson in order to complete the **Standard Model (SM)** and to study **Beyond the Standard Model (BSM)** scenarios, in particular to verify or refute the hypothesis of the existence of Super-symmetric particles and to search for new phenomena such as Dark Matter. Many **BSM** processes are already expected at the TeV scale but are characterized by very low cross-sections. A theoretical introduction on this topic can be found in **Chapter 1**. At the **LHC**, part of the physics program is also dedicated to heavy ion physics with the main goal of studying the weak bound among quarks and gluons under conditions similar to those in the early universe, the so-called *quark-gluon plasma*.

The most relevant parameters for any particle accelerator are the centre-of-mass energy and the luminosity. The former is important in order to produce new potentially heavy particles and to study their features in depth, while the latter is related to the rate of interactions occurring, and therefore with the production of rare particles. The main motivations for choosing a hadron collider over a lepton one are:

- the cross-sections of several predicted processes are higher for strong interaction processes;
- the interacting particles — quarks and gluons — carry variable fraction of the proton’s total energy, and therefore allow a wide range of centre-of-mass energies to be probed simultaneously;
- the centre-of-mass energy can be significantly larger because the hadron collider suffers less from energy loss due to synchrotron radiation.

For these reasons, hadron accelerators have been used as “discovery machines” that can break records in terms of centre-of-mass energy, while lepton colliders served as “high-precision measurement machines” in a lower energy range. The development of the centre-of-mass energy for several particle colliders shown over time can be found in Figure 2.1 for both hadron and lepton colliders.

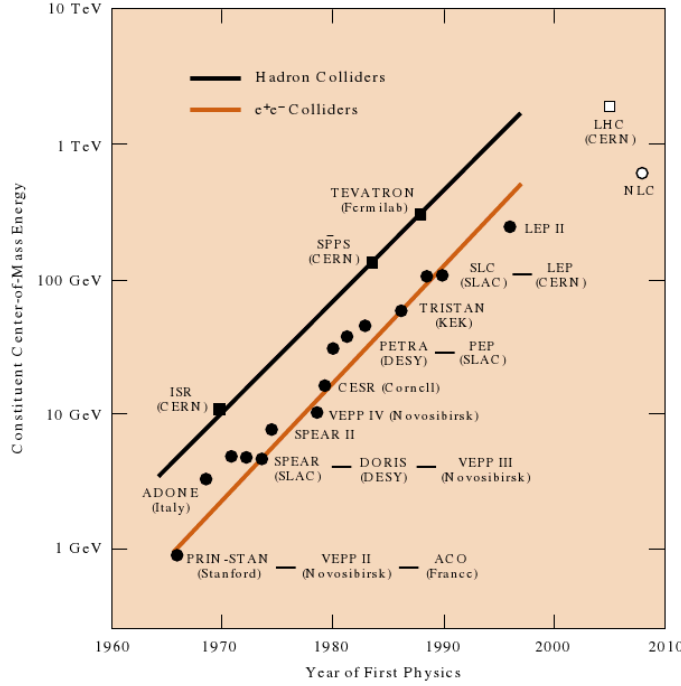


Figure 2.1: The centre-of-mass energy of several particle colliders plotted over time. The maximal centre-of-mass energy increased almost exponentially until the end of the past century. Hadron colliders have provided significantly higher centre-of-mass energies than lepton accelerators at any given time [30].

The luminosity L can be expressed by the following formula:

$$L = \gamma \frac{n_b N^2 f_{rev}}{4\pi \beta^* \epsilon_n} R \quad (2.1)$$

where γ is the proton beam energy in unit of rest mass; n_b is the number of bunches per beam, 2808 for the nominal LHC value in case of for 25 ns bunch spacing; N is the bunch population, 1.15×10^{11} protons at 25 ns; f_{rev} is the revolution frequency, 11.2 kHz at nominal value; β^* is the beam beta function, also called focal length, which is 0.55 m at the collision point at nominal design; ϵ_n is the transverse normalized emittance, $3.75 \mu\text{m}$ for the nominal design; R is a luminosity geometrical reduction factor, amounting to 0.85 at a β^* of 0.55 m and can be expressed by

$$R = 1 / \sqrt{1 + \frac{\theta_c \sigma_z}{2\sigma_T}} \quad (2.2)$$

where θ_c is the full crossing angle between colliding beam, $285 \mu\text{rad}$ as nominal design; and σ_T and σ_c are the transverse and longitudinal sizes, nominally $16.7 \mu\text{m}$ and 7.55 cm , respectively. With the nominal parameter values listed above and a centre-of-mass energy of 14 TeV, an instantaneous luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ for proton-proton physics is expected at the LHC. Slightly different parameters are chosen for the heavy ion program, which uses a design energy of 2.76 TeV per nucleon and a luminosity of $10^{27} \text{ cm}^{-2} \text{ s}^{-1}$.

2.1.1 The injection chain

The beam is delivered to the **LHC** for its final acceleration by an injection chain consisting of multiple stages, shown in **Figure 2.2**. The first step in the injection chain is a linear accelerator: the **Linear accelerator 2 (LINAC2)** for proton beams, and the **Linear accelerator 3 (LINAC3)** in the case of heavy ions. In a second stage, the proton beam is accelerated from 50.0 MeV to 1.4 GeV in the **Proton Synchrotron Booster (PSB)**, while the heavy ion beam is accelerated from 4.2 MeV to 72.2 MeV in the **Low Energy Ion Ring (LEIR)**. The acceleration process for the proton beam continues in the **PSB** which is composed of four stacked rings, where each ring injects a bunch into the **Proton Synchrotron (PS)**. Here the proton bunches are prepared and accelerated up to energies of 26 GeV before being injected into the **Super Proton Synchrotron (SPS)**. Heavy ion beam is accelerated in **LEIR**, which forms two bunches that are injected into **PS** using the same mechanism as for proton bunches. **PS** splits the two heavy ion bunches in half to provide four bunches with 5.9 GeV to **SPS**. The **SPS** accelerator is the one before **LHC** and the bunches achieve the energy of 450 GeV (177 GeV) for the protons (heavy ion) beam. Finally, they are injected into the main **LHC** ring, which consists of two adjacent parallel beam pipes where the two beams travel in opposite directions. Eight radio-frequency resonant cavities which oscillate at 400 MHz accelerate the bunches to their nominal peak energy, 7 TeV for protons and 5.1 TeV in the case of heavy ions [29, 31, 32]. At the design instantaneous luminosity, the final proton beam is composed of approximately 3000 bunches, each containing 10^{11} protons, separated in time by 25 ns.

A strong magnetic field is required in order to guide the beams around the accelerator ring. This is provided by 1232 dipole magnets that keep the beams on their circular path as well as 392 quadrupole magnets that focus the beams in order to maximize the luminosity at the intersection points, where the two beams cross. The magnets are made of a Niobium-Titanium alloy and work in the superconducting regime. Therefore, they are kept at a temperature of 1.9 K by means of super-fluid Helium. Since collisions occur between particles of the same charge, two separate beam pipes are required, with two opposite magnetic field configurations.

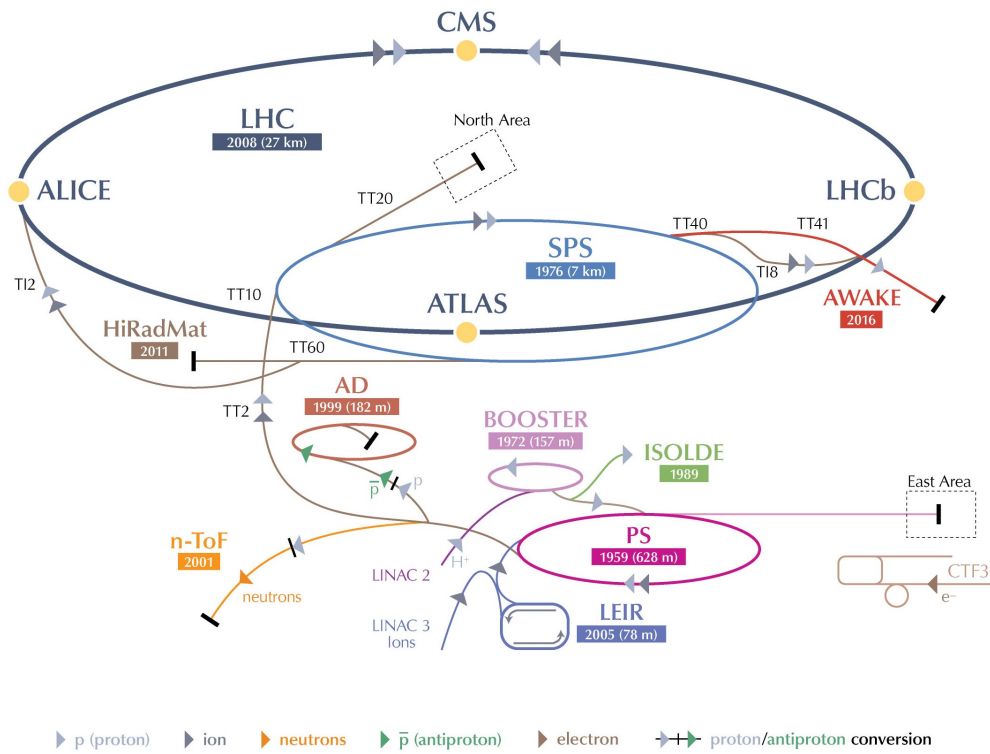
The four main experiments operating at the **LHC** are described in **Section 2.1.4**. **PSB**, **PS** and **SPS** also provide particle beams to smaller experiments located at **CERN**, such as the **Isotope mass Separator On-Line facility (ISOLDE)** in the case of **PSB**, the **Antiproton Decelerator (AD)** ring for **PS**, and several fixed-target experiments in the so-called *North Area* as well as the **Proton Driven Plasma Wakefield Acceleration Experiment (AWAKE)** for the **SPS**.

2.1.2 Operational periods of the LHC

The **LHC** operation has been split in so-called *running periods*, divided by *long shutdowns* of multiple years in which upgrade work is performed. As can be seen in **Figure 2.3**, **LHC** did not start operating with the design parameters of 14 TeV centre-of-mass energy and $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ instantaneous luminosity, but was initially operated at significantly lower energies and luminosity.

In 2010 head-on collisions of the two proton beams started with the first running period, called *Run 1*. During this first running period the initial centre-of-mass energy of 7 TeV was increased to 8 TeV in 2012 and the instantaneous luminosity reached about $8 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. Moreover, the **LHC** was operated with a time separation between bunches of 50 ns. The data volume delivered was 45 pb^{-1} and 6.1 fb^{-1} for 2010 and 2011 respectively, while in 2012 an integrated luminosity of 23.3 fb^{-1} was delivered.

Thanks to the successful upgrade of the accelerator done during the **Long Shutdown 1 (LS1)**, since spring 2015 the machine operates at a centre-of-mass energy of 13 TeV with a bunch crossing separation of 25 ns and increased luminosity. During this shutdown, all **LHC** experiments underwent upgrades in order to prepare for an instantaneous luminosity of up to $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, which significantly exceeds the **LHC's** nominal luminosity. In the second running period, *Run 2*, the integrated luminosities delivered during 2015 and 2016 were 4.2 fb^{-1} and 41.1 fb^{-1} respectively. *Run 2* will continue until the end of 2018, when the **Long Shutdown 2 (LS2)** will start. The **LS2** will be followed by *Run 3*. At the end of *Run 3* it is expected that about 300 fb^{-1} will have been collected. **Long Shutdown 3 (LS3)** is scheduled to last from 2024 to mid 2026 and the main preparation of the accelerator and the experiments will take place for the High Luminosity phase of the **LHC (HL-LHC)**, which will provide an unprecedented instantaneous luminosity of $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. A



LHC Large Hadron Collider SPS Super Proton Synchrotron PS Proton Synchrotron
 AD Antiproton Decelerator CTF3 Clic Test Facility AWAKE Advanced WAKEfield Experiment ISOLDE Isotope Separator OnLine DEvice
 LEIR Low Energy Ion Ring LINAC LINear ACcelerator n-ToF Neutrons Time Of Flight HiRadMat High-Radiation to Materials

Figure 2.2: Illustration of the CERN accelerator complex including LHC and non-LHC experiments ((©CERN, 2013–2016) [33]).

more detailed description of the HL-LHC period and the upgrade of the four main experiments is given in Chapter 5. The evolution of the integrated and the instantaneous luminosity versus time from 2010 to 2017 is shown in Figure 2.4.

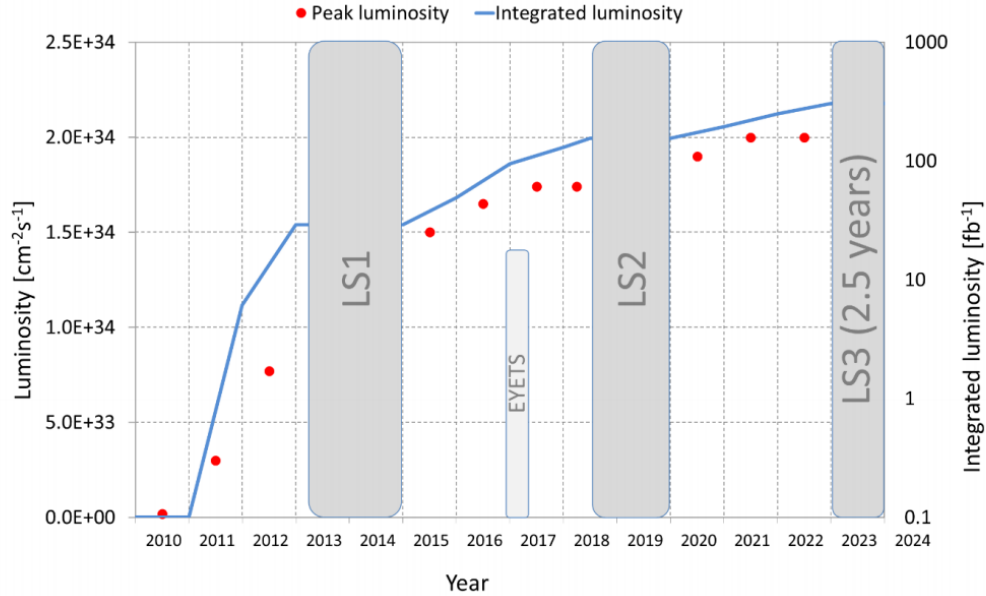


Figure 2.3: LHC's planned schedule until 2024. Running periods are interspersed with long shutdowns that are used to consolidate and upgrade both the accelerator and the experiments [34].

2.1.3 LHC operations

At the nominal instantaneous luminosity of $1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, the beams in the LHC accelerator consist of 2808 bunches of 1.1×10^{10} protons each. The beams have a small transverse size of $15 \mu\text{m}$ and collide every 25 ns. At the nominal conditions, around 1000 charged particles will emerge from the interaction region and about 20 inelastic collisions on average will be superimposed on the event of interest. Already from the start of 2015, nominal conditions were exceeded achieving a mean of about 50 superimposed inelastic collisions, referred to as *pileup*. This large number is one of the most serious challenges for the experimental operation at the LHC. It can come from multiple collisions produced in the same bunch crossing (*in-time pileup*) or it can be due to events overlapping between successive bunch crossings (*out-of-time pileup*). An example of a typical Fill¹ of LHC in 2017 is shown in Figure 2.5: At the beginning of the physics collisions, the peak luminosity and the respective number of in-time pileup events are at the maximum and then decrease almost linearly with time.

In order to recognize interesting events for the physics analyses, a good knowledge of the phenomenology of the collisions and a good choice of detectors technologies is required. In a proton-proton scattering two main situations can occur:

1. *soft collision*: a small momentum is transferred between two incoming protons interacting at large distances. The particles originating from this kind of collisions have a small transverse momentum but a large longitudinal one. These are the most likely events to occur.
2. *hard collision*: the distance between two protons is very small and a large amount of momentum is transferred. The particles produced in these events have a large scattering angle with respect to the beamline. These are rarer but more interesting events because heavy particles can be produced.

¹The Fill is part of the operational period of a collider where the beams are declared stable and collisions can occur.

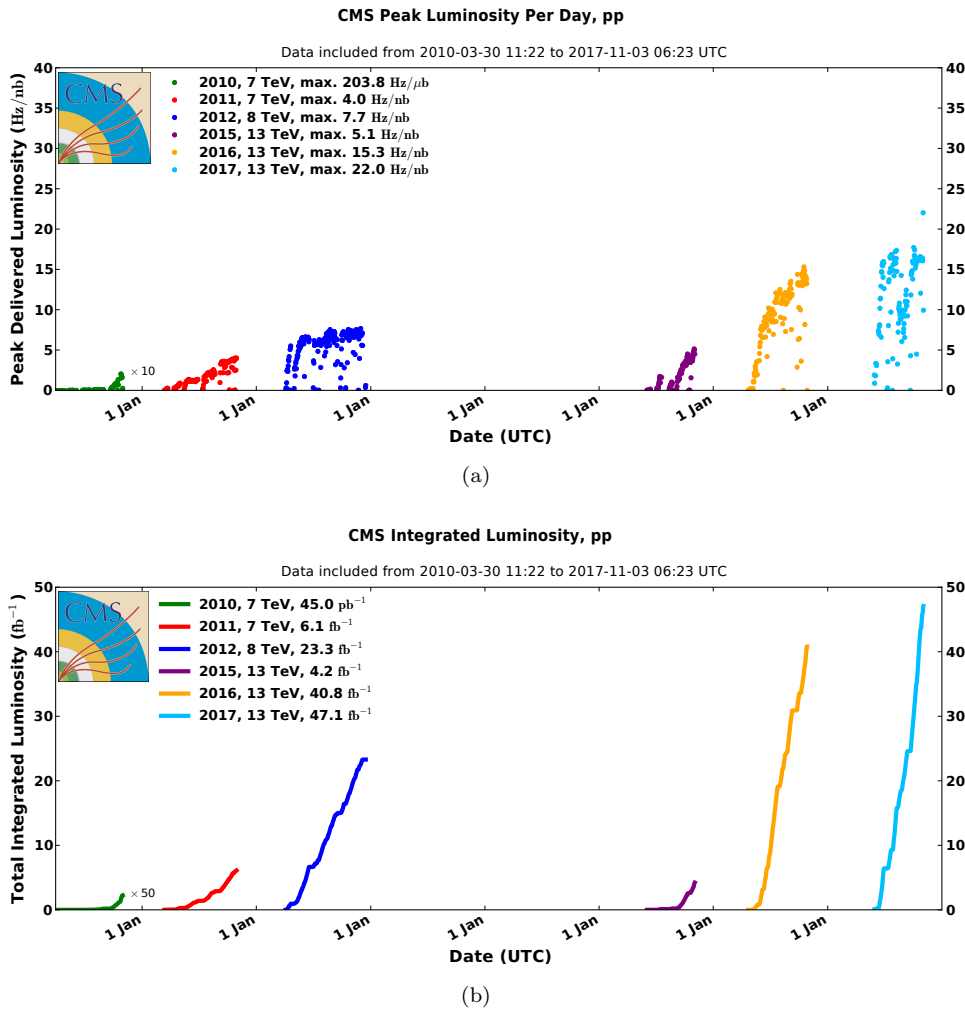


Figure 2.4: Evolution of the instantaneous (a) and total (b) integrated luminosity as a function of time from 2010 to 2017 [35].

Concerning detectors technologies, each experiment has made different choices where it was possible, in order to detect particles efficiently and to cope with the high radiation environment of LHC. In general, the large flux of particles coming from the interaction region is dealt with using high-granularity and radiation-hard detectors with good time resolution.

2.1.4 Experiments at the LHC

The LHC provides particle collisions at four interaction points where four large experiments have been installed: A Toroidal LHC Apparatus (ATLAS) [36] and Compact Muon Solenoid (CMS) [37] are so-called general purpose experiments, while LHC Beauty (LHCb) [38] and A Large Ion Collider Experiment (ALICE) [39] have a more specialized physics program. These four large experiments are complemented by smaller ones sharing their interaction point.

This thesis focuses mainly on track reconstruction. Therefore, the physics goal together with the tracking system choice of each large experiment is described in this section. The CMS experiment is described in more detail in Section 2.2.

ALICE and its tracking system

ALICE is primarily devoted to the investigation of high energy ion physics. It is designed to study the physics of strongly interacting matter and of quark-gluon plasma at extreme energy densities

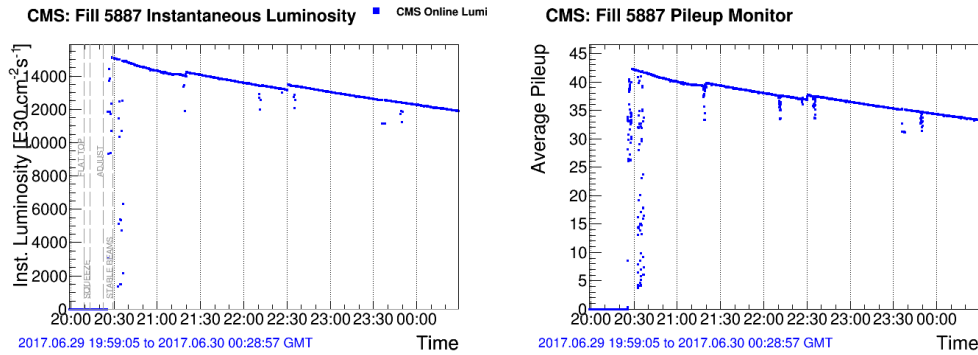


Figure 2.5: Instantaneous luminosity and pileup recorded by the **CMS** experiment as a function of the time during an **LHC** Fill in 2017.

and temperatures in nucleus-nucleus collisions. Given the complexity of ion physics events, **ALICE** is equipped with the largest number of detector types. Among them, an ensemble of cylindrical sub-detectors is focused on measuring the passage of particles carrying an electric charge using the bending power of the magnetic field. From inside out, **ALICE** tracking detectors are: the **Inner Tracking System (ITS)** consisting of six planes of high-resolution silicon pixel, drift, and strip detectors, the cylindrical **Time-Projection Chamber (TPC)**, and the **Transition Radiation Detector (TRD)**. The main goals of the **ITS** are to identify secondary vertices of heavy flavour and strange particle decays, to track low-momentum particles, and to improve the impact parameter and momentum resolution. The **TPC** detector can provide an efficient and robust tracking due to its time information also in a very high multiplicity environments ($\mathcal{O}(10000)$ charged particles). Finally, the **TRD** is also used for tracking in the central region, improving the p_T resolution at high momentum. In Figure 2.6(a) the **ITS-TPC** matching efficiency as a function of the transverse momentum for 2010–2013 data and Monte Carlo for heavy ion collisions is shown [40].

LHCb and its tracking system

As the name indicates, **LHCb** focuses on physics involving bottom quarks and investigation of CP violation phenomena. These studies require the measurement of the rare decays of B_d , B_s , and D mesons which are produced with a large cross-section at the **LHC**. Given the fact that b hadrons are predominantly produced in the forward or backward cone, the **LHCb** experiment is a single-arm spectrometer in contrast to the other three experiments. In order to exploit the large number of b hadrons, it requires a robust and flexible trigger as well as a **DAQ** that allows high bandwidth data taking and provides powerful online data processing. Furthermore, superior vertex and momentum resolution are crucial to study the rapidly oscillating $B_s - \bar{B}_s$ meson system. **LHCb** is thus equipped with a highly specialized silicon microstrip detector close to the interaction point, the **Vertex Locator (VELO)**. It can be moved to only 7 mm distance from the proton beams and therefore measures the position of the primary vertices and the impact parameters of the track with extremely high precision. A further silicon microstrip detector, the **Tracker Turicensis (TT)**, is placed before the dipole magnet. Its task is to improve the momentum resolution of reconstructed tracks and reject duplicate tracks. The magnet is positioned behind the **TT** detector. It bends the flight path of the particles in the $x - z$ plane and therefore allows the determination of their momenta. The tracking system is completed by the T stations (T1-T2-T3), which, together with the information from the **VELO**, determine the momentum and flight direction of the particles. The T stations are composed of silicon microstrip sensors close to the beam pipe and by straw-tubes in the outer regions. The track reconstruction efficiency for the 2015 data and for weighted simulation as a function of p_T can be seen in Figure 2.6(b) [41]. **LHCb** was designed for luminosities of $10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, which requires luminosity levelling when the **LHC** is producing higher luminosities: The beams are focused less compared to **ATLAS** or **CMS** before entering the collision area of **LHCb** to reduce the instantaneous luminosity.

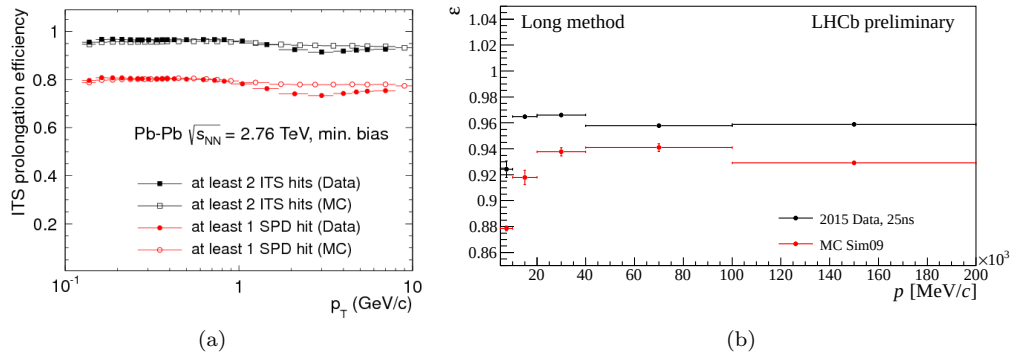


Figure 2.6: (a) **ITS-TPC** matching efficiency vs. p_T for data and Monte Carlo for Pb-Pb collisions for **ALICE** experiment [40].

(b) **LHCb** track reconstruction efficiency for the 2015 data and for weighted simulation as a function of p_T . The method used to establish the tracking performance is so-called “long-method” and it is described in Ref. [41].

ATLAS and its tracking system

The leading principle in the design of the remaining two **LHC** experiments, **ATLAS** and **CMS**, was the aim to cover as wide a range of physics measurements and searches as possible in order to be able to take full advantage of the discovery potential of the **LHC**. They were devised to operate at the highest luminosities the **LHC** can provide.

ATLAS is the largest of the four **LHC** experiments, measuring 25 m in diameter and 44 m in length. Its main distinguishing feature is the magnet system, composed by a Central Solenoid Magnet (2 T), a Barrel Toroid and End-cap Toroids (4 T each). The **ATLAS** inner detector is very compact and highly sensitive in order to accurately measure the decay products of each collision. It consists of three different systems of sensors immersed in the solenoid magnetic field: the Pixel Detector, the **Semiconductor Tracker (SCT)**, and the **Transition Radiation Tracker (TRT)**. The Pixel Detector is situated the closest to the interaction point and has the highest granularity. There are about 80 million readout channels in the entire Pixel Detector. The intrinsic spatial resolution of individual Pixel Detector modules is $10 \mu\text{m}$ in $r - \phi$ and $115 \mu\text{m}$ in z . The **SCT** is also a silicon detector with microstrips and surrounds the Pixel Detector. It provides eight measurements per track with resolution of $16 \mu\text{m}$ in $r - \phi$ and $580 \mu\text{m}$ in z . In the outermost region, the **TRT** provides a large number of tracking measurements (typically more than 30 hits per track), good pattern recognition, and contributes to particle identification. It is composed of gaseous proportional counters (70% Xe, 27% CO₂ and 3% O₂ straws) embedded in a radiator material and therefore is a light-weight detector. Its operational drift radius accuracy is about $130 \mu\text{m}$. **ATLAS** track reconstruction efficiency as a function of pseudorapidity for simulated data for a centre-of-mass energy of 7 TeV is shown in Figure 2.7(a) [42]. During **LS1** a fourth layer for the pixel detector, the **Insertable B-Layer (IBL)** was added between the original pixel detector and a new beam pipe with smaller diameter. In Figure 2.7(b) the transverse impact parameter resolution as a function of track momentum measured from data in 2015 at 13 TeV, with the Inner Detector including the **IBL**, for values of pseudorapidity between 0 and 0.2 is shown compared to the value measured from data in 2012 with 8 TeV. The data in 2015 is collected with a minimum bias trigger. The data in 2012 is derived from a mixture of jet, tau and missing E_T triggers [43].

Other LHC experiments

The **TOTal Elastic and diffractive cross-section Measurement (TOTEM)** experiment [44] is located on either side of the **CMS** experiment in the very forward region. It is designed to measure the total proton-proton cross-section and to understand better the proton structure by studying elastic scattering over a wide range in momentum transfer, and via the diffractive processes. It consists of

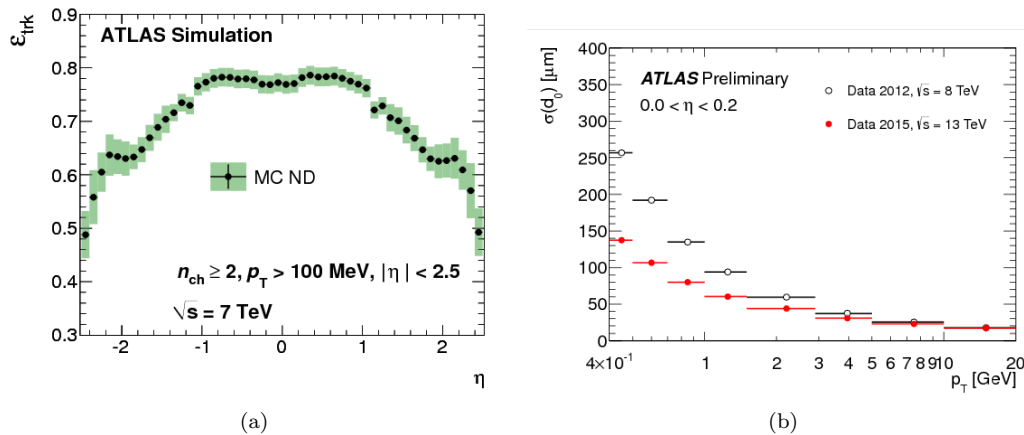


Figure 2.7: (a) The track reconstruction efficiency as a function of pseudorapidity derived from non-diffractive (ND) Monte Carlo [42].
 (b) Unfolded transverse impact parameter resolution measured from data in 2015 at 13 TeV with the Inner Detector including the **IBL**, as a function of track p_T for values of $0.0 < \eta < 0.2$, compared to that measured from data in 2012 at 8 TeV [43].

four detector pairs at different positions to cover the forward physics that escape the **CMS** detector. Similarly, **LHC forward (LHCf)** [45], is installed 140 m away from the interaction point on both sides of the **ATLAS** detector. It is intended to measure the forward production spectra of photons and neutral pions and the leading particle spectrum. These measurements are used to calibrate the Monte Carlo nuclear interaction model that is applied for several important measurements in cosmic ray physics such as the derivation of the cosmic ray energy spectra.

2.2 The CMS detector

The **Compact Muon Solenoid (CMS)** experiment is installed at one of the four interaction points in the **LHC** tunnel, near the French village of Cessy, between Lake Geneva and the Jura mountains. As the name itself suggests, the main distinguishing feature of **CMS** is a 3.8 T superconducting solenoid. With its 13 m length and its diameter of 6 m, it provides a high bending power to precisely measure the momentum of charged particles. Both the tracking system and most of the calorimeter systems are placed inside the magnetic coil. The solenoid magnetic field lines run parallel to the beam direction in the central region. The flux is closed by a return yoke made of a 12-sided iron structure, interspersed with four layers of muon stations. **CMS** itself has a diameter of 15 m and a length of 28.7 m. The different sub-detectors exploit the fact that different kinds of particles interact diversely and therefore can be distinguished based on the signals they leave in various detector components. A cutout view of the **CMS** detector is shown in **Figure 2.8**.

Since **CMS** is a general-purpose experiment, the detector is designed to cover the maximum possible solid angle around the collision point. However, for mechanical reasons and due to radiation requirements, the reconstruction of particles with very small angles with respect to the beam-line is not feasible in practice.

The reference frame consists of a right handed Cartesian system, which has the origin centered at the nominal collision point. **Figure 2.9(a)** shows the Cartesian system of reference: the x -axis pointing radially towards the center of the **LHC**, the y -axis pointing vertically upward, and the z -axis parallel to the beam direction such that the coordinate system is right-handed. The $x - y$ plane is also called the *transverse plane*. Given the cylindrical symmetry of the detector, the triplet (r, η, ϕ) can also be used as a coordinate system, where r is the radial distance from the z -axis, η is the so-called pseudorapidity and ϕ is the azimuthal angle in the transverse plane. The polar angle θ measured from the z -axis is not directly used since differences $\Delta\theta$ are not Lorentz-invariant. It is related to the pseudorapidity in the following way (**Figure 2.9(b)**):

$$\eta = -\log\left(\tan\frac{\theta}{2}\right) \quad (2.3)$$

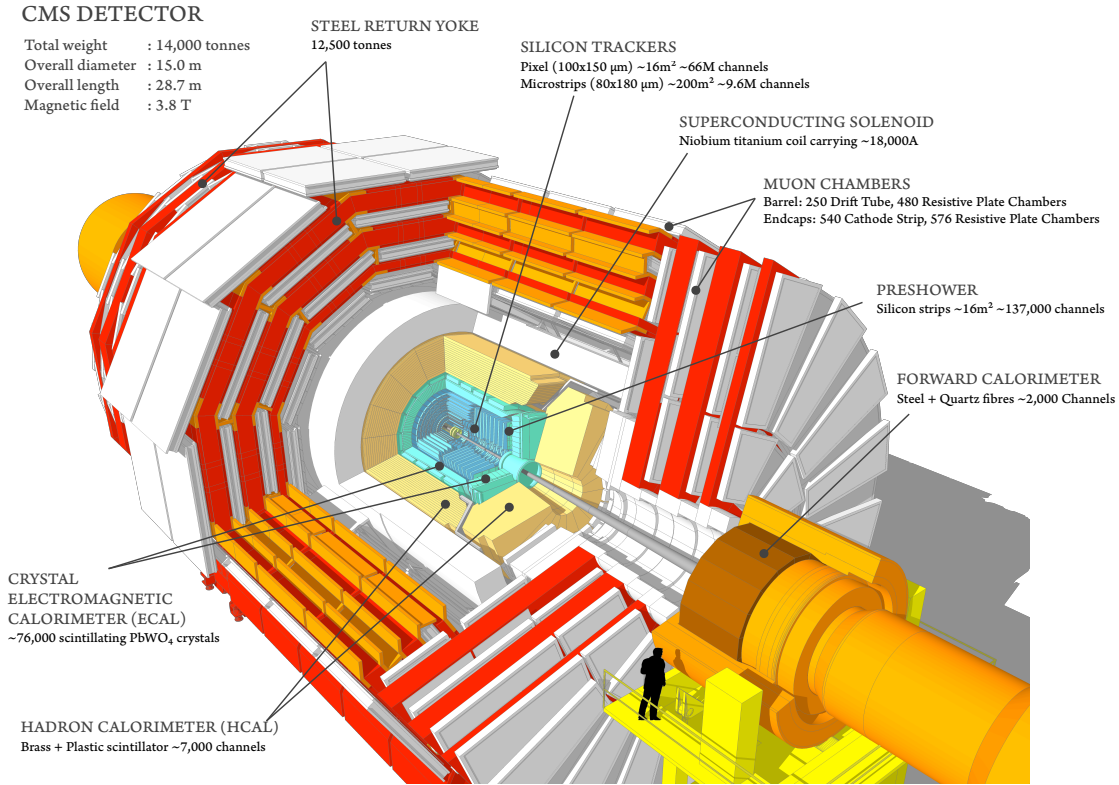


Figure 2.8: Perspective view of the CMS detector [46].

For high energy particles, this is a good approximation of the particle rapidity, y , defined as

$$y = \frac{1}{2} \log \left(\frac{E + p_z}{E - p_z} \right) \quad (2.4)$$

where E is the energy and p_z the longitudinal momentum of the particle. The pseudorapidity results from taking the rapidity to the limit of massless particles.

Apart from pseudorapidity, there are other useful variables used to describe particle properties which will be mentioned in this document. An often used quantity is the distance ΔR in the η - ϕ plane, which is defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (2.5)$$

Another very important quantity is the *transverse momentum* p_T calculated from the components of the momentum vector in the transverse plane as

$$p_T = \sqrt{p_x^2 + p_y^2} \quad (2.6)$$

An analogous relation holds for the *transverse energy*, E_T . To account for invisible particles like neutrinos, another important variable is defined: the *missing transverse momentum*. If it is assumed in first approximation that all incoming protons have momentum parallel to the beam axis only, then momentum conservation requires the sum of the transverse momenta of all particles in the final state to be zero. Therefore if the sum of momenta in the transverse plane for the visible particles is not zero, then the negative of this sum ($-E_T^{\vec{miss}}$) is the missing transverse momentum with magnitude

$$E_T^{\text{miss}} = \sqrt{E_{x,\text{miss}}^2 + E_{y,\text{miss}}^2} \quad (2.7)$$

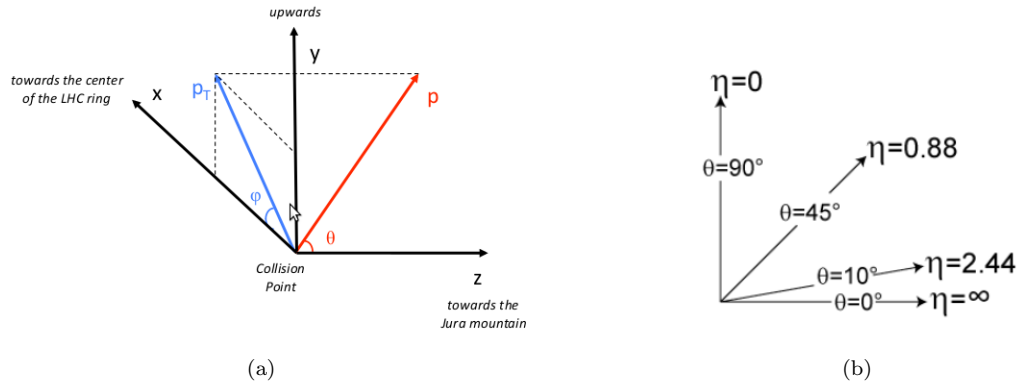


Figure 2.9: CMS coordinate system (a) and values for pseudorapidity in correspondence with various polar angles (b).

and is referred to as *missing transverse energy*. For example, the transverse mass m_T of a τ lepton decaying into a charged lepton l and a neutrino ν can be computed as

$$m_T = \sqrt{2p_T^l E_T^{miss} (1 - \cos \Delta\phi(l, E_T^{miss}))} \quad (2.8)$$

using the missing transverse energy and the transverse momentum of the lepton p_T as well as the angle between the two.

2.2.1 The tracking system

Precise particle momentum reconstruction is crucial to study events produced in proton-proton collisions. The trajectory of a particle with an electrical charge $e = \pm 1$ in a magnetic field B is bent by the Lorentz force. In a homogeneous magnetic field, the particle moves on a path described by a helix. Its transverse momentum p_T can be determined (non-relativistically) by:

$$p_T = e \int B dL \simeq eBR \quad (2.9)$$

with R being the radius of the curved track projected to the transverse plane. This phenomenon is described in detail in [Section 3.1.1](#).

The CMS tracking system [47] is designed to provide a precise and efficient measurement of particle trajectories using position-sensitive detectors like silicon pixel (two-dimensional) and silicon strip (one-dimensional) modules located in several layers around the interaction point. Through the trajectory reconstruction, the particle momentum and its projection on the transverse plane can be calculated as well as the charge sign.

Due to the challenging environment in LHC, the main requirements on the tracking system are:

- high granularity and fast response, given the high track multiplicity produced in each bunch crossing;
- high radiation hardness to face the intense particle flux in the region close to the proton beam;
- minimum material budget in order to be the least invasive as possible and therefore to limit processes that can mislead the trajectory reconstruction, such as multiple scattering, bremsstrahlung, photon conversion and nuclear interactions.

A silicon-based tracking system fulfills all these requirements.

The CMS tracking system is split into two parts, the Pixel Tracker and the Strip Tracker, and covers a pseudorapidity range up to $|\eta| = 2.5$. The modules assembly is shown in [Figure 2.10](#).

The material budget for different part of the tracking system is shown in Figure 2.12(a) in terms of radiation length, X_0 , which is described in Section 2.2.2. The position resolution of the Pixel and the Strip Tracker has an upper limit of $\frac{p}{\sqrt{12}}$, where p is the pitch of the sensor crossed by the particle. Currently, the Strip Tracker nominal temperature is -20°C , while the Pixel Tracker one is -25°C , in order to minimize degradation due to irradiation effects.

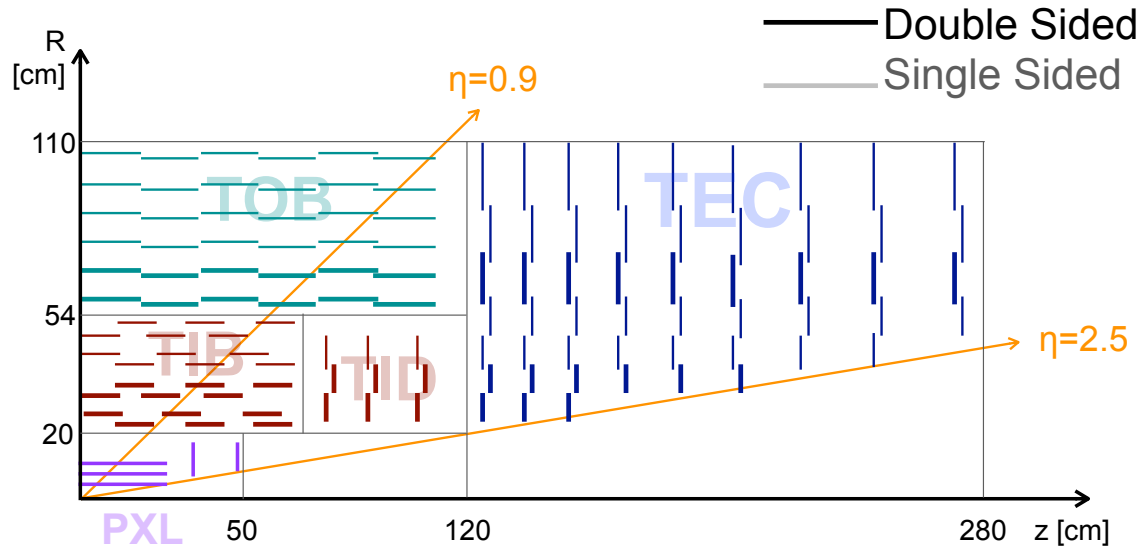


Figure 2.10: Schematic section through the CMS tracker formed by Pixel Tracker (PXL) and silicon Strip Tracker (TIB, TID, TEC, TID). Each short line represents a silicon module. Modules with double-sided strip sensors are represented in bold lines in the Strip Tracker.

Pixel Tracker

The Pixel Tracker is the innermost CMS detector sub-system. It is located at a radial distance less than 10 cm from the beam-line, where the rate of particles is about 10 million particles per $\text{cm}^{-2}\text{s}^{-1}$. It is composed of 66 million silicon pixels with dimensions $100\text{ }\mu\text{m} \times 250\text{ }\mu\text{m} \times 250\text{ }\mu\text{m}$, covering a total area of about 1 m^2 . The modules are arranged in three barrel layers (PXB) at mean radii of $\simeq 4.4\text{ cm}$, $\simeq 7.3\text{ cm}$ and $\simeq 10.2\text{ cm}$, and in four disks (PXF) at $z = \pm 34.5$ and $z = \pm 46.5\text{ cm}$. The devices are made of n^+ pixels on n type oxygenated silicon substrate with a resistivity of approximately $2\text{ k}\Omega\text{ cm}$. The electronic silicon chip is attached, using a microscopic spot of solder using the so-called bump bonding technique, which amplifies the signal. A schematic view of the sensors is shown in Figure 2.11(a).

In the barrel layers the magnetic field induces a Lorentz angle which increases charge sharing between neighbouring pixels. Charge sharing in conjunction with analog readout allows to achieve $10\text{ }\mu\text{m}$ position resolution for the (r, ϕ) coordinate and $15\text{ }\mu\text{m}$ in the z direction. The pixel detectors in the forward direction are tilted at an angle of 20° to induce charge sharing which allows to achieve $15\text{ }\mu\text{m}$ and $20\text{ }\mu\text{m}$ resolution respectively. This resolution is not only necessary for a precise track reconstruction, but also for the determination of both the vertices produced in the primary interaction and the decay vertices of short lived particles. An example of the achieved resolution and the total efficiency for the Pixel Tracker is shown in Figure 2.12(c,d).

Strip Tracker

The Strip Tracker constitutes the outer layers of the tracking system. The basic building blocks of the CMS Strip Tracker are silicon strip modules. Each module is equipped with one or two silicon sensors and a so-called Front-End (FE) hybrid containing readout electronics. The modules are arranged in ten layers in the barrel region extending up to a radius of 1.1 m. It is separated in an inner and outer part, called Tracker Inner Barrel (TIB) and Tracker Outer Barrel (TOB), respectively. Within one layer, there is a sub-structure of modules grouped together called strings

(**TIB**) or rods (**TOB**). In the endcap region, the silicon Strip Tracker consists of two blocks of disks, three disks belonging to the **Tracker Inner Disks (TID)** and nine disks to the **Tracker End Caps (TEC)**. In one disk of the **TEC**, 16 sub-structures of modules are installed, the so-called petals. Examples of Strip Tracker modules used in the **TEC** is shown in **Figure 2.11(b)**. In total, the **CMS** silicon strip tracker has 9.3 million strips and covers 198 m^2 of active silicon area. The expected resolutions grow to $\simeq 30 \mu\text{m}$ in (r, ϕ) and $\simeq 300 \mu\text{m}$ along the z coordinate. The modules in the innermost two layers of both the **TIB** and the **TOB**, as well as the modules in rings 1 and 2 of the **TID**, and 1, 2 and 5 of the **TEC** (**Figure 2.10**) are made of double-sided strip sensors, one of which is rotated by a stereo angle of 100 mrad , achieving a resolution along the z coordinate of about $230 \mu\text{m}$ and allowing the reconstruction of the hit position in 3-D. An example of the signal over noise ratio for the **TIB** detector is shown in **Figure 2.12(b)**.

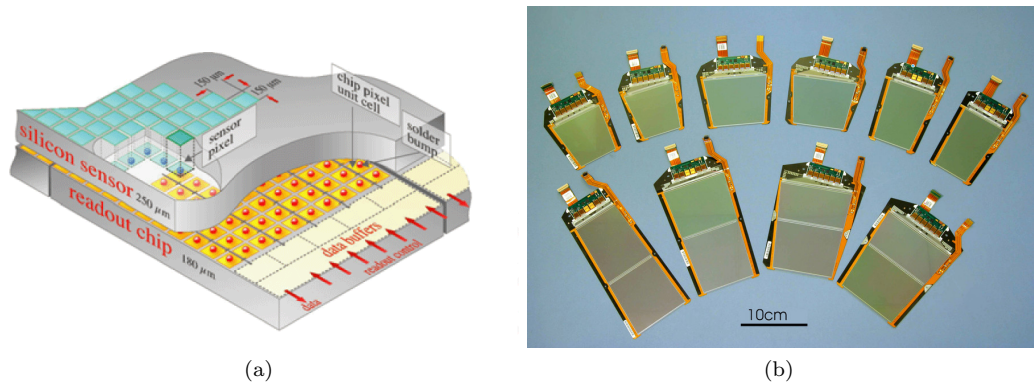


Figure 2.11: (a) Schematic view of the pixel sensor and the bump-bonded readout chip [47].
(b) Examples of silicon Strip Tracker modules used in the **TEC** [48].

Upgrade of the Pixel Tracker

As mentioned in **Section 2.1.2**, the original design goal of the **LHC** was to operate at $1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ with 25 ns bunch spacing. The pixel readout electronics were designed and optimized for the data rates and occupancy expected up to the **LHC** design luminosity. For these beam parameters, the dynamic inefficiency is about 4% at the nominal **CMS** read-out rate. With the current **LHC** operational parameters, the luminosity and pileup are more than double with respect to the nominal ones and therefore this data loss in the inner layer increased to 16%. In order to maintain the excellent performance, the Pixel Tracker was replaced in the year-end technical stop of 2016/2017. The removal of the old Pixel Tracker and the re-installation of the new one has been done a record time of only few months, also thanks to the **CMS** design which allows fairly easy access to the central detector systems.

The new Pixel Tracker is a new high efficiency and low mass detector. It is composed of four barrel layers and three forward disks providing four-hit pixel coverage up to $|\eta| = 2.5$. A schematic view of the new Pixel Tracker geometry is displayed in **Figure 2.13** [50]. The upgraded system is foreseen to withstand the luminosity conditions of **LHC** until **LS3**, providing excellent tracking performance at pileup value up to and exceeding 50. Tracking algorithm and performance of the new Pixel Detector as well as the entire **CMS** tracking system is discussed extensively in **Chapter 3**.

2.2.2 The electromagnetic calorimeter

The **CMS Electromagnetic Calorimeter (ECAL)** [51] is located between the tracking system and the hadron calorimeter, inside the superconducting solenoid. Its main task is to allow the reconstruction of electrons and photons. In **Section 1.2.2** some of the most important Higgs decay channels are listed. Among them, there are $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^{(*)} \rightarrow l^+l^-l^+l^-$ and $H \rightarrow WW^{(*)} \rightarrow l\nu_l l\nu_l$, which all need to reconstruct electrons or photons in their final state. In particular, for $m_H \simeq 125 \text{ GeV}$, an

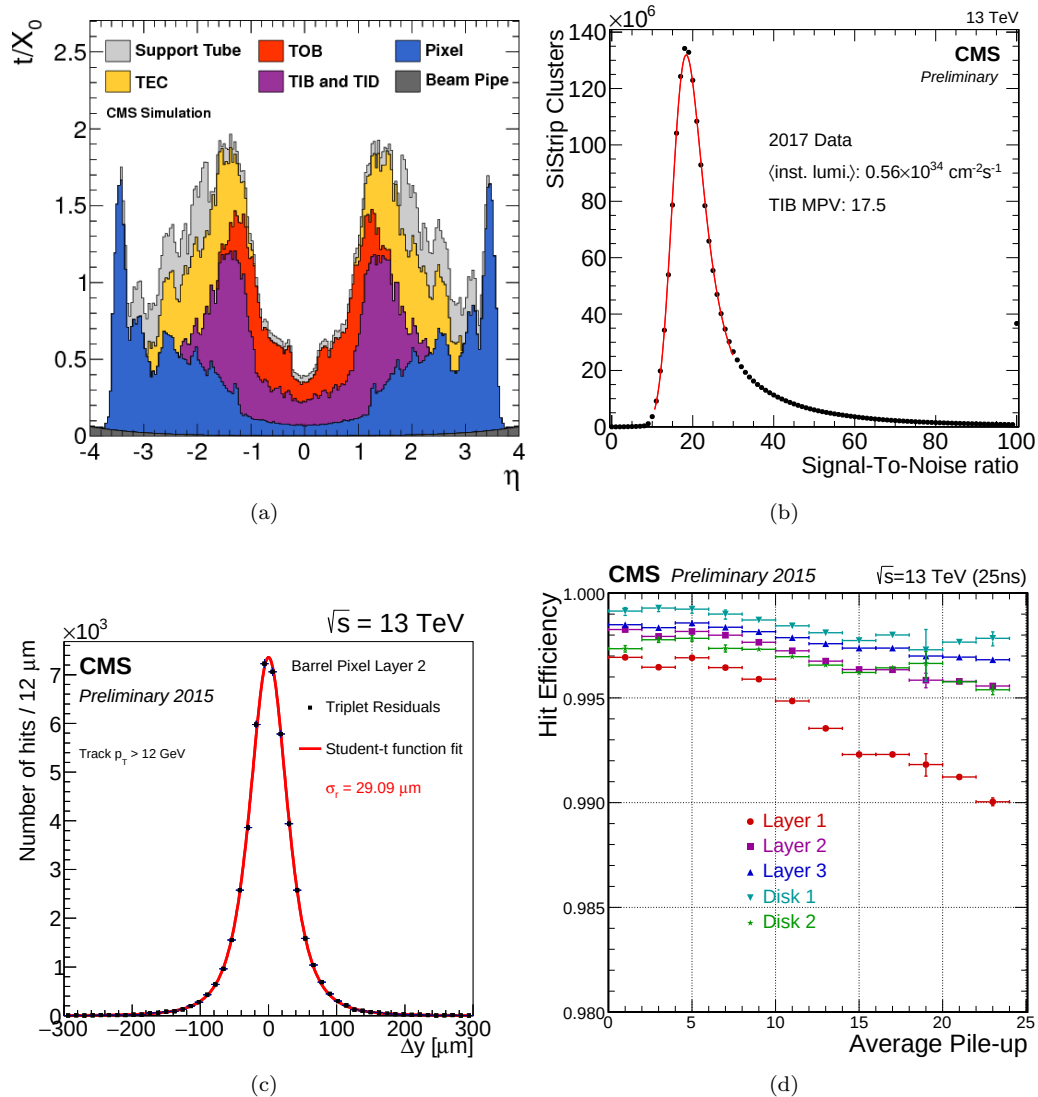


Figure 2.12: (a) Material budget of the CMS tracker in units of radiation length (X_0) as a function of η for the different sub-detectors [49]. (b) Example of signal over noise ratio for the TIB detector planes [49]. (c) Distribution of hit residuals on Pixel barrel layer 2 in the longitudinal direction (parallel to the beam) [49]. The distribution is fitted with a student's t-function for which sigma is shown on the plot. (d) Hit efficiency as a function of the average number of inelastic proton-proton collisions for the barrel layers and the forward disks of the Pixel Tracker [49].

excellent energy resolution is necessary to fully profit from the narrow peak of the $H \rightarrow \gamma\gamma$ decay. Additional requirements are a good angular resolution and an accurate pion/photon separation. The CMS choice was a hermetic, homogeneous scintillating detector.

High energy electrons and photons interact in the calorimeter producing a cascade of lower energy electrons and photons, called an *electromagnetic shower*. The process develops due to the subsequent pair production by photons as well as bremsstrahlung by electrons, and it stops when the energy of secondary particles falls below the *critical energy*. The longitudinal development of the shower is governed by the radiation length of the material, X_0 , which is defined as the average distance traveled by an electron before losing $1 - \frac{1}{e}$ of its initial energy, where e is the Euler's number. The transverse development of the shower is described by the Molière radius, r_M , related to X_0 . On average, 95% of showers are contained inside a cylinder with a radius equal to two r_M .

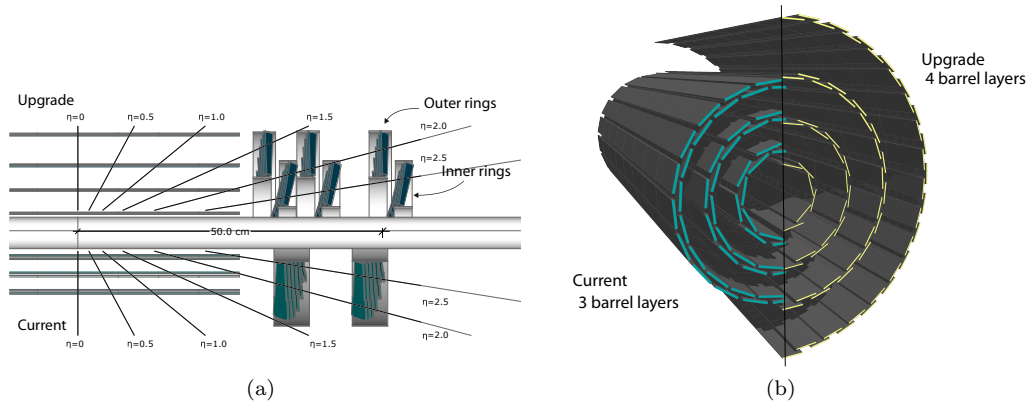


Figure 2.13: $r-z$ view layout comparison between the old Pixel Tracker and the new one (a) and transverse-oblique view comparing only the barrel layers (b) [50].

ECAL is constituted of 75848 lead tungstate (PbWO_4) crystals organized in structures of modules, 61200 in the **Electromagnetic calorimeter Barrel (EB)** and 7324 crystals in the **Electromagnetic calorimeter Endcap (EE)**. A schematic view of the **ECAL** design is shown in Figure 2.14. The **EB** has an inner radius of 1.29 m. Each crystal in **EB** has a truncated pyramid shape with a depth of 230 mm and it exposes a front face cross section of about $22 \times 22 \text{ mm}^2$, $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174^\circ$. Modules in **EB** are structured in so-called *supermodules*. There are in total 36 supermodules, 18 for each side of the interaction point. The supermodules are organized in a way so that crystal axes form a 3° angle with respect to a straight line from the interaction point, both in η and in ϕ directions. This set-up is used in order to avoid holes in the geometrical acceptance. The two **EE** are placed at the distance of 3.17 m from the interaction point. They cover the pseudorapidity range of $1.479 < |\eta| < 3.0$. The 7324 endcap crystals have different dimension with respect to the barrel ones: each of them has a front face cross section of $28.62 \times 28.62 \text{ mm}^2$ and a length of about 22 cm.

The choice of lead tungstate as the active detector material is dictated by some particular features: the short radiation length ($X_0 = 0.89 \text{ cm}$) and a small Molière radius ($r_M = 21.9 \text{ mm}$), the high resistance to radiation and the fast response, with 80% of the light emitted within the time of a **LHC** bunch crossing. The result is a compact calorimeter that can be housed inside the magnet and collects most of the electromagnetic shower in few crystals. Due to the relatively low light yield ($\simeq 30 \gamma/\text{MeV}$), an amplification stage with high-gain photodetectors is required both in the barrel and in the endcaps. In the barrel, Avalanche PhotoDiodes are glued in pairs on the rear face of each crystal; they are operated at gain 50 and are insensitive to the magnetic field. Instead, Vacuum PhotoTriodes are used in endcap regions: Despite their lower gain, they are more resistant to radiation damage. The light yield is highly dependent on the operating temperature, requiring constant cooling to 18°C . The radiation damage to the PbWO_4 crystals leads to a wavelength-dependent loss of light transmission, however without changes to the scintillation mechanism. This effect can thus be measured and corrected for by monitoring the optical transparency with injected laser light.

In addition to **EB** and **EE** system, a preshower detector (**ES**) is located just before the endcaps to help identify neutral pions that decay into two photons which can mimic a high-energy photon. It covers the pseudorapidity region $1.653 < |\eta| < 2.6$ and has a total material thickness of about $3X_0$. **ES** is made of two planes of lead interspersed between 2 mm wide silicon strip sensors.

The energy resolution σ_E , as measured in electron beams test, can be parameterized as follows [52]:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E(\text{GeV})}} \oplus \frac{12\%}{E(\text{GeV})} \oplus 0.3\% \quad (2.10)$$

where the three different contributions correspond to a stochastic, noise and constant term, respectively. For electrons coming from the Z boson decay with E_T of about 45 GeV, the energy

resolution is better than 2% in the central region of the ECAL barrel ($|\eta| < 0.8$), and is between 2% and 5% elsewhere.

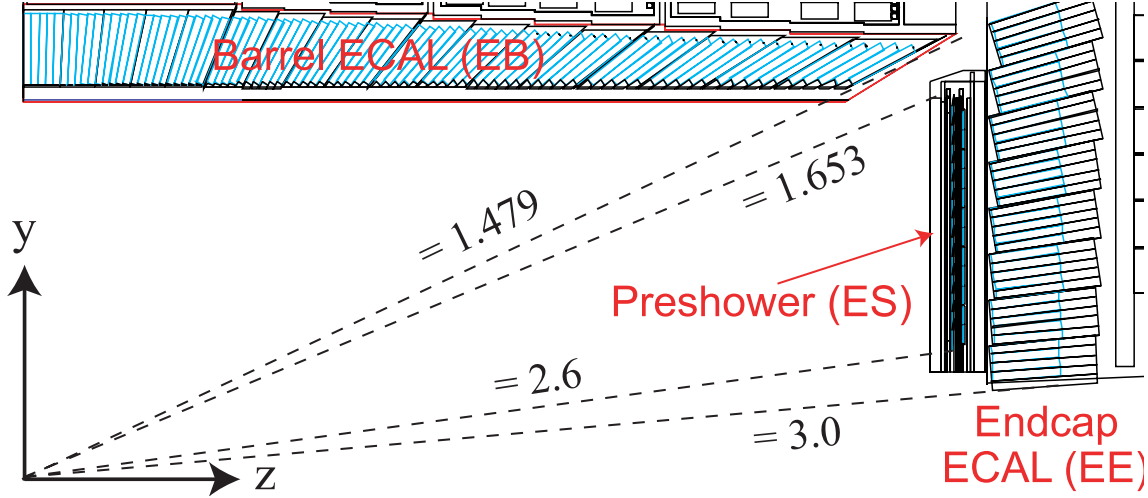


Figure 2.14: View in (y, z) of a quarter of the CMS electromagnetic calorimeter.

2.2.3 The hadron calorimeter

The **Hadron Calorimeter (HCAL)** [53] measures the energy of hadronic jets and is particularly important for detection of neutrinos or exotic particles that are manifest in the detector as an apparent missing transverse energy. It has a large hermetic coverage up to $\eta < 5$ as well as fine lateral segmentation. A schematic view of the calorimeter location within **CMS** is given in **Figure 2.15**.

HCAL is a sampling calorimeter, meaning that the jet position, energy and arrival time is reconstructed using alternating layers of absorber material, in the case of **CMS** brass or steel, and a fluorescent scintillator. When hadrons interact with the absorber material, they produce a significant amount of secondary particles. These secondary particles interact with the successive layers of absorber and create a cascade of particles, a so-called *hadronic shower*. As the shower develops, the particles constituting it pass through scintillator material. Due to this interaction, the scintillator emits blue-violet light, which is collected and carried by tiny optical **Wavelength-shifting Fibers (WLF)**. **WLF** shift the blue-violet light into green, which is detected with the use of readout boxes placed in several locations within the **HCAL** volume. The layers of the **HCAL** are placed in a staggered fashion so that there are no gaps in direct lines that a particle might escape through.

HCAL is organized into barrel (**Hadronic calorimeter Barrel (HB)** and **Hadronic calorimeter Outer (HO)**), endcap (**Hadronic calorimeter Endcap (HE)**) and forward (**Hadronic calorimeter Forward (HF)**) sections. The hadron calorimeter barrel is restricted between the outer radius of the electromagnetic calorimeter ($r = 1.77$ m) and the inner radius of the magnet coil ($r = 2.95$ m). There are 36 barrel wedges, composed of flat brass alloy absorber plates parallel to the beam axis. Each of them has a transverse granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087^\circ$ and weighs 26 tons. The innermost and outermost absorber layers are made of stainless steel for structural reasons. There are 17 active plastic scintillator tiles interspersed between the stainless steel and brass absorber plates. Similarly, 36 endcap wedges in **HE** measure particle energies with a granularity of $\Delta\eta \times \Delta\phi = 0.17 \times 0.17^\circ$.

The total absorber thickness at $\eta = 0$ is 5.82 interaction lengths, λ_I . To improve the energy resolution and to extend the coverage, the **HO** and the **HF** were added. The **HO** is placed outside the solenoid and complements the barrel calorimetry, ensuring no energy leaks out the back of the **HB** undetected in the region $|\eta| < 1.26$ and increasing the interaction length to $\simeq 10.6\lambda_I$ in total. The **HF** is placed 11.2 m from the interaction point and covers $3 < \eta < 5$. It is designed to survive

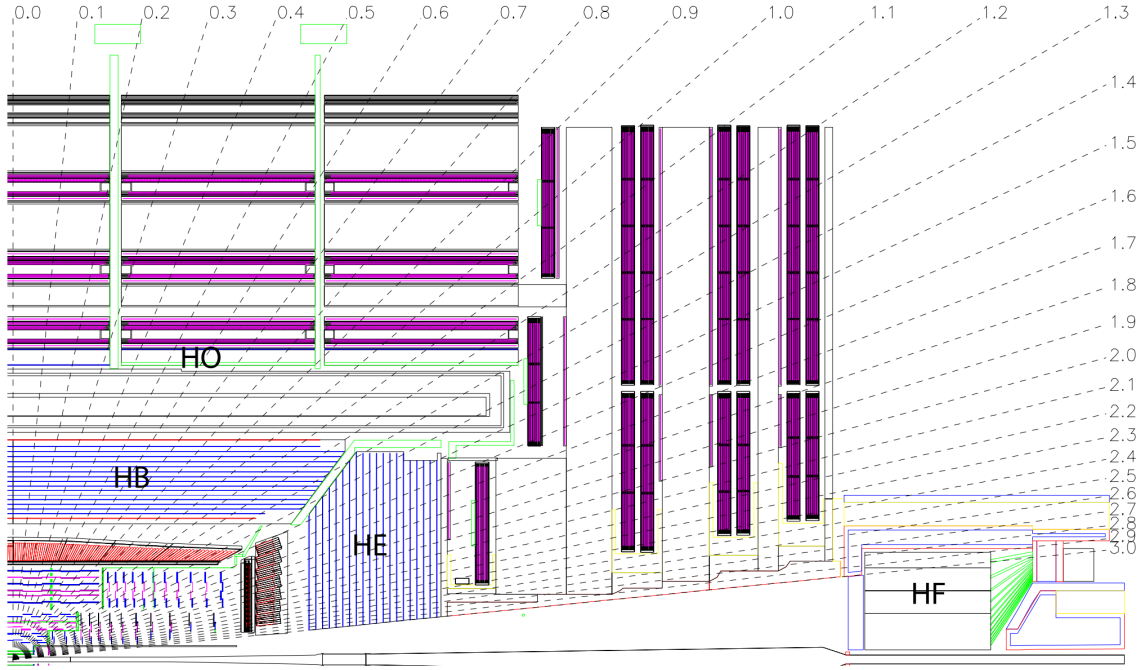


Figure 2.15: Longitudinal view in (r, z) of the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters.

for at least a decade in the hostile environment with very high particle fluxes. It consists of a steel absorber structure composed of 5 mm thick grooved plates interspersed between quartz fibers which constitute the calorimeter active medium. These fibers detect energy emitted by particles via Cherenkov radiation and, as a consequence, HF is practically insensitive to neutral hadrons, but very sensitive to the electromagnetic component of showers. Its granularity is $\Delta\eta \times \Delta\phi = 0.175 \times 0.175^\circ$.

The expected energy resolution for single pions interacting in the central part of the calorimeter according to test-beam results is

$$\frac{\sigma_E}{E} = \frac{94\%}{\sqrt{E}} \oplus 4.5\% \quad (2.11)$$

where the energy is measured in GeV. A significant degradation of the resolution is expected at $|\eta| = 1.4$, due to the presence of services and cables. The expected energy resolution for the very forward calorimeter is worse,

$$\frac{\sigma_E}{E_{\text{had}}} = \frac{172\%}{\sqrt{E_{\text{had}}}} \oplus 9\% \quad , \quad \frac{\sigma_E}{E_{\text{em}}} = \frac{100\%}{\sqrt{E_{\text{em}}}} \oplus 5\%. \quad (2.12)$$

2.2.4 The superconducting solenoid

The superconducting solenoid of the CMS experiment provides a magnetic field of 3.8 T at the interaction point. Its field lines are parallel to the beam-line. It measures 6 m in diameter and 12.5 m in length and weighs 220 tonnes in total. A distinctive feature is the four-layer winding made from a stabilised, reinforced NbTi conductor that provides an energy density of 11.6 kJ/kg, which can lead a deformation of up to 0.15% during energizing.

2.2.5 The muon system

The muon system [54] is of central importance to the CMS physics goals, for example for the $H \rightarrow ZZ^{(*)} \rightarrow l^+l^-l^+l^-$ decay channel, as already mentioned in Chapter 1. Among the main

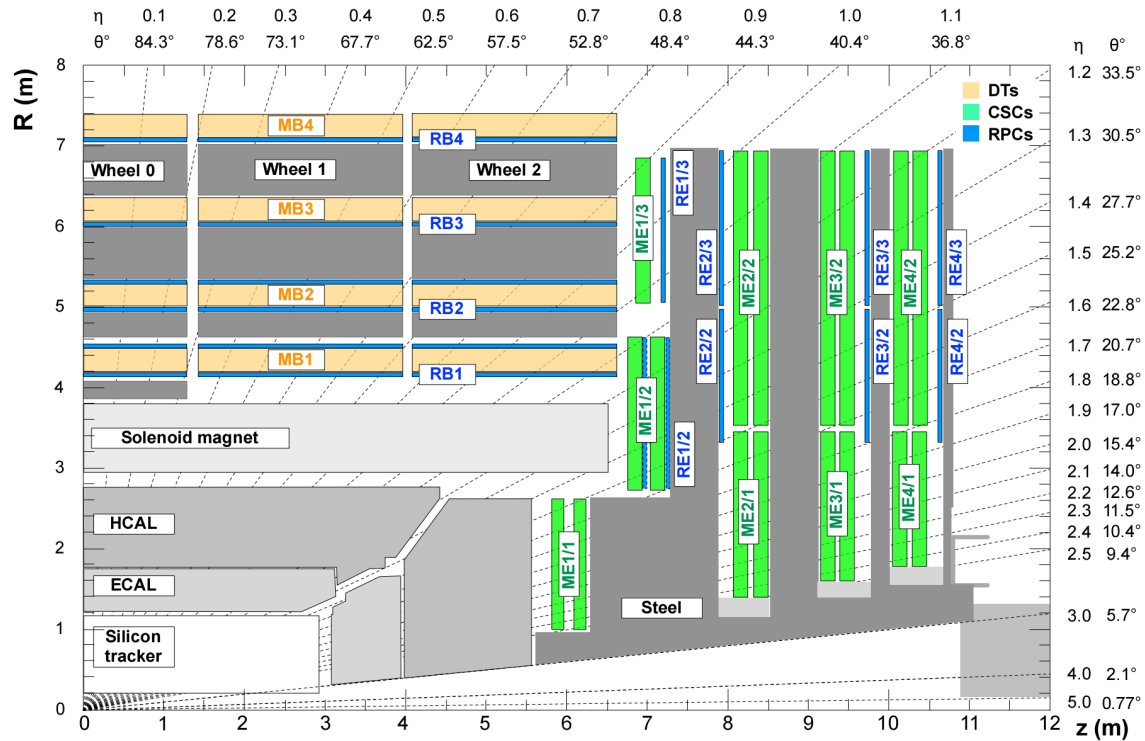


Figure 2.16: View in (r, z) of a quarter of the CMS muon system layout after LS1 [55].

design goals for the CMS muon system are: a robust muon identification, a precise momentum resolution, a good double muon mass resolution and the possibility to determine the charge of muons with p_T up to 1 TeV.

In order to fulfill these requirements, the muon system is located inside the magnet return yoke and is composed of three types of gaseous detectors to account for different radiation environments: Drift Tube (DT) chambers in the barrel and Cathode Strip Chambers (CSC) in the endcaps, both complemented by a Resistive Plate Chambers (RPC) system (Figure 2.16). The muon system consists of four layers of muon chambers in the barrel part and four in each endcap region, each one providing track segments reconstructed from distributed hits. These tracks are combined with the information coming from the tracker to form a complete muon track. This geometrical scheme permits to cover a pseudorapidity range up to 2.4. The high return yoke field (from $\simeq 1.8$ T in the barrel region up to $\simeq 2.5$ T in the endcaps) guarantees a good momentum resolution and charge measurement, even without the inner tracker information. In this way, the information of the muon system is also helpful in the L1 trigger selection.

Due to the neutron-induced background, the low muon rate and the uniform magnetic field, the DT system is used in the barrel region within $|\eta| < 1.2$. DTs are long aluminum cells operated with an Ar/CO₂ mixture at atmospheric pressure in a percentage of 85% and 15%, and with an anode wire in the center collecting ionization charges. Each DT is segmented into 5 wheels along the z -direction (YB/+2,+1,0,-1,-2), each subdivided into 12 sectors. A sector is shown in Figure 2.17, where four DTs (MB1, MB2, MB3 and MB4) are interspersed between the steel return yoke (YB/z/1,2,3). The first three stations contain eight chambers, in two groups of four each. They are separated as much as possible to achieve the best angular resolution. The two groups are used to measure the muon coordinate in the $r-\phi$ bending plane and in the z direction. The fourth station does not contain z -measuring planes. Each station is designed to measure muon positions with a spatial resolution of about 150 μm in z and 1 mrad in ϕ .

A high number of background events, a high muon flux and a high non-uniform magnetic field are the factors that determined the choice of CSC as muon detectors in the two end-caps. They cover the range of pseudorapidity between 0.9 and 2.4. These chambers are contained in disks (ME1,

ME2, ME3 and ME4) and overlap in ϕ to avoid holes in the geometrical acceptance. The **CSCs** are trapezoidal multi-wire proportional chambers in which six anodic planes are interspersed between seven cathode planes, which are segmented into strips running across wires. The spatial dimensions of these detectors vary throughout the system from 1.7 m to 3.4 m in length and subtend an angle from 10° to 20° . The chambers use an $\text{Ar}/\text{CO}_2/\text{CF}_4$ gas mixture in a percentage of 40%, 50% and 10%, respectively. A charged particle passing through the gas mixture generates an avalanche of electrons on a wire which then produce an electric signal. Each single hit has a resolution of about $100\ \mu\text{m}$ in r and $10\ \text{mrad}$ in ϕ . During **LS1** the **CSC** system underwent a significant upgrade [55, 56]: the outer ring of the ME4 was installed together with an additional layer of yoke disks that shield the chambers from environmental noise. This additional layer increases redundancy in the pseudorapidity region of $1.2 < |\eta| < 1.8$ improving considerably the L1-trigger efficiency. A further upgrade was the replacement of the read-out electronics for station ME1/1, allowing the full detector granularity to be used.

The **RPC** system provides a complementary muon system in **CMS** up to $|\eta| < 1.6$. The space resolution of a single chamber is poor due to the large width of the strips in which each plane is segmented (from 2.3 to 4.1 cm), but it has excellent timing resolution, about 1 ns, which allows unambiguous bunch crossing identification essential for triggering purposes. In the barrel there are six layers of **RPCs** installed. Two layers in each of the two innermost stations and one layer in each of the two outer stations. In the endcap only a plate in every disk in the endcaps was installed until Run 2. During the upgrade in **LS1**, an additional layer of **RPC** detectors was added in the endcap, significantly improving reconstruction efficiency as well as background rejection in the **L1** trigger [55, 56]. **RPCs** consist of double-gap Bakelite chambers, with 2 mm spacings filled with a mixture of $\text{C}_2\text{H}_2\text{F}_4/\text{C}_4\text{H}_{10}/\text{SF}_6$, with percentage of 96.2%, 3.5% and 0.3%, respectively.

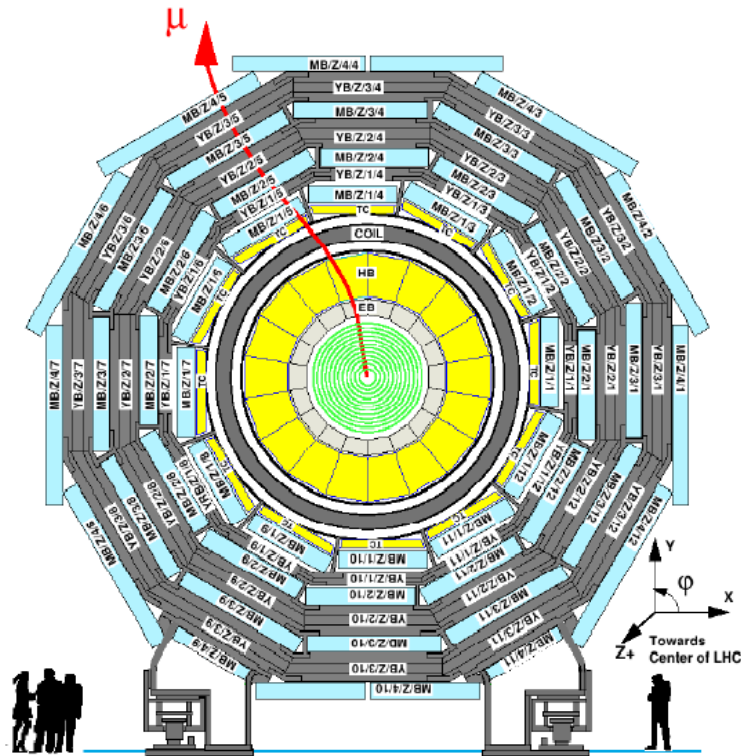


Figure 2.17: Layout of the CMS barrel muon **DT** chambers in one of the 5 wheels.

2.2.6 Trigger and data acquisition

As already described, most of the events emerging from the interaction point are created in soft collisions, which are not interesting for most of the CMS physics program. These events are so-called *minimum bias events*. Figure 2.18 shows the cross-sections for different processes studied at the LHC: The cross-section of interesting events, such as the production of the Higgs boson, is many orders of magnitude smaller than the cross-section of the minimum bias events, for example the production of $b\bar{b}$ pairs.

Around 1 MB of data is produced by CMS per event in nominal pileup condition. If every event is read out at the rate of 40 MHz, the total amount of data rate would be 40 TB/s, which cannot be either stored or analyzed. Furthermore, the detector FEs only support a read-out rate of 100 kHz at maximum for the full event information. Therefore, an online selection system in CMS is introduced, called trigger, with the purpose of selecting the potentially interesting events providing a large rate reduction factor and keeping at the same time high efficiency.

In CMS, the trigger process [57] is divided into two steps: a **Level-1 (L1)** trigger, which consists of a hardware system with largely programmable electronics, and a **High-Level Trigger (HLT)**, which is a software system implemented in a farm of about one thousand commercial processors. The L1 trigger, built from specialized, low latency hardware, was designed to analyze each 25 ns bunch crossing and take decisions in no more than 4 μ s. During the decision time, events are stored in pipelines of processing elements. The L1 trigger is organized into a calorimeter and a muon trigger, which reconstruct full objects such as muons or jets and provide them to the global trigger which takes the accept-reject decision. If the L1 accepts the event, data are readout by the DAQ and then processed by the HLT. Due to restrictions in the detector FE electronics, bandwidth and computing power, the read-out rate has to be kept below 100 kHz. The challenging idea of the HLT is to further refine the L1 selections in order to save interesting events for the different analyses while reducing the total output rate about 1 kHz. Thus it is realized operating on the event data on a large computing farm with algorithms very similar to the ones used for offline analysis. Three virtual trigger levels are implemented: at the first level the information coming from muon system and calorimeters are used, in the second level these signals are unified with the tracker pixels ones and in the final level the full event is available.

2.2.7 Reconstruction software

The CMS software is based on a single, shared code base that supports multiple applications: online trigger system, reconstruction and simulation executables, data quality monitoring, analysis-based event skimming and user data analysis. The requirements of specific applications are very different from each other, but having a single code framework has the advantage to provide an efficient code structure with high standards and the possibility of sharing algorithms development among users. The code framework in CMS is referred to as **CMSSW** [59, 60]. Frequent software releases and a web-based version control repository facilitate the integration of new features and bug fixes. CMSSW uses Git as version control system to track changes in files and to coordinate work among multiple people. Git stores each change with respect to the code base in a so-called *commit* and provides *branches* and *tags* to organize them efficiently. In order to integrate code changes in the main development branch, so-called *pull requests* are used as a code review system. The open-source CMS software is hosted by the GitHub website [61].

Figure 2.19 summarizes the algorithmic flow used to reconstruct raw data. In a first step, signals are recorded in each single detector and a local reconstruction is performed. The local reconstructed objects, such as tracker hits or ECAL clusters, are used to form physics objects such as tracks, muons, jets and electrons. Then, high-level algorithms are used to estimate more complicated quantities such as missing transverse energy and vertices. The building block of the event data model is the *Event*, which holds all data that is taken during a triggered physics event. Auxiliary information, such as changes in the detector environment and status, is accessible via the *EventSetup*. Algorithms, which build physics objects from raw data, are formed by one or more *modules* in the CMSSW framework. Events are processed by passing the Event object through a sequence of modules specified directly by the users via a python interface. Typically, each component of the event reconstruction is implemented in tens of modules. All events collected

by the **CMS** trigger system are reconstructed by the **CMS** prompt reconstruction system soon after being collected. In a subsequent step, also improvements to software algorithms and updated detector calibrations are incorporated in the reconstruction.

In the development of the code, many technical requirements have to be considered, such as memory consumption and processing time per event. Moreover, the code must reproduce results consistently, and must meet financial constraints. **Figure 2.20(a)** shows the processing time per-system. Tracking and the electron and photon finding algorithms are the main offline contributors

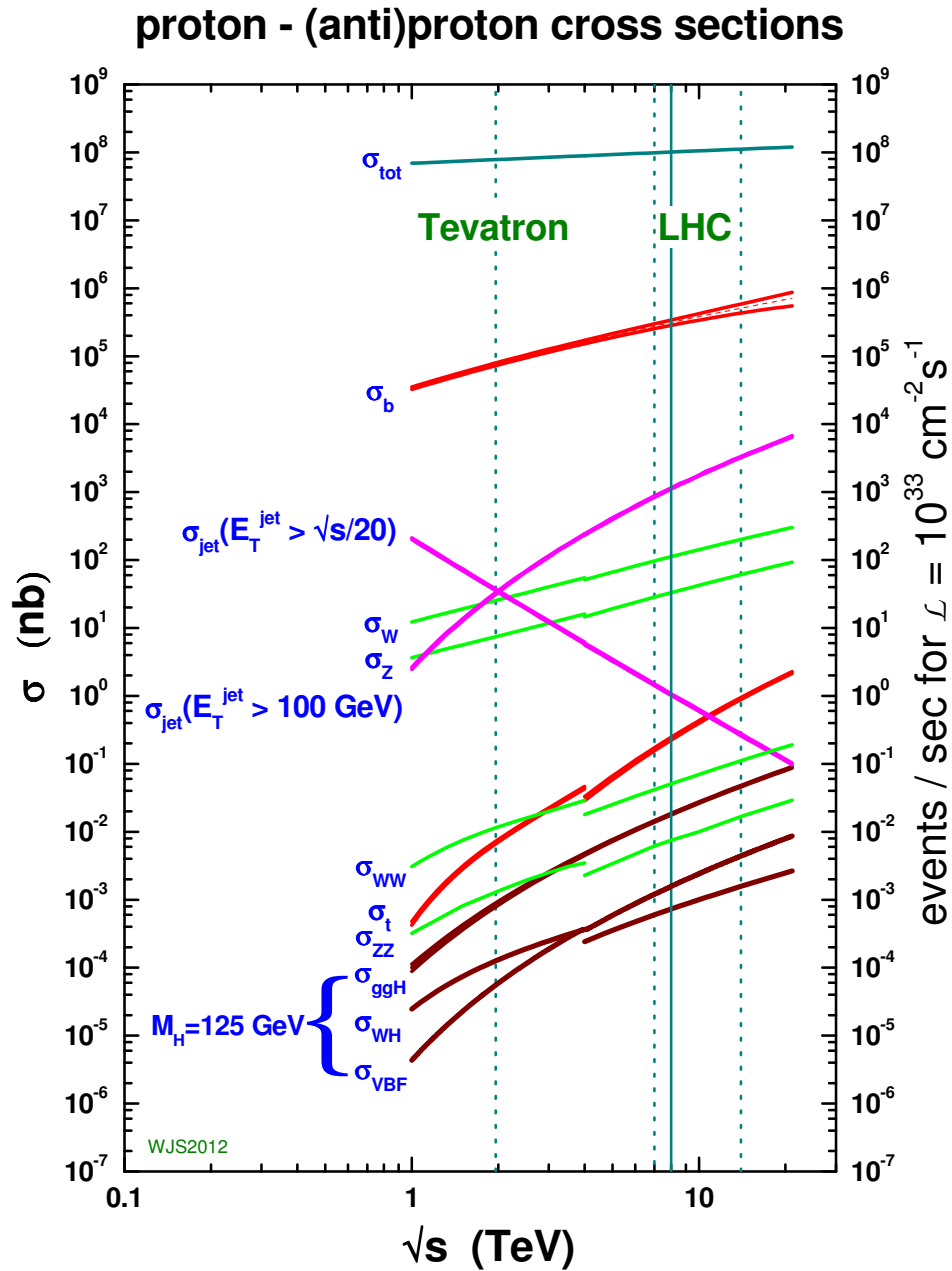


Figure 2.18: The cross-sections for processes studied at the **LHC** vary over orders of magnitude. The right hand side gives the interaction rates at nominal **LHC** luminosity [58].

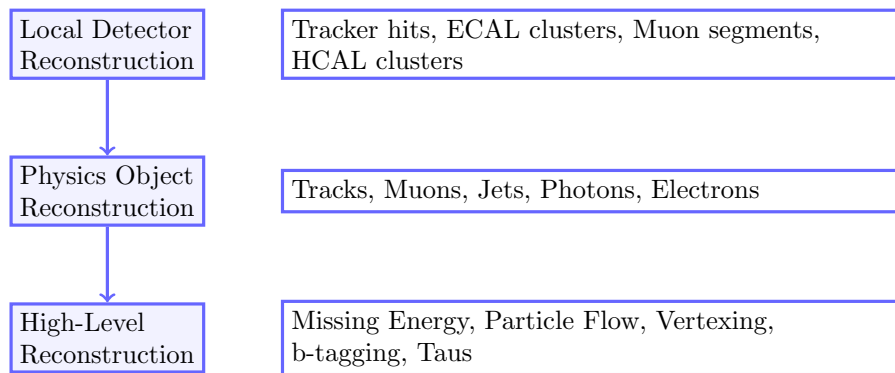


Figure 2.19: Schematic representation of the algorithmic flow of the **CMS** reconstruction software [60].

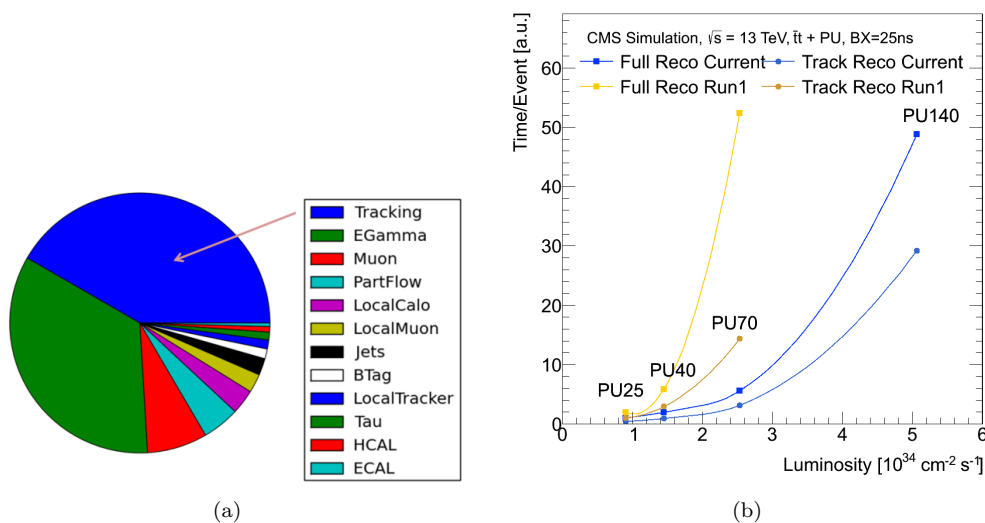


Figure 2.20: (a) Fractional breakdown of processing time into reconstruction components [60]. (b) Reconstruction time for a $t\bar{t}$ event as a function of instantaneous luminosity for two different versions of CMS reconstruction software. The time for track reconstruction only is broken out separately [63].

due to their complexity.

A significant effort was made during **LS1** in order to introduce technical innovations in **CMSSW** and enhance its reconstruction performance [62]. The biggest change was moving from a software with one-thread per event model to a multi-threaded one, both for the simulation and the real data reconstruction executable. This has the main advantage of reconstructing multiple events simultaneously in CPU-based technology. As a consequence, the demands on the computing infrastructure as well as the computing time process are significantly reduced. A result of this effort is shown in Figure 2.20(b) where the event reconstruction time is shown as a function of the pileup in the event. The sensitivity with respect to the pileup is reduced. The most significant improvements can be seen in the track reconstruction, which include not only technical code changes but also a new fully optimized algorithm. The tracking developments between Run-1 and Run-2 are fully described in Section 3.3.

2.3 Physics object reconstruction

As described in Section 2.2.7, the **CMS** code is designed to combine measurements coming from different sub-detectors producing physics object in a limited amount of CPU time and memory consumption. In this section the reconstruction of the main physics objects used in **CMS** analysis

is presented.

2.3.1 Muons

The **CMS** experiment is optimized to reconstruct muon particles with high efficiency. The tracking system and the muon chambers play a big role in muon reconstruction [59, 64]. The reconstruction algorithm starts with a local reconstruction step, where single hits in each muon sub-detector, **DTs** in the barrel and **CSCs** in the endcaps, are combined into three dimensional segments associated with a single muon layer. The information coming from different layers is then combined and two different categories of muon objects are built: Muon tracks reconstructed only in the muon spectrometer are called *standalone muons*, while if also the tracker information is taken into account, muons are referred to as *global muons* or *tracker muons*. The combination of these categories provide an efficient and robust identification of the final candidates collection.

In the case of the standalone muons, the segments associated with each single layer are used to seed the muon trajectory. From each seed, the **Kalman Filter (KF)** algorithm is applied to build the muon track from the innermost layer to the outermost surface of the muon system. Material effects are also taken into account. A detailed description of the **KF** algorithm can be found in [Section 3.1.2](#). Once the trajectory is built, a backward fit is performed to evaluate the muon track parameters. The inclusion of **RPC** measurements in this step helps significantly to reconstruct low p_T muons and muons escaping through the module gaps. The standalone muon track is then propagated to the nominal interaction region and the vertex-constrained fit is performed. Due to the large amount of material traversed by the muon particle, the momentum resolution as measured in the muon chambers can be degraded by multiple scattering.

Global muons are reconstructed starting from standalone muon tracks. The algorithm selects the tracks compatible with the tracker region and extrapolates the information into the silicon tracker. If a standalone track is found to be compatible with the silicon tracker one, the entire set of hits is merged and a final fit is performed providing the final muon trajectory at the interaction point. Global muon reconstruction is very efficient for muon tracks with p_T more than 6 – 7 GeV, provided that a minimal number of hits and segments are measured in the muon system.

A complementary approach is also used to reconstruct tracker muons. In this case, all the silicon tracker tracks are considered in the reconstruction process and the search of a muon signature is performed in calorimeters and muon systems. In this approach, tracks are reconstructed with an inside-out flow, starting from the subset of the tracker tracks considered possible muon candidates. They are then extrapolated to the muon system, taking into account possible multiple scattering and energy loss effects. Tracker muon reconstruction provides a higher efficiency for low p_T muons.

The relative p_T resolution as a function of the muon momentum is shown in [Figure 2.21](#). The additional information provided by the muon chambers improves the momentum resolution of high-energy muons, for which the tracker-only momentum measurement degrades. In the low momentum range, the resolution of the tracking system dominates.

2.3.2 Electrons

A high energy electron has two well-defined features: its trajectory can be reconstructed in the tracking system, and most of its energy is released in a small matrix of **ECAL** crystals around the impact point. Both parts of the reconstruction procedure can be complicated due to the fact that often the electron interacts with the tracker material emitting part or all of its energy via bremsstrahlung photons. For low energy electrons in particular, the magnetic field bends their trajectories causing a spread of radiated photons along the ϕ coordinate. Therefore, to obtain an accurate measurement of the electron energy and trajectory, it is essential to take into account the bremsstrahlung effect in both sub-detectors. Electron reconstruction in **CMS** is a three-step process where the information coming from both sub-detectors is used to maximize the reconstruction efficiency. At first, the energy deposited in **ECAL** is measured with a dedicated cluster algorithm. The position and energy measurement of these clusters, also called *ECAL-driven* seeds, is used to predict the position of the electron track in the innermost layer of the tracker. Another kind of seed collection is also built, the so-called *tracker-driven* seeds, which allow to extrapolate the trajectory information to the **ECAL**. The seeds obtained by both approaches are then combined in

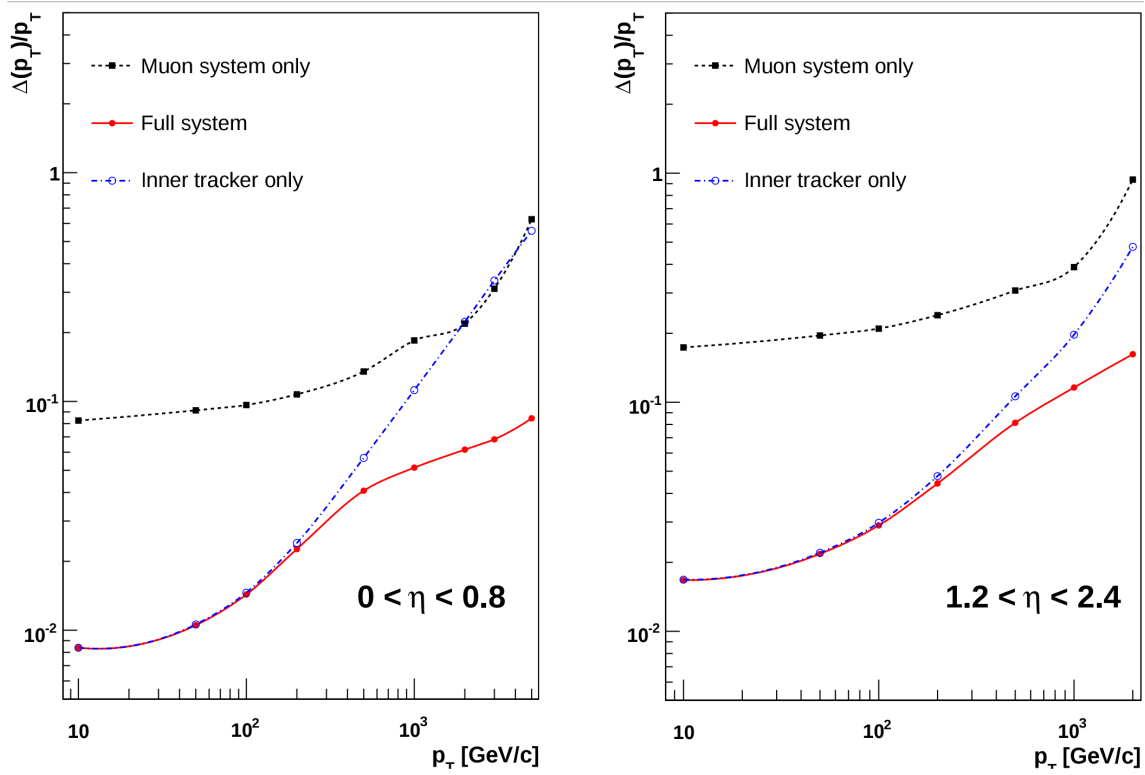


Figure 2.21: Muon momentum resolution versus p_T using the muon system only (black), the inner tracker only (blue) or the full system (red) for $|\eta| < 0.2$ (left) and $1.2 < |\eta| < 2.4$ (right) [37].

a unique collection. A tracking step is run on this last collection and the trajectory is built. In the last step a preselection is applied to classify electron objects based on track-cluster compatibility criteria.

In Run 1, the measurement of the energy deposit in **ECAL** was done using the *standalone* approach, which consisted of two clustering algorithms depending on the pseudorapidity of the electron candidate [65]. In the barrel, the Hybrid algorithm was used. It exploited the arrangement in $\eta \times \phi$ geometry of the crystals and electromagnetic shower spread along ϕ . The same approach cannot be used in the endcap due to the different arrangement. Therefore, a 5×5 algorithm was developed to collect the energy deposits in a cluster of 5×5 crystals around a so-called *crystal seed*. If more than one cluster is found in the same region and the total energy exceeds a certain threshold, they are collected into a *supercluster*. In Run 2, the new approach is part of the **PF** reconstruction described in Section 2.3.3. The algorithm clusters contiguous crystals around a crystal seed. This procedure is referred to as *mustache* clustering [66]. The energy requirement for the latter is $E_{\text{seed}} > 230$ MeV in the barrel and $E_{\text{seed}} > 600$ MeV or $E_{T,\text{seed}} > 150$ MeV in the endcap. A crystal is added to the cluster only if it has an energy deposit more than two standard deviations above the electronic noise. The significant improvements to the energy resolution of using this approach at the **HLT** trigger level as shown in Figure 2.22(a), similar gain is obtained also in the offline reconstruction.

As already mentioned, the cluster energy and position are then back-propagated to the nominal vertex in the magnetic field to build the electron track. Both $+1$ and -1 charge hypotheses are used to look for compatible tracker hits in the pixel detector. If compatible hits are found, an electron track seed is built and the tracking process starts. Given that the energy loss due to bremsstrahlung is highly non-Gaussian, the **KF** is not optimal for reconstructing electron tracks. A **Gaussian-sum Filter** (**GSF**) algorithm has therefore been implemented [67]. This algorithm describes the electron energy loss distribution, well represented by the Bethe-Heitler function, using a Gaussian mixture rather than a single Gaussian. A detailed description of the algorithm is given in Section 3.1.2. The improvement on the momentum reconstruction compared to the standard **KF** can be seen in

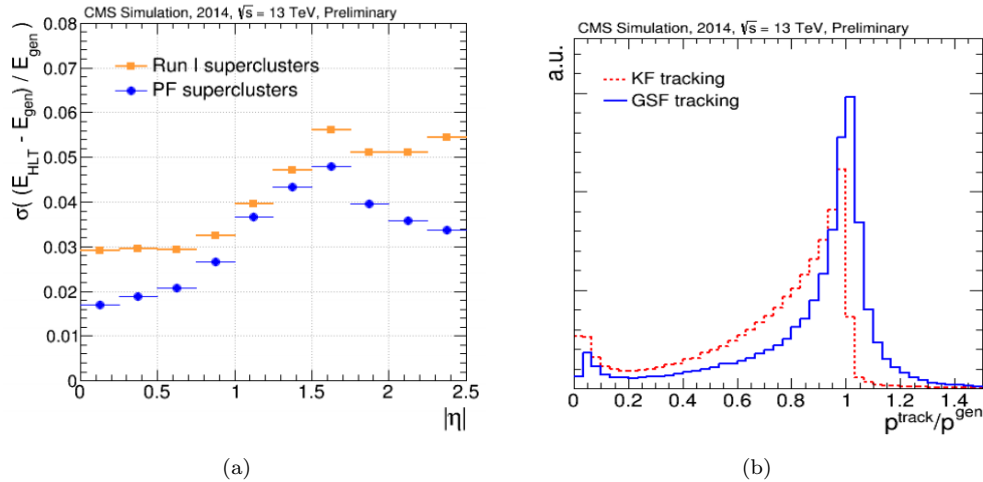


Figure 2.22: (a) Comparison of energy resolution of the two clustering algorithms as a function of the η of the electron after the energy correction procedure [66]. (b) Comparison of reconstructed particle momentum with respect to the simulated one using the **GSF** and **KF** track reconstruction algorithms [66].

Figure 2.22(b) using the full **CMS** simulation for the **HLT** trigger level, similar improvement is obtained also for the offline reconstruction.

In the final step, the electron candidate is built by associating reconstructed **GSF** tracks to its corresponding **ECAL** cluster. At this stage, the energy measurement provided by the electromagnetic calorimeter and the tracker momentum are merged. This improves significantly the estimation of the electron momentum at the interaction vertex for low energy particles, due to the fact that tracker momentum and **ECAL** energy estimation are differently affected by the bremsstrahlung.

2.3.3 Particle Flow reconstruction

The **Particle Flow** (**PF**) algorithm is a whole-event reconstruction technique [68]. It aims at reconstructing and identifying different particles in the event, i.e. electrons, muons, photons, charged hadrons and neutral hadrons, from a combination of all **CMS** sub-detector information. In this way the determination of particle types, directions and energies is optimized. While no substantial changes are expected for the reconstruction of high-energy electrons and muons, the resulting list of particles reconstructed with the **PF** algorithm represents the best description of the event at the particle level and can therefore be used to improve resolution and reconstruction for higher-level objects such as jets, missing transverse energy, τ_h etc. Stable particles are reconstructed in the **PF** algorithm as follows:

- Photons are identified as **ECAL** energy deposits without any linked charged particle trajectory. Their energy is directly obtained from the **ECAL** measurement.
- Electrons are identified as a primary charged particle track combined with **ECAL** energy clusters. The possibility of bremsstrahlung photons emission is also taken into account. Their energy is determined by a combination of the track momentum at the interaction point, the **ECAL** clusters energy, and the sum of all bremsstrahlung photons energy.
- Muons are identified as a collection of silicon tracker hits consistent with a track or several hits in the muon system. Muons are also associated with an energy deficit in the calorimeters and their energy is obtained from the entire track momentum.
- Charged particles that are not identified as electrons or muons are considered to be charged hadrons. Their energy is determined from a combination of the track momentum and the corresponding **ECAL** and **HCAL** energy.

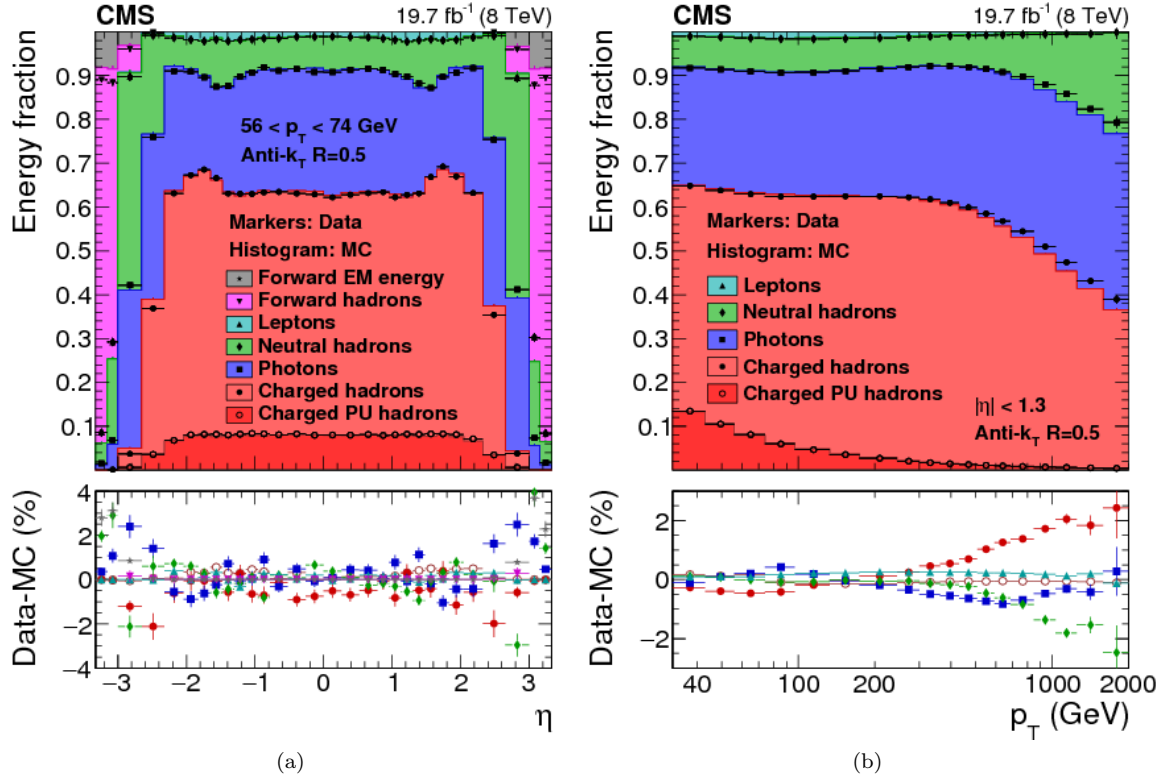


Figure 2.23: Reconstructed jet energy fractions in data and simulation as a function of pseudorapidity (a) and of p_T (b). The top part of each plot shows the stacked energy fractions, whereas the bottom part shows the percentage difference between observed and simulated events [68]. Charged hadrons associated with pileup vertices are denoted as *charged PU hadrons*.

- Neutral hadrons are identified as **HCAL** energy clusters not linked to any charged hadron trajectory, or as **ECAL** and **HCAL** energy excesses with respect to the expected charged hadron energy deposit. Their energy is obtained from the corresponding **ECAL** and **HCAL** energy.

Figure 2.23 shows the composition of typical minimum-bias events for both data and simulation. In the central part of the detector, the largest fraction of the event energy is carried by charged hadrons, while in the forward part the event energy is carried by hadronic candidates without distinction between charged and neutral particles. The figure underlines clearly the need of the tracker information to determine the hadron charge and that the electromagnetic objects contribute very little to the event energy. Therefore, once the well-isolated leptons are excluded from the particle list, the other particles produced by the **PF** can be clustered allowing a natural definition of “jet” and “missing energy” in the event.

2.3.4 Jets and missing transverse energy

In the **LHC** environment, a significant amount of jets are produced via **QCD** interactions for each event. Therefore, jet reconstruction and identification is one of the most important and challenging tasks in **CMS**. To allow for an accurate comparison between theory prediction and observation, well-defined jet finding methods are required. **CMS** uses two classes of algorithms to reconstruct jets: a cone-based one and a sequential-clustering one [69]. Cone-based algorithms assume that the jet extends in conical regions and therefore group particles based on $\eta - \phi$ space. Some examples are the Iterative Cone [59] and SIScone [70] algorithms. In contrast, sequential-clustering algorithms assume that particles within jets have small differences in transverse momenta and cluster them

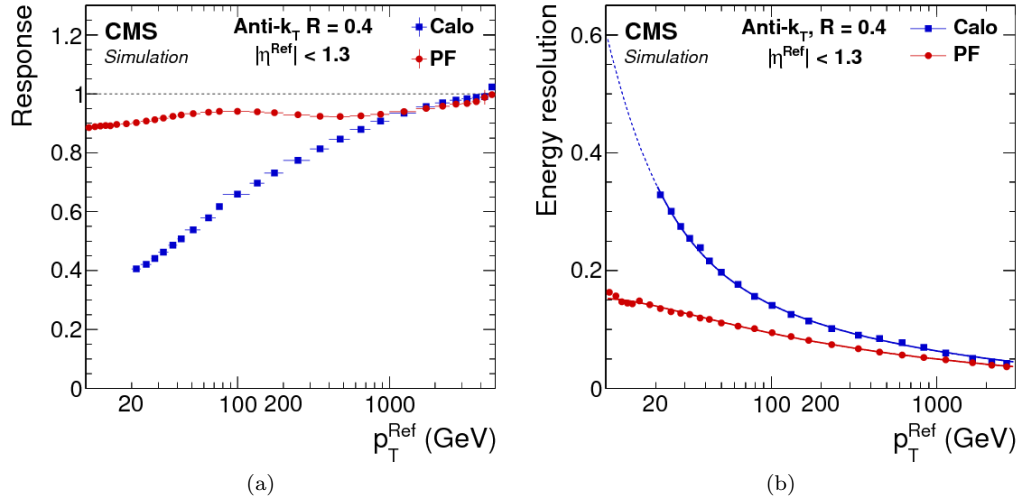


Figure 2.24: Jet response (a) and jet energy resolution (b) as a function of the p_T in the barrel regions for CaloJets (blue) and pfJets (red) collections. The lines, added to guide the eye, correspond to fitted functions with ad hoc parametrizations [68].

based on momentum space. k_t [71], the Cambridge/Aachen [72] and the Anti- k_t [73] algorithms are examples of this category. Three different categories of jets can be reconstructed depending on the sub-detector information used: *Calorimeter Jet (CaloJets)*, if electromagnetic and hadronic calorimeters energy deposits are used as input; *Calorimeter Jet-Plus-Tracks*, if CaloJets are further improved with the tracking system information; *Particle Flow Jets (pfJets)*, if the clustering algorithm is fed with PF particles.

The presence of particles invisible to the CMS detector is spotted using the missing transverse energy quantity, already described in Equation 2.7. The main idea is to search for an imbalance in the event transverse energy which can be caused by neutral particles such as neutrinos. The probability that this missing energy is associated with particles mis-reconstructed or falling outside the detector acceptance is expected to be small, due to the hermeticity and the high efficiency of CMS. Three techniques are used for the E_T^{miss} reconstruction depending on the sub-detector information: *Calorimetric MET*, *Track-Corrected MET* and *Particle Flow MET (pfMET)* respectively.

Using fully calibrated PF candidates as inputs for both jets and missing energy reconstruction has the advantage that little or no a-posteriori energy correction is needed. Moreover, the combination of the different sub-detectors allows for a significant degree of redundancy. In this way the reconstructed calorimetric objects are themselves less sensitive to the energy calibration. For instance, the pfMET does not require any further correction for muons or tracks, since they are already considered in the inclusive reconstruction approach. In Figure 2.24 the jet energy resolution and response as a function of the p_T is shown for simulated CaloJets and pfJets collections.

2.3.5 Hadronic τ lepton decays

As already mentioned in Section 1.2.3, the τ particle is the only known lepton heavy enough to decay into hadrons, which it does in about two thirds of the cases. In about 35% of the cases, a τ decays into an electron or muon and two neutrinos. Both muon and electron reconstruction is performed using the standard CMS algorithms already described in Section 2.3.1 and 2.3.2, respectively. In the case of the τ_h , the decay can be classified into the number of π^0 s and charged particles produced. If the decay contains a single track is defined *1-prong*, while it is defined *3-prong* if three tracks are generated. The hadronic decay is reconstructed using the hadron-plus-strips (HPS) algorithm [74, 75]. This algorithm is seeded with anti- k_t jets and uses the number of tracks and the number of ECAL energy deposits in the event to reconstruct the neutral pions that are present in most τ_h decays. The high probability for photons originating from $\pi^0 \rightarrow \gamma\gamma$ to

convert in e^+e^- is accounted for by collecting the photon and electron ($p_T > 0.5$ GeV) constituents into clusters (*strips*). The τ_h candidates are then divided in 1-prong, 1-prong + $\pi^0(s)$, and 3-prong decay modes by combining the strips with the charged-particle:

- a single charged particle without any strips: $h^\pm$;
- one charged particle and one strip: $h^\pm\pi^0$;
- combination of a single charged particle with two strips: $h^\pm\pi^0\pi^0$;
- combination of three charged particles: $h^\pm h^\pm h^\pm$.

Electrons and muons misidentified as τ_h are suppressed using dedicated criteria based on the consistency between the measurements in the tracker, the calorimeters, and the muon detectors. The main challenge in the τ_h reconstruction is the distinction with respect to the quark and gluon jets. In order to distinguish from this main background, some specific characteristic of the particles produced in τ_h decays are used in the algorithm, such as the lower multiplicity, the deposited energy confined in a narrow region, and the isolation with respect to other particles in the event. After the LHC Run-1, the HPS algorithm has been improved:

- an improved strip reconstruction algorithm, called dynamic strip reconstruction, was implemented. It takes into account also the electromagnetic energy leakage of the τ_h decay.
- an improved Multivariate Analysis (MVA) based discriminator [76] was implemented, including isolation as well as lifetime information, to suppress quark and gluon jets misidentification.
- further developments were added to the already existing MVA based discriminator to suppress $e \rightarrow \tau_h$ misidentification.

In Figure 2.25 some performance plots are shown. In Figure 2.25(a), the misidentification probabilities for the MVA based discriminator and for the isolation sum discriminators are shown as a function of τ_h identification efficiency. It is evaluated using $H \rightarrow \tau^+\tau^-$ and QCD simulation samples. Each point in the graph corresponds to a specific working point of the discriminators. In Figure 2.25(b), the probability for quark and gluon jets in W+jets events to pass the loose working point of the MVA based isolation discriminant as a function of the p_T of the jet is shown. The misidentification probabilities measured in observed events are compared to the expectation from simulation. The shaded bands include the systematic uncertainties related to the background subtraction and the jet energy scale.

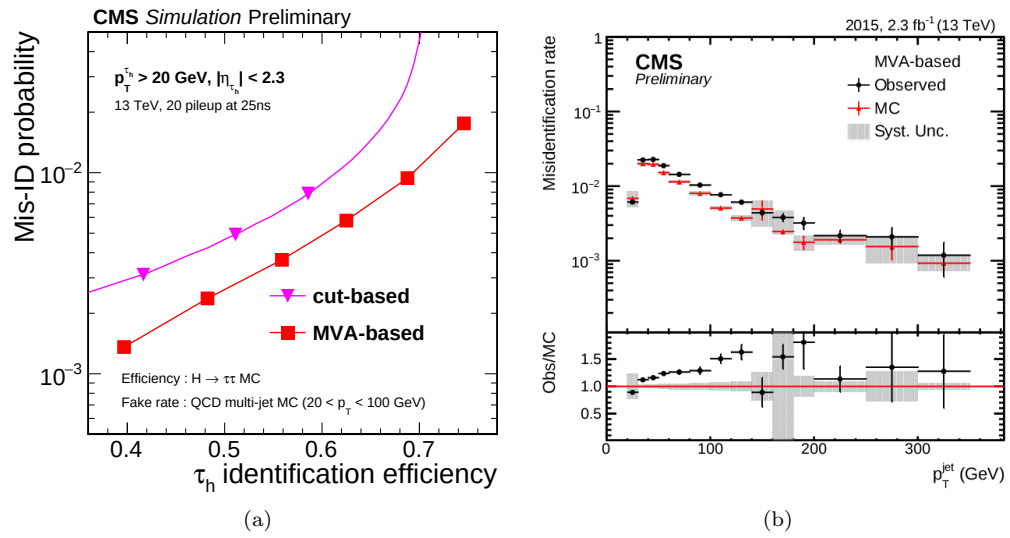


Figure 2.25: (a) Misidentification probability for the **MVA** based discriminator (red) compared to the cut-based isolation sum discriminators (purple) as a function of τ_h identification efficiency, evaluated using $H \rightarrow \tau^+\tau^-$ and **QCD** simulation samples [75].
 (b) The misidentification probabilities for quark and gluon jets to pass the *Loose* working point of the **MVA** based isolation discriminant as a function of the p_T of the jet for observed (black) and simulated (red) events [75].

CHAPTER 3

Track reconstruction in CMS

*To Marie Curie, first woman
Nobel Prize. (1904)*

Track reconstruction is one of the crucial parts in the event reconstruction because of its importance in the estimation of the particle momentum and in vertex identification. In proton-proton collisions at the **LHC**, this task is very challenging given the hundreds of particles generated in each bunch crossing. In **Section 3.1** a general theoretical and experimental overview of the track reconstruction in the **LHC** environment is given, focusing in particular on the different track fitting methods. A summary of the historical evolution of track reconstruction in the **CMS** experiment is then given. In **Section 3.2** the iterative tracking philosophy is described and first results on **Monte Carlo (MC)** are shown for Run 1. In the following section the developments introduced in the track reconstruction during **Long Shutdown 1 (LS1)** are described with particular emphasis on the implementation and results obtained with the **Deterministic Annealing Filter (DAF)** algorithm. Finally, the results on data and simulation for Run 2 are reported to verify the compatibility between data and **MC** and to assess the tracking performance after the upgrade of the Pixel Tracker at the end of the year 2016.

3.1 Theoretical and experimental context

The experimental scenario of today's high-energy particle physics experiments operating at a proton-proton collider can be summarized as follows:

- High particle multiplicities due to the high collision energy and luminosity provided by **LHC**, even higher in the case of heavy ion collisions.
- Large momentum range of particles in the final state from a few hundred MeV up to several hundred GeV.
- Multiple scattering is expected in non-sensitive material, such as detector frames, support and cooling structures.
- Large activity produced by the secondary interactions to be recognized, measured and included in the final event reconstruction. Some examples of secondary interaction are: high energy delta electrons produced by muons, hadronic interactions of pions, γ -conversion and bremsstrahlung for electrons.
- Different detector technologies are required to reconstruct distinct particles types in various ranges of p_T with the highest precision possible. It is also imperative that these technologies meet the radiation hardness requirements.

- High event rates leading to a large amount of data which needs to be selected and analyzed in the most efficient way.

The entire list needs to be taken into account in the track reconstruction process, which therefore becomes a very challenging task.

Track reconstruction in high-energy physics is usually divided into *track finding* and *track fitting*. In track finding, the pattern recognition problem is solved and the signals generated by charged particles in tracker detectors are grouped into track candidates. In track fitting, the best estimation of the track parameters is derived by error minimization. For both track finding and track fitting a large variety of methods have been developed in the past fifty years. A detailed overview of the history and an introduction into the subject can be found in Ref. [77].

From the physics point of view, there are a few aspects that are tightly connected with the quality of track fitting, such as:

- Invariant masses must be determined with the best precision and well-estimated errors. This is important not only to discover new particles in the mass spectra, but also to determine new physics contributions in already well-known resonances.
- Secondary vertices must be fully reconstructed to estimate lifetimes on the order of 10^{-13} sec, for example in the case of B mesons or τ leptons.
- Kink reconstruction must be efficient in order to recognize decays containing neutrinos.

An optimal track fitting requires a deep knowledge of the behavior of the detector as well as of the magnetic field, a precise treatment of multiple scattering and energy loss, and a track model which describes the trajectory of a charged particle.

3.1.1 The track model

A possible way to determine the momentum of a particle and its charge is to immerse it in a magnetic field and to measure the deviation from a straight trajectory with a position-sensitive detector. In order to determine particle charge and momentum, the track path must be expressed as a function of a finite number of parameters. The set of solutions of the equation of motion is the *track model* [77].

The equation of motion of a charged particle in a magnetic field is described by the Lorentz force [78]:

$$\vec{F} = \frac{d}{dt} \left(m \frac{d\vec{x}(s)}{dt} \right) = q\vec{v} \wedge \vec{B}(\vec{x}(s)) \quad (3.1)$$

where $\vec{F}$ is the Lorentz force, q and $\vec{v}$ are the particle charge and velocity, respectively, and $\vec{B}(\vec{x}(s))$ is the static magnetic field. Using the distance along the trajectory $s(t)$ and the absolute value of the velocity $v = ds/dt$, the following identities hold:

$$\frac{d\vec{x}(s)}{dt} = \frac{d\vec{x}(s)}{ds} \frac{ds}{dt} = \frac{d\vec{x}(s)}{ds} v, \quad (3.2)$$

$$\frac{d^2\vec{x}(s)}{dt^2} = \frac{d^2\vec{x}(s)}{ds^2} v^2. \quad (3.3)$$

Equation 3.1 can be thus rewritten as

$$\frac{d^2\vec{x}(s)}{ds^2} = \frac{q}{mv} \frac{d\vec{x}(s)}{ds} \wedge \vec{B}(\vec{x}(s)) \quad (3.4)$$

If the mass of the particle is known, in this equation there are six integration constants. With the identity

$$\left(\frac{dx}{ds} \right)^2 + \left(\frac{dy}{ds} \right)^2 + \left(\frac{dz}{ds} \right)^2 = 1 \quad (3.5)$$

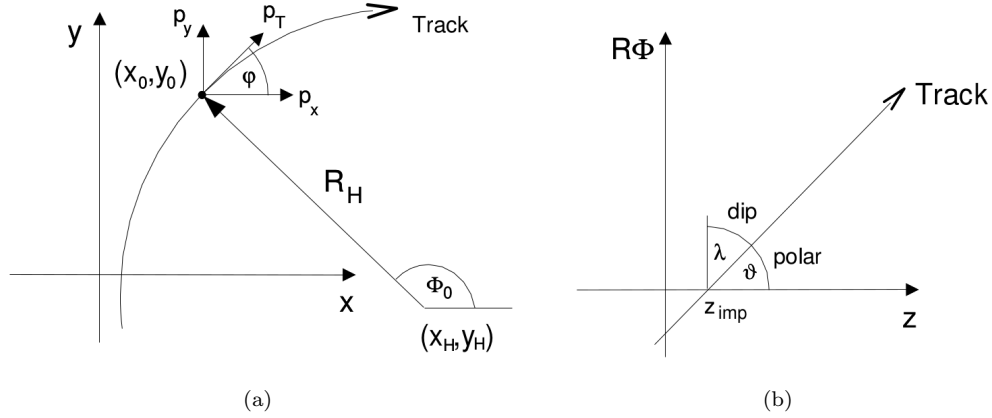


Figure 3.1: Helix parameters in the $x-y$ (a) and $R\Phi-z$ (b) plane for a constant magnetic field parallel to the z -axis.

and an arbitrary choice of one coordinate, the so-called *reference surface*, the number of free parameters that describes a track can be reduced to five: two position parameters for the impact point at the reference surface, two describing the direction of the track at the impact point and one for the particle momentum.

In the special case of a *homogeneous magnetic field*, $\vec{B}$ can be chosen parallel to the z -axis, and therefore can be written $\vec{B} = B\vec{e}_z$ where $\vec{e}_z = (0, 0, 1)^T$. This is the case for the central region of many accelerator experiments, such as **CMS** and **ATLAS** at the **LHC**. In this case, the equations of motion are

$$\frac{d^2x}{ds^2} = \frac{q}{mv} \frac{dy}{ds} B, \quad (3.6)$$

$$\frac{d^2y}{ds^2} = -\frac{q}{mv} \frac{dx}{ds} B, \quad (3.7)$$

$$\frac{d^2z}{ds^2} = 0. \quad (3.8)$$

The solution is a *helix* with an axis parallel to z :

$$x(s) = x_0 + R_H \left(\cos \left(\Phi_0 + \frac{hs}{R_H} \cos \lambda \right) - \cos \Phi_0 \right), \quad (3.9)$$

$$y(s) = y_0 + R_H \left(\sin \left(\Phi_0 + \frac{hs}{R_H} \cos \lambda \right) - \sin \Phi_0 \right), \quad (3.10)$$

$$z(s) = z_0 + s \sin \lambda. \quad (3.11)$$

where s is the path length along the track, $(x_0, y_0, z_0)^T$ is the starting point, R_H is the radius of the helix, $\lambda = \arcsin(dz/ds)$ is the dip angle ($-\pi/2 \leq \lambda \leq +\pi/2$), and Φ_0 is the azimuthal angle at the starting point with respect to the helix axis. The sense of rotation of the projected helix in the transverse plane is given by h , which is defined as $h = -\text{sign}(qB_z) = \pm 1$. A schematic representation of the helix parameters can be seen in **Figure 3.1**.

In the presence of an *inhomogeneous magnetic field*, the full trajectory of the particle can be estimated only from an analytical or numerical solution of **Equation 3.1**. A possible way to take into account the presence of an inhomogeneous magnetic field as well as the effect of material is to use the fourth-order Runge-Kutta method [77]. In order to solve the equations of motion numerically, this method uses a recursive formula to extrapolate the track trajectory in space.

Track parameters and reference frames

The track parameters can be defined in different coordinate frames [79]:

- **Global reference frame:** In this frame the position and the direction of the particle are defined using global coordinates. In global space two coordinate systems can be used: the Cartesian and the curvilinear system. The Cartesian parameters define a particle with a six coordinate vector:

$$(x, y, z, p_x, p_y, p_z) \quad (3.12)$$

The first three parameters represent the particle position, while the latter its momentum (direction) in global coordinates. The associated covariance matrix is a 6×6 matrix of rank 5.

The curvilinear parameters are defined by a vector with five components:

$$\left(\frac{q}{p}, \lambda, \phi, x_T, y_T \right) \quad (3.13)$$

where q/p is the signed inverse of momentum, and λ and ϕ are defined using the momentum components as following:

$$\tan \lambda = \frac{p_z}{p_T}, \quad (3.14)$$

$$\tan \phi = \frac{p_y}{p_x}. \quad (3.15)$$

The definition of p_T is given in Equation 2.6. The coordinates x_T and y_T are defined using a local Cartesian coordinate system defined by three orthogonal unit vectors $\vec{u}$, $\vec{v}$ and $\vec{t}$. The vector $\vec{t}$ is parallel to the track and pointing in the particle direction. Using the unit vector $\vec{z}$ parallel to the global z -direction, the two vectors $\vec{u}$ and $\vec{v}$ are

$$\vec{u} = \frac{\vec{z} \wedge \vec{t}}{|\vec{z} \wedge \vec{t}|} \quad (3.16)$$

$$\vec{v} = \vec{t} \wedge \vec{u} \quad (3.17)$$

Therefore, the z_T -axis is pointing along the particle direction, the x_T -axis is lying in the global $x - y$ plane, and the y_T -axis is perpendicular on the two others, in order to form a right-handed Cartesian coordinate system (Figure 3.2).

- **Local reference frame:** This frame is defined by a particular detector surface as shown in Figure 3.3. The five coordinates used are:

$$\left(\frac{q}{p}, \frac{dx}{dz}, \frac{dy}{dz}, x, y \right) \quad (3.18)$$

The local z -axis is perpendicular to the detector surface. The x coordinate is chosen as the one measured with highest precision, while y with the least known precision as shown in the figure.

3.1.2 The track fit

As already mentioned in Section 3.1, the track fit requires the knowledge of

- the geometrical layout of the detector and the resolution of the sensors which provide the track measurements and their associated errors;
- a stochastic model of material effects, such as multiple scattering and energy loss;

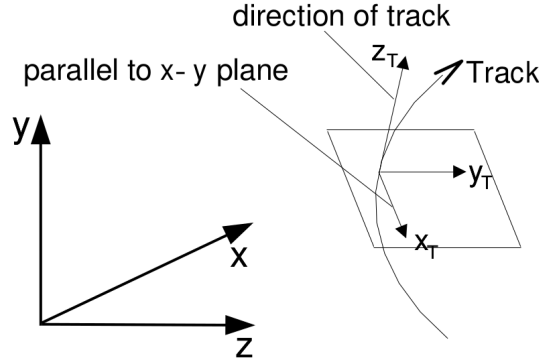


Figure 3.2: Schematic representation of the x_T , y_T and z_T coordinates defined in the curvilinear global frame.

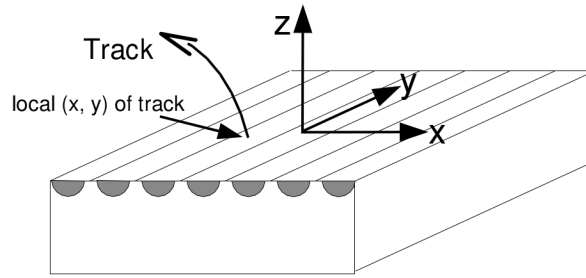


Figure 3.3: Schematic representation of the local frame with respect to a particular detector surface.

- a track model, depending on the magnetic field.

A track can be described at any point by a 5-component vector of track parameters $\vec{x}$, called the *state vector*, determined according to the constraints given by the equation of motion. This vector cannot be observed directly. The relation between the state vector and the observed n -measurements, described by the *measurement vector* $\vec{m}$, is

$$\vec{m} = \vec{f}(\vec{x}) + \vec{\epsilon} \quad (3.19)$$

where $\vec{\epsilon}$ is the vector of measurement errors and $\vec{f}$ is a deterministic function of $\vec{x}$:

$$\vec{f} : \vec{x} \rightarrow f_i(\vec{x}), i = 1, \dots, n \quad \text{or} \quad \vec{f}(\vec{x}) \quad (3.20)$$

The *covariance matrix* V of the measurements is defined as

$$\text{cov}(\vec{\epsilon}) = V \quad (3.21)$$

with the *weight matrix* G being

$$G = V^{-1} \quad (3.22)$$

The main task of the track fitting is to find a meaningful mapping $\vec{F}$ without bias and with minimum variance for the fitted parameters:

$$\hat{\vec{x}} = \vec{F}(\vec{m}) \quad (3.23)$$

The expectation value of the fitted vector $\hat{\vec{x}}$ is supposed to be the true value $\vec{x}_{\text{true}}$

$$E(\hat{\vec{x}}) = \vec{x}_{\text{true}} \quad (3.24)$$

with covariance matrix defined as

$$C(\hat{\vec{x}}) = E((\hat{\vec{x}} - \vec{x}_{\text{true}}) \cdot (\hat{\vec{x}} - \vec{x}_{\text{true}})^T) \quad (3.25)$$

The fit methods can be divided into two groups:

- **Hard assignment methods:** the track is fitted considering a hard assignment of the hits to the track, i.e. a hit either belongs or does not to the track. These are also called *classical methods* and some of them are: the Global Fit [77], the **Kalman Filter (KF)** [80], and the **Gaussian-sum Filter (GSF)** [81].
- **Soft assignment methods:** in this method several competing hits can contribute to the track, each of them with a respective assigned weight. The tracks can therefore share hits among each other. These methods are also called *adaptive methods* and one example is the **Deterministic Annealing Filter (DAF)** [82].

The Global Fit

The Global Fit is based on the **Least Squares Method (LSM)**. The errors are considered Gaussian distributed. The Global Fit is simpler and faster than other methods. The **LSM**-estimate $\hat{\vec{x}}$ of the state vector is the value which minimizes the following function:

$$M(\vec{x}) = (\vec{m} - \vec{f}(\vec{x}))^T G (\vec{m} - \vec{f}(\vec{x})) \quad (3.26)$$

The weight matrix G in this case contains also multiple scattering effects which are included in the correlation among measurements. If the track model can be approximated by a linear model in the neighbourhood of the measurements, the function $\vec{f}$ can be written as an expansion around the point $\vec{x}_0$

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) + H \cdot (\vec{x} - \vec{x}_0) + O((\vec{x} - \vec{x}_0)^2) \quad (3.27)$$

with

$$H = \frac{\partial \vec{f}(\vec{x})}{\partial \vec{x}}(\vec{x} = \vec{x}_0) \quad (3.28)$$

Differentiating $M(\vec{x})$ with respect to $\vec{x}$ and putting $\partial \vec{M}(\vec{x}) / \partial \vec{x} = 0$ yields the estimate of the vector state:

$$\hat{\vec{x}} = \vec{x}_0 + (H^T G H)^{-1} H^T G (\vec{m} - \vec{f}(\vec{x}_0)) \quad (3.29)$$

This procedure can be iterated to find also lower order estimates of the vector state. Among the different properties of the **LSM** for a linear model, the following ones can be found:

- If the measurement vector is unbiased, then the **LSM**-estimate of the state vector is also unbiased:

$$E(\hat{\vec{x}} - \vec{x}_{\text{true}}) = \vec{0} \quad (3.30)$$

- The covariance matrix of $\hat{\vec{x}}$, i.e. the error matrix of the fitted parameter, is

$$C(\hat{\vec{x}}) = (H^T G H)^{-1} \quad (3.31)$$

- The **LSM**-estimate is consistent.
- Several methods were developed in order to reduce the effect of measurements that deviate from the expected behavior on the evaluation of the **LSM**-estimate. One example is to modify the error matrix ensuring a correct propagation of the errors. If the estimate is insensitive to this effect, the **LSM**-estimate is considered robust against outliers.

In the case of non-Gaussian errors and a non-linear model the **LSM**-estimate is asymptotically unbiased and consistent.

The Kalman Filter

Originally the **KF** was developed to estimate the unobservable states of a stochastic model evolving in time (dynamic system). The **KF** is a recursive **LSM**-fit consisting of a succession of *prediction* and *filter* steps. In the prediction step the current state vector is extrapolated to a future time, taking into account multiple scattering interaction and possible energy loss. In the filter step the extrapolated state vector is updated with the information of the closest local measurement to the predicted state. The distance is usually computed using the χ^2 -statistic. This decision is purely local, and therefore does not take into account that other tracks can generate measurements even closer to the true one. In the case of low track density this poses no particular problem, but in the case of high track density global decision rules can be considered as an alternative.

The application of the **KF** to the track reconstruction problem is straightforward when the track is modeled as a dynamic system:

- The state of the track at a specific surface k is given by the 5-component state vector $\vec{x}_k$. The evolution in time, i.e. the trajectory of the particle between two adjacent surfaces, is described by the following equation:

$$\vec{x}_k = \vec{f}_{k-1}(\vec{x}_{k-1}) + \vec{w}_{k-1} \quad (3.32)$$

where $\vec{f}$ is a deterministic function which propagates the state vector from the surface $k-1$ to k and $\vec{w}_{k-1}$ includes the effect of the material on the trajectory, mostly multiple scattering. The latter is also called *noise process* in dynamic system language.

- The relation between the measurement $\vec{m}_k$ and the track state vector $\vec{x}_k$ at a surface k holds:

$$\vec{m}_k = \vec{h}_k(\vec{x}_k) + \vec{e}_k \quad (3.33)$$

with the measurement noise $\vec{e}_k$ and $\vec{h}_k$ the function that maps the track parameters on the measurements, often a simple projection on a subset of the state vector. In other words, the measured quantities are expressed as functions of the state vector, corrupted by a measurement noise.

If $\vec{f}$ and $\vec{h}$ are linear functions, the Equations (3.32) and (3.33) can be rewritten as:

$$\vec{x}_k = \vec{F}_{k-1}\vec{x}_{k-1} + \vec{w}_{k-1} \quad (3.34)$$

and

$$\vec{m}_k = \vec{H}_k\vec{x}_k + \vec{e}_k, \quad (3.35)$$

respectively. If $\vec{f}$ and $\vec{h}$ are non-linear, they can be approximated by their first-order Taylor expansion. The Jacobians $\vec{F}$ and $\vec{H}$ are computed as:

$$\vec{F}_{k-1} = \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x} = \vec{x}_{k-1}) \quad (3.36)$$

$$\vec{H}_k = \frac{\partial \vec{h}}{\partial \vec{x}}(\vec{x} = \vec{x}_k) \quad (3.37)$$

As already mentioned, the track fitting with the **KF** is performed in two steps:

- **Prediction:** using the information from all the measurements up to $k-1$ and the magnetic field extrapolation, an estimation of the future state vector along with its covariance matrix is given:

$$\hat{\vec{x}}_{k,\text{pred}} = \vec{f}_{k-1}(\hat{\vec{x}}_{k-1}) \quad (3.38)$$

$$C_{k,\text{pred}} = F_{k-1}C_{k-1}F_{k-1}^T + Q_{k-1} \quad (3.39)$$

where Q_{k-1} is the covariance matrix of w_{k-1} and is assumed to be known, and C_{k-1} is the covariance matrix defined in Equation 3.25.

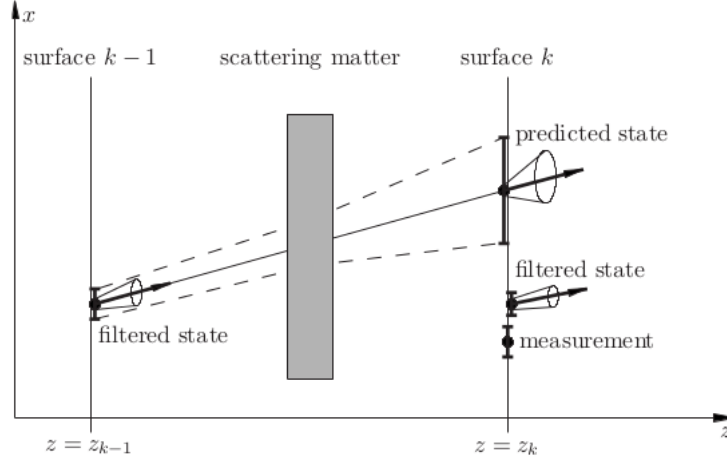


Figure 3.4: Schematic representation of the **KF** two-steps process.

- **Filtering (update)**: the predicted state vector $\hat{x}_{k,\text{pred}}$ is updated with a local measurement $\vec{m}_k$. The updated state vector on a detector surface with its covariance matrix can be written as following:

$$\hat{x}_{k,\text{update}} = \hat{x}_{k,\text{pred}} + K_k(\vec{m}_k - \vec{h}_k(\hat{x}_{k,\text{pred}})) \quad (3.40)$$

$$C_{k,\text{update}} = (I - K_k H_k) C_{k,\text{pred}} \quad (3.41)$$

where K_k is the Kalman gain matrix:

$$K_k = (C_{k,\text{pred}}^{-1} + H_k^T V_k^{-1} H_k)^{-1} H_k^T V_k^{-1} \quad (3.42)$$

and V_k is the covariance matrix of $\vec{m}_k$.

A schematic representation of both steps is given in Figure 3.4. If the dynamic system is strictly linear and $\vec{w}_k$ and $\vec{\epsilon}_k$ are Gaussian distributed, the **KF** is the optimal filter. In the case of non-Gaussian errors, the **KF** is the best linear filter.

When the last measurement is reached, the final estimate $\hat{x}_n$ contains the full information of the measurements $\vec{m}_1, \dots, \vec{m}_n$. The information can then be passed back to all previous estimate state vectors in the **smoother** step. The smoother step can also be done after each prediction/filter step using all measurements collected up to the present time. It can be implemented running two filters in opposite direction and combining both predictions with the local measurements.

The advantages of the Kalman filter/smoothen with respect to other fitting methods are:

- Thanks to its recursive nature, the **KF** can be used in track reconstruction not only as fitting procedure, but also as track finder.
- The computational cost is reduced because it is proportional to the number of surfaces crossed by the track and no large matrices have to be inverted.
- The estimated track parameters closely follow the real track.
- The linear approximation of the track model needs to be valid only between adjacent layers and not the entire length of the track.

It has to be noted that the **KF** needs an initial state vector, so-called *seed*, to start. Generally the seed is built where the track density is low or the granularity of the detector is high.

The Gaussian-sum Filter

LSMs have optimal properties if the track model can be approximated by a linear function and if the measurement errors and the process noise are Gaussian distributed. In the case of the **KF**, both momentum and its variance are corrected during each propagation step to take into account the radiative energy loss. This ensures unbiased estimates of the track parameters and of the associated uncertainties.

If the probability distributions encountered during the track reconstruction are not Gaussian, the **KF** cannot be considered optimal anymore and a non-linear estimator can achieve better results. An example of a non-linear generalisation of the **KF** is the **Gaussian-sum Filter (GSF)**, which describes the non-Gaussian errors distribution by mixtures of Gaussians. This algorithm consists of several **KFs** running in parallel but following one and the same physical track.

The **GSF** is used in particular for electron track reconstruction because of the highly non-Gaussian nature of electron energy loss process, mainly dominated by bremsstrahlung. The **GSF** algorithm has been implemented in the reconstruction software of the **CMS** tracker enhancing electron performance [67]. Major draw-backs of the **GSF** are its computational complexity and its missing protection against the assignment of a wrong hit when there is no competing “good” one from the actual track.

The Deterministic Annealing Filter

In the **DAF** algorithm the hard hit-to-track assignment is replaced by a soft assignment. This can be an effective way to protect against wrong hit assignments, for example hits that come from delta rays or hits belonging to a different track. It also reduces the combinatorics. Instead of a binary hit assignment (yes or no) to a track as it is in the **KF**, the **DAF** associates each hit with a certain probability, p , which can take any value between 0 and 1 and is computed using the hit residuals with respect to the predicted or smoothed state vector. It follows that if several hits are close enough to one track, all of them can compete for a contribution to the track with their respective assigned probability, also called *weight*.

To avoid wrong hit assignment due to insufficient information at the beginning of the fit, an iteration procedure is introduced. Initially a first complete track fit and smoothing is performed so that the assignment probability for each hit on a specific detector surface can be computed using the information from all other detector surfaces. If the hit probability falls below a certain threshold, usually computed using a χ^2 -cut, the hit is suppressed in the next iteration.

The **DAF** is an iterative **KF** with an annealing factor, T_I , which changes iteration by iteration. The update of the estimated state vector $\hat{x}$ and the covariance matrix on a surface k with n competing hits $\vec{m}_{k,1}, \dots, \vec{m}_{k,n}$ reasonably close to the same track are:

$$\hat{x}_{k,\text{update}} = \hat{x}_{k,\text{pred}} + K_k \sum_{i=1}^n p_i (\vec{m}_{k,i} - h_k(\hat{x}_{k,\text{pred}})) \quad (3.43)$$

$$C_{k,\text{update}} = (C_k^{-1} + p H_k^T V_{I,k}^{-1} H_k)^{-1} \quad (3.44)$$

where p is the sum over all weights p_i and the Kalman gain matrix K_k is defined in this case as

$$K_k = (C_{k,\text{pred}}^{-1} + p H_k^T V_k^{-1} H_k)^{-1} H_k^T V_{I,k}^{-1} \quad (3.45)$$

For each iteration the hit weight on a particular surface is computed based on the residual of the hit using the measurement $\vec{m}_i$ and the smoothed state $\hat{x}_{k,\text{smooth}}$ on a particular surface:

$$\phi_i = \exp \left[-\frac{1}{2} \frac{(\vec{m}_i - \vec{h}_k(\hat{x}_{k,\text{smooth}})) V_i^{-1} (\vec{m}_i - \vec{h}_k(\hat{x}_{k,\text{smooth}}))^T}{T_I} \right] \quad (3.46)$$

and then renormalized as following:

$$p_i = \frac{\phi_i}{\sum_{j=1}^n \phi_j + \sum_{j=1}^n \phi_j^{\text{cut}}} \quad (3.47)$$

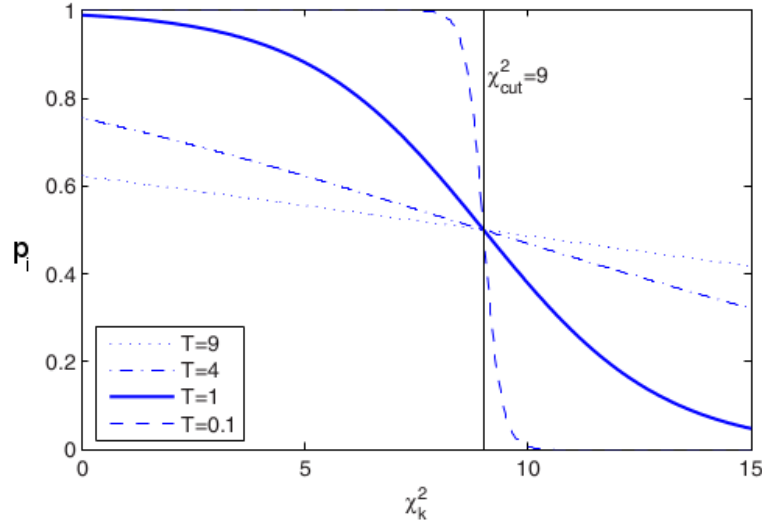


Figure 3.5: Example of hit assignment probability as a function of the χ^2 -threshold value for different annealing factors.

where ϕ_j^{cut} is a cut-off value which forces the assignment probability to be close to zero if the probability of a hit falls below a certain value expressed in terms of a χ^2 -threshold value:

$$\phi_{j,\text{cut}} = \frac{1}{(2\pi)^{n/2} \sqrt{\det V_{j,I}}} \exp \left[-\frac{1}{2} \frac{\chi_{\text{cut}}^2}{T_I} \right] \quad (3.48)$$

The cut-off value varies for each iterations with the annealing factor. The annealing factor is particularly important because it avoids dependences of the final assignment probabilities on their initial values. For each iteration the variance of the j -hit $V_{j,I}$ is re-computed as following:

$$V_{j,I} = T_I V_j \quad (3.49)$$

A typical annealing scheme is $T = \{81, 9, 4, 1, 1, 1\}$ in order to artificially increase the hit position errors in the first iterations and after few iterations bring them back to the nominal value. In this way, at the beginning of the algorithm all the hits contribute to the estimation of the track parameters, while in the end a weight near to 1 will be assigned to only one of the hits while the others will be assigned a very low weight. This can be seen in [Figure 3.5](#), where the assignment hit probability is plotted as a function of the χ^2 -threshold for different annealing factors.

The χ^2 of a track with n competing hits on a surface k is defined as the sum of the individual $\chi_{i,k}^2$:

$$\chi_k^2 = \sum_i \chi_{k,i}^2 \quad (3.50)$$

with

$$\chi_{k,i}^2 = p_i (\vec{m}_i - \vec{h}_k(x_{k,i,\text{pred}}))^T (V_i T_I + H C_{j,\text{pred}} H^T)^{-1} (\vec{m}_i - \vec{h}_k(x_{k,i,\text{pred}})) \quad (3.51)$$

The total χ^2 of the track is the sum over all the χ_k^2 :

$$\chi^2 = \sum_k \chi_k^2 \quad (3.52)$$

The number of degrees of freedom is the difference between the number of the measurements multiplied with the assignment probability and the number of the estimated parameters.

A possible way to implement the **DAF** is introducing a hit with virtual position and virtual covariance matrix, so-called *MultiRecHit*, which incorporates several competing hits. The virtual measurement $\vec{m}'$ is defined as the weighted mean of all competing measurements $\vec{m}_i$:

$$\vec{m}' = \left(\sum_i p_i G_i \right)^{-1} \sum_i p_i G_i \vec{m}_i \quad (3.53)$$

with $G_i = V_i^{-1}$. The covariance matrix V' associated to $\vec{m}'$ is:

$$V' = \left(\sum_i p_i G_i \right)^{-1} \quad (3.54)$$

In this way, it is possible to use identical formulas for the update step of a track state with corresponding covariance matrix and for the Kalman Gain matrix replacing $\vec{m}$ with $\vec{m}'$ and V with V' as follows:

$$\hat{x}_{\text{update}} = \hat{x}_{\text{pred}} + K'(\vec{m}' - \vec{h}(\hat{x}_{\text{pred}})) \quad (3.55)$$

$$K' = (C_{\text{pred}}^{-1} + H^T V'^{-1} H)^{-1} H^T V'^{-1} \quad (3.56)$$

$$C_{\text{update}} = (C_{\text{pred}}^{-1} + H^T V'^{-1} H)^{-1} \quad (3.57)$$

3.1.3 Specific choices in the CMS experiment

In many collider experiments the magnetic field in the region close to the proton-proton interaction point is approximately constant and rotationally invariant with respect to the beam axis. As described in [Section 3.1.1](#), the trajectories of charged particles in these conditions are helices and thus the natural choice of the global system of reference is often cylindrical. For the specific context of the reconstruction software of the **CMS** experiment, the global reference is defined in [Section 2.2](#). The five curvilinear parameters used in the pattern recognition and in the fit with respective units are:

- $q/|p|$, signed inverse momentum measured in GeV^{-1}
- $\lambda = \pi/2 - \theta$, where θ is the polar angle
- ϕ , the azimuthal angle
- $d_{xy} = -v_x \sin \phi + v_y \cos \phi$, measured in cm
- $d_{sz} = v_z \cos \lambda - (v_x \cos \phi + v_y \sin \phi) \sin \lambda$, measured in cm

where (v_x, v_y, v_z) is the reference position of the track in the global Cartesian system. The d_{xy} parameter represents the signed distance in the transverse plane between the straight line passing through (v_x, v_y) with azimuthal angle ϕ and the point $(0,0)$. The d_{sz} parameter is defined using an s -axis, which is defined by the projection of the straight line passing through (v_x, v_y, v_z) with angles (ϕ, λ) . The convention is to assign the s coordinate for (v_x, v_y) as the value $v_x \cos \phi + v_y \sin \phi$, which becomes zero when (v_x, v_y) is the point of minimum transverse distance to $(0,0)$. The d_{sz} parameter is thus the signed distance in the $s - z$ plane between the straight line and the point $(s = 0, z = 0)$ onto the $x - y$ plane.

As already briefly mentioned in [Section 2.2.1](#), the transverse momentum of a charged particle in a magnetic field is related to the radius R_c of curvature in the transverse plane via [Equation 2.9](#). With the following choices of units p_T in GeV/c , R in meter and B_z in Tesla, the equation can be rewritten as:

$$p_T = 0.29979 \cdot R_c \cdot B_z \quad (3.58)$$

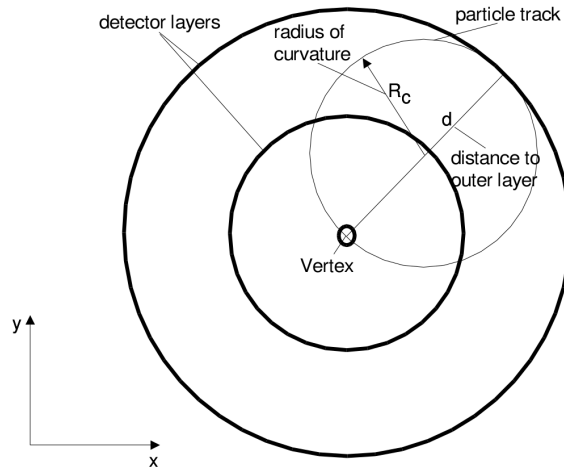


Figure 3.6: Definition of the minimum transverse momentum a full circle linking the origin and the outermost measurement.

where B_z is the magnetic field component on the z -axis. p_T can be determined by fitting a circle to (x, y) measurements of a set of points along the track in the plane transverse to the magnetic field. The minimum p_T is defined by the radius of a full circle linking the origin and the outermost measurement on a specific detector with a given distance d to the origin (Figure 3.6)

$$p_T = 0.29979 \cdot \frac{d}{2} \cdot B_z \quad (3.59)$$

In the specific case of CMS, considering a distance of 1 m between the vertex region and the outermost layer of the particle and a magnetic field strength of 3.8 T, all charged particles with $p_T \lesssim 600$ MeV will never reach the calorimeter region. Tracks in this range of p_T will bend back into the tracker and are called *loopers*.

The maximum p_T is defined using the *sagitta's method*. The sagitta, s , is defined in Figure 3.7 where the following relation holds:

$$\left(\frac{d}{2}\right)^2 + (R_c - s)^2 = R_c^2 \quad \rightarrow \quad R_c = \frac{\left(\frac{d}{2}\right)^2 + s^2}{2s} \quad (3.60)$$

The sagitta is thus related with the particle p_T in the following way:

$$p_T = 0.29979 \cdot \frac{\left(\frac{d}{2}\right)^2 + s^2}{2s} \cdot B_z \quad (3.61)$$

Assuming that the sagitta is equal to the pitch in the middle layer ($s \simeq 100 \mu\text{m}$), the maximum p_T of a particle reconstructed with the CMS tracker detector is of the order of 1.5 TeV.

3.2 Track reconstruction during Run 1

As mentioned in Section 2.1.2, during Run 1 the LHC operated with an instantaneous luminosity of $3.9 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ and with the proton bunches crossing at intervals of 50 ns. In these conditions, the CMS tracker is expected to be crossed by about 1000 charged particles produced by more than 20 collisions per bunch crossing on average. In order to obtain a precise and efficient measurement of charged particle momentum in this environment, CMS installed the largest all-silicon tracker ever built accompanied by a fast and robust track reconstruction software. The CMS tracker consists of a silicon Pixel Tracker close to the interaction point and a Strip Tracker surrounding it. It is described in detail in Section 2.2.1. A summary of the principal characteristics of the tracker is given in Table 3.1.

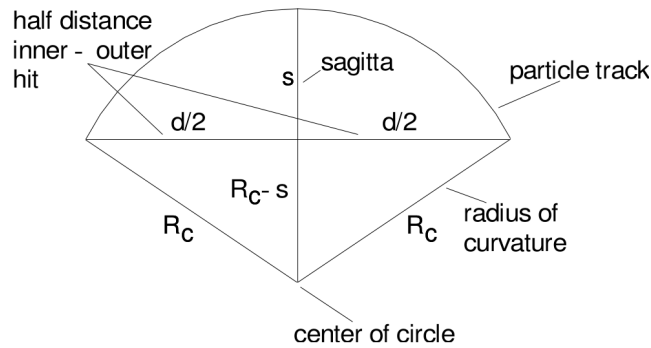


Figure 3.7: Definition of the sagitta as the distance of the midpoint of the arc to the chord defined by two points of the circle.

Table 3.1: Summary of the principal characteristics of the Pixel Tracker and the Strip Tracker.

Subsystem	Layers	Pitch	Location in CMS
PXB	3 cylinders	$100 \times 150 \mu\text{m}^2$	$4.4 < r < 10.2 \text{ cm}$
TIB	4 cylinders	$80 - 120 \mu\text{m}$	$20 < r < 55 \text{ cm}$
TOB	6 cylinders	$120 - 180 \mu\text{m}$	$55 < r < 116 \text{ cm}$
PXF	2 disks	$100 \times 150 \mu\text{m}^2$	$34.5 < z < 46.5 \text{ cm}$
TID	3 disks	$100 - 140 \mu\text{m}$	$58 < z < 124 \text{ cm}$
TEC	9 disks	$100 - 185 \mu\text{m}$	$124 < z < 282 \text{ cm}$

A summary of the software algorithms developed for the CMS track reconstruction used during Run 1 is reported here. A more detailed description is given in Ref. [83] along with the complete list of Run-1 results obtained with a centre-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$.

3.2.1 Local reconstruction

The first step of the reconstruction process is the *local reconstruction* mainly performed in the local system of reference of each sensor. In the case of the Pixel Tracker the x -axis is oriented parallel to the shorter pixel edge, while in the case of the Strip Tracker the x -axis is chosen perpendicular to the central strip in each sensor. In the local reconstruction step, the zero-suppressed signals above specified thresholds are clustered into *hits* in both pixel and strips channels [84]. A first estimation of the hit position and its uncertainty is given by summing the charge collected in neighbouring sensors in the x -coordinate, corrected for the Lorentz drift due to the CMS magnetic field and for a possible noise contribution. In the case of the double-sided modules in the Strip Tracker described in Figure 2.2.1, the information coming from the two hits, known as ‘ $r\phi$ ’ and ‘stereo’ hits, can be combined in so-called *matched hits*. These specific layers can therefore provide a measurement of the second coordinate, z in the barrel and r in the endcap. In order to assess the local reconstruction performance, the hit efficiency and the hit resolution are evaluated using MC simulation. The hit efficiency is defined as the probability to find a cluster in a given silicon sensor that has been traversed by a charged particle. For Run-1 simulated data, the hit efficiency is more than 99% both in the Pixel Tracker and in the Strip Tracker when excluding the defective components (Figure 3.8(a,b)). All defective modules of the tracker are taken into account both in the simulation of the detector and in the reconstruction of tracks. The hit resolution is defined using the measurement residuals, i.e. the difference between the measured and the expected hit position predicted by the fitted track. To minimize possible biases, each trajectory is refitted not using the information coming from the hit under study. The measured pixel hit resolution is about $9.4 \mu\text{m}$ in the $r\phi$ coordinate, while it is shown in Figure 3.8(c) in the longitudinal direction for both data and simulation as a function of the track-sensor angle. The best pixel hit resolution

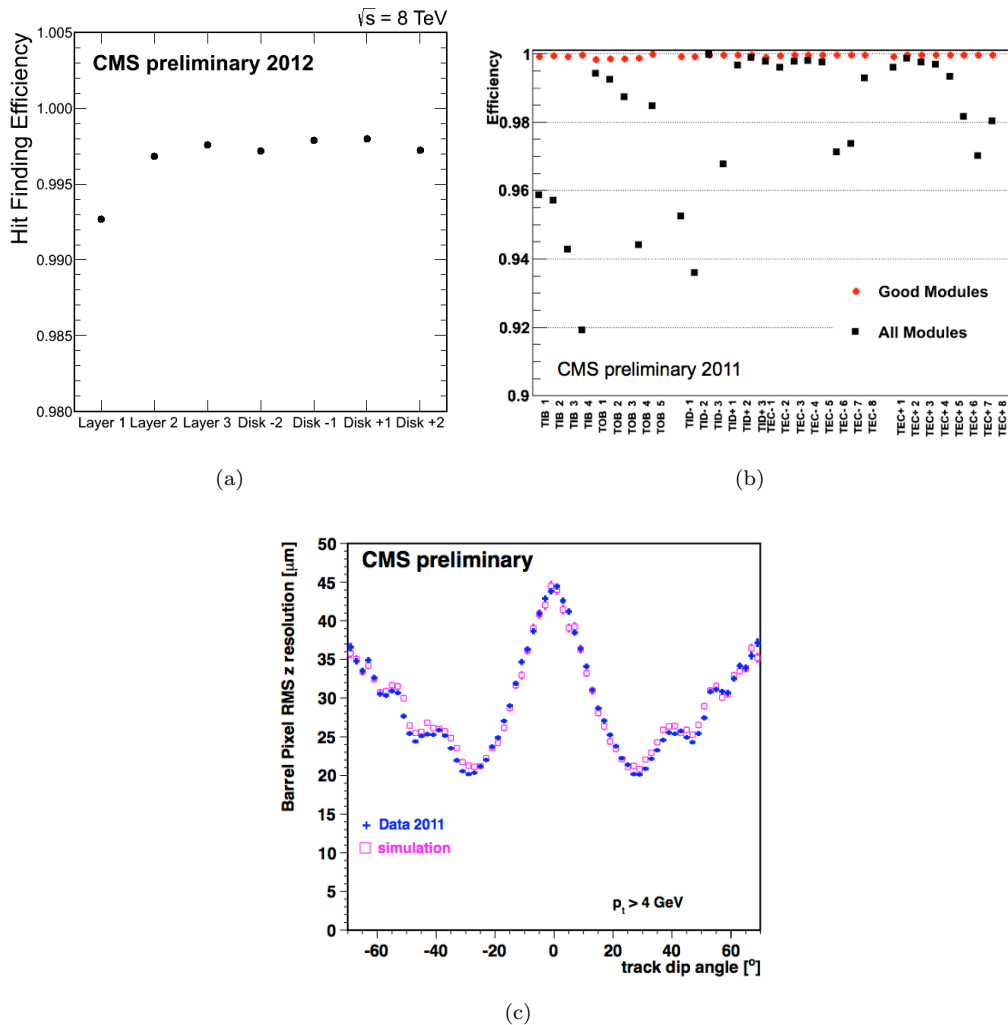


Figure 3.8: Average hit efficiency for layers or disks in the Pixel Tracker (a) and in the Strip Tracker (b) excluding defective modules. In the Strip Tracker plot the black squares show the hit efficiency in all modules, while the red dots for modules included in the readout. (c) Measured pixel hit resolution as a function of the track-sensor angle for data (blue crosses) and simulation (purple squares).

is obtained for incident angles around $\pm 30^\circ$, where the clusters are longer and thus the charge sharing improves the resolution.

3.2.2 Iterative tracking

The CMS software reconstructs tracks by using an adaptation of the combinatorial Kalman Filter [85], an extension of the **KF**, to allow pattern recognition and track fitting to occur in the same framework. It is referred to as **Combinatorial Track Finder (CTF)**. The *collection* of reconstructed tracks is produced in several iteration of the **CTF**. The entire track reconstruction process is called *iterative tracking*. The basic idea is to search for the easiest tracks first, to exclude from consideration the hits found in previous iterations (*masking*), and to continue with the next iteration. In this way the combinatorial complexity of the tracking problem is reduced and simplified for more difficult classes of tracks.

Each iteration can be divided in four different steps:

1. The **seed generation** step provides the starting trajectory parameters and the associated uncertainties. A first estimation of the 5-components of the state vector can be given using

either three 3-D hits¹ (triplets) or two 3-D hits (pairs) with a constraint on the trajectory origin. The seed is generated using two main parameters: the *seeding layers*, which correspond to the two or three detector layers where the hits are searched; and the *tracking regions*, which limit the possible number of seeds specifying the limits on the acceptable seed parameters. Seeds composed of hit pairs are found by combining hits in different layers of the detector. In case there are three seeding layers and a pair is already found in the two inner layers, a triplet is created by extending the hit search to the outer detector layer. In both cases, the compatibility of the seed parameters within the tracking region requirements is checked. A possible example of tracking regions can be a constraint on their minimum p_T or the consistency with the hypothesis that the track originates in the beam spot region. In most of the iterations, seeds are formed in the Pixel Tracker for the following reasons. Firstly, the high granularity of the pixel detector ensures a lower average channel occupancy in spite of the higher track density, as can be seen in [Figure 3.9](#). Moreover, the three-dimensional pixel measurements provide more constraints and more precise estimates of the trajectory parameters. Finally, generating seeds in the innermost part of the tracker leads to a higher efficiency for reconstructing tracks: the effect of the particle-material interaction is minor in this region and it facilitates reconstruction of low-momentum tracks that are then deflected by the strong magnetic field before reaching the outermost tracker region. However, seeds are also built in the Strip Tracker to find tracks produced outside of the pixel volume, such as tracks produced after a particle decay. In this case, only matched hits are used to improve the speed and quality of the seeding algorithm.

2. In the **track finding** step the **KF** algorithm is run for each seed. It begins using the coarse estimate of the track parameters provided by the trajectory seed to extrapolate the state vector to the next layer and then builds track candidates by searching for other hits in adjacent layers of the detector. If the found hits are considered to be compatible with the trajectory, the algorithm includes them in the track and updates the respective trajectory parameters. A χ^2 -test with respect to the extrapolated trajectory is used for the compatibility check. If a hit is not created in a module, for example due to module inefficiency, a configurable parameter provides the possibility of adding a *ghost hit* to the track and continue the search. To avoid a rapid increase in the number of candidates, only a limited number of them² are kept at each step, with the best candidates chosen based on the normalized χ^2 , a bonus given for each valid hit, and a penalty for each ghost hit. The track finding step for a specific track candidate continues until either there are no more compatible layers, or more than one missing hit is found, or its p_T drops below a certain value.
3. After the hits are collected and a first estimation of the track parameters is given, a **track fitting** step is performed using a **KF** or **GSF** filter/smoother to obtain the final estimate of the track parameters at the interaction point exploiting the full trajectory information. First, the **KF** algorithm is run in the innermost hits of the track and the corresponding covariance matrix is scaled up by a large factor in order to limit the biases. Then the fit proceeds on the full list of hits updating the estimate of the track trajectory iteratively: The hits are included from the innermost to the outermost. It is important to notice that for each valid hit, the estimated hit position and uncertainty can be reevaluated using the current values of the track parameters³. Once this first filter is run, a second one is initialized with the results obtained and run backward towards the beam-line. From the weighted average of the track parameters of these two filters, the optimal parameters are found for each surface associated with any track hits. During this step, matched hits are split as two single hits to improve track resolution.
4. The last step is the **track selection**. The main idea is to introduce quality cuts on several track variables to reduce the number of tracks produced by random combination of hits. The selections take into account the number of layers that have hits, the quality of the estimated track parameters using the χ^2 of the track, and the compatibility of the track with respect

¹Any hit that provides a 3-D position measurement.

²The default number of candidates is five.

³In the case of the pixel hits, both position and errors are reevaluated, in the case of strip hits just the uncertainty.

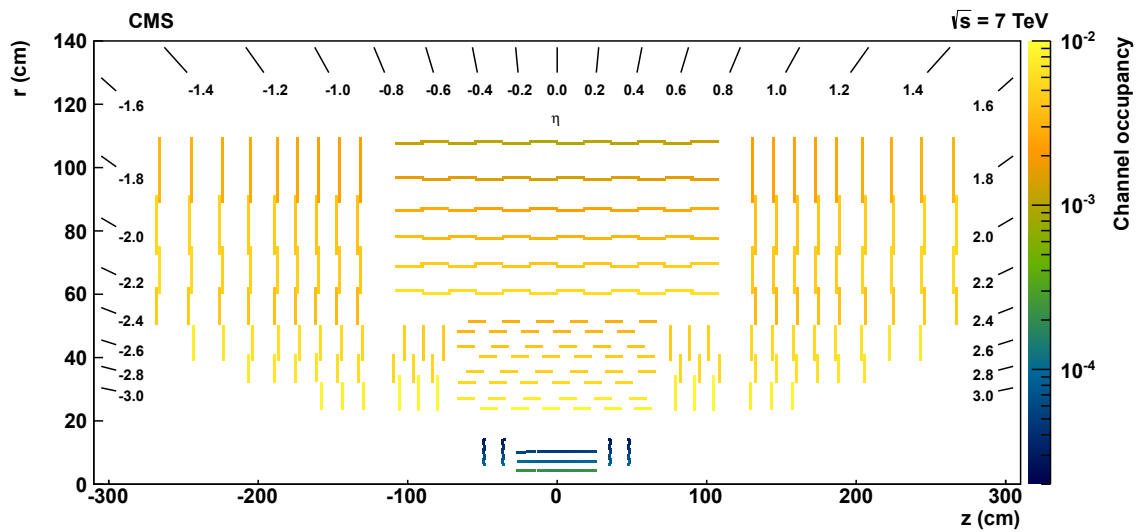


Figure 3.9: Channel occupancy (labeled by the scale on the right) for CMS silicon detectors in events taken with unbiased triggers with an average of nine pp interactions per beam crossing, displayed as a function of η , r , and z .

to the signal vertex. If more than one signal vertex is found, the compatibility with all vertices is considered. To optimize the performance, several requirements are also imposed as a function of other variables, such as the pseudorapidity and the p_T of the track. The entire list of variables and the respective quality cuts used during Run 1 is given in Ref. [83]. The quality cuts are used to set the track quality flags: *Loose*, *Tight* or *High Purity*. The *Loose* criteria corresponds to the minimum requirements for a track to be kept, while *Tight* and *High Purity* provide progressively more stringent requirements to enhance the purity of the sub-set of tracks. Most of the results reported in this thesis use the *High Purity* quality cut, which is the one used in most of the physics analyses.

Three different propagators are used in the track reconstruction procedure depending on the level of accuracy needed. The *analytical propagator* is the fastest one and assumes a uniform magnetic field without taking into account multiple scattering or energy loss. It is used, for example, to determine which adjacent layers of the detector can be intersected by the trajectory. The *material propagator* differs from the analytical propagator because it uses the predicted RMS scattering angle to increase the uncertainty in the trajectory parameters. Moreover, it also adjusts the momentum of the trajectory by the predicted mean energy loss using the Bethe-Bloch equation. It is used, for example, to search for compatible hits and update the trajectory. The *Runge-Kutta propagator* is the most precise propagator because takes into account the effect of material and the possible inhomogeneity of the magnetic field. It divides the distance to be extrapolated into many small steps. In this way, the extrapolation of the track trajectory can be solved numerically with fourth order accuracy. This propagator is used in the filtering and smoothing procedure to obtain the result with the best precision.

In Table 3.2 the six tracking iterations developed for Run-1 data as described in Ref. [83] are listed along with the corresponding seeding configuration and the target tracks. After the last step of the last iteration, the tracks found by each of the six iterations are merged into a single collection.

The first three iterations use the precise 3-D hits provided by the pixel modules to reconstruct the seed of the track. The first iteration is designed to find the tracks produced by prompt particles⁴ with p_T more than 800 MeV, the second iteration recovers prompt tracks that have a pair seed, and the third one is configured to find low- p_T prompt tracks. The last three iterations use also the Strip Tracker information in the seed step to find tracks produced outside of the pixel volume and to recover tracks not found in the previous iterations. These iterations are particularly important

⁴A prompt track is a track originating close to the beam crossing region.

Table 3.2: List of different tracking iterations used in Run 1 with the corresponding seeding configuration used and target tracks.

Iteration	Step name	Seeding	Target tracks
0	Initial	pixel triplets	prompt, high p_T
1	PixelPair	pixel pairs	high p_T recovery
2	LowPtTriplet	pixel triplets	prompt, low p_T
3	MixedTriplet	pixel+strip triplets	displaced
4	PixelLess	inner strip pairs	displaced
5	TobTec	outer strip pairs	displaced

for B physics studies, photon reconstruction and for improving energy resolution for particle-flow reconstruction. Moreover, it allows to search for long-lived particles that decay into the tracker volume.

Later versions of the software used in Run 1 have the same structure, with different iterations and tuned values to adapt to the higher pileup conditions. Main changes of the iteration scheme targeting a specific set of physics scenarios or objects were developed during the [LS1](#) and are reported in [Section 3.3](#).

3.2.3 Performance

The performance of the track reconstruction is evaluated using two different simulated samples: isolated particles and proton-proton-like collision events. Comparing the results obtained for these two samples allows to study the effect of the interaction between different particle and detector material and to better understand the level of performance degradation due to the superimposed pileup. The first sample is used to assess the track finding efficiency as a function of different variables, such as the pseudorapidity and the p_T , and to determine the resolution of the estimated track parameters. In Ref. [\[83\]](#) the Run-1 tracking performance using simple events with just a single generated muon, charged pion or electron are described in detail. These samples are generated according to a flat distribution in η inside the tracker acceptance, i.e. $|\eta| < 2.5$. Their transverse momenta are either fixed to 1, 10 and 100 GeV or generated according to a flat distribution in $\ln(p_T)$. The second sample is used to re-create a typical proton-proton collision and consists of events with $t\bar{t}$ pairs. In this case a pileup of minimum-bias events is added, where the number of pileup events is drawn from a Poisson distribution with mean equal to eight.

Efficiency and fake rate

The reconstructed tracks are associated to simulated particles using the following definition: A reconstructed track is associated to a specific simulated particle if at least 75% of its hits originate from this simulated particle. If a reconstructed track is not associated with any simulated particle, it is considered as a combination of unrelated hits and designated as a *fake* track. The tracking efficiency is defined as the fraction of charged particles associated to at least one reconstructed track, while the fake rate is the fraction of fake tracks in the set of reconstructed tracks. Both efficiency and fake rate strongly depend not just on the algorithms and track selections used, but also on the intrinsic tracker properties, such as its layout, its material budget and the precision of its measurements. In order to select the prompt particles, further criteria are applied on the tracks contributing to the efficiency and fake rate computation. In both cases, tracks are selected if they are produced within the tracker acceptance and within 3.5 cm and 30 cm from the centre of the luminous region for r and $|z|$, respectively. Moreover, the distributions contain only tracks passing the *High Purity* requirement, given that it is the default track selection for most of the physics analyses in [CMS](#). All the performance plots in terms of efficiency and fake rate shown in this thesis contain these selections if not indicated otherwise.

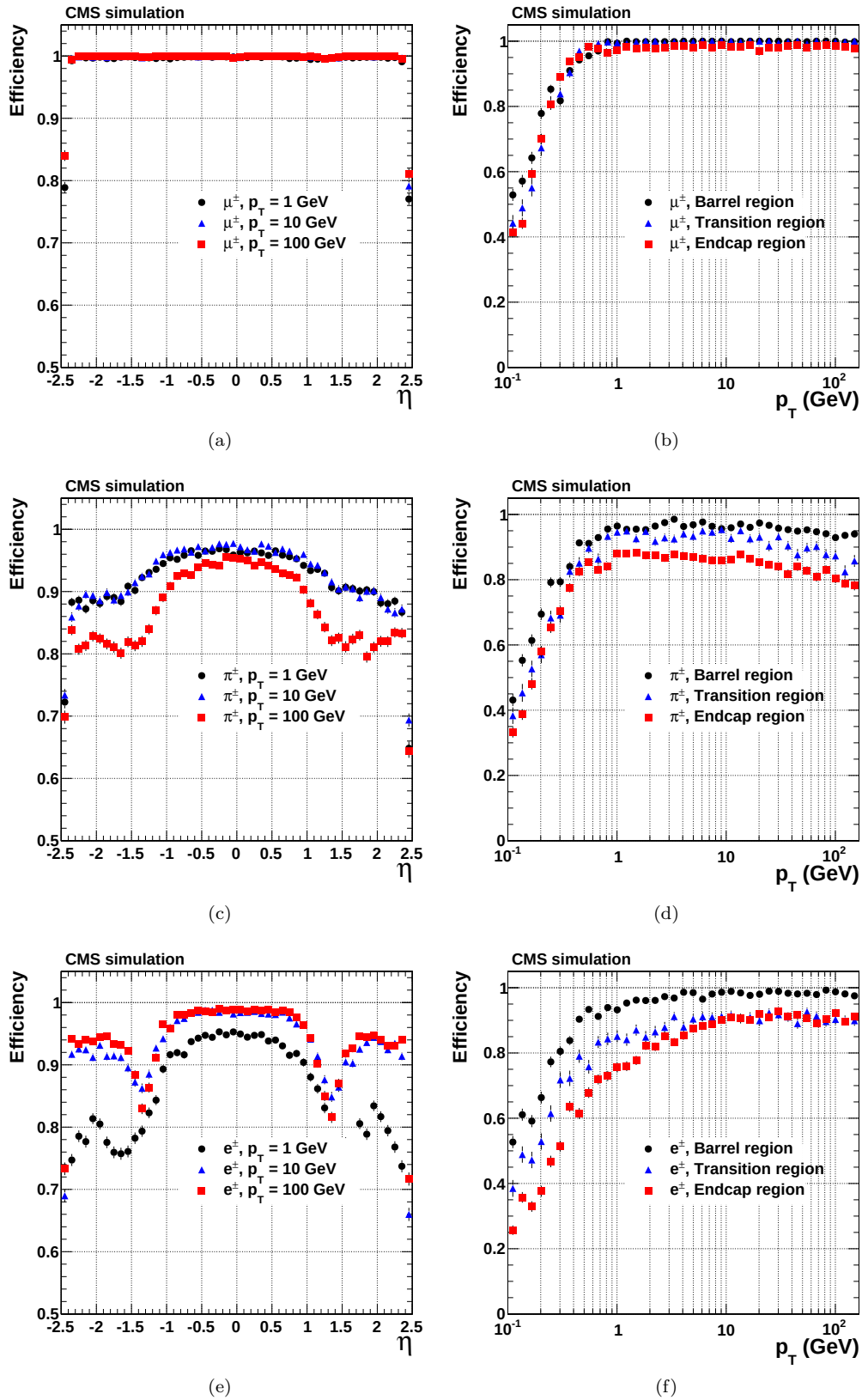


Figure 3.10: Track reconstruction efficiencies for single, isolated muons (a,b), pions (c,d) and electrons (e,f) as a function of η for $p_T = 1, 10, 100$ GeV and as a function of p_T for the barrel, transition, and endcap regions. The selections applied on these tracks are reported in the text [83].

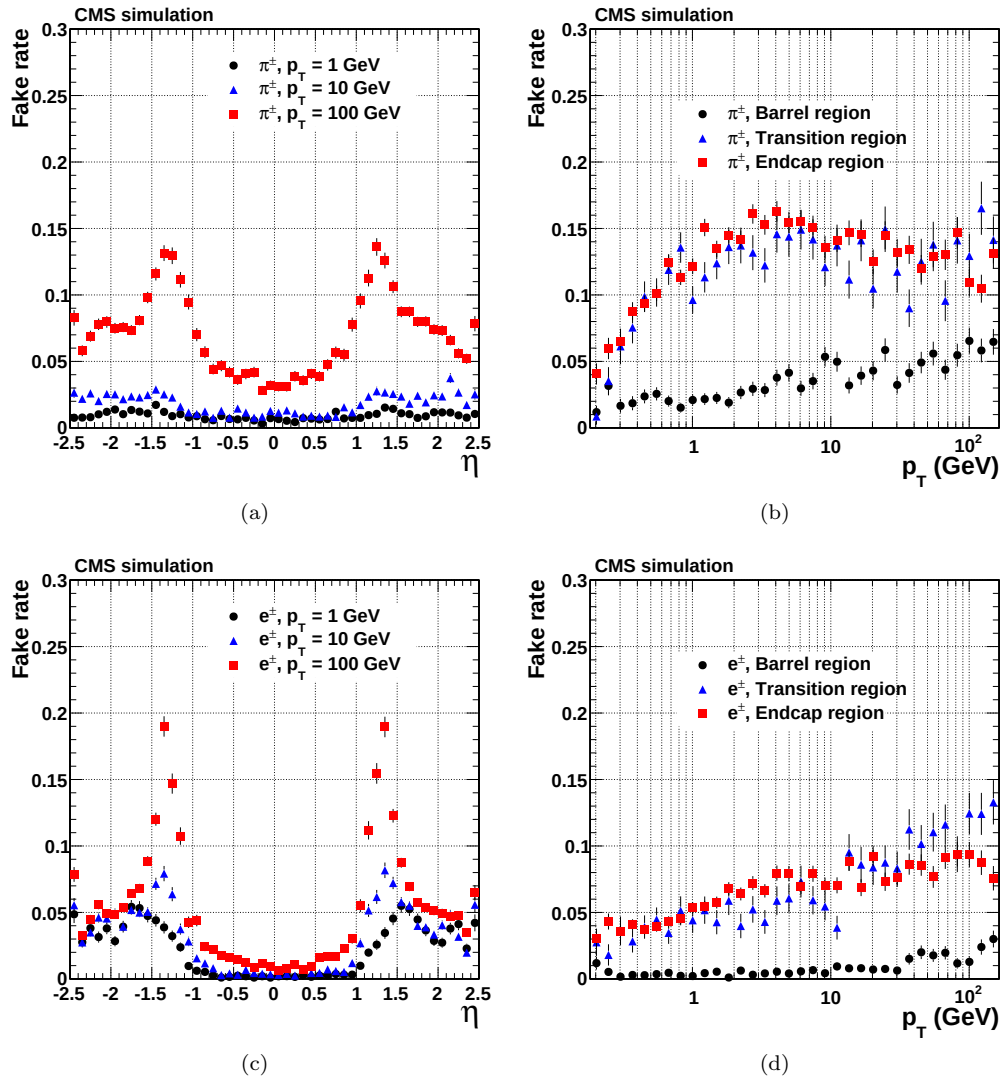


Figure 3.11: Tracking fake rate for single, isolated pions (a,b) and electrons (c,d) as a function of η for $p_T = 1, 10, 100$ GeV and as a function of p_T for the barrel, transition, and endcap regions. The selections applied on these tracks are reported in the text [83].

Figures 3.10 and 3.11 show the efficiency and fake rate for different samples of isolated particles as a function of the pseudorapidity and the p_T of the simulated particle and the reconstructed track, respectively. A further requirement on the p_T is introduced: $p_T > 0.9$ GeV for the study of efficiency as a function of the pseudorapidity, and $p_T > 0.1$ GeV for studying efficiency over the entire p_T spectrum.

In the case of isolated muons, the efficiency is almost 100% independent of the generated transverse momentum. The fake rate is completely negligible. This is due to the intrinsic nature of muons, which interact with the silicon detector through ionization and have negligible energy loss due to bremsstrahlung. Moreover, muon trajectories are affected almost exclusively by Coulomb scattering and energy loss, both relatively easy to include in the KF algorithm.

Charged pions interact with the detector not only through multiple scattering and energy loss, but they are also subject to elastic and inelastic nuclear interactions. The elastic nuclear interaction introduces long tails in the distribution of the scattering angle which are difficult to include in the KF algorithm. The charged pion track can be interrupted or the number of hits reduced bringing a clear degradation of the performance also in the estimation of the track parameters. Inelastic nuclear interactions are also another source of tracking inefficiency for hadrons, particularly in

those regions of the tracker with large material content. Due to these effects, the efficiency drops to 80% in the overlapping region, where the material budget is at the maximum, and does not fully recover in the endcap region. An inefficiency can also be seen for charged pions with $p_T < 700$ MeV because of the larger cross sections for nuclear interactions at low energies [5]. The efficiency is also affected for large momenta due to the nuclear interaction of the charged pions with the material which as a consequence produce other tracks mostly aligned with the charged pion track, and are thus difficult to discriminate from the real trajectory of the pion. The fake rate is also related to this effect and increases from around 2–3% for low- p_T pions up to 15% in the transition region for higher- p_T charged pions. In this tracker region, another degradation effect is correlated with the distances between successive hits on the track which are bigger here than in other tracker regions. This effect brings larger uncertainties in the extrapolation of the trajectory and therefore more spurious hits are incorrectly assigned to the track.

In the case of isolated electrons, the problems are once again different. As already mentioned in Section 2.3.2, the main source of energy loss for high-energy electrons is bremsstrahlung. The effect of the emitted radiation is very similar to the effect of the inelastic nuclear interaction described above for charged hadrons: the number of hits belonging to the electron track can be significantly reduced and, if a radiated photon converts, the hits produced by the electron-positron pair can be mixed up with the primary track hits. The latter effect is considered to be the principal source of misidentification of charge for electrons. The effect on the tracking efficiency and fake rate using the CTF algorithm is clearly visible. The efficiency for electrons drops from almost 100% in the barrel region to 90% in the endcap, with a significantly worse performance in the transition region because of the larger material budget. In this region the fake rate also increases for electrons with p_T equal to 100 GeV given the fact that secondary particles are emitted aligned with the primary electron. As already described in Section 2.3.2, the energy loss distribution is highly non-Gaussian and therefore a dedicated GSF algorithm in combination with the ECAL information is used to achieve better performance in the CMS electron reconstruction.

Figure 3.12 shows the efficiency and fake rate in the case of non-isolated charged particles. The challenge in the reconstruction stage is increased compared to isolated particles due to the large number of hits produced by the hundreds of primary particles and their interactions within the tracker volume. Another problem appearing in the $t\bar{t}$ sample is the effect produced by loopers, which move along a helix in the tracker without reaching the calorimeter region and increasing the possible hit combinatorics. Given the fact that most charged particles produced in the proton-proton collisions are hadrons, the same sources of inefficiency described for single isolated charged pions applies also to the $t\bar{t}$ results. The fact that the efficiency results for the non-isolated charged particles are very similar to the ones obtained for hadrons is overall very positive, because it indicates that the reconstruction algorithm is robust against the larger track multiplicity in the event. For high p_T particles, the efficiency drops due to the fact that charged particles in that p_T spectrum are mostly produced inside the core of collimated jets, where the reconstruction is more difficult due to the extremely high density of particles. The fake rate is found to be lower than 1% in the entire tracker pseudorapidity range and deteriorates for high- and low- p_T particle momenta due to two main effects: the large cross section for nuclear interactions at low energies, and the production of secondary particles by inelastic scattering, which are very close to the mother particle at higher energies. Efficiency and fake rate for $t\bar{t}$ events with and without pileup is shown in the figure. The result underlines the robustness of the algorithm: the presence of pileup degrades the efficiency and fake rate only for tracks with a $p_T < 1$ GeV.

Track parameter resolution

In the CMS convention the five parameters which fully describe the trajectory of the particle at the interaction point are listed in Section 3.1.3 with the corresponding unit of measurement. The resolution of the following parameters is presented: d_0 , the signed transverse distance from the beam spot; z_0 , the signed longitudinal distance from the beam spot; ϕ , the track angle in the transverse plane; $\cot \theta$, the cotangent of the polar angle; and p_T , the transverse momentum. The resolution of track parameters is studied using simulated isolated particles. It is evaluated using the distribution of track residuals and is defined as the half-width of the interval that contains 68% (or

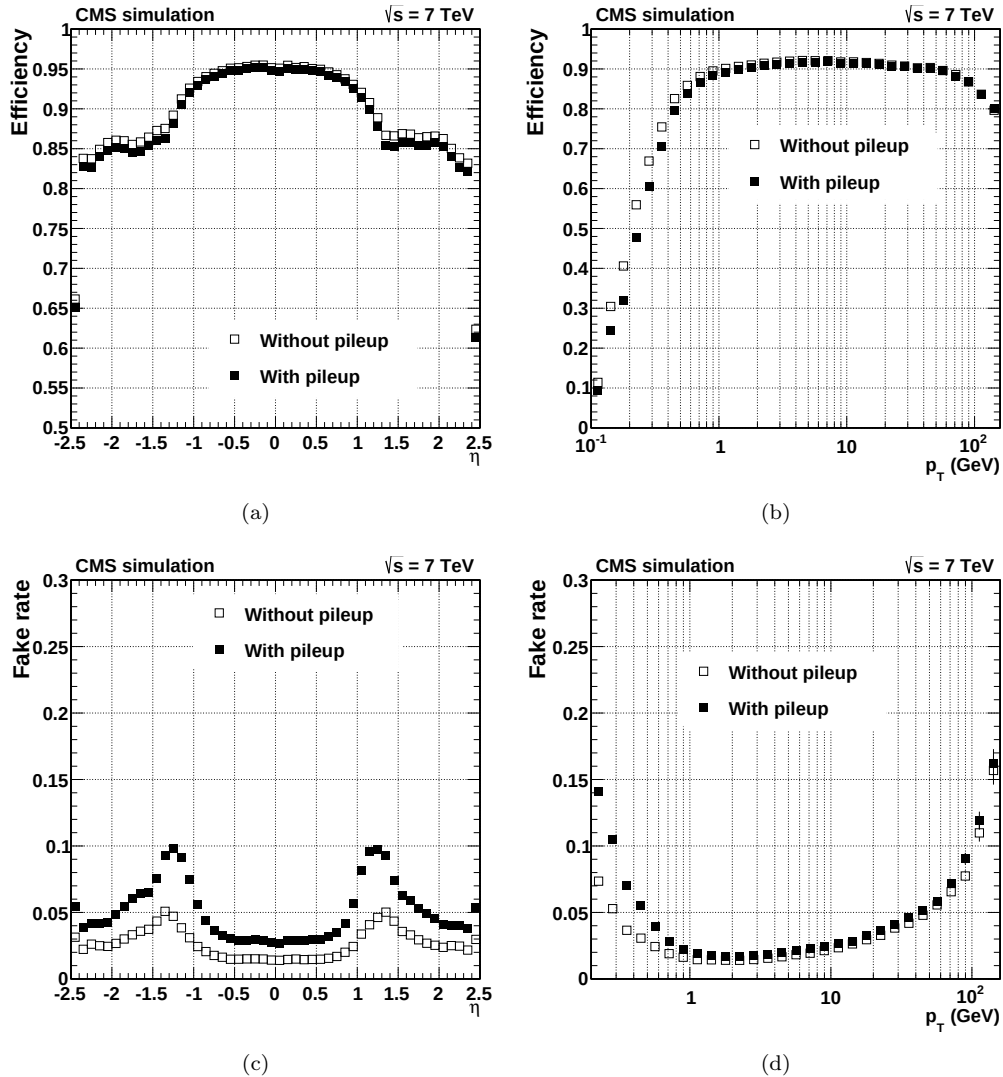


Figure 3.12: Tracking efficiency (a,b) and fake rate (c,d) for simulated $t\bar{t}$ events that include superimposed pileup collisions drawn randomly from a Poisson distribution with mean value of eight. The selections applied on these tracks are reported in the text [83].

90%) of all entries and that is centered on the most probable value of the residuals⁵. Figures 3.13 and 3.14 show the resolution of each of the five parameters for single muons plotted as a function of the η and the p_T of the simulated charged particle, respectively, with $p_T = 1, 10, 100$ GeV.

At high momentum, the resolution of the impact parameter is dominated by the position resolution of the innermost hit in the pixel detector; at lower momentum the resolution gets progressively worse because of multiple scattering. The deterioration for large pseudorapidity in both cases is explained by the larger extrapolation distance from the innermost hit to the beam axis. Moreover, the precision of the longitudinal impact parameter benefits from charge sharing in the pixel clusters for $|\eta| \leq 0.4$, and this effect is clear especially in the case of high p_T muons. The resolutions in the ϕ and $\cot \theta$ parameters are easily comparable to those of d_0 and z_0 , respectively, for very similar reasons. In particular, in the case of the polar angle, the representation of the resolution for $\cot \theta$ instead of θ is chosen because it is less dependent on the pseudorapidity. The p_T resolution is also found to be excellent for different p_T sample. In particular, the resolutions for the isolated muons generated with p_T equal to 1 and 10 GeV are well below the percent level in

⁵This value is extrapolated using the peak of a double-tailed Crystal Ball function fitted to the residuals.

the barrel region of the tracker detector, and its value reflects the amount of material traversed by the track interacting through multiple scattering. For $p_T = 100$ GeV muons the resolution rises to $2 - 3\%$. The relative precision in p_T is measured to be best for tracks with p_T around 3 GeV. In all p_T simulated sample, the deterioration of the resolution for larger η can be explained because of the shorter lever arm for the estimation of the curvature in the transverse plane. The degradation is even more accentuated in the region of $|\eta| \sim 1$ due to the gap between the barrel layers and the endcap disks.

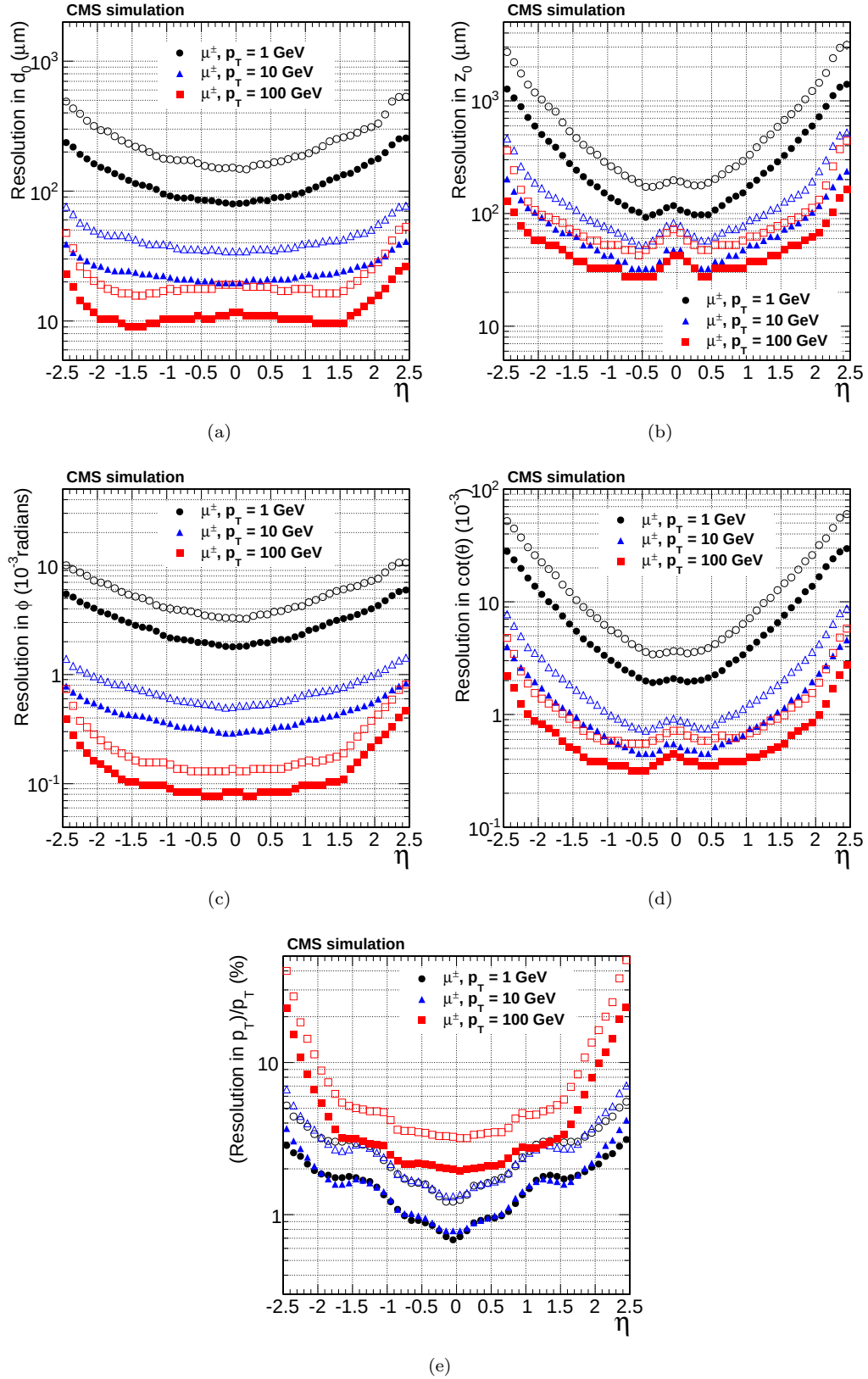


Figure 3.13: Resolution, as a function of pseudorapidity, in the five track parameters for single, isolated muons with generated p_T equal to 1 (black), 10 (blue) and 100 (red) GeV. From top to bottom and left to right: transverse and longitudinal impact parameters, ϕ , $\cot \theta$ and transverse momentum. The solid (open) marks correspond to the half-width for 68% (90%) intervals centered on the mode of the distribution in residuals [83].

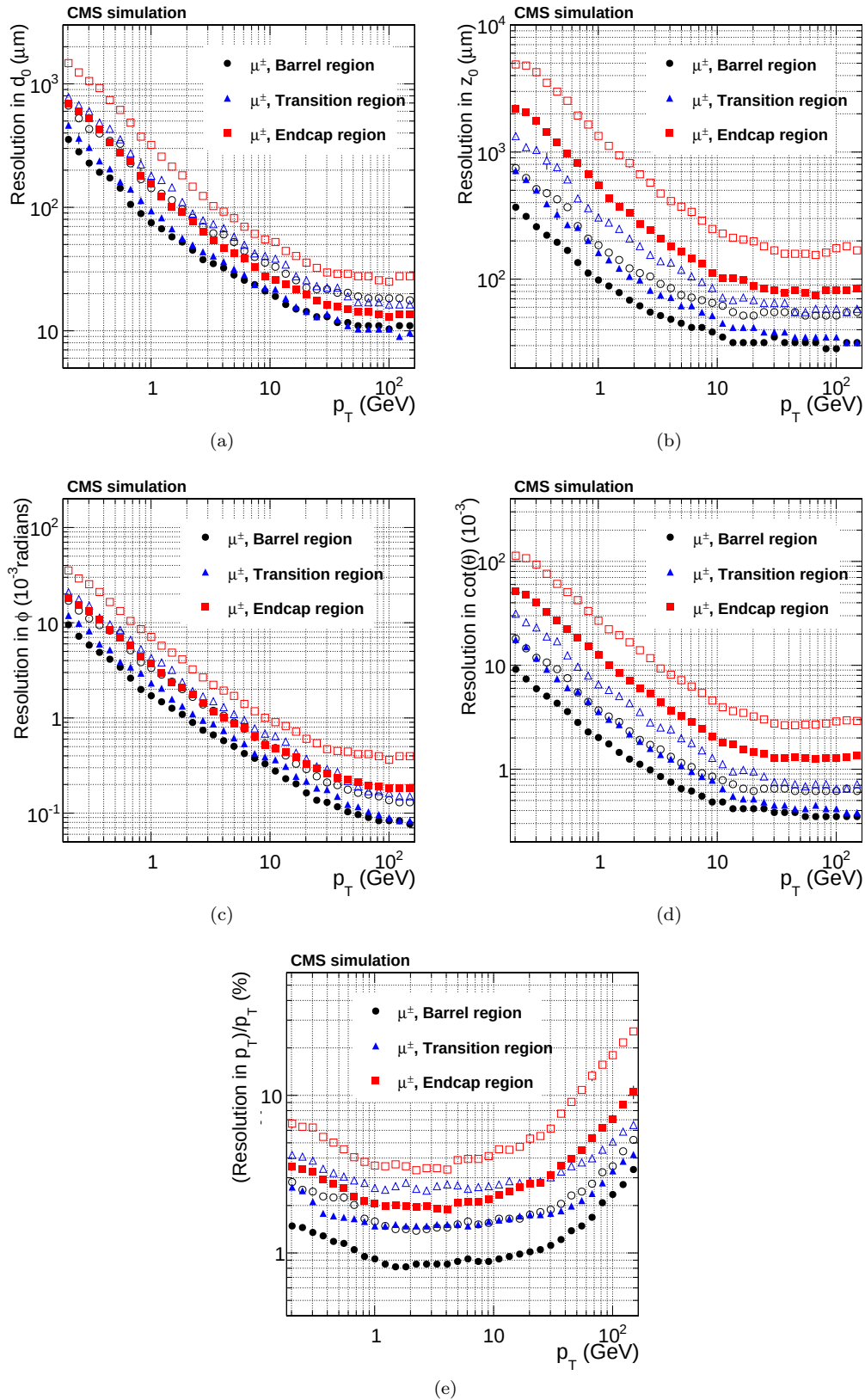


Figure 3.14: Resolution, as a function of the generated particle p_T , in the five track parameters for single, isolated muons with generated p_T equal to 1 (black), 10 (blue) and 100 (red) GeV. From top to bottom and left to right: transverse and longitudinal impact parameters, ϕ , $\cot \theta$ and transverse momentum. The solid (open) marks correspond to the half-width for 68% (90%) intervals centered on the mode of the distribution in residuals [83].

3.3 Developments towards Run 2

As described in [Section 2.1.2](#), the LHC upgrade during [LS1](#) improves the performance of the machine significantly with respect to Run 1:

- the centre-of-mass energy increases from 7 – 8 TeV to 13 – 14 TeV;
- the bunch time separation drops from 50 ns to 25 ns;
- the instantaneous luminosity scales from a peak of about $8 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ up to $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, double the amount expected from nominal conditions.

The difference of the pileup between the years 2012 and 2016 is shown in [Figure 3.15](#). It is supposed to rise further to a mean of 50 at the end of Run 2.

These new conditions and the subsequent rise in occupancy in the tracker pose a challenge for the track reconstruction because of its intrinsic combinatorial nature [\[63\]](#). Several problems are raised by this situation, and the iterative tracking technique is only a partial solution because the fraction of unmasked hits after all iterations is still above 50% already in the Run-1 pileup scenario. Aside from this, also several detector effects can compromise the high quality reconstruction of tracks. Firstly, when more than one track crosses a double-sided strip module, ambiguities may produce ghost hits. The number of ghost hits is strictly correlated with the pileup condition, and when the pileup achieves a mean of 40, the ghost rate in the innermost layer becomes larger than the rate of true hits. In this scenario, the effect on the time spent in reconstructing pairs seed in the Strip Tracker is clearly visible. Moreover, with 25 ns bunch crossing, charged particles coming from the out-of-time pileup increase even more the occupancy of the tracker. The effect in the Strip Tracker is about 45%, much larger than in the Pixel Tracker, which is only about 5%. Finally, the pixel detector is affected by dynamic inefficiency, which is mainly due to saturation of the chip readout buffer and reaches the 2% level in the first pixel barrel layer, already below the nominal instantaneous luminosity.

3.3.1 Physics- and timing-oriented developments

In order to cope with the new conditions in Run 2, some developments were made in different aspects of the reconstruction on top of the already high-quality software developed for Run 1 [\[63\]](#):

- Due to the increase in pileup, a loss of muon-track reconstruction efficiency in the tracker was observed during 2012 runs. In order to recover this loss, two new iterations dedicated exclusively to muons were added: an outside-in track reconstruction step seeded in the muon system, and an inside-out iteration that re-reconstructs muon-tagged tracks with looser requirements. In this way the efficiency is fully recovered, as can be seen in [Figure 3.16](#).

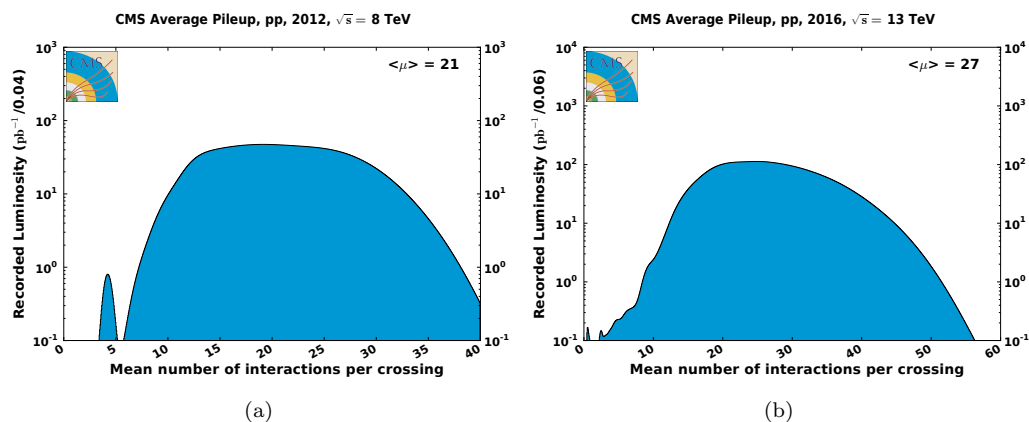


Figure 3.15: Mean number of interactions per bunch crossing for the year 2012 (a) and 2016 (b).

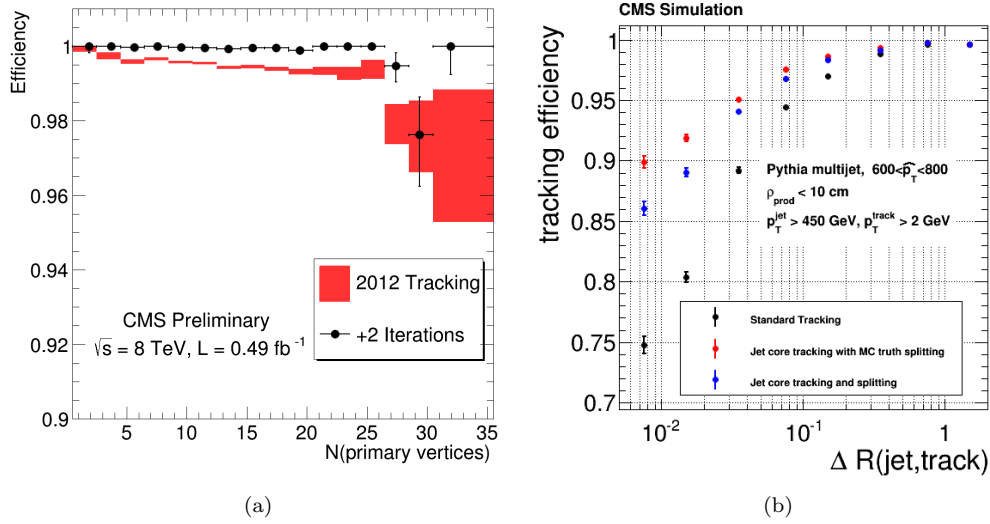


Figure 3.16: (a) Muon reconstruction efficiency in the silicon tracker with (black points) and without (red bands) the introduction of the two additional muon-specific iterations for part of the dataset collected in 2012. The efficiency is computed for all tracks passing a *Loose* selection with either $p_T > 2$ GeV or $|\eta| > 1$ and $p > 2$ GeV to ensure they reach the muon system [63]. (b) The track reconstruction efficiency as function of the ΔR between the geometrical centre of the jet and the track considered. Different scenarios are presented: standard tracking (black), jet core tracking with a cluster splitting based on **MC** truth information (red), and jet core tracking with new splitting algorithm (blue) [63].

The data efficiency is computed using the tag and probe method described in detail in [Section 3.4.1](#). Moreover, the muon momentum reconstruction is more precise and the identification is more robust. Given the very different nature of these two iterations with respect to the ones using only the information coming from the tracker, the tracks that are seeded only using the silicon detector are kept in a specific collection, called **Tracker-only seeded** collection.

- Further physics-oriented developments are also introduced to enhance the performance of the reconstruction inside jet cones. Firstly, a new iteration is included which targets jets with $E_T > 100$ GeV and searches for tracks around the calorimeter jet direction. This new iteration starts reconstructing the seed inside of a cone with $\Delta R = 0.1$ around the jet direction. The seed consists of a pair of hits either only belonging to the Pixel Tracker or also in combination with the first two layers of the **TIB**. The seed must be also compatible with the proton-proton interaction region. The pattern recognition step in this iteration allows using an order of magnitude more possible track candidates than in the other iterations. A pixel cluster splitter is also introduced in order to split merged clusters in the core of the jets. It exploits the information of the jet direction to predict the expected shape and charge of the clusters in the Pixel Tracker. Moreover, this splitter is based on a k-means like clustering which takes also into account the average expected charge per cluster. Thanks to these two developments, the track reconstruction efficiency inside jet cores has risen by about 20% as can be seen in [Figure 3.16\(b\)](#).
- In order to reduce the time spent in the iteration that uses the Strip Tracker information in the seeding step, a new seeding algorithm dedicated to this topic is developed. This new algorithm has the main goal of obtaining a purer seed collection with consequent reduction of time in trying to reconstruct tracks that can already be classified as fake and of estimating better the seed parameters and their uncertainties. The goal is achieved by moving from a pair to a triplet seed. The triplets are obtained with a straight line fit in combination with a dynamic re-estimation of the hit positions, using the seed direction hypothesis. The triplets are then selected on a χ^2 basis. The beam spot constraints are tightened as much as possible

while maintaining a stable physics output. In this way half of the seeds are rejected, but the same number of tracks is reconstructed.

- Part of the effort has also been put in reducing the fake tracks produced by out-of-time pileup. This is done by introducing a threshold on the cluster charge which takes into account the single sensor thickness as well as the trajectory crossing angle. The cut can be applied with increasing thresholds at different stages of the track reconstruction. The fake rate is thus reduced by $\simeq 50\%$ while keeping the efficiency unchanged.
- Many other technical improvements were also included to speed up the code and to optimize the performance. For example, the iterative tracking logic was globally optimized, the pixel dynamic inefficiency mentioned in Section 2.2.1 was included in the simulation, the tracking algorithms were successfully transitioned from the original single-threaded application to a multi-threaded one capable of processing multiple events concurrently [86], and a multivariate technique was implemented in the track selection step.

In Figure 3.17(a,b) the efficiency and the fake rate for high p_T prompt tracks are shown for both Run 1 and Run 2 in nominal conditions. In Figure 3.17(c) the iterative tracking sequence timing is plotted as a function of the pileup. The success of the developments is clear: the expected performances are very similar, or even better, for the second run and a large time reduction is obtained with respect to Run 1. The gain in time is almost twice for lower pileup scenario and up to four times for pileup around 70.

3.3.2 Deterministic Annealing Filter results

As described in Section 3.1.2, the standard KF does not have any protection against the wrong assignment of a hit to a track. This can be a problem, especially in high density tracks environment or in noisy environments. A possible way to improve the track reconstruction is to use the **Deterministic Annealing Filter (DAF)** algorithm. Its main characteristics can be summarized as follows:

- the hit-to-track-assignment is no longer binary but is expressed by an assignment probability $p \in [0, 1]$;
- a single hit can be assigned to one or several tracks;
- the computation of the weights on a particular sensor surface is based on the residual of the hit;
- the update of the estimated state vector and the covariance matrix on a detector depends on all the competing hits;
- different iteration cycles are run in order to reach the global minimum. For each cycle the annealing factor is changed and, as a consequence, the assignment probability of the track.

Thanks to these features, the wrong assignment of a hit to a given track can be avoided. As already mentioned in Section 3.1.2, a possible way to implement the DAF algorithm is to create a MultiRecHit object which incorporates the competing hits. In a first step, all the trajectories reconstructed in the event by using the iterative tracking algorithm are refitted with the DAF using exactly the KF formulas for the update of the state vector where the different competing hits are replaced with a MultiRecHit. For each annealing step, a full KF fitting-smoothing is applied, the weights are recomputed with the new annealing factor and the MultiRecHits are updated. At the end of the annealing program the isolated hits with very low weight are replaced with invalid hits and the fit is repeated for the last time. The tracks refitted with the DAF algorithm are then collected in a different collection to perform validation and studies.

Different samples are used to assess the correct working principles of the DAF: single muons events generated with p_T equal to 10 GeV, $t\bar{t}$ events and QCD jets events generated with different p_T ranges. Figure 3.18 shows the number of hits inside a MultiRecHit and the weight distribution of the competing hits if the MultiRecHit has more than one hit component. As expected, the number

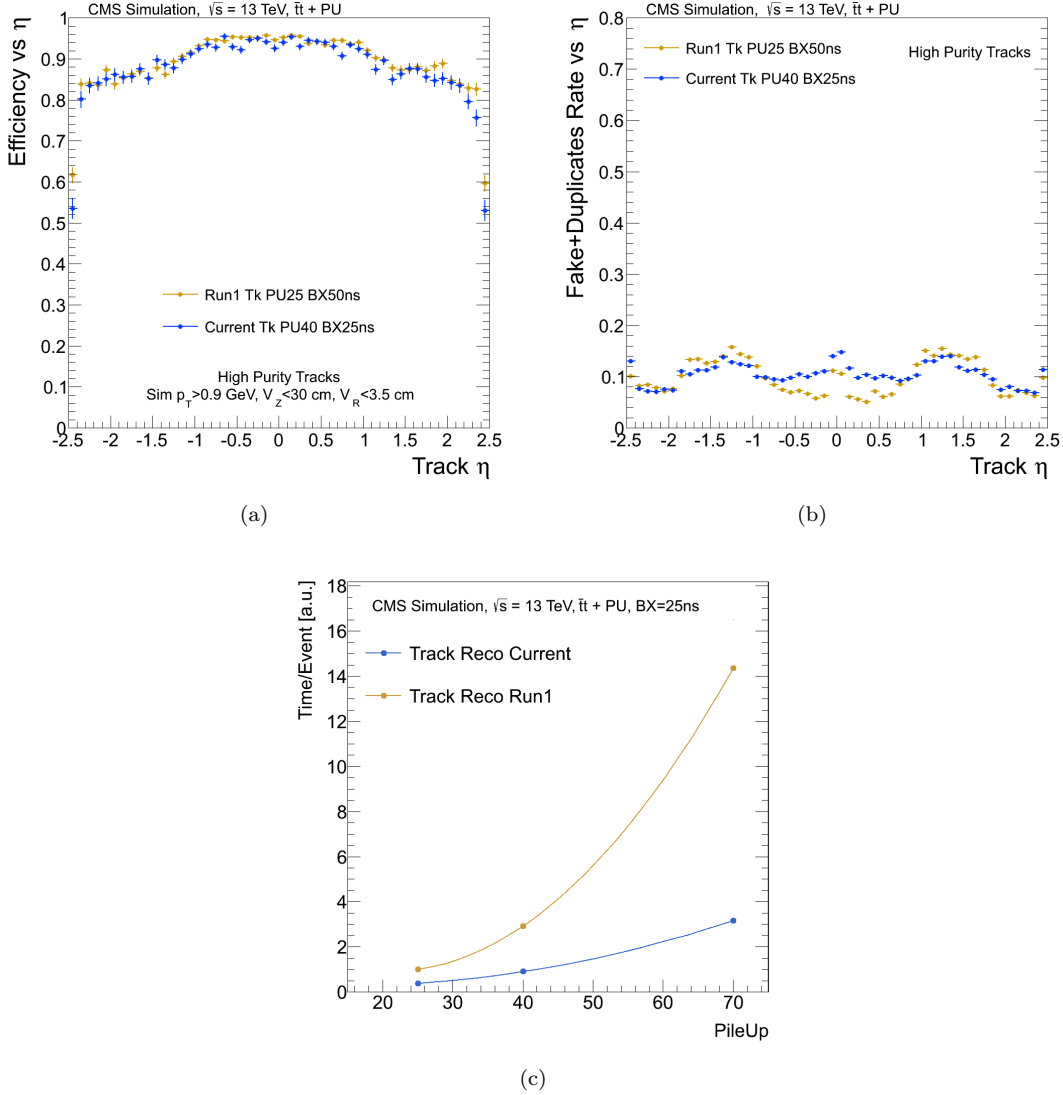


Figure 3.17: Track reconstruction efficiency (a), fake rate (b) and timing of iterative tracking sequence for Run 1 and Run 2 nominal condition in color yellow and blue, respectively [63].

of competing hits in the MultiRecHits increases with higher density tracks and the distribution of the weights is peaked at 0 and 1. The performance in terms of efficiency, fake rate, pulls and resolution are evaluated in all samples and are very similar with respect to the results obtained with the current reconstruction. The efficiency and fake rate as a function of the pseudorapidity of the track and of the ΔR between the track and its nearest neighbour is shown in Figures 3.19 and 3.20 for QCD jets events with $p_T \in [80, 120]$ GeV and $\in [3000, 35000]$ GeV, respectively. As expected, these results show a small improvement in the fake rate for the sample containing higher- p_T jets without any loss in terms of efficiency.

3.4 Run-2 results using data and simulation

In order to perform physics analyses on real data, it is essential that all relevant physical processes are well modeled by the simulation. The simulation reproduces the conditions during data taking in the most accurate way possible. It includes for example complicated effects of the detector geometry, a detailed representation of the magnetic field in the tracker region, the tracker alignment,

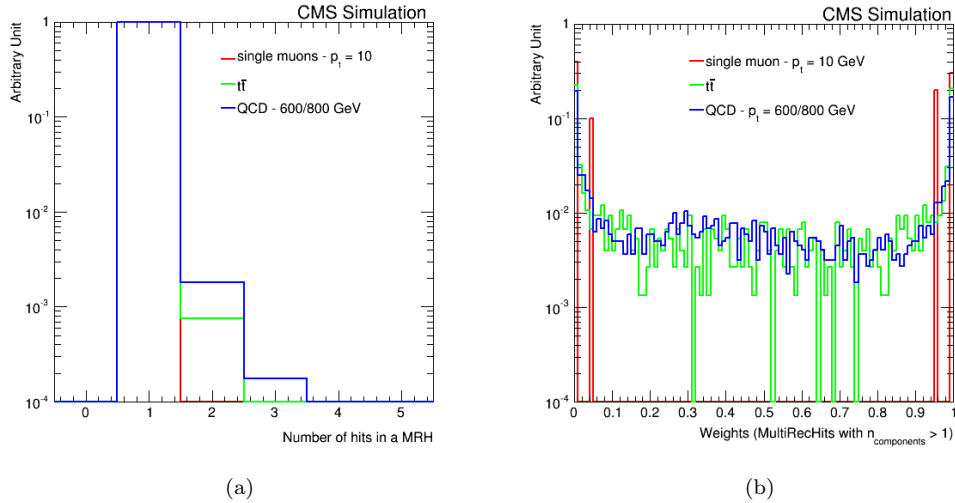


Figure 3.18: Distribution of the number of competing hits inside a MultiRecHit for different samples (a) and final distribution of weights for competing hits selected by the **DAF** algorithm (b). Both distributions are shown for different samples: single muon (red), $t\bar{t}$ (green) and **QCD** jets (blue).

etc. In the first part of this section, real collision data are used to validate the **MC** simulated data. This task is performed in **CMS** by directly comparing the particle reconstruction efficiency in real data with the efficiency predicted by the **MC** simulation. Both direct and indirect measurements are used. A brief summary of some of the methods is given using data collected in 2016. In the second part of this section, the commissioning of the tracking used after the upgrade of the Pixel Tracker in the year-end technical stop of 2016/2017 is presented. A description of the upgrade of the Pixel Tracker can be found in [Section 2.2.1](#).

3.4.1 Validation of simulation using 2016 data

Indirect Measurements

An example of an indirect measurement is the relative tracking efficiency for charged pions [87]. It is computed using the ratio of neutral charm-meson decays to final states of four or two charged particles. Two decay chains from $D^* \rightarrow D_0 \pi_{\text{slow}}$ are reconstructed in data and **MC**:

- $D_0 \rightarrow K^- \pi^+ \pi^- \pi^+$ as a “four-body final state”, labeled as $K3\pi$
- $D_0 \rightarrow K^- \pi^+$ as a “two-body final state”, labeled as $K\pi$

Taking into account that the two signal yields are proportional to the production cross section, the total integrated luminosity of the data collected, the branching fraction of the final state, and the selection efficiency, and assuming that the decay chains are well represented in the **MC** samples, the ratio R between their branching ratios can be measured using the formula

$$R = \frac{N_{K3\pi}}{N_{K\pi}} \frac{\epsilon_{K\pi}}{\epsilon_{K3\pi}} \quad (3.62)$$

where N is the number of events and ϵ is the selection efficiency in the respective decay chain. If it is assumed that the main difference in the selection efficiency between the $K\pi$ and $K3\pi$ processes is the reconstruction efficiency of the two π s, the following relations hold

$$\frac{N_{K3\pi}}{N_{K\pi}} \propto R_{\text{ref}} \cdot \epsilon_{\text{DATA}}^2 \quad \frac{\epsilon_{K\pi}}{\epsilon_{K3\pi}} \propto \frac{1}{\epsilon_{\text{MC}}^2} \quad (3.63)$$

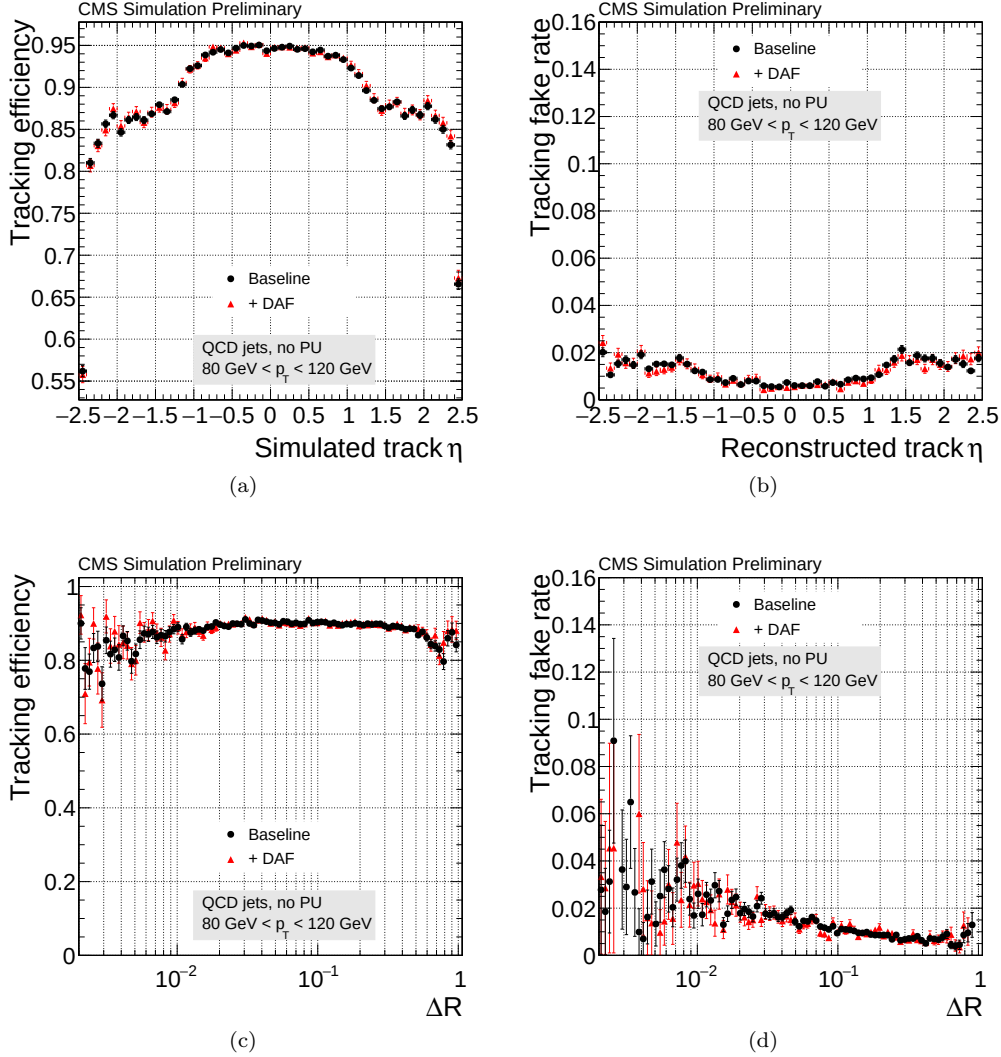


Figure 3.19: Track reconstruction efficiency as a function of simulated track pseudorapidity (a) and of the ΔR between the generated particle and its nearest neighbour (c) using the baseline track reconstruction (black) and refitting with the **DAF** algorithm (red) for **QCD** jets events with $p_T \in [80, 120]$ GeV. Track reconstruction fake rate as a function of reconstructed track pseudorapidity (b) and of the ΔR between the generated particle and its nearest neighbour (d) using the baseline track reconstruction (black) and refitting with the **DAF** algorithm (red) for **QCD** jets events with $p_T \in [80, 120]$ GeV.

where ϵ_{DATA} and ϵ_{MC} are the hadronic reconstruction efficiency in observed and simulated data, respectively. The pion relative efficiency is thus found comparing R with the nominal value of $R_{\text{ref}} = 2.08 \pm 0.05$ [5]:

$$\epsilon_{\text{rel}} = \frac{\epsilon_{DATA}}{\epsilon_{MC}} \propto \sqrt{\frac{R}{R_{\text{ref}}}} \quad (3.64)$$

Figure 3.21(a) shows the ratio for charged pions as a function of the minimum D^* candidates p_T threshold. Good agreement between R and R_{ref} is measured in all three pseudorapidity ranges for 2016 data [88]. It implies that the efficiency computed using the **MC** correctly reproduces the data.

Another indirect measurement uses low-mass resonances such as $K_S^0 \rightarrow \pi^+\pi^-$ and $\Lambda_0 \rightarrow \pi^\pm p^\mp$. These particles have a narrow intrinsic width mainly dominated by the limited knowledge and modeling of the detector material, the magnetic field and the reconstruction algorithms used to fit

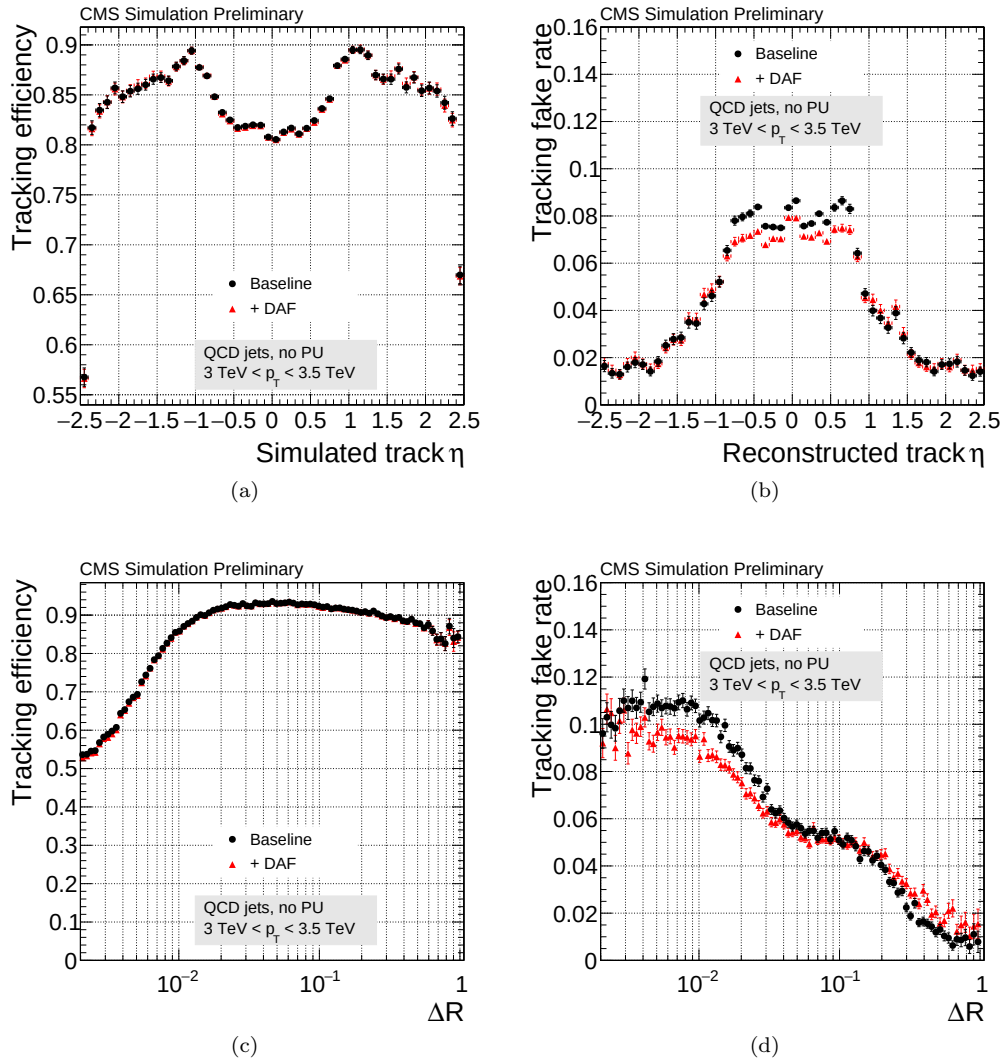


Figure 3.20: Track reconstruction efficiency as a function of simulated track pseudorapidity (a) and of the ΔR between the generated particle and its nearest neighbour (c) using the baseline track reconstruction (black) and refitting with the **DAF** algorithm (red) for **QCD** jets events with $p_T \in [3000, 3500]$ GeV. Track reconstruction fake rate as a function of reconstructed track pseudorapidity (b) and of the ΔR between the generated particle and its nearest neighbour (d) using the baseline track reconstruction (black) and refitting with the **DAF** algorithm (red) for **QCD** jets events with $p_T \in [3000, 3500]$ GeV.

the track trajectory. If these effects are fully understood, the systematic uncertainties which affect the momentum resolution and scale of charged tracks can be determined with good precision and the modeling of the detector and the track reconstruction can be improved. A study of the average mass as a function of different single track kinematics both in simulation and data has been carried out, and the entire set of results can be found in Ref. [89]. In Figure 3.21(b), the invariant mass of the K_S^0 particle as a function of the pseudorapidity is presented for data and simulation. The shape is well described by the simulation, and data and simulation agree at the level of few per mill in the tracker volume. The same conclusions apply also for the Λ_0 particle results.

Direct Measurements

An example of a direct measurement is the muon track reconstruction efficiency at the Z and at the J/ψ mass peak. The tag and probe method allows to measure the efficiency from the data itself,

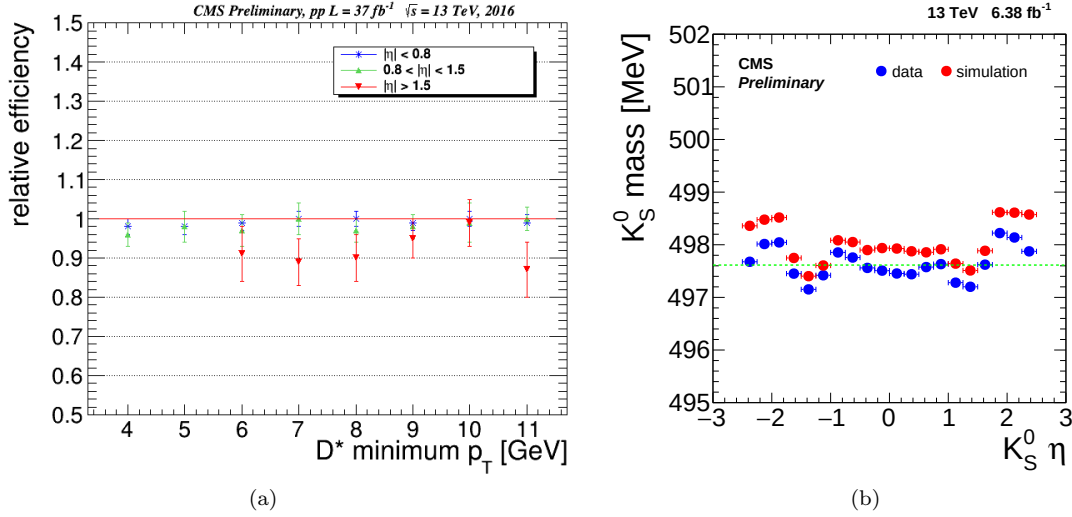


Figure 3.21: (a) Ratio of tracking relative efficiencies for charged pions as a function of the D^* minimum p_T for three different pseudorapidity regions: barrel (blue), transition (green) and endcap (red) [88]. (b) K_S^0 invariant mass as a function of the pseudorapidity for data (blue points) and simulation (red points) [89].

with no reference to simulation. Well-known resonances are reconstructed as pairs of particles, one of which has passed a tight identification, the so-called *tag*, while the other one has passed a loose identification, the so-called *probe*. Probes are then divided into two categories, depending on whether they pass or fail a specific selection criterion. The efficiency of the single object is then computed as the ratio of the number of the *passing probes* over the total number of probes in the sample.

In the specific case of this analysis, the muon-tracking efficiency is estimated using di-muon resonances, such as $Z \rightarrow \mu^+ \mu^-$ and $J/\psi \rightarrow \mu^+ \mu^-$, under the assumption that muon reconstruction in the silicon tracker and in the muon system is independent. As a starting point, only standalone-muons, defined in Section 2.3.1, are reconstructed with no silicon tracker information. The efficiency is thus defined as the probability that a track reconstructed in the silicon tracker is associated to a given standalone-muon probe. The following definitions are used:

- a tag muon is a global muon, defined in Section 2.3.1, above a certain threshold in p_T , associated to one of the two legs of the resonance and matched to a single muon trigger.
- a probe muon is any stand-alone muon with at least one valid hit.
- a passing probe muon is any probe muon that can be matched to at least one track.

The matching is defined by comparing the directions at the point of closest approach to the beam line of the two tracks. The matching is given in terms of the size of the cone in $(\Delta R, \Delta \eta)$ space drawn around the stand-alone muon. In the case of the Z and the J/ψ , the matching is chosen to be (0.30, 0.30) and (0.30, 0.15), respectively. The motivation of the choice for the J/ψ particle is discussed in more detail later in this section.

As described in Section 3.3, at the end of the iterative tracking process the tracks that fulfill some minimum requirements are kept in the track collection. Tracks that are seeded only using the silicon tracker are kept in the so-called **Tracker-only seeded** collection.

In both resonances, it is required that the invariant mass of the muon pairs is in a certain interval around the resonance mass, i.e. $70 < m_Z < 130$ and $2.3 < m_{J/\psi} < 4.3$. A simultaneous fit is performed on the invariant mass spectra of the pairs for both passing and failing probes in order to separate the signal and the non-resonant background. The signal parameters for passing and failing probes are constrained to be equal in the two datasets. In this way the net number

of passing and failing di-muon probes is estimated and the ratio can be computed. In the case of the Z resonance, the signal is modelled as a Breit-Wigner convoluted with a Gaussian, while the background is an exponential. In the case of the J/ ψ resonance, the fit models used are a simple Gaussian for the signal and a 3rd order Chebyshev polynomial for the background. An example of invariant mass distributions for passing and failing probes is shown in Figure 3.22 for both Z and J/ ψ events. The procedure is repeated applying different cuts on event variables in order to compute the efficiency as a function of a specific variable, for example the pseudorapidity of the tag muon.

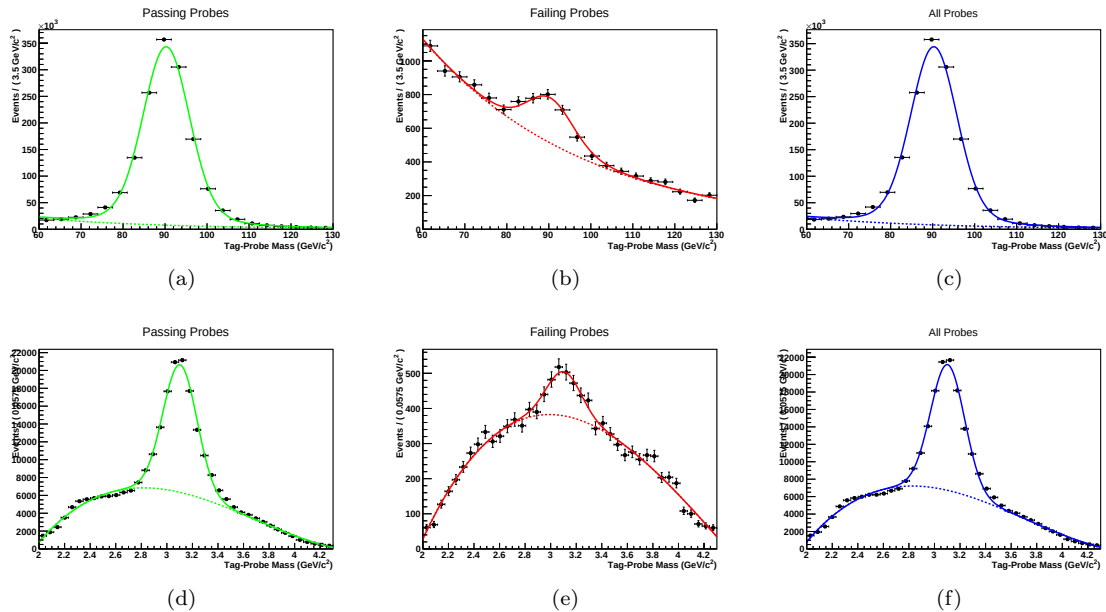


Figure 3.22: The black dots represents the data invariant mass distribution of the Z particle (upper row) and of the J/ ψ particle (lower row) for tag + passing probes (a, d), tag + failing probes (b, e) and tag + all probes (c, f). The fitted curves are also shown for each distribution (colored line). This plot is for $|\eta| < 1.0$.

The method was applied on muons coming from the two resonances produced in the 13 TeV collision data crossing at intervals of 25 ns recorded. The total amount of data corresponding to an integrated luminosity of about 36 fb^{-1} .

In the case of the $Z \rightarrow \mu^+ \mu^-$ decay, events were then selected from the HLT_IsoMu20 trigger path, which contains all the events with at least one isolated muon (tag) with a p_T above 20 GeV that has been reconstructed inside the range of pseudorapidity equal to $|\eta| < 2.4$. The track efficiency obtained from the data is compared for every binned variable to a sample of simulated Z events generated using the Madgraph [90] at LO QCD accuracy. A re-weighting is applied to the MC sample to match the pileup distribution in data.

In the case of the $J/\psi \rightarrow \mu^+ \mu^-$ decay, a specific trigger path (HLT_Mu7p5_L2Mu2_Jpsi) was used in order to select all events that contain one muon fully reconstructed with p_T above 7.5 GeV (the tag) and a second one with p_T above 2 GeV reconstructed without using the inner silicon tracker (the probe). Moreover, this trigger requires that the di-muon invariant mass is around the peak of the J/ ψ . The trigger is prescaled to keep the rate of accepted events low. Results from data are compared to simulated prompt J/ ψ events, generated with Pythia8 [91] with a filter requiring $p_T(J/\psi) > 8 \text{ GeV}$. The trigger requirements used in the data are also applied in the simulation, to avoid biases in the measurement of the efficiency.

With a first part of the 2015 J/ ψ data, different matching criteria have been studied in order to estimate the impact of this choice. The values of the cone thresholds ($\Delta R, \Delta\eta$) were varied in a wide range, and each time the full set of results was recomputed. The muon tracking efficiency and the fake matching rate are shown in Figure 3.23 for five different cones. The fake matching rate is computed taking into account all the events where a passing probe is found even if before

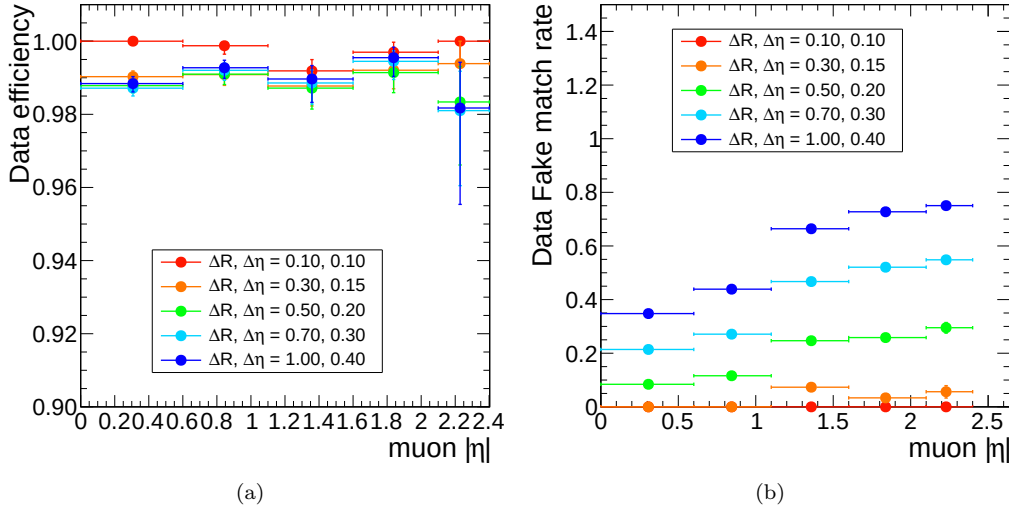


Figure 3.23: Muon tracking efficiency (a) and ratio with simulation (b) for different matching criteria for the J/ψ resonance.

the matching between probes and tracker tracks, all the tracks that combined with the tag give an invariant mass in a specific window around the resonance peak have been removed. Although the matching criterion $\Delta R < 0.1$ and $|\Delta\eta| < 0.1$ seems to have the highest efficiency keeping the lowest fake matching rate estimation, the fit performed on the invariant mass spectra of the failing pairs fails to find a signal. Thus it cannot be compared to the other results. Among the other matching criteria considered there is no big difference between the efficiencies, while the fake matching rate estimation (in both the definitions) is much more dependent on the cone choice. As a consequence, the matching criterion that has been chosen is $\Delta R < 0.3$ and $|\Delta\eta| < 0.15$, which is a good compromise between efficiency and robustness against fake matches.

Figure 3.24 and 3.25 shows the muon tracking efficiency and the corresponding ratios between real and simulated data for 2016 collisions data coming from the Z and the J/ψ resonance, respectively. On the left side, the collection containing all tracks is used, while on the right side only tracks reconstructed using the silicon tracker information are considered. The efficiency is presented as a function of different variables: the absolute pseudorapidity and the azimuthal angle of the probe muon and the number of signal vertices reconstructed in the event.

In Figure 3.24 the measured track efficiency as a function of $|\eta|$ and ϕ is found to be between 99.5% and 100% for the collection including all tracks, while it is lower by a few percentage points in the **Tracker-only seeded** collection. In both collections, the efficiency degrades with increasing number of signal vertices. As opposed to the Z resonance, most of the reconstructed J/ψ have a transverse momentum that is lower than 10 GeV, and in this case the muons are reconstructed mainly in the inner tracking system without making use of the new muon chamber seeded iterations. This result can clearly be seen in Figure 3.26, where the comparison between the simulation and the data muon tracking efficiency as a function of the probe muon p_T for both track collections is plotted for both Z and J/ψ resonance. In the case of the J/ψ particle, the efficiency is recovered for muons with p_T above 10 GeV due to the cuts imposed in the two dedicated muon iterations. For this reason, in the case of the J/ψ particle the results considering the two track collections are very similar in Figure 3.25. Also in the case of muon reconstructed from the J/ψ, the efficiency has a strong dependence on the number of primary vertices.

The agreement between data and simulation reconstruction efficiency is very good. The ratio is between 97% and 100% for both the collections and the data sample considered with the only exception of the high number of primary vertices, where the agreement is between 95% and 97%.

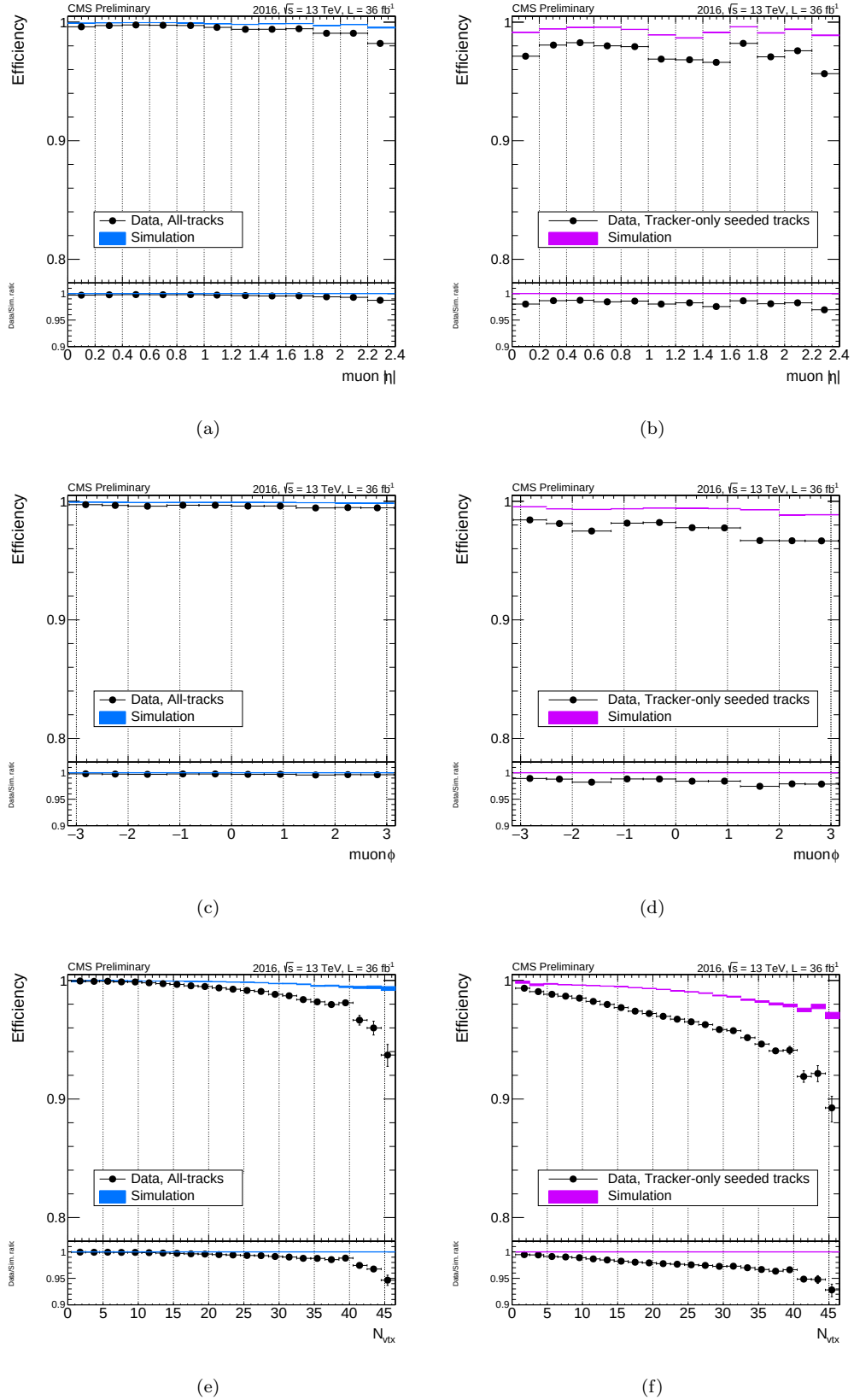


Figure 3.24: Data (black dots) and simulation (rectangles) tracking efficiency and respective ratio for muons coming from the Z decay as a function of the absolute pseudorapidity (a,b), the ϕ of the probe muon (c,d) and the number of primary vertices (e,f). On the left, only tracks in the **Tracker-only seeded** collection are used.

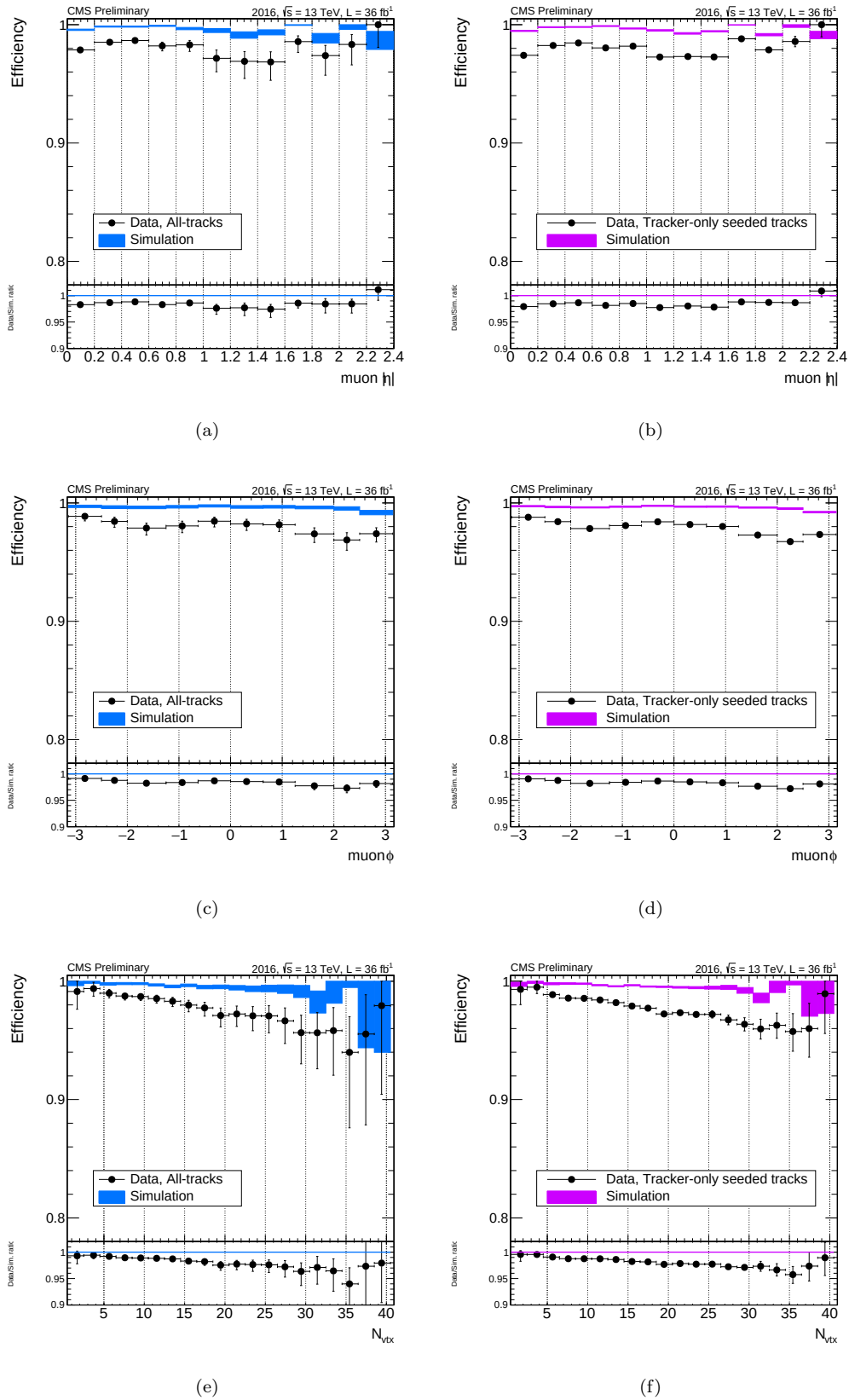


Figure 3.25: Data (black dots) and simulation (rectangles) tracking efficiency and respective ratio for muons coming from the J/ψ decay as a function of the absolute pseudorapidity (a,b), the ϕ of the probe muon (c,d) and the number of primary vertices (e,f). On the left, only tracks in the **Tracker-only seeded** collection are used.

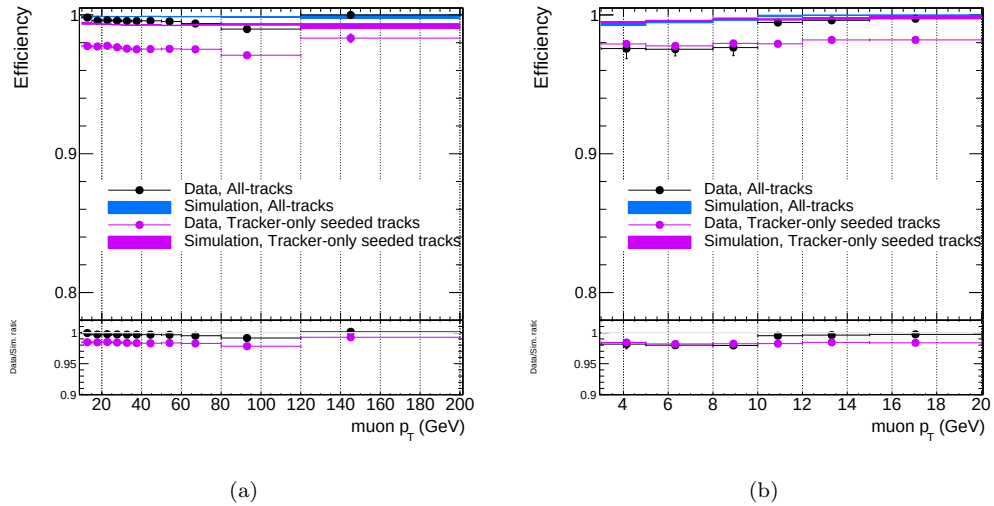


Figure 3.26: Data (dots) and simulation (rectangles) tracking efficiency and their respective ratio for muons coming from the Z (a) and the J/ψ decay (b) as a function of the transverse momentum of the probe muon. The result for the collection containing all the tracks is represented in black and blue, while in purple only tracks in the **Tracker-only seeded** collection are used.

3.4.2 Simulation results for the Pixel Tracker upgrade

A schematic representation of the CMS tracker after the Pixel Tracker upgrade is shown in Figure 3.27 and described in detail in Section 2.2.1. In order to profit from the new Pixel Tracker, the pixel seeding is extended from a triplet configuration to a quadruplet one. Different options are considered [93]:

- The *triplet propagation* seeding is the natural continuation of the pixel triplet seeding approach. In this algorithm, quadruplets are created by propagating triplets to the next layer and finding possible compatible hits. Moreover, the total χ^2 is computed for each quadruplet, and quadruplets with χ^2 above a certain threshold are rejected.
- In the *pixel seed extension* algorithm the idea is propagating the triplet seed at the level of the pattern recognition. In other words, it requires that the first hit after the triplet seed

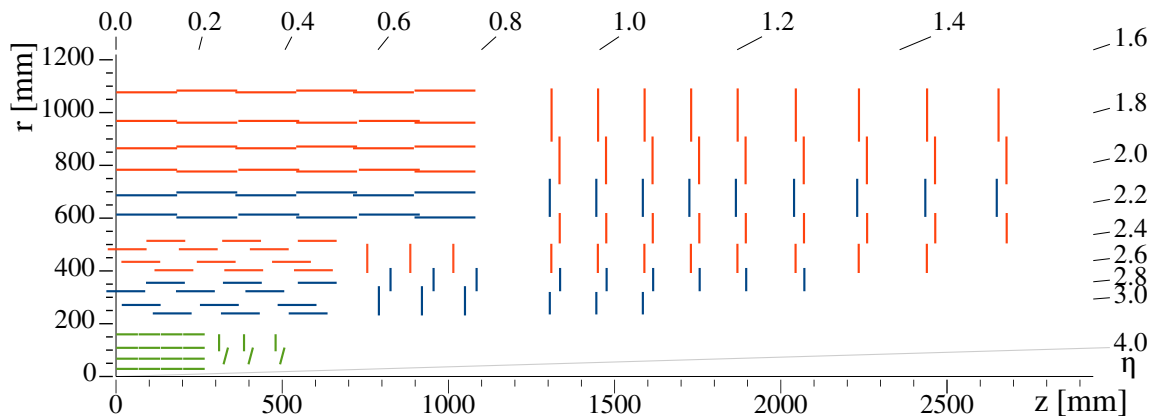


Figure 3.27: Schematic section in $r - z$ view of the CMS tracker after the Pixel Tracker upgrade at the end of 2016. The Pixel Tracker is represented in green, while single-sided and double-sided modules in the silicon Strip Tracker are shown as red and blue segments, respectively [92].

must be a pixel hit.

- A third algorithm is also implemented and evaluated as a possible track seeding algorithm: the **Cellular Automaton-based Hit Chain-Maker (CA)** [94]. In the **CA**, a network of cells is defined and evolves in discrete time steps from an initial state according to predefined rules. In the context of the track seed reconstruction, a *cell* is defined as a segment linking three hits in different layers. A graph including possible connections is created considering *neighborhoods* two cells that have in common a pair of hits and similar η . The *evolution rule* in time of this graph is the following: at each time step a cell can increase its state if on its left it has a neighbor with the same state. At the end of the network evolution, the neighbor fit triples are joined in a longer seed. The **CA** has the advantage that it includes a very fast computation of the compatibility between two connected cells and that it is easy to parallelize, given that the single element state depends only on the values of the cells in the local neighborhood. Thanks to these features, the **CA** can be designed for parallel architectures, such as **Graphics Processing Units (GPU)**. The resulting many-threads-per-event approach scales very well with the pileup by offloading the combinatorics to as many threads as are available on the **GPU**.

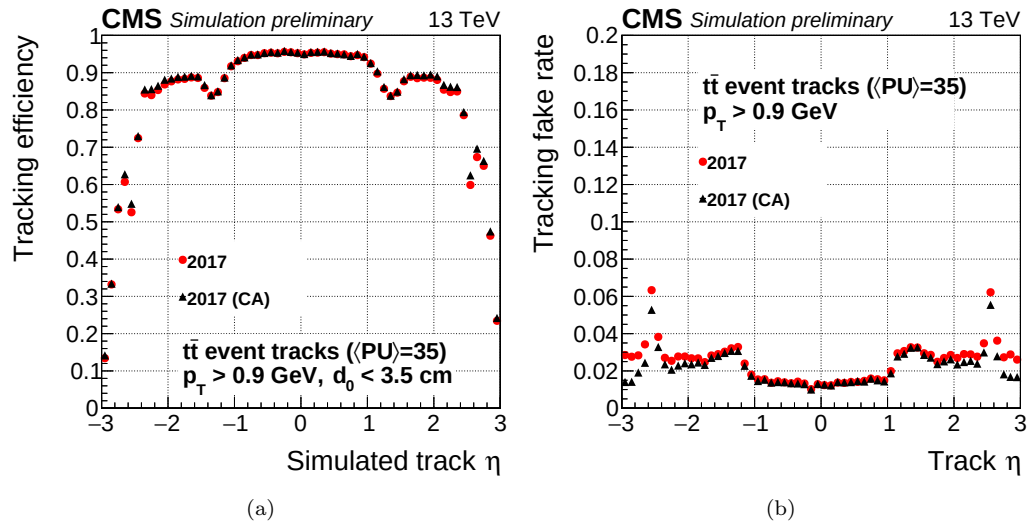
The iterations used for the track reconstruction after the Pixel Tracker upgrade are listed in **Table 3.3**. Each iteration was re-evaluated in the context of the new layout and a full optimization on the selection criteria and on the iterative sequence was also performed. In a first stage, the triplet propagation was introduced in the LowPtQuad and DetachedQuad iterations while the pixel seed extension algorithm was used in the Initial iteration. This configuration is labeled as *conventional quadruplet seeding*. This seeding configuration is evaluated with respect to the **CA** algorithm. The final choice of the seeding algorithm is not only dictated by the performance in terms of efficiency and fake rate, defined in **Section 3.2.3**, but also in terms of time consumption. The efficiency and fake rate for the two approaches is shown in **Figure 3.28** for $t\bar{t}$ events simulated with $\sqrt{s} = 13$ TeV and a superimposed pileup with average 35. The two configurations are compatible with each other. The time spent in the seeding and in the tracking sequence as a function of the average pileup is shown in **Figure 3.29** for 2016 tracking, 2017 tracking with conventional quadruplet seeding, and 2017 tracking with **CA** seeding. The **CA** seeding is shown to be faster than the conventional quadruplet seeding. The gain is visible and remarkable given the increased number of layer combinations involved in the seeding phase with respect to the 2016 case. It was already noted in the developments during the **LS1** that the increase in the combinatoric problem can be partially overcome in the pattern recognition step if the seeds chosen are already high quality. Therefore, a possible explanation is in the intrinsic nature of the **CA** algorithm, which needs fewer and simpler calculations localized in memory to perform as well or even better than the conventional quadruplet seeding approach. The approach chosen as a baseline is therefore the **CA** which guarantees much better seeding time, while preserving if not improving the overall outcome in terms of good tracks reconstructed.

A clear improvement in the efficiency and in the fake rate is shown in **Figure 3.30** for the 2017 tracking compared to 2016 tracking as a function of the pseudorapidity and the p_T of the simulated particle and reconstructed track, respectively [93]. A standard sample of $t\bar{t}$ events simulated with $\sqrt{s} = 13$ TeV and a superimposed pileup with average 35 is used. The contribution of different iterations for 2017 track reconstruction is shown in **Figure 3.31(a)** as a function of p_T : It can be noted how iterations targeting low- p_T tracks are more efficient in the region between 100 and 500 MeV. In **Figure 3.31(b)** the cumulative distribution as a function of the simulated track production vertex radius shows the capability of the tracker to reconstruct highly displaced tracks, such as pions from K_S^0 decay, or particles produced in nuclear interactions and photon conversions. The degradation of the performance with harsher pileup conditions is shown in **Figure 3.32**: The effect is more visible in the fake rate performance, but it remains tolerable.

In conclusion, the iterative algorithm in combination with a cutting-edge technology tracker allows to estimate the track parameters with high precision and to reconstruct tracks in even more challenging pileup conditions.

Table 3.3: List of different tracking iterations used after the Pixel Tracker upgrade with the corresponding seeding configuration used and target tracks.

Iteration	Step Name	Seeding	Target Track
0	Initial	pixel quadruplets	prompt, high p_T
1	LowPtQuad	pixel quadruplets	prompt, low p_T
2	HighPtTriplet	pixel triplets	prompt, high p_T recovery
3	LowPtTriplet	pixel triplets	prompt, low p_T
4	DetachedQuad	pixel quadruplets	from b hadron decay, $r \leq 5$ cm
5	DetachedTriplet	pixel triplets	from b hadron decay, $r \leq 10$ cm
6	MixedTriplet	pixel+strip triplets	displaced, $r \leq 7$ cm
7	PixelLess	inner strip pairs	displaced, $r \leq 25$ cm
8	TobTec	outer strip pairs	displaced, $r \leq 60$ cm
9	JetCore	pixel pairs in jets	high- p_T jets
10	Muon inside-out	muon-tagged tracks	muons
11	Muon outside-in	standalone muons	muons

Figure 3.28: Track reconstruction efficiency (a) and fake rate (b) as a function of simulated track η for standard seeding and CA seeding.

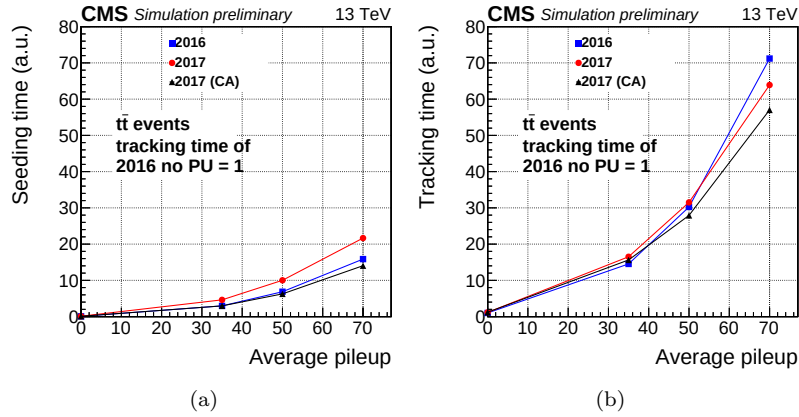


Figure 3.29: Time spent in seeding (a) and in the entire tracking sequence (b) as a function of average pileup for 2016 tracking (blue), 2017 tracking with conventional seeding (red), and 2017 tracking with CA seeding (black). All time measurements are normalized taking the tracking time of 2016 without pileup as a reference.

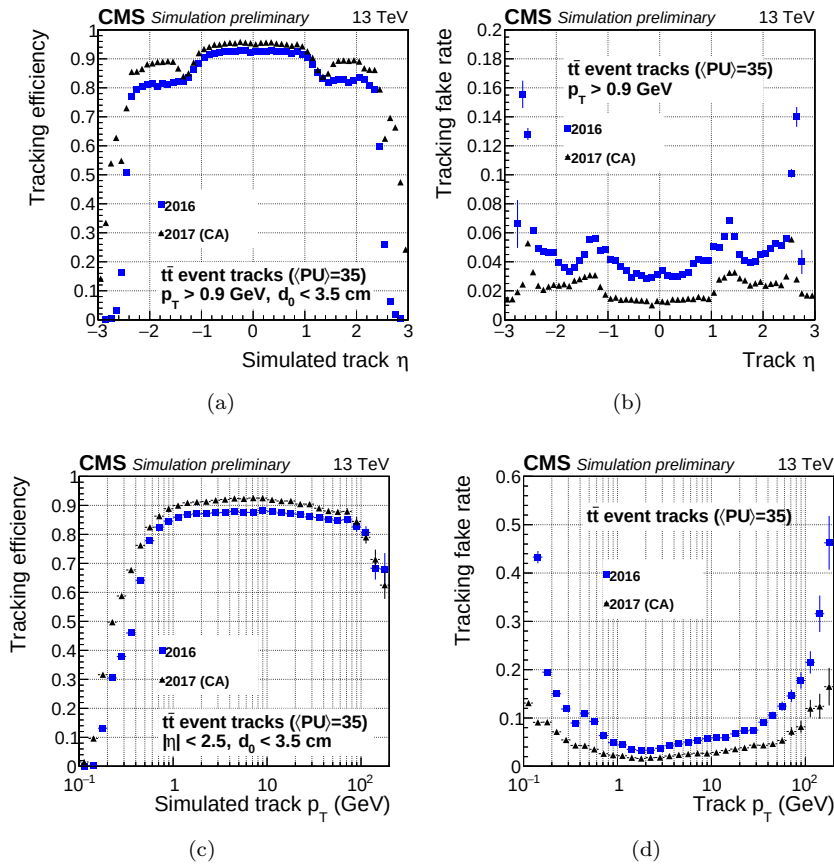


Figure 3.30: Track reconstruction efficiency as a function of simulated track pseudorapidity (a) and of the simulated track p_T (c) for 2016 (blue) and 2017 (black) detectors. Track reconstruction fake rate as a function of reconstructed track pseudorapidity (b) and of reconstructed track p_T (d) for 2016 (blue) and 2017 (black) detectors. The 2017 tracking reconstruction includes the CA seeding.

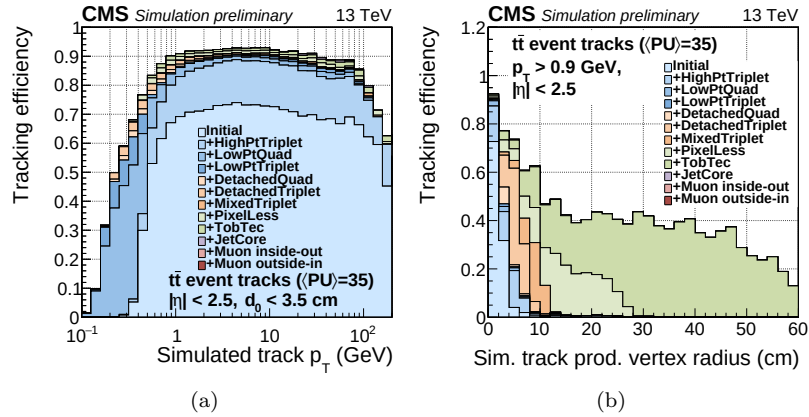


Figure 3.31: Cumulative contributions to the overall tracking performance from the 12 iterations in 2017 track reconstruction. The tracking efficiency for simulated $\bar{t}t$ events is shown as a function of the simulated track p_T (a) and of the simulated track production vertex radius (b).

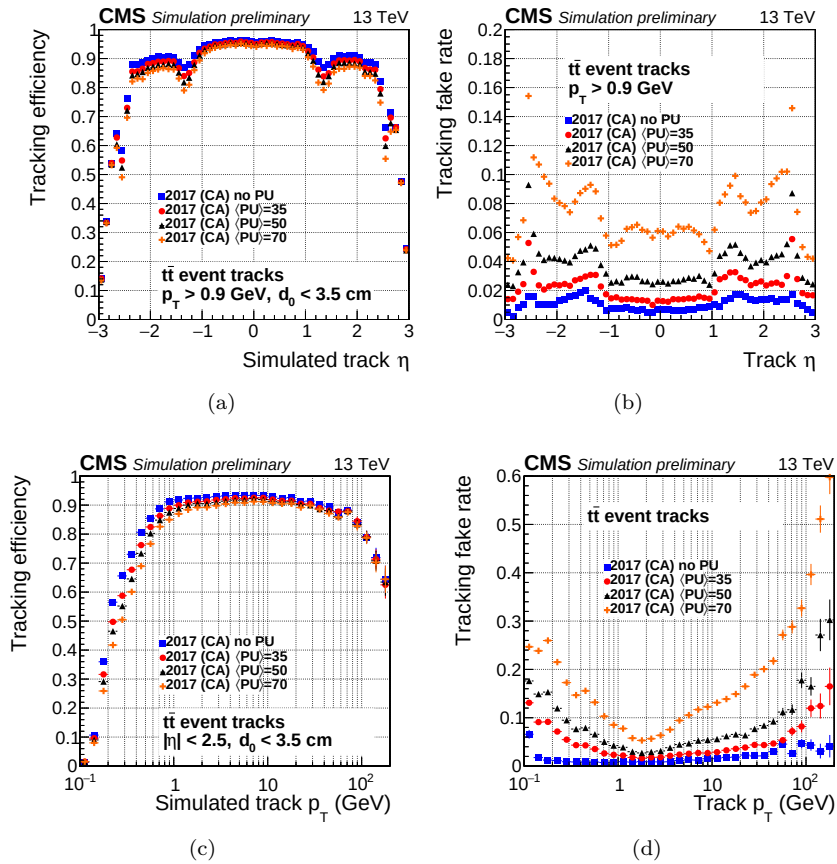


Figure 3.32: Track reconstruction efficiency as a function of simulated track pseudorapidity (a) and of the simulated track p_T (c) for 2017 tracker at different pileup conditions. Track reconstruction fake rate as a function of reconstructed track pseudorapidity (b) and of reconstructed track p_T (d) for 2017 tracker at different pileup conditions. The 2017 tracking reconstruction includes the CA seeding.

CHAPTER 4

Multivariate Analysis of $H \rightarrow \tau^+ \tau^-$

*To Emmy Noether, pioneer of
theoretical physics. (1918)*

In this chapter a study in the context of the $H \rightarrow \tau^+ \tau^-$ signal extraction is presented. The goal of this study is to determine whether a multi-layer perceptron machine learning technique is able to enhance the performance of the signal-versus-background classifier. An introduction of the main analysis approaches in finding a good signal-versus-background classifier is given in [Section 4.1](#). The [Multivariate Analysis \(MVA\)](#) approach implemented using machine learning techniques is described in detail, in particular the use of a neural network technique implemented through the NeuroBayes package. At the end of the section, a [figure of merit](#) is defined in order to compare different [MVAs](#) to each other and to the baseline. In [Section 4.2](#) the result of the NeuroBayes neural network using the $H \rightarrow \tau^+ \tau^-$ simulated dataset with centre-of-mass energy of 13 TeV is presented as well as the comparison with other classification approaches. A description of the simulation samples, the input variables and the optimization process of the NeuroBayes neural network is also given in the same section.

4.1 Signal-versus-background classifiers

In most analyses which aim to find new particles, the signature of the new particle is usually searched for in the invariant mass spectrum. The signature is usually represented by an excess of events over a continuum distribution at a certain invariant mass value. One of the biggest challenges in these analyses is therefore to discriminate the signature of a new particle, also called *signal*, from the continuum distribution, which is referred to as *background*. Three possibilities in order to efficiently separate signal events from background events are: increasing the resolution on the particle energy so that the peak can be well identified over the large background, requesting a signal with features easier to discriminate, and performing an event classification based on some measured event properties which have high separation power.

In the case of the $H \rightarrow \tau^+ \tau^-$ analysis, the first possibility is unfortunately not a feasible option because of the missing energy associated to neutrinos coming from the τ decay. As already explained in [Section 1.2.3](#), the uncertainty on the missing energy is one of the main factors that degrades the invariant mass resolution. This problem can be mitigated by a careful choice of the final state: in the case of this study a $\tau_\mu \tau_h$ pair is chosen. The second possibility listed above to separate signal over background can be implemented in the $H \rightarrow \tau^+ \tau^-$ analysis studying events primarily produced via the [Vector Boson Fusion \(VBF\)](#) process. This process can be easily recognized due to its clear signature characterized by two jets with large separation in pseudorapidity. The last possibility is used in all [Standard Model \(SM\)](#) Higgs analyses: the idea is to find a set of final-state

observables of the event, also called *input variables*, whose distribution is very different for the signal and the background events. These variables are then used to classify events into several categories with different discrimination power. The choice of these variables strongly depends on the processes which are included in the analysis.

Finding a good signal-versus-background classifier that can extract the signal events in the full set of recorded data efficiently is not a simple task, but it is very important because it can enhance the discovery potential of the analysis. Two different approaches to obtain this classifier in the $H \rightarrow \tau^+\tau^-$ analysis are currently under study: a *orthogonal selection cuts approach* also referred to as *cut-based approach*, and a *multivariate approach*. Both approaches are described in this thesis and the results are compared. A particular emphasis is given to the multivariate approach using artificial neural networks.

4.1.1 Cut-based approach

In this approach, the separation of signal events from background events is performed by defining a signal and a background region for every final-state observable. This region is used to discriminate signal-like events, falling in the signal region, from background-like ones. The approach follows an iterative scheme: if the value of the observable for a particular event falls in the signal region, then the event proceeds to the next observable. If an event falls in the background region, it is discarded. This process is repeated for all input variables considered.

In order to improve the quality of the classification, using more information from the input variables can be an option, such as the correlation between different variables, or using non-linear boundaries. Implementing these concepts in the cut-based classification approach is not straightforward, but can be addressed by using a **Multivariate Analysis (MVA)** approach.

4.1.2 Multivariate approach

In order to exploit correlations between final-state observables, an **MVA** technique can be used. With **MVA** techniques the classification task is solved by combining the information extracted from all input variables into one single classification output quantity, which summarizes how likely the event is to be signal or background. In this thesis three different **MVA** approaches are evaluated in order to perform the classification task on the same dataset:

- **Boosted Decision Tree (BDT)**: In a decision tree, the selection process can be described as a sequence of binary decisions (yes/no). This decision splits the multi-dimensional space of the final observables into two subsets and the process is then repeated for each set. The entire decision process can be summarized by a 2-D tree structure [95, 96, 97]. In contrast with the cut-based approach, the **BDT** makes use of the information contained in all events, instead of just a smaller part of it.
- **NeuroBayes neural network**: This **MVA** algorithm is a multi-layer feedforward neural network implemented using the NeuroBayes neural network package [98, 99]. An introduction into artificial neural networks, their features and the software package used is given in the next section.
- **Deep Neural Network (DNN)**: The main idea of deep learning algorithms is to use a neural network with many hidden layers. The concept of hidden layers in the artificial neural network architecture is described in the next section in detail. The idea is to solve complex classification problems exploiting the power of different layers in extracting different low-level features and then combining them in complex high-level features [100, 101].

The first two are interfaced to the multivariate analysis implementation of ROOT called **Toolkit for Multivariate Analysis (TMVA)** [76]. This thesis concentrates on the neural network technique implemented using the NeuroBayes software package, describing in detail its implementation features, the optimization process and the final result. A summary of the results obtained with all three **MVAs** methods and with the cut-based approach is given in **Section 4.2.4**.

Artificial Neural Networks

An **MVA** algorithm is an algorithm that is able to learn from data [100]. As the name suggests, artificial neural networks perform this process using a system vaguely inspired by the human brain: several artificial neurons, which constitute the basic building blocks of the network, are connected to each other and have the possibility to exchange information through links, which evolve in time.

An artificial neuron is an information-processing unit which calculates the output based on the input signals connected to it. It has the following three basic elements [102]:

- A set of *connecting links* or *synapses* which connect either a signal at the input x_j and the neuron k through a synaptic *weight* w_{kj} or different neurons among themselves.
- An *adder* which sums all the input signals weighted by the respective synapses; in most of the neural networks the adder is a *linear combiner*.
- An *activation function* which limits the amplitude of the output of the neuron, typically in the unit interval $[0, 1]$ or $[1, -1]$.

The schematic representation of a generic neuron is given in **Figure 4.1**. Apart from the elements described above, it also contains a *bias*, which can be used to set a specific offset on the adder. The output of a neuron k can be also represented using mathematical terms as follows:

$$y_k = \varphi \left(\sum_{j=1}^m w_{kj} x_j + b_k \right) = \varphi(v) \quad (4.1)$$

where $x_1, \dots, x_m$ are the m input signals; $w_{k1}, \dots, w_{km}$ are the synaptic weights of the neuron k ; b_k is the external bias; $\sum$ is the linear combiner; $\varphi(\cdot)$ is the activation function; and y_k is the output signal. The v parameter includes the external bias and is also called *activation potential*.

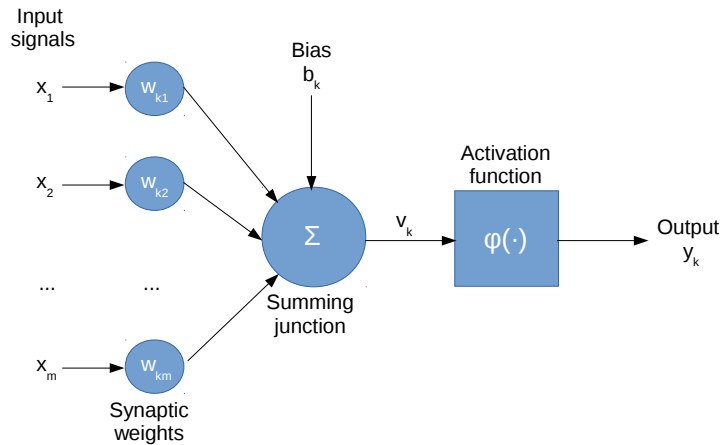


Figure 4.1: Basic model of a neuron in an artificial neural network.

The activation function $\varphi(v)$ plays a significant role in the neural network because it defines the output of a single neuron. Among the different activation functions that can be used, the Sigmoid function, whose graph is represented in **Figure 4.2**, is the most common. It is a strictly increasing function where the output can take values between the lower and upper bound of 0 and 1. An example of this kind of activation function is the logistic function:

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (4.2)$$

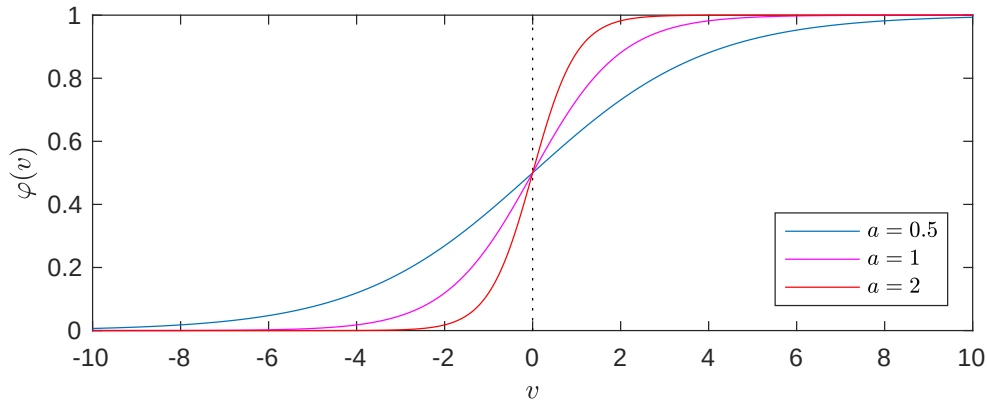


Figure 4.2: Sigmoid function as activation function $\varphi(v)$.

where a is the slope parameter of the Sigmoid function. This function with $a = 1$ is the activation function used in this thesis, if not specified otherwise.

The way how the the neurons are connected in a neural network is referred to as *network architecture*. Neural architectures can be divided into three fundamental classes:

- **Single-layer feedforward networks:** A single-layer feedforward network is the simplest layered network¹. It is composed of an *input layer* containing the input nodes that projects onto an *output layer* of neurons. This network is strictly feedforward because the output layer is only connected with the input layer. In a feedforward network, no loop connections and no connections in the same layer are allowed.
- **Multi-layer feedforward networks:** This is a second class of feedforward networks and it differs from the single-layer network because of the presence of one or more *hidden layers*. The hidden layers are composed of *hidden neurons* and, thanks to this feature, this class of networks has the power to solve more complex problems than the single-layer feedforward networks. In the example in Figure 4.3, the structure of a possible multi-layer neural network with one hidden layer is shown: the input layer is only connected to the hidden layer and the latter is only connected to the output layer. All the connections are described by weights. The set of output signals in the output layer constitutes the final response of the network. The network is called *fully connected* if all the neurons in each layer are connected to every node in the adjacent layers, otherwise the network is called *partially connected*.
- **Recurrent networks:** In these networks at least one feedback loop is present. An example is a network with just one layer where all the output neurons feed all the neurons in the input layer. Hidden layers can also be included in this class of networks. Their internal structure allows to exhibit dynamic temporal behavior.

Training and testing

After the network architecture has been defined, the *training* process starts. Machine learning algorithms can be divided in two categories depending on the training process: in the *supervised learning* networks, each single data point in the input dataset is associated with a label or a target, while in the *unsupervised learning* algorithms data points do not have any associated label. All the cases examined in this chapter belong to the first category.

In supervised neural networks that perform classification tasks, the main goal of the training process is that the neural network learns the properties of the input samples and how to distinguish samples with different labels. The training process consists of iteratively adjusting the weights of the network in order to obtain the output value as close as possible to the desired value [99]. This

¹A layered network is a network where all neurons are organized in layers.

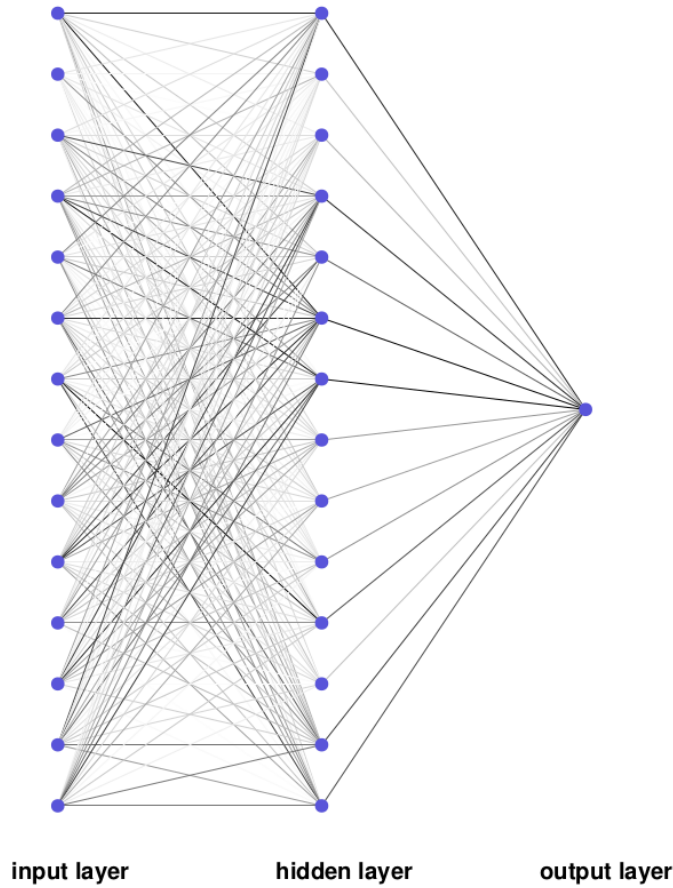


Figure 4.3: Example of fully connected multi-layer feedforward network [99].

task is performed by minimizing a *cost function* $C(\theta)$. The cost function is defined through a *loss function* L , that compares the output after a forward propagation with the target value providing information on how well these two values agree with each other. The cost function can be written as follows:

$$C(\theta) = \frac{1}{n} \sum_{j=1}^n L(\hat{y}_{\theta}^{(j)}, y^{(j)}) \quad (4.3)$$

where $\theta = w_1, b_1, \dots, w_n, b_n$ represents the weights and the biases of the network; the values $\hat{y}_{\theta}^{(j)}$ is computed performing a feedforward propagation from the input vector $\vec{x}^{(j)}$ using the configuration θ ; and the value of $y^{(j)}$ represents the target value of $\vec{x}^{(j)}$.

If $C(\theta)$ is differentiable, then the optimum of the neural network configuration θ is obtained minimizing the cost function with respect to the parameter θ :

$$\frac{\partial C(\theta)}{\partial \theta} = 0. \quad (4.4)$$

Training of the neural network by finding the minimum of the cost function is equivalent to finding the best set of weights and biases for the network. Possible choices for the loss functional are the quadratic deviation or a measure of the entropy. The value of the cost function after the training process is called *training error*. In order to improve the speed of learning, a *momentum* term can be added to the gradient of the cost function [103]. An example of a momentum is a fraction of the previous weight update to the current one. The momentum can be either chosen constant or as increasing over time.

The challenge in machine learning is to perform well not only on the training dataset, but especially on new, previously unseen, inputs. This feature is also called *generalization* [100]. The network is thus tested on an independent second sample at the end of the training, the so-called *test sample*. This step is performed under the assumption that the training and the test samples are generated independently from each other, but that their input variables are drawn from the same probability density function. Ideally, the neural network shows similar performance for the training dataset and the test set, with small deviations due to possible statistical fluctuations. If this is not the case, the reason can be found in the *underfitting* or the *overfitting* phenomenon described below.

Two main factors influence the machine learning performance: its ability to minimize the training errors as well as the difference between training and testing error. Underfitting happens when the model is not capable of achieving a low error value on the training dataset. This phenomenon is usually correlated with a low capacity² neural network, which may struggle to fit the training sample. Overfitting, also called *overtraining*, occurs when the gap between the training error and test error is too large and is usually related with high capacity models. In this case the neural network memorizes features of the training sample which are not useful on the test dataset, such as the inherent statistical fluctuations. The biggest challenge is therefore to find a good compromise in the size of the network capacity which must be appropriate for the complexity of the task and the size of the input dataset.

A possibility in order to avoid overfitting is to use additional techniques, such as the *regularization* method. In regularization theory [102] the main idea is to stabilize the solution introducing a non-negative function that includes prior information about the solution itself. In the context of machine learning, the regularization process includes any modifications of the learning algorithm which reduce the difference between training and testing error but keep the training error invariant. Practically, this can be done by introducing a regularization term in the neural network loss function, λL_C , where λ is a positive real number called regularization parameter and L_C depends on the distributions of the input signal.

The NeuroBayes package

The NeuroBayes neural network package [99] is a sophisticated tool developed for high energy physics to perform **MVA**. The package was developed by the University of Karlsruhe and is available under the NeuroBayes® from Phi-T GmbH [104]. It implements a feedforward neural network with one hidden layer and includes automated preprocessing of the input variables. The preprocessing step transforms and de-correlates the input variables before they are fed into the network. In the NeuroBayes package, this is done in two steps: a global preprocessing and a single-variable preprocessing. The first determines whether to perform the de-correlation and renormalization of the input variables. All possible global preprocessing options are listed in Table 4.1. In the single-variable preprocessing, the original input variable can be either used in the original distribution form or transformed to a flat distribution by a nonlinear transformation or into a Gaussian distribution with mean zero and $\sigma = 1$. The user can choose among a wide range of options to perform the preprocessing stage of the neural network.

The user also has the possibility to set some of the parameters of the NeuroBayes network to improve the analysis, such as the learning speed, the momentum or the regularization function. All the global neural network parameters are listed in Table 4.1, including their default values. Most of them are either self-explanatory or they have already been described in this chapter. Among the different techniques for the regularisation process, the following options are considered: the "REG" option uses the Bayesian regularisation procedure to apply different regularisation constants depending on the position of the weights inside the network; in the "ARD" option and in the "REG" option all the input nodes and the output nodes get individual regularisation constants, respectively; with the "ALL" option, the "ARD" and the "REG" options are combined to achieve the maximal amount of regularisation [99].

An implementation of the NeuroBayes package in the **TMVA** tool of the ROOT analysis framework is used in this thesis [76].

²The capacity of a network is defined as the complex measurement of the maximum amount of data that may be transferred between network locations over a link or network path.

Table 4.1: NeuroBayes neural network parameters that can be modified by the user.
 * default value

Parameter	Description and possible options
Regularization	Definition of the regularization scheme used in the training "OFF", "REG"*, "ARD", "ASR", "ALL"
LossFunc	Type of loss-function to be minimized "ENTROPY"*, "QUADRATIC"
ShapeTreat	Existence of direct connections between input/output layer "INCL", "OFF"*
TrainingMethod	Optional use of the BFGS algorithm [105] "","", "BFGS"
NTrainingIter	Number of complete training iterations any value, 100*
LearningSpeed	Multiplicative factor of the network learning speed any value, 1.0*
LimitLearningSpeed	Upper limit on the LearningSpeed parameter any value, 1.0*
Momentum	Define the training momentum term any value $\in [0.0, 1.0]$, 0.0*
WeightUpdate	Number of events processed before the weights are updated any value, 200*
NodesNumb	Number of nodes in hidden layer $NVAR + 1$ *
Global Preprocessing	0 = do not perform de-correlation 1 = de-correlate input variables and normalize 2 = de-correlate input variables

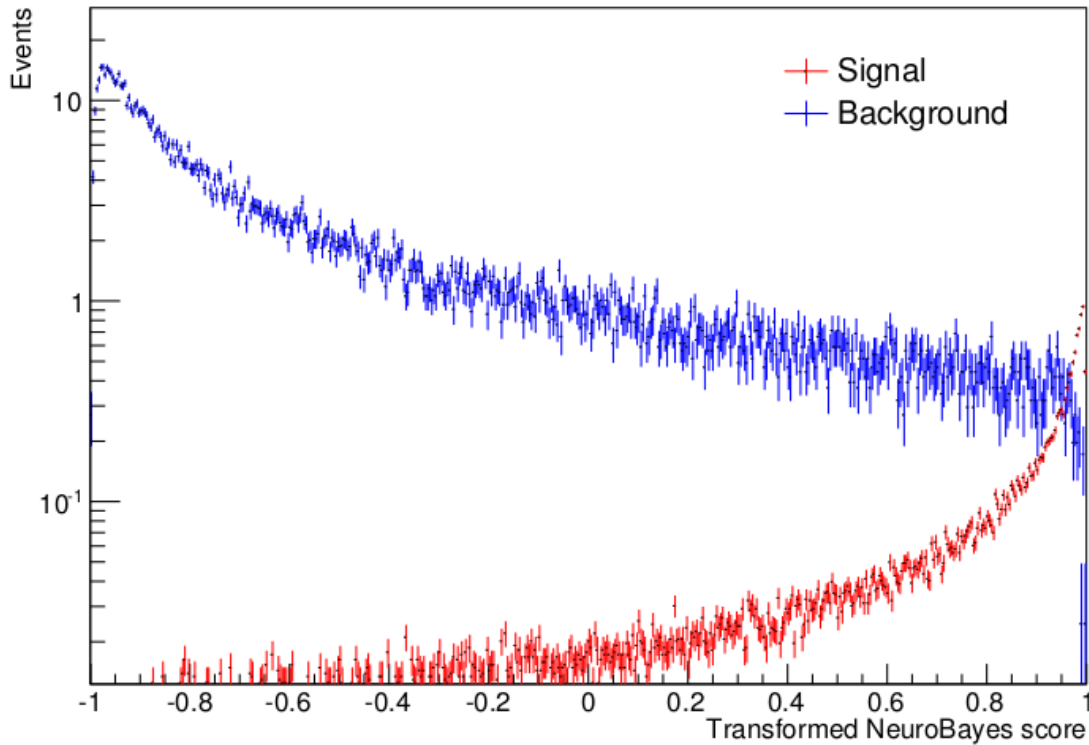


Figure 4.4: Example of score distribution of a **MVA** based on a NeuroBayes neural network. The signal-like events (red) peak at 1, while the background-like ones (blue) peak at -1.

4.1.3 Quantification of performance

In order to choose the classification method with the highest discrimination power, a unique **figure of merit** (**FoM**) must be defined. The first step in order to calculate the **FoM** is to train the neural network with the training dataset. At the end of the training process a so-called score is assigned to each single event. In the case of this thesis, the score corresponds to the output value for each event. The **FoM** of each **MVA** approach analyzed is computed taking into account the values in the score distribution.

The output distribution in the case of the NeuroBayes neural network is a real number between -1 and 1 . The events with an output value close to -1 are more likely to be background events, while the events with a score around 1 are signal-like events. In **Figure 4.4**, an example of the score distribution for background-like events and signal-like events in the case of the NeuroBayes neural network **MVA** is shown renormalized with respect to the number of expected events for a specific luminosity condition.

The two histograms with signal and background score are filled with the same binning. For each bin i of this histogram the Approximate Median Significance (AMS) is computed [106]. Without taking into account uncertainties, the AMS simplifies to

$$AMS_i(s_i, b_i) = \sqrt{2 \left[(s_i + b_i + b_{reg,i}) \ln \left(1 + \frac{s_i}{b_i + b_{reg,i}} \right) - s_i \right]} \quad (4.5)$$

where s_i and b_i are the number of signal-like events and background-like events, respectively.

The **FoM** is evaluated as the square root of the quadratic sum of the AMS_i for each bin:

$$AMS = \sqrt{\sum_{i=bin} AMS_i(s_i, b_i)} \quad (4.6)$$

In this formula, statistical and systematic uncertainties are not explicitly included for simplicity,

but they are taken into account in the actual calculation of the **FoM**. Some properties of this **FoM** are:

- It includes a parameter b_{reg} which is a regulation term to prevent the overestimation of AMS_i when the amount of background-like events is very low. During the optimization process the b_{reg} value is chosen to be 1, while once the optimization conditions are selected, this factor is set to almost 0.
- It converges to $s/\sqrt{b}$ for $s \ll b$.
- The binning of the final score distribution must be chosen carefully: if the size of the single bin is too large, part of the information contained in the distribution is not preserved, on the other hand, the result obtained with a too narrow binning can be less relevant due to statistical fluctuations.

In order to overcome this last issue, the **FoM** is evaluated with a narrow binning in a first step, and in a second step a rebinning is performed with respect to the statistical uncertainties and the amount of background events in the bins.

4.2 Performance of NeuroBayes neural network $H \rightarrow \tau^+\tau^-$

4.2.1 13 TeV simulation samples

In a first step, events are generated using the mathematical formalism and models describing the **SM** with different event generators: Powheg [107] and Pythia6 [108] for signal events and Madgraph [90] and Pythia8 [91] for background events. The signal events originate from a Higgs boson produced through **VBF** and decaying into $\mu\tau_h$ according to the requirements of the previous section. Drell-Yan processes are used to simulate background events. In this specific case only the $Z/\gamma^* \rightarrow \tau\tau$ process is included. In a second step, the **CMS** detector is simulated including its response to the passage of particles. In **CMSSW** this step is done using the Geant4 toolkit [109]. Finally, the reconstruction of each event is performed using the detector response information.

Data are simulated with a centre-of-mass energy of 13 TeV and normalized to total integrated luminosity corresponding to 10 fb^{-1} . On both the signal and the background dataset, a preselection is applied in order to select only potentially interesting events:

- Only events with a final state containing an isolated μ and one hadronically decaying τ lepton, τ_h , are selected.
- Only events originating from a Higgs boson produced through the **VBF** mode are selected. This is done requiring at least two jets with a certain threshold in p_T and a certain difference in pseudorapidity.
- Several kinematic selections are required, for example on the p_T of the isolated μ and τ_h which must be above 20 GeV.

All requirements are listed in **Table 4.2**. A total of 78404 events and 89764 events are selected after preselection for background and signal, respectively. Both datasets are split in half for the training and the testing procedure.

4.2.2 Input variables

In **Table 4.3** the input variables used in the **MVAs** are listed. It is important to choose suitable variables with high discrimination power between signal and background events. The discrimination power of each single variable can be seen in Figures 4.5 and 4.6, where all the 12 input variables are plotted for both signal and background. The distributions are normalized to 1 in order to keep a fair comparison independent of the number of expected events for each process.

Table 4.2: Summary of preselection requirements.

Variable	Definition	Requirement
$p_T(\tau_h)$	Transverse momentum of τ_h	> 20 GeV
$p_T(\mu)$	Transverse momentum of μ	> 20 GeV
$p_T(j_{1,2})$	At least 2 jets with transverse momentum	$> 45/30$ GeV
$\Delta\eta(j_1, j_2)$	Pseudorapidity gap between leading jets	> 2.0
m_{vis}	Invariant mass of the μ - τ_h -system	> 30 GeV
m_T	Invariant transverse mass of μ - E_T^{miss} -system	< 80 GeV
CJV	Central jet veto	✓
$q(\mu) \cdot q(\tau_h)$	Product of the charges of μ and τ_h	< 0
b veto	Veto on b-tagged jet in the event	✓
lepton veto	Veto on additional lepton in the event	✓

Table 4.3: Summary and description of the input variables used.

Symbol	Definition
m_{SV}	SVFit mass estimation [26]
$\Delta\eta(j_1, j_2)$	η gap between the leading jets
$p_{T,\text{sum}}$	Scalar p_T sum of τ_h , μ and the two jets
$p_{T,\text{tot}}$	Magnitude of the p_T vector sum of τ_h , μ , the two jets and E_T^{miss}
E_T^{miss} centrality	Relative angular E_T^{miss} position with respect to μ and τ_h [24]
μ centrality	Lepton pseudorapidity with respect to the two jets [24]
m_{vis}	Invariant mass of the μ - τ_h -system
$\Delta R(\mu, \tau_h)$	ΔR between μ and τ_h
m_{j_1, j_2}	Invariant mass of the two leading jets
m_T	Invariant transverse mass of μ - E_T^{miss} -system
$\eta_{j_1} \cdot \eta_{j_2}$	Product of the leading jet pseudorapidities
S	Sphericity - 3-D isotropy of the energy flow [24]

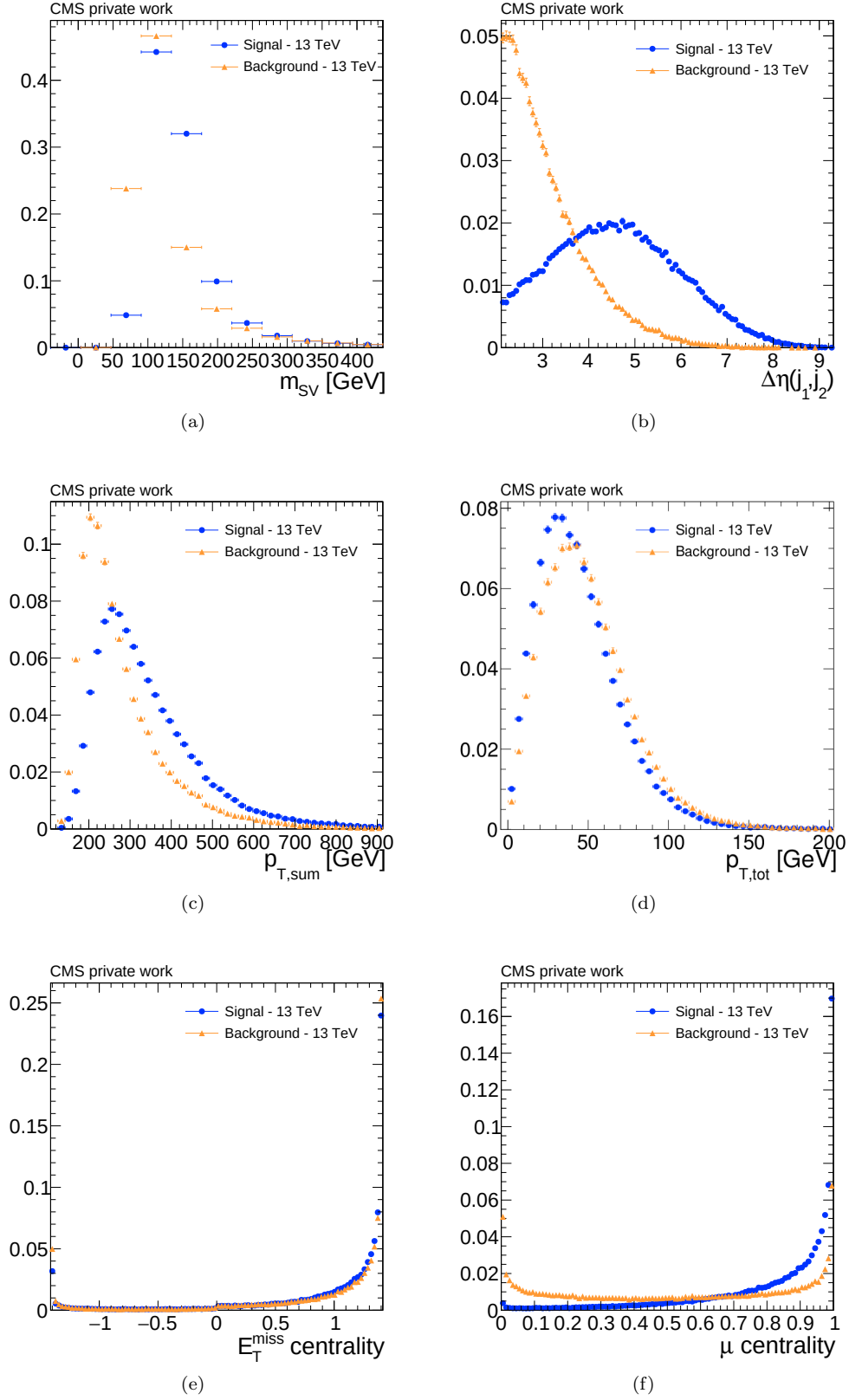


Figure 4.5: Distributions of the variables m_{SV} (a), $\Delta\eta(j_1, j_2)$ (b), $p_{T,sum}$ (c), $p_{T,tot}$ (d), E_T^{miss} centrality (e), and μ centrality (f) for signal (blue) and background (orange) events.

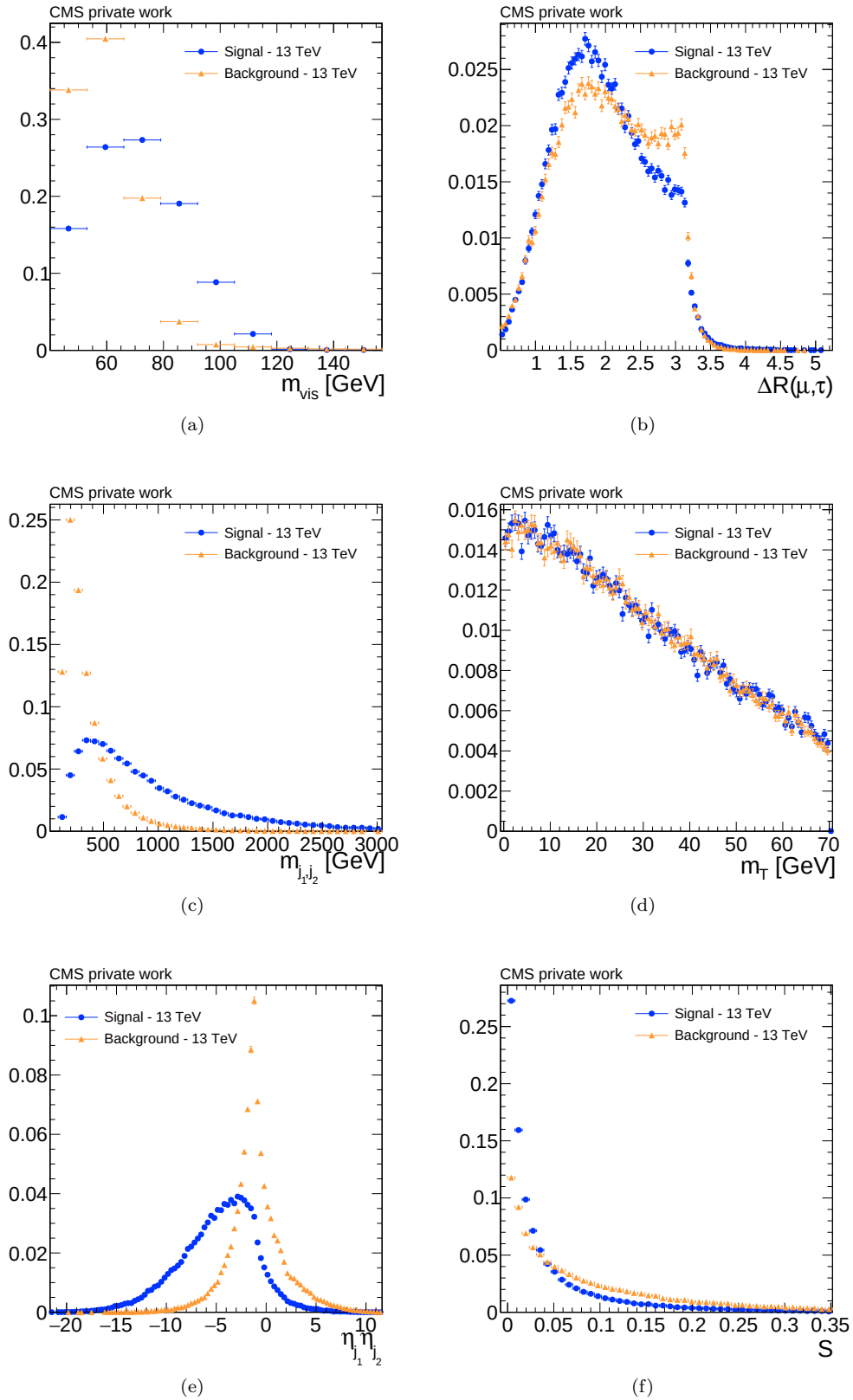


Figure 4.6: Distributions of the variables m_{vis} (a), $\Delta R(\mu, \tau_h)$ (b), m_{j_1, j_2} (c), m_T (d), $\eta_{j_1} \cdot \eta_{j_2}$ (e), and S (f) for signal (blue) and background (orange) events.

Table 4.4: Optimized NeuroBayes neural network parameters in the $H \rightarrow \tau^+\tau^-$ analysis for 13 TeV.

Parameter	Description and possible options
Regularization	"OFF"
LossFunc	"ENTROPY"
ShapeTreat	"OFF"
TrainingMethod	"BFGS"
NTrainingIter	220
LearningSpeed	1.0
LimitLearningSpeed	2.0
Momentum	0.0
WeightUpdate	200
Global Preprocessing	De-correlate input variables

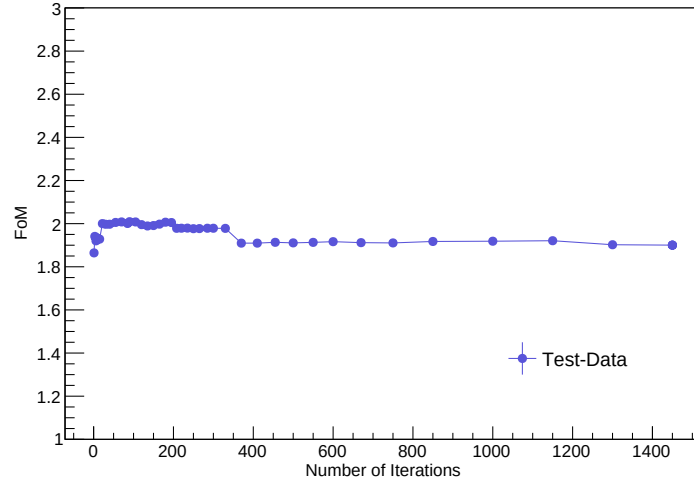
4.2.3 Optimization of the network

The NeuroBayes neural network and its parameters are extensively described in [Section 4.1.2](#). The optimization of this **MVA** is performed on three different aspects of the network:

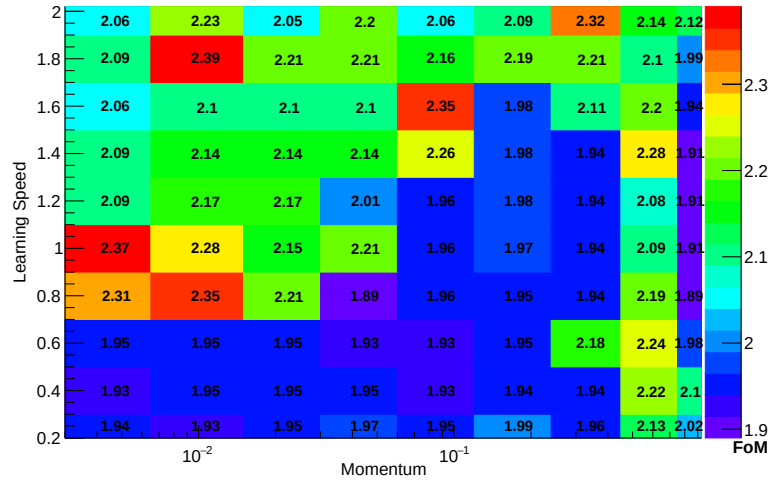
- The first step is performing an optimization of the NeuroBayes neural network global parameters. They are listed in [Table 4.1](#) and they can be divided into continuous and discrete parameters. The continuous parameters are the number of iterations, the learning speed and the momentum. They are optimized using both 1-D and 2-D scans as shown in [Figure 4.7](#) as example. The discrete parameters of the network are the regularization, the loss function, the shape treatment, the training method and the global preprocessing. Each combination of all possible values taken by the discrete parameters is evaluated in terms of the **FoM** and the combination with the highest **FoM** is chosen. In [Table 4.4](#) the final choice of all the global parameters is listed. In some cases the global maximum is not chosen as working point, in particular if it is located in an unstable region with a high variation in the **FoM** values in the neighbourhood.
- In a second step, the preprocessing procedure is optimized for each input variable. Each individual variable is optimized keeping fixed all the others. The optimization of the preprocessing also is done to optimize the **FoM** and does not include a-priori considerations of the physical properties of the variables. In [Table 4.5](#) the final choice of the preprocessing parameters is listed for each observable.
- The last step is to choose the optimal variable subset that leads to the highest performance of the signal extraction. In order to reduce the number of input variables, testing every single possibility is not feasible for such a high-dimensional problem. Therefore two iterative methods are used: the *up method* and the *down method*. The up method starts evaluating the **FoM** for each input variable individually. In this way the variable with the highest discrimination power is selected. The method proceeds testing all possible combinations of two variable including the one already selected and keeping the combination with highest **FoM**. The procedure is repeated for every input variable and the **FoM** of the best combinations is plotted in [Figure 4.8\(a\)](#). In the down method, the selection of the subset is performed removing in each iteration the variables one by one and choosing the combination with highest **FoM** as new starting point. The result for the down algorithm is shown in [Figure 4.8\(b\)](#). Both evaluations show that the highest performance is obtained when all input variables are used. It is interesting to note that the first variable in the up method, m_{j_1, j_2} , does not correspond to the last variable of the down method, m_{vis} . This fact underlines the importance of the variable correlations in the **MVA**.

Table 4.5: Optimized single preprocessing for each input variable.

Variable	Single preprocessing
μ centrality $\Delta R(\mu, \tau_h)$ m_{j_1, j_2} m_T $\eta_{j_1} \cdot \eta_{j_2}$ S	} Flatten the distribution and transform to Gaussian.
$\Delta\eta(j_1, j_2)$ m_{vis}	} Flatten the distribution and use result of regularised fit to mean values of the target.
$p_{T, \text{sum}}$ E_T^{miss} centrality	} Use original distribution and transform to flat distribution.
m_{sv} $p_{T, \text{tot}}$	} Use original distribution + use result of regularised monotonous fit to mean values of the target.

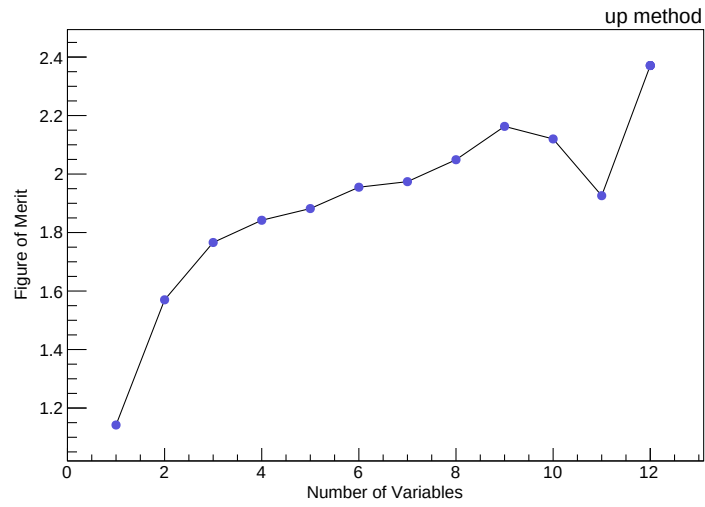


(a)

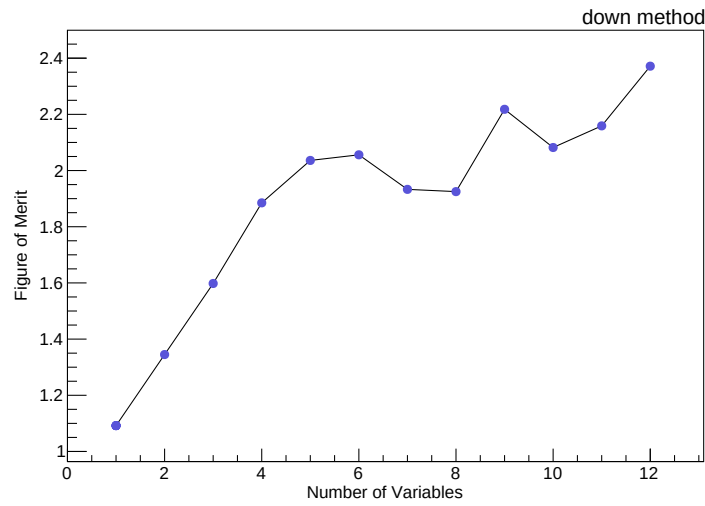


(b)

Figure 4.7: Examples of the optimization of the NeuroBayes neural network parameters during the testing procedure: 1-D scan of the FoM as a function of the number of iterations (a) and 2-D scan of the FoM in the learning speed vs momentum plane (b).



(a)



(b)

Figure 4.8: Best FoM for each combination of input variable as a function of the number of variable in the combination itself. Both methods are shown, the up method (a) and the down method (b).

Table 4.6: Highest FoM for different MVAs used in the $H \rightarrow \tau^+\tau^-$ analysis. The result found for the NeuroBayes technique using the 13 TeV dataset with integrated luminosity of $L = 10\text{fb}^{-1}$ is compared not only to other MVAs but also to the result for the 8 TeV dataset assuming an integrated luminosity of $L = 19.7\text{fb}^{-1}$. The b_{reg} term in Equation 4.6 is set close to 0 for this calculation.

Approach	Cut-based	NeuroBayes	DNN	BDT
8 TeV	2.1 [97]	2.9 [110]	2.1 [101]	3.5 [97]
13 TeV	1.9 [97]	3.1	2.1 [101]	3.7 [97]

4.2.4 Result and comparison with other classification approaches

The final distribution for background events and signal plus background events using the NeuroBayes neural network MVA is shown in Figure 4.9. The total figure of merit is found to be 3.1. As mentioned already in Section 4.1.3, in the calculation of the final FoM the b_{reg} parameter is set to a value close to 0, precisely 10^{-6} . In the distribution the concentration of signal-like events in the region around 1 is clearly shown, demonstrating the high discrimination power of this MVA.

In Table 4.6 the NeuroBayes MVA classification approach is compared to three other approaches: the cut-based approach, the BDT MVA approach and the deep neural network MVA approach. All the MVA approaches produce better performance than the cut-based approach. Among the different MVAs analyzed, the BDT performs best, while the DNN only performs slightly better than the cut-based. The NeuroBayes analysis has some potential but does not perform as well as the BDT. Probably both the NeuroBayes and the DNN require a larger dataset in order to get a significant improvement in performance with respect to the BDT. Moreover, in the table also the results for the 8 TeV dataset are reported, and the same conclusions are reached. The optimization and the results of these approaches are described in detail in Refs. [110, 97, 101].

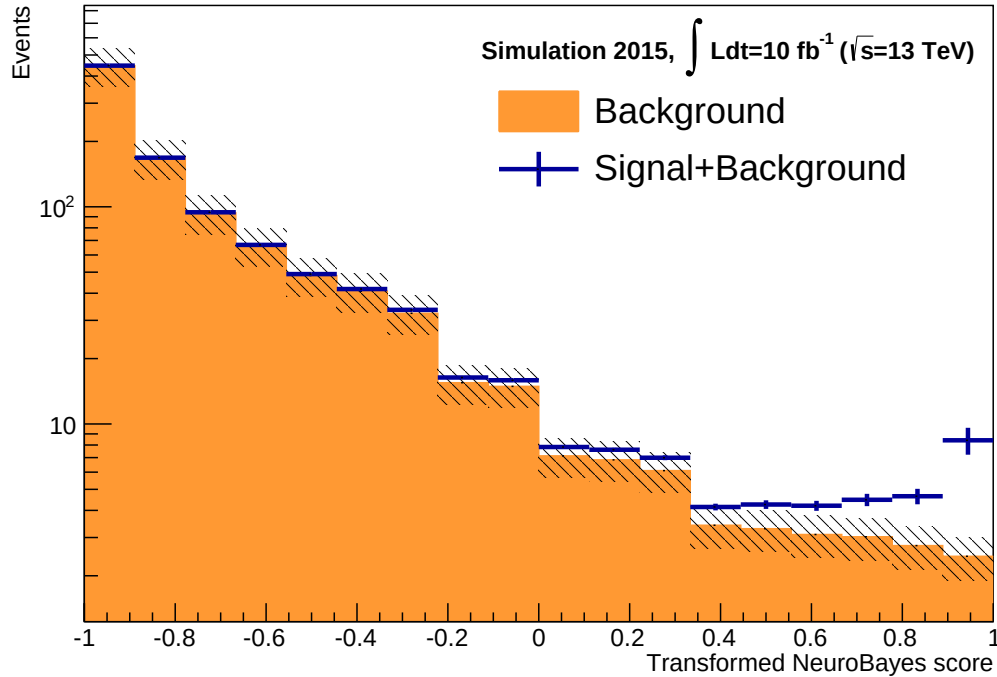


Figure 4.9: Final NeuroBayes MVA distribution for background (orange) and signal plus background (blue) events. The FoM computed for this final distribution is found to be $\simeq 3.1$, imposing $b_{reg} = 10^{-6}$.

CHAPTER 5

The CMS detector at the High Luminosity LHC

*To Grace Hopper, first woman
using “debugging”. (1947)*

During the **Long Shutdown 3**, scheduled from 2024 to mid 2026, **CERN** is planning an upgrade program in preparation of the **High Luminosity LHC (HL-LHC)**, which will bring the luminosity up to $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, or even $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ in the ultimate performance scenario. In the latter scenario, up to 200 inelastic collisions on average will be superimposed on the event of interest. An overview of this upgrade program including the main theoretical motivations is given in **Section 5.1**. In order to face the challenging conditions expected during the entire **LHC** operation period, two upgrade phases are foreseen by the **CMS** experiments and they are described in **Section 5.2**. In particular, **CMS** will build a completely new silicon tracking detector to reconstruct charged particle momenta with high precision in this high-occupancy environment. In **Section 5.3**, a detailed description of the new **CMS** tracking detector is given.

5.1 The High Luminosity LHC

5.1.1 Introduction

As already mentioned in **Section 2.1.2**, during the **Long Shutdown 3 (LS3)** the **LHC** will be upgraded to the **High Luminosity LHC (HL-LHC)** [34]. An overview of the current schedule from 2015 up to 2035 is shown in **Figure 5.1**. For both **LHC** and injectors, all the different phases are reported: running periods for data taking, long shutdowns, beam commissioning periods and technical stops.

The **HL-LHC** program aims to achieve an instantaneous luminosity of $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, or even $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ in the ultimate performance scenario, at a rate of 40 MHz. This is five to seven times beyond the nominal instantaneous luminosity of the **LHC**. About 140 pileup events on average per bunch crossing are expected, increasing up to 200 pileup events in the ultimate luminosity scenario. The total integrated luminosity collected by the **ATLAS** and **CMS** experiments for an estimated projected lifetime of the **HL-LHC** of about ten years is around 3000 fb^{-1} , and up to 4000 fb^{-1} in the ultimate scenario. The **HL-LHC** runs with the same centre-of-mass energy of 14 TeV and bunch spacing conditions of 25 ns as the **LHC**. The project was approved by CERN Council in June 2016 and is considered a top priority effort by the European Strategy for Particle Physics [111] and the US Particle Physics Project Prioritization Panel [112].

5.1.2 Theoretical motivations for the HL-LHC

Plenty of highly relevant physics results have been already achieved by the experiments **ATLAS**, **ALICE**, **CMS**, and **LHCb** using the amount of data provided by **LHC** up to now. Two highlights of these results are: the **SM** Higgs boson discovery by the **ATLAS** and **CMS** experiment in 2012 [18, 19], and the first measurement of the branching ratios of the rare decays of the neutral B_s^0 and B^0 mesons to two muons by **CMS**, **LHCb** [113] and **ATLAS** [114]. Moreover, a large variety of new particle physics models have been already tested experimentally, and both **ATLAS** and **CMS** have been able to place stringent limits on many of them. Despite this success, there are still a lot of open questions in particle physics. In this context, the amount of data collected by the **LHC**'s experiments is not sufficient to answer them.

The **HL-LHC** upgrade program focuses on precision measurements of the **Standard Model** (SM) Higgs boson, on vector-boson scattering, and on searches for signatures of physics **Beyond the Standard Model** (BSM) [115, 116]. It expands greatly the physics potential of the **LHC** and allows the experiments to prove rare and statistically limited **SM** and **BSM** processes. For example, given that all the properties of the Higgs boson measured up to now are compatible with those expected from the **SM**, one of the main goals is using the 3000 fb^{-1} provided by the **HL-LHC** to gain a deep understanding of the potential responsible for the realization of the Brout-Englert-Higgs mechanism. In Figure 5.2 the couplings to fermions and bosons for the **SM** Higgs boson as a function of the particle mass are shown using **CMS** Run 1 data, along with the projection for the integrated luminosity of 3000 fb^{-1} . A substantial improvement can be seen in particular for the Higgs coupling to muons, which is a particularly challenging measurement given the branching fraction of only about 10^{-4} . A second crucial aspect is the Higgs boson self-coupling, whose cross section is about a factor 1000 smaller than the one for single Higgs boson production. A first measurement of the trilinear Higgs coupling is possible with the **HL-LHC** data collected by **ATLAS** and **CMS**.

The processes of weak vector boson scattering are particularly important in the **HL-LHC** analysis program because they are tightly linked to electroweak symmetry breaking and the **SM** Higgs boson. Up to now first measurements have been provided but they are experimentally very challenging, due to the large and irreducible backgrounds and the small cross sections of this process. New possibilities in this sector are opened by the 3000 fb^{-1} provided by the **HL-LHC**.

In the context of new physics searches, the sensitivity for searches for **BSM** particles is highly correlated with the luminosity, and therefore the discovery potential of these processes at the **HL-LHC** increases, such as searches for Supersymmetry, Dark Matter, extra dimensions, and extra gauge bosons. With a ten-fold increase in luminosity, the discovery potential is extended to higher masses for many **BSM** scenarios. For example, in Figure 5.3 the 95% **CL** exclusion limits and 5σ discovery contours for 300 fb^{-1} and 3000 fb^{-1} luminosity is shown in a simplified squark-gluino model with massless neutralino as well as in $(\tilde{t}, \tilde{\chi}_1^0)$ mass plane with the assumption of the $\tilde{t} \rightarrow t + \tilde{\chi}_1^0$

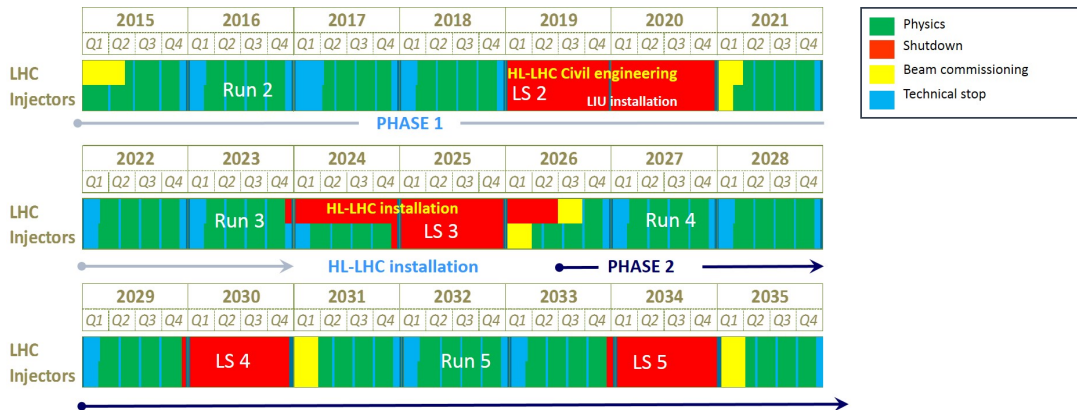


Figure 5.1: Overview of the **LHC**'s planned schedule from 2015 to 2035 [92]. The running periods are indicated in green and long shutdowns in red. Phases of beam commissioning and technical stops are also indicated in yellow and blue, respectively.

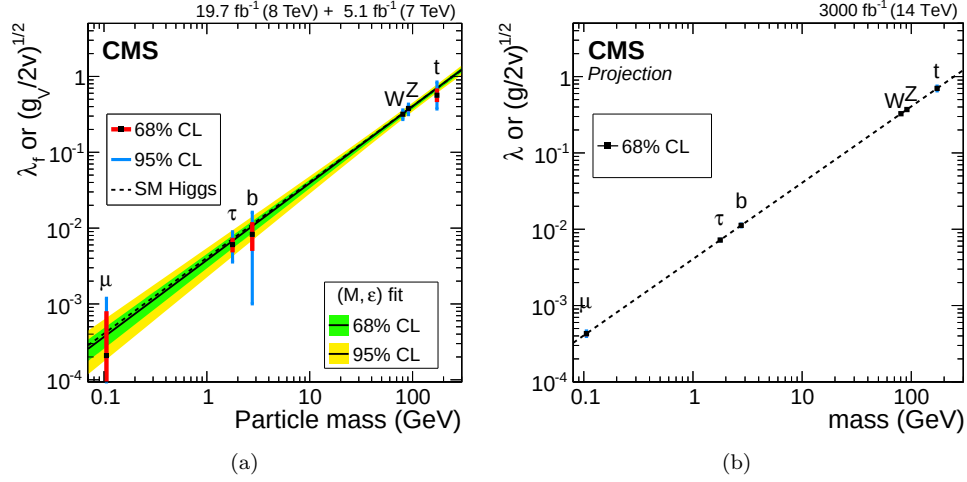


Figure 5.2: Higgs couplings to fermions (Yukawa coupling, λ_f) and bosons (parametrized as $(g_V/2v)^{1/2}$, where v is the SM Higgs boson vacuum expectation value) as a function of the particle mass [116]. In (a) Run 1 data are used, while in (b) the projection for the uncertainties assuming standard model couplings is plotted for an integrated luminosity of 3000 fb⁻¹.

or the $\tilde{t} \rightarrow b + \tilde{\chi}_1^\pm$, $\tilde{\chi}_1^\pm \rightarrow W^\pm + \tilde{\chi}_1^0$ decay modes. Moreover, new channels with low production cross sections or small coupling strengths open up, giving the possibility to search indirectly for BSM physics in rare decays.

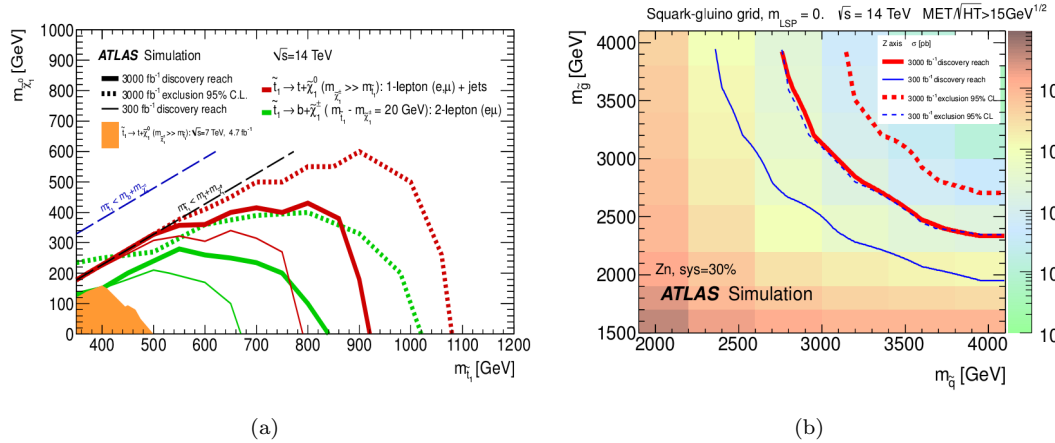


Figure 5.3: 95% CL exclusion limits and 5σ discovery contours for 300 fb⁻¹ and 3000 fb⁻¹ luminosity in a simplified squarkgluino model with massless neutralino (a) and in $\tilde{t}, \tilde{\chi}_1^0$ mass plane with the assumption of the $\tilde{t} \rightarrow t + \tilde{\chi}_1^0$ or the $\tilde{t} \rightarrow b + \tilde{\chi}_1^\pm$, $\tilde{\chi}_1^\pm \rightarrow W^\pm + \tilde{\chi}_1^0$ decay modes (b).

5.1.3 The machine

The main goal of the HL-LHC design study reported in Ref. [34] is to find a set of hardware configurations and beam parameters that allow to:

- Achieve a peak luminosity of $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ with luminosity levelling (described below);
- Allow the ATLAS and CMS experiments to collect an integrated luminosity of about 300 fb⁻¹ per year, with a total integrated luminosity of about 3000 fb⁻¹ in ten years after the upgrade installation;

Table 5.1: HL-LHC beam parameter with respect to the LHC ones.

Parameter	Nominal LHC	Nominal HL-LHC
Beam energy in collision [TeV]	7	7
N	1.5×10^{11}	2.2×10^{11}
n_b	2808	2748
θ_c [μrad]	285	590
Beam separation [σ]	9.4	12.5
Minimum β^* [m]	0.55	0.15
ϵ_n [μm]	3.75	2.50
r.m.s. energy spread	1.13×10^{-4}	1.13×10^{-4}
r.m.s. bunch length	7.55×10^{-2}	7.55×10^{-2}
R without crab cavities	0.84	0.30
R with crab cavities	0.98	0.83
Average number of events/bunch crossing	27	135
Peak line density of pile-up event [event/mm]	0.21	1.25
Luminosity/IP [$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$]	1	5

- Operate the accelerator with high efficiency until at least the year 2035.

The *luminosity levelling* consists in operating at constant luminosity below the maximum achievable value during the physics collisions period, as shown in Figure 5.4. It is used to limit the amount of pileup events which otherwise is not sustainable by the interaction region magnets and the detectors.

The studies for such a highly complex machine have already begun in 2010, and it is estimated that the necessary developments require about ten years of prototyping, testing and realization. A strong R&D effort is put into a number of key innovative technologies which can allow the HL-LHC machine to achieve the goal. In Table 5.1, the beam parameters of the HL-LHC including levelling are listed in comparison with the nominal design of the LHC. Some of these parameters are described in the instantaneous luminosity formula in Equation 2.1.

The increase of the instantaneous luminosity is achieved using novel cutting-edge technology. A first step consists of replacing the inner triplet quadrupole magnets in the insertion regions with new, advanced and more radiation tolerant ones based on Nb₃Sn superconducting technology. This intervention is necessary due to the amount of radiation dose received after about 300 fb^{-1} , which is close to the maximum dose that the inner triplet quadrupoles and their corrector magnets can withstand. A total amount of 24 such Nb₃Sn quadrupoles is needed for the HL-LHC project: They feature a higher magnetic field strength and a larger magnet aperture, which significantly lowers the β^* parameter in the collision region, and therefore increases the instantaneous luminosity. The significant decrease of β^* corresponds to a larger crossing angle θ_c and a smaller geometrical reduction factor R . A second step is thus needed in order to compensate for this reduction. A possible solution is to increase the crossing angle of the proton beams using so-called *crab cavities*, which

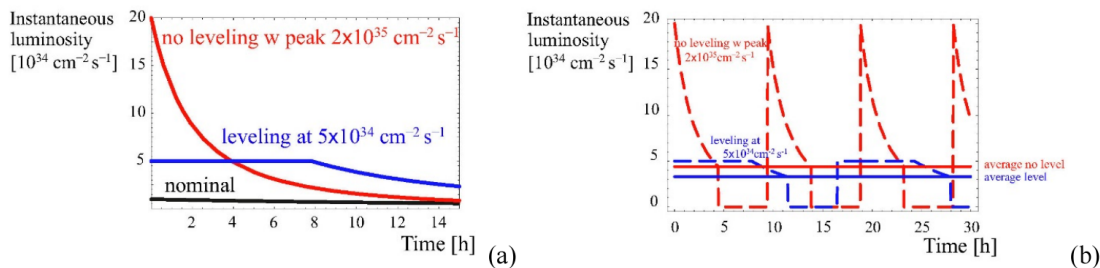


Figure 5.4: (a) Luminosity profile for a single long fill: starting at nominal peak luminosity (black line), with upgrade no levelling (red line), with levelling (blue line) [34].
(b) Luminosity profile with optimized run time, without and with levelling (blue and red dashed lines), and average luminosity in both cases (solid lines) [34].

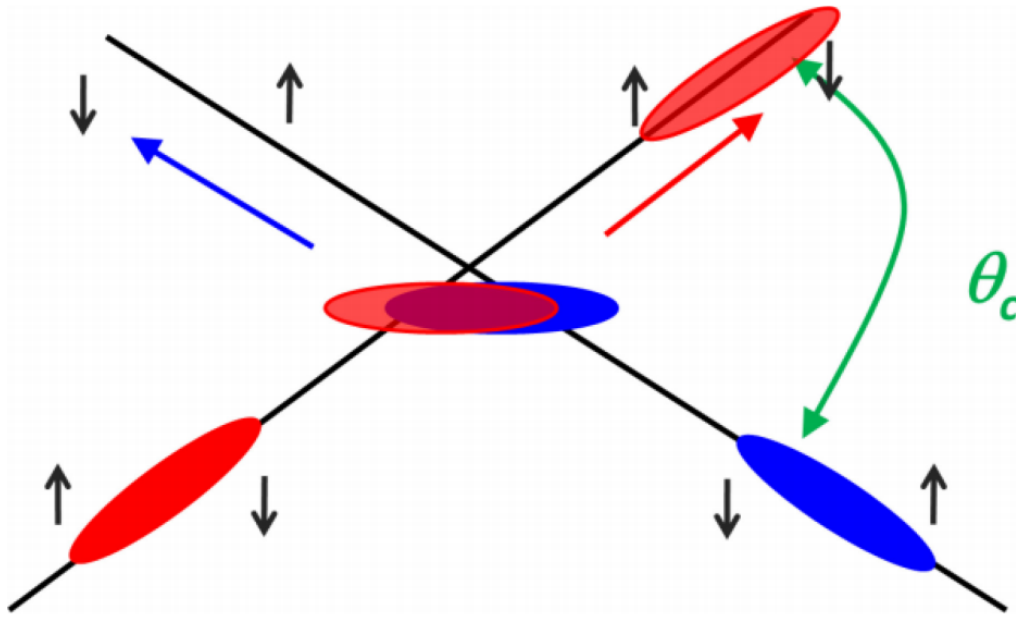


Figure 5.5: Effect of the crab cavity on the beam [34]. Small arrows indicate the torque on the bunch generated by the transverse radiofrequency field.

will be installed in the interaction regions. Special superconducting radiofrequency crab cavities have been developed and tested in the last few years with the idea of rotating both beams longitudinally by $\theta_c/2$, so that they effectively collide head on. A simplified representation is illustrated in Figure 5.5. They can also be used in order to provide a dynamic β^* variation, and therefore allow further optimization of the luminosity region during collisions. Further technological HL-LHC improvements are: a new technology for beam collimation and long high-power superconducting links with zero energy dissipation [34].

The LHC will also have a high-luminosity heavy-ion operation which will start already in Run 3 given that it does not require either the upgraded inner triplet quadrupoles or the crab cavities. The ultimate goal of this second program is to accumulate 10 nb^{-1} of Pb-Pb luminosity during the LHC operations after Run 2. Nucleus-nucleus collisions will thus be provided to ALICE, CMS and ATLAS, and proton-nucleus collisions also to LHCb.

If, as already mentioned, the performance of the HL-LHC will demonstrate that it can go beyond the design levelled luminosity, the HL-LHC program is also considering an ultimate performance scenario with a levelling luminosity of $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. All systems and detectors are therefore already designed with some safety margin to allow the machine to go further depending on the behavior of the machine itself and of the upgraded detectors in the first running year at the HL-LHC. This scenario is represented in Figure 5.6 and will allow the ATLAS and CMS experiments to collect almost 4000 fb^{-1} by 2037.

5.1.4 Experiments at the HL-LHC

In Table 5.2 the target instantaneous luminosities with levelling for proton-proton operation for the HL-LHC with respect to the LHC one is shown for all four main experiments. For ALICE and LHCb the peak levelled luminosity achieved will be lower than the one expected for ATLAS and CMS: ALICE is expected to collect integrated luminosities of 100 pb^{-1} of proton-proton data per year, while LHCb will collect from 5 fb^{-1} to 10 fb^{-1} per year. A short description of the upgrade plan before and during the HL-LHC era is summarized here for ALICE [117], LHCb [118] and ATLAS [119, 115]. The CMS upgrade program is described in more detail in Section 5.2. The current operating detectors at the LHC are already described in Section 2.1.4.

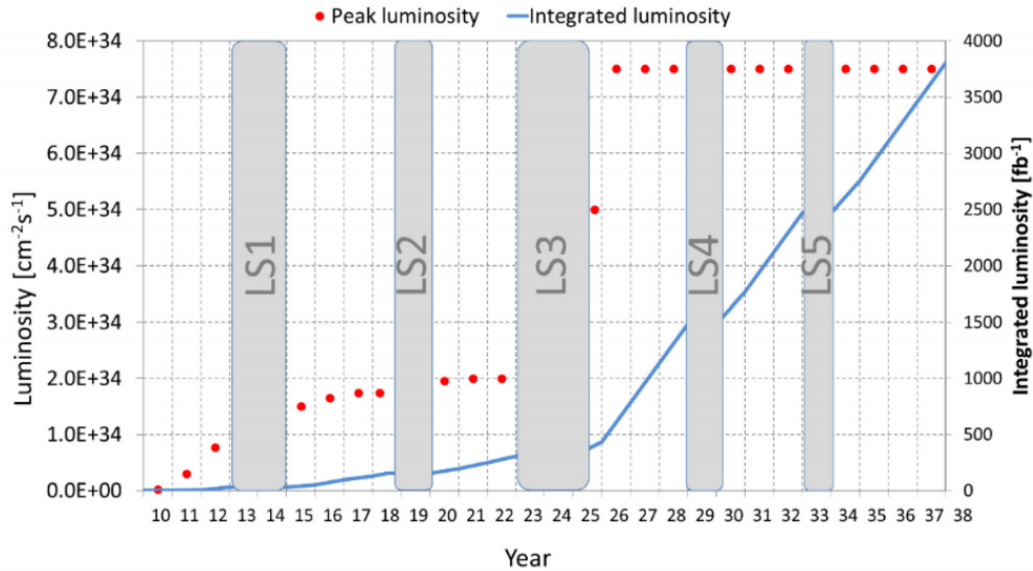


Figure 5.6: Forecast for peak luminosity (red dots) and integrated luminosity (blue line) in the era, in the case of ultimate **HL-LHC** parameters [34].

Table 5.2: Target luminosities for proton-proton operation for the **HL-LHC** for each experiment with respect to the **LHC** ones.

Experiment	Peak levelled luminosity [$\text{cm}^{-2} \text{s}^{-1}$]	
	at the LHC	at the HL-LHC
ALICE	1×10^{31}	1×10^{31}
ATLAS	2×10^{34}	5×10^{34}
CMS	2×10^{34}	5×10^{34}
LHCb	4×10^{32}	2×10^{33}

Upgrade plan of ALICE

During the **LS2**, **ALICE** will undergo an upgrade program to prepare for the high-luminosity heavy-ion operation and to fully exploit the **LHC** Runs 3 and 4. The main goals of this program are to improve the tracking precision with a new silicon tracker system, to replace completely the **Time-Projection Chamber (TPC)** wire chambers with **Gas Electron Multiplier (GEM)** readout, and to install a read-out system which will accept Pb-Pb interactions at a rate of up to 50 kHz. The upgrade of the **TPC** defines the installation strategy and schedule. The idea is to record an integrated luminosity of 10 nb^{-1} after **LS2** in minimum-bias trigger mode.

Upgrade plan of LHCb

The current **LHCb** detector will continue to operate until the end of Run 2 in 2018. Afterwards the **LHCb** experiment will be upgraded, referred to as the Phase-1 upgrade, to allow the experiment to function efficiently at the instantaneous luminosity of $2 \times 10^{33} \text{ cm}^{-2} \text{s}^{-1}$. Most of the sub-detectors will be replaced as well as the entire **Front-End (FE)** electronics and the data-acquisition system. At the end of Run 4, **LHCb** will have presumably collected around 50 fb^{-1} , and many of the components will not be able to sustain any more radiation. During **LS3**, the detector will be completely re-designed with the idea of operating at an instantaneous luminosity of 2×10^{34} , ten times that of the Phase-1 upgrade detector. This second upgrade of the **LHCb** experiment, the so-called Phase-2 upgrade, aims to fully exploit the flavour-physics opportunities of the **HL-LHC** project. A strong R&D effort has already started, addressing possible challenges and solutions. In particular, the following aspects are under study: the introduction of a high granularity tungsten sampling electromagnetic calorimeter, the enlargement of the tracking acceptance to reconstruct

soft tracks, and downstream fast-timing capabilities.

Upgrade plan of ATLAS

ATLAS foresees two upgrade phases to adapt to the challenges of the **LHC**. During **LS2** a first upgrade, the so-called Phase-1 upgrade, will allow the experiment to withstand the expected radiation dose until the end of Run 3, while retaining the same physics performance. In this upgrade, the granularity of the calorimeters involved in the **Level-1 (L1)** trigger will be increased, and a new muon trigger and tracking detectors (new small wheel) in the forward region will be introduced into the muon spectrometer. Moreover, a dedicated hardware trigger system will be developed using the **FastTracKer (FTK)** system. The execution time for global tracking in this system averages only 25 μs .

After **LS3** a second upgrade is foreseen, the Phase-2 upgrade, to support the full **HL-LHC** physics program at peak instantaneous luminosities of $7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and for a total integrated luminosity of $3000 - 4000 \text{ fb}^{-1}$. The high luminosity conditions present challenges to the operation and performance of the entire detector as well as the trigger. A complete replacement of the current Inner Tracker system is foreseen due to the accumulated radiation damage. The new system will be a new full silicon tracker composed of silicon pixels and silicon strip detectors. The silicon technology is chosen to maintain tracking performance in the high occupancy environment and to cope with the increase of the radiation fluence. Moreover, an upgrade of the forward calorimeter system and muon spectrometer is also foreseen. Furthermore, the entire trigger system will be re-designed: the current **L1** trigger will move to a new stage, called Level-0 trigger, using the information from the calorimeters and the muon detectors, while the **L1** trigger will gain new functionality and will receive information also from the tracking systems. The Level-0 trigger rate will accept a rate of at least 500 kHz with a latency of 6 μs , while the **L1** system reduces the rate to 200 kHz with an additional latency of 14 μs . Finally, the computing and software will be upgraded to meet the challenges of the increased luminosity and changes in computer architectures.

5.2 The upgrade of the CMS experiment

5.2.1 Overview of the upgrade plan

The **CMS** experiment foresees two different upgrade phases to face the **LHC** challenges: a Phase-1 upgrade phase consisting of the period from **LS1** to **LS2**, and a Phase-2 upgrade during the **LS3** with the main goal of getting the detector ready for operations at the **HL-LHC**.

The work foreseen during the Phase-1 upgrade is very broad. It involves muon detectors, hadron calorimeters, the pixel detector, the trigger and data acquisition, and the beam radiation monitoring and luminosity measurement system [120]. As already mentioned in **Chapter 2**, during the extended year end technical stop between 2016 and 2017, part of the Phase-1 upgrade was already carried out: the new pixel detector was installed, the readout of the **HF** was upgraded and the new **L1** trigger was included. Some of the upgrades are still to be performed: changing the photo detectors in the **HB** and **HE** of the hadron calorimeter, exchange of scintillating tiles in **HE** if needed, and including new **FE** electronics for the **CSC** endcap muon chambers.

The two biggest challenges dictating the need of a second upgrade during **LS3** are the unprecedented level of radiation and the very extreme pileup conditions at the **HL-LHC** [116]. For the design integrated luminosity of 3000 fb^{-1} a 1 MeV neutron equivalent fluence of $2.3 \times 10^{16} n_{eq}/\text{cm}^2$ and a total ionizing dose of 12 MGy are expected at the centre of **CMS**. The amount of absorbed dose in different regions of the **CMS** detector is shown in **Figure 5.7**. This requires improved radiation hardness, especially for the sub-detectors close to the interaction point. Concerning the pileup challenge, the type of pileup and the possible problems arising in the event reconstruction are described in **Section 2.1.3**. An example of a high pileup event close to the interaction region is shown in **Figure 5.8**, where 102 vertices are reconstructed. On top of the effect on track reconstruction, described in **Chapter 3**, harsh pileup conditions can also add extra energy to the calorimeter measurements, confuse the trigger and the offline event reconstruction, and increase the amount of data that has to be read out in each bunch crossing. Some possible solutions to counteract the larger amount of pileup and the associated increase in the particle density are either increasing the

granularity of the sub-detectors or correcting the energy deposition using timing or pulse shape information provided by the readout electronics. Higher data rates will require an increase of the bandwidth of the new CMS detector and improved trigger capability to keep the trigger rate at an acceptable level while not compromising physics potential.

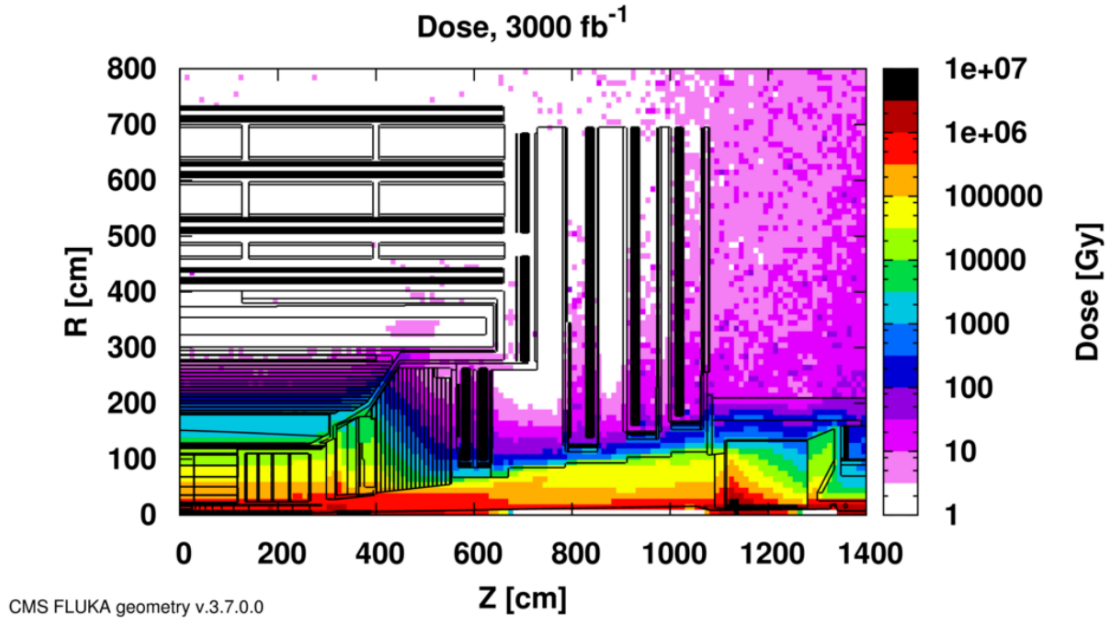


Figure 5.7: Absorbed dose in the CMS cavern after an integrated luminosity of 3000 fb^{-1} in the $r - z$ view [116].

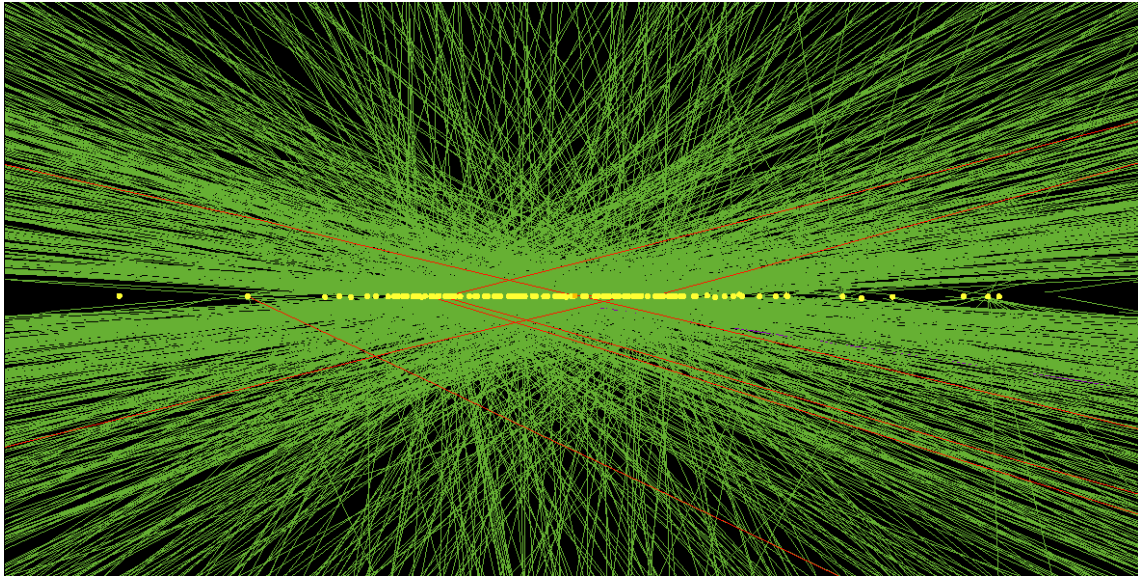


Figure 5.8: High pileup event with 102 reconstructed vertices.

5.2.2 The CMS detector at the HL-LHC

The primary goal of the Phase-2 upgrade program is to keep the excellent performance of the Phase-1 detector under the challenging **HL-LHC** conditions. In Ref. [116] a detailed description of the entire Phase-2 upgrade of the **CMS** experiment is presented. Studies on performance projections using a combination of data measurements and the exposure of test components to radiation levels similar to the ones expected for the **HL-LHC** show clearly that both the current tracker and the current endcap calorimeters cannot withstand the future conditions. A short description of the upgrade of these two sub-detectors including also upgrades of the other **CMS** sub-detectors is listed here.

- **Tracker:** The current **CMS** tracking system must be replaced with a new one during **LS3** due to the radiation damage [92]. The new sub-detector consists of silicon sensors, pixels and strips, with significantly increased granularity and covering a larger forward acceptance. Moreover, a significant new feature of the Phase-2 tracker is the capability of providing tracking information at the level of the **L1** trigger at 40 MHz. In the current situation, this information is only available at the **HLT**. This allows the trigger rates to be kept at a sustainable level without sacrificing physics potential. A more detailed description of the upgrade of the **CMS** Phase-2 tracker can be found in Section 5.3.
- **Barrel calorimeters:** Data collected in the **LHC** runs suggests that the radiation damage on the **EB** and **HB** sub-detectors will be acceptable for the full **HL-LHC** operations [121]. However in order to maintain exactly the same physics performance, the **FE** electronics of both sub-detectors must be improved to satisfy the new **L1** trigger requirements. Moreover, in the case of the **EB**, the new **FE** boards will allow the exploitation of the information from single crystals in the **L1** trigger, providing more precise timing resolution and helping to mitigate the increasing noise from the photodetectors.
- **Endcap calorimeters:** Similarly to the tracker, also the endcap electromagnetic and hadronic calorimeters will suffer significant radiation damage by **LS3** [122]. The sub-detector in replacement will be a sampling calorimeter with electromagnetic and hadronic sections with excellent transverse and longitudinal segmentation, the so-called **High Granularity Calorimeter (HGC)**. Thanks to its high granularity, it will provide the possibility to reconstruct a very detailed 3-D image of the electromagnetic and hadronic particle showers. A schematic representation in the $r - z$ view of the **HGC** detector is shown in Figure 5.9, divided in **Electromagnetic Compartment (CE-E)** and **Hadronic Compartment (CE-H)**. The **CE-E** and a large fraction of **CE-H** use hexagonal silicon sensors of 0.5 to 1 cm² cell size as active elements, while the remainder of the **CE-H** features highly segmented scintillators with SiPM readout. The **CE-E** has a depth of approximately $26X_0$ and $1.7\lambda_I$ and consists of 28 layers of 1.4 mm thick with 25%Cu-75%W baseplate as absorber. The **CE-H** consists of 12 planes of stainless steel plates which are 35 mm thick and are followed by another 12 stainless steel planes with a thickness of 68 mm as absorber. The **HGC** detector accounts for a depth of $10.7\lambda_I$ in total, including **CE-E** and **CE-H**.
- **Muon detectors:** Several studies have been performed since 2015 in order to understand whether the current muon chambers are able to cope with the increased particle rates of the **HL-LHC** [123]. No significant deterioration of key chamber parameters has been observed, and therefore the chambers themselves can be used until the end of the **HL-LHC** operation. Some developments are instead foreseen for the **FE** electronics in the case of the **DTs** and the **CSCs**, which will be replaced with improved versions to increase radiation tolerance, readout speed, and performance. Moreover, in the pseudorapidity region $1.5 < |\eta| < 2.4$ the muon system will be enhanced with improved **RPCs** chambers and new chambers based on the **GEM** technology. The main goals of this forward upgrade are: adding redundancy, improving the triggering and reconstruction performance, and increasing the acceptance in the forward detector region. A $r - z$ view of the muon chambers in the Phase-2 **CMS** detector is shown in Figure 5.10.
- **Trigger and data acquisition:** The selection of interesting physics events at the first trigger stage becomes extremely challenging at the **HL-LHC** because of the increased trigger rate

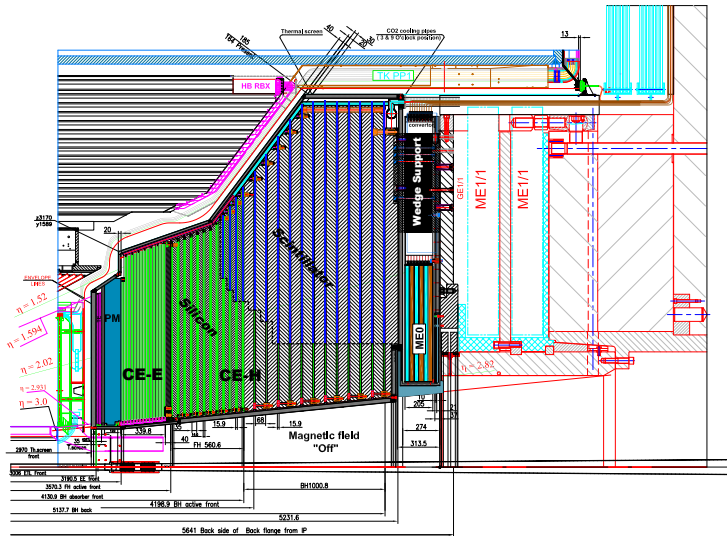


Figure 5.9: Longitudinal cross section of the upper half of one endcap calorimeter [122]. The green region is instrumented with silicon detectors and the blue region with scintillator.

and the larger complexity of the events in high pileup conditions. The trigger and the **DAQ** system will thus undergo an upgrade to face the **HL-LHC** challenges [124, 125]. The **L1** trigger latency will increase from about $4\text{ }\mu\text{s}$ in the current system to a maximum of $12.5\text{ }\mu\text{s}$, allowing to reconstruct tracks in programmable hardware. Additionally, the maximal read-out rate will increase of up to 750 kHz compared to the current 100 kHz . As a consequence, the **DAQ** system will be upgraded to implement the increase of the bandwidth and the computing power required by the larger size of the events and by the new **L1** trigger rate.

- **Timing detector:** The possibility of adding a timing sub-detector both in the barrel and in the endcap is also under examination [126]. The idea is that this additional layer will provide a very precise time measurement for each individual track crossing it. This information can be used to improve the physics performance and to counteract the extreme pileup conditions. Several studies have already shown the benefits that this highly specialized sub-detector can bring in terms of object reconstruction and physics analysis, such as improved track and vertex reconstruction abilities, higher lepton efficiencies and diphoton vertex location, and better jet identification. In Figure 5.11 the advantage of including the timing information in vertex reconstruction is clearly shown: the 4-D vertices are reconstructed adding an extra dimension in the vertex algorithm currently used in **CMS** [127] and can be easily discriminated from the 3-D vertices, i.e. their projection on the z -axis. Several technologies are still under study which can withstand the radiation conditions, provide a timing resolution of about $10 - 30\text{ ps}$ and have a minimal impact on the neighboring sub-detectors. In the barrel, a detector based on lutetium-yttrium orthosilicate crystals activated with cerium (LYSO:Ce) read out with SiPMs is proposed, while in the endcap, **MIP**-sensitive silicon technology is proposed arranged in an hermetic single layer.
- **Software and Computing:** In the **HL-LHC** scenario at 140 (200) pileup, the offline software and the computing areas would fail to reach the resources needed by a factor of 4 (12), assuming only technology improvements. A significant R&D program has thus already started to improve the algorithms used for data reconstruction, analysis, storage and access, and to adapt the **CMS** software and computing model to new future technologies.

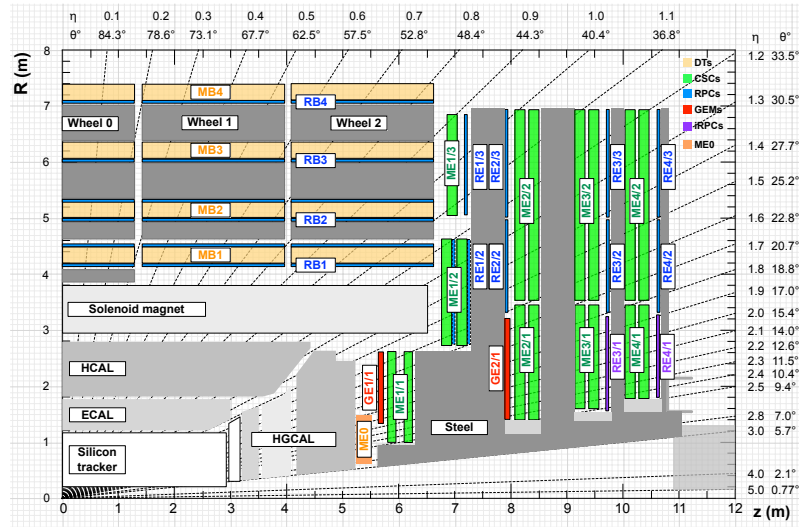


Figure 5.10: An $r - z$ cross section of a quadrant of the CMS detector, including the Phase-2 upgrade of the muon detectors [123].

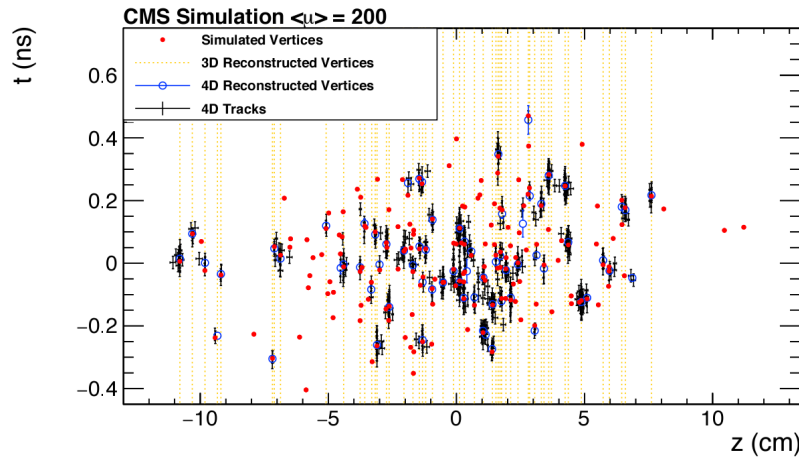


Figure 5.11: Space-time structure of simulated (red dots) and reconstructed vertices with (blue dots) and without (yellow lines) the information of a fully covering fast-timing layer with timing resolution per track of about 20 ps [128]. A $H \rightarrow \gamma\gamma$ event is simulated with superimposed pileup of 200.

5.3 The Phase-2 upgrade of the CMS tracker

The present Strip Tracker is designed to operate efficiently up to an integrated luminosity of 500 fb^{-1} and an average pileup of 30 to 50 collisions per bunch crossing. Figure 5.12 shows the expected performance of the current tracker in the HL-LHC conditions. The unacceptable level of degradation due to radiation effects beyond 1000 fb^{-1} is one of the main driving forces for a substantial upgrade of the CMS tracking system [116, 92]. A second driving force is the need to include the track information in the L1 trigger. The L1 trigger system foresees to have an output rate of 750 kHz, and this goal appears to be impossible to achieve using only the information coming from the calorimeters and the muon detectors. In order to face both these challenges, CMS will install a new silicon-based tracker. The new Phase-2 tracking detector consists of an Inner Tracker (IT) composed of silicon pixels modules, and an Outer Tracker (OT) composed of silicon modules with strip and macro-pixel sensors.

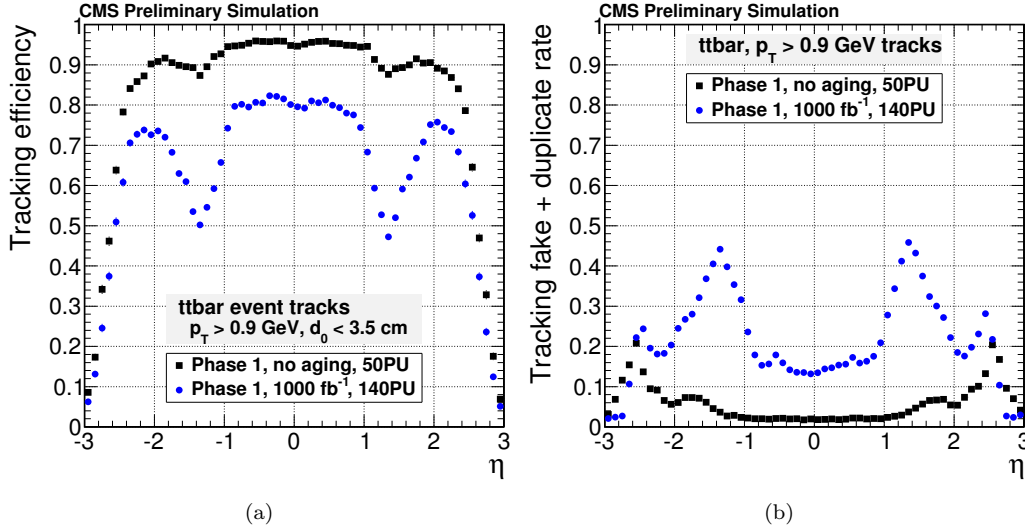


Figure 5.12: Tracking efficiency (a) and fake rate (b) as a function of pseudorapidity for a $t\bar{t}$ sample with nominal pileup conditions for the Phase-1 detector before (black) and after (blue) the Strip Tracker has been aged by an equivalent of 1000 fb^{-1} [116].

5.3.1 Requirements at the HL-LHC

The Phase-2 **CMS** tracker has to fulfill the following requirements:

- **High radiation tolerance:** The new tracker must be able to operate efficiently for the entire **HL-LHC** operation period up to an integrated luminosity of 3000 fb^{-1} . In [Figure 5.13](#) the expected particle fluence within the tracker region is shown. In the case of the **OT**, no maintenance interventions are foreseen during the entire **HL-LHC** operation, while the pixel modules of the **IT** are foreseen to remain accessible and can be replaced as they accumulate substantial radiation damage.
- **Increased granularity:** In the harsher conditions prevailing at the **HL-LHC**, at every beam crossing the tracker will be traversed by about 6000 charged particles with transverse momentum above 300 MeV produced by about 200 collisions on average. In order to ensure efficient tracking performance and to limit the problem due to combinatorics in both the seeding and the track finding phase, the channel occupancy must be kept at around or below the per cent level in the **OT** and below the per mille level in the **IT**. This requirement allows also to avoid cluster merging and to maintain a good two-track separation in the **IT**, which is particularly important for track finding performance in highly energetic jets.
- **Extended tracking acceptance:** The **HL-LHC** physics program mentioned in [Section 5.1.2](#) involves many analyses which will definitely benefit from a larger tracker acceptance region, such as the ones including the vector-boson scattering measurements. Therefore the new tracker is designed to extend the coverage in η up to $|\eta| = 4$.
- **Reduced material:** The tracking performance and the overall physics performance depends strongly on the amount of material in the tracking volume. Therefore a lighter tracker option has been chosen. The material budget is reported in [Figure 5.14\(a\)](#): At $\eta = 0$ and in the most forward region, i.e. $3.0 < |\eta| < 4.0$, the total amount of material inside the tracking volume is estimated as 30% of a radiation length. In the **OT** region, it rises up to 80%.
- **Contribution to the **L1** trigger:** As already mentioned, the **CMS** trigger will operate at the **HL-LHC** with substantially increased latency and output rate. In order to maintain an acceptable output rate despite the tenfold luminosity, it is clear that the tracker must be involved into the **L1** trigger decision. In [Figure 5.14\(b\)](#) the effect of the inclusion of the tracking information is shown in term of trigger rates as a function of the trigger threshold

for track-matched electron objects: The combination of calorimeter and tracking information provides a further reduction in the trigger rate, while keeping an acceptable efficiency [124].

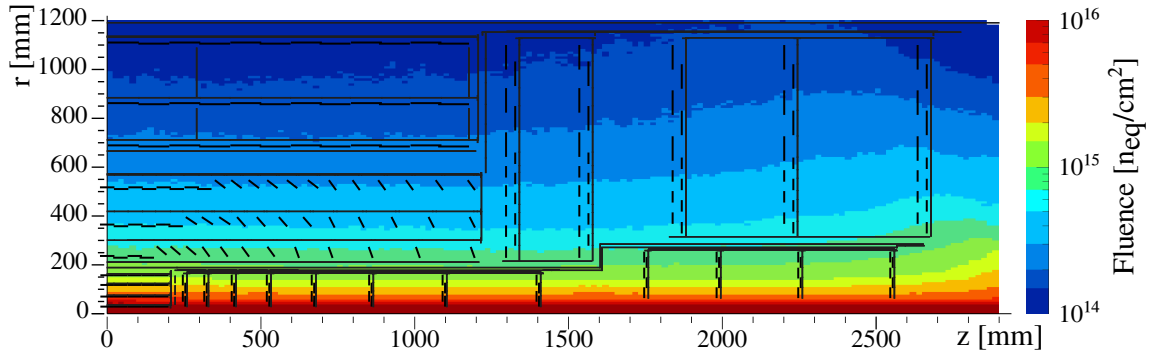


Figure 5.13: Integrated particle fluence in 1 MeV neutron equivalent per cm^2 within the Phase-2 tracker region. The estimates shown correspond to a total integrated luminosity of 3000 fb^{-1} of pp collisions at $\sqrt{s} = 14 \text{ TeV}$ [92].

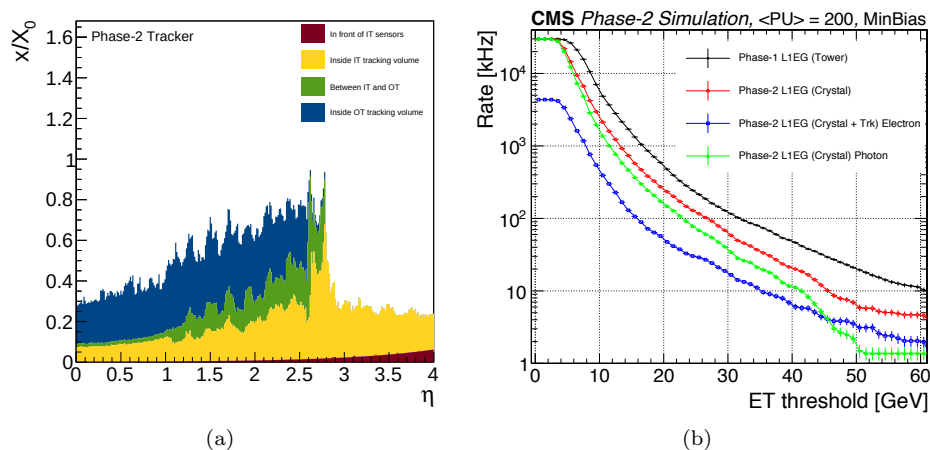


Figure 5.14: (a) Material budget inside the tracking volume estimated in units of radiation lengths for the Phase-2 tracker. The stacked histogram shows the material in front of the **IT** (brown), inside the **IT** (yellow), between **IT** and **OT** sensors (green), and inside the **OT** tracking volume (blue) [92]. (b) Expected rate for minimum bias events using the single electron calorimeter trigger as a function of trigger threshold [124].

Figure 5.15 shows a schematic view of a possible layer configuration of the new Phase-2 tracker including both the **IT** and the **OT**.

In this layout, the **IT** has four barrel layers, as already implemented in the Phase-1 upgrade, and twelve endcap discs. The new pixels have an area of about $2500 \mu\text{m}^2$, which is smaller by a factor of six compared to the current **CMS** Pixel Tracker. This feature enables to achieve low occupancy in this region, particularly important in the context of the pixel-based track seeding and improved track separation in track dense environments. Different pixel sizes and thicknesses in the **IT** sub-detector are under consideration. More information about this topic can be found in Section 5.3.2. It is important to notice that the number of detection layers in the forward part assures a robust performance over the whole rapidity acceptance up to $|\eta| \simeq 4.0$.

In the **OT** sub-detector, six barrel layers and five end-cap discs are foreseen. As clearly shown in Figure 5.15, part of the **OT** modules in the three innermost barrel layers are arranged in a tilted scheme. The reason for this choice is discussed in more detailed in Section 5.3.5. The modules in the **OT** are designed to give a contribution to the **L1** trigger. To this purpose, each module is a stack of two closely spaced sensors and is able to correlate hits produced in the two sensors in real time. This allows the rejection of low- p_T particles using the strong bending power of the

3.8 T magnetic field. A detailed description of the two different kinds of modules, **Pixel-strip** (PS) and **2-strips** (2S), used in the **OT** can be found in Section 5.3.3. The six layers of the **OT** are the minimum required to ensure robust track finding at the **L1** trigger in the rapidity acceptance of $|\eta| < 2.4$.

The global number of layers and the overall geometrical layout has been optimized using the tkLayout tool [129, 130] as well as the full simulation to ensure that performance are unaffected when one detecting layer is not operating or some parts are lost. The resulting excellent performance expected to be achieved with the upgraded tracker for the offline tracking is discussed in Chapter 6.

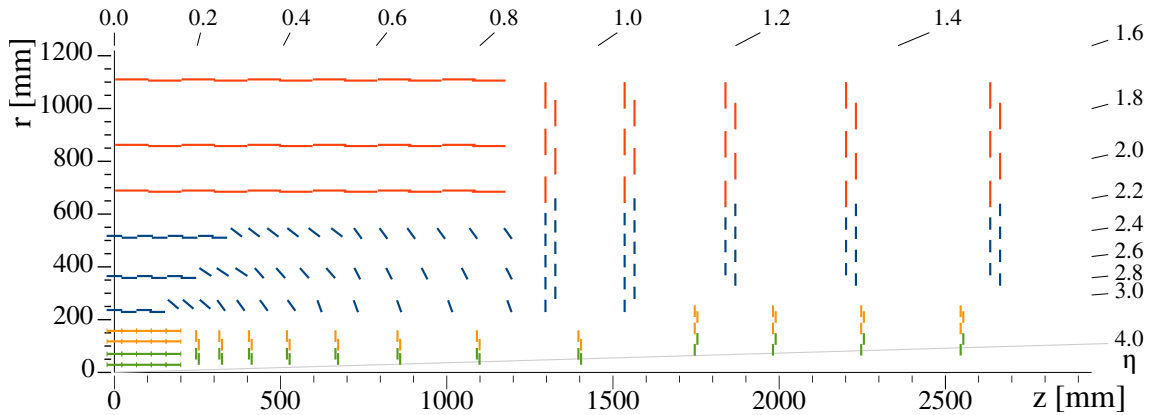


Figure 5.15: $r-z$ view of the Phase-2 CMS tracking system. The **IT** modules made with two readout chips are shown in green, while the one with four readout chips in yellow. Each green line corresponds to pixel modules made of two readout chips, while each yellow line to pixel modules with four readout chips. The **OT** is represented with blue and red lines for **PS** modules and for the **2S** modules, respectively [92].

5.3.2 The Inner Tracker

The **IT** sub-detector is designed to maintain or improve the tracking and vertexing capabilities under **HL-LHC** conditions. It is particularly important, given its crucial role in the track seeding procedure and in the two-track separation problem. Extended pseudorapidity coverage, low occupancy and reduced material budget are the driving forces of the design of the sub-detector layout, the sensors and the services [92].

- Layout:** The **IT** structure is composed of **Tracker Barrel Pixel** (TBPX) layers in the barrel region and **Tracker Barrel Forward** (TFPX) and **Tracker Barrel Endcap** (TEPX) discs in the endcap region. The four cylindrical layers of **TBPX** extend from $r \simeq 2.9$ cm up to $r \simeq 16$ cm radially and up to $|z_{\max}| \simeq 20$ cm along the beamline. Modules in **TBPX** are arranged in *ladders* with no overlap in z . In each layer, ladders which are neighbours are mounted staggered in radius with the aim of achieving a $r-\phi$ overlap between different ladders. In order to avoid a gap at $|\eta| = 0$, an odd number of modules is foreseen along z . In the forward region, the eight discs of **TFPX** are placed between $|z| = 25$ cm and $|z| = 140$ cm, while the four discs of **TEPX** cover the very forward region up to $|z| \simeq 255$ cm. The **TEPX** discs have a larger inner radius compared to the **TFPX** ones, $r \simeq 6.5$ cm, and extend up to $r \simeq 25$ cm in order to guarantee a uniform coverage in the entire spectrum of η . The modules in the **TFPX** and **TEPX** are arranged in concentric *rings*. Each disc is physically made by two twin discs which facilitate to introduce overlaps in r and in $r-\phi$. Each mechanic disc is made of two flat “D”-shaped carbon fiber supports, referred to as *dees*. In total, the number of modules in the barrel and the encap part of the **IT** amounts of 864 and 3488, respectively. The total instrumented surface is about 4.5 m^2 .
- Sensor design:** The idea is to use thin planar n-in-p type silicon sensors with the option of introducing 3-D sensors in limited part of the sub-detector. The 3-D sensors are particularly interesting in the **HL-LHC** context because they are intrinsically more radiation hard thanks

to the shorter charge collection distance, but they have the disadvantage that their production process is significantly more expensive and therefore not suitable for large volumes. A possible solution is, for example, to use them only in the regions of highest particle fluences. Regarding the planar sensors, several possibilities about the final choice of the thickness and the size are still under examination [131]. The pixel thickness is foreseen to be in the range of $100 - 200 \mu\text{m}$. The $100 \mu\text{m}$ thick sensors are as efficient in collecting charge as the $200 \mu\text{m}$ thick ones, but operate at a 200 V smaller bias voltage. A drawback of the $100 \mu\text{m}$ thick sensors is the smaller signal before irradiation and that they are more prone to bowing. Two possibilities of pixel segmentation have been studied: $25 \times 100 \mu\text{m}^2$ and $50 \times 50 \mu\text{m}^2$. The rectangular pixels are found to have a spatial resolution better than square pixels in the transverse plane and perform similarly in the z -coordinate. However, these studies do not take into account possible effects of radiation damage and technical challenges and, as a consequence, the discussion is still open.

- **The Read-Out Chip (ROC):** The ROC is one of the challenging key elements of the entire IT sub-detector, because it must be highly radiation tolerant, have a low noise figure to process the small signal from small pitch pixels on thin sensors, and a deep readout buffer and a fast readout rate to comply with the latency and the rate of the trigger. A possibility to overcome the last problem is to store hits in a pixel array within multi-pixel regions, such as 2×2 or 4×4 . A first prototype of this ROC using the 65 nm CMOS technology was developed by the RD53 group, a joint ATLAS-CMS collaboration [132]. The active size of the ROC will be around $16.4 \times 22.0 \text{ mm}^2$.
- **Modules:** A IT pixel module consists of a pixel sensor, several ROCs, a flex circuit¹, and a mechanical support. The overall concept is similar to the technology used currently in the Pixel Tracker: pixel sensors are bump bonded to the ROCs; the flex circuit is glued onto the sensor and wire bonded to the ROCs; a module support strip is glued to the back side of the ROCs to facilitate the assembly of the module itself on the mechanical support structure and to ensure an efficient heat removal. The pixel module is connected to the global readout, control and powering system through low mass electrical cables. Moreover, CO₂ cooling pipes are used to remove the heat generated on the module keeping the pixel chips and sensors at an operating temperature of about -20°C . A detailed view of the two types of pixel modules foreseen in the IT are shown in Figure 5.16. They differ in the sensor surface and in the number of ROCs, which can be two or four depending on the match of the input specifications with the output rate of a module. More details on this subject can be found in Ref. [131].

5.3.3 The Outer Tracker

The main requirements for the OT are partially the same already mentioned for the IT, such as high radiation tolerance, temperature operation around -20°C , increased granularity, large readout bandwidth and deep FE buffers. On top of these requirements, the OT must have trigger capabilities to contribute to the L1 trigger [92].

- **Layout:** In the OT layout structure three sub-detectors are distinguished: the Tracker Barrel with PS modules (TBPS); the Tracker Barrel with 2S modules (TB2S); and the Tracker Endcap Double-Discs (TEDD). The subdivision in ladders in the barrel and dees in the endcap is very similar to the IT case. This layout ensures that at least six layers in the rapidity range $|\eta| < 2.4$ are crossed by particles emerging from the interaction region with $|z| < 70 \text{ mm}$. The only exception is the transition region between barrel and endcap, around $|\eta| \simeq 1.0$, where the number of crossed modules drops to five. This feature is very important to provide a robust track finding for the L1 trigger [133]. For the same reason three layers of the TBPS are chosen: at least three macro-pixel layers are needed to measure the polar angle of the track and provide some first measurement of the vertex, including a minimal

¹A flex circuit is a technology for assembling electronic circuits by mounting electronic devices on flexible plastic substrates.

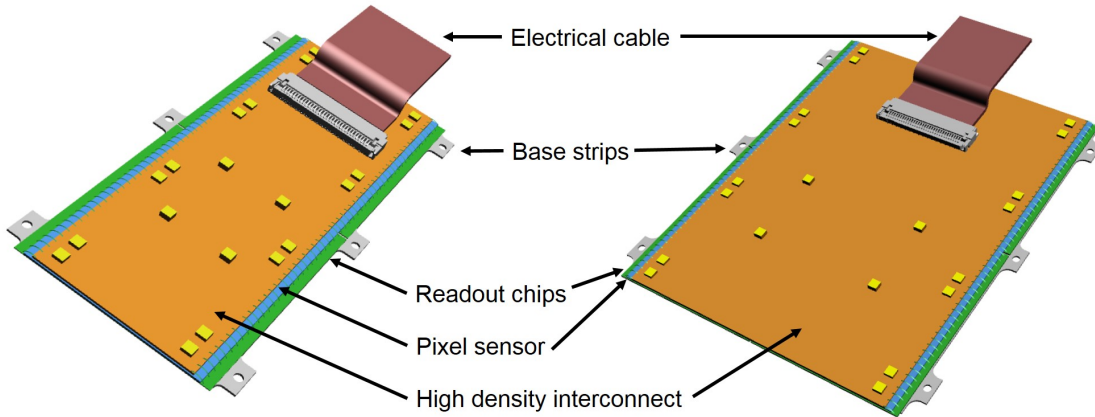


Figure 5.16: Sketch of the two types of pixel modules foreseen in the **IT**. The modules with two number of **ROCs** is shown on the left, while the ones with four number of **ROCs** is shown on the right. The yellow elements symbolize passive electrical components [92].

redundancy. Figure 5.15 shows one possible arrangement of the foreseen 13296 **OT** modules with a total active surface of about 198 m².

- **The p_T modules:** The p_T modules are the building blocks of the entire **OT** and play a key role in the **L1** trigger capability of the Phase-2 tracker. To include tracks in the **L1** trigger the main idea is to correlate two signals coming from two closely-spaced sensors and to check their compatibility with the hypothesis of both coming from a track with a p_T above a certain threshold. In order to implement this idea in a mechanical structure, the p_T modules are composed of two silicon sensors with a gap between the mid-planes of the active volumes of the sensors. Both sensors are read out by a common **FE** electronics with is capable to combine signal pairs. The compatibility is defined in hardware through a selection window. If the signal pairs are inside the selection window, they are combined into *stubs*. Stubs are then dispatched to the **L1** trigger system at every bunch crossing. All other hits in the **OT** modules are stored in the **FE** pipelines and read out when a trigger is received. A simplified illustration of the concept is shown Figure 5.17.

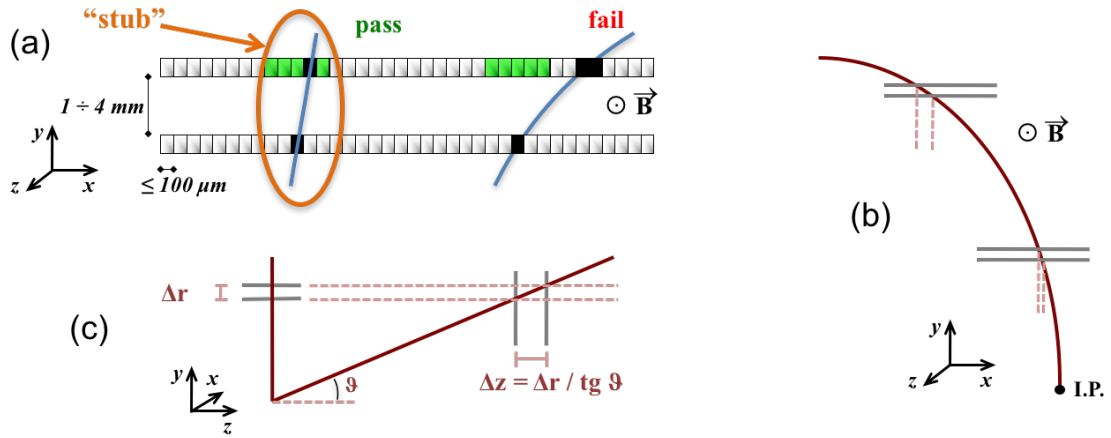
As already mentioned, two different kind of modules for the final configuration of the **OT** are under study: the **PS** modules composed of a bottom sensor with $1.5 \text{ mm} \times 100 \mu\text{m}$ pixels and a top sensor with $2 \times 25 \text{ mm}$ strips with $100 \mu\text{m}$ pitch; and the **2S** modules composed of two sensors with $2 \times 5 \text{ cm}$ long strips with $90 \mu\text{m}$ pitch. In both cases, the sensitive elements of the modules are planar silicon sensors. The **2S** sensors are mounted such that the strips of the two sensors are parallel to each other. The precise measurement of the z -coordinate coming from the macro-pixel in the first type of modules is particularly important in the **L1** trigger for primary vertex discrimination and for robust pattern recognition. Moreover, to achieve efficient rejection of low- p_T particles in the entire tracker volume the selection window must be programmable in the **FE** electronics, and the modules must have different sensor gaps depending on their position. A sensors gap of 1.8 and 4.0 mm is foreseen for **PS** modules, while in the case of **2S** modules the gap can be 1.6, 2.6 and 4.0 mm. The main parameters of **2S** and **PS** modules are summarized in Table 5.3. Images of the assembled modules and further details of their components are shown in Figure 5.18.

- **The readout:** The electronic system is designed to deliver trigger data with the latency and rate constraints imposed by the 40 MHz trigger. The **FE** readout in both modules delivers binary data, and, as a consequence, the electronic system is fully digital with the exception of the analogue ASIC **FEs**. The **2S** module electronics is based on the CBC2 (CMS Binary Chip), a full-scale prototype **FE** readout ASIC; instead the readout for the two sensor types of the **PS** module requires the development of two **FE** ASICs: the Short Strip ASIC (SSA) and the Macro-Pixel ASIC (MPA), both implemented in a 65 nm CMOS process. Data generated

Table 5.3: Main parameters of the 2S module and the PS module of the CMS tracker.

PS module	2S module
$\simeq 2 \times 45 \text{ cm}^2$ active area	$\simeq 2 \times 90 \text{ cm}^2$ active area
32×960 macro-pixels: $\simeq 1.5 \text{ mm} \times 100 \mu\text{m}$ 2×960 strips: $\simeq 2.4 \text{ cm} \times 100 \mu\text{m}$	2×1016 strips: $\simeq 5 \text{ cm} \times 90 \mu\text{m}$ 2×1016 strips: $\simeq 5 \text{ cm} \times 90 \mu\text{m}$
1.8 and 4.0 mm sensors gaps	1.6, 2.6, and 4.0 mm sensors gaps

by the readout hybrid chips (CBC2 in the 2S module, and SSA and MPA in the PS module) are buffered, aggregated and formatted by the concentrator ASIC (Concentrator Integrated Circuit, CIC) that acts as a database to the service hybrid. The task of the service hybrid is thus to host all connections to/from the back-end and associated electronic components: a 5 Gb/s data link, an optical converter, and the DC/DC converter that provides power to the module electronics. The service hybrid connecting different individual services makes each CMS OT module an autonomous element in the entire tracker system. The 2S module carries one service hybrid, while for the PS modules the service hybrid has to be split into two physical boards. In both modules, the sensors are connected to the readout hybrid and to the auxiliary electronics for powering and readout using wire bounding at the edge of the modules. Sketches of the connectivity between sensors and readout hybrid circuits is shown in Figure 5.18 for both modules type.

Figure 5.17: Illustration of the p_T module concept.

- (a) The correlation of the two signals in the sensors allows the rejection of low- p_T track using the selection window represented in green.
- (b) At larger radii the two signals coming from the same p_T track have larger distance.
- (c) For the endcap discs, a larger spacing between the sensors is needed to achieve the same discriminating power as in the barrel at the same radius. [92].

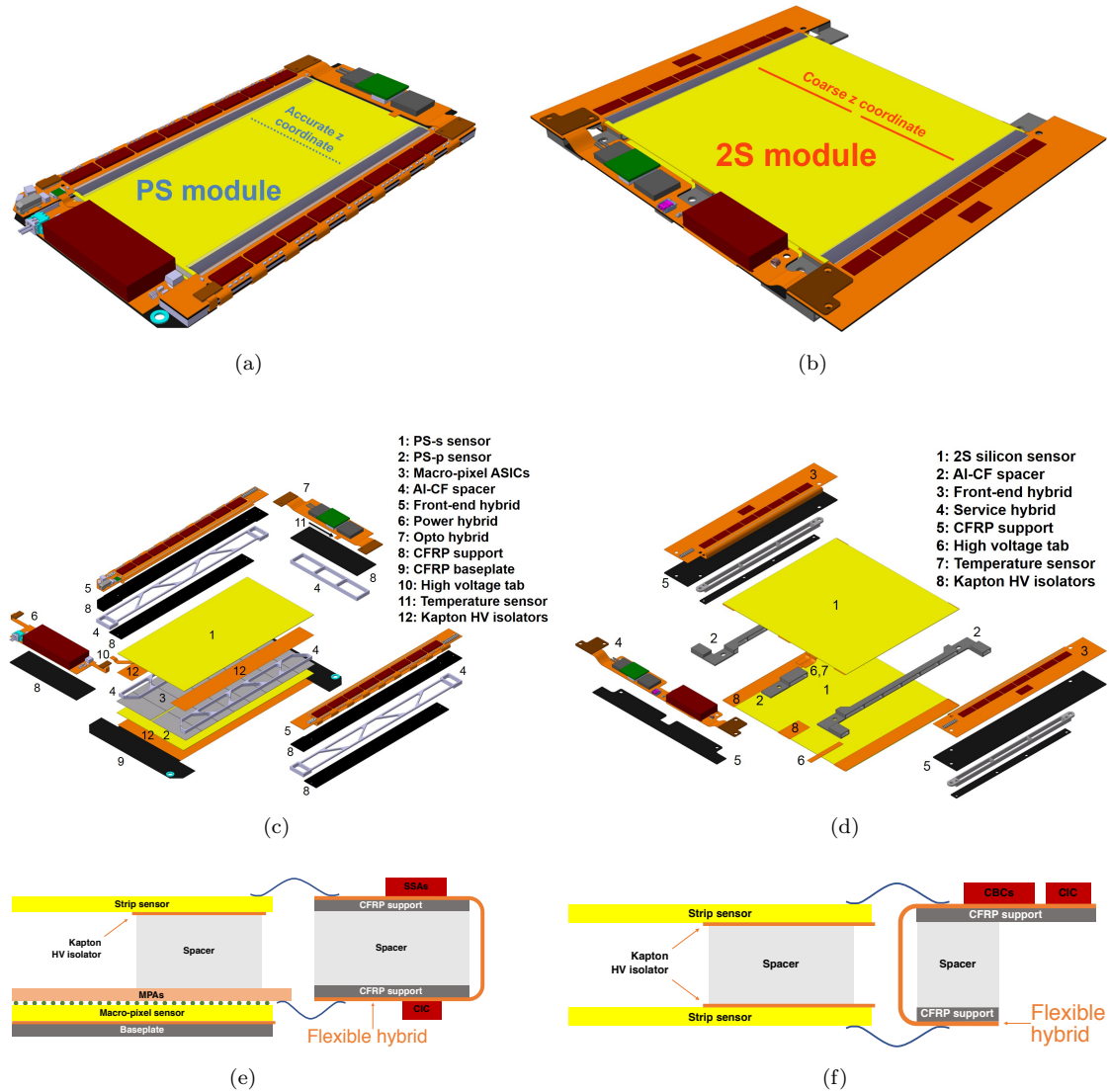


Figure 5.18: The PS module (a,c,e) and 2S module (b,d,f) of the OT. Shown are views of the assembled modules (a,b), details of the module parts (c,d), and the FE hybrid folded assembly and connectivity(e,f) [92].

5.3.4 Services

In order to mitigate radiation damage, silicon sensors need to be operated at a temperature of -20°C or lower [92]. A two-phase CO_2 cooling system is designed to remove up to 100 kW from the OT and 50 kW from the IT, taking into account the power dissipated by the front-end electronics, the leakage current in the silicon sensors (increasing with irradiation), as well as heat leaks from the surroundings.

The services connecting the current tracker to the back-end systems will be completely removed and replaced with new ones, designed for the upgrade tracker. They include low and high voltage power cables, cooling pipes, optical fibre bundles, dry gas pipes and sniffing pipes, as well as some additional cables for hardwired temperature and humidity measurements.

For the IT, all services have connections at Tracker bulkheads, to facilitate the removal of the detector for maintenance. For the Outer Tracker instead, power cables and optical fibre bundles have pigtails reaching the patch panels located on the inner wall of the magnet cryostat, while cooling and dry gas pipes are the only services having connections at the bulkheads.

5.3.5 The Level-1 Track Trigger

As already partially discussed in Section 5.3.3, the design of the layout and the modules of the OT is largely driven by the L1 trigger needs. The p_T modules described above are the fundamental blocks of the future L1 track finder in CMS. Their concept implies that both top and bottom sensors must be connected to the same readout electronics to perform stub finding, which in this specific context is the selection of all the pairs of signals consistent with tracks with $p_T > 2$ GeV with high efficiency². In order to implement the connectivity between the upper and lower sensors with reliable and affordable technologies, each row of strips of the two halves of each module is read out by a separate FE hybrid. This lack of communication in the module prevents the reconstruction of stubs when particles cross the module near the centre with a large incident angle. The problem that arises in the case of a flat layout is represented schematically in Figure 5.19.

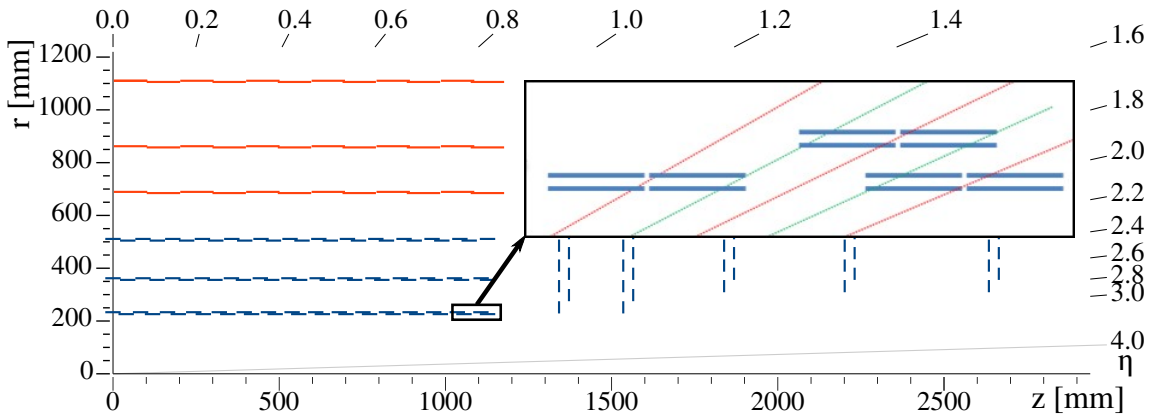


Figure 5.19: Schematic representation of the geometrical inefficiency for PS modules for on-module stub finding in the case of a flat barrel [92].

The stub finding inefficiency in the case of a flat layout is more severe in the innermost part of the OT because of bigger incident angles, the shorter strip length and the larger sensor spacing. This problem can be significantly alleviated by introducing a tilted TBPS sub-detector. In Figure 5.20(a) the stub finding efficiency is presented in the case of the flat and tilted layout for stubs with $p_T > 10$ GeV in the innermost TBPS layer. An efficiency above 90% is obtained over the full range for the tilted geometry, while the flat geometry leads to $\simeq 40\%$ efficiency at large pseudorapidities. Studies on the the expected performance of the offline tracking for these two layouts are presented in Chapter 6.

In Figure 5.20(b) the stub reconstruction efficiency in the barrel layers for simulated muons with $1 < p_T < 10$ GeV is shown as a function of the p_T of the simulated particle. The turn-on curves are sharp for all layers, providing high efficiency at the target threshold of 2 GeV and showing the good transverse momentum discrimination of the tilted layout. A uniform p_T threshold throughout the tracking volume is achieved by tuning the acceptance window, which is programmable in the ASICs of each module, depending on the module location in the detector. Once the stubs are found, they are sent to the off-detector L1 tracking system, which reconstructs tracks for input to the CMS L1 trigger.

Track reconstruction in the L1 trigger is very challenging. Preliminary studies estimate an order of magnitude of 10^4 stubs at pileup 200 received by the L1 tracking system, which must reconstruct tracks and take decisions in approximately $5 \mu\text{s}$, including $1 \mu\text{s}$ required to transmit data from the detector to the system. In order to simplify this challenge, the track finding problem can be divided into data organization, pattern recognition, track fitting, and duplicate removal stages. Three different approaches for the L1 track trigger are being pursued by different teams in CMS. All approaches include the use of FPGAs, they all organize data regionally either in η or in ϕ , and distributed in a round-robin time-multiplexing fashion. The main difference between them is the technique used in the pattern recognition process. The approaches are:

²The possibility of a higher threshold in p_T is also under study.

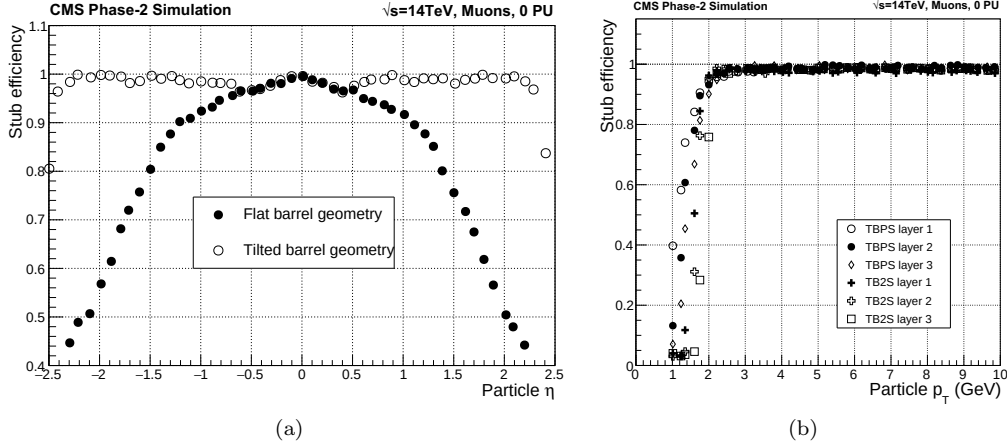


Figure 5.20: (a) Stub reconstruction efficiency for stubs with $p_T > 10$ GeV in TBPS layer 1 as a function of the pseudorapidity of the simulated particle for the flat (solid points) and tilted (open circles) layouts [92]. (b) Stub reconstruction efficiency for muons as a function of p_T in the barrel (left) and in the endcaps (right) [92].

- **Associative Memory plus FPGA approach:** The possible track candidates are found in a very fast way using the associative memory ASICs technology. This chip uses coarse stub position information to perform the pattern recognition with a match against pre-computed track patterns. Studies have shown that around 100 million reference patterns are needed for the full OT to achieve robust tracking. In a second step, the stubs belonging to the selected track candidates are used in full-resolution to perform the track fit within the FPGA.
- **FPGA-based Hough transform approach:** All data from a single event are sent through a single data processing module for assembly and processing. The Hough transform algorithm [134] is used in a highly parallelised first stage which creates track candidates by coarsely grouping stubs consistent with the hypothesis of a high p_T track. In a second stage the Kalman Filter algorithm is used to fit the track candidates and remove possible fakes.
- **FPGA-based tracklet approach:** FPGAs are used to implement the traditional Kalman Filter road search where the search starts from track seeds, referred to as tracklets, built from pairs of stubs in adjacent layers or double-discs. The track candidate reconstruction then proceeds including compatible stubs. A linearized χ^2 fit determines the final track parameters, and the process concludes with a duplicates removal step.

For each approach hardware demonstrators have been constructed, and the algorithm has been tested with a $t\bar{t}$ sample with pileup of about 200. All three approaches are able to deliver tracks with the needed performance and within the timing requirements. Some examples of L1 track trigger performance are shown in Figures 5.21 and 5.22. These results are based on the flat barrel geometry. The first set of plots shows the L1 tracking efficiency for prompt muons and electrons reconstructed in the $t\bar{t}$ sample mentioned above. In the case of muons, the sharp turn-on curve is clearly visible at around 2 – 3 GeV, and the efficiency is approximately flat in the entire range of pseudorapidity at the level of about 98%. In the case of electrons, the efficiency turns up more slowly and flattens out at $\simeq 90\%$, mostly due to interactions with the detector material. The second set of plots shows the excellent resolutions of the p_T and z_0 parameters of muons with $p_T > 10$ GeV in $t\bar{t}$ events for different pileup scenarios. The resolutions are defined in terms of an interval centered on the residual distribution that contains 68% or 90% of the tracks. As expected, a degradation of the resolution appears due to a corresponding increase of multiple scattering.

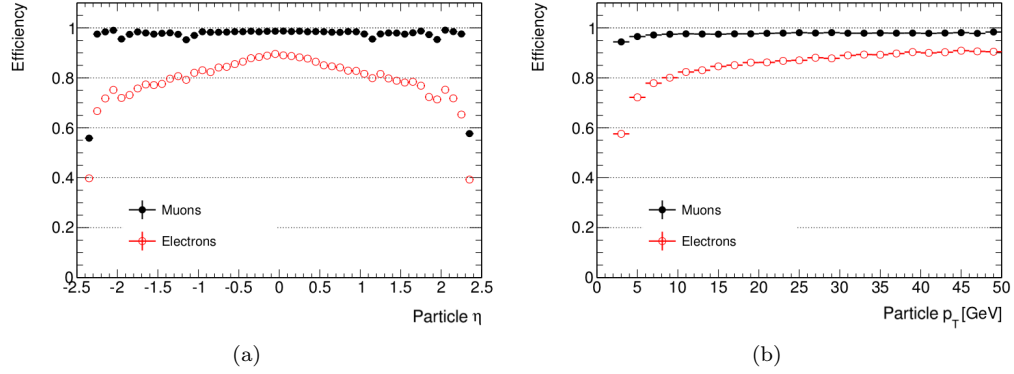


Figure 5.21: **L1** tracking efficiency for muons (filled black dots) and electrons (open red dots) as a function of the generated particle η (a) and p_T (b). In the first plot a cut on $p_T > 3$ GeV is applied; while in the second one only particles within $|\eta| < 2.4$ are considered [92].

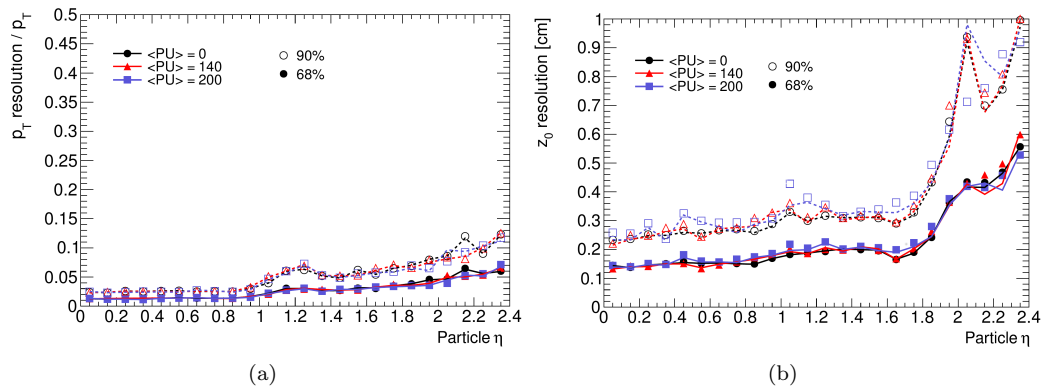


Figure 5.22: Relative p_T resolution (a) and z_0 resolution (b) as a function of the pseudorapidity for muons in $t\bar{t}$ events with zero (black dots), 140 (red triangles), and 200 (blue squares) pileup events on average. Only track with $p_T > 3$ GeV are considered. Different truncation intervals in the track parameters distributions are used: 68% (filled markers and solid lines) or 90% (open markers and dashed lines) [92].

CHAPTER 6

Track reconstruction in Phase-2 CMS

*To Chien-Shiung Wu, the First
Lady of Physics Research. (1957)*

In [Section 6.1](#) the current status of the track reconstruction in the Phase-2 upgrade of the [CMS](#) experiment is described: an introduction to the local reconstruction and the Phase-2 iterative tracking is given as well as the full set of results for different [HL-LHC](#) pileup conditions. Among the possible developments discussed at the end of the section, the possibility of exploiting the new [Outer Tracker](#) (OT) is studied in [Section 6.2](#). Thanks to the OT stacked layout, a new type of hits, so-called vector hits, can be implemented which contain both position and direction information. Their use can be essential to facilitate distinguishing real tracks from a purely random combination of hits. Additionally, a new iteration with seeds only in the OT can be introduced using these vector hits. The aim is to reconstruct tracks coming from displaced vertices. The performance gain obtained by this development is described in detail in [Section 6.3](#) including some results for the long-lived neutral particle K_S^0 .

6.1 Track reconstruction during Phase 2

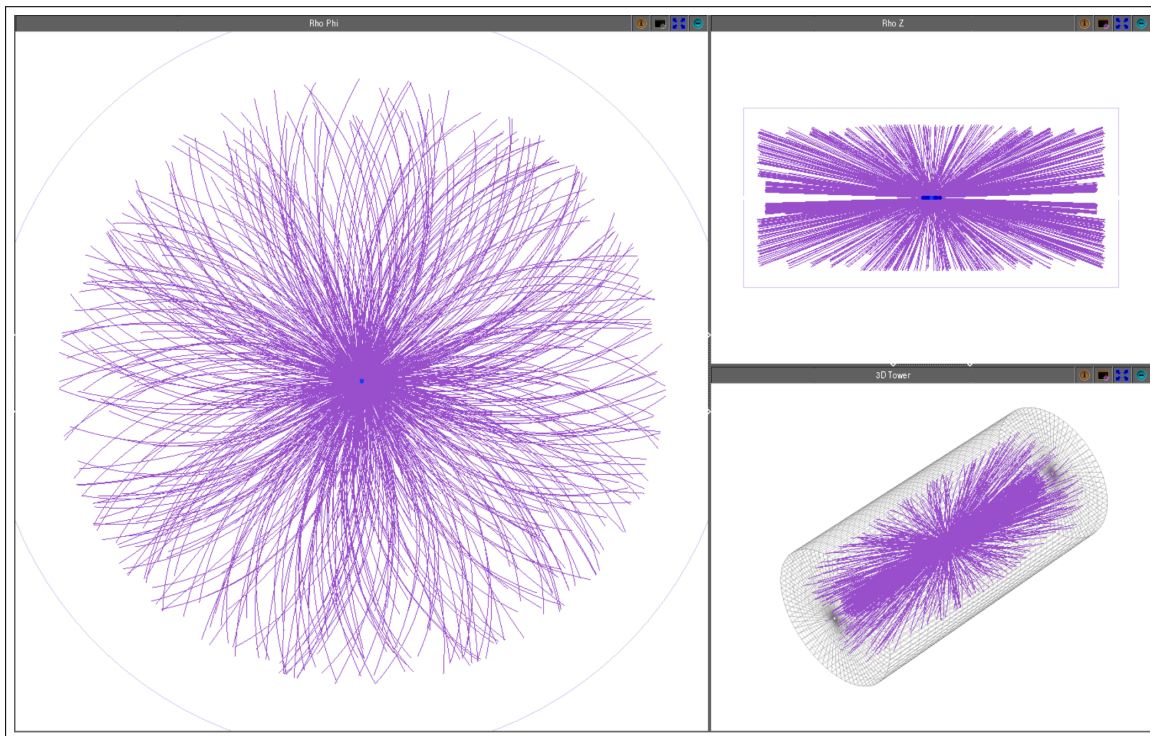
Starting from 2026, the [LHC](#) will achieve a peak of instantaneous luminosity of $5-7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ with a bunch spacing of 25 ns. This new upgraded accelerator, called [High Luminosity LHC](#) ([HL-LHC](#)), is described in [Chapter 5](#) together with the upgrade of the [LHC](#) experiments, with particular emphasis on the Phase-2 upgrade of the [CMS](#) experiment.

In each bunch crossing, the [CMS](#) tracker will be traversed by about 6000 charged particles with transverse momentum above 300 MeV, produced by about 200 pileup events per bunch crossing on average. An example of a reconstructed event simulated with superimposed pileup of about 200 is shown in [Figure 6.1](#). It is zoomed in the tracker region and shown from different perspectives. Track reconstruction and parameter estimation are very challenging tasks in this scenario. In order to provide satisfactory performance under the conditions at the [HL-LHC](#), a new tracker will be installed in [CMS](#). A detailed description of the [Inner Tracker](#) (IT) and [Outer Tracker](#) (OT) sub-detectors in the new tracker along with its design features is given in [Section 5.3](#). A summary of the principal characteristics of the future tracker is given in [Table 6.1](#).

The [CMS](#) collaboration started to approach the track reconstruction problem at [HL-LHC](#) using the same algorithms developed in Run 1 for both local and global track reconstruction. A first set of very preliminary results for the expected performance of the offline tracking for the upgraded tracker

Table 6.1: Summary of the principal characteristics of the Phase-2 **Inner Tracker** and the **Outer Tracker**.

Subsystem	Layers	Pitch	Location in CMS
TBPX	4 cylinders	$25 \times 100 \mu\text{m}^2$	$2.9 < r < 16 \text{ cm}$
TBPS	3 cylinders	P: $100 \mu\text{m}$ S: $100 \mu\text{m}$	$25 < r < 52 \text{ cm}$
TB2S	3 cylinders	$90 \mu\text{m}$	$68 < r < 111 \text{ cm}$
TFPX	8 disks	$25 \times 100 \mu\text{m}^2$	$25 < z < 140 \text{ cm}$
TEPX	4 disks	$25 \times 100 \mu\text{m}^2$	$175 < z < 255 \text{ cm}$
TEDD	5 disks	PS and 2S modules pitch	$131 < z < 265 \text{ cm}$

Figure 6.1: Example of a typical Phase-2 **CMS** reconstructed event simulated with average pileup of about 200. The number of reconstructed tracks and vertices is 5934 and 126, respectively.

was presented in the Technical Proposal [116]. Specific optimizations on the local reconstruction as well as on the iterative tracking were introduced later on in order to move closer to a more realistic description of the interaction of charged particles with the tracker and, as a consequence, also closer to a more realistic reconstruction of the entire event. A second, more detailed, set of results can be found in the Technical Design Report [92].

In the results described in this section, the capability of the new **OT** is not yet exploited in the reconstruction. This topic is explored in **Section 6.2**.

6.1.1 Local reconstruction

All the studies presented in this thesis have been produced with the official **CMSSW** package, where the **CMS** detector response is simulated together with the effect of the 3.8 T magnetic field. The detector geometry and materials are described and simulated using the Geant4 toolkit. The $25 \times 100 \mu\text{m}^2$ modules for the **IT** have been simulated with a thickness of $150 \mu\text{m}$, while for the **OT** modules the thickness chosen is $200 \mu\text{m}$. Moreover, the strip sensors have simulated pitch of $90 \mu\text{m}$ and a length of 50.25 mm in the case of **2S**, while $100 \mu\text{m}$ pitch and 23.13 mm length in the case of the **PS** sensors, respectively. Finally, the macro-pixels in the **PS** modules were simulated with a

Table 6.2: Parameters used in the digitizer for the simulation of the **IT** and **OT** response.

Sensor type	Pixel	PS-pixel	PS-strip	2S
Threshold [e^-]	1200	6300	6300	5800
Noise [e^-]	0	200	700	1000
Cross-talk	5%	5%	5%	5%

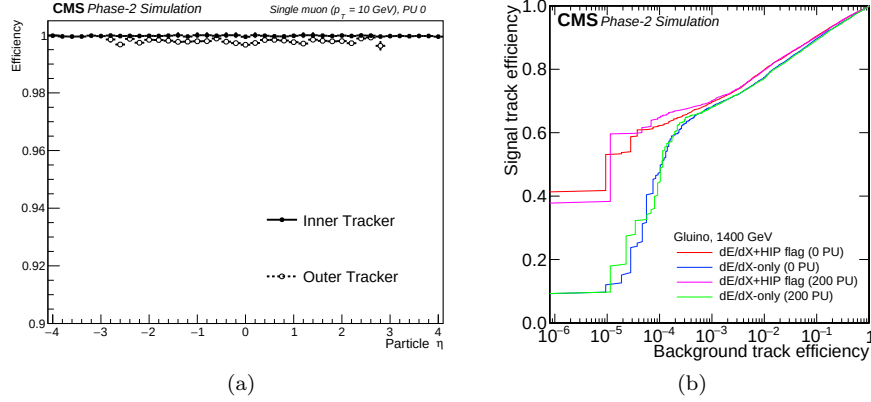


Figure 6.2: (a) Digitization efficiency as a function of pseudorapidity for single muons events generated with p_T equal to 10 GeV, for the **IT** (filled dots) and **OT** (open circles) [92].
(b) Effect on the background rejection for a gluino search without pileup and with 200 pileup with and without the use of the HIP flag in the analysis [92].

pitch of 100 μm and a length of 1446 μm .

In a first step, the data acquisition performed by the readout electronics of the tracker modules is simulated. This step is called *digitization step* and starts from the simulated hit positions taking into account the migration of the charges within the sensors, the Lorentz drift, the diffusion process and the possible energy loss in the sensors. Moreover, also the noise from the readout electronics is added for each channel and to other channels taking into account cross-talk. The main idea behind this first step is to obtain an output, the so-called *digi*, as close as possible to the real data. If the total charge associated with a given channel exceeds a predefined threshold, a digi is created. Different thresholds and readout parameters can be defined for each sensor type; a summary of the parameters chosen is given in Table 6.2. The parameters are continuously tuned taking into account results from the latest beam tests. The digitization step is performed through a specific simulation package developed for both sub-detectors, **IT** and **OT**. The high digitization efficiency is shown in Figure 6.2(a) as a function of pseudorapidity evaluated with single muons with p_T equal to 10 GeV without superimposed pileup.

In the simulation, an extra flag is used in the **PS** modules when the total charge is above the threshold of 1.4 **MIPs**. This flag, also called **Highly Ionizing Particle (HIP)** flag, is particularly important in searches for heavy stable, or quasistable, charged particles (HSCPs) where highly ionizing particles must be distinguished from minimum ionizing particles. The effect of using the HIP flag combined with the energy loss measurement in the **IT**, noted with dE/dx , is shown in Figure 6.2(b).

In Figure 6.3 the occupancy obtained for a minimum bias event with an average of 200 pp interactions per beam crossing is presented. It remains at an acceptable level, compared to the one measured with the current detector, despite the higher level of pileup. It is important to notice that the occupancy in the first **IT** layer and in the macro-pixels of the first **OT** layer in the barrel region are similar.

After the digitization step, the digis are used to build clusters. In the **IT**, clusters are built from contiguous pixels, whose position is defined as the barycenter of the single cluster. In the **OT**, clusters are built in the high-resolution direction only, by aggregating adjacent digi hits. In this

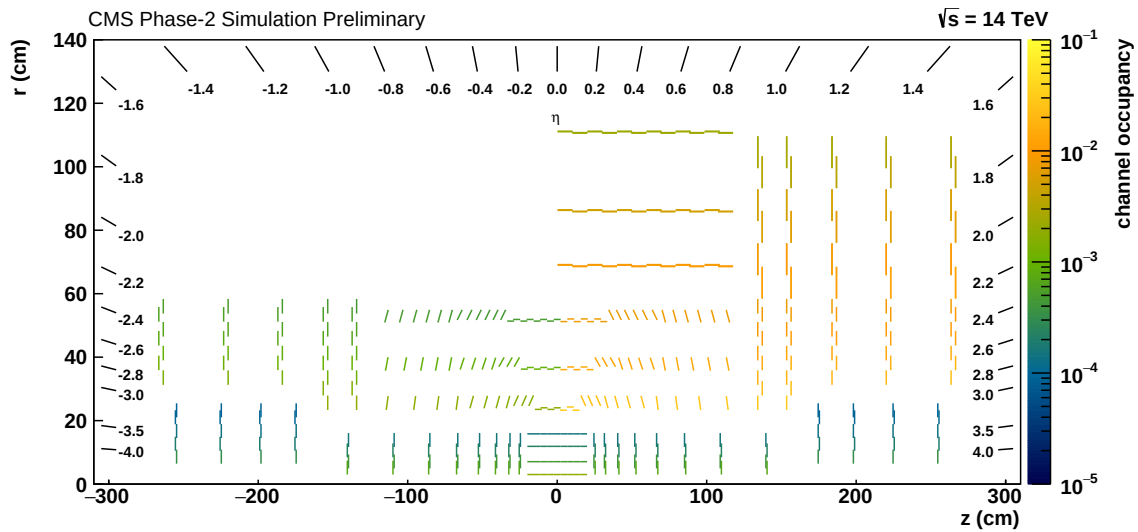


Figure 6.3: Maps of the average fraction of occupied channels per sensor for a minimum bias events with an average of 200 pp interactions per beam crossing, showing one half of the tracker in $r - z$ view. For the OT, the occupancy for the macro-pixel sensors in PS modules is depicted for negative values of z , while the occupancy for the strip sensors in 2S and PS modules is depicted for positive values of z .

Table 6.3: List of different tracking iterations used in the Phase-2 upgrade with the corresponding seeding configuration used and target tracks.

Iteration	Step Name	Seeding	Target Track
0	HighPtQuad	pixel quadruplets	prompt, high p_T
1	HighPtTriplet	pixel triplets	prompt, high p_T recovery
2	LowPtQuad	pixel quadruplets	prompt, low p_T
3	LowPtTriplet	pixel triplets	prompt, low p_T
4	DetachedQuad	pixel quadruplets	from b hadron decay, $r \leq 5$ cm
5	Muon inside-out	muon-tagged tracks	muons
6	Muon outside-in	standalone muons	muons

case, their position is chosen to be their geometrical centre. This cluster position is then corrected for the Lorentz drift, and a position uncertainty is estimated.

6.1.2 Iterative tracking

A detailed description of the iterative tracking method developed in the CMS experiment to perform track reconstruction is given in Section 3.2.2. The seven iterations used in Phase-2 tracking are listed in Table 6.3. As already extensively discussed in Chapter 3, the main difference between iterations is the configuration of the seed generation and the target tracks. As can be seen, some of the special iterations developed for Run 2 are also included, such as the two final iterations using the information from muon chambers to enhance the performance on the muon track reconstruction. In all iterations the seeding procedure is performed in the IT for very similar reasons as the ones described already in the case of the Run 1 track reconstruction. Thanks to its high granularity, the IT has lower average channel occupancy compared to the OT, as can be seen in Figure 6.3. The same seeding algorithm used in the present CMS tracking is also used for Phase 2: triplets are found by combining pixel hits in different layers of the detector, while quadruplets are formed using the Cellular Automaton-based Hit Chain-Maker algorithm. A description and the performance of this algorithm were already given in Section 3.4.2.

6.1.3 Performance

The track reconstruction performance is quantified by the track finding efficiency, the fake rate, and the resolution of the estimated track parameters. The exact definitions of these quantities are given in [Section 3.2.3](#). The performance of the Phase-2 track reconstruction is evaluated using three different simulated samples. The first sample consists of single isolated muons, generated according to a flat distribution in η inside the Phase-2 tracker acceptance, $|\eta| < 4.0$, with fixed transverse momenta of 1, 10 and 100 GeV. The second sample consists of events with $t\bar{t}$ pairs, used to re-create a typical proton-proton collision at the [HL-LHC](#). The last sample is produced using highly energetic [QCD](#) jets events, in particular with $3 < p_T < 3.5$ TeV. Minimum-bias samples can also be superimposed to the events: two pileup scenarios are considered, where the number of pileup events is drawn from a Poisson distribution with mean equal to 140 and 200, respectively. They are denoted as 140PU and 200PU, respectively. The results are shown for the *High Purity* track selection, if not specified otherwise. They can be considered as a conservative lower limit and are expected to improve with careful tuning of the reconstruction algorithms for the high pileup scenario.

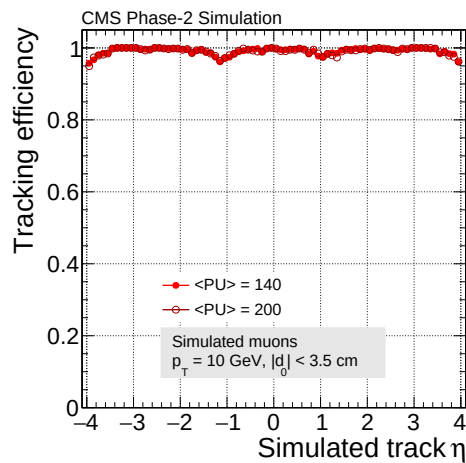


Figure 6.4: Tracking efficiency as a function of the pseudorapidity for single muons with p_T equal to 10 GeV, with 140 (full circles) and 200 (open circles) pileup events. The efficiency is shown for tracks produced within a radius of 3.5 cm from the centre of the luminous region [92].

The robustness of the tracking algorithm is underlined in [Figure 6.4](#), where the tracking efficiency for single isolated muons with a transverse momentum of 10 GeV is shown for the two pileup scenarios. The efficiency is flat around 100% in the entire range of acceptance of the tracker, independently of the amount of pileup superimposed. The same conclusion can be applied to the efficiency results obtained from the $t\bar{t}$ sample. [Figure 6.5](#) shows the efficiency as a function of the pseudorapidity and the p_T of the simulated particle, and the fake rate as a function of the pseudorapidity and the p_T of the reconstructed track. In the efficiency as a function of η , a selection on the p_T of the simulated particle is applied. The efficiency is around 90%, falling off at $|\eta| > 3.8$ and dropping for large p_T values. This behavior for high p_T values is due to the fact that charged particles in that p_T region are mostly produced inside the core of collimated jets, where the reconstruction is more difficult due to the extremely high density of particles. The fake rate is found to be around 2-4% and almost doubles for higher pileup, but is still tolerable.

As mentioned in [Section 5.3.5](#), some studies have also been carried out in order to understand the effect of different tracker layouts on the tracking performance. For example, in [Figure 6.6](#) the efficiency and fake rate are plotted for a flat layout and a tilted layout using the $t\bar{t}$ sample with 140PU. The tilted OT is introduced to overcome the stub finding efficiency problem at the level of the L1 trigger described in [Section 5.3.5](#) and does not influence the performance of the offline tracking. It can be noted that the average performance of these two specific plots is different from the ones presented in the rest of this section, because they were produced with an older version of

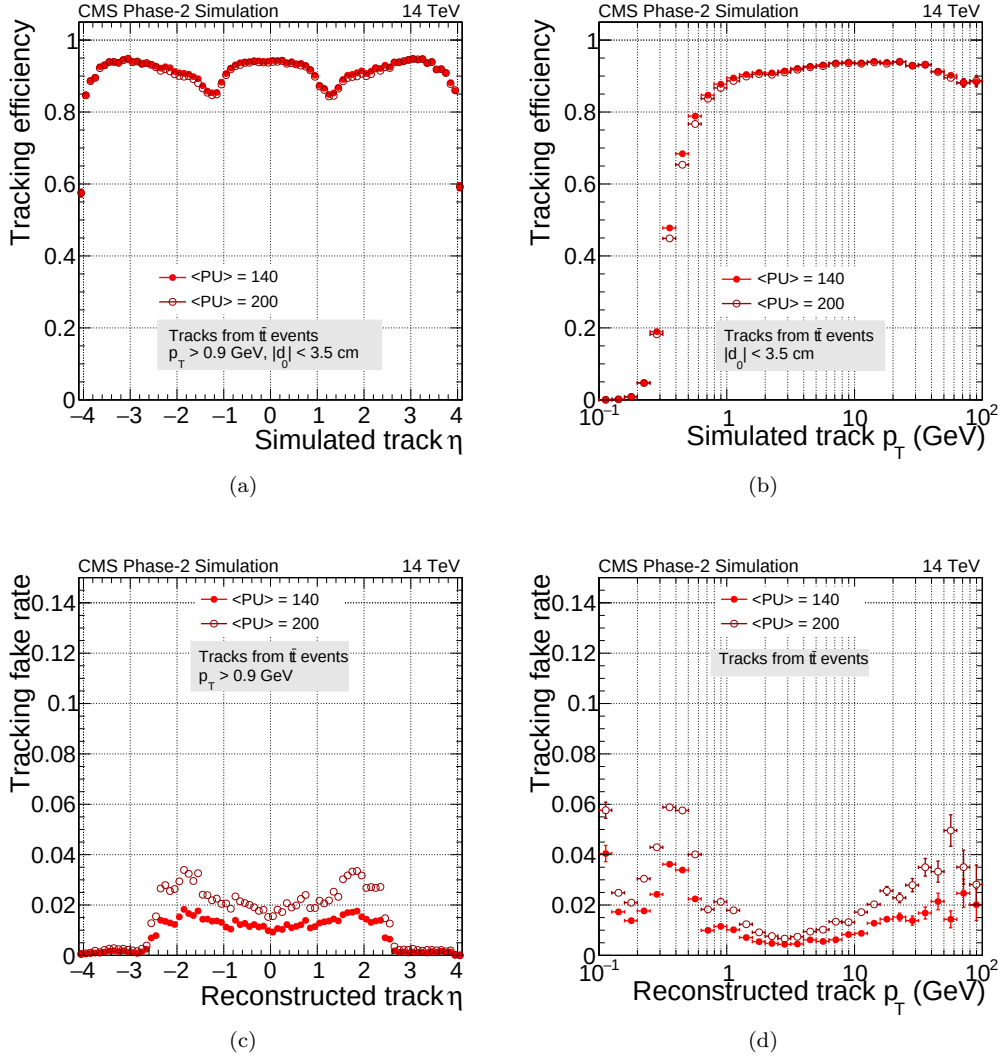


Figure 6.5: Tracking efficiency (a,b) and fake rate (c,d) for simulated $t\bar{t}$ events that include superimposed pileup collisions. They are plotted as a function of the pseudorapidity (a,c) and the p_T (b,d) of the simulated particle with 140 (full circles) and 200 pileup events (open circles). The tracks are required to have $p_T > 0.9$ GeV. The efficiency is shown for tracks produced within a radius of 3.5 cm from the centre of the luminous region [92].

the Phase-2 tracking, and therefore the performance is worse than the most recent ones presented in this thesis. This also shows the continuing work and improvement that have been carried out in the last few years with the introduction of some optimization that were not in place before. Another example of studies of the tracker layout using the full simulation is given in Figure 6.7 where the forward part of the IT sub-detector is simulated with either 11 disks or 12. In this case the 200PU scenario is used. The efficiency and fake rate results are very similar in the entire η range underlining the robustness of the forward tracking.

Figure 6.8(a) shows the contribution of the individual iterations used in Phase-2 tracking to the final tracking efficiency. The $t\bar{t}$ sample is used with superimposed pileup of 200. As already described in Section 3.2.3, the tracking efficiency is at least 30% down to about 300 MeV, thanks to the iterations targeting low transverse momentum particles.

An important requirement of the CMS Phase-2 tracker is to have good two-track separation. This requirement is achieved by increasing the granularity of the IT modules with respect to the Pixel Tracker in the Phase-1 tracker. In Figure 6.8(b) the performance for solving the two-

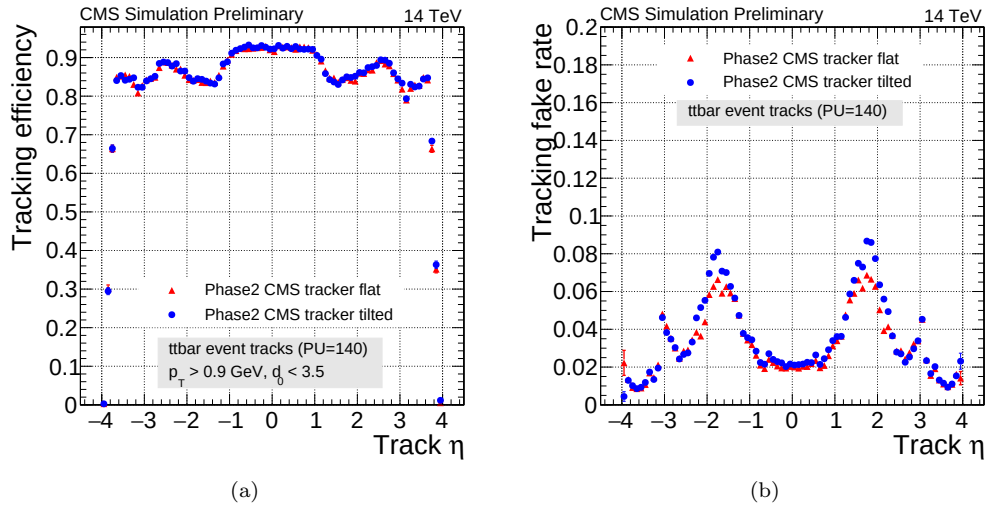


Figure 6.6: Tracking efficiency (a) and fake rate (b) for simulated $t\bar{t}$ events with 140 superimposed pileup collisions using a flat layout (red triangles) or a tilted layout (blue circles).

track separation problem for both trackers is shown, the tracking efficiency is computed as a function of ΔR , which is the distance between a simulated track and its nearest neighbour, defined in Equation 2.5. In this case the QCD sample with high p_T is chosen, because the problem is more visible in the case of tracks produced very close to each other, which is typically found in the core of high-energy jets. It is important to note that the Phase-1 reconstruction already includes some optimizations to perform better in this scenario, such as a special algorithm to split clusters and a special iteration to perform robust tracking in jet cores, as explained in Section 3.3.1. Although these optimizations have not yet been integrated into the Phase-2 reconstruction, a significant improvement can already be seen for small values of ΔR thanks to the higher granularity of the new detector. Further improvements are expected for large values of ΔR after applying a similar tuning to the one for the Run-2 tracking.

In Figures 6.9 and 6.10 the resolution of the five estimated track parameters is shown as a function of the pseudorapidity. The track parameters used in CMS are described in detail in Section 3.2.3 together with their dependency on the pseudorapidity. Figure 6.9 shows the comparison of the parameter resolution for single muons with $p_T = 10$ GeV between the Phase-1 tracker and the future Phase-2 tracker. For all parameters the resolution of the Phase-2 tracker is the same or better than the Phase-1 detector thanks to the improved hit resolution and the reduction in material budget. A significant improvement can be noticed in particular for the transverse impact parameter and the p_T resolution, which are directly connected with these quantities. In the case of the transverse impact parameter, it is found to be below $10 \mu\text{m}$ in the central region and about $20 \mu\text{m}$ at the edge of the acceptance.

In Figure 6.10 the resolution of the track parameters of the future Phase-2 tracker is shown for single isolated muons generated with p_T values of 1, 10 and 100 GeV. A very similar conclusion to the one already reported in Section 3.2.3 can be drawn.

Comparing these results to the ones presented in Section 3.4.2 for the Phase-1 upgrade, it can be noted that the excellent track finding performance is preserved despite the more difficult HL-LHC conditions. The only quantity that suffers from increasing pileup is the fake rate, but it is still tolerable. It is therefore clear that further tuning of the algorithm is needed to exploit the entire capability of the new tracker in case of very high levels of pileup.

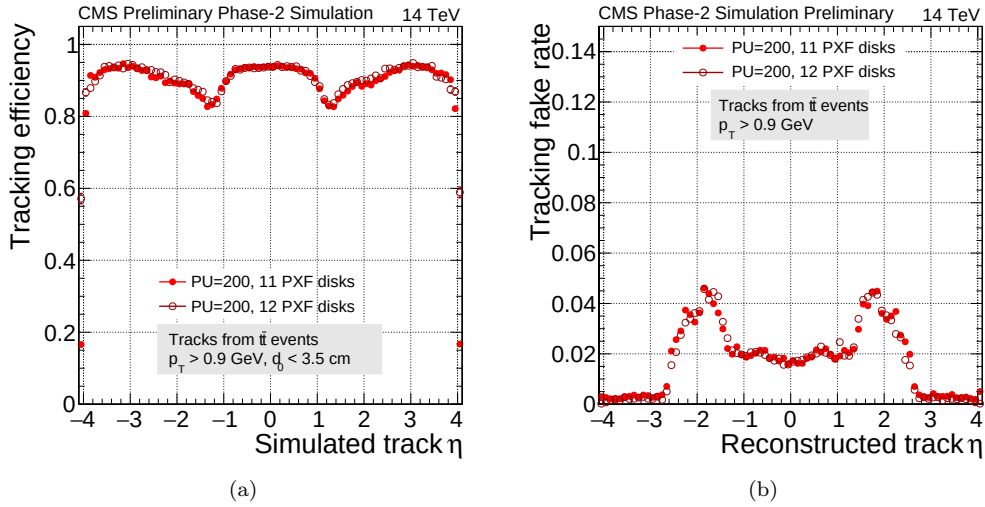


Figure 6.7: Tracking efficiency (a) and fake rate (b) for simulated $t\bar{t}$ events with 140 superimposed pileup collisions using 11 disks (full circles) or 12 disks (open circles) in the forward part of the IT sub-detector.

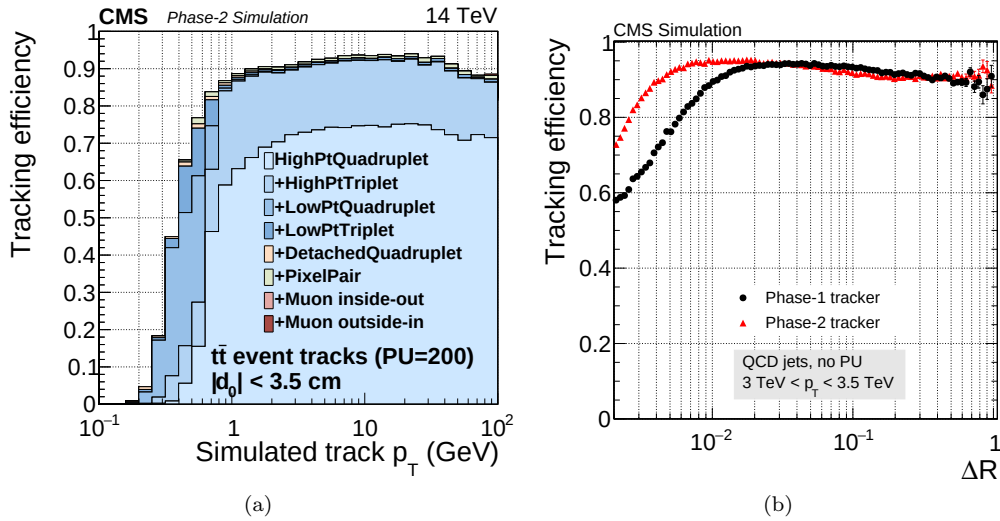


Figure 6.8: (a) Cumulative contributions to the overall tracking performance for the $t\bar{t}$ sample with an average pileup of 200 as a function of the simulated track p_T . The label of each of the seven iterations used in Phase-2 track reconstruction identifies the seeding algorithm used (see Table 6.3). (b) Tracking efficiency in the cores of jets with $3 < p_T < 3.5$ TeV as a function of the distance between a simulated track and its nearest neighbour, ΔR , for the Phase-1 (black) and the Phase-2 (red) tracker, without superimposed pileup [92].

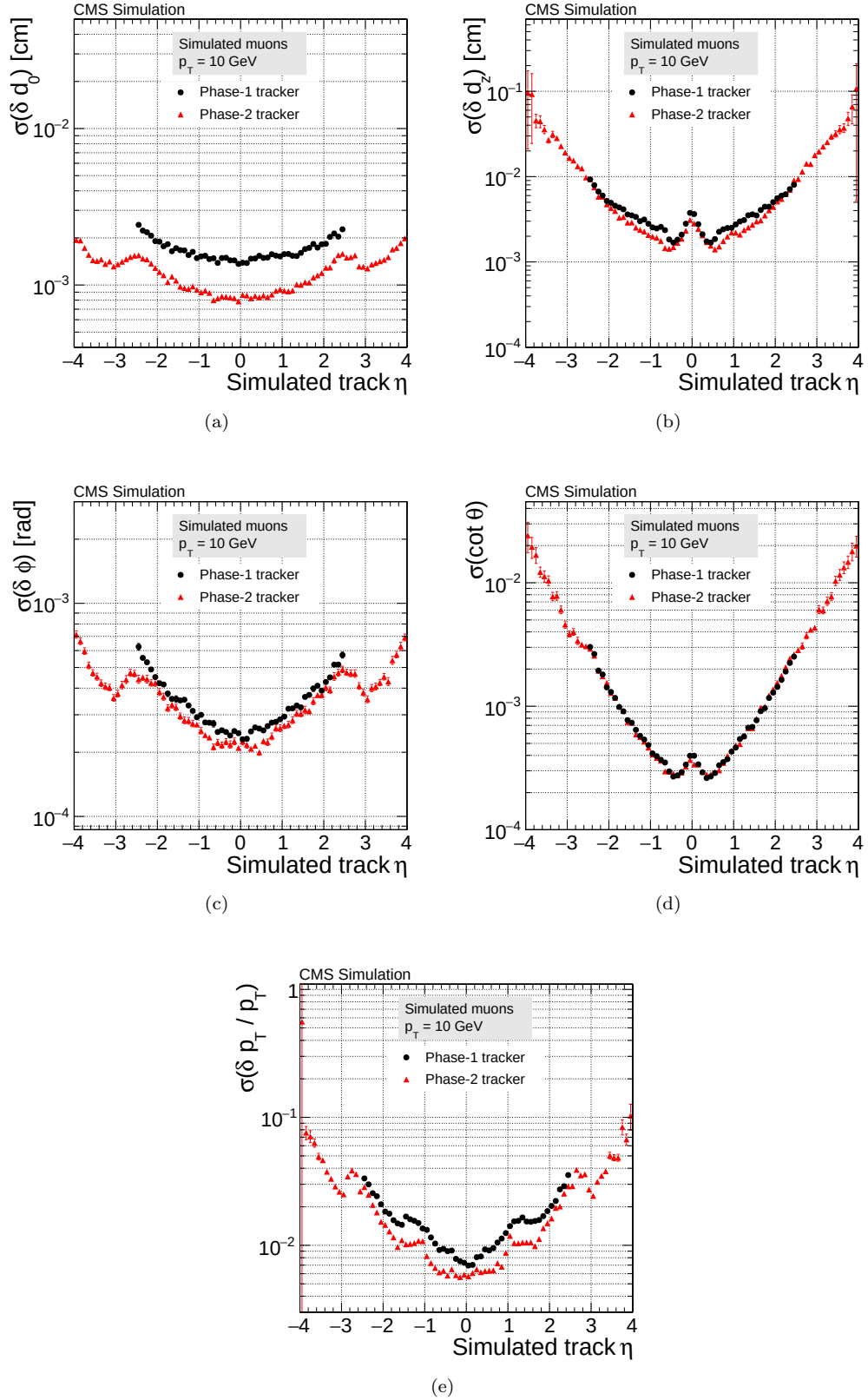


Figure 6.9: Resolution as a function of pseudorapidity, in the five track parameters for single, isolated muons with generated p_T equal to 10 GeV for the Phase-1 tracker (black) and the Phase-2 tracker (red). From top to bottom and left to right: transverse and longitudinal impact parameters, ϕ , $\cot \theta$ and transverse momentum [92].

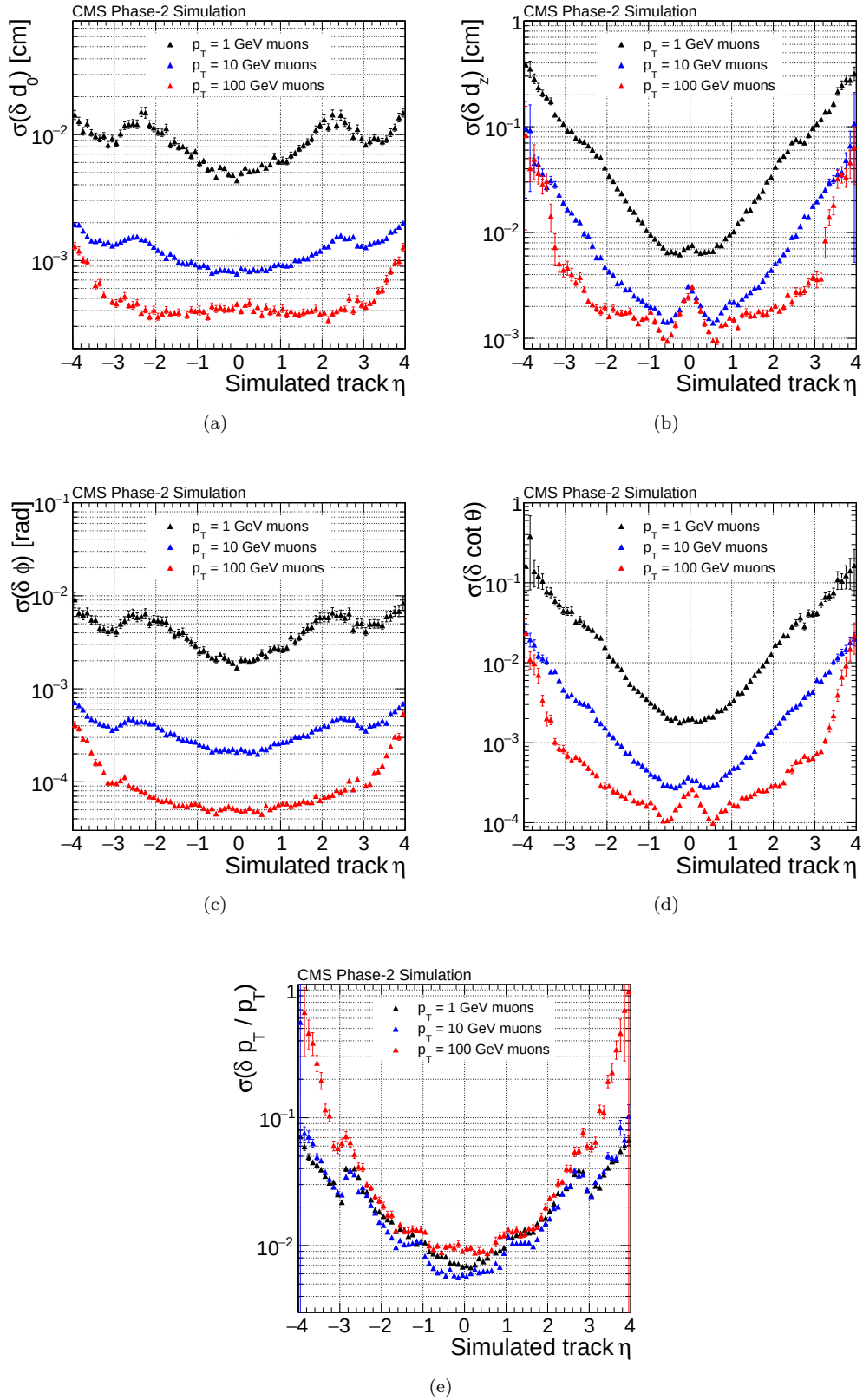


Figure 6.10: Resolution as a function of pseudorapidity, in the five track parameters for single, isolated muons with generated p_T equal to 1 (black), 10 (blue) and 100 (red) GeV. From top to bottom and left to right: transverse and longitudinal impact parameters, ϕ , $\cot \theta$ and transverse momentum [92].

Primary vertex reconstruction and resolution

The Phase-2 vertex reconstruction is based on the algorithm used for the current CMS tracker [83]. In a first step, the Deterministic Annealing algorithm [135] is used to find vertices. In a second step, the Adaptive Vertex Fit [136] is run to compute the best estimate for the vertex position and the parameters of the associated tracks. In the last step, the reconstructed vertices are sorted to identify the hard interaction point, the so-called *primary vertex*, among the pileup vertices. The sorting is based on $\sum p_T^2$, but using a more sophisticated technique: the algorithm replaces the individual tracks contributing to the $\sum p_T^2$ by jets obtained by clustering tracks originating from the same vertex using the anti- K_t jet clustering algorithm [73]. The final $\sum p_T^2$ is computed using these jets, the remaining isolated single tracks and the missing transverse momentum, re-weighted accordingly.

For $t\bar{t}$ events with an average pileup of 140 and 200, the probability of reconstructing the signal vertex and to tag it correctly is about 94% and 89%, respectively. In Figure 6.11(a) the number of reconstructed vertices as a function of the number of simulated vertices is shown. The dependence is almost linear, underlining the high quality of the vertex reconstruction. In Figure 6.11(b) the merging rate¹ is shown as a function of their distance in the z coordinate. The merging rate is almost independent of the pileup and drops to zero at about 0.6 mm. Another important aspect is the resolution of the vertex position in the x , y and z coordinates. In Figure 6.12 this is shown as a function of the number of tracks associated with the vertex. It can be seen that also in this case the result is almost independent of the amount of pileup in the event. Moreover, the longitudinal resolution is only 50% worse than the transverse one, as expected given the pixel dimensions of Inner Tracker modules.

As in the case of the tracking performance, all results should be considered as a lower limit of the expected performance.

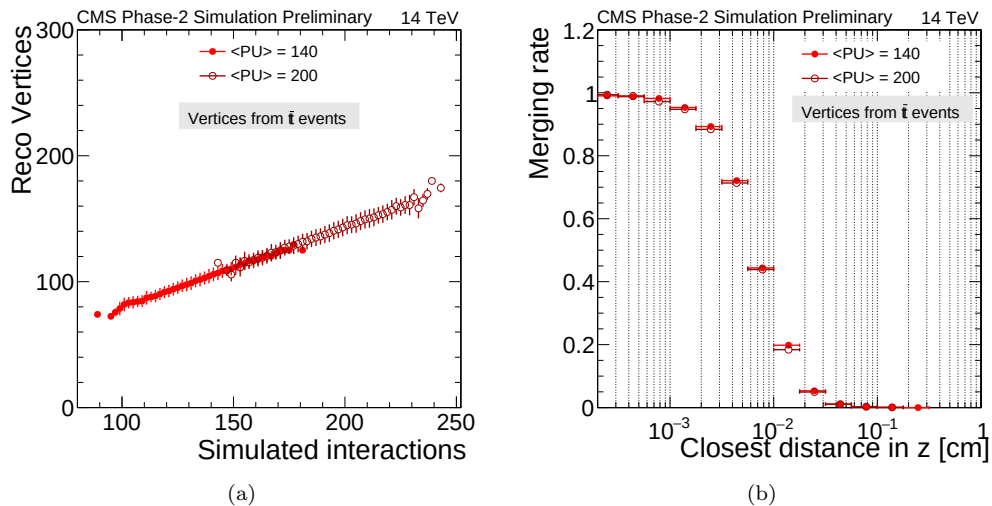


Figure 6.11: Number of reconstructed vertices as a function of simulated vertices (a) and merging rate of vertices as a function of their distance in the z coordinate (b) for $t\bar{t}$ events with 140 superimposed pileup events (full circles) and 200 pileup events (open circles) [92].

6.1.4 Possible developments

Some possible developments are:

- **Iterations targeting physics objects:** Further studies are required to gain a better understanding of the performance of the new tracker for different physics objects. For example, the two iterations dedicated exclusively to muons have already been included, but given the

¹The merging rate is the probability of two adjacent vertices to merge into a single one.

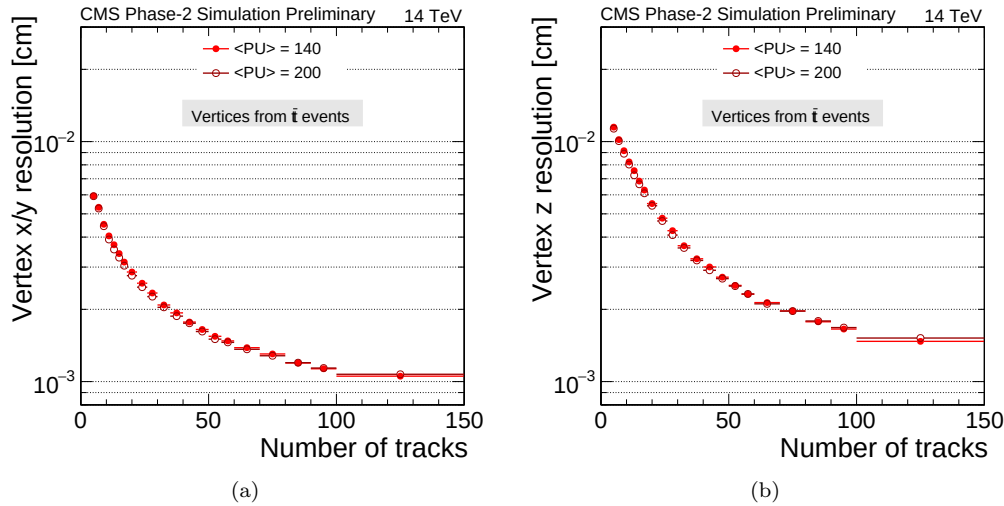


Figure 6.12: Vertex position resolution in x and y (a), and z (b) as a function of the number of tracks associated to the vertex, for $t\bar{t}$ events with 140 pileup events (full circles) and 200 pileup events (open circles) [92].

high two-tracks separation power shown by the future tracker, it is yet not clear if also the iteration specialized for high- p_T jets is needed. Therefore, it is important to first understand which physics objects need particular tuning with the new tracker and in a second step develop highly specialized iterations.

- **Multiple algorithms:** A potentially interesting development is to introduce different algorithms in different iterations. Different algorithms could be included either to reconstruct better specific physics objects, as already mentioned in the previous bullet, or to be specialized for certain tasks, or even to exploit different parts of the tracker layout.
- **Specific optimization for the very forward region:** The Phase-2 tracker is the first detector in CMS that explores the region $3.0 < \eta < 4.0$ with such a high precision. The main idea is to better understand the features of the physics processes in this very forward region and to develop a track reconstruction highly specialized for that.
- **Tracking at HLT using GPUs:** The software for the CMS High-Level Trigger (HLT) tracking has been moved in 2017 from a one-event-per-thread strategy to a many-threads-per-event approach. This massive parallelism within each event scales with the pileup by offloading the combinatorics to as many threads as are available. This task has been implemented in CMS by designing a HLT based on GPUs. The same valid approach will be used also for the future Phase-2 HLT.
- **New hardware technology:** As explained in Section 2.2.7, the track reconstruction is traditionally the most time consuming part of the event reconstruction in CMS, and this will become an even more pressing problem in the high pileup scenario for the HL-LHC. It is thus imperative to continue to monitor time and memory consumption and to introduce optimization and algorithms which can be easily parallelizable. In this context, some studies have already been performed to exploit the vectorization and parallelization possibilities in the Kalman Filter (KF) algorithm for the HL-LHC [137].
- **Tracking in the Outer Tracker:** As described in Section 5.3.3, the main goals of the new OT are: to resist the higher radiation level, to keep the channel occupancy low, and to contribute to the L1 trigger. The last point is achieved by designing modules that are able to reject low- p_T particles by taking advantage of the 3.8 T magnetic field. Each module is a stack of two closely spaced sensors and is able to correlate hits produced in the two sensors

in real time. A possibility to further exploit the new OT design is to profit from the L1 track trigger information at the offline tracking level, for instance by using L1 tracks as seeds or incorporating L1 algorithms. A second possibility is to build short track segments in the OT and develop new algorithms that exploit this information to search for a seed in the IT and to reconstruct displaced tracks without using any information coming from the IT. The latter possibility is evaluated in the rest of this chapter.

6.2 Vector Hits

A possibility to exploit the new tracker layout in the offline tracking process is to combine the two compatible clusters from the OT stacked modules in an offline stub, called *vector hit*. In Figure 6.13 a schematic representation of the vector hit concept is shown. The vector hit is a short track segment and carries not only the hit position, but also the direction information. The impact of using vector hits instead of single hits to solve the pattern recognition problem in the tracker was investigated in Ref. [138] for the first time. The article concludes that vector hits can be useful in track finding and gives indications on the importance of the stack separation parameter in this context. However, this study was pursued using a simplified fast simulation that contains many simplifications and thus does not allow to give any definitive conclusion. Given the impact of this preliminary study in the future tracking, this thesis investigates the possibility of exploiting vector hits in a more realistic simulation and using them directly in the Combinatorial Track Finder (CTF) developed for Phase 2.

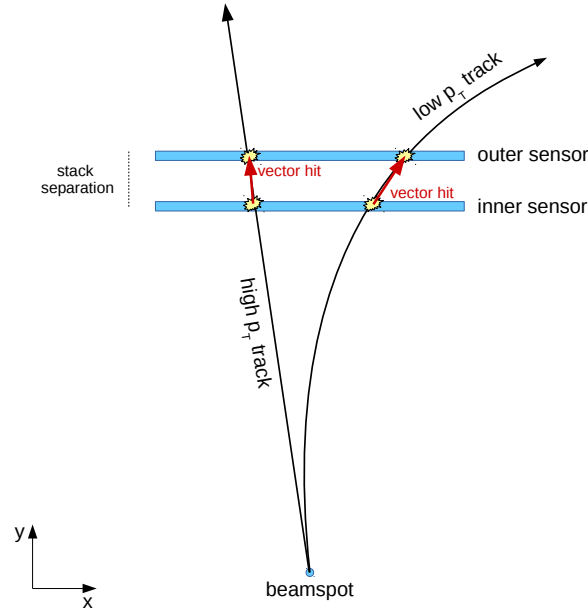


Figure 6.13: Schematic representation of the vector hit for high- p_T and low- p_T track in the bending plane.

6.2.1 Local reconstruction

Before introducing a mathematical representation of the vector hit concept and its characteristics, the nomenclature referring to the different parts of the OT geometry is reported:

- **Stack module:** the stack module is the equivalent to the p_T module described in Section 5.3.3. They can be either a PS or 2S module.

- **Lower and upper sub-modules:** These sub-modules are defined in the system of reference of the stack module. In the **PS** module, the lower sub-module is composed of macro pixels, while the upper sub-module uses strips. The centre of each sub-module is used as the (0,0) point of the $x - y$ axes in the local system of reference.
- **Inner and outer sub-modules:** These sub-modules are defined in the global system of reference of **CMS**. The inner is the sub-module closer to the proton-proton interaction point, while the outer is the further one.

The difference between lower and upper sub-modules in the global system of reference is schematically represented by **Figure 6.14**. The stack module and the lower and upper sub-modules are represented in **CMSSW** with C++ classes that describe the geometry of the sub-modules and have a unique detector Id to differentiate each other and with other parts of the Phase-2 tracker.

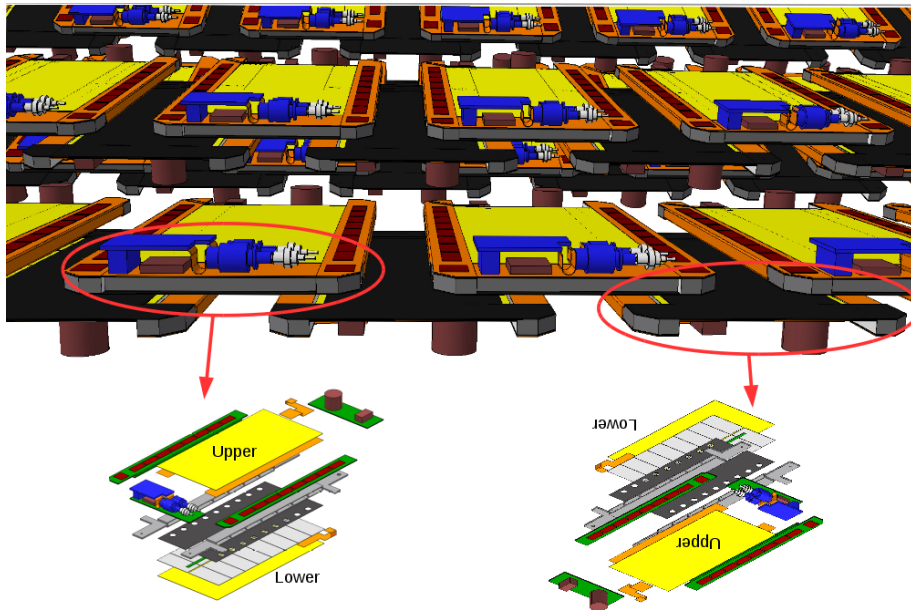


Figure 6.14: Schematic representation of lower and upper sub-modules in a barrel layer in the global system of reference.

A vector hit can be mathematically represented in the local system of reference by a 4-components vector:

$$\left(x, y, \frac{dx}{dz}, \frac{dy}{dz} \right) \quad (6.1)$$

where x and y are the most and least precise coordinate, respectively, while dx/dz and dy/dz are the respective direction components on the two axes. The covariance matrix associated to the vector hit is symmetric.

The algorithm used to build vector hits starts from all the clusters found in the event that belong to the lower and the upper sub-module of a particular stack module. The x - and y -coordinate of the vector hit and the corresponding errors are the ones associated to the lower cluster. The dx/dz and dy/dz are computed by a linear fit using the positions of the lower and upper clusters in the $x - z$ and $y - z$ plane. The linear fit gives also a first estimate of the variances of the direction parameters and their covariance with respect to the local position.

After the implementation of the vector hit concept in **CMSSW**, some studies were performed at the level of the local reconstruction. In **Figures 6.15** and **6.16** the cluster distribution for a $t\bar{t}$ sample with superimposed pileup of about 200 is shown in each layer in the **OT** barrel and in some rings in the **OT** endcap, respectively. From these distributions it is clear that a pure combination of all clusters in the lower and upper sub-modules causes big difficulties in terms of memory and time consumption.

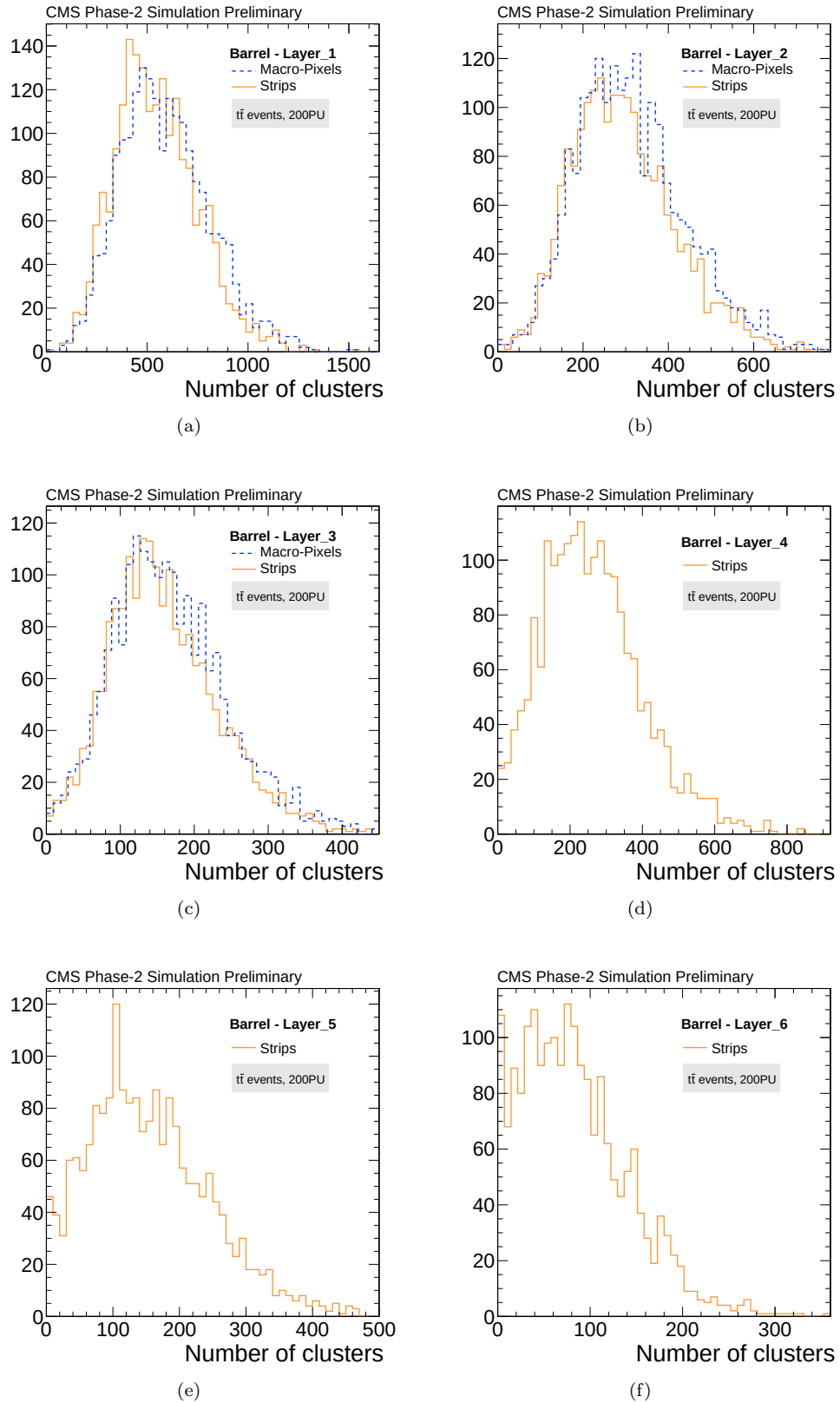


Figure 6.15: Distributions of the number of cluster for all the six layers in the OT barrel.

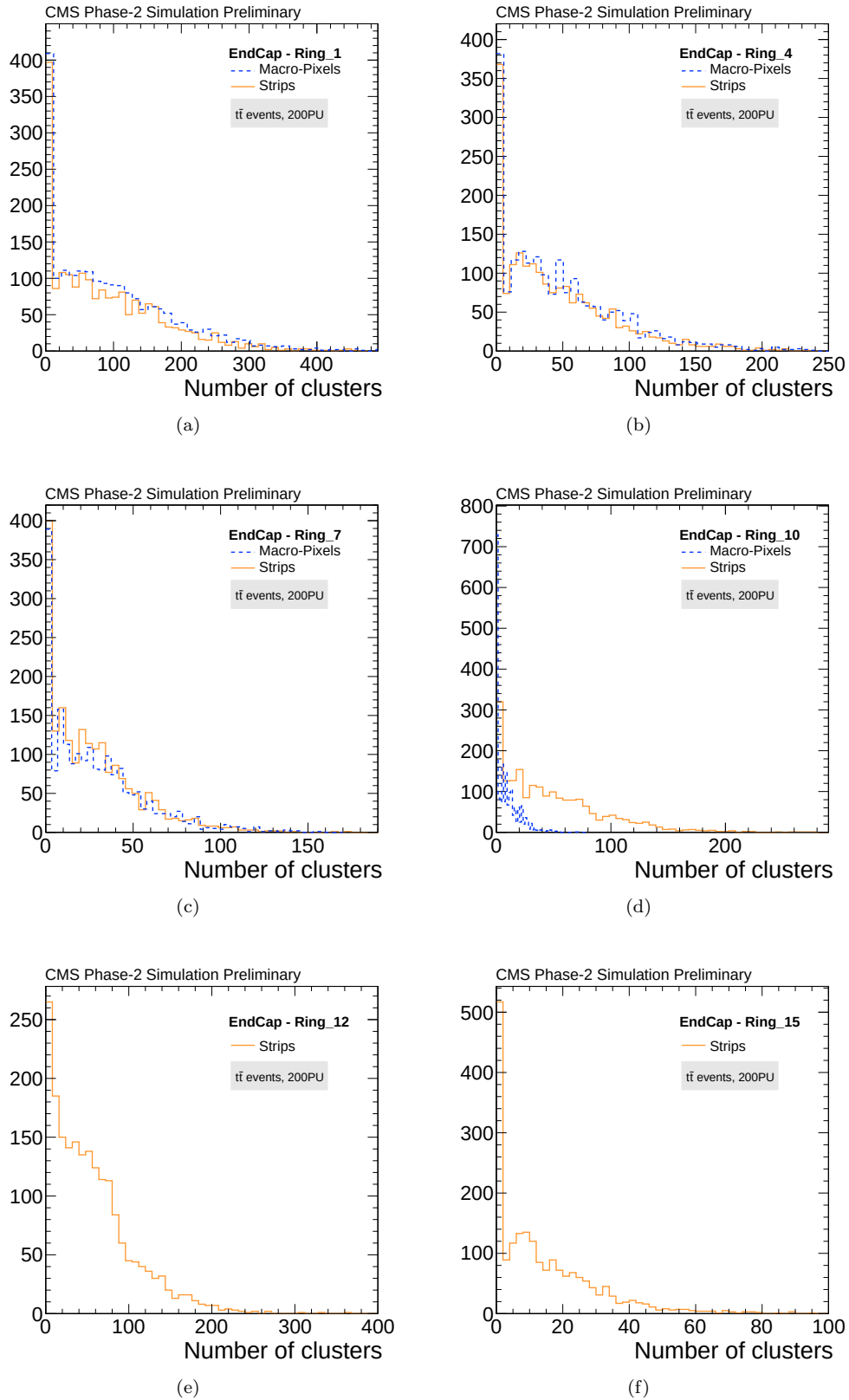


Figure 6.16: Distribution of the number of cluster for some rings in the OT endcap.

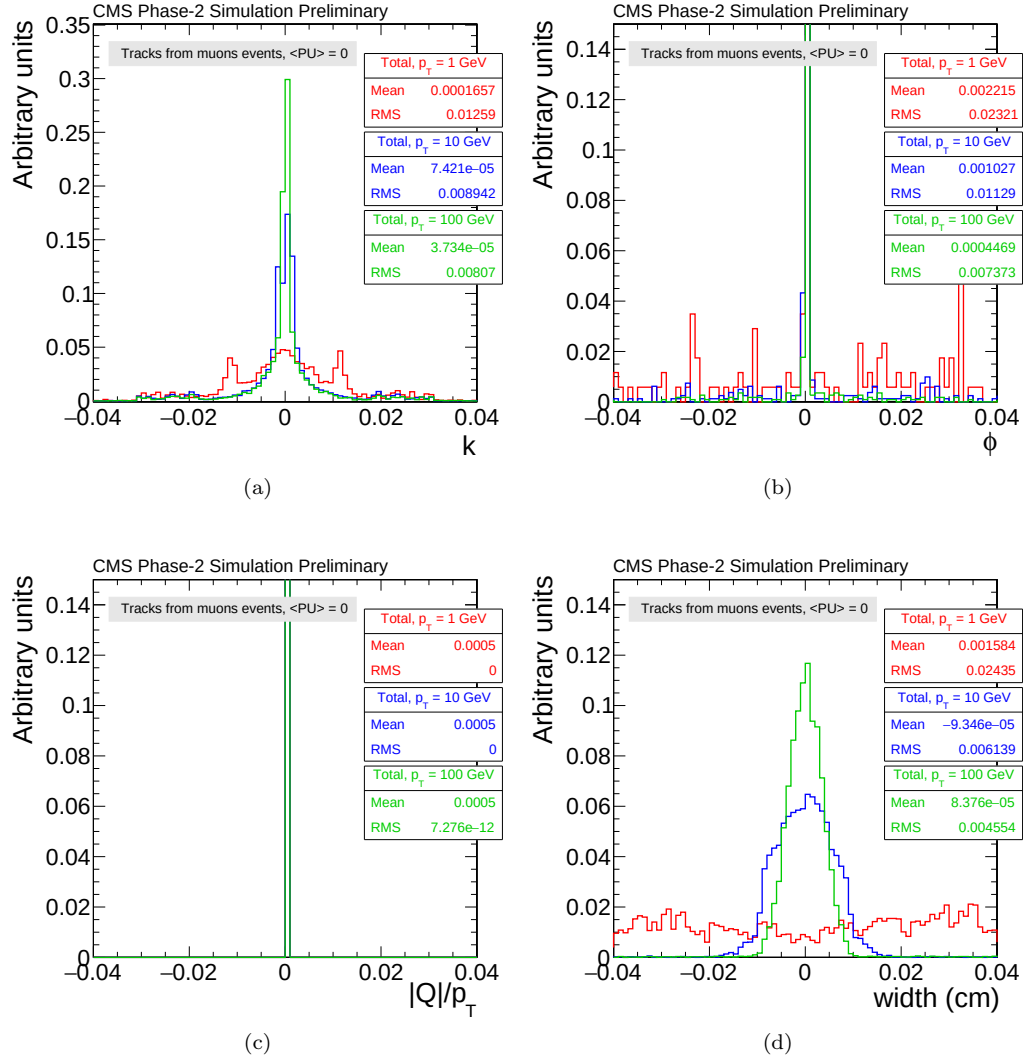


Figure 6.17: Distribution of some vector hit variables produced by single isolated muon events generated with p_T equal to 1 GeV (red), 10 GeV (blue), 100 GeV (green). From top to bottom and left to right: curvature, ϕ , $|Q|/p_T$, and width.

The combinatorial problem arising from combining all clusters found in two sub-modules can be overcome by introducing a constraint in the vector hit reconstruction. This constraint can be tuned to reject the tracks coming from pileup, and therefore can exploit the p_T rejection power of the vector hits. In Figure 6.17 some variables which can achieve this goal are computed for each vector hit for a single isolated muon sample generated with different p_T . The variables are: curvature (k), azimuthal angle (ϕ), ratio between the absolute charge of the particle and the p_T ($|Q|/p_T$), and width.

Defining $(x_{\text{inner}}, y_{\text{inner}})$ and $(x_{\text{outer}}, y_{\text{outer}})$ as the x and y global coordinates for the inner and the outer clusters, respectively, the curvature, the azimuthal angle and the transverse momentum for each vector hit can be extracted using the radius of circle, ρ :

$$\rho = \frac{\sqrt{(x_{\text{inner}}^2 + y_{\text{inner}}^2)(x_{\text{outer}}^2 + y_{\text{outer}}^2)} h_3}{2h_1} \quad (6.2)$$

with

$$h_1 = x_{\text{inner}} y_{\text{outer}} - x_{\text{outer}} y_{\text{inner}} \quad (6.3)$$

$$h_3 = x_{\text{inner}}^2 - 2x_{\text{inner}}x_{\text{outer}} + x_{\text{outer}}^2 + y_{\text{inner}}^2 - 2y_{\text{inner}}y_{\text{outer}} + y_{\text{outer}}^2 \quad (6.4)$$

The curvature is then defined as:

$$k = \frac{1}{\rho} \quad (6.5)$$

while the azimuthal angle is computed as follows:

$$\phi = \arctan 2(x_{\text{inner}} - x_m, y_{\text{inner}} - y_m); \quad (6.6)$$

where x_m and y_m are the center of circle:

$$x_m = \frac{x_{\text{inner}}^2 y_{\text{outer}} - x_{\text{outer}}^2 y_{\text{inner}} + y_{\text{inner}}^2 y_{\text{outer}} - y_{\text{inner}} y_{\text{outer}}^2}{2h_1} \quad (6.7)$$

$$y_m = \frac{-x_{\text{inner}}^2 x_{\text{outer}} + x_{\text{inner}} x_{\text{outer}}^2 + x_{\text{inner}} y_{\text{outer}}^2 - x_{\text{outer}} y_{\text{inner}}^2}{2h_1} \quad (6.8)$$

It is important to note that if h_1 defined in Equation 6.3 is zero, the track is considered a straight track and as a consequence, no curvature or azimuthal angle can be extracted. The transverse momentum is related to the radius of the curvature of the vector hit as follows:

$$p_T = 0.003B\rho \quad (6.9)$$

where B is the intensity of the magnetic field in the centre of the CMS detector. The last variable considered, the width, can be defined as follows:

$$w = |x_{\text{lower}} - x_{\text{upper}}| \quad (6.10)$$

where x_{lower} and x_{upper} are the local precise coordinates of the lower and the upper cluster, respectively. Among the four variables considered, the Figure 6.17 shows that the vector hit width is the one which can discriminate best between different p_T muon samples.

In order to distinguish the vector hits produced by a track and the ones that are purely background, a vector hit is defined as *true* if both upper and lower clusters are associated to the same simulated track. Moreover, a vector hit passing the width requirement is labeled as *accepted*, otherwise it is labeled as *rejected*. The number of vector hits in different OT barrel layers and OT endcap disks for a typical event at the HL-LHC can be seen in Figure 6.18. If no selection is applied, all vector hits created in the event are accepted. The number of vector hits produced in this case is some orders of magnitude higher than the number of true vector hits. The effect is even more clear if only simulated particles with p_T about 1 GeV are considered.

Parallax correction

Before implementing a threshold on the vector hit width, the cluster positions are corrected taking into account the *parallax correction*. In this context, the parallax correction p is defined as the correction which must be applied to the cluster position in the local reference frame to take into account of the incident angle of a straight track on the detector surface. In Figure 6.19 an example of this kind of problem in the bending plane is shown.

If no correction is applied, the width of the vector hit produced by a high- p_T track will be different from 0, not because of a real bending of the particle but purely due to geometrical effects. If the global coordinates of the lower cluster are denoted as

$$\vec{a}_{\text{lower}} = \begin{pmatrix} X_{\text{lower}} \\ Y_{\text{lower}} \\ Z_{\text{lower}} \end{pmatrix} \quad (6.11)$$

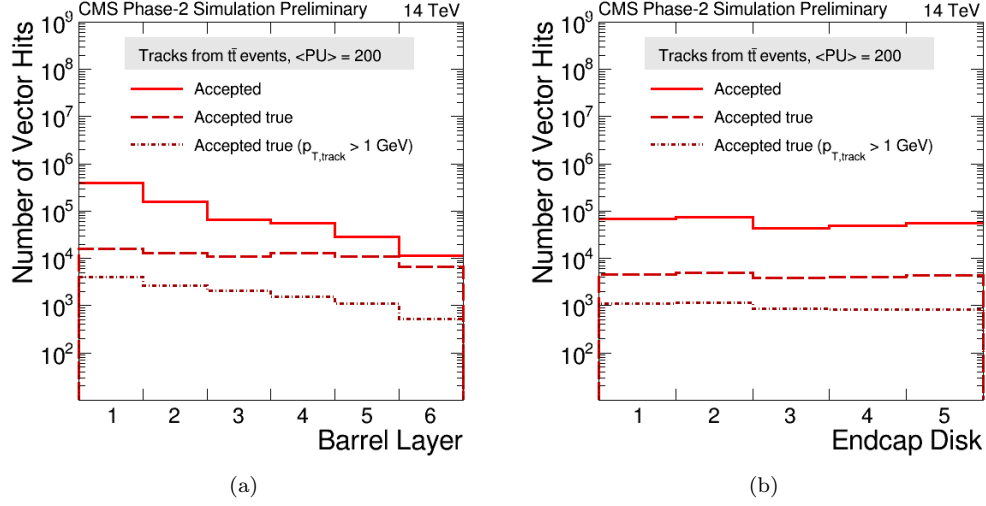


Figure 6.18: Distribution of vector hits in different Phase-2 OT layers in the barrel (a) and the endcap (b) for a single $t\bar{t}$ event simulated with superimposed pileup collisions of 200. A vector hit is defined as true when both upper and lower clusters are associated to the same simulated track.

a global vector $\vec{b}$ can be created using the difference between $\vec{a}_{\text{lower}}$ and the axis origin $(0,0,0)$. Moving $\vec{b}$ into the local system of reference of the lower detector and re-normalizing for the third component, a new vector can be obtained:

$$\vec{c} = \begin{pmatrix} c_x \\ c_y \\ 1 \end{pmatrix} \quad (6.12)$$

The parallax correction can be thus computed as

$$p = \Delta w \cdot c_x \quad (6.13)$$

where Δw is the stack separation of the module.

Depending on the position in the global reference system of the lower and upper sub-modules, four cases can arise:

- **Case 1a and 1b:** The upper sub-detector is the outer one. The correction is applied as follows:

- if the local position of the upper is positive, then the new position of the upper cluster is

$$\text{if } x_{\text{upper}} > 0 \quad \rightarrow \quad x'_{\text{upper}} = x_{\text{upper}} - p \quad (6.14)$$

- if the local position of the upper is negative, then the new position of the upper cluster is

$$\text{if } x_{\text{upper}} < 0 \quad \rightarrow \quad x'_{\text{upper}} = x_{\text{upper}} + p \quad (6.15)$$

- **Case 2a and 2b:** The lower sub-detector is the outer one. The correction is applied as follows:

- if the local position of the upper is positive, then the new position of the lower cluster is

$$\text{if } x_{\text{upper}} > 0 \quad \rightarrow \quad x'_{\text{lower}} = x_{\text{lower}} - p \quad (6.16)$$

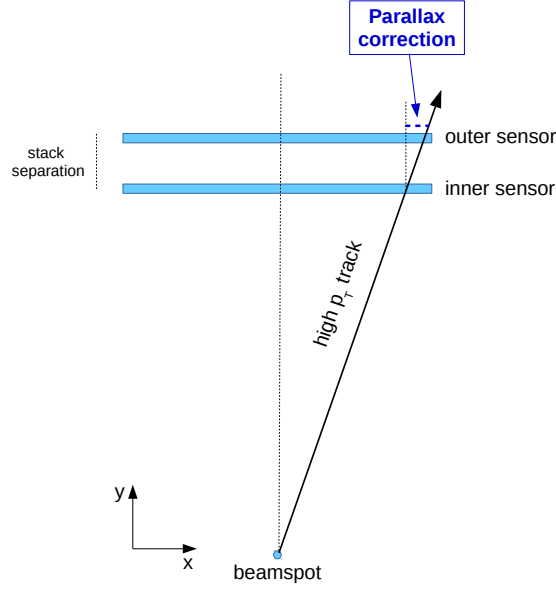


Figure 6.19: Schematic representation of the vector hit width depending on the position of the high- p_T track. The definition of the parallax correction is also represented.

- if the local position of the upper is negative, then the new position of the lower cluster is

$$\text{if } x_{\text{upper}} < 0 \quad \rightarrow \quad x'_{\text{lower}} = x_{\text{lower}} + p \quad (6.17)$$

The four different cases are shown schematically in [Figure 6.20](#). The corrected positions using the parallax correction are then used in [Equation 6.10](#).

Width requirement

In a first approach, the compatibility test between the lower and the upper cluster is carried out in the following way: the clusters belonging to the lower and the upper sub-modules are ordered with respect to their precise local coordinate; a fixed threshold is computed using the error on the precise coordinate; and finally the vector hit is built only if the lower and the upper cluster position are found within the width requirement. The fixed threshold is computed as follows

$$w < 10\sigma \quad \text{where} \quad \sigma = \sqrt{\sigma_{xx,\text{lower}}^2 + \sigma_{xx,\text{upper}}^2} \quad (6.18)$$

The parameter σ_{xx} is the standard deviation on the most precise local position of the cluster. The value of σ is typically around 0.004 cm. [Figure 6.17\(d\)](#) shows the width distribution for accepted vector hits derived from a single isolated muon sample generated with different p_T . This requirement is thus found to be good for high- p_T tracks, but not for the low- p_T track case.

As shown in [Figure 6.21\(a,b\)](#), the width scatter plots for muons with p_T equal to 1 GeV is more than the typical value of 10σ in the outermost layers of the **OT** barrel and for all the disks in the **OT** endcap. It is important that the vector hits belonging to these tracks are not rejected at this selection stage. In [Figure 6.21\(c,d\)](#) the width distribution for muons with p_T equal to 1 GeV is shown for different layers of the barrel and for different disks of the endcap in the **OT**. From these distributions, different thresholds are chosen for different **OT** layers, while in the case of the **OT** endcap, the same threshold is set for all disks. A summary of the chosen values is presented

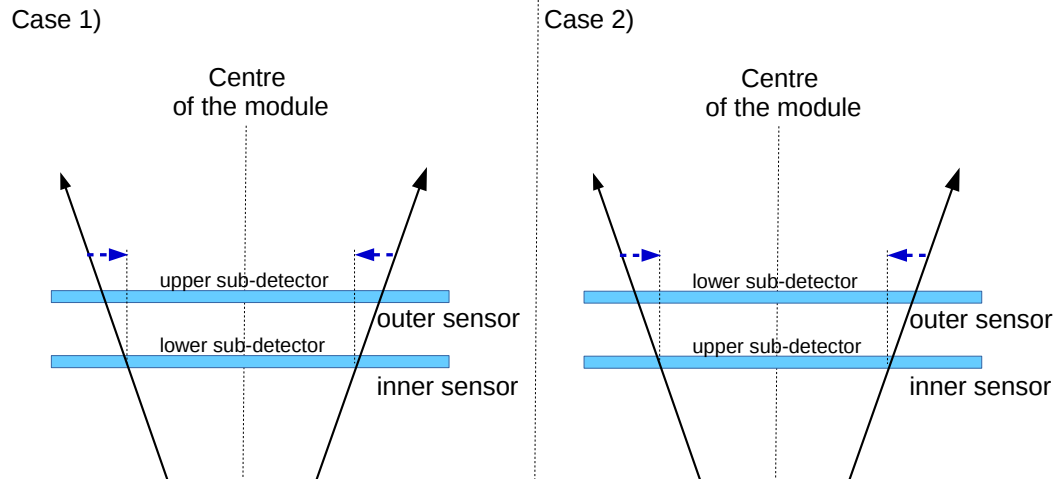


Figure 6.20: Schematic representation of the four different cases of parallax correction depending on the position of the lower and upper sub-modules in the global system of reference.

in Table 6.4. In Figure 6.22 three different distributions of accepted vector hits considering the two possible thresholds are shown for different OT barrel layers and endcap disks. A single $t\bar{t}$ event simulated with 200PU is used in this study.

Table 6.4: Width requirements chosen for the six layers in the OT barrel and the five disks in the OT endcap.

Barrel	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6
Threshold [cm]	0.05	0.06	0.08	0.09	0.1	0.12
Endcap	Disk 1	Disk 2	Disk 3	Disk 4	Disk 5	
Threshold [cm]	0.1	0.1	0.1	0.1	0.1	

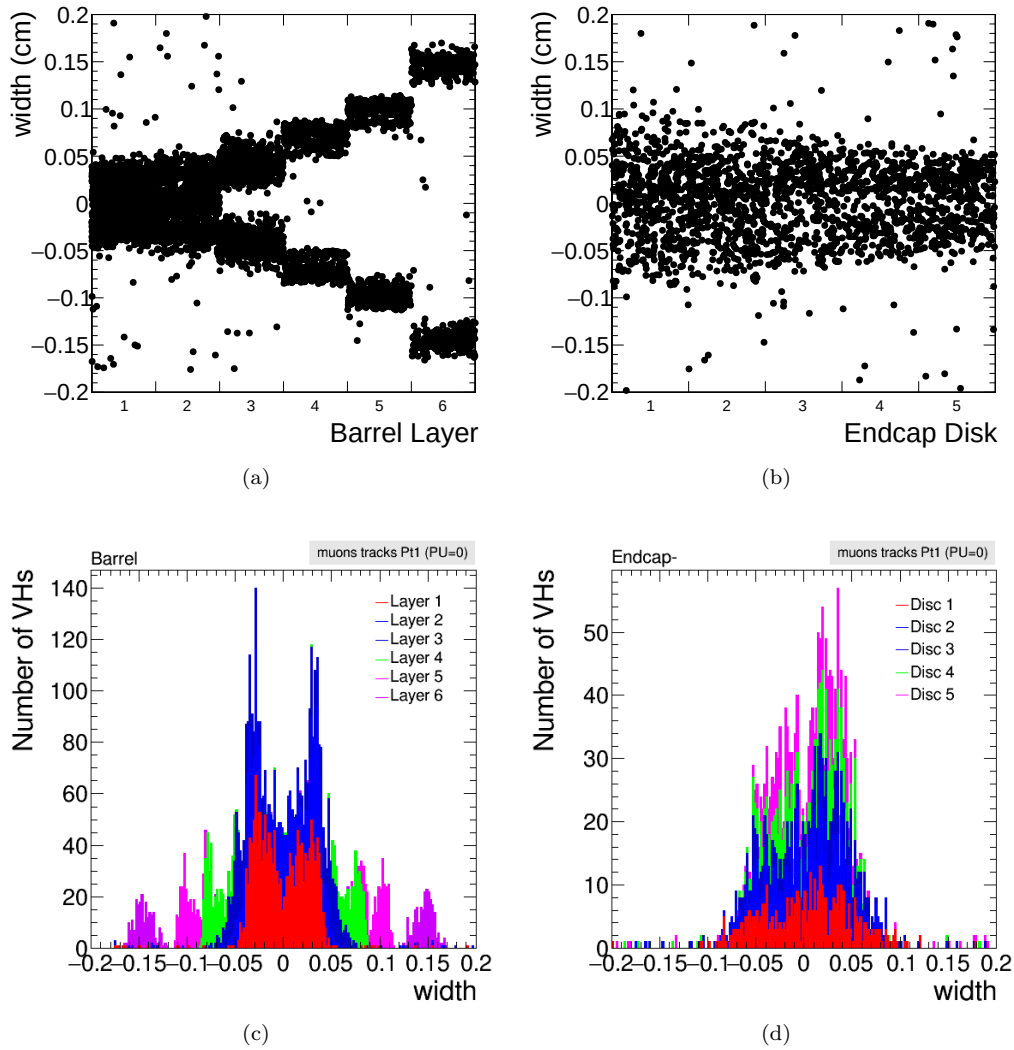


Figure 6.21: Width scatter plots (a,b) and histograms (c,d) produced only for muons with p_T equal to 1 GeV for the six layers in the **OT** barrel and the five disks in the **OT** endcap, respectively.

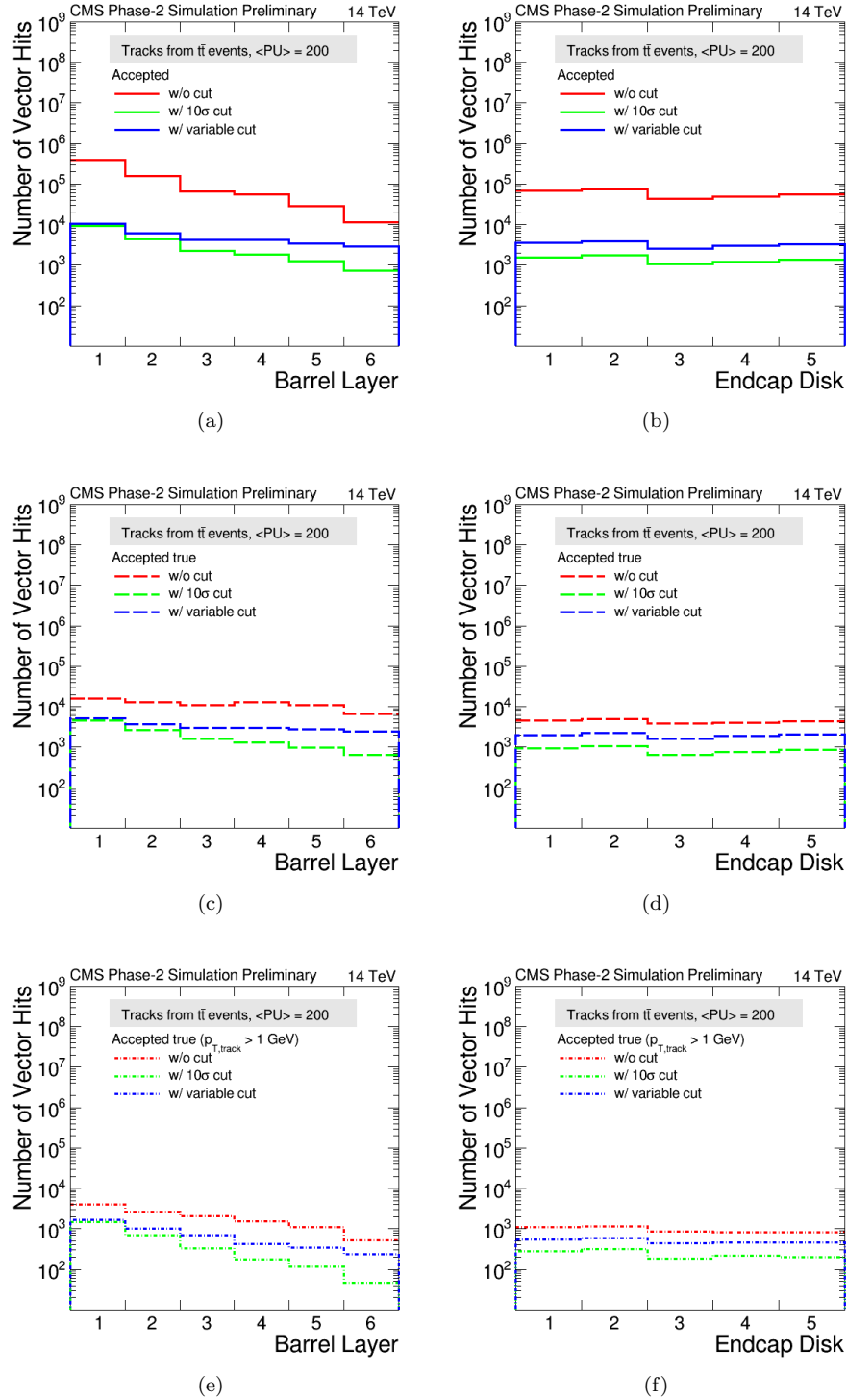


Figure 6.22: Number of accepted vector hits in different OT barrel layers (a) and endcap disks (b). Number of true accepted vector hits in different OT barrel layers (c) and endcap disks (d). Number of true accepted vector hits considering only the simulated tracks with p_T more than 1 GeV in different OT barrel layers (e) and endcap disks (f). All plots are created using a single $t\bar{t}$ event simulated with 200PU. The thresholds considered are: a constant threshold at 10σ for each layer (green) and a variable threshold for each layer (blue).

6.2.2 Global reconstruction

In order to use the vector hits in CMS track reconstruction, the CTF must be adapted. In all iterations the single hits in the OT are replaced by vector hits and the KF is extended to make use of the full information contained in the vector hits. The iterations listed in Table 6.3 are used. A first sanity check is to control the number of hits belonging to reconstructed tracks. The expected drop in the number of hits in the entire OT region of the Phase-2 tracker is shown in Figure 6.23(c) for the $t\bar{t}$ sample without any pileup scenario. The same sample is used to measure the performance in terms of efficiency and fake rate. As shown in the other plots of Figure 6.23, the track reconstruction efficiency is almost unchanged, as is the fake rate. The small loss in efficiency compared to the baseline is due to the fact that at the current stage vector hits belonging to different modules are not taken into account yet, as well as the possibility that just one hit is found in the stack module. Two possible solutions to this problem are either to include vector hits only in some iterations or to introduce a recovery mechanism which allows the tracks to be reconstructed also with single hits in the OT.

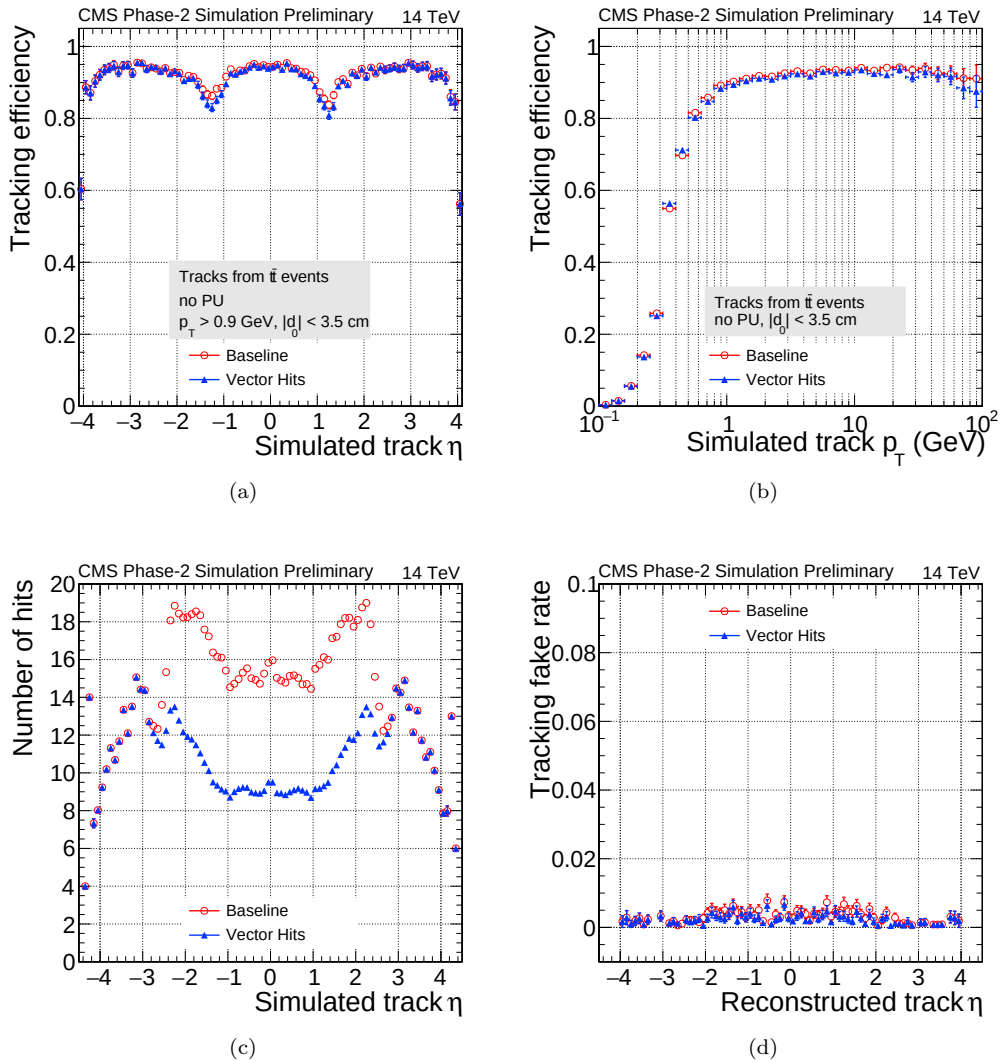


Figure 6.23: Tracking efficiency (a,b) and fake rate (d) for simulated $t\bar{t}$ events without superimposed pileup for the current Phase-2 tracking (open red circles) and the future tracking using vector hits (full blue triangles). The comparison in terms of number of hits as a function of η is also shown (c).

6.2.3 Seeding in the Outer Tracker

As already mentioned in Section 3.2, the track reconstruction developed for Run 1 and Run 2 provides two iterations specialized for finding tracks produced outside of the pixel volume. As the current Phase-2 tracking does not include these two iterations, its performance in reconstructing displaced tracks is limited. In Figure 6.24 a comparison in terms of tracking efficiency as a function of the simulated track production vertex radius is shown for Phase-1 reconstruction and the Phase-2 baseline. The $t\bar{t}$ sample with pileup at respective nominal conditions is used.

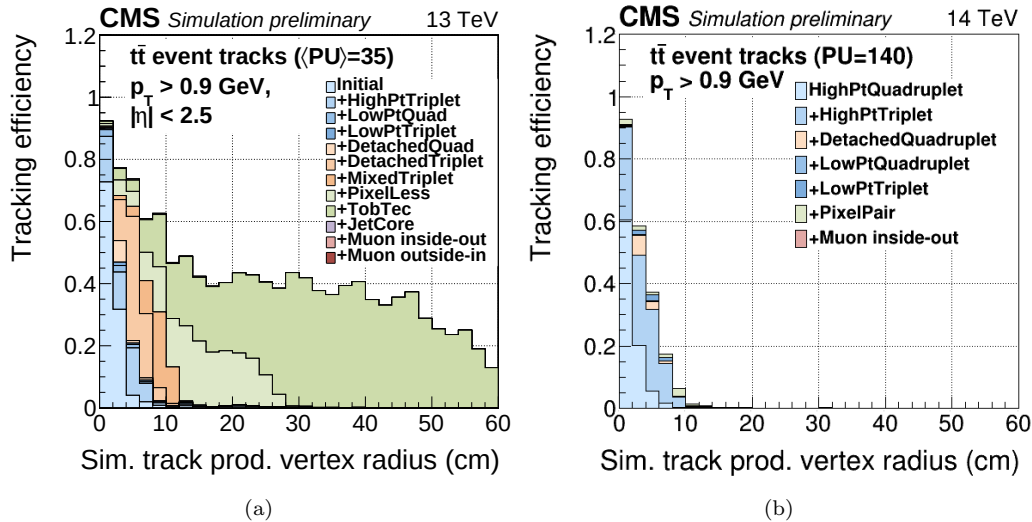


Figure 6.24: Cumulative contributions to the overall tracking performance as a function of the simulated track production vertex radius in 2017 track reconstruction (a) and in Phase-2 track reconstruction (b). Simulated $t\bar{t}$ events are used at nominal pileup conditions of 35 (a) and 140 (b).

In order to recover displaced tracks in the Phase-2 scenario, vector hits are used in a new iteration, called PixelLess, whose seeds are created in the OT. The new iteration has the following features:

- **Hit Doublets:** Each seed is made of pairs of vector hits reconstructed in adjacent layers.
- **Tracking regions:** A hit doublet is accepted as seed if the estimated p_T is more than 0.4 GeV and if the transverse and longitudinal distances of closest approach to the reconstructed beam spot are less than 1 cm and 12 cm, respectively.
- **Seeding layers:** Currently only the PS modules in the OT are included: only the first three layers of the barrel and the first two discs of each endcap in the OT are used. In each endcap, only the first eight rings are included.

After the seeding step has been performed, the usual track reconstruction algorithm is run. The seeding efficiency for this new iteration as a function of the pseudorapidity is shown in Figure 6.25 for single isolated muons with p_T in the range between 0 and 200 GeV and for the $t\bar{t}$ sample. In both datasets no pileup is added. The efficiency is flat at almost 100% for the first sample analyzed in the pseudorapidity range spanned by the OT, while it is around 90% for the second sample for simulated tracks with p_T more than 1 GeV. The degradation for the $t\bar{t}$ sample can be explained with nuclear interaction effects which increases with the amount of material budget.

6.3 Results

A possible development for Phase-2 track reconstruction is described in Section 6.2. It exploits the implementation and use of a new kind of hits, vector hits, in the new Outer Tracker layout.

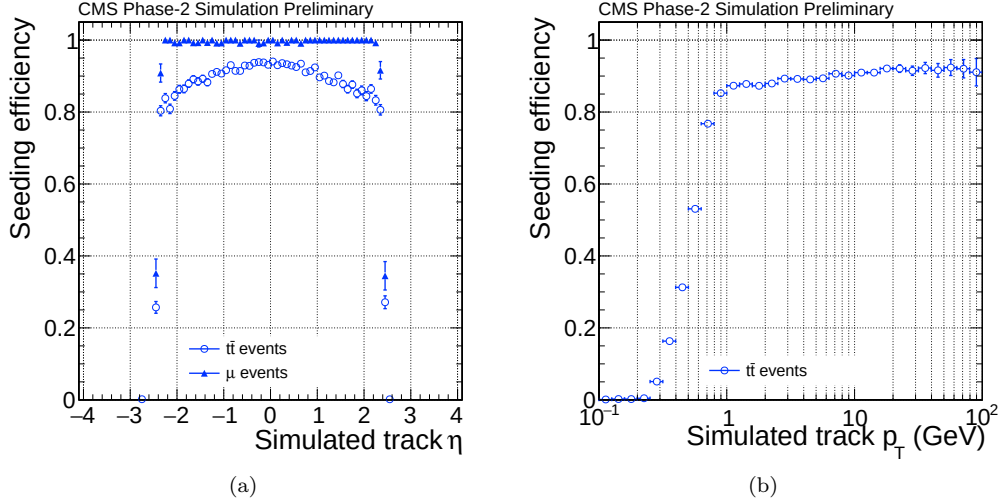


Figure 6.25: Seeding efficiency of the **OT**-seeded iteration using a single muon sample generated with variable p_T (full triangles) and a $t\bar{t}$ sample (open circles) without any pileup superimposed as a function of the pseudorapidity (a) and the p_T of the simulated particle (b).

Vector hits are built in the **OT** stacked modules and are used in a new iteration which uses only seeds produced in the **OT** referred to as PixelLess. In Figure 6.26 the current Phase-2 tracking is compared to track reconstruction using vector hits in all iterations and including the new PixelLess iteration. A $t\bar{t}$ sample with 200PU is used in both scenarios. The same selections that are described in Section 3.2.3 are applied for both efficiency and fake rate and are indicated for each plot. The efficiency distributions as a function of the pseudorapidity and the p_T of the simulated particle show that the new development has similar or better performance with respect to the current Phase-2 tracking. Moreover, the number of misidentified tracks is shown to be reduced by almost an order of magnitude in the **OT** pseudorapidity range and especially for low- p_T tracks. This effect is probably due to the fact that the tracks are now required to have two single hits for each layers, which have already been selected as compatible with each other under the width requirement.

In Figure 6.27 the efficiency is shown as a function of the simulated track production vertex radius: due to the introduction of the PixelLess iteration, the performance for the future tracking improves significantly for large vertex radius compared to the current Phase-2 tracking.

Related to this improvement, it can be noted that the new Phase-2 tracking increases dramatically the reconstruction performance of the tracks coming from displaced vertices. As an example, the performance for a long-living neutral particle, the K_S^0 , is presented here. The invariant mass distribution of the K_S^0 particle, its reconstruction efficiency and misidentification rate are shown in Figure 6.28. The K_S^0 particles are directly produced in the $t\bar{t}$ sample already used for the performance plots in Figures 6.26 and 6.27. In this specific case the requirement on the *High Purity* of the track is dropped as well as any requirement on the track p_T or on the distance with respect to the centre of the luminous region. A displaced vertex is required with two daughter tracks. Some preliminary conclusions can already be drawn: The future tracking has better resolution of the invariant mass peak, the efficiency is almost doubled and the number of misidentified particles is significantly reduced. It is clear that further studies will benefit from higher statistics to understand the performance in more detail.

This preliminary development can be improved upon in the future, introducing more **OT** layers or including a second **OT** seeded-only iteration which makes use also of the **2S** modules in the **OT**. In this case the tracking regions would have to be enlarged.

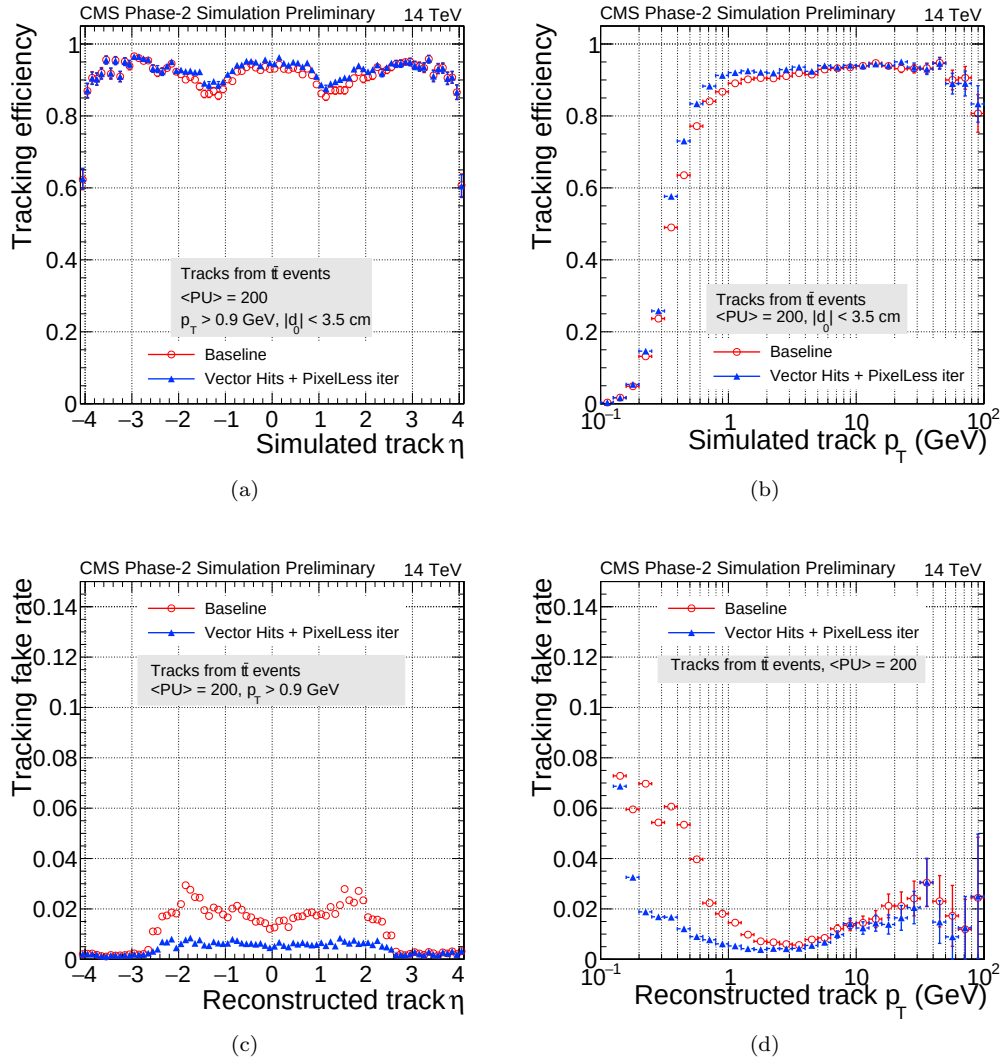


Figure 6.26: Tracking efficiency (a, b) and fake rate (c, d) as a function of the pseudorapidity and the track p_T for the current Phase-2 tracking (open red circles) and the future tracking using vector hits and including the OT-seeded iteration (full blue triangles). Simulated $t\bar{t}$ events are used with 200PU.

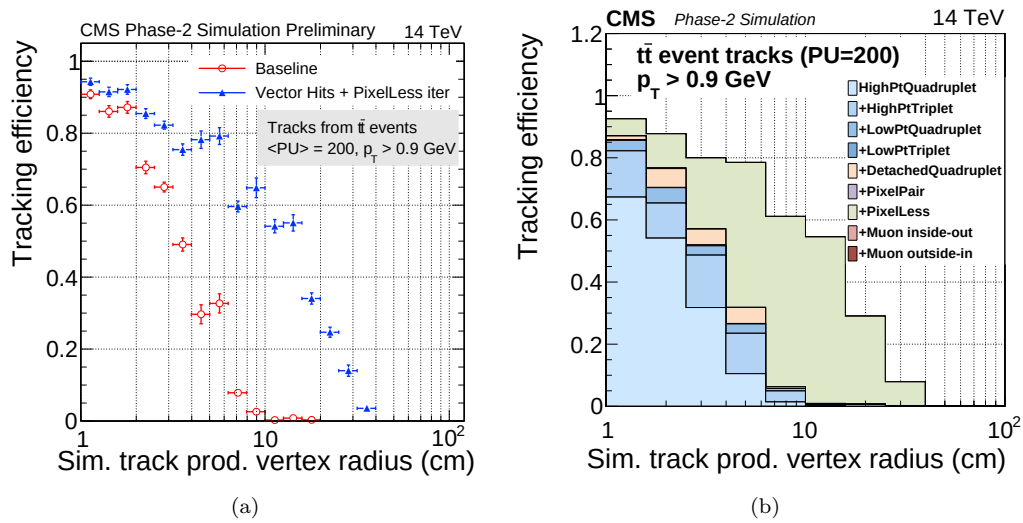


Figure 6.27: (a) Tracking efficiency as a function of the simulated track production vertex radius for the current Phase-2 tracking (open red circles) and the future tracking using vector hits and including the OT-seeded iteration (full blue triangles).

(b) Cumulative contributions to the overall tracking performance as a function of the simulated track production vertex radius using the Phase-2 track reconstruction including vector hits and the OT-seeded iteration.

Simulated $t\bar{t}$ events are used in both plots with 200PU.

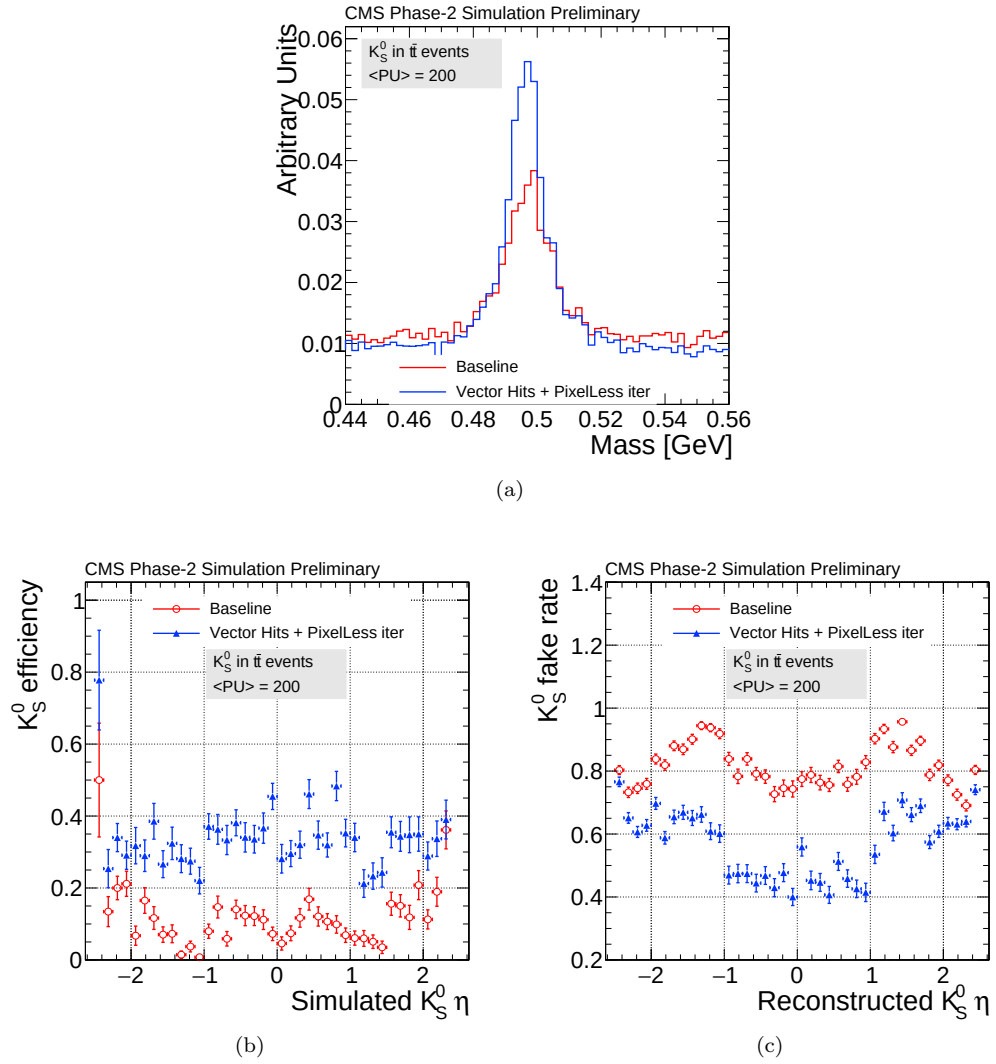


Figure 6.28: Invariant mass distribution (a), reconstruction efficiency (b) and misidentification rate (c) using K_S^0 particles reconstructed in $t\bar{t}$ events with 200PU for the current Phase-2 tracking (open red circles) and the future tracking using vector hits and including the OT-seeded iteration (full blue triangles).

CHAPTER 7

Summary and outlook

*To Fabiola Gianotti, first woman
director general of CERN. (2016)*

In the previous chapters, the work performed during my Ph.D. project in Higgs physics and particle reconstruction in the CMS collaboration was presented.

On the Higgs analysis side, a study in the context of the $H \rightarrow \tau^+\tau^-$ analysis was described with the aim of finding a good signal-versus-background classifier. Higgs boson events produced through **Vector Boson Fusion (VBF)** and decaying into $\mu\tau_h$ were simulated as signal, while events generated from the $Z \rightarrow \tau^+\tau^-$ process were used as background events. The **Multivariate Analysis (MVA)** approach was chosen over a cut-based approach to enhance the discrimination power of the final-state observables used in the analysis, exploiting possible correlations and non-linear decision boundaries. In the analysis reported here, a multi-layer perceptron was chosen as the classifier, in order to profit from its ability to learn from data under a supervised training process. A feedforward neural network with one hidden layer was implemented through the NeuroBayes package, which is a sophisticated tool that includes also automated preprocessing of the input variables. The input dataset was simulated with a centre-of-mass energy of 13 TeV and a total integrated luminosity corresponding to 10 fb^{-1} and was equally divided for the training and testing process. After that the neural network and its parameters were optimized, the classification performance was quantified using a unique **figure of merit (FoM)**, and the final result was compared with other classifiers. Although the **FoM** was found to be better than the cut-based approach, it is not the best result among the different **MVAs** analyzed. This is very likely due to the fact that the NeuroBayes neural network requires a larger dataset in order to get a significant improvement in performance.

In the field of event reconstruction, my work has revolved around the track reconstruction. This is one of the crucial parts in the event reconstruction chain because of its importance in the estimation of the particle momentum and in vertex identification. This task is very challenging in the context of the proton-proton collisions at the **LHC**, given the hundreds or even thousands of particles generated in each bunch crossing. During **Long Shutdown 1**, the **Deterministic Annealing Filter (DAF)** algorithm was implemented in the CMS software to protect track reconstruction in noisy environments or in high track density environments against wrong hit assignments. The **DAF** algorithm showed small improvements for a sample of **QCD** jets events with p_T between 3 and 3.5 TeV. The reduction of mis-identified tracks was even more evident inside of the core of high- p_T jets, without any loss in terms of efficiency.

During Run 2, I also measured the muon tracking efficiency by the tag and probe method on both simulated and real data. The consistency between data and **MC** was studied as a function of several variables and was found to be within a few per-mill in the $Z \rightarrow \mu^+\mu^-$ sample, and within a few per-cent in the $J/\psi \rightarrow \mu^+\mu^-$ sample.

In the context of the Phase-2 upgrade of the **CMS** experiment, this thesis has laid the foundation of the future track reconstruction at the **HL-LHC**. In a first step, the existing tracking algorithms were adapted to the new tracker, and several studies were performed with data from full simulation with different tracker geometries and pileup conditions. The performance of tracking and vertexing for different event types was evaluated and found to be excellent. The results are included in the Tracker Technical Design Report [92]. In a second step, a new type of hits so-called vector hits, was introduced and studied, with the aim of answering the question whether vector hits bring tangible benefits in the pattern recognition step at the **HL-LHC**. The work on the local reconstruction of the vector hits was described in order to study possible mitigation actions to deal with the harsh pileup condition at the **HL-LHC**. In a third step, the vector hits were introduced into the **CMS** track reconstruction, extending the **Kalman Filter** to make use of their full information. Moreover, a new iteration in the Phase-2 tracking was added whose seeds are built using only the **OT** information. Finally, the latest results were shown, where the single hits in the **OT** were replaced by vector hits in all iterations and the new iteration was included. A reduction of about an order of magnitude of the number of mis-identified tracks for high pileup environments was found with a significant improvement in the reconstruction of tracks coming from displaced vertices.

The natural development of my work on Phase-2 tracking is to analyze the time and memory consumption introduced by using vector hits and to introduce further technical- and physics-related optimizations. The first point is imperative in the track reconstruction because it is traditionally the most time consuming part of the entire reconstruction in **CMS**, and this will become an even more pressing problem in the high pileup scenario at the **HL-LHC**. For instance, the vector hits can be used in new highly parallelizable algorithms, such as the long short-term memory neural network [139] or the elastic arm algorithm [140]. In the physics analyses, further developments in the long-living neutral particle reconstruction can also be introduced along with more detailed studies using higher statistics. Possible developments are either including more **OT** layers or a second **OT** seeded-only iteration which makes use also of the **2S** modules in the **OT**. This work can be of fundamental importance for the **HL-LHC** analyses that will target more and more exotic final states.

Bibliography

- [1] J. J. Thomson, “Cathode rays”, *Phil. Mag. Ser.5*, vol. 44, pp. 293–316, 1897, [doi:10.1080/14786449708621070](https://doi.org/10.1080/14786449708621070).
- [2] E. Rutherford, “The scattering of alpha and beta particles by matter and the structure of the atom”, *Phil. Mag. Ser.6*, vol. 21, pp. 669–688, 1911, [doi:10.1080/14786440508637080](https://doi.org/10.1080/14786440508637080).
- [3] J. Chadwick, “The Existence of a Neutron”, *Proc. Roy. Soc. Lond.*, vol. A136, no. 830, pp. 692–708, 1932, [doi:10.1098/rspa.1932.0112](https://doi.org/10.1098/rspa.1932.0112).
- [4] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version*. Physics textbook, New York, NY: Wiley, 2008.
- [5] C. Patrignani *et al.* (Particle Data Group Collaboration), “Review of Particle Physics”, *Chin. Phys.*, vol. C40, no. 10, p. 100001, 2016, [doi:10.1088/1674-1137/40/10/100001](https://doi.org/10.1088/1674-1137/40/10/100001).
- [6] Q. R. Ahmad *et al.* (SNO Collaboration), “Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory”, *Phys. Rev. Lett.*, vol. 89, p. 011301, 2002, [doi:10.1103/PhysRevLett.89.011301](https://doi.org/10.1103/PhysRevLett.89.011301), e-print: [arXiv:nuc1-ex/0204008](https://arxiv.org/abs/nuc1-ex/0204008) [nuc1-ex].
- [7] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964, [doi:10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508).
- [8] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964, [doi:10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- [9] G. S. Guralnik *et al.*, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964, [doi:10.1103/PhysRevLett.13.585](https://doi.org/10.1103/PhysRevLett.13.585).
- [10] J. Goldstone *et al.*, “Broken Symmetries”, *Phys. Rev.*, vol. 127, pp. 965–970, 1962, [doi:10.1103/PhysRev.127.965](https://doi.org/10.1103/PhysRev.127.965).
- [11] F. J. Hasert *et al.* (Gargamelle Neutrino Collaboration), “Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment”, *Phys. Lett.*, vol. 46B, pp. 138–140, 1973, [doi:10.1016/0370-2693\(73\)90499-1](https://doi.org/10.1016/0370-2693(73)90499-1).
- [12] G. Arnison *et al.* (UA1 Collaboration), “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV”, *Phys. Lett.*, vol. 122B, pp. 103–116, 1983, [doi:10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2).
- [13] G. Arnison *et al.* (UA1 Collaboration), “Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/ c^2 at the CERN SPS Collider”, *Phys. Lett.*, vol. 126B, pp. 398–410, 1983, [doi:10.1016/0370-2693\(83\)90188-0](https://doi.org/10.1016/0370-2693(83)90188-0).
- [14] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction”, *Prog. Theor. Phys.*, vol. 49, pp. 652–657, 1973, [doi:10.1143/PTP.49.652](https://doi.org/10.1143/PTP.49.652).

- [15] N. Cabibbo, “Unitary Symmetry and Leptonic Decays”, *Phys. Rev. Lett.*, vol. 10, pp. 531–533, 1963, [doi:10.1103/PhysRevLett.10.531](https://doi.org/10.1103/PhysRevLett.10.531).
- [16] D. Wang *et al.* (PVDIS Collaboration), “Measurement of parity violation in electronquark scattering”, *Nature*, vol. 506, no. 7486, pp. 67–70, 2014, [doi:10.1038/nature12964](https://doi.org/10.1038/nature12964).
- [17] F. Mandl and G. Shaw, *Quantum field theory*. Wiley, 1984.
- [18] ATLAS Collaboration (ATLAS Collaboration), “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett.*, vol. B716, pp. 1–29, 2012, [doi:10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020), e-print: [arXiv:1207.7214](https://arxiv.org/abs/1207.7214) [[hep-ex](#)].
- [19] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys.Lett.*, vol. B30, 2012, [doi:10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021), e-print: [arXiv:1207.7235](https://arxiv.org/abs/1207.7235).
- [20] LHC Higgs Cross Section Working Group, S. Dittmaier, *et al.*, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”, CERN Report CERN-2011-002, 2011.
- [21] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”, Tech. Rep. CERN-2013-004, CERN, Geneva, 2013.
- [22] G. Aad *et al.* (ATLAS Collaboration), “Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”, *Phys. Rev.*, vol. D90, no. 11, p. 112015, 2014, [doi:10.1103/PhysRevD.90.112015](https://doi.org/10.1103/PhysRevD.90.112015), e-print: [arXiv:1408.7084](https://arxiv.org/abs/1408.7084) [[hep-ex](#)].
- [23] S. Chatrchyan *et al.* (CMS Collaboration), “Measurement of the properties of a Higgs boson in the four-lepton final state”, *Phys. Rev.*, vol. D89, no. 9, p. 092007, 2014, [doi:10.1103/PhysRevD.89.092007](https://doi.org/10.1103/PhysRevD.89.092007), e-print: [arXiv:1312.5353](https://arxiv.org/abs/1312.5353) [[hep-ex](#)].
- [24] G. Aad *et al.* (ATLAS Collaboration), “Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector”, *JHEP*, vol. 4, p. 117, 2015, [doi:10.1007/JHEP04\(2015\)117](https://doi.org/10.1007/JHEP04(2015)117), e-print: [arXiv:1501.04943](https://arxiv.org/abs/1501.04943) [[hep-ex](#)].
- [25] A. Sirunyan *et al.*, “Observation of the SM scalar boson decaying to a pair of τ leptons with the CMS experiment at the LHC”, Tech. Rep. CMS-PAS-HIG-16-043, CERN, Geneva, 2017.
- [26] S. Chatrchyan *et al.* (CMS Collaboration), “Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons”, *JHEP*, vol. 05, p. 104, 2014, [doi:10.1007/JHEP05\(2014\)104](https://doi.org/10.1007/JHEP05(2014)104), e-print: [arXiv:1401.5041](https://arxiv.org/abs/1401.5041) [[hep-ex](#)].
- [27] M. Roos, “Dark Matter: The evidence from astronomy, astrophysics and cosmology”, 2010, e-print: [arXiv:1001.0316](https://arxiv.org/abs/1001.0316) [[astro-ph.CO](#)].
- [28] J. F. Donoghue, “The effective field theory treatment of quantum gravity”, *AIP Conf. Proc.*, vol. 1483, pp. 73–94, 2012, [doi:10.1063/1.4756964](https://doi.org/10.1063/1.4756964), e-print: [arXiv:1209.3511](https://arxiv.org/abs/1209.3511) [[gr-qc](#)].
- [29] L. Evans and P. Bryant, “LHC Machine”, *JINST*, vol. 3, no. 08, p. S08001, 2008.
- [30] W. K. H. Panofsky, “The evolution of particle accelerators & colliders”, 1997.
- [31] K. Schindl, “The injector chain for the LHC”, Tech. Rep. CERN-OPEN-99-052, 1999.
- [32] J. M. Jowett, M. Schaumann, and R. Versteegen, “Heavy ion operation from run 2 to HL-LHC”, Tech. Rep. CERN-2014-006.167.
- [33] F. Marcastel, “CERN’s Accelerator Complex. La chaîne des accélérateurs du CERN.” ["https://cds.cern.ch/record/1621583"](https://cds.cern.ch/record/1621583).

- [34] G. Apollinari *et al.*, “High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report”, Tech. Rep. CERN-2015-005, Geneva, 2015.
- [35] CMS Collaboration, “Public CMS Luminosity Information.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [36] G. Aad *et al.* (ATLAS Collaboration), “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST*, vol. 3, p. S08003, 2008.
- [37] S. Chatrchyan *et al.* (CMS Collaboration), “The CMS Experiment at the CERN LHC”, *JINST*, vol. 3, p. S08004, 2008, doi:10.1088/1748-0221/3/08/S08004.
- [38] A. A. Alves *et al.* (LHCb Collaboration), “The LHCb Detector at the LHC”, *JINST*, vol. 3, p. S08005, 2008.
- [39] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST*, vol. 3, no. 08, p. S08002, 2008.
- [40] B. B. Abelev *et al.* (ALICE Collaboration), “Performance of the ALICE Experiment at the CERN LHC”, *Int. J. Mod. Phys.*, vol. A29, p. 1430044, 2014, doi:10.1142/S0217751X14300440, e-print: [arXiv:1402.4476](https://arxiv.org/abs/1402.4476) [nucl-ex].
- [41] LHCb Collaboration, “Measurement of the track reconstruction efficiency at LHCb”, *JINST*, vol. 10, no. 02, p. P02007, 2015.
- [42] E. Stanecka, “The ATLAS Inner Detector operation, data quality and tracking performance”, in *Proceedings, 32nd International Symposium on Physics in Collision (PIC 2012): Strbske Pleso, Slovakia, September 12-15, 2012*, pp. 383–388, 2013.
- [43] K. Potamianos, “The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider”, *PoS*, vol. EPS-HEP2015, p. 261, 2015, e-print: [arXiv:1608.07850](https://arxiv.org/abs/1608.07850) [physics.ins-det].
- [44] V. Berardi *et al.*, “Total cross-section, elastic scattering and diffraction dissociation at the large hadron collider at cern: Totem technical design report”, Tech. Rep. TOTEM-TDR-001-add-1; CERN-LHCC-2004-020, Geneva, 2004.
- [45] O. Adriani *et al.*, “LHCf experiment: Technical Design Report”, Tech. Rep. CERN-LHCC-2006-004, Geneva, 2006.
- [46] T. Sakuma and T. McCauley, “Detector and Event Visualization with SketchUp at the CMS Experiment”, Tech. Rep. CMS-CR-2013-379, 2014.
- [47] CMS Collaboration, “The CMS tracker system project: Technical Design Report”, Tech. Rep. CMS-TDR-005, 1997.
- [48] Institut für Hochenergiephysik, “Picture gallery of CMS experiment.” <http://www.hephy.at/project/cmstracker/gallery/>.
- [49] CMS Collaboration, “CMS Tracker Detector Performance Results.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK>.
- [50] M. Dominguez *et al.*, “CMS Technical Design Report for the Pixel Detector Upgrade”, Tech. Rep. CERN-LHCC-2012-016, CMS-TDR-011, FERMILAB-DESIGN-2012-02, 2012.
- [51] CMS Collaboration, “The CMS electromagnetic calorimeter project: Technical Design Report”, Tech. Rep. CMS-TDR-004, 1997.
- [52] CMS Collaboration, “Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV”, Tech. Rep. CMS-PAS-EGM-11-001, 2011.
- [53] CMS Collaboration, “The CMS hadron calorimeter project : Technical Design Report”, Tech. Rep. CMS-TDR-002, 1997.

- [54] CMS Collaboration, “The CMS muon project : Technical Design Report”, Tech. Rep. CMS-TDR-003, 1997.
- [55] G. Abbiendi (CMS Collaboration), “The CMS muon system in Run 2: preparation, status and first results”, *PoS*, vol. EPS-HEP2015, p. 237, 2015.
- [56] L. Guiducci, “CMS muon system towards LHC Run 2 and beyond”, Tech. Rep. CMS-CR-2014-333, CERN, Geneva, 2014.
- [57] CMS Collaboration, “CMS TriDAS project : Technical Design Report; 1, the trigger systems”, Tech. Rep. CMS-TDR-006, 2000.
- [58] W. Stirling. Private communication.
- [59] CMS Collaboration, “CMS Physics: Technical Design Report Volume 1: Detector Performance and Software”, Tech. Rep. CERN-LHCC-2006-001; CMS-TDR-8-1, Geneva, 2006.
- [60] D. J. Lange, “The CMS Reconstruction Software”, *Journal of Physics: Conference Series*, vol. 331, no. 3, p. 032020, 2011.
- [61] “Github website.” <https://github.com/>.
- [62] K. Bloom, “CMS software and computing for LHC Run 2”, *PoS*, vol. ICHEP2016, p. 185, 2016, e-print: [arXiv:1611.03215](https://arxiv.org/abs/1611.03215) [[physics.ins-det](#)].
- [63] G. Cerati (CMS Collaboration), “Vertexing and Tracking Algorithms at High Pile-Up”, *PoS*, vol. Vertex2014, p. 37, 2015.
- [64] S. Chatrchyan *et al.* (CMS Collaboration), “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”, *JINST*, vol. 7, p. P10002, 2012, [doi:10.1088/1748-0221/7/10/P10002](https://doi.org/10.1088/1748-0221/7/10/P10002), e-print: [arXiv:1206.4071](https://arxiv.org/abs/1206.4071) [[physics.ins-det](#)].
- [65] E. Meschi *et al.*, “Electron Reconstruction in the CMS Electromagnetic Calorimeter”, *CMS Notes*, 2001.
- [66] A. A. Anuar, “Electrons and photons at High Level Trigger in CMS for Run II”, *J. Phys. Conf. Ser.*, vol. 664, no. 8, p. 082001, 2015, [doi:10.1088/1742-6596/664/8/082001](https://doi.org/10.1088/1742-6596/664/8/082001).
- [67] W. Adam *et al.*, “Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC”, Tech. Rep. CMS-NOTE-2005-001, CERN, Geneva, 2005.
- [68] A. M. Sirunyan *et al.* (CMS Collaboration), “Particle-flow reconstruction and global event description with the cms detector”, *JINST*, vol. 12, p. P10003, 2017, [doi:10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003), e-print: [arXiv:1706.04965](https://arxiv.org/abs/1706.04965) [[physics.ins-det](#)].
- [69] R. Atkin, “Review of jet reconstruction algorithms”, *J. Phys. Conf. Ser.*, vol. 645, no. 1, p. 012008, 2015, [doi:10.1088/1742-6596/645/1/012008](https://doi.org/10.1088/1742-6596/645/1/012008).
- [70] G. P. Salam and G. Soyez, “A Practical Seedless Infrared-Safe Cone jet algorithm”, *JHEP*, vol. 5, p. 86, 2007, [doi:10.1088/1126-6708/2007/05/086](https://doi.org/10.1088/1126-6708/2007/05/086), e-print: [arXiv:0704.0292](https://arxiv.org/abs/hep-ph/0704.0292) [[hep-ph](#)].
- [71] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, *Phys. Rev.*, vol. D48, pp. 3160–3166, 1993, [doi:10.1103/PhysRevD.48.3160](https://doi.org/10.1103/PhysRevD.48.3160), e-print: [arXiv:hep-ph/9305266](https://arxiv.org/abs/hep-ph/9305266) [[hep-ph](#)].
- [72] S. Bentvelsen and I. Meyer, “The Cambridge jet algorithm: features and applications”, *Eur. Phys. J.*, vol. C4, pp. 623–629, 1998, [doi:10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063), e-print: [arXiv:hep-ph/9803322](https://arxiv.org/abs/hep-ph/9803322).
- [73] M. Cacciari *et al.*, “The Anti- k_T jet clustering algorithm”, *JHEP*, vol. 804, p. 63, 2008, [doi:10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063), e-print: [0802.1189](https://arxiv.org/abs/hep-ph/0802.1189).

- [74] V. Khachatryan *et al.* (CMS Collaboration), “Reconstruction and identification of lepton decays to hadrons and τ at CMS”, *JINST*, vol. 11, no. 1, p. P01019, 2016, [doi:10.1088/1748-0221/11/01/P01019](https://doi.org/10.1088/1748-0221/11/01/P01019), e-print: [arXiv:1510.07488](https://arxiv.org/abs/1510.07488) [physics.ins-det].
- [75] “Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2”, Tech. Rep. CMS-PAS-TAU-16-002, CERN, Geneva, 2016.
- [76] A. Hocker *et al.*, “TMVA - Toolkit for Multivariate Data Analysis”, *PoS*, vol. ACAT, p. 40, 2007, e-print: [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039) [PHYSICS].
- [77] R. Frühwirth and M. Regler, *Data Analysis Techniques for High-Energy Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 2000.
- [78] J. D. Jackson, *Classical Electrodynamics*. Wiley, 1998.
- [79] A. Strandlie and W. Wittek, “Propagation of Covariance Matrices of Track Parameters in Homogeneous Magnetic Fields in CMS”, Tech. Rep. CMS-NOTE-2006-001, CERN, Geneva, 2006.
- [80] R. Frühwirth, “Application of Kalman filtering to track and vertex fitting”, *Nucl. Instrum. Methods Phys. Res., A*, vol. 262, p. 444, 1987.
- [81] R. Frühwirth, “Track fitting with non-Gaussian noise”, *Comput. Phys. Commun.*, vol. 100, pp. 1–16, 1997, [doi:10.1016/S0010-4655\(96\)00155-5](https://doi.org/10.1016/S0010-4655(96)00155-5).
- [82] R. Frühwirth and A. Strandlie, “Track fitting with ambiguities and noise: A study of elastic tracking and nonlinear filters”, *Comput. Phys. Commun.*, vol. 120, no. 2-3, pp. 197–214, 1999, [doi:10.1016/S0010-4655\(99\)00231-3](https://doi.org/10.1016/S0010-4655(99)00231-3).
- [83] S. Chatrchyan *et al.* (CMS Collaboration), “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST*, vol. 9, no. 10, p. P10009, 2014, [doi:10.1088/1748-0221/9/10/P10009](https://doi.org/10.1088/1748-0221/9/10/P10009), e-print: [arXiv:1405.6569](https://arxiv.org/abs/1405.6569) [physics.ins-det].
- [84] H. C. Kastli *et al.*, “Design and performance of the CMS pixel detector readout chip”, *Nucl. Instrum. Meth.*, vol. A565, pp. 188–194, 2006, [doi:10.1016/j.nima.2006.05.038](https://doi.org/10.1016/j.nima.2006.05.038), e-print: [arXiv:physics/0511166](https://arxiv.org/abs/physics/0511166) [physics].
- [85] P. Billoir and S. Qian, “Simultaneous pattern recognition and track fitting by the Kalman filtering method”, *Nucl. Instrum. Meth.*, vol. A294, pp. 219–228, 1990, [doi:10.1016/0168-9002\(90\)91835-Y](https://doi.org/10.1016/0168-9002(90)91835-Y).
- [86] C. D. Jones *et al.* (CMS Collaboration), “Using the CMS Threaded Framework In A Production Environment”, *J. Phys.: Conf. Ser.*, vol. 664, no. 7, 2015.
- [87] V. Mariani, “The RunII tracking performances of the CMS detector”, *Nuovo Cimento C*, vol. 40, no. 1, p. 23, 2017.
- [88] CMS Collaboration, “CMS Tracking POG Performance Plots for year 2016.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPlots2016>.
- [89] CMS Collaboration, “CMS Tracking POG Performance Plots 2016 (targeting ICHEP conference).” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPlotsICHEP2016>.
- [90] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, “MadGraph 5 : Going Beyond”, *JHEP*, vol. 6, p. 128, 2011, [doi:10.1007/JHEP06\(2011\)128](https://doi.org/10.1007/JHEP06(2011)128), e-print: [arXiv:1106.0522](https://arxiv.org/abs/1106.0522) [hep-ph].
- [91] T. Sjöstrand *et al.*, “An Introduction to PYTHIA 8.2”, *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015, [doi:10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024), e-print: [arXiv:1410.3012](https://arxiv.org/abs/1410.3012) [hep-ph].

- [92] CMS Collaboration, “The Phase-2 Upgrade of the CMS Tracker”, Tech. Rep. CERN-LHCC-2017-009. CMS-TDR-014, CERN, Geneva, 2017. To be published.
- [93] “2017 tracking performance plots”, Tech. Rep. CMS-DP-2017-015, Apr 2017.
- [94] F. Pantaleo, “New Track Seeding Techniques for the CMS Experiment”, Ph.D thesis, Universität Hamburg, 2017.
- [95] L. Breiman *et al.*, *Classification and regression trees*. The Wadsworth statistics/probability series, Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 1984.
- [96] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting”, *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997, doi:10.1006/jcss.1997.1504.
- [97] V. Schamböck, “Potential of the CMS $H \rightarrow \tau\tau$ analysis using boosted decision trees”, Master’s thesis, TU Wien, 2016.
- [98] M. Feindt, “A Neural Bayesian Estimator for Conditional Probability Densities”, Tech. Rep. IEKP-KA-04-05, 2004.
- [99] M. Feindt and U. Kerzel, “The NeuroBayes neural network package”, *Nucl. Instrum. Meth.*, vol. A559, pp. 190–194, 2006, doi:10.1016/j.nima.2005.11.166.
- [100] I. Goodfellow *et al.*, *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [101] M. Spanring, “Study of potential improvements of the CMS $H \rightarrow \tau\tau$ analysis using artificial neural networks with multiple layers”, Master’s thesis, TU Wien, 2015.
- [102] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2nd ed., 1998.
- [103] N. Qian, “On the momentum term in gradient descent learning algorithms”, *Neural Netw.*, vol. 12, pp. 145–151, Jan. 1999, doi:10.1016/S0893-6080(98)00116-6.
- [104] Phi-T, “The NeuroBayes Users Guide”, 2010.
- [105] R. H. Byrd *et al.*, “A limited-memory algorithm for bound constrained optimization”, *Siam Journal on Scientific Computing*, vol. 16, pp. 1190–1208, 1994.
- [106] G. Cowan and E. Gross, “Discovery significance with statistical uncertainty in the background estimate.” <https://www.p.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf>.
- [107] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP*, vol. 11, p. 70, 2007, doi:10.1088/1126-6708/2007/11/070, e-print: arXiv:0709.2092 [hep-ph].
- [108] T. Sjöstrand *et al.*, “PYTHIA 6.4 Physics and Manual”, *JHEP*, vol. 5, p. 26, 2006, doi:10.1088/1126-6708/2006/05/026, e-print: arXiv:hep-ph/0603175 [hep-ph].
- [109] S. Agostinelli *et al.* (GEANT4 Collaboration), “GEANT4: A Simulation toolkit”, *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003, doi:10.1016/S0168-9002(03)01368-8.
- [110] K. Simon, “Improving the Higgs to $\tau\tau$ measurement with NeuroBayes.” Bachelor’s thesis, TU Wien, 2015.
- [111] CERN Council, “The European Strategy for Particle Physics, Update 2013”, CERN-Council-S/106. <https://cds.cern.ch/record/1567258/files/esc-e-106.pdf>.
- [112] Particle Physics Project Prioritization Panel (P5), “Building for Discovery - Strategic Plan for U.S. Particle Physics in the Global Context”, May 2014. http://science.energy.gov/~media/hep/hepap/pdf/May-2014/FINAL_P5_Report_053014.pdf.

- [113] CMS and LHCb collaborations, “Observation of the rare $B_s^0 \rightarrow \mu^+ \mu^-$ decay from the combined analysis of CMS and LHCb data”, *Nature*, vol. 522, p. 68, 2015, doi:10.1038/nature14474, e-print: [arXiv:1411.4413 \[hep-ex\]](#).
- [114] M. Aaboud *et al.* (ATLAS Collaboration), “Study of the rare decays of B_s^0 and B^0 into muon pairs from data collected during the LHC Run 1 with the ATLAS detector”, *Eur. Phys. J. C*, vol. 76, p. 513, 2016, doi:10.1140/epjc/s10052-016-4338-8, e-print: [arXiv:1604.04263 \[hep-ex\]](#).
- [115] ATLAS Collaboration, “Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment”, Tech. Rep. CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, 2012.
- [116] D. Contardo *et al.*, “Technical Proposal for the Phase-II Upgrade of the CMS Detector”, Tech. Rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Geneva, 2015.
- [117] R. Tieulent, “ALICE Upgrades: Plans and Potentials”, Tech. Rep. arXiv:1512.02253, 2015.
- [118] R. Aaij *et al.*, “Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era”, Tech. Rep. CERN-LHCC-2017-003, CERN, Geneva, 2017.
- [119] ATLAS Collaboration, “Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment”, Tech. Rep. CERN-LHCC-2011-012. LHCC-I-020, CERN, Geneva, 2011.
- [120] CMS Collaboration, “Technical proposal for the upgrade of the CMS detector through 2020”, Tech. Rep. CERN-LHCC-2011-006. LHCC-P-004, 2011.
- [121] CMS Collaboration, “The Phase-2 Upgrade of the CMS Barrel Calorimeters Technical Design Report”, Tech. Rep. CERN-LHCC-2017-011. CMS-TDR-015, CERN, Geneva, 2017. To be published.
- [122] CMS Collaboration, “The Phase-2 Upgrade of the CMS Endcap Calorimeter”, Tech. Rep. CERN-LHCC-2017-023. CMS-TDR-019, CERN, Geneva, 2017. To be published.
- [123] CMS Collaboration, “The Phase-2 Upgrade of the CMS Muon Detectors”, Tech. Rep. CERN-LHCC-2017-012. CMS-TDR-016, CERN, Geneva, 2017. To be published.
- [124] CMS Collaboration, “The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report”, Tech. Rep. CERN-LHCC-2017-013. CMS-TDR-017, CERN, Geneva, 2017. To be published.
- [125] CMS Collaboration, “The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report”, Tech. Rep. CERN-LHCC-2017-014. CMS-TDR-018, CERN, Geneva, 2017. To be published.
- [126] CMS Collaboration, “Technical Proposal for a MIP Timing Detector in the CMS experiment Phase 2 Upgrade”, Tech. Rep. CMS-TDR-17-006-001, CERN, Geneva, 2017. To be published.
- [127] E. Chabanat and N. Estre, “Deterministic Annealing for Vertex Finding at CMS”, Tech. Rep. CERN-2005-002, 2005.
- [128] CMS Collaboration, “Initial Report of the Fast Timing Working Group”, Tech. Rep. CMS-DP-2016-008, 2016.
- [129] G. Bianchi, “tklayout: a design tool for innovative silicon tracking detectors”, *JINST*, vol. 9, p. C03054, 2014, doi:10.1088/1748-0221/9/03/C03054.
- [130] S. Mersi *et al.*, “CMS Tracker layout studies for HL-LHC”, *Physics Procedia*, vol. 37, p. 1070, 2012, doi:10.1016/j.phpro.2012.03.729.
- [131] E. Migliore, “CMS Pixel Detector design for HL-LHC”, Tech. Rep. CMS-CR-2016-230, CERN, Geneva, 2016.

- [132] J. Chistiansen and M. Garcia-Sciveres, “RD collaboration Proposal: Development of pixel readout integrated circuits for extreme rate and radiation”, Tech. Rep. CERN-LHCC-2013-008. LHCC-P-006, CERN, Geneva, 2013.
- [133] CMS Collaboration, “CMS Phase II Upgrade Scope Document”, Tech. Rep. CERN-LHCC-2015-019. LHCC-G-165, CERN, Geneva, 2015.
- [134] P. Hough, “Method and means for recognizing complex patterns”, 1962. US Patent 3,069,654.
- [135] K. Rose, “Deterministic annealing for clustering, compression, classification, regression and related optimisation problems”, *Proceedings of the IEEE*, vol. 86/11, p. 2210, 1998, [doi:10.1109/5.726788](https://doi.org/10.1109/5.726788).
- [136] R. Frühwirth, W. Waltenberger, and P. Vanlaer, “Adaptive Vertex Fitting”, Tech. Rep. CMS-NOTE-2007-008, CERN, Geneva, 2007.
- [137] G. Cerati *et al.*, “Parallelized Kalman-Filter-Based Reconstruction of Particle Tracks on Many-Core Processors and GPUs”, *EPJ Web Conf.*, vol. 150, p. 6, 2017, [doi:10.1051/epjconf/201715000006](https://doi.org/10.1051/epjconf/201715000006), e-print: [arXiv:1705.02876](https://arxiv.org/abs/1705.02876) [physics.comp-ph].
- [138] R. Frühwirth, “Pattern recognition with vector hits”, *JINST*, vol. 7, no. 09, p. C09002, 2012.
- [139] S. Hochreiter and J. Schmidhuber, “Long short-term memory”, *Neural Comput.*, vol. 9, pp. 1735–1780, Nov. 1997, [doi:10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
- [140] M. Ohlsson, “Extensions and explorations of the elastic arms algorithm”, *Comput. Phys. Commun.*, vol. 77, no. 1, pp. 19 – 32, 1993, [doi:https://doi.org/10.1016/0010-4655\(93\)90033-9](https://doi.org/10.1016/0010-4655(93)90033-9).

Acronyms

CMSSW CMS Software. 36, 38, 99, 132, 144

2S 2-strips. 123–128, 132–134, 143, 156, 162, 175

AD Antiproton Decelerator. 19

AIS Acronyms In Science. 173

ALICE A Large Ion Collider Experiment. 22–24, 110, 113, 114

ATLAS A Toroidal LHC Apparatus. 8, 10–15, 22–25, 49, 109–111, 113–115

AWAKE Proton Driven Plasma Wakefield Acceleration Experiment. 19

BDT Boosted Decision Tree. 13, 15, 92, 107

BSM Beyond the Standard Model. 14, 17, 110, 111

CA Cellular Automaton-based Hit Chain-Maker. 84–87, 89, 134

CE-E Electromagnetic Compartment. 118

CE-H Hadronic Compartment. 118

CERN European Organisation for Nuclear Research. 8, 17, 19, 20, 109

CL Confidence Level. 110, 111

CMS Compact Muon Solenoid. I, 8, 10–15, 17, 22–25, 27–30, 32–39, 41–43, 47, 49, 55, 57–59, 62, 63, 66, 75, 83, 99, 109–111, 113–123, 126, 128, 129, 131, 132, 134, 136, 137, 141, 142, 144, 148, 154, 161, 162, 173

CSC Cathode Strip Chambers. 34, 35, 39, 115, 118

CTF Combinatorial Track Finder. 60, 66, 143, 154

DAF Deterministic Annealing Filter. I, 47, 52, 55, 57, 73, 75–77, 161

DAQ Data Acquisition. 23, 36, 119

DIA Dolor In Asinum. 173

DNN Deep Neural Network. 92, 107

DT Drift Tube. 34, 35, 39, 118

EB Electromagnetic calorimeter Barrel. 31, 118

- ECAL** Electromagnetic Calorimeter. 29, 31, 36, 39–43, 66
- EE** Electromagnetic calorimeter Endcap. 31
- EWK** Electroweak. 1–6, 8
- FE** Front-End. 28, 36, 114, 115, 118, 124, 125, 128
- FoM** figure of merit. 91, 98, 99, 103, 105–107, 161
- FPGA** Field-Programmable Gate Array. 129
- FTK** FastTrackR. 115
- GEM** Gas Electron Multiplier. 114, 118
- GPU** Graphics Processing Units. 84
- GSF** Gaussian-sum Filter. 40, 41, 52, 55, 61, 66
- HB** Hadronic calorimeter Barrel. 32, 33, 115, 118
- HCAL** Hadron Calorimeter. 32, 41, 42
- HE** Hadronic calorimeter Endcap. 32, 33, 115
- HF** Hadronic calorimeter Forward. 32, 33, 115
- HGC** High Granularity Calorimeter. 118
- HIP** Highly Ionizing Particle. 133
- HL-LHC** High Luminosity LHC. I, 19, 21, 109–115, 118–121, 123, 131, 135, 137, 142, 148, 162
- HLT** High-Level Trigger. 36, 40, 41, 118, 142
- HO** Hadronic calorimeter Outer. 32, 33
- IBL** Insertable B-Layer. 24, 25
- ISOLDE** Isotope mass Separator On-Line facility. 19
- IT** Inner Tracker. 120–125, 127, 131–134, 136, 138, 143
- ITS** Inner Tracking System. 23, 24
- KF** Kalman Filter. I, 39–41, 52–55, 60, 61, 65, 73, 129, 142, 154, 162
- L1** Level-1. 34–36, 115, 118–125, 128–130, 135, 142, 143
- LEIR** Low Energy Ion Ring. 19
- LEP** Large Electron-Positron collider. 17
- LHC** Large Hadron Collider. I, 1, 8, 11, 14, 17–25, 27, 29, 31, 37, 42, 44, 47, 49, 58, 71, 109, 110, 112–115, 118, 131, 161
- LHCb** LHC Beauty. 22–24, 110, 113, 114
- LHCf** LHC forward. 25
- LINAC2** Linear accelerator 2. 19
- LINAC3** Linear accelerator 3. 19

- LO** leading-order. 8, 9, 79
- LS1** Long Shutdown 1. I, 19, 24, 34, 35, 38, 47, 63, 71, 84, 115, 161
- LS2** Long Shutdown 2. 19, 114, 115
- LS3** Long Shutdown 3. I, 19, 29, 109, 114, 115, 118
- LSM** Least Squares Method. 52, 53, 55
- MC** Monte Carlo. I, 47, 59, 72, 75, 76, 79, 161
- MIP** Minimum Ionizing Particle. 119, 133
- MVA** Multivariate Analysis. I, 12, 44, 45, 91–93, 96, 98, 99, 103, 107, 161
- NLO** next-to-leading-order. 8, 9
- NNLO** next-to-next-to-leading-order. 8
- NNNLO** next-to-next-to-next-to-leading-order. 8
- OT** Outer Tracker. 120–129, 131–135, 142–146, 148–159, 162
- PF** Particle Flow. 40–43
- PS** Pixel-strip. 123–128, 132–134, 143, 144, 155, 175
- PS** Proton Synchrotron. 19
- PSB** Proton Synchrotron Booster. 19
- PXB** Pixel Tracker Barrel. 28, 59
- PXF** Pixel Tracker Forward. 28, 59
- QCD** Quantum Chromodynamics. 2, 4, 8, 9, 11, 42, 44, 45, 73–77, 79, 135, 137, 161
- QED** Quantum Electrodynamics. 2, 4
- ROC** Read-Out Chip. 124, 125
- RPC** Resistive Plate Chambers. 34, 35, 39, 118
- SCT** Semiconductor Tracker. 24
- SM** Standard Model. 1–4, 7–13, 15, 17, 91, 99, 110, 111
- SPS** Super Proton Synchrotron. 19
- SPS** Super Proton Synchrotron. 19
- TB2S** Tracker Barrel with 2S modules. 124, 132
- TBPS** Tracker Barrel with PS modules. 124, 128, 129, 132
- TBPX** Tracker Barrel Pixel. 123, 132
- TEC** Tracker End Caps. 28, 29, 59
- TEDD** Tracker Endcap Double-Discs. 124, 132
- TEPX** Tracker Barrel Endcap. 123, 132

TFPX Tracker Barrel Forward. 123, 132

TIB Tracker Inner Barrel. 28–30, 59, 72

TID Tracker Inner Disks. 28, 29, 59

TMVA Toolkit for Multivariate Analysis. 92, 96

TOB Tracker Outer Barrel. 29, 59

TOTEM TOTal Elastic and diffractive cross-section Measurement. 24

TPC Time-Projection Chamber. 23, 24, 114

TRD Transition Radiation Detector. 23

TRT Transition Radiation Tracker. 24

TT Tracker Turicensis. 23

VBF Vector Boson Fusion. 8, 11–13, 15, 91, 99, 161

VELO Vertex Locator. 23

WLF Wavelength-shifting Fibers. 32

Thanks

Firstly, I would like to thank my supervisor, as well as the Institut für Hochenergiephysik and the Doktoratskolleg for Particles and Interactions to give me the chance to pursue my PhD.

Rudi Frühwirth. Thanks Rudi for all the support you gave me in these four years. I really enjoyed working with you on the different problems related (and not) to tracking, such as debugging the code, understanding the physical reasons behind each result, reevaluating ideas and implementing complicated solutions. I am grateful too because you also got me involved in the organization of other activities, such as the Connecting the Dots workshop, and you never refuse to give me a chance to prove myself, both in small groups and huge collaborations. Thanks also for caring about my formation and education, supporting me in schools and conferences. I have learned really a lot from you, from tracking to opera, from statistics to German. I was very lucky to have you as supervisor. Thanks again.

HEPHY. When I arrive the first time at HEPHY, I was one of the few foreign people in the institute. I must admit that it has not been easy at the beginning, but all the people in the institute have always been more than supportive in helping, translating, and giving advice. During these four years at HEPHY I met incredible people in this institute, and if I look up the meaning of “passion” in the dictionary, this is all I need to describe their relationship with physics: a strong and barely controllable emotion. In the institute, I have not just found extremely professional colleagues, but also friends - thank you all. To conclude this part, I would like to thank in particular Jochen, who made this environment possible and continues every day to improve it.

DK-PI. The experience in the DK-PI was very intense for many reasons: I have met a lot of young physicists who were working in my field, and I had intense exchanges of opinions with them; I have met a lot of physicists who were NOT working in my field, and also with them I had intense exchanges of opinions. Due to this very peculiar recipe and a careful supervision of the lectures offered, I have definitively broadened my knowledge in many aspects of particle physics. Thanks a lot.

Moreover, I would like to include a special thanks to my colleagues at CERN.

Tracking POG. Vincenzo, because with you, nobody can stop learning - even when the mind is so full that someone can think that with a single additional piece of information it will explode. Marco, for all the psychological, technical and social support. Matti, if I should quantify my gratefulness using the number of emails you sent me, it would be exactly 772 thanks. Giuseppe, thanks for helping me in making my first baby-steps in the crazy tracking code. Felice, for always very interesting and useful discussions. And all people involved in other activities, it was a lot of fun.

Upgrade group. This group includes a lot of very impressive (and important!) people. I would like to thank Andrea, Duccio, Giacomo, Luca, Patrizia, Meenakshi and Katja for all the time they have spent brainstorming, correcting, writing emails, organizing, and allowing me to play such an intense and important role. In other words, thanks for believing in me. It was an honor for me to work with all of you.

Tracker group. From the hardware to the software, I would like to thank all the people involved in this amazing project. In particular, I would like to thank Gaele, Viktor, Alessia, Christophe and Enrico for their enormous support on the local reconstruction side; Stefano and Gabrielle for the very intense collaboration; Erik and Chris for the amount of fun and knowledge

I gained during my shifts.

Finally, I would like to conclude with some more private acknowledgments.

Family. Cara family, ho dedicato questa tesi a tutti voi: nonni, genitori, fratello, zii e cugini perché questo mio traguardo é anche un po' il vostro. Mi avete accompagnata in tutto il mio percorso e supportata in tutte le mie decisioni con sorrisi, consigli e intense discussioni. Vi ringrazio con tutto il cuore. In particolare vorrei ringraziare mamma e papà perché, fin da quando sono bambina, sono la tenacia e la tecnica che mi ha portata fino a questo traguardo. Non so bene se vi rendete conto di quanto c'è di voi dentro di me, io penso di iniziare a realizzarlo. Grazie.

New and old relationships. I would like to thank my old-and-always-there friends, which followed me in my experience at 100% from the beginning, and the not-anymore-so-new ones I have collected during these years. It is extremely nice feeling at home even if you are far away from it. This helped me a lot during difficult times of my thesis and it would not be the same without all of you around. In particular, I would like to thank my very-intellectual-but-also-funny partner, Dinyar. You helped me in facing every challenge during my PhD and gave me more motivation to develop my skills and try to do my best every day. Most of the times, talking with you about my problems was already a step in the direction to solve them. Thanks.

Strangers. I would like to conclude thanking a young scientist that came into my primary school when I was around 7. He showed us some experiments with physics and when I answered an easy question, he told me "Good! You are going to become a physicist.". Here I am. Thanks to you, stranger, and all other people that have just encouraged me without getting anything back.

Last but not least I would like to thank all **women scientists** that have made it possible for me to complete my *Philosophiae Doctor* degree. I have dedicated each chapter of my thesis to one of them - it was the bare minimum.